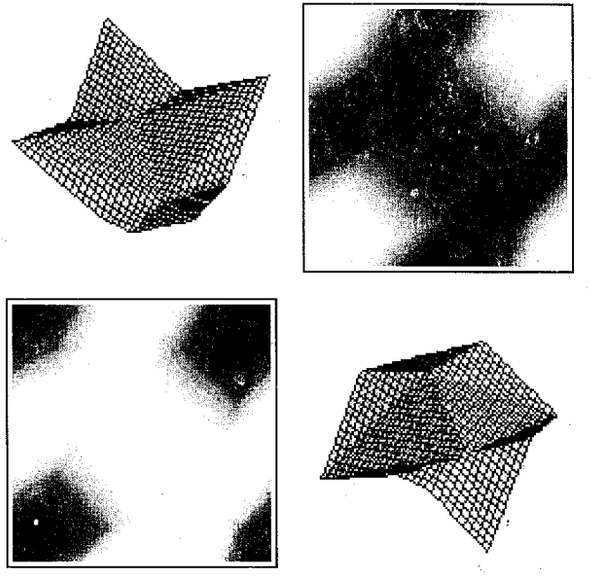
L'Insegnamento della Logica

MINISTERO DELLA PUBBLICA ISTRUZIONE
DIREZIONE GENERALE ISTRUZIONE
CLASSICA, SCIENTIFICA E MAGISTRALE

in collaborazione con
ASSOCIAZIONE ITALIANA DI LOGICA E SUE APPLICAZIONI



In copertina:
una proposizione e la sua negazione
nella logica ad infiniti valori

Elaborazione grafica con software *Mathematica*®
eseguita presso il laboratorio **Codici**
del *Dipartimento di Scienze dell'Informazione*
dell'*Università degli Studi* di Milano

Stampa:
Tipografia Veronese - Padova
Laser fotocomposizioni & C. - Padova

L'Insegnamento della Logica

MINISTERO DELLA PUBBLICA ISTRUZIONE
DIREZIONE GENERALE ISTRUZIONE
CLASSICA, SCIENTIFICA E MAGISTRALE

in collaborazione con
ASSOCIAZIONE ITALIANA DI LOGICA E SUE APPLICAZIONI

a cura di
LUCIA CIARRAPICO e DANIELE MUNDICI

Con la collaborazione del
LICEO GINNASIO STATALE "FRANCESCA CAPECE"
Maglie (Lecce)

Questa pubblicazione è stata realizzata con il contributo dell' IRRSAE PUGLIA

Desidero rivolgere un cordiale saluto ai collaboratori di questo studio con il quale hanno voluto dare ai Docenti della Scuola Secondaria uno strumento attento sui concetti logici fondamentali per l'approfondimento della scienza matematica.

Si tratta di risultati certamente positivi che hanno arricchito la preparazione dei docenti e ancor più l'arricchiranno in futuro. Un ringraziamento sentito ai rappresentanti della Direzione Generale dell'Istruzione Classica e dell'AILA che hanno reso possibile tutto ciò.

ROMANO CAMMARATA
Direttore Generale dell'Istruzione Classica, Scientifica e Magistrale

Presentazione

Negli ultimi decenni di questo secolo i rapporti della logica con le scienze dell'informazione sono andati sempre più assomigliando ai rapporti tra analisi matematica e scienze fisiche nel secolo scorso.

Il crescente ruolo dell'analisi ebbe conseguenze importanti anche dal punto di vista didattico: per esempio per quanto riguarda l'insegnamento dei numeri reali e del calcolo differenziale, che furono innestati in una millenaria tradizione didattica di stampo euclideo, pur essendo estranei alla concezione matematica greca.

Passando dal calcolo differenziale al calculus ratiocinator, le realtà della logica contemporanea si chiamano: linguaggio formale, sintassi e semantica, regola deduttiva, conseguenza, completezza e incompletezza, enumerabilità e indecidibilità, algoritmo, macchina di Turing universale, passo di calcolo, complessità. Alcune delle applicazioni si chiamano:verifica di software e hardware, programmazione logica, deduzione automatizzata, rappresentazione della conoscenza imprecisa e suo controllo mediante logiche aventi molti valori di verità.

Dei concetti logici fondamentali parlano queste pagine, rivolgendosi a Docenti della Scuola Secondaria.

Non è facile dire se il contributo maggiore alla realizzazione del testo finale sia stato dato da insegnanti oppure da logici professionisti. Sicuramente l'interazione positiva tra queste componenti, iniziata in un primo incontro a Lecce (22-26 novembre 1993) e proseguita in un successivo incontro a Otranto (21-23 aprile 1994), è stata determinante.

Questi incontri sono parte di un progetto di aggiornamento che la Direzione Generale dell'Istruzione Classica, Scientifica e Magistrale sta portando avanti da alcuni anni al fine di fornire un'adeguata preparazione ai docenti di matematica sulle molteplici tematiche innovative che, seppure in via sperimentale, fanno parte dei programmi di insegnamento.

Il corso di Lecce e Otranto, previsto dal Decreto Ministeriale 25 novembre 1992, è frutto di una felice collaborazione tra la Direzione Generale dell'Istruzione Classica, Scientifica e Magistrale e dell' AILA, Associazione Italiana di Logica e sue Applicazioni. Alla buona riuscita del corso ha contribuito l' efficienza del Preside Vito Papa del Liceo Ginnasio "Francesca Capece" di Maglie e dei suoi collaboratori.

Al corso hanno partecipato 50 docenti di scuole di tutto il territorio nazionale. Essi rappresentano un numero estremamente esiguo rispetto ai moltissimi che insegnano matematica e che amano insegnarla bene.

Il libro vuole perciò essere un piccolo supporto per altri docenti che vogliono sapere più logica e per quanti desiderano realizzare momenti di formazione sul tema. Alla sua realizzazione ha dato un contributo finanziario anche l'IRRSAE Puglia, a cui va la nostra gratitudine. Alla sua unità di stile di stampa hanno contribuito, con molte ore di sapiente lavoro al computer, il signor Andrea Antonini, il dottor Carlo Mereghetti, del Dottorato in Informatica del Dipartimento di Scienze dell'Informazione dell'Università di Milano e il professor Ugo Solitro dello stesso Dipartimento.

Al lettore che, in prima lettura, trovasse difficoltà in certi capitoli chiediamo la benevola considerazione del fatto che:

(i) La logica dei predicati fu introdotta nella seconda metà del secolo scorso, i teoremi di completezza e incompletezza di Gödel, la macchina di Turing risalgono agli anni trenta di questo secolo. Taluni risultati fondamentali di cui parla questo testo non hanno più di vent'anni di vita. E' dunque comprensibile che, a differenza di quanto avvenuto per le nozioni di spazio e numero, che hanno avuto due o tre millenni di levigatura su banchi, cattedre e lavagne, non si sia ancora consolidata una tradizione didattica per la logica.

(ii) E dunque, il testo naturalmente si presta a molti diversi percorsi di lettura: molti capitoli sono indipendenti tra di loro, molti non richiedono prerequisiti. Ai lettori attenti non sfuggirà poi l' eterogeneità delle provenienze culturali e professionali di coloro che hanno contribuito al testo; né passerà inosservata la varietà di generi letterari di queste pagine: la conferenza, la lezione introduttiva, la lezione non introduttiva, il syllabus-eserciziario, il reportage di un viaggio didattico, la rassegna stampa.

Per una volta, non tutto il testo va letto con carta e penna, non tutti gli esercizi assegnati vanno svolti. E in effetti, questa ricchezza espressiva è la migliore testimonianza che i giochi non sono ancora stati fatti per la logica e il suo insegnamento; e che, per una buona didattica della logica formale, nulla è più illogico del formalismo didascalico.

LUCIA CIARRAPICO
Ispettrice
Ministero della Pubblica Istruzione

DANIELE MUNDICI
Presidente dell'Associazione
Italiana di Logica e sue Applicazioni

Indice

PARTE PRIMA

Lezioni introduttive con Complementi ed Esercizi

Insiemi finiti e infiniti

DARIO PALLADINO
Dipartimento di Filosofia
Università di Genova
via Balbi 4
16126, Genova

1. Premessa

La creazione della teoria degli insiemi ad opera di Georg Cantor (1845-1910) va considerata uno degli eventi più importanti della storia del pensiero matematico; essa, oltre ad essere stata l'origine di una imponente sequenza di risultati tecnici, ha comportato, unitamente all'evoluzione del metodo assiomatico, la vera e propria svolta concettuale che consente di contraddistinguere la matematica del nostro secolo da quella delle epoche precedenti. Essa può essere fatta risalire ai lavori che Cantor pubblicò a partire dal 1872 - in una memoria del 1873 sono presentate le definizioni del concetto di potenza (cardinalità) di un insieme e di insieme numerabile - e soprattutto alla serie di articoli dallo stesso titolo (*Über unendliche, lineare Punktmannigfaltigkeiten*) comparsi fra il 1879 e il 1884 sui *Mathematische Annalen*.

La teoria cantoriana non ha né in sé, né nelle intenzioni dell'autore, quel significato fondazionale che è venuta assumendo a partire dall'inizio del nostro secolo, ma scaturì piuttosto dalle ricerche del matematico tedesco nel campo dell'analisi infinitesimale. Infatti, lo studio della rappresentazione delle funzioni in serie trigonometriche condusse Cantor a considerare insiemi infiniti di punti della retta e del piano e ad avvertire l'esigenza di estendere ad insiemi arbitrari i concetti di grandezza e di "numero"; in tal modo, attraverso generalizzazioni molto ardite per quei tempi[1], pervenne ai concetti di numero cardinale e

[1] Le idee di Cantor hanno stentato ad affermarsi: uno dei caratteri distintivi della nuova "mentalità insiemistica", vale a dire l'accettazione dell'infinito attuale, si veniva a scontrare non solo con le potenti correnti costruttiviste fiorenti in Germania nella seconda metà del secolo scorso (molte incomprensioni bloccarono la carriera accademica di Cantor ed ebbero effetti devastanti sulla sua salute mentale), ma anche con la secolare diffidenza verso l'infinito che aveva attraversato tutta la storia del pensiero occidentale. L'accettazione delle generalizzazioni cantoriane (in verità spesso accompagnate da considerazioni metafisiche e teologiche di scarsa presa presso il mondo matematico) si è lentamente realizzata quando queste hanno prodotto risultati che

ordinale e alla classificazione degli insiemi infiniti. In sintesi, si può affermare che con Cantor l'infinito attuale è divenuto oggetto di studio matematico, dato che gli insiemi infiniti, venendo assoggettati all'aritmetica del transfinito, sono entrati nel dominio della matematica tradizionalmente intesa come scienza della quantità e del contare.

L'aspetto "fondazionale" della teoria degli insiemi si è manifestato quando ci si è resi conto che tutti i principali concetti che intervengono nello sviluppo delle teorie matematiche (funzione, proprietà, relazione, numero, struttura, ecc.) possono venire ricondotti a pochissimi concetti assunti come primitivi (essenzialmente quelli di insieme e di appartenenza ad un insieme). Il linguaggio insiemistico, quindi, costituisce uno degli elementi unificanti dell'intero corpo delle discipline matematiche (e, come si vedrà nello sviluppo di questo corso di lezioni, anche della logica).

Varie sono le ragioni che consigliano di far precedere un corso di logica dall'analisi di alcune nozioni insiemistiche:

a) i linguaggi formali con cui si sviluppano i calcoli logici sono insiemi di formule generate partendo da un insieme di simboli e iterando l'applicazione di determinate operazioni: lo studio della sintassi dei calcoli logici si identifica quindi con quello delle proprietà di questi insiemi generati ricorsivamente;
b) il significato delle formule della logica viene fissato attraverso una semantica di tipo insiemistico[2];
c) la teoria dei modelli[3], uno dei settori più importanti della logica matematica, è fondata sulla teoria degli insiemi e ne costituisce per molti aspetti uno sviluppo.

Queste tematiche saranno tutte adeguatamente trattate nelle prossime lezioni di questo corso. In questo intervento dapprima riprenderemo in modo schematico i più elementari concetti insiemistici, per poi soffermarci più diffusamente sulla caratterizzazione degli insiemi finiti e infiniti e sulla gerarchia dei numeri cardinali, evidenziando sia gli aspetti rilevanti per il seguito, sia quelli aventi un intrinseco interesse didattico[4].

avevano importanti ricadute nella teoria delle funzioni di variabile reale e complessa, ossia nel settore da cui Cantor era partito nelle sue ricerche.

[2] L'elaborazione della semantica modellistica ad opera di Tarski negli anni trenta ha coronato una fase di ricerche logiche che ha condotto ad una sistemazione dei capitoli centrali della logica matematica rimasta pressoché inalterata fino ai giorni nostri.

[3] Dal punto di vista didattico lo studio della logica assume rilevanza non solo per il suo interesse intrinseco e come preliminare per l'introduzione dei concetti informatici, ma quando viene finalizzato all'analisi delle teorie matematiche formalizzate e dei loro modelli.

[4] Gli insiemi finiti e infiniti sono citati esplicitamente nei nuovi programmi di matematica per la scuola media superiore.

2. Definizioni preliminari

Un *insieme*, parafrasando la definizione originale di Cantor, è una qualsiasi collezione[5] di individui, detti *elementi* dell'insieme, determinati e distinti della realtà o del pensiero.

Nel seguito indicheremo insiemi generici con lettere maiuscole A, B, C,... e i loro elementi con lettere minuscole a, b, c,...[6].

La relazione che un individuo ha con un insieme di cui è elemento è detta *relazione di appartenenza* e si indica con il simbolo $\in$:

$a \in A$ significa che a è elemento di A (a appartiene ad A);
$a \notin A$ significa che a non è elemento di A (a non appartiene ad A).

Se due insiemi A e B sono uguali allora hanno gli stessi elementi; si assume anche il viceversa, ossia che due insiemi sono uguali se hanno gli stessi elementi, quindi:

$A = B$ se e solo se per ogni x, $x \in A$ se e solo se $x \in B$.

Negli insiemi, quindi, non contano l'ordine degli elementi né le eventuali ripetizioni di elementi. La relazione di uguaglianza è riflessiva, simmetrica e transitiva.

Per descrivere un insieme finito di solito si elencano gli elementi racchiudendoli fra parentesi graffe. Ad esempio:

$A = \{a, e, i, o, u\}$ $B = \{0, 2, 4, 6, 8\}$

Questa tecnica si adopera anche per insiemi infiniti quando è chiaro come si possono determinare tutti gli elementi dell'insieme:

$C = \{1, 3, 5, 7, 9, 11,...\}$ $D = \{1, 2, 4, 8, 16, 32,...\}$

Una seconda possibilità è quella di indicare una proprietà P(x) che caratterizza gli elementi dell'insieme; in tal caso si usa la notazione $\{x : P(x)\}$. Negli esempi precedenti:

[5] Questa definizione è palesemente circolare (collezione, famiglia, aggregato, ecc. sono sinonimi di insieme). Nell'ambito della teoria degli insiemi il concetto di insieme è assunto come primitivo. La nostra esposizione non è condotta con metodo assiomatico, ma sviluppata in modo intuitivo: l'assiomatizzazione della teoria degli insiemi è posteriore alla trattazione cantoriana e si intreccia con la problematica delle antinomie (sulla quale in questa sede preferiamo sorvolare). Rinviamo il lettore interessato a una esposizione della teoria assiomatica degli insiemi e a un approfondimento delle tematiche affrontate in questa lezione a G. Lolli, *Dagli insiemi ai numeri,* Boringhieri, Torino, 1994.

[6] Nonostante questa differenza notazionale, non vi è alcuna particolare differenza concettuale tra insiemi ed elementi, in quanto gli elementi di un insieme possono a loro volta essere insiemi e gli insiemi essere elementi di altri insiemi. Ad esempio, un certo triangolo è elemento dell'insieme dei poligoni del piano, ma può anche essere visto come insieme di punti di un piano. Un segmento è un insieme di punti, ma può essere elemento dell'insieme dei tratti di una spezzata, la quale, a sua volta, può essere considerata elemento dell'insieme delle spezzate del piano.

$A = \{x : x \text{ è vocale dell'alfabeto}\}$
$B = \{x : x \text{ è numero naturale pari minore di } 10\}$
$C = \{x : x \text{ è numero naturale dispari}\}$
$D = \{x : x \text{ è numero naturale potenza di } 2\}$

Vi è un unico insieme, detto *insieme vuoto*, che non ha elementi e che si indica con $\varnothing$.

Dati due insiemi A e B, si dice che A è *sottoinsieme* di B (è *contenuto in* B), in simboli, $A \subseteq B$, se e solo se ogni elemento di A è anche elemento di B:

$A \subseteq B$ se e solo se per ogni x, se $x \in A$, allora $x \in B$.

$\subseteq$ è detta anche relazione di *inclusione* tra insiemi, ed è riflessiva, transitiva ed antisimmetrica (se $A \subseteq B$ e $B \subseteq A$, allora $A = B$).

Si ha inoltre che, per ogni A, $A \subseteq A$ e $\varnothing \subseteq A$.

Le usuali operazioni insiemistiche di *unione* ($\cup$), *intersezione* ($\cap$) e *differenza* ($-$) si definiscono come segue:

$A \cup B = \{x : x \in A \text{ oppure (vel) } x \in B\}$
$A \cap B = \{x : x \in A \text{ e } x \in B\}$
$A - B = \{x : x \in A \text{ e } x \notin B\}$

(se $A \subseteq B$, $B - A$ è detto *complementare* di A in B e, quando B si può sottintendere, si indica anche con $-A$ o con $\overline{A}$).

Le operazioni di unione e di intersezione si estendono facilmente anche a più insiemi: l'unione di una collezione di più insiemi è l'insieme formato dagli elementi che appartengono ad almeno un insieme della collezione, mentre l'intersezione di una collezione di insiemi è formata dagli elementi che appartengono a tutti gli insiemi della collezione stessa.

Le proprietà algebriche delle operazioni insiemistiche sono molteplici. Richiamiamo prima quelle che caratterizzano la struttura di *algebra di Boole*. In un insieme non vuoto U, se A, B e C sono sottoinsiemi qualsiasi di U, si ha:

$A \cup B = B \cup A$ (proprietà commutativa di $\cup$)
$A \cap B = B \cap A$ (proprietà commutativa di $\cap$)
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (proprietà distributiva di $\cap$ su $\cup$)
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (proprietà distributiva di $\cup$ su $\cap$)
$A \cup \varnothing = A \quad A \cap U = A$
$A \cup (-A) = U \quad A \cap (-A) = \varnothing;$

a queste si possono ricondurre tutte le altre, tra le quali:

$A \cup (B \cup C) = (A \cup B) \cup C$ (proprietà associativa di $\cup$)
$A \cap (B \cap C) = (A \cap B) \cap C$ (proprietà associativa di $\cap$)
$A \cap (A \cup B) = A$ (legge di assorbimento)
$A \cup (A \cap B) = A$ (legge di assorbimento)
$A \cup A = A$ (legge di idempotenza)
$A \cap A = A$ (legge di idempotenza)
$A \cap \varnothing = \varnothing \quad A \cup U = U \quad -(-A) = A$
$A \cup B = -((-A) \cap (-B))$ (legge di De Morgan)
$A \cap B = -((-A) \cup (-B))$ (legge di De Morgan)

3. Insieme potenza e proprietà

Dato un insieme A, indichiamo con P(A) l'*insieme potenza* o *insieme delle parti* di A, ossia l'insieme dei sottoinsiemi di A:

$P(A) = \{x : x \subseteq A\}$

Il nome deriva dal fatto che, se A contiene n elementi, allora P(A) contiene 2^n elementi.

Ad esempio:

se $A = \{a, b\}$, allora $P(A) = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$;
se $A = \{a, b, c\}$, allora $P(A) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$[7].

L'insieme potenza riveste un ruolo fondamentale sia in teoria degli insiemi - come vedremo più avanti - sia in logica.

Nella semantica tarskiana si adotta l'atteggiamento *estensionale*, vale a dire si identificano le proprietà in un certo insieme A con i sottoinsiemi di A. Chiariamo questo importante aspetto con un esempio.

Sia $A = \{1, 3, 4, 5, 7, 8, 12, 15, 17, 20\}$.

Intuitivamente parlando, corrispondono a proprietà in A espressioni linguistiche che indicano attributi degli elementi di A, quali ad esempio, "essere pari", "essere dispari", "essere multiplo di 4", "essere primo", "essere minore di 10", ecc. Più precisamente si può dire che una proprietà è indicata da una espressione linguistica contenente una variabile

[7] Si osservi che tra gli elementi di P(A) figurano sempre $\varnothing$ e A e tutti i singoletti, ossia tutti i sottoinsiemi costituiti da un solo elemento di A. Occorre tener distinto l'elemento a e $\{a\}$ che è l'insieme che ha a come unico elemento:

$a \in A$ se e solo se $\{a\} \subseteq A$ se e solo se $\{a\} \in P(A)$.

individuale che diviene vera o falsa - che diviene una proposizione - quando la variabile è sostituita con il nome di un individuo; nell'esempio, "x è pari", "x è dispari", "x è multiplo di 4", "x è primo", "x è minore di 10", e così via. Secondo il punto di vista estensionale si identifica la proprietà in A con il sottoinsieme di A costituito dagli elementi che verificano la proprietà stessa:

"essere pari" = {4, 8, 12, 20}
"essere dispari" = {1, 3, 5, 7, 15, 17}
"essere multiplo di 4" = {4, 8, 12, 20}[8]
"essere primo" = {3, 5, 7, 17};
"essere minore di 10" = {1, 3, 4, 5, 7, 8}

Si assume inoltre che valga anche il viceversa, ossia che ogni sottoinsieme di A sia una proprietà in A (indipendentemente dall'aver associato ad esso una espressione linguistica del tipo ora considerato). Così, ad esempio, anche {1, 3, 7, 18} e {4, 7, 8, 15, 17} sono proprietà in A. In definitiva:

proprietà in A = sottoinsieme di A = elemento di P(A)

e quindi P(A) *è l'insieme delle proprietà in* A[9].

4. Prodotto cartesiano e relazioni

Dati due individui a e b, l'insieme {a, b} i cui elementi sono a e b è detto, se $a \neq b$, *coppia non ordinata* di a e b (come si è già osservato, {a, b} = {b, a}). Se a = b allora evidentemente {a, b} = {a} = {b}.

Se si vuole evidenziare l'ordine con cui si succedono i due elementi, ossia precisare che a precede b, si introduce la *coppia ordinata* (a, b), in cui a è il primo elemento e b il secondo elemento della coppia (e quindi, in generale, $(a, b) \neq (b, a)$)[10].

[8] Si può osservare che "essere pari" e "essere multiplo di 4" sono proprietà diverse come significato intuitivo - dal punto di vista *intensionale* - ma, in A, determinano lo stesso sottoinsieme e, quindi, sono *estensionalmente* coincidenti.

[9] Il punto di vista adottato ha il vantaggio di caratterizzare il concetto intuitivo di proprietà, precisandolo con una definizione esplicita e consentendo di applicare alle proprietà le operazioni insiemistiche. D'altra parte, esso crea un distacco fra le potenzialità espressive del linguaggio e le proprietà che, alla luce di quanto vedremo più avanti, è alla radice di molti dei risultati più significativi della logica matematica.

[10] Nelle trattazioni più approfondite il concetto di coppia ordinata non viene assunto come primitivo, ma viene definito ponendo, ad esempio:

$(a, b) = \{\{a\}, \{a, b\}\}$.

Si dimostra che in base a tale definizione vale la proprietà:

$(a, b) = (c, d)$ se e solo se $a = c$ e $b = d$,

che caratterizza appunto le coppie ordinate.

Dati due insiemi A e B, si dice *prodotto cartesiano* di A e di B, e si indica con $A \times B$, l'insieme delle coppie ordinate aventi primo elemento in A e secondo elemento in B:

$$A \times B = \{ (x, y) : x \in A \text{ e } y \in B\}$$

In particolare, $A \times A$ o A^2 è l'insieme delle coppie ordinate aventi primo e secondo elemento in A.

In analogia a quanto esposto nel paragrafo precedente a proposito delle proprietà, si adotta un corrispettivo atteggiamento a proposito delle relazioni. Una relazione a due argomenti (binaria) in A è indicata con una espressione linguistica contenente due variabili individuali che diviene vera o falsa quando le due variabili sono sostituite con i nomi di due elementi (non necessariamente distinti) di A.

Ad esempio, se A = {3, 4, 6, 7, 9, 12}, corrispondono a relazioni in A "x è minore di y", "x è primo con y", "x è multiplo ma non sottomultiplo di y", ecc. Adottando l'atteggiamento estensionale si identifica la relazione con l'insieme delle coppie che la soddisfano:

"essere minore di" = {(3,4), (3,6), (3,7), (3,9), (3,12), (4,6), (4,7), (4,9), (4,12), (6,7), (6,9), (6,12), (7,9), (7,12), (9,12)}
"essere primo con" = {(3,4), (4,3), (3,7), (7,3), (4,7), (7,4), (4,9), (9,4), (6,7), (7,6), (7,9), (9,7), (7,12), (12,7)}
"essere multiplo ma non sottomultiplo di" = {(6,3), (9,3), (12,3), (12,4), (12,6)}

La relazione in A si identifica quindi con un sottoinsieme del prodotto cartesiano A^2. Anzi, si definisce relazione binaria in A un qualsiasi sottoinsieme di A^2:

relazione (binaria) in A = sottoinsieme di A^2 = elemento di $P(A^2)$.

Con considerazioni analoghe si arriva ad identificare una relazione tra due insiemi A e B con un sottoinsieme del prodotto cartesiano $A \times B$.

Quanto finora esposto si può facilmente generalizzare.

Il concetto di *terna ordinata* di tre elementi a, b e c si può ricondurre a quello di coppia:

$$(a, b, c) = ((a, b), c),$$

quello di *quaterna ordinata* a quello di terna e coppia:

$$(a, b, c, d) = ((a, b, c), d)$$

e, procedendo induttivamente, si perviene al concetto di n-*pla ordinata* $(a_1, a_2, ..., a_n)$ di n individui $a_1, a_2, ..., a_n$.

Così A^3 è l'insieme delle terne ordinate di elementi di A, A^4 è l'insieme delle quaterne ordinate di elementi di A,..., A^n è l'insieme delle n-ple ordinate di elementi di A; una relazione a tre argomenti in A è un qualsiasi sottoinsieme di A^3, una relazione a quattro argomenti un qualsiasi sottoinsieme di A^4,..., una relazione a n argomenti un qualsiasi sottoinsieme di A^n.

5. Funzioni e operazioni

Dati due insiemi A e B, si dice *funzione* di *dominio* A e *codominio* B un qualsiasi sottoinsieme f di $A \times B$ avente la seguente proprietà:

per ogni $a \in A$ esiste uno ed un solo $b \in B$ tale che $(a, b) \in f$.

Anziché $(a, b) \in f$ si può scrivere, conformemente all'uso comune in matematica, $f(a) = b$ e, per indicare che f è una funzione da A a B, si usa sovente la notazione $f:A \to B$.

Si può anche dire che una funzione associa ad ogni elemento di A uno e un solo elemento di B. Dunque, per ogni $a \in A$ abbiamo:

(a) esiste un $b \in B$ tale che $(a, b) \in f$, ed inoltre,

(b) se $(a, b) \in f$ e $(a, c) \in f$, allora $b = c$ [11].

Ad esempio, se $A = \{3, 5, 6, 7\}$ e $B = \{4, 8, 9, 10, 12\}$, sono funzioni da A a B:

$f = \{(3,8), (5,9), (6,9), (7,4)\}$
$f' = \{(5,10), (3,12), (7,9), (6,4)\}$

mentre *non* sono funzioni da A a B:

$p = \{(3,4), (5,6), (6,12)\}$ [12]
$p' = \{(3,8), (5,9), (5,10), (6,12), (7,12)\}$ [13]

Se $C = \{4, 6, 8, 10, 12\}$ e $D = \{3, 5, 7, 10\}$ sono funzioni da C a D:

[11] Intuitivamente "una funzione è una legge che associa...". In analogia a quanto visto a proposito delle proprietà e delle relazioni, nella nostra definizione la funzione è identificata con l'insieme delle coppie soddisfacenti tale "legge", (il *grafico* della funzione-legge). Prescindendo dalla nozione imprecisa di "legge", nella teoria degli insiemi una funzione $f:A \to B$ è semplicemente un sottoinsieme di $A \times B$ avente le proprietà (a) e (b).

[12] Vale la condizione (b) ma non la (a). Si usa dire che p è una funzione *parziale*, avente dominio il sottoinsieme {3, 5, 6} di A.

[13] Le coppie (5,9) e (5,10) appartenenti alla relazione binaria p' hanno la stessa *ascissa* 5 ma diverse *ordinate*, contraddicendo la condizione (b).

$g = \{(4,3), (12,10), (8,10), (10,5), (6,10)\}$
$g' = \{(4,3), (6,7), (8,5), (10,10), (12,3)\}$

Se $E = \{2, 4, 6, 8, 10, 12\}$ e $F = \{6, 8, 12, 14, 18, 20\}$ sono funzioni da E a F:

$h = \{(2,20), (4,14), (6,12), (8,6), (10,18), (12,8)\}$
$h' = \{(2,6), (4,8), (6,8), (8,8), (12,20), (10,12)\}$

Si dice che una funzione $f:A \to B$ è *suriettiva* se e solo se, per ogni $b \in B$, esiste $a \in A$ tale che $(a, b) \in f$ (ossia, $f(a) = b$). Sono suriettive le funzioni g' e h degli esempi precedenti. Se gli insiemi A e B sono finiti e il numero di elementi di B è maggiore di quello degli elementi di A, allora non può esistere alcuna funzione suriettiva da A a B.

Si dice che una funzione $f:A \to B$ è *iniettiva* se e solo se, per ogni $a, a' \in A$, se $(a, b) \in f$ e $(a', b) \in f$, allora $a = a'$ (se $a \neq a'$, allora $f(a) \neq f(a')$), ossia ad elementi distinti di A corrispondono elementi distinti di B. Sono iniettive le funzioni f' e h degli esempi precedenti. Se gli insiemi A e B sono finiti e il numero di elementi di A è maggiore del numero di elementi di B, allora non può esservi alcuna funzione iniettiva da A in B.

Le funzioni che sono sia iniettive che suriettive sono dette *biiettive* o *corrispondenze biunivoche*. E' biiettiva la funzione h degli esempi precedenti. Se A e B sono finiti esiste una corrispondenza biunivoca fra essi se e solo se essi hanno lo stesso numero di elementi[14].

Accanto alle funzioni ad un argomento si possono considerare anche funzioni a più argomenti: spesso una *funzione a due argomenti* è denotata $f:A^2 \to B$, e una funzione *a n argomenti* è denotata $g:A^n \to B$. Nel caso in cui B coincide con A, f è detta essere un'*operazione binaria* su A, e g un' *operazione* a n *argomenti* su A.

6. Finito e infinito

Per illustrare meglio uno degli aspetti più rilevanti della teoria degli insiemi apriamo una breve parentesi. Senza entrare nel dettaglio delle ragioni che già a partire dalla cultura greca hanno favorito l'accettazione dell'infinito in matematica solo nella sua accezione potenziale - sostanzialmente secondo i canoni più volte espressi da Aristotele - si può ricordare che il rifiuto della concezione attuale dell'infinito era dovuto, almeno in grande misura, alle difficoltà evidenziate dall'insorgere di vari "paradossi dell'infinito" che già a partire da Zenone di Elea hanno sollevato interrogativi ai quali non era facile trovare risposte soddisfacenti. Ad esempio, consegue dall'esistenza di segmenti incommensurabili, quali la diagonale e il lato del quadrato o il lato e l'altezza di un triangolo equilatero, dimostrata già

[14] Questa considerazione è alla base della teoria dei numeri cardinali e della definizione insiemistica dei numeri naturali alle quali accenneremo nel prossimo paragrafo.

dalla scuola pitagorica, che i segmenti contengono infiniti punti[15]. Allora, dati due segmenti disuguali, da un lato bisognerà affermare che un'infinità di punti (quella del segmento più corto) è minore dell'altra infinità di punti, ma, d'altro lato, mediante una semplice proiezione centrale, è possibile porre in corrispondenza biunivoca i punti dei due segmenti, per cui sembra doversi concludere che le due infinità di punti si equivalgono. E' ben noto l'esempio di Galileo relativo agli insiemi dei numeri naturali e dei loro quadrati: da un lato essi si equivalgono (poiché ogni numero ha uno ed un solo quadrato), dall'altro i numeri quadrati perfetti costituiscono un sottoinsieme proprio dell'insieme dei numeri naturali (anzi, i numeri quadrati vanno sempre più diradandosi nella sequenza dei numeri naturali)[16].

Anche coloro i quali non hanno assunto un atteggiamento di rifiuto dell'infinito attuale, hanno tuttavia identificato l'infinito come un'entità assoluta, alla quale non ci si può rivolgere con consapevolezza a causa dei limiti della ragione umana, e nella cui analisi non si possono far intervenire le categorie con le quali si opera nell'ambito del finito. Cantor, invece, è riuscito ad individuare le tecniche mediante le quali "dominare" gli insiemi infiniti, assoggettandoli a relazioni d'ordine e ad una articolata aritmetica.

Lo strumento principale è quello della corrispondenza biunivoca: se tra due insiemi A e B esiste una corrispondenza biunivoca, i due insiemi sono detti *equipotenti* (e si scrive $A \cong B$) e ad essi viene attribuito, per definizione, lo stesso *numero cardinale* (se $A \cong B$, allora Card A = Card B); se si può istituire una corrispondenza iniettiva da A in B - ossia immergere A in B -, allora il numero cardinale di A è minore o uguale di quello di B (se $A \cong B' \subseteq B$, allora Card $A \leq$ Card B).

L'idea quindi è quella di individuare dei rappresentanti canonici (numeri cardinali) in modo che a ciascun insieme sia associato uno ed un solo numero cardinale che, per così dire, ne "misura la grandezza". Restando a livello intuitivo, si può identificare il numero cardinale di un insieme A con la collezione degli insiemi equipotenti ad A[17]. Se si pone:

$$0 = \text{Card } \varnothing$$
$$1 = \text{Card } \{0\}$$
$$2 = \text{Card } \{0, 1\}$$
$$3 = \text{Card } \{0, 1, 2\}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

[15] Se i segmenti contenessero un numero finito ancorché enorme di punti, il punto sarebbe sottomultiplo comune di tutti i segmenti, e non potrebbero esistere segmenti incommensurabili.

[16] E' interessante la risposta che Galileo offre all'insorgere di questo paradosso: "Queste son di quelle difficoltà che derivano dal discorrere che noi facciamo col nostro intelletto finito intorno a gl'infiniti, dandogli quegli attributi che noi diamo alle cose finite e terminate, il che penso sia inconveniente perché stimo che questi attributi di maggioranza, minorità ed eguaglianza non convenghino agl'infiniti, de i quali non si può dire, uno esser maggiore o minore o eguale all'altro" (G. Galilei, *Ediz. Naz.*, Vol. VIII, pp. 77-78).

[17] Essendo $\cong$ una relazione di equivalenza tra insiemi, i numeri cardinali risultano le classi di equivalenza rispetto ad essa. Nelle trattazioni più rigorose, condotte per evitare le antinomie, si sceglie in ciascuna classe di equivalenza un "rappresentante" che viene assunto come "numero cardinale" di tutti gli insiemi della classe.

e, in generale:

$$n+1 = \text{Card } \{0, 1, \ldots, n\}$$

si ottiene induttivamente la definizione insiemistica dei numeri naturali.

In teoria degli insiemi si ammette l'esistenza dell'insieme **N** dei numeri naturali (*assioma dell'infinito*), i.e. si accetta l'infinito attuale. Una volta introdotti i numeri naturali, si può definire *finito* un insieme equipotente a un numero naturale (e, quindi, *infinito* un insieme che non può essere posto in corrispondenza biunivoca con alcun numero naturale).

D'altra parte, seguendo Dedekind, si possono caratterizzare gli insiemi *infiniti* come quelli che possono essere posti in corrispondenza biunivoca con una loro parte propria, facendo così divenire una *definizione* quella che, come si è accennato, è stata considerata per secoli una caratteristica paradossale che rendeva razionalmente impraticabile l'infinito.

In questa ottica un insieme è *finito* se e solo se non può essere posto in corrispondenza biunivoca con una sua parte propria[18].

L'insieme **N** dei numeri naturali è infinito secondo entrambe le accezioni menzionate in precedenza in quanto equipotente a una sua parte propria (esempio di Galileo) e non equipotente ad alcun numero naturale[19]. Il suo numero cardinale si indica solitamente con $\aleph_0$ (aleph con zero).

7. Insiemi numerabili

Si dicono *numerabili* gli insiemi aventi numero cardinale $\aleph_0$, ossia che possono essere posti in corrispondenza biunivoca con l'insieme **N** dei numeri naturali.

[18] Queste due diverse caratterizzazioni degli insiemi finiti e infiniti sono equivalenti se si assume un principio (detto *assioma di scelta*) sulla cui accettazione si sono accese numerose discussioni, ma che riveste un ruolo fondamentale per lo sviluppo della matematica dell'infinito (e anche della logica matematica). Esso afferma, in una delle sue numerose formulazioni equivalenti, che, data una collezione qualsiasi di insiemi non vuoti a due a due disgiunti (con intersezione vuota), esiste un insieme, detto *insieme di scelta*, che ha in comune uno ed un solo elemento con ciascun insieme della collezione. Un altro modo di formulare l'assioma di scelta consiste nell'affermare che esiste una *funzione di scelta* su P(A), vale a dire una funzione f che a ciascun sottoinsieme non vuoto X di A (A insieme non vuoto) associa un elemento del sottoinsieme stesso, ossia tale che $f(X) \in X$.

[19] Dato un insieme infinito secondo Dedekind, esso non può essere equipotente a un numero naturale, altrimenti anche quest'ultimo risulterebbe equipotente a una sua parte, cosa che si dimostra essere impossibile. Per dimostrare l'implicazione inversa occorre impiegare l'assioma di scelta (cfr. la nota precedente). Sia A un insieme non equipotente a un numero naturale. Si sceglie un elemento x_0 in A, poi un elemento x_1 in $A - \{x_0\}$, poi un elemento x_2 in $A - \{x_0, x_1\}$, e così via, sfruttando una funzione di scelta in P(A). Si costruisce in tal modo una successione $X = \{x_0, x_1, x_2, x_3, \ldots\}$ di elementi di A; tale processo, infatti, non si arresta poiché, ad ogni stadio, $A - \{x_0, x_1, \ldots, x_n\}$ non è vuoto (se fosse $A - \{x_0, x_1, \ldots, x_n\} = \varnothing$, allora $A = \{x_0, x_1, \ldots, x_n\}$ e A sarebbe equipotente ad un numero naturale) e la funzione di scelta individua il successivo elemento della successione. Si ha allora che la funzione f che lascia fissi gli elementi di $A - X$ e associa ad ogni elemento di X il suo successivo ($f(x_n) = x_{n+1}$) è una corrispondenza biunivoca fra A e la sua parte propria $A - \{x_0\}$. Quindi, un insieme non equipotente a un numero naturale è equipotente a una sua parte propria.

Il ragionamento qui proposto consente anche di concludere che, conformemente all'intuizione, l'insieme dei numeri naturali è l'insieme infinito "più piccolo", e il suo numero cardinale il minimo cardinale transfinito.

Se A è un insieme numerabile e f una corrispondenza biunivoca $f : \mathbf{N} \to A$, ponendo $f(0) = a_0$, $f(1) = a_1$, $f(2) = a_2$,..., $f(n) = a_n$,..., dato che f è biiettiva, si ha:

$$A = \{a_0, a_1, a_2, ..., a_n, ...\},$$

ossia gli elementi di A possono essere posti in una successione infinita senza ripetizioni (e viceversa, se gli elementi di un insieme possono essere posti in una successione infinita senza ripetizioni, allora l'insieme è numerabile). Si constata allora facilmente che sono numerabili l'insieme dei numeri pari, dei numeri dispari, dei numeri primi (e, in generale, un sottoinsieme infinito di un insieme numerabile è numerabile), come pure l'insieme **Z** dei numeri interi che si può scrivere $\{0, +1, -1, +2, -2, +3, -3, ...\}$.

Sia ora $A_0, A_1, A_2, ..., A_n, ...$ una successione infinita senza ripetizioni di insiemi numerabili e poniamo in successione gli elementi di ciascuno di essi indicando con a_{hk} il k-esimo elemento dell'insieme A_h:

A_0 $a_{00}, a_{01}, a_{02}, a_{03}, a_{04}, ...$
A_1 $a_{10}, a_{11}, a_{12}, a_{13}, a_{14}, ...$
A_2 $a_{20}, a_{21}, a_{22}, a_{23}, a_{24}, ...$
A_3 $a_{30}, a_{31}, a_{32}, a_{33}, a_{34}, ...$
... .
... .

Gli elementi che figurano in questo quadro doppiamente infinito (verso destra e verso il basso) possono essere disposti in un'unica successione mettendo prima gli elementi con minore somma degli indici e, a parità di somma degli indici, mettendo prima quello con primo indice minore:

$a_{00}, a_{01}, a_{10}, a_{02}, a_{11}, a_{20}, a_{03}, a_{12}, a_{21}, a_{30}, a_{04}, a_{13}, a_{22}, a_{31}, a_{40}, a_{05}, a_{14}, a_{23}, a_{32}, ...$

L'unione di tutti gli insiemi della successione, essendo un sottoinsieme infinito di quest'ultima successione, è numerabile:

l'unione di una collezione numerabile di insiemi numerabili è un insieme numerabile.

Analogamente si vede che il prodotto cartesiano di due insiemi numerabili è ancora un insieme numerabile. Infatti, se i due insiemi sono:

$$A = \{a_0, a_1, a_2, ..., a_n, ...\} \qquad B = \{b_0, b_1, b_2, ..., b_n, ...\}$$

si può scrivere il prodotto cartesiano $A \times B$ come un quadro doppiamente infinito analogo al precedente in cui al posto di a_{hk} figura la coppia ordinata (a_h, b_k).

Pertanto, è numerabile l'insieme delle frazioni (essendo l'unione dell'insieme numerabile delle frazioni di denominatore 1, di quello delle frazioni di denominatore 2, di quello delle frazioni di denominatore 3, ecc.) e, quindi, è numerabile l'insieme dei numeri razionali.

Sia S l'insieme delle successioni infinite dei numeri 0 e 1 che, da un certo punto in poi, sono costituite da tutti 0. Sono elementi di S, ad esempio:

$s_1 = 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, ...$
$s_2 = 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, ...$

Dimostriamo che *l'insieme* S *è numerabile* individuando una funzione $f: S \to \mathbf{N}$ che sia biiettiva.
Un generico elemento s di S si può scrivere:

$$s = b_0, b_1, b_2, ..., b_n, ...$$

dove ciascun b_h è 0 o 1 ed è costantemente eguale a 0 da un certo punto in poi.

Poniamo $f(s) = \sum_{h=0}^{\infty} b_h \cdot 2^h$ (la somma contiene un numero finito di addendi).

Nei due esempi precedenti risulta:

$f(s_1) = 1\cdot2^0 + 1\cdot2^2 + 1\cdot2^3 + 1\cdot2^5 + 1\cdot2^6 + 1\cdot2^7 + 1\cdot2^9 = 749$
$f(s_2) = 1\cdot2^1 + 1\cdot2^2 + 1\cdot2^3 + 1\cdot2^6 + 1\cdot2^8 + 1\cdot2^9 = 846$

Si vede facilmente che a successioni diverse corrispondono numeri naturali diversi, per cui f è iniettiva. Per vedere che f è suriettiva basta osservare che, dato un qualsiasi numero naturale n, esprimendolo in forma binaria, ponendo s la successione delle cifre in rappresentazione binaria di n in ordine inverso seguita da infiniti 0, allora f(s) = n.

Ad esempio:

se n = 270, allora n = 100001110 e s = 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0,...
se n = 92, allora n = 1011100 e s = 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0,...
se n = 35, allora n = 100011 e s = 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0,...

L'insieme S ora considerato è in corrispondenza biunivoca con l'insieme dei sottoinsiemi *finiti* dell'insieme dei numeri naturali: basta considerare la funzione f che ad

ogni sottoinsieme finito A di **N** associa la successione s di 0 e 1 che ha gli 1 in corrispondenza dei posti degli elementi che appartengono al sottoinsieme, vale a dire: $s_n = 1$ se e solo se $n \in A$.

Ad esempio:

se B = {2, 4, 10, 11, 13}, allora f(B) = 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0,...
se C = {1, 3, 5, 7, 9, 11}, allora f(C) = 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0,...

Trattandosi di sottoinsiemi finiti, ad ogni sottoinsieme finito è associata una successione di 0 e 1 che contiene tutti 0 da un certo punto in poi. Si riconosce subito che f è una corrispondenza biunivoca e, quindi, *l'insieme dei sottoinsiemi finiti di* **N** *è numerabile.*

Sia ora K l'insieme di tutte le n-ple di numeri naturali, per n =1,2,3,....
Sono elementi di K, ad esempio, (2, 4, 5), (3, 4, 0, 2), (7, 1, 2, 0, 2, 1).
Dimostriamo che *l'insieme* K *è numerabile.*
Ad ogni numero naturale n associamo la sequenza n* costituita da n+1 uno.

Ad esempio:

5* = 1, 1, 1, 1, 1, 1; 8* = 1, 1, 1, 1, 1, 1, 1, 1, 1; 0* = 1; 1* = 1, 1; 3* = 1, 1, 1, 1

Dato un qualsiasi elemento di K, $(n_0, n_1, n_2,..., n_k)$, ad esso associamo la sequenza di 0 e 1 così definita:

$$f((n_0, n_1, n_2,..., n_k)) = 0, n_0^*, 0, n_1^*, 0, n_2^*, 0, ..., 0, n_k^*, 0, 0, 0, 0, 0,...$$

Ad esempio:

f((2, 4, 5)) = 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0,...
f((3, 4, 0, 2)) = 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0,...
f((7, 1, 2, 0, 2, 1)) = 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0,...

Evidentemente f è una funzione iniettiva da K nell'insieme S prima considerato, per cui, essendo S numerabile e K in corrispondenza biunivoca con un sottoinsieme infinito di S, anche K è numerabile.

Dalla numerabilità di K segue subito quella dell'insieme delle equazioni polinomiali a coefficienti interi e, quindi, dato che ogni siffatta equazione, per il teorema fondamentale dell'algebra, ha un numero finito di radici, si ottiene facilmente la numerabilità dell'insieme dei numeri reali (o complessi) algebrici, ossia quei numeri, come la radice di 2, che sono soluzione di una equazione polinomiale a coefficienti interi. Sempre dalla numerabilità di K

segue facilmente che è numerabile l'insieme delle formule che sono stringhe finite di simboli di un alfabeto numerabile[20].

8. Insiemi più che numerabili

Una delle scoperte più importanti di Cantor è l'individuazione di insiemi infiniti che hanno cardinalità maggiore di $\aleph_0$.

Consideriamo l'insieme S° costituito da tutte le successioni dei numeri 0 e 1 (senza la restrizione che da un certo punto in poi siano costituite da tutti 0).
Dimostriamo che S° *è più che numerabile.*
Procediamo per assurdo, supponendo che S° sia numerabile. Si può porre allora:

$$S^\circ = \{s_0, s_1, s_2,..., s_n,...\}$$

e nella successione in parentesi devono figurare *tutte* le possibili successioni di 0 e di 1. Indicando con s_{hk} il k-esimo elemento (che è o 0 o 1) della successione s_h, si può scrivere:

s_0 $s_{00}, s_{01}, s_{02}, s_{03}, s_{04},\cdots$
s_1 $s_{10}, s_{11}, s_{12}, s_{13}, s_{14},\cdots$
s_2 $s_{20}, s_{21}, s_{22}, s_{23}, s_{24},\cdots$
s_3 $s_{30}, s_{31}, s_{32}, s_{33}, s_{34},\cdots$
... .
... .

Consideriamo la successione di 0 e 1 costituita dagli elementi della diagonale principale del quadro doppiamente infinito:

$$s_{00}, s_{11}, s_{22}, s_{33},..., s_{nn},\cdots$$

Definiamo una nuova successione $d = d_0, d_1, d_2, d_3,..., d_n,...$ ponendo: $d_n = 1 - s_{nn}$ (ossia scambiando nella successione precedente ogni 0 con 1 e ogni 1 con 0).
Si constata immediatamente che d è una successione di 0 e 1 diversa da tutte le s_n: è diversa da s_0 almeno per il primo elemento (essendo $d_0 \neq s_{00}$), è diversa da s_1 almeno per il secondo elemento (essendo $d_1 \neq s_{11}$), è diversa da s_2 almeno per il secondo elemento (essendo $d_2 \neq s_{22}$), e così via. Ma ciò è assurdo poiché nella successione $s_0, s_1, s_2, s_3,...$ avrebbero dovuto comparire tutte le successioni di 0 e 1.

[20] Nel §4 della prossima lezione introdurremo il linguaggio della logica dei predicati: esso è costituito da un alfabeto (che è un insieme numerabile di simboli), dai termini e dalle formule ben formate, i quali sono stringhe finite di simboli dell'alfabeto. Pertanto, l'insieme dei termini e l'insieme delle formule ben formate sono insiemi numerabili.

S° è un esempio di insieme infinito che non si può porre in corrispondenza biunivoca con l'insieme **N** dei numeri naturali e, quindi, ha un numero cardinale maggiore di $\aleph_0$ (è più che numerabile).

Si dimostra facilmente che S° è in corrispondenza biunivoca con l'insieme potenza P(**N**) dell'insieme **N** dei numeri naturali.

Basta considerare la funzione f già introdotta nel paragrafo precedente che ad ogni sottoinsieme A di **N** associa la successione s di 0 e 1 che ha gli 1 in corrispondenza dei posti degli elementi che appartengono al sottoinsieme: $s_n = 1$ se e solo se $n \in A$.

Trattandosi di sottoinsiemi qualsiasi di **N** (e non solo di quelli finiti), le successioni corrispondenti di 0 e 1 non sono più costituite da tutti 0 da un certo punto in poi; e.g.:

se B = {1, 3, 5, 7, 9, 11,...}, allora f(B) = 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1,...
se C = {0, 5, 10, 15,...}, allora f(C) = 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0,...

Si riconosce facilmente che f è una corrispondenza biunivoca. Quindi:

P(**N**) *è un insieme più che numerabile* e, ricordando quanto si è visto nel §3, *l'insieme delle proprietà di* **N** *è più che numerabile*[21].

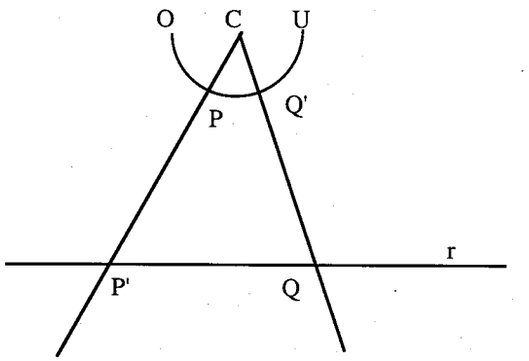
Si dimostra inoltre che S° è in corrispondenza biunivoca con l'insieme **R** dei numeri reali compresi fra 0 e 1 [22]. Vi è poi una corrispondenza biunivoca tra l'insieme dei numeri reali compresi fra 0 e 1 e l'insieme **R** dei numeri reali, ad esempio quella individuata dalla legge $y = \text{tang } \pi (x - 1/2)$. Quindi:

l'insieme **R** *dei numeri reali è più che numerabile.*

Un modo geometrico per visualizzare la corrispondenza fra i numeri reali compresi fra 0 e 1 e l'insieme **R** dei numeri reali è il seguente. E' noto che, fissando su una retta r un'origine O e un punto unità U, ad ogni punto della retta è associato uno ed un solo numero reale, e viceversa, ossia che vi è una corrispondenza biunivoca tra i numeri reali e i punti di una retta (e quindi *l'insieme dei punti di una retta è più che numerabile).* Consideriamo il segmento unitario OU, incurviamolo a semicirconferenza di centro C come nella figura seguente, e consideriamo la corrispondenza tra OU e la retta così definita: ad ogni punto P di OU associamo il punto P' di r ottenuto intersecando r con la semiretta CP (e ogni punto Q di r

[21] Questa considerazione, unita a quanto detto alla fine del §7, evidenzia l'impossibilità dei linguaggi usuali (con un insieme numerabile di espressioni) di esprimere tutte le proprietà dei numeri naturali.
[22] Basta osservare che, se si usa la base 2, ossia solo le cifre 0 e 1, ogni numero reale compreso tra 0 e 1 ha uno sviluppo decimale infinito che è una successione di S° e, viceversa, ogni successione di S° è lo sviluppo di un numero reale compreso fra 0 e 1. Si può sistemare questa corrispondenza in modo che sia biunivoca e, quindi, l'insieme dei numeri reali compresi tra 0 e 1 è più che numerabile.

è il corrispondente del punto Q' della semicirconferenza ottenuto intersecandola con la semiretta CQ).



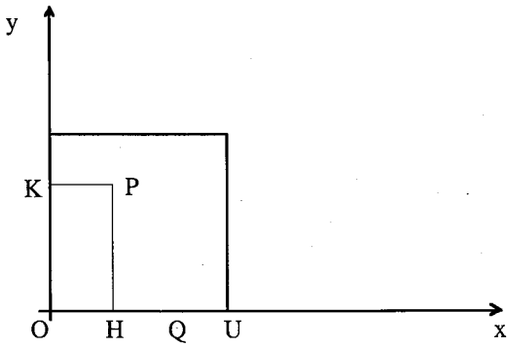
Si vede facilmente che si tratta di una corrispondenza biunivoca e, data l'arbitrarietà della scelta dei punti O e U, si può concludere che vi sono tanti punti su una retta quanti in ogni suo segmento.

Pertanto, oltre alla cardinalità del numerabile, vi è la cardinalità infinita, detta *potenza del continuo*, dell'insieme **R** dei numeri reali, dell'insieme dei punti di una retta, dell'insieme dei punti di un segmento, dell'insieme P(**N**), dell'insieme delle proprietà dei numeri naturali.

Molti altri insiemi matematicamente significativi hanno la potenza del continuo. Poiché si dimostra facilmente che, togliendo da un insieme che ha la potenza del continuo un insieme finito o numerabile di elementi, si ottiene un insieme che ha ancora la potenza del continuo, dato che l'insieme dei numeri razionali è numerabile, ne segue che l'insieme dei numeri irrazionali ha la potenza del continuo e, dato che l'insieme dei numeri algebrici è numerabile, ne segue che l'insieme dei numeri trascendenti ha la potenza del continuo. Si dimostra poi che hanno la potenza del continuo l'insieme dei numeri complessi, l'insieme dei punti del piano e dello spazio (anche a più dimensioni) e delle figure piane e solide, delle funzioni continue di variabile reale.

Vediamo, in breve, come si può stabilire che *l'insieme dei punti di un segmento e l'insieme dei punti del quadrato costruito su tale segmento sono equipotenti.*

Costruiamo su due lati del quadrato due assi cartesiani e assumiamo il lato come unità di misura. Vediamo come è possibile definire una corrispondenza biunivoca tra i punti del quadrato e i punti del segmento OU. Come è ben noto ogni punto P del quadrato è individuato da due numeri reali, l'ascissa e l'ordinata, che sono le misure dei segmenti OH e OK rispetto a OU.



Ogni punto P del quadrato individua due numeri reali, l'ascissa e l'ordinata, che, per come abbiamo posto gli assi, sono numeri reali compresi fra 0 e 1:

ascissa: $0,a_1a_2a_3a_4.....$
ordinata: $0,b_1b_2b_3b_4.....$

Al punto P associamo il punto Q del segmento OU che ha come ascissa il numero reale $0,a_1b_1a_2b_2a_3b_3a_4b_4.....$

Il punto Q' del segmento OU di ascissa $0,c_1c_2c_3c_4c_5c_6c_7c_8.....$ è immagine del punto P' del quadrato avente ascissa $0,c_1c_3c_5c_7.....$ e ordinata $0,c_2c_4c_6c_8.....$ In tal modo, a parte alcune precisazioni sulle rappresentazioni decimali dei numeri reali sulle quali si può ora sorvolare, si è istituita una corrispondenza biunivoca tra i punti del quadrato e quelli del suo lato[23].

Questo procedimento è immediatamente generalizzabile considerando i punti di un cubo e quelli del suo spigolo o i punti di un ipercubo a n dimensioni e quelli del suo spigolo. Ciò significa che il concetto di *dimensione*, mediante il quale distinguiamo fra loro le figure lineari, piane e solide, non si può basare su una valutazione quantitativa del "numero" di punti delle figure stesse.

Uno dei risultati più significativi delle ricerche di Cantor è che, oltre ai due tipi di infinito - il numerabile e il continuo - su cui ci siamo brevemente soffermati, ne esistono infiniti altri.

[23] Una dimostrazione più rigorosa fa ricorso al teorema di Cantor-Bernstein (che enunceremo nel § 9). Vedi ad esempio le osservazioni di V. Villani in *Archimede* **41** (1994), pp. 119-121.

Ciò consegue da un importante teorema:

per ogni insieme A, Card A < Card P(A).

Vediamo la dimostrazione del teorema.
Se consideriamo la funzione h che ad ogni $a \in A$ associa l'elemento $\{a\}$ di P(A) ($h:A\to P(A)$, $h(a) = \{a\}$), essa è evidentemente iniettiva, per cui:

Card A ≤ Card P(A).

Per dimostrare che la disuguaglianza vale in senso stretto bisogna dimostrare che tra A e P(A) non si può istituire alcuna corrispondenza biunivoca. Procediamo per assurdo supponendo che esista una $f:A\to P(A)$ che sia biiettiva e, quindi, in particolare, suriettiva.
Consideriamo il sottoinsieme B di A ($B \in P(A)$) costituito dagli elementi di A che non appartengono al sottoinsieme ad essi associato dalla funzione f:

$B = \{x \in A: x \notin f(x)\}$.

Essendo, per ipotesi, f suriettiva, esiste $b \in A$ tale che $f(b)= B$ (*).
Si ricava allora la seguente contraddizione:

$b \in B$ se e solo se $b \notin f(b)$ (per def. di B) se e solo se $b \notin B$ (per (*))

e la dimostrazione per assurdo è conclusa.

Sfruttando quanto appena dimostrato si ottiene facilmente una gerarchia crescente di numeri cardinali di grandezza sempre maggiore:

0,1,2,3,....., Card **N**, Card P(**N**), Card P(P(**N**)), Card P(P(P(**N**))),...

Si dimostra facilmente che l'unione **M** della collezione **N**, P(**N**), P(P(**N**)),... ha numero cardinale maggiore di tutti quelli della sequenza precedente, per cui Card **M** dà origine ad una nuova sequenza crescente di numeri cardinali:

Card **M**, Card P(**M**), Card P(P(**M**)),...

e così via, e si ottiene una gerarchia di numeri cardinali transfiniti assai più ricca di quella dei cardinali finiti[24].

[24] Molte ricerche nella teoria degli insiemi riguardano i cosiddetti grandi cardinali, ossia numeri cardinali che occupano posizioni molto elevate nella gerarchia dei numeri cardinali.

Un problema tuttorà al centro di discussioni consiste nello stabilire se tra **N** e P(**N**) (in generale, tra A e P(A), se A è infinito) esistano insiemi di cardinalità intermedia. La situazione è per così dire aperta, nel senso che, come dimostrato da Gödel e da Cohen, è coerente con gli usuali principi della teoria degli insiemi sia assumere che tali insiemi non esistano, e allora si dice che si accetta l'*ipotesi del continuo* (in generale, l'*ipotesi generalizzata del continuo*), sia assumere che esistano.

9. Cenni all'aritmetica dei numeri cardinali e ai numeri ordinali

Nella parte finora svolta abbiamo visto come i numeri cardinali indichino una "misura della grandezza" degli insiemi. D'altra parte, affinché si possa parlare di "numeri", occorre introdurre una relazione d'ordine e delle operazioni fra essi.

Diamo ora qualche cenno sulla confrontabilità dei numeri cardinali.
Dati due insiemi qualsiasi A e B, a priori si possono verificare i seguenti casi:

(a) A e B sono equipotenti (e allora Card A = Card B);
(b) A è equipotente a un sottoinsieme di B, ma non viceversa (e allora Card A<Card B);
(c) B è equipotente a un sottoinsieme di A, ma non viceversa (e allora Card A>Card B);
(d) A è equipotente a un sottoinsieme di B e B è equipotente a un sottoinsieme di A. In questo caso si dimostra (*teorema di Cantor-Bernstein*) che A è equipotente a B (e allora Card A = Card B);
(e) A non è equipotente a un sottoinsieme di B e B non è equipotente a un sottoinsieme di A.

Se si potesse realizzare il caso (e) non vi sarebbe alcuna possibilità di porre a confronto i due insiemi A e B e Card A e Card B non sarebbero confrontabili (né uguali, né uno minore dell'altro). Si può dimostrare[25] che il caso (e) non si può verificare e quindi che, dati due numeri cardinali, essi sono sempre confrontabili (*legge di tricotomia*); detto in altre parole, l'ordinamento dei numeri cardinali è *totale.*

Mediante le operazioni insiemistiche di unione, prodotto cartesiano e elevamento a potenza (A^B è l'insieme delle funzioni di dominio B e codominio A) si estendono ai numeri cardinali transfiniti le operazioni di addizione, moltiplicazione e elevamento a potenza, le quali godono della gran parte delle proprietà formali delle corrispettive operazioni fra numeri naturali, oltre ad altre peculiari, come ad esempio

$$\aleph_0 + \aleph_0 = \aleph_0 + n = \aleph_0;\ \aleph_0 \times \aleph_0 = \aleph_0 \times n = \aleph_0.$$

Se i numeri cardinali transfiniti consentono di valutare comparativamente le quantità di elementi degli insiemi infiniti, essi non costituiscono tuttavia una generalizzazione del

[25] Usando in modo essenziale l'assioma di scelta .

processo del contare. Per esempio, la legge di cancellazione cessa di valere ($\aleph_0 = \aleph_0 + 1 = \aleph_0 + 2 = \aleph_0 + 3 = ...$).

Due insiemi totalmente ordinati A e B sono *simili* se tra essi esiste una corrispondenza biunivoca f *monotona*: se $x \leq y$ allora $f(x) \leq f(y)$. Allora, automaticamente, anche la funzione inversa di f sarà monotona. Si dice che due insiemi simili hanno lo stesso *tipo d'ordine*. Mentre tutti gli insiemi totalmente ordinati con n elementi sono simili (per cui talora si dice che i numeri naturali rivestono sia la funzione cardinale, sia la funzione ordinale), le cose cambiano quando si considerano insiemi infiniti. Se ordiniamo l'insieme **N** dei numeri naturali nei due modi seguenti:

0, 1, 2, 3, 4,.....
1, 2, 3, 4,.....:., 0

si constata immediatamente che nessuna corrispondenza biunivoca monotona può trasformare il primo nel secondo ordinamento: infatti nel secondo caso vi è un massimo elemento e due elementi non hanno predecessore immediato; invece nel primo caso non vi è massimo elemento e vi è un solo elemento senza predecessore immediato. Pur dissimili, i due ordinamenti rimangono tuttavia confrontabili in questo senso: il primo è simile al segmento iniziale del secondo ottenuto abolendo l'elemento finale 0. La funzione successore fornisce l'opportuna corrispondenza.

Non sempre due insiemi totalmente ordinati sono confrontabili (non sempre, cioè, essi sono simili o uno è simile a un segmento iniziale dell'altro). Se consideriamo l'insieme **N** dei numeri naturali ordinato nei due modi seguenti:

0, 1, 2, 3, 4,.....
.....4, 3, 2, 1, 0

si riconosce immediatamente che tra i due ordinamenti non si può istituire alcun confronto: infatti ogni elemento del primo insieme ha un numero finito di predecessori e infiniti successori, mentre ogni elemento del secondo insieme ha infiniti predecessori e un numero finito di successori.

Per salvaguardare la possibilità di un confronto e per generalizzare il procedimento del contare, si considerano i tipi d'ordine degli insiemi *bene ordinati*, ossia quegli insiemi totalmente ordinati in cui ogni sottoinsieme non vuoto ha un minimo elemento. Essi vengono detti *numeri ordinali.*

Il numero ordinale dell'insieme vuoto si indica con 0. Per semplicità di notazione, per ogni $n \geq 1$ il numero ordinale dell'insieme totalmente ordinato $\{0, 1,..., n-1\}$ si indica con n. Il numero ordinale di **N** con l'ordine usuale (che è un buon ordinamento) si indica con

ω. Non è difficile vedere che anche l'insieme $\{1, 2, 3, 4,...\}$ ha numero ordinale ω. Con $\omega + 1$ si denota il numero ordinale dell'insieme totalmente ordinato 1, 2, 3, 4,....., 0, ottenuto aggiungendo all' ordinale ω un elemento finale. Si noti che $\omega + 1$ ha massimo elemento, mentre ω non l' ha.

Partendo dall' ordinale 0, tutti gli ordinali si possono ottenere mediante due principi generatori. Il primo è il passaggio al successivo: posponendo un nuovo elemento a un (insieme ben ordinato avente) ordinale α si ottiene un nuovo (insieme bene ordinato cui si associa il successivo numero) ordinale $\alpha + 1$. Il secondo principio generalizza la costruzione dell' ordinale ω come "limite" degli ordinali finiti: la sua applicazione richiede alcuni dettagli preliminari su cui qui non ci soffermiamo (ad esempio, la relazione "essere simile a un segmento iniziale di" induce un buon ordinamento in ogni insieme di ordinali). Per i nostri scopi è sufficiente formulare il secondo principio in questo modo: qualunque insieme non vuoto X di ordinali che non abbia un massimo elemento individua un nuovo ordinale $\sup X$, il minimo ordinale maggiore di ciascun ordinale in X.

Così, ad esempio, $\omega = \sup \{ 0, 1, 2, 3,... \}$.

Da ω, con il primo principio, si ottengono gli ordinali: ω, $\omega + 1$, $\omega + 2$, $\omega + 3$,...

Con il secondo principio si ottiene l'ordinale $\omega + \omega = \sup\{\omega, \omega + 1, \omega + 2, \omega + 3,...\}$, che è caratterizzabile anche come l'ordinale dell'insieme bene ordinato $\{0, 1, 2, 3,..., \omega, \omega + 1, \omega + 2, \omega + 3,... \}$. Inoltre $\omega + \omega$ è il numero ordinale dell' insieme **N** bene ordinato nel modo seguente: 0, 2, 4, 6,..., 1, 3, 5, 7,... . Anziché $\omega + \omega$ si può scrivere equivalentemente $\omega \times 2$. Analogamente si ottengono gli ordinali $\omega \times 2$, $\omega \times 3$, $\omega \times 4$,... e si scrive $\sup\{\omega \times 2, \omega \times 3, \omega \times 4,...\} = \omega \times \omega = \omega^2$. Si procede poi con ω^3, ω^4, ecc.

Se si considerano tutti i modi in cui si può bene ordinare l'insieme **N** dei numeri naturali, si ottiene l'insieme degli ordinali numerabili e si dimostra che il suo limite, che si indica con $\aleph_1$, è un ordinale più che numerabile[26].

Anche per i numeri ordinali si possono introdurre delle operazioni aritmetiche di addizione, moltiplicazione e elevamento a potenza (peraltro già sfruttate nelle precedenti notazioni), sulle quali non è ora il caso di soffermarsi. Segnaliamo solo che queste operazioni non godono di alcune delle proprietà formali delle corrispettive operazioni aritmetiche.

Ad esempio, non vale la proprietà commutativa dell'addizione; si ha che $1 + \omega \neq \omega + 1$. Infatti, per definizione di addizione ordinale, $1 + \omega$ è l' ordinale dell'insieme ottenuto posponendo a un singolo elemento una successione simile a **N** con l' ordine usuale; siccome tale insieme è simile a **N** segue che $1 + \omega = \omega$; invece, come già notato in precedenza, $\omega + 1$, il successivo di ω, è un ordinale diverso da ω.

[26] Se si accetta l'ipotesi del continuo tale ordinale ha la potenza del continuo e corrisponde ad un buon ordinamento dell'insieme **R** dei numeri reali, i quali, a loro volta, possono essere bene ordinati in molteplici modi cui corrisponde un insieme di ordinali di cardinalità maggiore della potenza del continuo, e così via.

Così $\omega \times 2 \neq 2 \times \omega$, e non vale la proprietà commutativa della moltiplicazione. Infatti, per definizione di moltiplicazione ordinale, $2 \times \omega$ è il numero ordinale dell'insieme totalmente ordinato $\{0', 0'', 1', 1'', 2', 2'',...\}$ che è simile a **N**. Dunque $2 \times \omega = \omega$. Invece, come abbiamo visto, $\omega \times 2 = \omega + \omega \neq \omega$ si ottiene prendendo due copie di **N** e ponendole una dopo l'altra.

Linguaggio naturale e Formalizzazione
Connettivi e Quantificatori

DARIO PALLADINO
Dipartimento di Filosofia
Università di Genova
via Balbi 4
16126, Genova

1. Premessa

Prima di affrontare il tema centrale di questa lezione, vale a dire la presentazione del linguaggio artificiale con il quale sono espressi i calcoli logici, è opportuna qualche considerazione preliminare sulla natura della logica che evidenzi come mai si faccia esplicito riferimento al linguaggio (piuttosto che al pensiero) e che illustri il significato e il ruolo della formalizzazione. Infatti, non solo il termine *logica* ha assunto nella storia del pensiero significati profondamente differenti, ma è anche difficile proporne una caratterizzazione che possa essere universalmente condivisa. Senza entrare in questioni storiche, troppo intricate per essere affrontate in questa sede, né cercare di proporre una definizione esauriente della disciplina, ci è sufficiente introdurre il discorso con alcune considerazioni di carattere molto generale e, almeno in larga misura, condivisibili. E' solo dopo avere esaminato e studiato i contenuti di tutto il corso di lezioni che sarà possibile farsi un'idea più precisa sugli obiettivi e sui metodi dell'indagine logica come è venuta configurandosi nel nostro secolo.

E' bene segnalare subito che la logica si occupa dei ragionamenti dopo che essi sono stati espressi in qualche forma di linguaggio (e quindi non dell'attività del pensare, dei meccanismi interni della nostra mente, ma piuttosto del pensato dopo che questo è stato comunicato) e che uno dei suoi scopi è quello di caratterizzare quali sono i ragionamenti corretti.

Un ragionamento si presenta come una sequenza finita di proposizioni, dove con *proposizione* si intende una espressione linguistica per la quale ha senso chiedersi se è vera o falsa (a prescindere dal fatto che si sappia quale delle due circostanze si verifica). Una

proposizione può assumere, come si usa dire, uno ed un solo dei due *valori di verità*, il vero (**V**) e il falso (**F**). Questa ipotesi preliminare è detta *principio di bivalenza* [1].

L'ultima proposizione di un *ragionamento*, detta *conclusione*, è preceduta solitamente da "quindi" (o espressioni analoghe, quali "ne segue che", "allora", "pertanto", ecc.). Le altre proposizioni sono le *premesse* (e l'ordine delle premesse è inessenziale). Nel seguito scriveremo spesso la conclusione separandola dalle premesse mediante una linea orizzontale.

Che cosa distingue un ragionamento da una sequenza generica di proposizioni? Chi propone un ragionamento vuole ricondurre la verità della conclusione a quella delle premesse. Tipici esempi di ragionamento sono le dimostrazioni matematiche, le quali si sviluppano mediante passaggi che "conservano la verità". Il nesso chiave, quindi, è quello di *conseguenza logica*: un ragionamento è corretto quando la conclusione è conseguenza logica dell'insieme delle premesse, ossia quando non può darsi il caso che le premesse siano tutte vere e la conclusione falsa. Questa caratterizzazione, pur restando per ora a livello intuitivo, è sufficiente a evidenziare il nesso dell'indagine logica con la matematica e il metodo assiomatico classico[2] e per le nostre considerazioni preliminari.

Cominciamo con l'osservare che siamo tutti disposti a ritenere corretto il seguente ragionamento:

2 è pari o 3 è pari
3 non è pari
---
2 è pari

Infatti, quando supponiamo vera una alternativa fra due proposizioni e riconosciamo che è vera la negazione di una delle due, ci sentiamo autorizzati a ritenere vera l'altra proposizione in alternativa.

I seguenti ragionamenti sono corretti in forza delle stesse considerazioni:

[1] Cogliamo l'occasione per segnalare che, attualmente, la logica è articolata in svariati settori essendo molteplici i contesti nei quali si procede deduttivamente articolando dei ragionamenti. In molti campi il principio di bivalenza si rileva inadeguato. Senza entrare in dettagli, è chiaro che, dal punto di vista intuitivo, varie proposizioni possono essere verosimili, plausibili, altamente probabili, quasi certe, incerte, ecc., ossia tra il vero e il falso vi è tutta una gamma di possibilità intermedie. Il settore della logica del quale ci occupiamo in questa sede modella da vicino il ragionamento matematico e, in ogni caso, è preliminare allo studio di altre logiche (*polivalenti*, *fuzzy*, *probabiliste*) in cui si rinuncia al principio di bivalenza o di logiche *parziali*, in cui non a tutte le proposizioni si attribuisce un valore di verità.

[2] Come è ben noto, già a partire da Euclide, l'assetto assunto dalle teorie matematiche, in primo luogo la geometria, è stato quello assiomatico, caratterizzato dalla scelta di alcuni principi evidenti (veri di per sé) i quali hanno il ruolo di "sostenere" la verità di tutte le proposizioni da essi dedotte: le "dimostrazioni", i ragionamenti logici, assicurano la trasmissione della verità dai principi (assiomi e postulati) ai teoremi. Il metodo assiomatico ha subito, a partire dal secolo scorso, delle profonde trasformazioni collegate con quanto ci apprestiamo ad illustrare.

Carlo è ligure o piemontese
Carlo non è piemontese

Carlo è ligure

Il triangolo T è rettangolo o isoscele
Il triangolo T non è isoscele

Il triangolo T è rettangolo

Si dice che i tre ragionamenti hanno la stessa *forma logica*. E' allora del tutto naturale evidenziare tale forma sottolineando che si prescinde dai "contenuti" delle singole proposizioni poste in alternativa:

A o B
non B (*)

A

Analoghe considerazioni si possono svolgere a proposito dei tre seguenti ragionamenti:

3 è minore di 7
Se un numero è minore di un altro, allora il secondo è maggiore del primo
---
7 è maggiore di 3

La retta r è perpendicolare alla retta s
Se un retta è perpendicolare ad un'altra, allora la seconda è incidente alla prima
---
s è incidente a r

Ugo è nipote di Giuseppe
Se una persona è nipote di un'altra, allora quest'ultima è zio della prima
---
Giuseppe è zio di Ugo

In tutti e tre i ragionamenti una premessa afferma il sussistere di una certa relazione fra un individuo e un altro; l'altra premessa che il sussistere di questa relazione comporta, qualsiasi siano i due individui, il sussistere di un'altra relazione fra il secondo e il primo individuo; la conclusione afferma che quest'ultima relazione sussiste fra il secondo e il primo individuo menzionati nella prima premessa. Per il sussistere del nesso di conseguenza logica non è importante la natura degli elementi coinvolti (numeri, rette, persone), né delle relazioni (minore e maggiore, perpendicolare e incidente, nipote e zio), quanto i nessi stabiliti tra essi nelle premesse e nella conclusione.

Anche in questo caso appare naturale rendere trasparente quanto è sufficiente a garantire il riconoscimento del nesso di conseguenza logica con una opportuna riscrittura. Se indichiamo con a e b i due individui, con R la prima relazione e con S la seconda (ovvero indichiamo con lettere gli elementi "variabili" nei tre ragionamenti), si ottiene:

Rab
per ogni due individui x e y, se Rxy allora Syx (**)

Sba

Il nesso di conseguenza logica acquista significato proprio in relazione a scritture come questa in cui figurano termini il cui valore è indeterminato.

Riprendendo, ad esempio, il primo dei tre ragionamenti, le due premesse sono vere e la conclusione è vera e, quindi, non appare molto sensato affermare che il ragionamento è corretto poiché, se le premesse sono vere, lo è anche la conclusione, dato che, di fatto, le premesse sono vere.

Nel caso di (**), invece, ha perfettamente senso dire, ad esempio, "se Rab è vera", in quanto, a seconda del significato assunto da a, da b e da R, Rab potrà essere vera (ad esempio se a = 3, b = 5 e R = minore, oppure a = 8, b = 4 e R = doppio) oppure essere falsa (ad esempio se a = 9, b = 7 e R = minore, oppure a = 8, b = 5 e R = doppio).

Il nesso di conseguenza logica, e quindi la correttezza del ragionamento, sussiste perché, qualsiasi siano gli individui denotati da a e da b, e le relazioni denotate da R e da S, se le premesse sono vere, allora è vera anche la conclusione.

Inoltre, sono corretti anche i due seguenti ragionamenti formalizzati da (**):

3 è multiplo di 7
Se un numero è multiplo di un altro, allora quest'ultimo è maggiore del primo
---
7 è maggiore di 3

3 è multiplo di 7
Se un numero è multiplo di un altro, allora quest'ultimo è minore del primo
---
7 è minore di 3

anche se nel primo le premesse sono false e la conclusione vera e nel secondo le premesse sono false e la conclusione falsa (e quindi la verità della conclusione va tenuta distinta dalla correttezza del ragionamento).

L'individuazione della forma logica avviene con riferimento a un linguaggio le cui espressioni non sono né vere né false, ma che sono suscettibili di molteplici interpretazioni. Inoltre, come vedremo fra breve, date le molte ambiguità del linguaggio naturale, si rivela

opportuno simbolizzare anche le restanti parti di (*) e (**) le quali, dopo le precisazioni che presenteremo nel seguito, assumeranno le seguenti configurazioni:

$$\frac{A \vee B \quad \neg B}{A} \quad (*) \qquad \frac{Rab \quad \forall x \forall y(Rxy \rightarrow Syx)}{Sba} \quad (**)$$

Quanto finora esposto giustifica ampiamente l'attributo "formale" che accompagna usualmente il termine "logica" e l'esigenza di ricorrere a un linguaggio artificiale (suscettibile di svariate interpretazioni). L'impiego sistematico di simboli ha fatto sì che per molto tempo sia stato usato frequentemente l'attributo "simbolica" (e la più prestigiosa rivista di logica è il *Journal of Symbolic Logic*)[3].

A questo proposito è bene tener presente che *formalizzazione* e *simbolizzazione* indicano procedimenti diversi. Un sistema di simboli è usualmente introdotto per ottenere maggiore chiarezza, precisione o concisione (ad esempio, la stenografia, la segnaletica stradale), ma conserva un suo preciso contenuto. Formalizzare, invece, significa prescindere dai contenuti, e ciò non richiede necessariamente l'uso di una simbolizzazione, anche se questa, facilitando enormemente l'indagine di tipo formale, diviene uno strumento praticamente indispensabile.

2. Proposizioni atomiche e composte

Consideriamo alcuni esempi di ragionamento:

(a) Se il quadrilatero Q' è un rombo, allora i due triangoli T' e T" sono isosceli. Se almeno uno dei due triangoli T' e T" è isoscele, allora il quadrilatero Q" è un rettangolo. Quindi, se il quadrilatero Q" non è un rettangolo, allora il quadrilatero Q' non è un rombo.

(b) Se il quadrilatero Q' è un rombo, allora almeno uno dei due triangoli T' e T" è isoscele. Se i triangoli T' e T" sono isosceli, allora il quadrilatero Q" non è un rettangolo. Quindi, se il quadrilatero Q" è un rettangolo, allora il quadrilatero Q' non è un rombo.

(c) Tutti i numeri divisibili per 4 sono pari. 8 è divisibile per 4. Quindi 8 è pari.

[3] La locuzione "logica matematica" ha un duplice significato. Da un lato significa "logica della matematica" per sottolineare la stretta dipendenza dell'indagine logica da quella condotta nella matematica: l'enorme sviluppo della logica nei primi decenni del secolo è avvenuto a stretto contatto con la nascita dell'assiomatica moderna (Peano, Hilbert) e con la problematica dei fondamenti della matematica. D'altro lato la logica si sviluppa anch'essa con il metodo deduttivo tipico della matematica, nel senso che lo scopo, come si vedrà in seguito, non è tanto quello di elencare le forme corrette di ragionamento, ma di ricondurle (possibilmente) tutte ad alcune particolarmente semplici assunte come primitive (e quindi "logica matematica" significa anche "logica sviluppata con metodo matematico").

(d) Tutti i quadrati sono rombi. Tutti i rombi sono quadrilateri. Quindi, tutti i quadrati sono quadrilateri.

(e) Vi è un numero pari primo. Un numero pari è multiplo di 2. Quindi, almeno un numero multiplo di 2 è primo.

Se esaminiamo le proposizioni che intervengono in questi ragionamenti possiamo constatare che sono di due tipi: quelle che esprimono fatti semplici, vale a dire che non contengono al loro interno altre proposizioni, e quelle più complesse, ossia che contengono parti che sono a loro volta proposizioni.

Le proposizioni del primo tipo sono dette *atomiche* o *semplici* e le altre proposizioni *composte*.

Le proposizioni atomiche, a loro volta, sono di due tipi.

(**A**) Le proposizioni atomiche del primo tipo contengono nomi di individui e esprimono il fatto che un individuo ha una certa proprietà ("8 è pari") oppure che tra più individui sussiste una certa relazione ("8 è divisibile per 4")[4]. Nelle proposizioni atomiche possono essere coinvolte, oltre che proprietà e relazioni binarie, anche relazioni a più argomenti:

(1) "8 è pari"; "6 è primo";

(2) "8 è divisibile per 4"; "10 è maggiore di 7"; "9 è minore di 3";

(3) "5 è compreso tra 3 e 7"; "3 è la semisomma di 2 e 5"; "M è il punto medio del segmento di estremi A e B";

(4) "I punti A, B, C e D sono i vertici di un rombo"; "2,6,3 e 9 formano una proporzione";

(5) "I punti A, B, C, D e E sono vertici di un pentagono convesso"[5].

In logica si usa il termine *predicato* per indicare proprietà e relazioni a un numero qualsiasi di argomenti. Il primo tipo di proposizioni atomiche è quindi caratterizzato dalla presenza di uno o più nomi di individui specifici e di un predicato che si dichiara sussistere degli individui menzionati.

Per formalizzare le proposizioni atomiche di questo tipo introduciamo dei simboli:
- si indicano i nomi degli individui con le lettere: a, b, c, d, a', b', c', d', a",... dette *costanti individuali.* (Quando si vuole essere più precisi si sceglie una volta per tutte una determinata successione di costanti individuali: $a_1, a_2, a_3, a_4, ..., a_n, ...$).

[4] Per la caratterizzazione delle proprietà e delle relazioni si ricordi quanto esposto nella prima lezione.
[5] Evidentemente si possono proporre esempi tratti dal linguaggio ordinario: "Carlo è medico"; "Giuseppe corre"; "Aldo ama Silvia"; "Elisa è figlia di Mario e Luisa"; "Savona è tra Genova e Imperia", ecc. Per ragioni didattiche preferiamo proporre prevalentemente esempi del linguaggio della matematica, che è più preciso di quello naturale.

- si indicano i nomi dei predicati con le lettere: P, Q, R, S, P', Q', R', S', P",... dette *costanti predicative*. (Quando si vuole essere più precisi si usa una notazione con doppio indice: per ogni n si assume una successione di costanti predicative a n argomenti: $R_1^n, R_2^n, \ldots, R_i^n, \ldots$).

Dato il carattere introduttivo di questo nostro intervento usiamo la notazione più semplice, senza indici, per cui proposizioni come le (1)-(5) si formalizzano con una costante predicativa seguita dall'opportuno numero di costanti individuali, ad esempio nel modo seguente:

(1') Pa (oppure Pb, oppure Qa, oppure Qc,...);

(2') Pab (oppure Rba, oppure Sbc,...);

(3') Pabc (oppure Rabc, oppure Rbcd,...);

(4') Rabcd (oppure Sdcab,...);

(5') Rabcde (oppure Sadecb,...)[6].

Sempre restando nell'ambito delle proposizioni atomiche del primo tipo va osservato che non sempre gli individui sono menzionati direttamente mediante il loro nome, ma, a volte, attraverso un loro legame con altri individui.

Così, ad esempio, anziché "8" si può dire:

(6) "il doppio di 4";

(7) "la radice quadrata di 64";

(8) "il prodotto di 2 e 4";

(9) "il medio proporzionale tra 4 e 16";

(10) "il successivo della somma di 3 e 4";

(11) "il doppio del quadrato di 2";

(12) "la somma del quadrato di 2 e della somma di 1 e 3";

(13) "il successivo della somma di 3 e del prodotto di 1 con il quadrato di 2" [7].

Come è noto, "il doppio di...", "la radice quadrata di...", "il prodotto di... e di...", "la somma di... e di..." caratterizzano funzioni a uno o due argomenti. Si rivela opportuno, quindi, avere a disposizione dei simboli per denotare funzioni:
- si indicano i nomi delle funzioni con le lettere: f, g, h, l, f', g', h', l', f", ... dette *costanti funzionali.* (Quando si vuole essere più precisi si usa una notazione con doppio indice: per ogni n si assume una successione di costanti funzionali a n argomenti: $f_1^n, f_2^n, \ldots, f_i^n, \ldots$).

[6] Se compare più volte lo stesso individuo occorre avere l'avvertenza di usare la stessa costante; così "5 è multiplo di 5" si formalizza, ad esempio, con Raa (o Rbb, o Saa), "8 è la somma di 4 e 4" con, ad esempio, Sabb (o Rbcc), "1 è il prodotto di 1 e 1" con Saaa (o Rbbb).
[7] Esempi tratti dal linguaggio comune sono: "il padre di Carlo"; "il figlio di Giovanna e Enrico"; "la capitale della Lombardia"; "il figlio dello zio di Piero e della nuora di Sergio"; "la cuccia del cane dello zio di Pietro"; ecc.

Continuando a servirci della notazione più semplice, le precedenti espressioni si formalizzano, ad esempio, nel modo seguente:

(6') e (7') f(a) (oppure f(b), oppure g(a),...);
(8') e (9') f(a,b) (oppure g(b,c), oppure h(a,c),...);
(10') f(g(a,b));
(11') f(g(a));
(12') f(g(a),f(b,c));
(13') f(g(g(a,b),h(a,l(b)))).

Quindi, una formalizzazione delle proposizioni atomiche del primo tipo:

(14) "4 è maggiore del doppio di 6";
(15) "5 è compreso tra il quadrato di 2 e il quadrato di 3";
(16) "La somma di 2 e 2 è minore del prodotto di 2 e 3";
(17) "Il quadrato di 2 è medio proporzionale tra 1 e il quadrato di 4";
(18) "La radice quadrata del doppio di 8 è minore della radice quadrata di 25";
(19) "2 elevato alla terza è minore di 3 elevato alla seconda";
(20) "Il quadrato di 2 è compreso tra la radice quadrata di 9 e la somma di 2 e 9",

è la seguente:

(14') R(a,f(b));
(15') S(a,f(b),f(c));
(16') Q(f(a,a),g(a,b));
(17') S(f(a),b,f(c));
(18') R(f(g(a)),f(b));
(19') Q(f(a,b),f(b,a));
(20') S(f(a),g(b),h(a,b))[8].

(B) Esempi di proposizioni atomiche del secondo tipo sono:

(21) "Tutti i quadrati sono rombi";
(22) "Vi è un numero pari primo";
(23) "Tutti i numeri divisibili per 4 sono pari".

A differenza degli esempi precedenti si tratta di proposizioni (in quanto suscettibili di essere vere o false) atomiche (non contenendo parti che a loro volta sono proposizioni) in cui

[8] Le parentesi, come in algebra, servono a rendere univocamente leggibili le formule. Il loro impiego, comunque, non è strettamente necessario.

compaiono i termini "tutti", "vi è un", detti *quantificatori*. La prima proposizione afferma che tutti gli individui che hanno la proprietà di "essere un quadrato" hanno anche la proprietà di "essere un rombo"; la seconda che esiste almeno un individuo che ha la proprietà di "essere numero pari" e "essere numero primo"; la terza che ogni individuo che ha la proprietà di "essere divisibile per 4" (la relazione "essere divisibile" con l'individuo 4 [9]) ha anche quella di "essere pari", e quindi riguardano la "quantità" di individui generici che possiedono proprietà o entrano in relazione con altri.

Nei ragionamenti sia di carattere matematico, sia del linguaggio ordinario, intervengono molti altri quantificatori (ad esempio, "esiste un numero finito di...", "quasi tutti...", "la maggior parte di...", "vi sono tanti...quanti...").

Tuttavia, per l'analisi della gran parte dei ragionamenti che intervengono sia in matematica, sia in altri contesti deduttivi è sufficiente prendere in considerazione solo i due quantificatori "per ogni" ed "esiste"[10] detti rispettivamente *quantificatore universale* e *quantificatore esistenziale* [11] che si indicano con i simboli $\forall$ e $\exists$.

La formalizzazione delle proposizioni atomiche quantificate universalmente ed esistenzialmente richiede strumenti che saranno presentati nel prossimo paragrafo e, quindi, la rimandiamo al §4.

Le proposizioni non atomiche sono dette *composte*. Esse si ottengono combinando proposizioni atomiche mediante termini detti *connettivi*. Nei primi due esempi di ragionamento (a) e (b) di inizio paragrafo figurano quattro proposizioni atomiche, vale a dire "Il quadrilatero Q' è un rombo", "il triangolo T' è isoscele", "il triangolo T" è isoscele", "il quadrilatero Q " è un rettangolo", combinate mediante i connettivi "se..., allora...", "e", "o" (a due argomenti, che collegano due proposizioni) e "non" (a un argomento, che si applica ad una singola proposizione).

[9] Segnaliamo che la formalizzazione non è un'operazione univoca; ad esempio "2 è minore di 3" si può leggere "2 ha la proprietà di essere minore di 3" e quindi rendere con Pa (a = 2, P = essere minore di 3), oppure "3 ha la proprietà di essere dopo 2" e quindi rendere con Qb (b = 3, Q = essere dopo 2), oppure "2 e 3 sono nella relazione minore" e quindi rendere con Rab (a = 2, b = 3 e R = essere minore).

[10] In effetti, come è immediato riconoscere intuitivamente, questi due quantificatori sono interdefinibili in quanto "per ogni" ha lo stesso significato di "non esiste non" ("Ogni quadrato è rombo" equivale a "Non esiste un quadrato che non sia rombo") e "esiste" ha lo stesso significato di "non per ogni non" ("Esiste un numero primo pari" equivale a "Non ogni numero primo non è pari").

[11] Come emerge già dagli esempi proposti, nel linguaggio comune i due quantificatori possono essere espressi con varie locuzioni, quali "tutti", "ogni", "un" e "almeno", "qualche", "un" rispettivamente. Si osservi che "un" può fungere da quantificatore universale (come in "Un numero pari è divisibile per 2") o esistenziale (come in "Un numero pari è primo") e solo il contesto chiarisce il suo ruolo. Questo è un esempio particolarmente efficace per evidenziare le ambiguità del linguaggio ordinario, in cui lo stesso termine può assumere significati differenti.

3. I connettivi vero-funzionali. Basi di connettivi

Nei ragionamenti possono intervenire molteplici connettivi. Al fine di graduare l'approfondimento dell'analisi logica delle proposizioni e del rapporto di conseguenza logica, si è soliti considerare inizialmente un solo tipo di connettivi, i *connettivi vero-funzionali*, peraltro sufficienti a esprimere la quasi totalità dei ragionamenti di carattere matematico: si dice che un connettivo è *vero-funzionale* se e solo se il valore di verità della proposizione composta ottenuta tramite esso dipende unicamente dai (è funzione dei) valori di verità delle proposizioni componenti.

L'unico connettivo vero-funzionale a un argomento significativo è quello che inverte il valore di verità della proposizione a cui si applica. Esso è detto *negazione* e per esso si usa il simbolo ¬. Il suo comportamento può essere visualizzato mediante la seguente tabella:

A	¬A
V	**F**
F	**V**

• Il simbolo ¬ si legge "non", poiché, nel linguaggio comune, è proprio tramite la negazione "non" che di solito si inverte il valore di verità di un enunciato. Un connettivo, quindi, si può intendere sia come un operatore tra proposizioni, sia una funzione da valori di verità a valori di verità.

Consideriamo ora i possibili connettivi vero-funzionali *a due argomenti*. Essi sono 16. Se indichiamo con A e B i due argomenti e con c_1,..., c_{16} i sedici connettivi, si ha la seguente tabella (in cui, per brevità, scriviamo c_i anziché $c_i(A,B)$):

A	B	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}
V	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**	**F**	**F**	**F**	**F**
V	**F**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**
F	**V**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**
F	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**

Se trascuriamo i connettivi che hanno un valore costante, indipendente da quelli di A e B (cioè c_1 e c_{16}), e quelli che si riducono ad A, o a B (c_4 e c_6), o alle loro negazioni (c_{13} e c_{11}), restano dieci connettivi a due argomenti significativi:

A	B	c_2	c_3	c_5	c_7	c_8	c_9	c_{10}	c_{11}	c_{14}	c_{15}
V	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**	**F**
V	**F**	**V**	**V**	**F**	**F**	**F**	**V**	**V**	**V**	**F**	**F**
F	**V**	**V**	**F**	**V**	**F**	**F**	**V**	**V**	**F**	**V**	**F**
F	**F**	**F**	**V**	**V**	**V**	**F**	**V**	**F**	**F**	**F**	**V**

Concentriamo la nostra attenzione sui primi cinque, dato che i secondi cinque si possono ottenere dai primi attraverso la negazione: negando c_2 si ottiene c_{15}, negando c_3 si ottiene c_{14}, e così via.

• $c_2(A,B)$ ha valore di verità **V** quando almeno una delle due proposizioni A e B ha valore **V**, ed ha valore **F** quando entrambe A e B hanno valore **F.**

$c_2(A,B)$ si dice *disgiunzione* o *alternativa* (*non esclusiva*) di A e di B (e A e B sono i *disgiunti*), e per esso introduciamo il simbolo ∨ (anziché $c_2(A,B)$ scriviamo A ∨ B). Tale simbolo si legge "o", poiché, nel linguaggio comune, per indicare che riteniamo vera una proposizione composta di due quando almeno una è vera, usiamo "o" o "oppure" nel senso di *vel*. Si ha allora la seguente tavola di verità per la disgiunzione:

A	B	A ∨ B
V	**V**	**V**
V	**F**	**V**
F	**V**	**V**
F	**F**	**F**

• $c_8(A,B)$ ha valore di verità **V** quando A e B hanno entrambe valore **V** e ha valore **F** negli altri tre casi. Esso si dice *congiunzione* di A e B (e A e B sono i *congiunti*) e lo indichiamo con A ∧ B. Il simbolo ∧ si legge "e" poiché, nel linguaggio comune, per asserire la verità simultanea di due proposizioni si adopera proprio la congiunzione "e".

• $c_5(A,B)$ ha valore di verità **F** solo quando A ha valore di verità **V** e B ha valore di verità **F** (e **V** negli altri casi). Esso si dice *condizionale* (*materiale*) di *antecedente* A e *conseguente* B e si indica con A → B. Queste condizioni di verità corrispondono a quelle impiegate nei ragionamenti matematici quando si afferma "se A, allora B" (o, equivalentemente, "A solo se B"), ossia che non può darsi il caso che A sia vera e B falsa[12].

[12] Su queste "letture" dei connettivi torneremo diffusamente più avanti nel §5.

• Si osservi che $c_3(A,B)$ non è altro che il condizionale materiale di antecedente B e conseguente A (cioè $B \rightarrow A$) e quindi non corrisponde a un connettivo concettualmente diverso dal precedente.

• $c_7(A,B)$ ha valore di verità **V** quando A e B hanno stesso valore di verità e ha valore **F** quando A e B hanno valore diverso. Esso è detto *bicondizionale* di A e B e si indica $A \leftrightarrow B$, che si legge "A se e solo se B".

Si tenga presente che un connettivo si può identificare con una funzione di verità. La funzione associata a $\leftrightarrow$ si ottiene facilmente componendo quelle relative a $\wedge$ e a $\rightarrow$; come è facile calcolare:

$A \leftrightarrow B$ equivale a $(A \rightarrow B) \wedge (B \rightarrow A)$[13]

Ciò si può esprimere dicendo che il bicondizionale si può definire in termini di congiunzione e condizionale.

Così, analogamente:

$A \rightarrow B$ equivale a $(\neg A) \vee B$[14]

Abbiamo già osservato che gli ultimi cinque connettivi della tabella si possono esprimere come negazione dei primi cinque. Si può comunque rilevare che:

• $c_{10}(A,B)$ corrisponde alla *disgiunzione esclusiva* (ha valore **V** se e solo se una sola delle proposizioni A e B ha valore di verità **V**), che indichiamo con A + B, ossia a quella combinazione vero-funzionale che nel linguaggio comune esprimiamo con "o" (nel senso di *aut*).

• $c_9(A,B)$ ha valore **F** quando A e B sono entrambe vere e valore **V** negli altri tre casi; esso si indica con NAND - essendo la negazione della congiunzione - ed è detto "o incompatibile" in quanto, nel linguaggio ordinario spesso usiamo "o" in questo senso, ossia per sottolineare che le due proposizioni in alternativa non possono essere entrambe vere.

• $c_{15}(A,B)$ è la negazione della disgiunzione ed è detta NOR (e nel linguaggio comune viene espressa con "né A, né B").

NAND e NOR sono detti anche funtori di Sheffer e hanno una interessante proprietà che enunceremo tra breve.

---

[13] Se b(x,y) è la funzione associata a $\leftrightarrow$ (b(**V**,**V**) = b(**F**,**F**) = **V**, b(**V**,**F**) = b(**F**,**V**) = **F**), m(x,y) quella relativa a $\rightarrow$ (m(**V**,**V**) =m(**F**,**V**)= m(**F**,**F**) =**V**, m(**V**,**F**) = **F**) e c(x,y) quella relativa a $\wedge$ (c(**V**,**V**) = **V**, c(**V**,**F**) = c(**F**,**V**) = c(**F**,**F**) = **F**), allora:

per ogni x,y $\in$ {**V**,**F**} b(x,y) = c(m(x,y),m(y,x))

[14] Come nella nota precedente, detta n(x) la funzione associata alla negazione $\neg$ (n(**V**) = **F**, n(**F**) = **V**) e d(x,y) quella della disgiunzione $\vee$ (d(**V**,**V**) = d(**V**,**F**) = d(**F**,**V**) =**V**, d(**F**,**F**) = **F**):

per ogni x,y $\in$ {**V**,**F**} m(x,y) = d(n(x),y)

Ricapitoliamo nella seguente tabella il comportamento dei connettivi a due argomenti esaminati:

A	B	A $\vee$ B	A $\wedge$ B	A $\rightarrow$ B	A $\leftrightarrow$ B	A + B	A NAND B	A NOR B
V	**V**	**V**	**V**	**V**	**V**	**F**	**F**	**F**
V	**F**	**V**	**F**	**F**	**F**	**V**	**V**	**F**
F	**V**	**V**	**F**	**V**	**F**	**V**	**V**	**F**
F	**F**	**F**	**F**	**V**	**V**	**F**	**V**	**V**

Come si è accennato in precedenza tra i connettivi intercorrono molteplici rapporti di interdefinibilità. Senza entrare in dettagli (per i quali rinviamo a un qualsiasi manuale di logica) elenchiamo alcune delle equivalenze più significative:

(1) $A \wedge B$ equivale a $\neg((\neg A) \vee (\neg B))$ (legge di De Morgan)
(2) $A \vee B$ equivale a $\neg((\neg A) \wedge (\neg B))$ (legge di De Morgan)
(3) $A \wedge B$ equivale a $\neg(A \rightarrow (\neg B))$
(4) $A \vee B$ equivale a $(\neg A) \rightarrow B$
(5) $\neg(\neg A)$ equivale a A (legge di doppia negazione)
(6) $\neg A$ equivale a A NAND A
(7) $A \vee B$ equivale a (A NAND A) NAND (B NAND B)
(8) $\neg A$ equivale a A NOR A
(9) $A \wedge B$ equivale a (A NOR A) NOR (B NOR B)
(10) $A \wedge B$ equivale a $B \wedge A$
(11) $(A \wedge B) \wedge C$ equivale a $A \wedge (B \wedge C)$
(12) $A \vee B$ equivale a $B \vee A$
(13) $(A \vee B) \vee C$ equivale a $A \vee (B \vee C)$

La (1) indica che la congiunzione si può definire in termini di negazione e disgiunzione, la (2) che la disgiunzione si può definire in termini di negazione e di congiunzione, e così via.

Queste considerazioni sono particolarmente significative nel senso che si può dimostrare che *tutti* i connettivi vero-funzionali (a qualsiasi numero di argomenti) possono essere definiti in termini di alcuni di essi. Più precisamente, si definisce *base di connettivi* un insieme di connettivi vero-funzionali che consente di definire tutti gli altri. Senza entrare in troppi dettagli tecnici, vediamo, attraverso un esempio, come si riconosce che $\{\neg, \wedge, \vee\}$ è una base di connettivi.

Consideriamo il connettivo vero-funzionale a tre argomenti $\Diamond(A,B,C)$ caratterizzato dalla seguente tabella:

A	B	C	$\Diamond(A,B,C)$
V	**V**	**V**	**F**
V	**V**	**F**	**V**
V	**F**	**V**	**V**
V	**F**	**F**	**F**
F	**V**	**V**	**F**
F	**V**	**F**	**F**
F	**F**	**V**	**V**
F	**F**	**F**	**F**

Si considerano le righe della tabella in cui $\Diamond(A,B,C) = \mathbf{V}$ e, in corrispondenza di ciascuna di esse, scriviamo una congiunzione in cui ciascuna delle tre lettere A, B e C è scritta una ed una sola volta senza o preceduta dal segno di negazione a seconda se nella riga sotto di essa compare **V** o **F** rispettivamente.

In corrispondenza della seconda riga si ottiene:

$A \wedge B \wedge \neg C$[15]

In corrispondenza della terza e della settima riga si hanno:

$A \wedge \neg B \wedge C$;

$\neg A \wedge \neg B \wedge C$

Si forma poi la disgiunzione delle congiunzioni ottenute[16]:

$(A \wedge B \wedge \neg C) \vee (A \wedge \neg B \wedge C) \vee (\neg A \wedge \neg B \wedge C)$ (*)

Il valore di verità che questa espressione assume in corrispondenza delle possibili assegnazioni di valori di verità a A, B e C coincide, come si verifica facilmente, con quello di $\Diamond(A,B,C)$. Quindi, usando i connettivi $\neg$, $\wedge$, $\vee$ si ottiene la combinazione (*) che, dal punto di vista del valore di verità, equivale alla proposizione composta mediante il connettivo dato[17].

Il procedimento illustrato nell'esempio precedente si applica con generalità. Ad esempio:

[15] Si tenga presente che, per l'equivalenza (11), vale la proprietà associativa della congiunzione, per cui si possono scrivere senza ambiguità congiunzioni iterate: una congiunzione di più proposizioni ha valore **V** se e solo se ogni congiunto ha valore **V**.

[16] Anche per la disgiunzione vale la proprietà associativa (13) e, quindi, si possono scrivere disgiunzioni iterate: una disgiunzione di più proposizioni ha valore di verità **V** quando almeno uno dei disgiunti ha valore **V**.

[17] Qualora nella tabella di $\Diamond(A,B,C)$ non figurasse alcuna **V**, una espressione contenente solo i connettivi $\neg$, $\wedge$, $\vee$ che assume sempre valore **F** qualsiasi siano i valori assunti dalle lettere A, B e C è, ad esempio:
$(A \wedge \neg A) \vee (B \wedge \neg B) \vee (C \wedge \neg C)$.

$A + B$ equivale a $(A \wedge \neg B) \vee (\neg A \wedge B)$

A NAND A equivale a $(A \wedge \neg B) \vee (\neg A \wedge B) \vee (\neg A \wedge \neg B)$

A NOR A equivale a $\neg A \wedge \neg B$

$A \leftrightarrow B$ equivale a $(A \wedge B) \vee (\neg A \wedge \neg B)$

Dal fatto che $\{\neg, \wedge, \vee\}$ è una base di connettivi e da quanto osservato a proposito della (1), si deduce che $\{\neg, \vee\}$ è una base di connettivi, mentre, da quanto osservato a proposito della (2), si deduce che $\{\neg, \wedge\}$ è una base.

Dalle (3) e (4) si deduce che anche $\{\neg, \rightarrow\}$ è una base.

Dalle (6) e (7), e dalle (8) e (9), per quanto appena osservato, segue che ciascuno dei funtori di Sheffer, NAND e NOR, da solo, forma una base di connettivi, ossia che mediante ciascuno di essi si possono definire tutti gli altri connettivi vero-funzionali (a qualsiasi numero di argomenti).

Per queste ragioni, quando si introduce il linguaggio artificiale con cui si costruiscono i calcoli logici[18], si assumono come primitivi solo i connettivi di una base. E' importante sottolineare che, mediante una base di connettivi, si riesce a "dominare" l'intero orizzonte della vero-funzionalità, ossia si possono formalizzare tutti i ragionamenti che coinvolgono proposizioni comunque complesse, purché ottenute combinando proposizioni semplici mediante connettivi vero-funzionali.

4. Il linguaggio della logica dei predicati

I connettivi intervengono nella formalizzazione delle proposizioni atomiche del secondo tipo introdotte nel §2. Le proposizioni (21), (22) e (23) si parafrasano nel modo seguente:

(21) per ogni individuo x (se x è un quadrato, allora x è un rombo);
(22) esiste un individuo x tale che x è un numero pari e x è numero primo;
(23) per ogni individuo x (se x è divisibile per 4, allora x è pari);

e si formalizzano, ad esempio, nel modo seguente:

(21') $\forall x(Px \rightarrow Qx)$; (22') $\exists x(Px \wedge Qx)$; (23') $\forall x(Rxa \rightarrow Qx)$

Per formalizzare le proposizioni quantificate occorre quindi introdurre delle lettere che stanno per individui generici:

- si indicano i nomi degli individui generici con le lettere: x, y, z, x', y', z', x",... dette *variabili individuali.* (Quando si vuole essere più precisi si sceglie una volta per tutte una successione di variabili individuali: $x_1, x_2, x_3, x_4, \ldots, x_n, \ldots$).

[18] Vi sono dei calcoli logici, ad esempio quello intuizionista, o quello polivalente, in cui ai connettivi si attribuisce un significato diverso da quello qui illustrato e in cui non valgono i risultati di interdefinibilità ora enunciati.

Vediamo alcuni esempi di proposizioni più complesse e la loro formalizzazione[19].

(24) "Tutti i numeri maggiori di 8 sono maggiori sia di 6 che di 7"
$\forall x(Rxa \rightarrow (Rxb \wedge Rxc))$

(25) "Per ogni numero ne esiste uno maggiore"
$\forall x \exists y Rxy$ $\quad (\forall x(Px \rightarrow \exists y(Py \wedge Rxy))$, dove P = essere numero)

(26) "Esiste un numero maggiore o uguale di tutti i numeri"
$\exists x \forall y Rxy$ $\quad (\exists x(Px \wedge \forall y(Py \rightarrow Rxy))$, dove P = essere numero)

(27) "Nessun numero minore di 3 è compreso tra 5 e 10"
$\neg \exists x(Rxa \wedge Sxbc)$

(28) "Per due punti qualsiasi passa una retta"
$\forall x \forall y(Px \wedge Py \rightarrow \exists z(Qz \wedge Rzx \wedge Rzy))$

(29) "Se un punto A sta tra due punti B e C allora sta anche tra C e B"
$\forall x \forall y \forall z(Px \wedge Py \wedge Pz \rightarrow (Sxyz \rightarrow Sxzy))$

(30) Ogni numero pari è somma di due numeri primi"
$\forall x(Px \rightarrow \exists y \exists z(Qy \wedge Qz \wedge Sxyz))$

(31) "Il quadrato di 5 è maggiore del quadrato di 4 e non è il quadrato di alcun numero pari"
$R(f(a),f(b)) \wedge \neg \exists x(Px \wedge S(f(a),f(x)))$

In definitiva, gli ingredienti del linguaggio della logica dei predicati sono: l'*alfabeto*, l'insieme dei *termini* e l'insieme delle *formule ben formate.*

L'*alfabeto* è costituito da:

(A) costanti individuali: $a_1, a_2, a_3, a_4, ..., a_n, ...$ (a, b, c, a',...)

(B) variabili individuali: $x_1, x_2, x_3, x_4, ..., x_n, ...$ (x, y, z, x',...)

(C) costanti predicative: per ogni n, $R_1^n, R_2^n, \ldots, R_i^n, \ldots$ (P, Q, R, S, P',...)

(D) costanti funzionali: per ogni n, $f_1^n, f_2^n, \ldots, f_i^n, \ldots$ (f, g, h, f',...)

(E) simboli per connettivi: $\neg, \wedge, \vee, \rightarrow$ [20]

(F) simboli per quantificatori: $\forall, \exists$ [21]

(G) simboli ausiliari: (,)

[19] Si è già accennato al fatto che la formalizzazione di proposizioni del linguaggio ordinario è un'operazione assai più articolata di quanto non appaia negli esempi qui proposti, non è univoca e richiede un'analisi del linguaggio in larga misura relativa ai fini che si intendono perseguire: solo quando ci si muove in un determinato contesto deduttivo e si precisa il linguaggio formale che si intende impiegare, la formalizzazione appare più agevole e più univoca. Per la formalizzazione di molte proposizioni matematiche si rivela utile introdurre un simbolo apposito (in genere il simbolo =) per il predicato di identità.

[20] Per quanto si è osservato nel §3 basterebbe limitarsi ad una base di connettivi.

[21] Come si è richiamato nella nota 10, in effetti basta assumere uno solo dei due quantificatori e definire l'altro.

I *termini* sono le sequenze finite di simboli dell'alfabeto mediante le quali si indicano gli individui (determinati o generici). L'insieme dei termini è definito induttivamente[22]:

(T1) Le costanti e le variabili individuali sono termini.

(T2) Se $t_1, t_2, ..., t_n$ sono termini e f_i^n è un simbolo di funzione a n argomenti, allora $f_i^n t_1 t_2 ... t_n$ è un termine.

(T3) Nient'altro è un termine.

Le *formule atomiche* sono le sequenze finite di simboli del tipo $R_i^n t_1 t_2 ... t_n$, dove $t_1, t_2, ..., t_n$ sono termini e R_i^n è un simbolo di predicato a n argomenti.

La definizione delle *formule ben formate* (fbf) è di tipo induttivo:

(F1) Le formule atomiche sono fbf.

(F2) Se A è una fbf, allora $(\neg A)$ è una fbf.

(F3) Se A e B sono fbf, allora $(A \wedge B)$, $(A \vee B)$, $(A \rightarrow B)$ sono fbf.

(F4) Se A è una fbf e x è una variabile individuale, allora $(\forall xA)$ e $(\exists xA)$ sono fbf.

(F5) Nient'altro è fbf.

Le fbf non atomiche sono dette *composte*.

$(\neg A)$ è detta *negazione* di A, si legge "non A", $\neg$ è il suo segno logico (connettivo) principale e A è la sua sottoformula immediata.

$(A \wedge B)$ è detta *congiunzione* di A e B, si legge "A e B", $\wedge$ è il suo segno logico (connettivo) principale e A e B sono le sue sottoformule immediate.

$(A \vee B)$ è detta *disgiunzione* di A e B, si legge "A o B", $\vee$ è il suo segno logico (connettivo) principale e A e B sono le sue sottoformule immediate.

$(A \rightarrow B)$ è detta *condizionale* di A e B (A è l'*antecedente* e B è il *conseguente*), si legge "se A, allora B", $\rightarrow$ è il suo segno logico (connettivo) principale e A e B sono le sue sottoformule immediate.

$(\forall xA)$ è detta *quantificazione universale* di A con *indice* x, si legge "per ogni x, A", $\forall$ è il suo segno logico (quantificatore) principale e A, che è detta anche *campo d'azione* del quantificatore, è la sua sottoformula immediata.

$(\exists xA)$ è detta *quantificazione esistenziale* di A con *indice* x, si legge "esiste x tale che A", $\exists$ è il suo segno logico (quantificatore) principale e A, che è detta anche *campo d'azione* del quantificatore, è la sua sottoformula immediata.

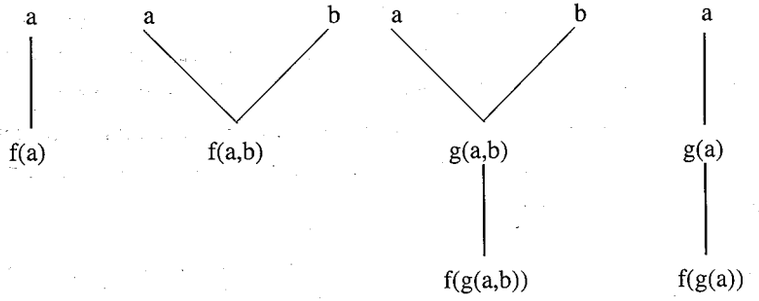
Le *sottoformule* di una fbf si ottengono assumendo come base che ogni fbf è sottoformula di se stessa e iterando ricorsivamente la nozione di sottoformula immediata.

[22] Sulle definizioni induttive si veda la lezione III, in particolare il §4.

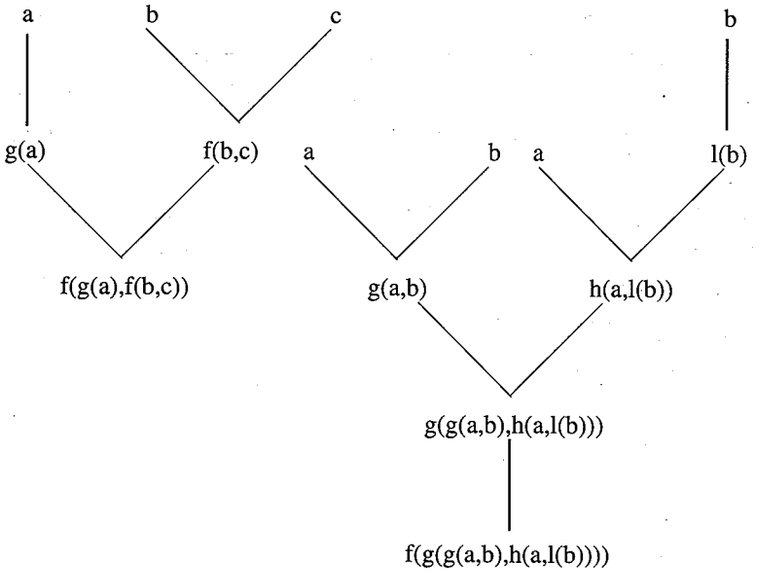
Le occorrenze di una variabile x nel campo d'azione di un quantificatore con indice x sono *vincolate* dal quantificatore. Una fbf in cui tutte le occorrenze delle variabili individuali sono o indice di un quantificatore o vincolate da un quantificatore è detta *chiusa* o *enunciato* (e tutte le fbf con le quali si formalizzano proposizioni del linguaggio comune o matematico sono enunciati).

Definizioni come quelle di termine e di fbf consentono di associare a ciascun termine o fbf il suo *albero di formazione*, una struttura[23] che indica i passi con i quali il termine o la fbf è costruito partendo dagli elementi iniziali (costanti e variabili individuali nel caso dei termini, fbf atomiche nel caso delle fbf). Senza entrare ora in troppi dettagli, scriviamo direttamente gli alberi di formazione di alcuni termini introdotti nel §2:

(1) f(a)
(2) f(a,b)
(3) f(g(a,b))
(4) f(g(a))
(5) f(g(a),f(b,c))
(6) f(g(g(a,b),h(a,l(b))))



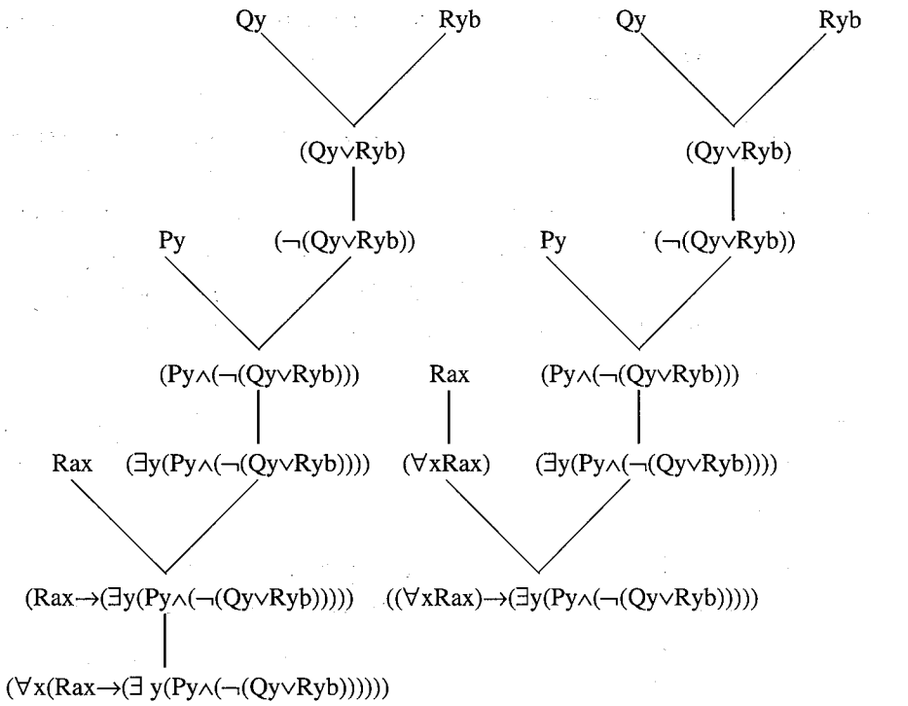
[23] Un *albero* è un insieme, i cui elementi sono detti *nodi*, dotato di una relazione binaria R (se aRb si dice che "a sta sotto b" o "b sta sopra a"; e si dice che "a sta immediatamente sotto b" se a sta sotto b e non esiste alcun nodo che sta sopra a e sotto b) che soddisfa le seguenti proprietà:
(1) R è transitiva (se a sta sotto b e b sta sotto c allora a sta sotto c);
(2) R è irriflessiva (nessun nodo sta sotto se stesso);
(3) esiste un unico nodo, detto *radice*, che sta sotto tutti gli altri;
(4) se a sta sotto b esiste un'unica successione finita di nodi c_1, c_2,..., c_k tale che c_1 = a, c_k = b e c_h sta immediatamente sotto c_{h+1}, per ogni h da 1 a k – 1.

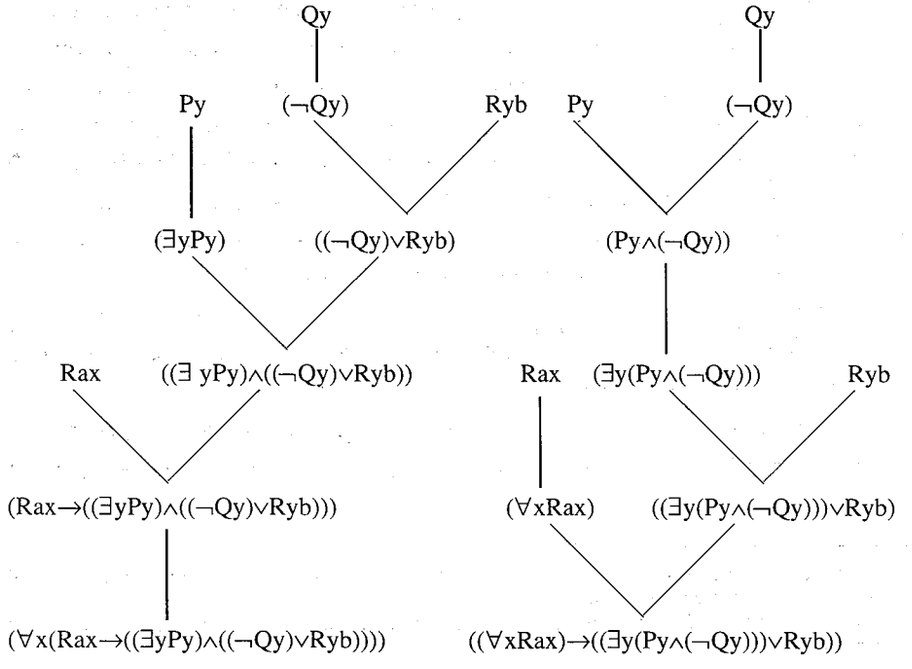


Alla radice di questi alberi figura il termine costruito, nelle foglie (ossia i nodi che non stanno sotto ad alcun altro nodo) i termini iniziali (costanti o variabili individuali) e negli altri nodi i termini intermedi. L'albero di formazione di un termine, se si precisano le modalità di costruzione, è essenzialmente unico. Ad ogni albero, e quindi ad ogni termine, si può associare un numero naturale, detto *altezza* dell'albero, che è il massimo numero di tratti che figurano nei rami dell'albero (un ramo è un tragitto dalla radice a una foglia). Gli ultimi due alberi hanno altezza 2 e 4 rispettivamente.

Analogo discorso si può ripetere a proposito delle fbf. Vediamo gli alberi di formazione delle seguenti fbf:

(1) $(\forall x(Rax\rightarrow(\exists y(Py\wedge(\neg(Qy\vee Ryb))))))$
(2) $((\forall xRax)\rightarrow(\exists y(Py\wedge\neg((Qy\vee Ryb)))))$
(3) $(\forall x(Rax\rightarrow((\exists yPy)\wedge((\neg Qy)\vee Ryb))))$
(4) $((\forall xRax)\rightarrow((\exists y(Py\wedge(\neg Qy)))\vee Ryb))$





Alla radice figura la fbf, nelle foglie le fbf atomiche e negli altri nodi le fbf intermedie e, quindi, nell'albero figurano tutte le sottoformule della fbf. Ogni nodo "sta sotto" alle sue sottoformule immediate. La (1) ha altezza 6, la (2), la (3) e la (4) hanno altezza 5 [24].

Si può fin d'ora osservare come nelle fbf compaiano molte coppie di parentesi. Si possono introdurre delle convenzioni per eliminare alcune coppie di parentesi, conservando tuttavia la "lettura" univoca della fbf: si suppone che la negazione e i quantificatori leghino più strettamente della congiunzione e della disgiunzione, che, a loro volta, leghino più strettamente del condizionale; inoltre, le parentesi più esterne possono essere soppresse[25]. Con queste convenzioni le quattro fbf precedenti si possono scrivere nel modo seguente:

(1) $\forall x(Rax \rightarrow \exists y(Py \wedge \neg(Qy \vee Ryb)))$
(2) $\forall x Rax \rightarrow \exists y(Py \wedge \neg(Qy \vee Ryb))$
(3) $\forall x(Rax \rightarrow (\exists y Py \wedge (\neg Qy \vee Ryb)))$
(4) $\forall x Rax \rightarrow (\exists y(Py \wedge \neg Qy) \vee Ryb)$

Osserviamo infine che (1) e (2) sono enunciati (tutte le occorrenze di x e y sono vincolate), la (3) è una fbf aperta (le ultime due occorrenze di y sono libere) e la (4) è una fbf aperta (l'ultima occorrenza di y è libera).

5. Complemento: costanti logiche e linguaggio comune

Nei paragrafi precedenti abbiamo illustrato il linguaggio artificiale con il quale si esprimono vari tipi di proposizioni che intervengono nei ragionamenti. Nel seguito del corso si analizzerà come mediante questo linguaggio si possa "modellare" l'attività deduttiva. E' opportuno ora sottolineare come la "traduzione" dei ragionamenti nel linguaggio formale sia un'operazione complessa, che richiede l'esplicitazione dell'effettivo significato delle proposizioni del linguaggio naturale o matematico. Infatti, il linguaggio comune non è articolato primariamente per svolgere ragionamenti; anzi, da questo punto di vista, è alquanto carente a causa delle sue ridondanze, sfumature e persino ambiguità. Molto spesso la "forma logica" delle proposizioni non solo è tutt'altro che trasparente, ma anche in contrasto con la lettura sintattica superficiale. Ciò è particolarmente evidente se si esamina come sono impiegati i connettivi nel linguaggio naturale.

Consideriamo la congiunzione "e". In molti casi "e" congiunge due proposizioni e ha il significato vero-funzionale che corrisponde alla tavola di $\wedge$ (ed è per questo che leggiamo

[24] Le questioni relative agli alberi di formazione e alle proprietà sintattiche del linguaggio della logica dei predicati saranno riprese nella prossima lezione.
[25] In precedenza abbiamo già adottato queste convenzioni per scrivere le formalizzazioni delle proposizioni del linguaggio comune.

"e" il simbolo $\wedge$). *Non vale però il viceversa*, ossia molte occorrenze di "e" non possono essere "tradotte" con $\wedge$[26]. La proposizione:

"I numeri 4 e 8 sono pari"

equivale a:

"4 è pari e 8 è pari"

e la "e" è vero-funzionale e può essere "tradotta" con $\wedge$.

La proposizione "Gli angoli di 60° e 30° cono complementari", che ha la stessa struttura grammaticale, non ha il significato di "60° è complementare e 30° è complementare", ma di "60° è complementare di 30° e 30° è complementare di 60°" (e la "e" è vero funzionale).

In "Elisa e Franco sono una bella coppia" non si può distribuire il predicato sui singoli individui. Così, con "La maglia del Milan è rossa e nera" non si intende affermare "La maglia del Milan è rossa e la maglia del Milan è nera". In "Sono andato alla stazione e ho preso il treno" l'uso di "e" non è vero-funzionale in quanto la verità della proposizione non dipende unicamente dalla verità dei congiunti ("Sono andato alla stazione", "Ho preso il treno"), ma anche dalla loro successione temporale (e infatti è diverso dire "Ho preso il treno e sono andato alla stazione", e la non validità della proprietà commutativa testimonia la non vero-funzionalità di "e" e la sua non "traducibilità" con $\wedge$).

Nel caso della disgiunzione "o" la situazione è ancora più complessa. Si è gia visto che, anche restando nell'ambito della vero-funzionalità, con "o" si può intendere sia $\vee$ (vel), sia + (aut), sia NAND (o incompatibile). L'uso di "o" nel linguaggio comune, quindi, è intrinsecamente ambiguo.

Persino l'uso del "non" può riservare sorprese. "Non ho visto niente" non equivale a "Ho visto qualcosa" (come sancito dalla regola di doppia negazione), "Non desidero partire" non è la negazione di "Desidero partire", ma significa "Desidero non partire". Si tenga presente che la negazione di "Ogni quadrato è un rettangolo" non è "Nessun quadrato è un rettangolo", ma "Esiste un quadrato che non è un rettangolo".

Il connettivo che solleva il maggior numero di difficoltà è il condizionale $\rightarrow$, il quale viene letto "se..., allora...". Il fatto è che, nel linguaggio comune, "se..., allora..." viene solitamente impiegato proprio per esprimere un nesso (causale) tra i contenuti

[26] Con $\wedge$ si possono tradurre altri connettivi del linguaggio comune. In genere, "ma" si può tradurre con $\wedge$ ("3 è dispari, ma 4 è pari"), come pure "mentre" ("3 è dispari, mentre 4 è pari") quando non ha anche il significato temporale (come in "Mentre mangio, guardo la TV").

dell'antecedente e del conseguente, e quindi il valore di verità di una proposizione del tipo "se A, allora B" deriva dal sussistere o meno di tale nesso (e non solo dal valore di verità di A e di B); in altri termini, gli usi del "se..., allora..." sono molto spesso non vero-funzionali. La tavola di verità del condizionale materiale corrisponde, come si è detto, all'uso del "se..., allora..." che si adotta nel ragionamento matematico. E' noto infatti che la gran parte del teoremi matematici ha proprio la forma "se A (ipotesi), allora B (tesi)" e che per dimostrare la validità di un teorema si accetta di procedere "per assurdo". Si suppone che il teorema sia falso (cioè che "se A, allora B" abbia valore di verità **F**) e si pone che A (l'ipotesi) sia vera e B (la tesi) falsa. Si ammette quindi che l'*unico caso* in cui "se A, allora B" ha valore di verità **F** è quando A ha valore **V** e B ha valore **F** (proprio come sancito nella tavola di verità di $\rightarrow$).

A ulteriore riprova della adeguatezza della tavola di verità di $\rightarrow$, consideriamo la proposizione aritmetica:

"per ogni x e per ogni y, se x è divisore di y, allora x è divisore di 2y".

Essa, come è noto, è vera nei numeri naturali, quindi, trattandosi di una proposizione quantificata universalmente, devono essere veri tutti i suoi casi particolari (ottenuti sostituendo i nomi dei numeri naturali al posto delle variabili individuali x e y).
Sono quindi vere:

"se 3 è divisore di 9, allora 3 è divisore di 18" (i)
"se 6 è divisore di 9, allora 6 è divisore di 18" (ii)
"se 5 è divisore di 9, allora 5 è divisore di 18" (iii)

Pertanto il condizionale ha valore **V** quando antecedente e conseguente hanno valore **V** (i), quando l'antecedente ha valore **F** e il conseguente ha valore **V** (ii) e quando sia l'antecedente che il conseguente hanno valore **F** (iii).

Al posto di "se A, allora B", si usano anche:

"se A, B"; "B, se A"; "A solo se B"; "A è condizione sufficiente per B";
"B è condizione necessaria per A"; "B supposto che A"; "solo se B, allora A".

Il bicondizionale "A se e solo se B" equivale a "se A, allora B" e "se B, allora A". Si osservi che, in "A se e solo se B", il primo "se" ("A, se B") indica il secondo condizionale ("se B, allora A"), mentre il "solo se" indica il primo condizionale ("se A, allora B").

"A se e solo se B", essendo la congiunzione delle due proposizioni "A è condizione necessaria per B" ("se B, allora A") e "A è condizione sufficiente per B" ("se A, allora B") si

esprime anche "A è condizione necessaria e sufficiente per B" ("condizione necessaria e sufficiente affinché B è A").

A differenza del condizionale, il bicondizionale è commutativo, per cui le precedenti proposizioni equivalgono anche a "B è condizione necessaria e sufficiente per A" ("condizione necessaria e sufficiente affinché A è B") in cui la condizione necessaria è il condizionale "se A, allora B" e la condizione sufficiente è il condizionale "se B, allora A".

Quando vale il condizionale "se A, allora B" (che è equivalente al suo contronominale "se non B, allora non A"), ma non vale il condizionale inverso "se B, allora A", per sottolineare la dissimmetria che si instaura fra le proposizioni A e B, si dice talora "A è condizione sufficiente, ma non necessaria, per B", "B è condizione necessaria, ma non sufficiente, per A".

Notiamo ancora che, spesso, sui testi si fa confusione tra il condizionale (che è un connettivo) e l'implicazione o deducibilità (che è una relazione fra proposizioni). Spetta proprio all'indagine logica evidenziare i rapporti tra i due concetti.

Potremmo fare molti altri esempi per sottolineare come nel linguaggio logico artificiale si riflettano solo alcuni aspetti del linguaggio naturale. Si può dire che la logica studia "modelli" ideali in modo per certi aspetti analogo a come la fisica studia i corpi rigidi e i gas perfetti, i quali non esistono in natura, ma possono fornire informazioni preziose, tanto più utili nella pratica quanto più i corpi e i gas reali si trovano in determinate situazioni fisiche. Il "modello", infatti, contiene dei costrutti teorici che si possono far corrispondere a entità più o meno osservabili degli oggetti reali. Così il linguaggio della logica dei predicati contiene degli elementi che si possono far corrispondere a espressioni del linguaggio naturale e, mediante la formalizzazione, si ottengono informazioni sui nessi logici fra le proposizioni, tanto più utili quanto più adeguata è stata la "traduzione" formale[27].

6. Conclusione

Il problema di stabilire la correttezza dei ragionamenti proposti all'inizio del §2 si può ora tradurre in quello di determinare se:

(a) da $Pa \rightarrow (Qb \wedge Qb')$, $(Qb \vee Qb') \rightarrow Ra'$ segue logicamente $\neg Ra' \rightarrow \neg Pa$;
(b) da $Pa \rightarrow (Qb \vee Qb')$, $(Qb \wedge Qb') \rightarrow \neg Ra'$ segue logicamente $Ra' \rightarrow \neg Pa$;
(c) da $\forall x(Rxa \rightarrow Px)$, Rba segue logicamente Pb;
(d) da $\forall x(Px \rightarrow Qx), \forall x(Qx \rightarrow Rx)$ segue logicamente $\forall x(Px \rightarrow Rx)$;

[27] Altre considerazioni riguardano i quantificatori. L'uso di "alcuni" è emblematico. La sua traduzione con $\exists$ è lecita solo quando è usato con il significato di "esiste almeno un". Spesso "alcuni" è invece usato con il significato di "più di uno e non tutti": quando diciamo "Alcuni di noi sono minorenni" intendiamo dire che vi è *più di un* minorenne e anche che vi sono tra noi alcuni che sono maggiorenni (altrimenti avremmo detto "Tutti siamo minorenni"); in questo caso non è lecita la traduzione con $\exists$.

(e) da $\exists x(Px \wedge Qx), \forall x(Px \rightarrow Rxa)$ segue logicamente $\exists x(Rxa \wedge Qx)$.

Le tecniche per affrontare queste questioni saranno approfondite nel seguito del corso. Per concludere questa lezione esaminiamo brevemente come si può procedere nel caso dei primi due.

In (a) e (b) figurano quattro formule atomiche chiuse (Pa, Qb, Qb', Ra') le quali, a seconda del significato attribuito alle costanti individuali e predicative (tutte a un solo argomento) possono risultare vere o false.

Per vedere se (a) è corretto occorre esaminare se, ogniqualvolta sono vere le due premesse, è vera la conclusione. Servendoci delle tavole di verità dei connettivi[28] (§3) possiamo esaminare, in corrispondenza di tutte le possibili assegnazioni di valori di verità alle quattro formule atomiche, se, quando le premesse sono vere, è vera anche la conclusione.

Si può procedere in un altro modo supponendo che le premesse siano vere e la conclusione falsa. Se la conclusione è falsa, ossia $\neg Ra' \rightarrow \neg Pa$ ha valore di verità **F**, allora, ricordando la tavola di verità del condizionale, deve essere $\neg Ra'$ vera e $\neg Pa$ falsa, ossia Ra' falsa e Pa vera. Se Pa è vera, avendo supposta vera la prima premessa, sempre per la tavola di verità del condizionale materiale, deve essere vera $Qb \wedge Qb'$, ossia, per la tavola di verità della congiunzione, Qb e Qb' devono essere entrambe vere. Ma, allora, per la tavola della disgiunzione, è vera $Qb \vee Qb'$, e, avendo supposta vera la seconda premessa, si ottiene che Ra' deve essere vera. Si è ottenuta una contraddizione, ossia che Ra' è al contempo falsa e vera. Quindi, non può darsi il caso che le premesse siano vere e la conclusione falsa e, pertanto, sussiste il nesso di conseguenza logica.

Procediamo analogamente nel caso di (b) supponendo che le premesse siano vere e la conclusione falsa. Se la conclusione è falsa ($Ra' \rightarrow \neg Pa$ ha valore di verità **F**), allora Ra' è vera e $\neg Pa$ falsa, ossia Ra' e Pa sono entrambe vere. Se Pa è vera, avendo supposta vera la prima premessa, si deduce che $Qb \vee Qb'$ è vera e, quindi, che almeno una tra Qb e Qb' è vera. Ora, se Qb è vera e Qb' falsa, $Qb \wedge Qb'$ è falsa e ciò non contraddice la verità della seconda premessa. A differenza del caso precedente non si è pervenuti ad alcuna contraddizione: se Pa, Qb e Ra' sono vere e Qb' falsa le due premesse sono vere e la conclusione falsa. Pertanto, il nesso di conseguenza logica non sussiste e il ragionamento non è corretto.

[28] In effetti, nel caso di (a) e (b), si tratta di ragionamenti a "livello proposizionale" nel senso che il sussistere del nesso (eventuale) di conseguenza logica dipende solo dai nessi tra le proposizioni atomiche che intervengono nelle premesse e nella conclusione, e non dalla loro forma specifica. Sarebbe stata sufficiente una formalizzazione "meno fine":
(a) Da $A \rightarrow (B \wedge C)$, $(B \vee C) \rightarrow D$ segue logicamente $\neg D \rightarrow \neg A$;
(b) Da $A \rightarrow (B \vee C)$, $(B \wedge C) \rightarrow \neg D$ segue logicamente $D \rightarrow \neg A$.

Numeri naturali e Principio di Induzione

DARIO PALLADINO
Dipartimento di Filosofia
Università di Genova
via Balbi 4
16126, Genova

1. Il principio di induzione completa

Consideriamo la somma dei primi n numeri dispari:

n = 2: $1 + 3 = 4 = 2^2$;
n = 3: $1 + 3 + 5 = 9 = 3^2$;
n = 4: $1 + 3 + 5 + 7 = 16 = 4^2$;
n = 5: $1 + 3 + 5 + 7 + 9 = 25 = 5^2$;
n = 6: $1 + 3 + 5 + 7 + 9 + 11 = 36 = 6^2$;
....

Sorge allora spontanea la congettura:

"la somma dei primi n numeri dispari è uguale a n^2".

Tuttavia, per dimostrare che la congettura è un teorema, ossia vale qualsiasi sia il numero naturale n, non è sufficiente aver esaminato un numero finito, anche enorme, di casi particolari.

La storia della matematica è piena di abbagli induttivi presi anche da eminenti matematici: proprietà vere per certi valori di n non sono vere per ogni n. Vediamo alcuni esempi[1].

[1] La bibliografia sul principio di induzione è molto vasta. La maggior parte degli esempi che qui proporremo è tratta da I. S. Sominsky, *Il metodo di induzione matematica*, Progresso Tecnico Editoriale, Milano, 1964, e L. I. Golovina e I. M. Yaglom, *L'induzione in geometria*, Progresso Tecnico Editoriale, Milano, 1966. Rinviamo all'ampia bibliografia del saggio M. Ferrari, "L'induzione matematica: storia e didattica", *Atti del Convegno*

Consideriamo i numeri della forma $2^{2^n}+1$. Per n = 0, 1, 2, 3 e 4 si ottengono numeri primi:

$$2^{2^0}+1=3;\quad 2^{2^1}+1=5;\quad 2^{2^2}+1=17;\quad 2^{2^3}+1=257;\quad 2^{2^4}+1=65537.$$

Fermat congetturò che tutti i numeri di tale forma fossero primi, ma Eulero, un secolo più tardi, scoprì che:

$$2^{2^5}+1=4294967297=641\times 6700417,$$

e quindi non è primo.

Il matematico sovietico D. A. Grave formulò la congettura che tutti i numeri della forma $2^{p-1}-1$ con p numero primo non fossero divisibili per p^2 (dopo averla verificata per i numeri primi minori di 1000). In seguito si trovò che $2^{1092}-1$ è divisibile per 1093^2 (e 1093 è un numero primo) e quindi che la congettura è falsa.

Se consideriamo la proprietà:

"il numero $n^2-79n+1601$ è primo"

essa è vera per n = 1, 2, 3,... fino a n = 79. Tuttavia per n = 80 è falsa, in quanto $80^2-79\times 80+1601=1681=41^2$.

Come ultimo esempio, se nell'espressione $991n^2+1$ si sostituiscono i valori n = 1, 2, 3,... non si ottiene mai un quadrato perfetto fino a quando non si sostituisce un numero molto grande di 29 cifre [2].

Questi esempi dimostrano che una proprietà P(n) può valere per moltissimi casi particolari, senza essere valida in generale. A differenza di quanto avviene nelle scienze sperimentali in cui si accetta l'induzione per enumerazione semplice, ossia la constatazione empirica di un gran numero di casi è ritenuta sufficiente per stabilire leggi di validità generale, in matematica, per poter dichiarare vera una proposizione del tipo:

"per ogni numero naturale n, P(n)"

Internazionale "Cultura matematica e insegnamento" nel decimo anniversario della scomparsa di Luigi Campedelli (30-31 maggio e 1 giugno 1988), Firenze, 1989, pp. 119-138, il quale contiene interessanti notizie storiche e considerazioni didattiche sul principio di induzione.

[2] E precisamente n = 1205573579033135944744253876 7.

(ossia la verità di ciascuna proposizione nell' elenco infinito P(0), P(1), P(2), P(3),..., P(n),...) non è sufficiente averla verificata per un numero finito di casi particolari, ma si ricorre al seguente principio, detto di *induzione completa* :

se una proprietà P(n) *vale per il numero zero* (*vale* P(0))·*e, quando vale per un arbitrario numero naturale* k, *vale per il successivo* k + 1 (*se* P(k), *allora* P(k + 1)), *allora la proprietà vale per tutti i numeri naturali.*

Servendoci delle costanti logiche introdotte nella lezione precedente possiamo schematizzare tale principio nel modo seguente:

(**A**) $P(0)$
$\forall k(P(k)\rightarrow P(k+1))$
$\overline{\qquad\qquad}$
$\forall n\, P(n)$

dove k e n (da ora in poi) sono variabili numeriche.

La prima premessa, cioè P(0), è detta *base* dell'induzione, la seconda premessa il *passo* dell'induzione e l'ipotesi P(k) del passo prende il nome di *ipotesi induttiva.* Intuitivamente, la base assicura che il primo numero naturale, cioè 0, ha la proprietà in questione, e il passo che la proprietà si trasmette da ogni numero al successivo: poiché iterando l'operazione di passaggio al successivo partendo da 0 si raggiungono via via tutti i numeri naturali, si conclude che la proprietà vale per tutti i numeri naturali.

La validità del principio di induzione si basa sulla struttura dei numeri naturali e, quindi, non è un principio di natura puramente logica[3].

Nonostante il nome di "principio di induzione" si tratta di uno strumento per compiere deduzioni, per dimostrare teoremi sui numeri e, in un certo senso, caratterizza i numeri naturali, tanto è vero che è comunemente assunto tra gli assiomi delle teorie assiomatiche per l'aritmetica[4].

[3] Anche se lo abbiamo rappresentato schematicamente come i ragionamenti corretti della lezione precedente, esso non ha validità generale, ossia qualsiasi siano i suoi contenuti, ma, come si è detto, vale per i numeri naturali.

[4] E' il quinto assioma delle celebre assiomatizzazione dell'aritmetica di Peano. La formalizzazione dell'aritmetica e le questioni logiche collegate con la formalizzazione del principio di induzione saranno esaminate in una lezione successiva. Per una panoramica di queste questioni vedi G. Lolli, "Il principio di induzione", *Notiziario dell'Unione Matematica Italiana*, Supplemento n.5, maggio 1993, pp. 65-98 (contenente anche una breve storia del principio) e la relativa bibliografia. Vedi anche V. M. Abrusci, "Due note sul principio di induzione matematica", *Atti del Congresso Nazionale di Logica*, Montecatini Terme, 1-5 ottobre 1979, Bibliopolis, Napoli, 1981, pp. 55-77.

Il principio di induzione completa può essere formulato anche in termini di insiemi di numeri naturali[5]:

se S *è un insieme di numeri naturali tale che* $0 \in S$ *e se* $k \in S$, *allora* $k+1 \in S$, *allora* $S = \mathbf{N}$
(in formula: $S \subseteq \mathbf{N} \wedge 0 \in S \wedge \forall k(k \in S \rightarrow k+1 \in S) \rightarrow S = \mathbf{N}$)

Talvolta capita che la base, anziché P(0), sia P(1), o P(2) o, in generale, P(h); in questo caso la conclusione è rispettivamente per ogni $n \geq 1$ vale P(n), per ogni $n \geq 2$ vale P(n), per ogni $n \geq h$ vale P(n)[6]:

$$\frac{\begin{array}{l} P(h) \\ \forall k \geq h\ (P(k) \rightarrow P(k+1)) \end{array}}{\forall n \geq h\ P(n)}$$

Vediamo ora alcuni esempi di applicazione del principio di induzione completa.

ESEMPIO 1 *Dimostrare che per ogni numero naturale* $n \geq 1$, *la somma dei primi n numeri naturali è* $\frac{n(n+1)}{2}$:

$$\textit{per ogni } n \geq 1 \textit{ vale } P(n) \equiv 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}.$$

Base. P(1) è $1 = \frac{1 \times 2}{2}$, ossia 1 = 1, e la base è dimostrata.

Passo. Nel passo si assume come ipotesi (ipotesi induttiva) che valga P(k), ossia:

$$1 + 2 + 3 + \ldots + k = \frac{k(k+1)}{2},$$

e si deve dimostrare P(k + 1):

$$1 + 2 + 3 + \ldots + k + (k+1) = \frac{(k+1)(k+2)}{2}.$$

[5] Si ricordi quanto esposto nella prima lezione relativamente all'identificazione delle proprietà con i sottoinsiemi. Nell'ambito della teoria degli insiemi si pone:

$0 = \varnothing \quad n+1 = n \cup \{n\}$

e si assume l'assioma dell'infinito: esiste un insieme S tale che

$0 \in S \wedge \forall k(k \in S \rightarrow k+1 \in S)$.

Si *definisce* poi l'insieme dei numeri naturali come l'intersezione (ossia il più piccolo) di tutti gli insiemi che hanno quest'ultima proprietà. E' allora immediato verificare che i numeri naturali soddisfano il principio di induzione proprio per come sono definiti.

[6] Questa forma di induzione da h (compreso) in avanti si riconduce immediatamente alla precedente applicando quest'ultima alla proprietà P(h+n).

A tal fine basta aggiungere (k + 1) ad ambo i membri di P(k) e svolgere i calcoli nel secondo membro:

$$1 + 2 + 3 + \ldots + k + (k+1) = \frac{k(k+1)}{2} + (k+1) = \frac{k(k+1)+2(k+1)}{2} = \frac{(k+1)(k+2)}{2}.$$

Per il principio di induzione completa la proprietà è dimostrata.

ESEMPIO 2 *Dimostrare che la somma dei primi* n *numeri dispari è uguale a* n^2:

$$\textit{per ogni } n \geq 2 \textit{ vale } P(n) \equiv 1 + 3 + 5 + \ldots + (2n-1) = n^2.$$

Base. P(2) è $1 + 3 = 2^2$, ossia 4 = 4 e quindi è ovvia[7].
Passo. Se vale P(k):

$$1 + 3 + 5 + \ldots + (2k-1) = k^2,$$

allora, aggiungendo (2k + 1) ad ambo i membri, si ottiene:

$$1 + 3 + 5 + \ldots + (2k-1) + (2k+1) = k^2 + 2k + 1 = (k+1)^2,$$

ossia P(k + 1), e il passo è dimostrato.

ESEMPIO 3 *Dimostrare che, per ogni* $n \geq 1$, *vale* P(n) ≡ "$n^3 + 11n$ è divisibile per 6".

Base. P(1) vale poiché 1 + 11 = 12 è divisibile per 6.
Passo. Supponiamo che valga P(k), ossia che il numero $k^3 + 11k$ sia divisibile per 6, e consideriamo $(k + 1)^3 + 11(k + 1)$. Svolgendo i calcoli si ha:

$$k^3 + 3k^2 + 3k + 1 + 11k + 11,$$

ossia

$$k^3 + 11k + 3(k^2 + k + 4).$$

[7] La base, molto spesso, si riduce ad una banale verifica, in quanto si tratta di constatare che un dato numero (di solito 0 o 1) soddisfa la proprietà. Non bisogna tuttavia sottovalutarla in quanto costituisce il punto di inizio, il fondamento, del procedimento induttivo. Consideriamo, ad esempio, la proprietà: "Ogni numero n è uguale al suo successivo (n = n+1)" evidentemente falsa. Relativamente ad essa si può dimostrare il passo (se la proprietà vale per un certo k (k = k+1), ne segue subito che k+1 = (k+1)+1, ossia k+1 = k+2, e la proprietà vale per k+1), ma non vale la base.

Essendo, per ipotesi induttiva, $k^3 + 11k$ multiplo di 6, per dimostrare P(k + 1) basta dimostrare che $3(k^2 + k + 4)$ è multiplo di 6, ossia che $k^2 + k + 4$ è pari.

Per dimostrare che $k^2 + k + 4$ è pari (e quindi concludere la dimostrazione), si può osservare che, se k è pari, allora k^2 è pari e quindi $k^2 + k + 4$ è pari in quanto somma di tre numeri pari, mentre, se k è dispari, anche k^2 è dispari, e quindi $k^2 + k$ è pari, e allora lo è anche $k^2 + k + 4$.

Oppure si può nuovamente procedere per induzione:

Poniamo Q(n) ≡ "$n^2 + n + 4$ è pari".

Base. Vale Q(0) in quanto 4 è pari.
Passo. Se $k^2 + k + 4$ è pari, allora $(k + 1)^2 + (k + 1) + 4 = (k^2 + k + 4) + 2k + 6$ è pari essendo somma di $k^2 + k + 4$ (pari per ipotesi induttiva) e $2k + 6 = 2(k + 3)$ (anch'esso pari).

Quindi, se vale Q(k), vale Q(k + 1) e allora, per ogni n, $n^2 + n + 4$ è pari.

ESEMPIO 4 *Dimostrare che il prodotto di tre numeri naturali consecutivi è multiplo di* 6.

Posto n il numero di mezzo dei tre consecutivi, si tratta di dimostrare che, per ogni $n \geq 1$, vale P(n) ≡ "(n − 1) n (n + 1) è multiplo di 6", ossia, "$n^3 - n$ è multiplo di 6".

Base. Dato che 0 è multiplo di 6, P(1) è verificata.
Passo. Se vale P(k), ossia se $k^3 - k$ è multiplo di 6, consideriamo $(k + 1)^3 - (k + 1)$ che è uguale a $(k^3 - k) + 3k(k + 1)$. Ora, $k^3 - k$ è multiplo di 6 per ipotesi induttiva e $3k(k + 1)$ è multiplo di 6 poiché almeno uno dei due numeri consecutivi k e k + 1 è pari. Quindi il numero $(k + 1)^3 - (k + 1)$, essendo somma di due numeri multipli di 6, è multiplo di 6, e quindi vale P(k + 1)[8].

ESEMPIO 5 *Dimostrare che, per ogni* n, $n^5 - n$ *è multiplo di* 30.

La proprietà si verifica subito per i primi valori di n. Supponiamo che valga per k e consideriamo:

$$(k + 1)^5 - (k + 1) = k^5 + 5k^4 + 10k^3 + 10k^2 + 5k + 1 - k - 1 =$$
$$k^5 - k + 5(k^4 + 2k^3 + 2k^2 + k).$$

Essendo, per ipotesi induttiva, $k^5 - k$ multiplo di 30, basta dimostrare che $k^4 + 2k^3 + 2k^2 + k$ è multiplo di 6. Dimostriamo allora per induzione che, per ogni n, vale:

[8] Dal punto di vista didattico si osservi che non sempre, per dimostrare una proprietà valida per tutti i numeri naturali, il ricorso al principio di induzione sia necessario. In questo esempio la tesi si ottiene più rapidamente - e più elegantemente - osservando che di tre numeri consecutivi almeno uno è pari e uno è multiplo di 3.

Q(n) ≡ "$n^4 + 2n^3 + 2n^2 + n$ è multiplo di 6".

La proprietà vale per i primi valori di n. Supponiamo che valga per k e consideriamo:

$$(k + 1)^4 + 2(k + 1)^3 + 2(k + 1)^2 + (k + 1) =$$
$$k^4 + 4k^3 + 6k^2 + 4k + 1 + 2k^3 + 6k^2 + 6k + 2 + 2k^2 + 4k + 2 + k + 1 =$$
$$k^4 + 2k^3 + 2k^2 + k + 2(2k^3 + 6k^2 + 7k + 3).$$

Essendo, per ipotesi induttiva, $k^4 + 2k^3 + 2k^2 + k$ multiplo di 6, è sufficiente dimostrare che $2k^3 + 6k^2 + 7k + 3$ è multiplo di 3.
Dimostriamo allora, sempre applicando il principio di induzione, che, per ogni n, vale:

R(n) ≡ "$2n^3 + 6n^2 + 7n + 3$ è multiplo di 3".

La proprietà vale per i primi valori di n. Supponiamo che valga per k e consideriamo:

$$2(k + 1)^3 + 6(k + 1)^2 + 7(k + 1) + 3 =$$
$$2k^3 + 6k^2 + 6k + 2 + 6k^2 + 12k + 6 + 7k + 7 + 3 =$$
$$2k^3 + 6k^2 + 7k + 3 + 3(2k^2 + 6k + 5)$$

Essendo, per l'ipotesi induttiva, $2k^3 + 6k^2 + 7k + 3$ multiplo di 3 e $3(2k^2 + 6k + 5)$ multiplo di 3, R(k + 1) è dimostrata.

ESEMPIO 6 *Dimostrare che, per ogni* $n \geq 1$, $7 \times 5^{2n-1} + 2^{3n+1}$ *è multiplo di* 17.

Base. Per n = 1 si ha 35 + 16 = 51 che è multiplo di 17.
Passo. Supponiamo che la proprietà valga per k, ossia che si abbia:

$$7 \times 5^{2k-1} + 2^{3k+1} = h \times 17, \text{ cioè } 2^{3k+1} = h \times 17 - 7 \times 5^{2k-1}.$$

Per k + 1 si ha:

$$7 \times 5^{2(k+1)-1} + 2^{3(k+1)+1} = 7 \times 5^{2k+1} + 2^{3k+4} = 7 \times 25 \times 5^{2k-1} + 8 \times 2^{3k+1}$$

e, sostituendo:

$$7 \times 5^{2(k+1)-1} + 2^{3(k+1)+1} = 175 \times 5^{2k-1} + h \times 17 \times 8 - 56 \times 5^{2k-1} =$$
$$119 \times 5^{2k-1} + h \times 17 \times 8$$

che è multiplo di 17 (essendo $119 = 17 \times 7$).

ESEMPIO 7 *Se* a *e* b *sono due numeri naturali (con* $b > 0$*), esistono due numeri naturali* q *e* r *tali che* $a = bq + r$ *e* $r < b$.

Si procede per induzione su a, ossia si considera la proprietà:

$$P(a) \equiv \text{“per ogni } b > 0 \text{ esistono q e r con } r < b \text{ tali che } a = bq + r\text{”}.$$

Base. $P(1) \equiv$ “per ogni $b > 0$ esistono q e r con $r < b$ tali che $1 = bq + r$”.
Se $b = 1$ basta prendere $q = 1$ e $r = 0$ e allora $1 = bq + r$ con $r < b$.
Se $b > 1$ basta prendere $q = 0$ e $r = 1$ e allora $1 = bq + r$ con $r < b$.
Passo. Supponiamo che la proprietà valga per a e dimostriamola per $a + 1$.
Dato che per ogni $b > 0$ esistono q e r con $r < b$ tali che $a = bq + r$, ne segue che:

$$a + 1 = bq + r + 1.$$

Se $r+1<b$, ponendo $q' = q$ e $r' = r+1$, allora $a+1 = bq' + r'$ (con $r' < b$).
Se $r+1 = b$, allora $a+1 = bq+b = b(q+1)$ e, ponendo $q' = q+1$ e $r' = 0$, si ha $a+1 = bq' + r'$ (con $r' < b$).

La proprietà P(n), pur contenendo una variabile numerica, può anche non essere riferita ai numeri naturali.

ESEMPIO 8 *Sia* x *un numero reale* > -1. *Dimostrare che, per ogni* n, *risulta*:

$$(1 + x)^n \geq 1 + nx.$$

Base. Per $n = 0$ si ha $1 \geq 1$ che è vera.
Passo. Supponiamo che la proprietà valga per k:

$$(1 + x)^k \geq 1 + kx.$$

Dato che $x > -1$, $1 + x > 0$ e, quindi, moltiplicando ambo i membri per $1 + x$, si ha:

$$(1 + x)^{k+1} \geq (1 + kx)(1 + x) = 1 + (k + 1)x + kx^2 \geq 1 + (k + 1)x$$

e il passo è dimostrato.

ESEMPIO 9 *Dimostrare che, per ogni numero reale* $x \neq k\pi$, *e per ogni numero naturale* $n \geq 1$, *vale:*

$$P(n) \equiv \cos x + \cos 3x + \cos 5x + \ldots + \cos(2n-1)x = \frac{\operatorname{sen}(2nx)}{2\operatorname{sen} x}$$

Base. $P(1)$ è $\cos x = \dfrac{\operatorname{sen}(2x)}{2\operatorname{sen} x} = \dfrac{2\operatorname{sen} x \cos x}{2\operatorname{sen} x} = \cos x$, e quindi vale.

Passo. Supponiamo che valga P(k):

$$\cos x + \cos 3x + \cos 5x + \ldots + \cos(2k-1)x = \frac{\operatorname{sen}(2kx)}{2\operatorname{sen} x}$$

e dimostriamo P(k+1):

$$\cos x + \cos 3x + \cos 5x + \ldots + \cos(2k-1)x + \cos(2k+1)x = \frac{\operatorname{sen}(2k+2)x}{2\operatorname{sen} x}$$

Si ha:

$$\cos x + \cos 3x + \cos 5x + \ldots + \cos(2k-1)x + \cos(2k+1)x =$$

$$= \frac{\operatorname{sen}(2kx)}{2\operatorname{sen} x} + \cos(2k+1)x = \frac{\operatorname{sen}(2kx)}{2\operatorname{sen} x} + \cos(2kx + x) =$$

$$= \frac{\operatorname{sen}(2kx)}{2\operatorname{sen} x} + \cos(2kx)\cos x - \operatorname{sen}(2kx)\operatorname{sen} x =$$

$$= \frac{\operatorname{sen}(2kx) + \cos(2kx)2\operatorname{sen} x\cos x - \operatorname{sen}(2kx)2\operatorname{sen}^2 x}{2\operatorname{sen} x} =$$

$$= \frac{\operatorname{sen}(2kx)(1 - 2\operatorname{sen}^2 x) + \cos(2kx)2\operatorname{sen} x\cos x}{2\operatorname{sen} x} =$$

$$= \frac{\operatorname{sen}(2kx)\cos 2x + \cos(2kx)\operatorname{sen} 2x}{2\operatorname{sen} x} = \frac{\operatorname{sen}(2kx + 2x)}{2\operatorname{sen} x} = \frac{\operatorname{sen}(2k+2)x}{2\operatorname{sen} x}$$

ESEMPIO 10 *Dimostrare che, diviso il piano con* n *rette, si può colorare ciascuna regione in rosso o nero il modo che regioni adiacenti (con un segmento in comune) siano di colore diverso.*

Base. La proprietà è ovvia per $n = 1$, poiché in tal caso si hanno solo due regioni (i due semipiani) che possono essere colorati uno in rosso e l'altro in nero.
Passo. Supponiamo che la proprietà valga per $n = k$, ossia quando si sono tracciate k rette, e dimostriamola per $k + 1$ rette. Consideriamo una retta r qualsiasi delle $k + 1$ e, sfruttando l'ipotesi induttiva, coloriamo nel modo richiesto le regioni individuate dalle rette esclusa r (che sono in numero di k). Consideriamo ora i due semipiani individuati da r e, in uno dei

due, cambiamo il colore di tutte le regioni. Dimostriamo che la nuova colorazione soddisfa la condizione richiesta. Consideriamo due regioni adiacenti. Se la retta che contiene il segmento che le separa non è r, allora erano colorate in un modo diverso e continuano ad esserlo sia che i loro colori siano rimasti inalterati, siano che siano stati entrambi cambiati. Se la retta che le separa è r allora le due regioni avevano lo stesso colore prima del cambiamento e ora in una delle due parti il colore è stato invertito.

ESEMPIO 11 *Provare che il numero dei sottoinsiemi di un insieme finito di* n *elementi è* 2^n.

Base. Se l'insieme A ha 0 elementi (n = 0), allora è l'insieme vuoto e, quindi, ha come unico sottoinsieme l'insieme vuoto stesso, ossia il numero dei sottoinsiemi è 1 e $1 = 2^0$ (se A ha un elemento, allora ha due sottoinsiemi, l'insieme vuoto e A stesso e $2 = 2^1$).
Passo. Supponiamo che un insieme qualsiasi di k elementi abbia 2^k sottoinsiemi e sia A un insieme di k + 1 elementi. Si deve dimostrare che i sottoinsiemi di A sono 2^{k+1}. Consideriamo un qualsiasi $a \in A$. Dividiamo i sottoinsiemi di A in due classi disgiunte ponendo in una classe i sottoinsiemi che non hanno a come elemento e nell'altra i rimanenti (quelli che contengono a). I sottoinsiemi della prima classe sono evidentemente i sottoinsiemi di A – {a}, il quale ha k elementi e, quindi, per ipotesi induttiva, sono in numero di 2^k. Dato che vi è una corrispondenza biunivoca tra le due classi di sottoinsiemi (se si fa corrispondere ad un sottoinsieme della prima classe il sottoinsieme della seconda ottenuto aggiungendogli l'elemento a, tale corrispondenza è evidentemente iniettiva e suriettiva), anche i sottoinsiemi della seconda classe sono 2^k. In totale, i sottoinsiemi di A sono $2^k + 2^k = 2^{k+1}$

2. Induzione sul decorso dei valori, principio del minimo intero e della discesa infinita
Un'altra formulazione del principio di induzione, detta *induzione sul decorso dei valori*, che è poi quella più frequentemente usata nelle dimostrazioni dei teoremi di logica, sfrutta un'ipotesi induttiva (apparentemente) più forte (anziché "indurre" da n a n + 1 si "induce" da tutti i predecessori di n + 1 a n + 1):

se una proprietà P(n) *vale per* 0 *e, quando vale per tutti i numeri naturali da* 0 *a* k *(estremi inclusi) vale per* k + 1, *allora vale per tutti i numeri naturali.*

Schematicamente:

(**B**) $\dfrac{P(0) \quad \forall k(P(0) \wedge P(1) \wedge ... \wedge P(k) \rightarrow P(k+1))}{\forall n\, P(n)}$

o anche:

$$\frac{P(0) \quad \forall k(\forall m(m \leq k \rightarrow P(m)) \rightarrow P(k+1))}{\forall n\, P(n)}$$

Si può dimostrare che questa nuova formulazione (**B**) è equivalente alla (**A**) precedente[9]. Come nel caso precedente la base, anziché P(0), può essere P(h), il passo:

$$\forall k \geq h\ (\forall m(h \leq m \leq k \rightarrow P(m)) \rightarrow P(k+1))$$

e allora la conclusione è $\forall n \geq h\ P(n)$.

Vi sono poi altre varianti in cui si considerano due o alcuni predecessori di k + 1, anziché il solo predecessore immediato k (prima formulazione) o tutti i predecessori (seconda formulazione).

La seconda formulazione può essere proposta in modo da avere una sola premessa (senza la base):

se una proprietà P(n), *quando vale per tutti i numeri naturali minori di* k, *vale per* k, *allora vale per tutti i numeri naturali.*

Schematicamente:

(**B'**) $\dfrac{\forall k(P(0) \wedge P(1) \wedge ... \wedge P(k-1) \rightarrow P(k))}{\forall n\, P(n)}$

o anche:

$$\frac{\forall k(\forall\, m(m < k \rightarrow P(m)) \rightarrow P(k))}{\forall n\, P(n)}$$

[9] E' ovvio che la nuova formulazione (**B**) implica la precedente formulazione (**A**): le basi coincidono e, se vale il passo di (**A**), vale a maggior ragione il passo di (**B**). Quindi, se valgono le ipotesi di (**A**), valgono quelle di (**B**), e, per (**B**), si può concludere che la proprietà vale per tutti i numeri, che è anche la conclusione di (**A**).
Per dimostrare l'implicazione inversa, ossia che (**B**) segue da (**A**), supponiamo di avere una proprietà P(n) che soddisfa le ipotesi della seconda formulazione (**B**) del principio di induzione.
Poniamo $Q(n) = P(0) \wedge P(1) \wedge ... \wedge P(n)$. Per la prima ipotesi vale Q(0) (= P(0)). Supponiamo che valga $Q(k) = P(0) \wedge P(1) \wedge ... \wedge P(k)$. Per la seconda ipotesi si ha che vale P(k+1), e quindi vale $P(0) \wedge P(1) \wedge ... \wedge P(k) \wedge P(k+1) = Q(k+1)$. Si ha dunque Q(0) e $\forall k(Q(k) \rightarrow Q(k+1))$. Dalla prima formulazione (**A**) del principio di induzione (applicato a Q(n)) segue $\forall n Q(n)$. Poiché, evidentemente, si ha che $Q(n) \rightarrow P(n)$, da $\forall n Q(n)$ segue $\forall n P(n)$, ossia la conclusione di (**B**).

In questo caso l'induzione procede dai numeri minori di k a k e non è necessaria la base. Infatti, se vale il passo, dato che non vi sono numeri naturali minori di 0, segue banalmente P(0)[10].

La versione insiemistica è:

se S *è un insieme di numeri naturali tale che, se contiene i numeri naturali minori di* k *allora contiene* k, *allora* S *coincide con l'insieme dei numeri naturali*
(in formula: $S \subseteq \mathbf{N} \wedge \forall k(\forall m(m < k \rightarrow m \in S) \rightarrow k \in S) \rightarrow S = \mathbf{N}$)[11].

Una terza formulazione prende il nome di *principio del minimo intero* :

(**C**) *ogni insieme non vuoto di numeri naturali ha elemento minimo*
(in formula: $S \subseteq \mathbf{N} \wedge S \neq \varnothing \rightarrow \exists m(m \in S \wedge \forall k(k < m \rightarrow k \notin S))$).

Essa è equivalente alle precedenti. L'equivalenza, in breve, si dimostra come segue. Assumiamo il principio del minimo intero (**C**) e che una proprietà verifichi le premesse del principio di induzione completa (**A**). Supponiamo, per assurdo, che non valga la conclusione di (**A**), ossia che vi siano dei numeri naturali n per cui non valga P(n). Per (**C**) l'insieme S costituito dai numeri naturali per cui non vale P(n) non è vuoto, e allora ha un minimo elemento m (per cui non vale P(m)). Non può essere m = 0 in quanto, per ipotesi, vale P(0). Ma allora m ha un predecessore m–1 per cui vale P(m – 1) (essendo m il minimo numero per cui P(n) non vale). Ma, per la seconda ipotesi di (**A**), se vale P(m – 1), allora vale P(m), in contraddizione con quanto prima stabilito.

Per vedere l'implicazione inversa assumiamo, per assurdo, che (**C**) non valga, ossia che esista un sottoinsieme non vuoto S di **N** senza minimo elemento. Ragioniamo sul complementare S' di S in **N**. $0 \in S'$ in quanto, se $0 \in S$, allora ovviamente 0 sarebbe il minimo di S. Se tutti i numeri fino a k compreso sono in S', allora anche k + 1 è in S' (altrimenti k + 1 sarebbe il minimo di S). Ma allora, per (**B**) (che è equivalente ad (**A**)), S' = **N**, contro l'ipotesi che S non sia vuoto.

Una ulteriore formulazione del principio di induzione è il *principio della discesa infinita*: per dimostrare che tutti i numeri naturali hanno una proprietà P(n) si fa vedere che, se un numero naturale k non ha la proprietà (se vale ¬P(k)), allora vi è un numero naturale

[10] $\forall m(m < 0 \rightarrow P(m))$ vale perché la sua negazione equivale a $\exists m(m < 0 \wedge \neg P(m))$ che è falsa poiché non esiste alcun numero naturale minore di 0.

[11] Si dimostri che anche questa variante della seconda formulazione del principio di induzione consegue dalla prima formulazione (*Avviamento*: Dato l'insieme S che soddisfa $S \subseteq \mathbf{N} \wedge \forall k(\forall m(m < k \rightarrow m \in S) \rightarrow k \in S)$, si consideri l'insieme:
$S' = \{k: \forall m(m < k \rightarrow m \in S)\}$
e si dimostri che $S' \subseteq S$ e $0 \in S'$ e $k \in S' \rightarrow k+1 \in S'$. Dalla prima formulazione segue che S'=**N** e, quindi, S=**N**.

minore di k che non ha la proprietà. Siccome non esiste alcuna catena discendente infinita di numeri naturali, si conclude che k non può non avere la proprietà, e quindi, essendo k generico, che tutti i numeri la possiedono. In formula:

(**D**) $\forall k(\neg P(k) \rightarrow \exists m(m < k \wedge \neg P(m))) \rightarrow \forall n P(n)$

L'equivalenza con (**C**) è quasi immediata: se vale l'ipotesi di (**D**), allora l'insieme S dei numeri naturali per cui non vale la proprietà P(n), se non è vuoto, non ha minimo. Quindi, per (**C**), S è vuoto, ossia vale $\forall n P(n)$. Per il viceversa, procediamo per assurdo supponendo che vi sia un sottoinsieme non vuoto di S di **N** senza minimo elemento. Sia P(n) la proprietà di non appartenere a S ($n \notin S$). Se vale ¬P(k), ossia $k \in S$, dato che k non può essere il minimo di S, allora esiste m con $m < k$ tale che $m \in S$, ossia tale che ¬P(m). Per (**D**) segue $\forall n P(n)$, ossia $\forall n(n \notin S)$, in contraddizione con l'ipotesi che S non sia vuoto.

ESEMPIO 1 *Dimostrare che ogni numero naturale* ≥ 2 *si può esprimere come prodotto di numeri primi* (*prodotto che può ridursi a un solo fattore*).

La proprietà vale per 2 (che è primo). Procediamo per induzione sul decorso dei valori supponendo che la proprietà valga per 2, 3,..., k e dimostrandola per k + 1.
Se k + 1 è primo la proprietà è immediata. Se k + 1 non è primo, allora si ha $k + 1 = r \times s$ con $r, s \leq k$. Per ipotesi induttiva r e s si possono esprimere come prodotto di primi e allora, essendo un prodotto di prodotti ancora un prodotto, anche k + 1 è prodotto di primi.

ESEMPIO 2 *Provare che, per ogni numero naturale n, esiste un numero naturale* m *tale che*:

$$m^2 \leq n < (m + 1)^2 \quad (\textit{esistenza della radice quadrata intera})$$

La dimostrazione è immediata per n = 0 (basta porre m = 0) e per n = 1 (basta porre m = 1). Supponiamo allora $n \geq 2$.

Sia $S = \{k \in \mathbf{N}: k^2 > n\}$.

S non è vuoto in quanto $n \in S$ (per ogni $n \geq 2$, $n^2 > n$[12]). Per il principio del minimo intero S ha elemento minimo h. Dato che $h \in S$, $h^2 > n$ e quindi $h > 1$.
Se poniamo $m = h - 1$, m soddisfa le condizioni richieste:

$m^2 \leq n$ (essendo h il minimo numero naturale il cui quadrato supera n)
$n < (m + 1)^2 = h^2$ (dato che $h \in S$).

[12] Dimostrarlo per esercizio con il principio di induzione completa.

ESEMPIO 3 *Dimostrare che ogni insieme non vuoto di numeri naturali superiormente limitato è dotato di elemento massimo.*

Sia S un sottoinsieme non vuoto di **N** superiormente limitato. L'insieme S' dei maggioranti di S (S' = { k∈**N** : ∀x(x ∈ S→ x < k)}) non è vuoto e, quindi, per il principio del minimo intero, ha minimo elemento m. Non può essere m = 0 altrimenti, essendo m maggiorante di S, S dovrebbe essere vuoto contro l'ipotesi. Poniamo allora h = m − 1. Deve essere h ∈ S (in quanto, come è immediato verificare, se fosse h ∉ S, h sarebbe un maggiorante di S minore del minimo dei maggioranti). Quindi h è il massimo di S (se vi fosse in S un elemento n maggiore di h, sarebbe n ≥ h + 1 = m contro il fatto che m è un maggiorante di S).

ESEMPIO 4 *Dimostrare che $\sqrt{2}$ è irrazionale.*

Si tratta di far vedere che non esistono due numeri naturali a e b tali che $a^2 = 2b^2$. Sia P(n) ≡ $\neg\exists h(n^2 = 2h^2)$. Procediamo applicando il principio della discesa infinita. Supponiamo che valga ¬P(k), ossia che esista un numero h (che risulta minore di k) tale che $k^2 = 2h^2$. Ma allora k è pari; posto k = 2s, ne segue che $4s^2 = 2h^2$, ossia $h^2 = 2s^2$, e quindi vale ¬P(h). Per il principio della discesa infinita (**D**) segue $\forall n\neg\exists h(n^2 = 2h^2)$, ossia $\forall n\forall h(n^2 \neq 2h^2)$, che è quanto si voleva dimostrare.

3. Definizioni induttive

I procedimenti induttivi possono essere utilizzati anche per definire. Accanto al principio di induzione completa (e alle formulazioni ad esso equivalenti), il quale è essenzialmente uno strumento per *dimostrare* proposizioni quantificate universalmente, si può enunciare un principio di *definizione* per induzione (o ricorrenza) largamente usato in aritmetica e nelle applicazioni (soprattutto in Informatica).

Partiamo da un esempio molto elementare.

Supponiamo di voler calcolare le permutazioni P_n di n oggetti distinti. La soluzione non è immediata se già non si hanno nozioni preliminari di calcolo combinatorio. Si può allora cercare di vedere come cambia la situazione quando si passa da n a n + 1 oggetti, ossia di risolvere il problema per n + 1 oggetti supposto di averlo già risolto per n oggetti.

Ad esempio, se gli oggetti sono a, b e c una permutazione è:

bca

Se si ha un quarto oggetto d (ossia si passa da 3 a 4 oggetti), esso può essere collocato o davanti, o in una posizione intermedia, o al fondo (in totale in 4 posizioni):

dbca, bdca, bcda, bcad

Si riconosce immediatamente che, se gli oggetti sono n, le possibili posizioni in cui collocare l'ulteriore oggetto sono n + 1. Quindi, in corrispondenza di ogni permutazione di n oggetti, si ottengono n + 1 permutazioni di n + 1 oggetti. Pertanto:

$$P_{n+1} = (n + 1)P_n$$

Questa relazione, unita alla constatazione immediata che $P_1 = 1$, consente di affermare che il problema è risolto per n = 1 e, se risolto per n, è risolto per n + 1.

A questo punto si può calcolare che:

$$P_1 = 1; P_2 = 2\cdot1; P_3 = 3\cdot2\cdot1; P_4 = 4\cdot3\cdot2\cdot1; \ldots$$

e pervenire alla formula:

$$P_n = n\cdot(n - 1)\cdot\ldots\cdot4\cdot3\cdot2\cdot1 = n!$$

che si può dimostrare facilmente con il principio di induzione completa.

Nel §1 si è sfruttato il fatto che il numero dispari di posto n si può esprimere come 2n−1. Di solito, i numeri dispari si presentano come successione:

1, 3, 5, 7,...

e, quindi, mediante la relazione ricorrente:

$$d_1 = 1;\ d_{n+1} = d_n + 2.$$

Mediante queste relazioni si dimostra facilmente per induzione che:

$$d_n = 2n - 1.$$

Nelle trattazioni rigorose l'addizione si introduce mediante le seguenti clausole:

$$\begin{cases} a + 0 = a \\ a + s(b) = s(a + b) \end{cases}$$

dove con s si indica l'operazione di passaggio al successivo che viene assunta come primitiva (ossia si scrive s(k) anziché k + 1 [13]).

[13] Ciò per tenere distinta l'operazione di addizione (che viene ora definita) dalla funzione di passaggio al successivo (che viene assunta come primitiva). Nei contesti assiomatici, che saranno illustrati in una lezione successiva, il fatto che s(k) = k+1 viene dimostrato (dalla seconda clausola, essendo, per definizione, s(0) = 1, si ha: k+1 = k+s(0) = s(k+0) e s(k+0) = s(k) per la prima clausola).

La prima clausola stabilisce cosa significa addizionare 0, la seconda cosa significa addizionare il successivo s(b) di b quando si sa già addizionare b[14].

Così la moltiplicazione, una volta definita l'addizione, si introduce con le due seguenti clausole:

$$\begin{cases} a \times 0 = 0 \\ a \times s(b) = a \times b + a \end{cases}$$

e l'elevamento a potenza con le due clausole:

$$\begin{cases} a^0 = 1 \\ a^{s(b)} = a^b \times a \end{cases}$$

Le varie proprietà delle operazioni si dimostrano facendo ricorso al principio di induzione completa. Vediamo un paio di esempi.

ESEMPIO 1 *Dimostrare la proprietà associativa dell'addizione* :

per ogni a, b, c $(a + b) + c = a + (b + c)$.

Poniamo P(c) la proprietà "(a + b) + c = a + (b + c)", ossia lasciamo a e b come parametri e procediamo per induzione su c.

Base. P(0) è "(a + b) + 0 = a + (b + 0)". Infatti: (a + b) + 0 = a + b e a + (b + 0) = a + b per la prima clausola dell'addizione.
Passo. Supposto che valga P(c), dimostriamo P(s(c)), ossia:

$$(a + b) + s(c) = a + (b + s(c)).$$

Si ha:

(a + b) + s(c) =	(per la II clausola dell'addizione)
s((a + b) + c) =	(per l'ipotesi induttiva P(c))
s(a + (b + c)) =	(per la II clausola dell'addizione)
a + s(b + c) =	(per la II clausola dell'addizione)
a + (b + s(c)).	

[14] Nell'ottica assiomatica tutto ciò che non è assunto come assioma va dimostrato, e quindi anche affermazioni quali 3+2 = 5 possono essere dimostrate.
Posto, per definizione, s(0)=1, s(1)=2, s(2)=3, ...(ossia scegliendo la numerazione decimale), si ha: 3+2 = 3+s(1) = (per la II clausola) s(3+1) = s(3+s(0)) = (per la II clausola) s(s(3+0)) = (per la I clausola) s(s(3)) = s(4) = 5.

ESEMPIO 2 *Dimostrare che, per ogni* a, b e c:

$$a^{b+c} = a^b \times a^c.$$

Procediamo come nell'esempio precedente per induzione su c assumendo come P(c) la proprietà "$a^{b+c} = a^b \times a^c$".

Base. Vale P(0), cioè $a^{b+0} = a^b \times a^0$. Infatti, $a^{b+0} = a^b$ (per la I clausola della addizione) e $a^b \times a^0$ = (per la I clausola dell'elevamento a potenza) $a^b \times 1 = a^b \times s(0)$ = (per la II clausola della moltiplicazione) $= a^b \times 0 + a^b$ = (per la I clausola della moltiplicazione) $0 + a^b$ = (per la proprietà commutativa dell'addizione[15] e la I clausola della addizione) a^b.
Passo. Assumiamo P(c) e dimostriamo P(s(c)). Si ha:

$a^{b+s(c)}$ =	(per la II clausola dell'addizione)
$a^{s(b+c)}$ =	(per la II clausola dell'elevamento a potenza)
$a^{b+c} \times a$ =	(per l'ipotesi induttiva P(c))
$(a^b \times a^c) \times a$ =	(per la proprietà associativa della moltiplicazione[16])
$a^b \times (a^c \times a)$ =	(per la II clausola dell'elevamento a potenza)
$a^b \times a^{s(c)}$.	

Come si è visto a proposito del principio di induzione, si possono proporre definizioni induttive di carattere più generale.

ESEMPIO 3 La successione dei numeri di Fibonacci[17] è definita come segue:

$$\begin{cases} F_1 = 1 \\ F_2 = 1 \\ F_{n+1} = F_n + F_{n-1} \qquad \text{per } n \geq 2 \end{cases}$$

e i suoi primi elementi sono:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

Non è immediato trovare una espressione esplicita per F_n. Una formula per F_n, detta di Binet, è la seguente:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1+\sqrt{5}}{2}\right)^n\right]$$

[15] Che supponiamo di aver già dimostrato.
[16] Che supponiamo di aver già dimostrato.
[17] Cfr., ad esempio, N. N. Vorobyov, *I numeri di Fibonacci*, Progresso Tecnico Editoriale, Milano 1965.

Per dimostrare la validità di tale formula si può ricorrere al principio di induzione sul decorso sui valori.

Si verifica immediatamente che, per n = 1, si ha $F_1 = 1$.

Supponiamo che la proprietà valga per i numeri minori di k + 1 e dimostriamola per k + 1.

Poniamo, per brevità, $a = \frac{1+\sqrt{5}}{2}$ e $b = \frac{1-\sqrt{5}}{2}$.

Per ipotesi induttiva si ha: $F_k = \frac{1}{\sqrt{5}}(a^k - b^k)$ e $F_{k-1} = \frac{1}{\sqrt{5}}(a^{k-1} - b^{k-1})$.

Quindi:

$$F_{k+1} = F_k + F_{k-1} = \frac{1}{\sqrt{5}}(a^k - b^k + a^{k-1} - b^{k-1}) = \frac{1}{\sqrt{5}}(a^{k-1}(a+1) - b^{k-1}(b+1)).$$

D'altra parte si ha che:

$$a^2 = \frac{1+5+2\sqrt{5}}{4} = 1 + \frac{2+2\sqrt{5}}{4} = 1 + \frac{1+\sqrt{5}}{2} = 1 + a$$

e, analogamente, $b^2 = 1 + b$, per cui, sostituendo:

$$F_{k+1} = \frac{1}{\sqrt{5}}(a^{k-1}a^2 - b^{k-1}b^2) = \frac{1}{\sqrt{5}}(a^{k+1} - b^{k+1}),$$

e la formula di Binet vale per k + 1.

ESEMPIO 4 Il gioco della torre di Hanoi. Nel gioco della torre di Hanoi bisogna trasferire n anelli di diverse dimensioni da un piolo verticale P_1 (in cui sono infilati dal più largo al più stretto) a un piolo P_3 (inizialmente vuoto) servendosi di un terzo piolo P_2 (anch'esso inizialmente vuoto) in modo che non si sovrapponga mai un anello più grande a un anello più piccolo. Indichiamo con m_n il minimo numero di mosse necessario a completare il gioco.

Dato che, evidentemente, $m_1 = 1$, cerchiamo di calcolare m_n supponendo di avere già calcolato i valori m_k con $k < n$.

Ragioniamo come segue. Per completare il gioco quando vi sono n anelli è necessario che, ad un certo punto, si possa trasferire l'anello più largo dal piolo P_1 al piolo P_3 (che deve essere sgombro). Occorre quindi che i restanti n – 1 anelli siano posti (sempre in ordine crescente dall'alto in basso) sul piolo P_2. Ciò significa che si sono impiegate m_{n-1} mosse appunto per spostare gli n – 1 anelli da P_1 al piolo P_2 (è come aver svolto il gioco ignorando l'anello più grande e usando come piolo finale P_2 anziché P_3). A questo punto si può eseguire la mossa di spostare l'anello più grande dal piolo P_1 al piolo P_3. Bisogna poi trasferire gli n – 1 anelli da P_2 a P_3 (ed è come ripetere il gioco con n – 1 anelli con piolo iniziale P_2).

In definitiva si ha:

$$m_n = m_{n-1} + 1 + m_{n-1} = 2m_{n-1} + 1.$$

I primi termini sono:

$$m_1 = 1;\ m_2 = 3;\ m_3 = 7;\ m_4 = 15;\ m_5 = 31;\ \ldots$$

Viene allora spontanea la congettura:

$$m_n = 2^n - 1$$

che proviamo a dimostrare per induzione completa.

La formula vale per n = 1.

Supposto che valga per k ($m_k = 2^k - 1$), dimostriamola per k + 1:

$$m_{k+1} = 2m_k + 1 = 2(2^k - 1) + 1 = 2^{k+1} - 2 + 1 = 2^{k+1} - 1.$$

La giustificazione delle definizioni per induzione non si basa sul solo principio di induzione. Basta riflettere sul fatto che il principio di induzione è un principio per dimostrare, mentre le definizioni induttive, consentendo appunto di definire delle funzioni, si basano sull'esistenza delle funzioni aventi le caratteristiche desiderate[18]. Riferendoci al precedente esempio dell'addizione, si tratta di dimostrare che esiste una (ed una sola[19]) funzione f (di due variabili, essendovi il parametro a) tale che:

$$f(a,0) = a;\quad f(a, s(b)) = s(f(a,b))\ \ [20].$$

Questo problema fu risolto per la prima volta da Dedekind mediante la dimostrazione del teorema di ricorrenza che, nel caso più semplice (senza parametri), si può formulare nel modo seguente:

se X *è un insieme*, x *un elemento di* X *e* F *è una funzione di dominio* $\mathbf{N} \times X$ *e codominio* X, *allora esiste una ed una sola funzione* f *di dominio* **N** *e codominio* X *tale che* $f(0) = x$ *e* $f(s(n)) = F(n, f(n))$.

[18] Si può inoltre rilevare che le definizioni induttive non hanno forma "definiendum = definiens" che è richiesta per l' eliminabilità del definiendum. Considerando, ad esempio, la definizione dell'addizione, essa non ha la forma "a + b = ..." e, quindi, non si vede come si possa eliminare il simbolo definito + da una formula quale "a + b = b + a" (nelle clausole induttive il simbolo definito compare anche a destra del segno di =).

[19] L'unicità si dimostra facilmente con il principio di induzione.

[20] Una volta dimostrata l'esistenza di una f siffatta, allora a + b = f(a,b) e si può eliminare il definiendum dalle espressioni contenenti +.

Ponendo $X = \mathbf{N}$, $x = a$, $F(n, x) = s(x)$, in base al teorema esiste esattamente una funzione f_a tale che $f_a(0) = a$ e $f_a(s(n))=s(f_a(n))$; ponendo ora $f(a,b) = f_a(b)$, si ha la funzione di cui sopra.

Nel caso del fattoriale, poniamo $X = \mathbf{N}$, $x = 1$ e $F(n, x) = x(n + 1)$ e allora $n! = f(n)$, dove f è la funzione la cui esistenza e unicità è stabilita dal teorema di ricorrenza.

In tutti gli altri casi si può procedere in modo analogo.

4. Le definizioni e le dimostrazioni induttive in logica

Le definizioni e le dimostrazioni induttive non riguardano solo i numeri naturali, ma anche quegli insiemi i cui elementi sono ottenuti partendo da un insieme di elementi scelti inizialmente e iterando un numero finito di volte l'applicazione di alcune determinate operazioni. Esempi di insiemi di questo tipo li abbiamo incontrati nel §4 della Lezione II, e precisamente l'insieme dei termini e l'insieme delle fbf del linguaggio della logica dei predicati.

Per entrambi la prima clausola ((T1) o (F1)) stabilisce quali sono gli elementi iniziali, le altre clausole, tranne l'ultima, stabiliscono le operazioni con le quali si ottengono nuovi elementi a partire da elementi già dati, e l'ultima clausola ((T3) o (F5)) esprime una condizione di chiusura. Così (T3) afferma che sono termini *solo* le sequenze finite di simboli dell'alfabeto che sono o costanti o variabili individuali (T1) o che sono ottenute da n termini concatenati già dati anteponendo ad essi un simbolo di funzione a n argomenti (T2), ossia che l'insieme dei termini è il minimo insieme che soddisfa le condizioni (T1) e (T2). (F5) ha un analogo significato relativamente alle fbf.

Il principio di induzione si può riformulare per l'insieme dei termini: per dimostrare che tutti i termini hanno una certa proprietà P basta dimostrare che:

(a) *Base.* Le costanti e le variabili individuali hanno la proprietà P.

(b) *Passo.* Se i termini $t_1, t_2,..., t_n$ hanno la proprietà P (ipotesi induttiva), allora $f_i^n t_1 t_2 ... t_n$ ha la proprietà P.

Analogamente, per dimostrare che una proprietà P vale per tutte le fbf basta mostrare:

(a) *Base.* Tutte le fbf atomiche hanno la proprietà P.

(b_1) $Passo_1$. Se la fbf A ha la proprietà P (ip. induttiva), allora la fbf $(\neg A)$ ha la proprietà P.

(b_2) $Passo_2$. Se le fbf A e B hanno la proprietà P (ip. indutt.), allora le fbf $(A \wedge B)$, $(A \vee B)$ e $(A \rightarrow B)$ hanno la proprietà P.

(b_3) $Passo_3$. Se la fbf A ha la proprietà P (ip. indutt.), allora le fbf $(\forall xA)$ e $(\exists xA)$ hanno la proprietà P.

L'aver definito induttivamente il linguaggio della logica dei predicati consente un "controllo" sistematico delle proprietà sintattiche del linguaggio stesso.

L'alfabeto, come si è visto, è costituito da sei categorie di simboli. I vari simboli devono essere univocamente distinguibili e, per le costanti funzionali e predicative, deve essere calcolabile in modo effettivo il numero di argomenti di ciascuna di esse. Sono possibili diversi "alfabeti" e quello che abbiamo introdotto soddisfa queste condizioni. Le sequenze finite di simboli sono ottenute mediante la concatenazione, che consiste semplicemente nello scriverli uno dopo l'altro. E' fondamentale la proprietà di univocità di scomposizione, ossia che sia decidibile se una sequenza finita di simboli sia un termine o una fbf e, in caso affermativo, che sia univocamente determinabile in modo effettivo la sequenza di passi induttivi che consente di ottenerla dagli elementi iniziali (ossia che non vi sia ambiguità di lettura).

Si può dimostrare, ad esempio, il seguente teorema:

Data una sequenza finita di simboli che sono tutti costanti o variabili individuali o costanti funzionali, è possibile stabilire univocamente in modo effettivo se essa è una concatenazione di termini e di quali.

La dimostrazione è per induzione sul decorso dei valori della lunghezza della stringa.

Se la sequenza ha lunghezza 1, allora si riduce ad un solo simbolo e possiamo decidere se si tratta di una costante o di una variabile individuale. In caso affermativo la sequenza è la concatenazione di un solo termine e, in caso negativo (ossia se l'unico simbolo è una costante funzionale), non si tratta di una concatenazione di termini.

Ammettiamo che il teorema valga per tutte le sequenze di lunghezza fino a n e dimostriamolo per le sequenze di lunghezza n + 1. Consideriamo una sequenza S di lunghezza n + 1 ed esaminiamo il primo simbolo. Si hanno due casi.

(a) Se esso è una costante o una variabile individuale, esaminiamo la parte restante S' di lunghezza n. Per ipotesi induttiva possiamo decidere se S' è una concatenazione di r termini e di quali. In caso affermativo S è la concatenazione di r + 1 termini di cui il primo è costituito dalla costante o dalla variabile e gli altri sono gli r termini che compongono S' (e la scomposizione è unica poiché la costante o la variabile non può essere il primo simbolo di un altro termine). In caso negativo S non è una concatenazione di termini.

(b) Se il primo simbolo di S è una costante funzionale a k argomenti f_i^n, con $k < n + 1$, allora, a partire dal primo simbolo della parte restante S', esaminiamo le sequenze di k, k+1, k + 2 simboli per vedere se si trova una sequenza di k termini $t_1 t_2 ... t_k$. L'ipotesi induttiva garantisce che questa ricerca è possibile e dà una risposta univoca. In caso negativo S non è

una concatenazione di termini. In caso affermativo, se la sequenza dei k termini esaurisce S', allora S è costituita dall'unico termine $f_i^k t_1 t_2 ... t_k$; se non esaurisce S', allora S può essere una sequenza di termini di cui $f_i^k t_1 t_2 ... t_k$ è il primo. Si esamina allora la parte restante di S e, applicando l'ipotesi induttiva, si può determinare univocamente se questa è una concatenazione di termini (e di quali) che, unitamente al termine iniziale, completa l'eventuale scomposizione di S.

Questo teorema giustifica anche quanto abbiamo affermato nella lezione precedente a proposito dell'albero di formazione di un termine: ad ogni termine è associato in modo univoco un albero che visualizza la costruzione del termine a partire dalle costanti e variabili individuali attraverso i sottotermini[21]. La costruzione dell'albero avviene induttivamente: alla radice si pone il termine dato e, per ogni nodo, se esso contiene una costante o una variabile esso non ha successori immediati (è una foglia), mentre, se esso contiene un termine del tipo $f_i^k t_1 t_2 ... t_k$, allora ha come successori immediati k nodi contenenti rispettivamente i termini $t_1, t_2, ..., t_k$.

Mediante il concetto di altezza dell'albero di formazione di un termine (detta anche *altezza del termine*) si vede che il principio di induzione per i termini prima formulato è riconducibile al principio di induzione (sul decorso dei valori) enunciato per i numeri naturali. Basta osservare che (a) equivale a dimostrare la proprietà P per termini di altezza 0 e (b) equivale a dimostrare la proprietà P per termini di altezza n + 1 supposto che la proprietà valga per i termini di altezza ≤ n.

Così, per definire operazioni, funzioni o proprietà sui termini, si può procedere induttivamente, definendole prima per le costanti e le variabili e poi per i termini del tipo $f_i^k t_1 t_2 ... t_k$ supposto di averle già definite per i termini $t_1, t_2, ..., t_k$.

Quanto visto a proposito dei termini, si può ripetere quasi inalterato per le fbf. Il teorema precedente consente facilmente di stabilire che le fbf atomiche (quelle del tipo $R_i^n t_1 t_2 ... t_n$) sono univocamente ed effettivamente riconoscibili. Per le fbf vale un teorema di decomponibilità unica, che risulta più complesso del precedente per la presenza di più casi e delle parentesi[22]. Senza entrare in tutti i dettagli ci limitiamo a qualche ulteriore considerazione a integrazione di quanto esposto nella lezione precedente.

[21] Negli esempi della Lezione II abbiamo usato le parentesi poiché si è adottata una notazione più snella nella quale non viene evidenziato il numero di argomenti delle costanti funzionali. I sottotermini di un termine sono tutti i termini che figurano nell'albero di formazione.

[22] La necessità dell'impiego delle parentesi nelle fbf è dovuto alla scelta della notazione infissa per i connettivi a due argomenti (mentre per le costanti funzionali e predicative abbiamo adottato una notazione prefissa). L'abbondanza di parentesi nelle fbf è dovuta proprio alla necessità di garantire l'univocità di lettura.

Anzitutto, si può dimostrare che una sequenza finita di un numero pari di parentesi aperte e chiuse ammette un solo accoppiamento proprio[23] che, se esiste, si può determinare effettivamente. Le parentesi delle fbf ammettono, come si verifica facilmente per induzione, un accoppiamento proprio, che quindi è unico. Pertanto, data una sequenza finita di simboli, si può determinare effettivamente se le parentesi aperte e chiuse che figurano in essa ammettono un accoppiamento proprio (e, quindi, associare a ciascuna parentesi aperta la corrispondente parentesi chiusa)[24].

Ciò premesso, sia data una sequenza finita S di simboli. Per dimostrare che è possibile determinare univocamente ed effettivamente se essa è una fbf, quale è il suo segno logico principale, le sue sottoformule immediate, ecc., procediamo per induzione sulla lunghezza di S. Se la lunghezza è 1, S non è una fbf. Sia S una sequenza di lunghezza n + 1 e supponiamo che il teorema valga per le sequenze di lunghezza n. Esaminiamo il primo simbolo di S. Si hanno due casi a seconda se il primo simbolo è una costante predicativa o una parentesi aperta (in tutti gli altri casi S non è una fbf).

(a) Se si tratta di un simbolo di predicato a k argomenti si può stabilire se S è una fbf atomica (determinando, come si è illustrato in precedenza, se la parte restante è la concatenazione di k termini).

(b) Se si tratta di una parentesi aperta, allora l'ultimo simbolo deve essere una parentesi chiusa e si passa all'esame del secondo simbolo. Si hanno vari sottocasi.

(b_1) Se il secondo simbolo è $\neg$, o $\forall$ o $\exists$ seguiti da una variabile individuale, e quindi S è del tipo $(\neg A)$, $(\forall x A)$, $(\exists x A)$, si è ricondotti all'analisi di A, che ha lunghezza minore di n.

(b_2) Se il secondo simbolo è una costante predicativa a k argomenti occorre stabilire se è seguita dalla concatenazione di k termini (cosa che sappiamo essere univocamente eseguibile), subito seguita da un connettivo c ($o \wedge o \vee o \rightarrow$), ossia se è del tipo $(R_i^k t_1 t_2 ... t_k \, c \, A)$. Si passa allora all'esame di A che ha lunghezza minore di n.

(b_3) Se il secondo simbolo è un'altra parentesi aperta, occorre determinare, come si è accennato in precedenza, la parentesi chiusa corrispondente nell'accoppiamento proprio (se esiste) delle parentesi di S. Occorre poi che il simbolo che segue la parentesi chiusa sia un connettivo a due argomenti c; in altre parole si deve ottenere per S la seguente

[23] Più precisamente un *accoppiamento proprio* è una corrispondenza biunivoca f fra le parentesi sinistre e le parentesi destre tale che (i) se p è una parentesi sinistra, allora f(p) segue p nella sequenza, e (ii) se p' e p" sono due parentesi sinistre tali che p' precede p" nella sequenza, allora f(p') segue f(p") nella sequenza. La dimostrazione si ottiene facilmente per induzione sulla lunghezza 2n della sequenza. Se n = 1, allora vi sono solo due parentesi e il riconoscimento è immediato. Se si ha una sequenza di 2k+2 parentesi, basta osservare che la prima parentesi chiusa della sequenza in un accoppiamento proprio è associata alla parentesi aperta che la precede immediatamente. Eliminando questa coppia ci si riconduce ad una sequenza di lunghezza 2k alla quale si applica l'ipotesi induttiva.

[24] Se ciò non accade la sequenza finita di simboli non è una fbf.

configurazione: ((A) *c* B) e siamo di nuovo ricondotti a sequenze A e B di lunghezza inferiore. Se non si verifica nessuno di questi casi S non è una fbf.

Come in precedenza, resta giustificata la costruzione dell'albero di formazione della fbf che si può eseguire in concomitanza della verifica illustrata nella dimostrazione precedente: alla radice si pone la fbf, immediatamente sopra la o le sottoformule immediate (se esistono) e si procede fino alle foglie (che contengono le fbf atomiche). Come per i termini, si riconosce che il principio di induzione per le fbf precedentemente enunciato si può ricondurre ad una induzione sul decorso dei valori dell'altezza dell'albero di formazione delle fbf.

Così, per definire funzioni e concetti relativi ai termini e alle fbf si può procedere per via induttiva. Vediamo, a titolo di esempio, la definizione di *occorrenza libera di una variabile* in una fbf:

(a) una variabile y ha una occorrenza libera nella fbf atomica $R_i^n t_1 t_2 ... t_n$ se e solo se occorre in almeno uno dei termini $t_1, t_2, ..., t_n$;

(b_1) y ha una occorrenza libera in $(\neg A)$ se e solo se ha una occorrenza libera in A;

(b_2) y ha una occorrenza libera in $(A \wedge B)$, $(A \vee B)$, $(A \rightarrow B)$ se e solo se ha una occorrenza libera in almeno una fra A e B.

(b_3) y ha una occorrenza libera in $(\forall xA)$ e $(\exists xA)$ se e solo se $y \neq x$ e y ha una occorrenza libera in A.

Una occorrenza di una variabile è *vincolata* se non è libera. Il riconoscimento delle occorrenze libere e vincolate delle variabili è reso agevole se si esamina l'albero di formazione, in quanto il campo di azione di un quantificatore $\forall x$ o $\exists x$ è la sottoformula che sta immediatamente sopra al nodo nel quale il quantificatore appare per la prima volta, ed esso vincola tutte (e sole) le occorrenze libere di x in tale sottoformula.

Per vedere un ulteriore esempio di definizione induttiva, illustriamo l'operazione di *sostituzione* di un termine r a tutte le occorrenze di una variabile x in un termine t, il cui risultato indichiamo con t[x/r]. Si procede per induzione sull'altezza di t:

Base. t ha altezza 0:

(a) se t è una costante individuale, allora t[x/r] = t

(b) se t è una variabile individuale diversa da x, allora t[x/r] = t

(c) se t è x, allora t[x/r] = r

Passo. t ha altezza n + 1, e quindi $t = f_i^n t_1 t_2 ... t_k$, allora:

$$t[x/r] = f_i^n t_1[x/r] t_2[x/r] ... t_k[x/r].$$

In modo analogo si definisce l'operazione di *sostituzione* di un termine r a tutte le occorrenze libere di una variabile x in una fbf A, il cui risultato indichiamo con A[x/r], per induzione sull'altezza di A.

Base. Se A ha altezza 0, ossia A è $R_i^k t_1 t_2 ... t_k$, allora:

$$A[x/r] = R_i^k t_1[x/r] t_2[x/r] ... t_k[x/r].$$

Passo. A ha altezza n + 1:

(a) se $A = (\neg B)$, allora $A[x/r] = (\neg B[x/r])$

(b) se $A = (B \wedge C)$, allora $A[x/r] = (B[x/r] \wedge C[x/r])$

(c) se $A = (B \vee C)$, allora $A[x/r] = (B[x/r] \vee C[x/r])$

(d) se $A = (B \rightarrow C)$, allora $A[x/r] = (B[x/r] \rightarrow C[x/r])$

(e) se $A = (\forall xB)$ o $(\exists xB)$, allora A[x/r] = A

(f) se $A = (\forall yB)$ o $(\exists yB)$ con y diversa da x, allora:
$A[x/r] = (\forall yB[x/r])$ o $(\exists yB[x/r])$[25].

[25] La sostituzione di un termine r in una fbf A al posto delle occorrenze libere di una variabile x si esegue solitamente, in logica, quando il termine r è libero per x in A: se A è una fbf e r è un termine, allora si dice che r *è libero per* x *in* A se e solo se nessuna occorrenza libera di x in A si trova nel campo di azione di un quantificatore $(\forall y)$ o $(\exists y)$ dove y è una variabile che occorre in r.

Esercizi

DARIO PALLADINO

Lezione I

Equipotenza e numeri cardinali

ESERCIZIO 1 Dimostrare le seguenti equipotenze:

(a) $A \times B \cong B \times A$

(b) $(A \times \{a\}) \cong A \cong A^{\{a\}}$

(c) $(A \times B) \times C \cong A \times (B \times C)$

(d) $\{a\}^A \cong \{a\}$

(e) se $A \cong A'$ e $B \cong B'$, allora $A \times B \cong A' \times B'$

(f) se $A \cong A'$, $B \cong B'$, $A \cap A' = B \cap B' = \varnothing$, allora $A \cup B \cong A' \cup B'$

(g) $A^{B \times C} \cong (A^B)^C$

(h) $(A \times B)^C \cong A^C \times B^C$

(i) se $A \cap B = \varnothing$, allora $C^{A \cup B} \cong C^A \times C^B$

Ricordando che:

se $A \cap B = \varnothing$, allora $\text{Card}(A \cup B) = \text{Card } A + \text{Card } B$

$\text{Card}(A \times B) = \text{Card } A \cdot \text{Card } B$

$\text{Card}(A^B) = (\text{Card } A)^{\text{Card } B}$

dedurre da ciascuna delle precedenti equipotenze una proprietà dell'aritmetica dei numeri cardinali.

ESERCIZIO 2 Dimostrare che $\text{Card } P(A) = 2^{\text{Card } A}$

ESERCIZIO 3 Se $A \cong B' \subseteq B$ (ossia se $\text{Card } A \leq \text{Card } B$), allora:

$\text{Card } A + \text{Card } C \leq \text{Card } B + \text{Card } C$

$\text{Card } A \cdot \text{Card } C \leq \text{Card } B \cdot \text{Card } C$

$(\text{Card } A)^{\text{Card } C} \leq (\text{Card } B)^{\text{Card } C}$

$(\text{Card } C)^{\text{Card } A} \leq (\text{Card } C)^{\text{Card } B}$

ESERCIZIO 4 Posto κ il numero cardinale del continuo e ricordando che (vedi l'esercizio 2) tale cardinale è uguale a $2^{\text{Card } \mathbf{N}}$ $(= 2^{\aleph_0})$, sfruttando anche quanto dimostrato nell'esercizio 3, dedurre che:

$$n^{\aleph_0} = \aleph_0{}^{\aleph_0} = \kappa^{\aleph_0} = \kappa$$

ESERCIZIO 5 Dedurre dal risultato precedente che il piano cartesiano, lo spazio cartesiano, lo spazio a n dimensioni e lo spazio a una quantità numerabile di dimensioni sono insiemi aventi la potenza del continuo.

ESERCIZIO 6 Dimostrare (sfruttando il teorema di Cantor) che l'insieme **M** (§8) unione della successione **N**, P(**N**), P(P(**N**)), P(P(P(**N**))),... ha numero cardinale maggiore di tutti i numeri cardinali degli insiemi della successione.

ESERCIZIO 7 Dopo aver osservato che l'insieme degli intervalli ad estremi razionali della retta cartesiana è numerabile, dedurne che ogni famiglia disgiunta di segmenti della retta è numerabile.

ESERCIZIO 8 Ricordando che una funzione di variabile reale ha un estremo proprio nel punto a se e solo se esiste un intorno I(a) di a tale che o $f(x) < f(a)$ per ogni x in $I(a) - \{a\}$ o $f(x) > f(a)$ per ogni x in $I(a) - \{a\}$, dedurne che una funzione di variabile reale può avere al più una quantità numerabile di estremi propri.

ESERCIZIO 9 Dimostrare che l'insieme dei numeri reali algebrici è numerabile.

ESERCIZIO 10 Ricordando quanto ottenuto nell'esercizio 4 ($\kappa^{\aleph_0} = \kappa$) ossia che l'insieme delle successioni di numeri reali ha la potenza del continuo, dedurne che l'insieme delle funzioni continue $f:\mathbf{R}\to\mathbf{R}$ ha la potenza del continuo.

ESERCIZIO 11 Dimostrare che l'insieme delle funzioni $f:\mathbf{R}\to\mathbf{R}$ ha potenza superiore a quella del continuo (e precisamente 2^{κ}).

ESERCIZIO 12 Molto spesso, in matematica, si adoperano scritture del tipo: ∞^2, ∞^3, ∞^n (ad esempio in geometria analitica e nella risoluzione dei sistemi lineari). Queste scritture hanno dei nessi con quanto illustrato in questa lezione a proposito dei numeri cardinali?

ESERCIZIO 13 *Numeri naturali.* All'inizio del §6 si è accennato alla definizione insiemistica dei numeri naturali e al fatto che, nelle trattazioni più rigorose, occorre scegliere un rappresentante per ogni cardinalità. Per quanto riguarda i numeri naturali si pone:

$$0 = \varnothing \qquad n + 1 = n \cup \{n\}$$

Si dice poi che un insieme A è *induttivo* se e solo se:

$$0 \in A \text{ e per ogni } x(\text{se } x \in A, \text{ allora } x + 1 \in A)$$

e si assume come assioma (*assioma dell'infinito*) che esiste almeno un insieme induttivo. Infine, si definisce l'insieme **N** dei numeri naturali come l'intersezione di tutti gli insiemi induttivi. In questa impostazione si ha, ad esempio:

$$1 = \{0\}, 2 = \{0, 1\}, 3 = \{0, 1, 2\}, \text{ ecc.}$$

Dimostrare che:

(a) se $n \in N$, allora, se $x \in n$, allora $x \subset n$
(*suggerimento*: procedere per induzione[1] su n)
(b) se $m + 1 = n + 1$, allora $m = n$
(c) un numero naturale non può essere equipotente a un suo sottoinsieme proprio
(d) un sottoinsieme proprio di un numero naturale n è equipotente a un numero naturale k minore di n (si osservi che si può adottare la seguente definizione della relazione d'ordine tra naturali: $m < n$ se e solo se $m \in n$)

Dedurne che, detto *finito* un insieme equipotente ad un numero naturale, si ha che:
- un insieme finito non può essere equipotente a una sua parte propria;
- un sottoinsieme di un insieme finito è finito;
- un insieme finito è equipotente ad un unico numero naturale;
- un insieme finito linearmente ordinato è dotato di massimo e minimo.

ESERCIZIO 14 Se A e B sono insiemi disgiunti e bene ordinati dalle relazioni $<_A$ e $<_B$ rispettivamente, verificare che $A \cup B$ è bene ordinato dalla relazione $<$ così definita:

$$x < y \text{ se e solo se } x <_A y \text{ oppure } x <_B y \text{ oppure } (x \in A \text{ e } y \in B)$$

Sfruttare questo risultato per definire l'addizione fra numeri ordinali.

ESERCIZIO 15 Se A e B sono insiemi bene ordinati dalle relazioni $<_A$ e $<_B$ rispettivamente, verificare che $A \times B$ è bene ordinato dalla relazione $<$ così definita:

$$(a,b) < (a',b') \text{ se e solo se } b <_B b' \text{ oppure } b = b' \text{ e } a <_A a'.$$

Utilizzare questo risultato per definire la moltiplicazione fra numeri ordinali.

---

[1] Sul principio di induzione vedi la terza lezione.

Lezione II

ESERCIZIO 1 Angelo, Bruno e Carlo sono tre studenti che hanno sostenuto un esame. Se Pa, Pb e Pc formalizzano rispettivamente le proposizioni "Angelo ha superato l'esame", "Bruno ha superato l'esame", "Carlo ha superato l'esame", determinare le fbf composte che "traducono" le seguenti proposizioni:

(a) Solo Carlo ha superato l'esame
(b) Solo Angelo non ha superato l'esame
(c) Almeno uno di Angelo, Bruno e Carlo ha superato l'esame
(d) Almeno due fra Angelo, Bruno e Carlo hanno superato l'esame
(e) Al più due fra Angelo, Bruno e Carlo hanno superato l'esame
(f) Esattamente due fra Angelo, Bruno e Carlo hanno superato l'esame

ESERCIZIO 2 Angelo, Bruno e Carlo sono gli unici tre membri di una commissione che vota una proposta. Se Pa, Pb e Pc formalizzano rispettivamente le proposizioni "Angelo ha votato a favore della proposta", "Bruno ha votato a favore della proposta", "Carlo ha votato a favore della proposta", determinare le fbf composte che "traducono" le seguenti proposizioni:

(a) La votazione è stata unanime
(b) La proposta è passata a maggioranza
(c) La proposta ha ricevuto un numero dispari di voti favorevoli
(d) La proposta è stata respinta, ma non all'unanimità
(e) La proposta è stata respinta con il voto contrario di Bruno

ESERCIZIO 3 Usando costanti individuali e funzionali formalizzare le seguenti espressioni che denotano individui:

(a) la somma di 2 e del quadrato di 4
(b) il prodotto della somma di 3 e 5 e della radice quadrata di 9
(c) il doppio della radice quadrata della somma del quadrato di 3 e del quadrato di 4
(d) il prodotto di 2 alla terza e 3 alla seconda
(e) la somma del prodotto del quadrato di 2 e del cubo di 3 e della radice quadrata della somma di 6 e 3
(f) il prodotto di 2 elevato al doppio di 3 e del doppio di 5 elevato alla radice quadrata della somma di 1 e 15

ESERCIZIO 4 Scrivere gli alberi di formazione dei termini ottenuti nel precedente esercizio 3.

ESERCIZIO 5 Formalizzare le seguenti proposizioni:

(a) Il prodotto di 3 e 4 è minore della radice quadrata di 169
(b) Il doppio di 5 è compreso tra il quadrato di 2 e la somma di 3 e 9
(c) Il quadrato di 5 è maggiore del quadrato di 4 e minore del quadrato di 6
(d) Il prodotto di due numeri dispari è sempre un numero dispari
(e) Il prodotto di due numeri primi non è mai un numero primo
(f) Se il prodotto di due numeri è pari allora almeno uno dei due è pari

ESERCIZIO 6 Formalizzare le proposizioni che intervengono nei seguenti ragionamenti:

(a) Se il triangolo T è rettangolo, allora almeno uno dei quadrilateri Q' e Q" è un rombo. Q" non è un rombo. Quindi, se Q' non è un rombo, allora T non è rettangolo.
(b) Se il triangolo T è rettangolo, allora almeno uno dei quadrilateri Q' e Q" è un rombo. Q' è un rombo. Quindi, se T è rettangolo, allora Q" non è un rombo.
(c) Se il triangolo T è rettangolo, allora i due quadrilateri Q' e Q" sono rombi. Q' non è un rombo. Quindi T non è rettangolo.
(d) Se il triangolo T è rettangolo, allora, se il quadrilatero Q' è un rombo, il quadrilatero Q" è un rombo. Q' non è un rombo. Quindi, o T è rettangolo o Q" è un rombo.
(e) Se il quadrilatero Q' è un rombo, allora almeno uno dei due triangoli T' e T" è isoscele. Se T' è isoscele, allora Q' non è un rombo. Se il quadrilatero Q" è un rombo, allora T" non è isoscele. Quindi, se Q' è un rombo, allora Q" non lo è.
(f) Se nessuno dei due triangoli T' e T" è isoscele, allora il quadrilatero Q' è un rombo. Se T' è isoscele, allora Q' è un rombo. Se il quadrilatero Q" non è un rombo, allora T" non è isoscele. Quindi, se Q' non è un rombo, allora Q" lo è.

Ragionando come nel §6 determinare quali dei precedenti ragionamenti sono corretti.

ESERCIZIO 7 Formalizzare le proposizioni che intervengono nei seguenti ragionamenti:

(a) Se un insieme è contenuto strettamente in un altro e quest'ultimo è contenuto strettamente in un terzo, allora il primo insieme è contenuto strettamente nel terzo. Nessun insieme è contenuto strettamente in se stesso. Quindi, se un insieme è contenuto strettamente in un altro, allora quest'ultimo non è contenuto strettamente nel primo.
(b) Se una retta è parallela ad una seconda retta e questa è parallela a una terza, allora la prima è parallela alla terza. Se una retta è parallela a un'altra, allora questa è parallela alla prima. Quindi, ogni retta è parallela a se stessa.

(c) Se una retta è parallela ad una seconda retta e questa è parallela a una terza, allora la prima è parallela alla terza. Se una retta è parallela a un'altra, allora questa è parallela alla prima. Ogni retta ha almeno una parallela. Quindi, ogni retta è parallela a se stessa.

(d) Se un insieme è parte propria di un altro, allora quest'ultimo non è parte propria del primo. Quindi, nessun insieme è parte propria di se stesso.

(e) Un insieme è parte di un altro se e solo se ogni insieme che è parte del primo è parte del secondo. Un insieme interseca un altro se e solo se vi è un insieme che è parte di entrambi. Quindi:

1) Ogni insieme è parte di se stesso.
2) Ogni insieme interseca se stesso.
3) Se un insieme è parte di un altro e quest'ultimo è parte di un terzo, allora il primo insieme è parte del terzo.
4) Un insieme interseca ogni sua parte.
5) Se un insieme interseca ogni parte di un altro insieme, allora quest'ultimo è parte del primo.
6) Se un insieme è parte di un altro e interseca un terzo insieme, allora il secondo insieme interseca il terzo.

(f) Un barbiere di Savona rade tutte e sole le persone che non si radono da sé. Quindi, non esiste alcun barbiere a Savona.

(g) Per ogni insieme ve ne è uno di cardinalità maggiore. Se un insieme è incluso in un altro, la sua cardinalità non è maggiore di quella dell'altro. Ogni insieme è incluso nell'insieme di tutti gli insiemi. Quindi, l'insieme di tutti gli insiemi non è un insieme.

(Il problema di stabilire la correttezza o meno dei ragionamenti precedenti può essere affrontato con le tecniche che saranno esposte nelle lezioni successive).

ESERCIZIO 8 Formalizzare le seguenti proposizioni che possono essere assunte tra gli assiomi della geometria euclidea piana. Si introducano i predicati unari P(x) e R(x) ("x è punto" e "x è retta"), il predicato binario I(x,y) ("il punto x sta sulla retta y"), il predicato ternario S(x,y,z) ("i punti x, y e z sono allineati e y sta fra x e z") e il predicato quaternario U(x,y,z,t) ("il segmento di estremi x,y è congruente al segmento di estremi z,t"):

(a) Per due punti passa una retta.
(b) Per due punti distinti passa una retta.
(c) Per due punti distinti passa una ed una sola retta.

(d) Su una retta vi sono almeno due punti.
(e) Vi sono almeno tre punti che non giacciono su una retta.
(f) Se un punto B sta tra A e C, allora sta anche tra C e A.
(g) Per ogni coppia di punti A e C vi è almeno un punto B della retta AC tale che C sta tra A e B.
(h) Se due segmenti sono congruenti a un terzo, allora sono congruenti tra loro.
(i) Se B sta tra A e C e B' sta tra A' e C', allora se AB è congruente a A'B' e BC è congruente a B'C', allora AC è congruente a A'C'.
(l) (*Assioma delle parallele*) Dati una retta e un punto fuori di essa esiste al più una retta passante per il punto e non incidente la retta data.
(l') (*Assioma della geometria iperbolica*) Dati una retta e un punto fuori di essa esistono almeno due rette passanti per il punto e non incidenti la retta data.
(l") (*Assioma della geometria ellittica*) Non esistono rette parallele.

Notare che l'Assioma di Archimede ("Dati due segmenti non degeneri esiste un multiplo del minore che supera il maggiore") non può essere formalizzato con il linguaggio della logica del primo ordine. Notare che l'usuale formulazione dell'assioma di continuità di Cantor ("Dati due insiemi separati di punti di una retta esiste almeno un punto di separazione fra essi") non è esprimibile con il linguaggio della logica del primo ordine.

ESERCIZIO 9 Scrivere l'albero di formazione delle seguenti fbf:

(a) $(((\forall x(\neg Rxa))\wedge(\exists\, y Sby))\rightarrow Rxy)$
(b) $(\forall x(((\neg Rxa)\wedge(\exists\, y Sby))\rightarrow Rxy))$
(c) $(\forall x((\neg Rxa)\wedge(\exists\, y(Sby\rightarrow Rxy))))$
(d) $((\forall x(\neg Rxa))\wedge(\exists\, y(Sby\rightarrow Rxy)))$
(e) $(\exists x((Px\wedge(\exists y Rxy))\rightarrow(\neg(\forall z(Sxz\rightarrow Ryx)))))$
(f) $(((\exists x Px)\wedge(\exists y(Rxy\rightarrow(\neg(\forall z Sxz)))))\rightarrow Ryx)$
(g) $((\exists x Px)\wedge(\exists y(Rxy\rightarrow(\neg(\forall z(Sxz\rightarrow Ryx))))))$
(h) $(\exists x(Px\wedge(\exists y(Rxy\rightarrow(\neg(\forall z(Sxz\rightarrow Ryx))))))$
(i) $((((\exists x Px)\wedge(\exists y Rxy))\rightarrow(\neg(\forall z Sxz)))\rightarrow Ryx)$
(l) $(((\exists x Px)\wedge(\exists y Rxy))\rightarrow(\neg((\forall z Sxz)\rightarrow Ryx)))$

ESERCIZIO 10 Scrivere le fbf del precedente esercizio 9 con il minor numero possibile di parentesi.

ESERCIZIO 11 Individuare il segno logico principale delle seguenti fbf scritte già in notazione abbreviata:

(a) $\exists x(Px \wedge \forall y(Qy \vee Rxy) \rightarrow Syx)$
(b) $\exists xPx \wedge \forall y(Qy \vee Rxy) \rightarrow Syx$
(c) $\exists x(Px \wedge \forall yQy) \vee Rxy \rightarrow Syx$
(d) $(\exists xPx \wedge \forall yQy) \vee (Rxy \rightarrow Syx)$
(e) $\exists xPx \wedge \forall y(Qy \vee Rxy \rightarrow Syx)$
(f) $\exists xPx \wedge \forall y(Qy \vee (Rxy \rightarrow Syx))$

ESERCIZIO 12 Determinare le occorrenze libere e vincolate delle variabili individuali nelle fbf di cui agli esercizi 9 e 11.

ESERCIZIO 13 Nelle fbf di cui agli esercizi 9 e 11 operare la sostituzione:

(a) [x/a] (b) [y/f(b)] (c) [x/z]

ESERCIZIO 14 Stabilire se il termine

(a) b (b) y (c) f(x)

è libero per x nelle fbf di cui all'esercizio 11.

ESERCIZIO 15 Stabilire se il termine

(a) f(c) (b) x (c) f(y)

è libero per y nelle fbf di cui all'esercizio 11.

Lezione III

ESERCIZIO 1 Dimostrare che, per ogni $n \geq 1$, e $q \neq 1$:

$$1 + q + q^2 + q^3 + \ldots + q^n = \frac{1-q^{n+1}}{1-q}$$

ESERCIZIO 2 Dimostrare che, per ogni $n \geq 1$, $1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$

ESERCIZIO 3 Dimostrare che, per ogni n, $n < 2^n$.

ESERCIZIO 4 Dimostrare che, per ogni n, $11^{n+2} + 12^{2n+1}$ è divisibile per 133.

ESERCIZIO 5 Dimostrare che, per ogni n, $3^{2n} - 2^n$ è multiplo di 7.

ESERCIZIO 6 Dimostrare che, per ogni n, $3n^5 + 5n^3 - 8n$ è divisibile per 120.

ESERCIZIO 7 Dimostrare che la somma dei cubi di tre numeri naturali consecutivi è sempre divisibile per 9.

ESERCIZIO 8 Dimostrare che, per ogni $n \geq 1$, $5^n + 2 \times 3^{n-1} + 1$ è divisibile per 8.

ESERCIZIO 9 $$1^3 + 2^3 + 3^3 + \ldots + n^3 = \frac{n^2(n+1)^2}{4} = (1 + 2 + \ldots + n)^2$$

ESERCIZIO 10 $$\frac{1}{1\cdot2} + \frac{1}{2\cdot3} + \frac{1}{3\cdot4} + \ldots + \frac{1}{n(n+1)} = \frac{n}{n+1}$$

ESERCIZIO 11 $$\frac{1}{1\cdot3} + \frac{1}{3\cdot5} + \frac{1}{5\cdot7} + \ldots + \frac{1}{(2n-1)(2n+1)} = \frac{n}{2n+1}$$

ESERCIZIO 12 $$\frac{1}{1\cdot4} + \frac{1}{4\cdot7} + \frac{1}{7\cdot10} + \ldots + \frac{1}{(3n-2)(3n+1)} = \frac{n}{3n+1}$$

ESERCIZIO 13 $$\frac{1}{1\cdot5} + \frac{1}{5\cdot9} + \frac{1}{9\cdot13} + \ldots + \frac{1}{(4n-3)(4n+1)} = \frac{n}{4n+1}$$

ESERCIZIO 14 $$1\cdot2 + 2\cdot3 + 3\cdot4 + \ldots + n(n+1) = \frac{n(n+1)(n+2)}{3}$$

ESERCIZIO 15 $$1\cdot2 + 2\cdot5 + 3\cdot8 + \ldots + n(3n-1) = n^2(n+1)$$

ESERCIZIO 16 $$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \ldots + \frac{n}{2^n} = 2 - \frac{n+2}{2^n}$$

ESERCIZIO 17 $$\left(1-\frac{1}{2^2}\right)\cdot\left(1-\frac{1}{3^2}\right)\cdot\ldots\cdot\left(1-\frac{1}{n^2}\right) = \frac{n+1}{2n}$$

ESERCIZIO 18 $$\text{sen } x + \text{sen } 3x + \text{sen } 5x + \ldots + \text{sen}(2k-1)x = \frac{1-\cos(2kx)}{2\text{sen}\,x}$$

ESERCIZIO 19 Dimostrare la formula di De Moivre:

$$(\cos x + i \text{ sen } x)^n = \cos nx + i \text{ sen } nx.$$

ESERCIZIO 20 Dimostrare che, per ogni $n > 0$, $|\text{sen } nx| \leq n|\text{sen } x|$.

ESERCIZIO 21 Verificare che, se si provasse a dimostrare per induzione che, per ogni n, $10^n + 1$ è divisibile per 9, vale il passo, ma non la base. (*Traccia*. 10^{n+1} si può scrivere $9\cdot10^n + 10^n + 1$, ossia la somma di un numero divisibile per 9 e di un altro numero, $10^n + 1$, divisibile per 9 per ipotesi induttiva).

ESERCIZIO 22 Verificare che, se si provasse a dimostrare che ogni n è ≤ 107, vale la base, e vale anche il passo per ogni k, con l' eccezione di k = 107.

ESERCIZIO 23 Trovare l'errore nella seguente dimostrazione per induzione della proprietà "per ogni $n \geq 2$, se tra n persone vi è uno ed un solo bugiardo, esso può essere individuato mediante una sola domanda". Vale la base (infatti, se le persone sono 2 il bugiardo può essere individuato con la domanda: "Se io chiedo all'altro chi è il bugiardo, chi indicherebbe?" - la persona non indicata è il bugiardo). Per quanto riguarda il passo, consideriamo k + 1 persone tra cui vi è un bugiardo. Scegliamo una persona: se non è il bugiardo lo si può individuare, per ipotesi induttiva, tra le altre k persone con una sola domanda, se è il bugiardo allora lo si è individuato. (*Traccia*. L'ipotesi induttiva è applicata in modo scorretto in quanto non è detto che fra le k persone vi sia uno ed un solo bugiardo).

ESERCIZIO 24 Trovare l'errore nella seguente dimostrazione per induzione della proprietà "comunque dati n numeri naturali, essi sono uguali". La proprietà vale per n = 1 e quindi vale la base. Supponiamo che la proprietà valga per gli insiemi di k numeri e sia $\{a_1, a_2, ..., a_k, a_{k+1}\}$ un insieme di k + 1 numeri naturali. Consideriamo $\{a_1, a_2, ..., a_k\}$ e $\{a_2, ..., a_k, a_{k+1}\}$. Essi sono insiemi che contengono k numeri, quindi, per ipotesi induttiva, sono formati da numeri uguali: $a_1 = a_2 = ... = a_k$ e $a_2 = ... = a_k = a_{k+1}$. Ma allora $a_1 = a_2 = ... = a_k = a_{k+1}$ e il passo è dimostrato. (*Traccia*. Il passo vale se $k \geq 2$ e la proprietà non vale per 2...).

ESERCIZIO 25 Trovare l'errore nella seguente dimostrazione per induzione della proprietà "per ogni n, se max(a,b)=n, allora a=b". La proprietà vale per n = 0; infatti, se max(a,b) = 0, allora, a = 0 e b = 0, per cui a = b. Supponiamo che la proprietà valga per k e dimostriamola per k+1. Supponiamo che max(a,b)=k+1. Poniamo c = a – 1 e d = b – 1. Allora max(c,d) = k e, per ipotesi induttiva, c = d; ne segue che c + 1 = d + 1, ossia che a = b, e quindi la proprietà vale per k + 1. (*Traccia*. Si può togliere 1 solo se...)

ESERCIZIO 26 Dimostrare che la somma degli angoli di un poligono convesso o concavo di n + 2 lati è uguale a n angoli piatti.

ESERCIZIO 27 L'area di un triangolo rettangolo i cui lati hanno lunghezza intera non può essere un quadrato perfetto.

Traccia di svolgimento. Procediamo applicando il principio della discesa infinita. E' noto che (tutte) le terne pitagoriche a, b e c primitive (in cui a, b e c non hanno fattore comune maggiore di 1) si ottengono con le seguenti formule:

$$a = m^2 - n^2 \qquad b = 2mn \qquad c = m^2 + n^2$$

dove m e n sono interi positivi primi fra loro con m > n e di cui uno pari e l'altro dispari (a parte la possibilità di scambiare a con b e si può supporre che a sia dispari). Non vi è perdita di generalità supponendo che i lati del triangolo siano una terna pitagorica primitiva (perché?).

L'area del triangolo è $mn(m+n)(m-n)$.

Si verifica immediatamente che ciascun fattore è primo con gli altri tre. Quindi, se il prodotto è un quadrato perfetto, ciascun fattore deve essere un quadrato perfetto. Allora si può porre $m^2 - n^2 = p^2$, dove, per le ipotesi fatte, p e m sono dispari e n è pari. Quindi p, m e n sono una terna pitagorica primitiva e si può scrivere:

$$p = k^2 - h^2 \qquad n = 2kh \qquad m = k^2 + h^2,$$

con k, h primi fra loro e uno pari e l'altro dispari.

Essendo anche n un quadrato, o k o h deve essere un quadrato perfetto dispari e l'altro il doppio di un quadrato perfetto. Poiché anche m è un quadrato, ponendo $m = u^2$, la terza equazione diviene $k^2 + h^2 = u^2$.

Ma allora k, h e u sono i lati di un triangolo la cui area, che risulta $\frac{1}{2}$ k h, è, per quanto prima osservato, un quadrato perfetto e la cui ipotenusa u è minore di quella del triangolo iniziale (infatti: $u < m < m^2 < m^2 + n^2 = c$).

Il risultato segue applicando il principio della discesa infinita.

ESERCIZIO 28 Servirsi del risultato ottenuto nell'esercizio 16 per dimostrare che le equazioni $x^4 - y^4 = z^4$ e $x^4 + y^4 = z^4$ non hanno soluzioni intere positive. (*Suggerimento*: verificare che, se x, y e z fossero soluzione della prima equazione, i tre numeri $x^4 - y^4$, $2x^2y^2$, $x^4 + y^4$ sarebbero i lati di un triangolo rettangolo la cui area è un quadrato perfetto. L'irrisolubilità della seconda equazione si riconduce immediatamente a quella della prima).

ESERCIZIO 29 Definiti per ricorrenza i coefficienti binomiali mediante le formule:

$$\begin{cases} C_{n,0} = 1 \\ C_{n,n} = 1 \\ C_{n+1,r} = C_{n,r-1} + C_{n,r} \quad \text{per } 1 \leq r \leq n \end{cases}$$

dimostrare la formula esplicita:

$$C_{n,r} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{r \cdot (r-1) \cdot \ldots \cdot 2 \cdot 1}$$

ESERCIZIO 30 Dimostrare per induzione il teorema del binomio:

$$(a+b)^n = \sum_{k=0}^{n} C_{n,k}\, a^{n-k}\, b^k$$

ESERCIZIO 31 Dimostrare per induzione:

(a) la proprietà commutativa dell'addizione (si dimostri prima che, per ogni a, a+0=0+a);
(b) la proprietà associativa della moltiplicazione;
(c) la proprietà commutativa della moltiplicazione;
(d) la proprietà di cancellazione: per ogni a, b e c, se $a + c = b + c$, allora $a = b$;
(e) la proprietà distributiva della moltiplicazione rispetto alla addizione: $a \times (b + c) = a \times b + a \times c$;
(f) per ogni a, b e c, $(a^b)^c = a^{b \times c}$;
(g) per ogni a, b e c, $(a \times b)^c = a^c \times b^c$

ESERCIZIO 32 Dimostrare per induzione che, comunque dati due termini diversi t' e t", t' non può essere un segmento iniziale di t" (ossia, non esiste una sequenza finita di simboli S tale che t'S = t").

ESERCIZIO 33 Dimostrare per induzione che, comunque date due fbf diverse A' e A", A' non può essere un segmento iniziale di A".

ESERCIZIO 34 Dimostrare che il simbolo logico principale di una fbf è il primo simbolo logico della fbf per cui, contando da destra a sinistra, il numero delle parentesi aperte supera di uno il numero delle parentesi chiuse.

ESERCIZIO 35 Definire induttivamente la funzione V che associa a ciascun termine le variabili che occorrono nel termine e la funzione L che ad ogni fbf A associa l'insieme delle variabili che hanno una occorrenza libera in A.

ESERCIZIO 36 Dimostrare che, se x non ha occorrenze libere in A, allora A[x/t] = A.

ESERCIZIO 37 Verificare che non sempre A[x/y][y/x] = A.

Il Metodo ipotetico deduttivo
Concetti primitivi, Assiomi Definizioni, Teoremi
Coerenza e Indipendenza
Modelli di una Teoria

CLAUDIO BERNARDI
Dipartimento di Matematica
Università Roma "La Sapienza"
Piazzale Aldo Moro 2
00185 Roma

0. Premessa

I programmi Brocca suggeriscono, alla fine del Triennio, di rivedere e sistemare quanto studiato negli anni precedenti (soprattutto in geometria), per discutere e mettere in luce la struttura logica delle teorie matematiche. Si tratta, in sostanza, di rendere gli studenti più consapevoli dei metodi e dei procedimenti a loro già familiari. Naturalmente, esula dallo scopo di queste note scritte la discussione di quanta parte delle osservazioni che seguono vada presentata via via agli studenti, nell'arco degli anni precedenti, quanta vada invece rimandata all'ultimo anno e quanta, infine, non sia proponibile in una classe.

1. Le definizioni

Nelle trattazioni matematiche si incontrano due tipi di frasi "ufficiali": definizioni e teoremi.

In una definizione si introduce una parola (o una locuzione) nuova per indicare gli oggetti che godono di determinate proprietà. Una definizione è, essenzialmente, una abbreviazione: è più semplice dire che un numero naturale è "primo" invece di dover spiegare ogni volta che il numero "è maggiore di 1 ed ammette come divisori solo sé stesso e 1". Una definizione non fornisce nuove informazioni e potrebbe, a rigore, essere eliminata, anche se i successivi enunciati perderebbero in concisione e in chiarezza. In effetti, una

definizione si introduce non solo per comodità linguistica, ma anche per fissare l'attenzione su certi concetti che giocheranno un ruolo importante nel seguito della trattazione (si pensi alla derivata in analisi, agli ideali di un anello, alla compattezza in topologia, o, a livello elementare, all'altezza di un triangolo).

Le definizioni hanno una fondamentale importanza nella didattica della matematica, perché abituano ad esprimersi correttamente (i programmi delle Elementari contengono sagge raccomandazioni in proposito). In certi casi, è istruttivo chiedere agli studenti di "trovare" da soli una definizione e non soltanto di ripeterla (ad esempio, quando si deve descrivere una situazione intuitivamente chiara, come nel caso della perpendicolarità fra retta e piano).

Esaminiamo ancora un paio di definizioni.

- "Un'equazione è una uguaglianza soddisfatta solo da particolari valori attribuiti alle incognite." Il problema è il significato da dare all'aggettivo "particolari": se lo intendiamo come un quantificatore esistenziale, allora non è più corretto parlare di equazioni impossibili; per motivi analoghi non è lecito intendere "particolari" come "non tutti". In realtà, la definizione riportata può risultare didatticamente efficace, ma non è rigorosa: il fatto che l'uguaglianza sia "soddisfatta solo da particolari valori" è un avvertimento, non una condizione che distingue le equazioni da altre uguaglianze.
- "Un numero L è estremo superiore di un insieme H non vuoto di numeri reali se è maggiore o uguale di ogni elemento di H e se, per ogni ε positivo, esiste un elemento x di H tale che $L-\varepsilon \leq x \leq L$." Questa definizione è corretta e accettabile nella didattica, ma è preferibile dire che "Un numero L è estremo superiore di un insieme H non vuoto di numeri reali se è il minimo numero che sia maggiore o uguale di ogni elemento di H (o, se si preferisce, se è il minimo maggiorante di H)". Questa seconda definizione, infatti, fa riferimento solo alla struttura di ordine e non ad operazioni, e può quindi essere applicata in situazioni più generali.

2. I teoremi

Un teorema esprime una nuova proprietà che riguarda concetti già noti, e che pertanto va dimostrata.

Per indicare i teoremi sono usati spesso altri nomi (con differenze più che altro psicologiche): corollario, lemma, criterio (quest'ultimo è un teorema che esprime una condizione sufficiente), regola (ad esempio, di Ruffini, di De L'Hospital: si tratta di enunciati che forniscono indicazioni su come svolgere un calcolo), legge (termine per lo più usato in fisica), principio (talvolta anche con il significato di assioma, o comunque fatto molto

generale: p. di induzione, p. di Cavalieri, p. del terzo escluso, p. di identità dei polinomi, p. di equivalenza delle equazioni).

La maggioranza dei teoremi che si incontrano nelle trattazioni matematiche è di tipo universale ($\forall x \forall y \ldots$) e contiene, come connettivo principale, l'implicazione. Queste circostanze, tuttavia, non capitano sempre: ad esempio, il teorema in geometria dello spazio secondo cui "esistono rette sghembe" non contiene né quantificatori universali né implicazioni). Altre volte, l'implicazione non è esplicita nel linguaggio naturale: così, "ogni triangolo è inscrivibile in un cerchio" va schematizzato nella forma "se tre punti non sono allineati, allora ...".

Solo per i teoremi che sono sotto forma di implicazione è sensato (ed importante): 1) distinguere ipotesi e tesi, 2) costruire l'enunciato inverso (che può essere o non essere un teorema). Spesso, è lecito parlare di "inverso di un teorema" con riferimento a più enunciati, o perché vi sono più ipotesi e soltanto alcune si scambiano con la tesi (una volta si introduceva in proposito il "soggetto"), o perché un enunciato nella lingua corrente si può formalizzare in più modi. Ad esempio, si consideri l'enunciato seguente, che si riferisce a quadrilateri in geometria piana, "se i lati opposti sono paralleli e un angolo è retto, allora le diagonali sono uguali". Possiamo costruire (almeno) i seguenti inversi: "se i lati opposti sono paralleli e le diagonali sono uguali, allora un angolo è retto", "se un angolo è retto e le diagonali sono uguali, allora i lati opposti sono paralleli" (uno solo dei due è un teorema).

Fra i vari tipi di dimostrazione, vale la pena di menzionare esplicitamente le dimostrazioni per assurdo: per dimostrare α, si mostra che da $\neg \alpha$ segue un assurdo. Si noti che, nel caso in cui α sia del tipo $p \rightarrow q$, allora $\neg \alpha$ equivale a $p \wedge \neg q$: la dimostrazione della contronominale $\neg q \rightarrow \neg p$ rientra in questo schema, perché consiste nel mostrare appunto che $\neg q$ è in contrasto con p.

Segnaliamo ancora che la dimostrazione di un teorema che assicura l'esistenza di un oggetto con determinate proprietà, può essere "costruttiva" (quando si mostra come ottenere un oggetto nelle condizioni volute), ovvero non costruttiva (quando si ragiona per assurdo, oppure si ricorre a questioni di cardinalità). Esempi tipici in proposito: l'esistenza di numeri trascendenti; il teorema di punto fisso per similitudini (che non siano isometrie).

3. Concetti primitivi e assiomi

Per dimostrare un teorema, si ricorre ad altri risultati già noti, i quali, a loro volta, erano stati ottenuti da precedenti proprietà. Questo processo a ritroso non può, evidentemente, proseguire all'infinito. Pertanto, è necessario che alcuni enunciati siano accettati senza dimostrazione: questi vengono detti postulati o assiomi. Una situazione analoga si incontra a proposito delle definizioni: in ogni definizione, il nuovo termine viene spiegato mediante parole già note; queste, a loro volta, erano state definite in base ad altri concetti. Anche in tal

caso deve esserci un punto di partenza, che è costituito dagli enti (o concetti, o termini) primitivi, che non vengono definiti.

Il discorso si applica alla geometria (dove fra gli enti primitivi riconosciamo in primo luogo i classici concetti di punto, retta, piano, legati dalle relazioni di incidenza), ma vale per ogni teoria matematica, anche se non è sempre altrettanto facile individuare assiomi ed enti primitivi. Ad esempio, in una teoria per l'aritmetica (numeri naturali con le operazioni elementari) non è necessario introdurre come primitivo il concetto di numero (perché tutti gli oggetti che si considerano sono numeri), ma è sufficiente considerare come enti primitivi un elemento specifico (lo zero) e le operazioni di passaggio al successore, di addizione e moltiplicazione.

Vediamo ora, con riferimento alla geometria, come viene giustificata la presenza (e la scelta) degli assiomi e degli enti primitivi.

In una prima impostazione abbiamo le risposte più intuitive. Gli enti primitivi vanno scelti fra quei concetti così chiari e naturali per tutti, che non c'è alcun bisogno di definirli (ad esempio, ci troviamo in imbarazzo se ci viene chiesto che cosa è una retta, ma non c'è bisogno di fornire alcuna spiegazione perché tutti abbiamo l'idea di retta). Analogamente, gli assiomi sono quelle proprietà, relative ai concetti primitivi, così evidenti che una loro dimostrazione, oltre che impossibile, sarebbe anche inutile.

Questa impostazione fu accettata, più o meno esplicitamente, fino al secolo scorso, ma appare oggi piuttosto ingenua (se siamo disposti ad accettare senza dimostrazione i postulati perché evidenti, allora dovremmo comportarci nello stesso modo di fronte a qualunque enunciato che secondo l'intuizione sia corretto, con conseguenze incontrollabili). Nel 1800 varie novità imposero un ripensamento generale. Ne citiamo alcune:

- 1830 nascita delle geometrie non euclidee, in cui si parte da presupposti diversi da quelli usuali: se gli assiomi esprimono proprietà ovvie, non possono certo essere modificati;
- 1840-50 introduzione di spazi a più di tre dimensioni: è difficile sostenere che in questi ambienti certi concetti, anche i più semplici, siano naturali ed evidenti;
- 1850-60 von Staudt, nell'ambito dello studio della geometria proiettiva, enuncia il principio di dualità: è lecito attribuire a parole come punto e retta significati diversi da quelli intuitivi, mantenendo la correttezza degli enunciati.

Le nuove posizioni assunte da più matematici verso la fine del secolo vengono sistemate e chiarite con l'opera di Hilbert (I fondamenti della geometria, 1899). Nell'impostazione hilbertiana, che è oggi normalmente accettata, i postulati sono visti come definizione dei concetti primitivi: non ci interessa sapere che cosa è un punto o una retta, ma siamo disposti ad usare questi termini ogni volta che certi oggetti soddisfano le proprietà

espresse dai postulati. Le figure acquistano, in questo contesto, solo un ruolo schematico-mnemonico, mentre perdono la funzione di rappresentazione, sia pure approssimata, degli enti geometrici: la geometria non è più una "scienza qualificata dai suoi oggetti", ma una scienza puramente deduttiva. Naturalmente, resta il problema dei rapporti fra geometria ed esperienza fisica.

4. Coerenza, indipendenza e completezza di un sistema di assiomi

Per introdurre una teoria assiomatica T, occorre in primo luogo precisare i simboli specifici del linguaggio (costanti, predicati, simboli per funzioni): tali simboli corrispondono agli enti primitivi della teoria. Occorre poi fissare l'insieme (o sistema) degli assiomi di T, cioè un insieme Σ di formule chiuse del linguaggio (l'insieme Σ comprende sempre gli assiomi logici); come regole di deduzione si intendono, salvo diverso avviso, le regole del calcolo dei predicati.

Applicando le regole a partire dagli assiomi si ottengono via via i teoremi. Per indicare che A è un teorema della teoria T che ha Σ come insieme di assiomi si scrive $\Sigma \vdash A$ oppure $T \vdash A$.

DEFINIZIONE Un sistema di assiomi Σ (o la corrispondente teoria) si dice *coerente* (o *consistente*, o *non contraddittorio*) se da Σ non si può dedurre una contraddizione, cioè se non esiste una formula (chiusa) A tale che $\Sigma \vdash A$ e $\Sigma \vdash \neg A$ (o equivalentemente $\Sigma \vdash A \wedge \neg A$).

TEOREMA Σ è inconsistente se e solo se per ogni formula B si ha $\Sigma \vdash B$.
Dimostrazione La direzione ($\leftarrow$) è ovvia. Per l'altra, basta osservare che $(A \wedge \neg A) \rightarrow B$ è una tautologia ed è quindi un teorema di T: di conseguenza, se una teoria dimostra $A \wedge \neg A$, allora dimostra anche B ("Modus Ponens").

DEFINIZIONE Un sistema di assiomi Σ si dice *dipendente* se esiste un assioma che si può dedurre dagli altri; cioè se esiste un $A \in \Sigma$ tale che $\Sigma - \{A\} \vdash A$; Σ si dice indipendente in caso contrario.

DEFINIZIONE Un sistema di assiomi Σ (o la corrispondente teoria) si dice *completo* se per ogni formula chiusa A si ha $\Sigma \vdash A$ oppure $\Sigma \vdash \neg A$ (spesso, parlando di teoria completa si sottintende che la teoria sia consistente).

La completezza (non prevista nei programmi Brocca) esprime una proprietà che a priori ci aspetteremmo in molte teorie, cioè che non restino situazioni incerte. Tuttavia, come si vedrà nel seguito del corso, questa aspettativa è destinata a rimanere (fortunatamente) insoddisfatta.

Si noti la stretta analogia che intercorre fra sistemi di assiomi e sistemi di equazioni: la consistenza corrisponde al fatto che le varie equazioni non esprimano condizioni contrastanti; l'indipendenza al fatto che nessuna equazione sia conseguenza delle altre, cioè che ciascuna equazione dia informazioni nuove; la completezza al fatto che le equazioni permettano di rispondere ad ogni domanda relativa alle incognite.

TEOREMA $\Sigma \vdash A$ se e solo se $\Sigma \cup \{\neg A\}$ è contraddittorio.
Dimostrazione ($\rightarrow$) Se Σ permette di dimostrare A, allora $\Sigma \cup \{\neg A\}$ permette di dimostrare sia A sia $\neg A$.
($\leftarrow$) Supponiamo che nella teoria $\Sigma \cup \{\neg A\}$ si dimostri una contraddizione. Allora, assumendo come assiomi solo le formule di Σ, da $\neg A$ segue una contraddizione: abbiamo così (nella teoria Σ) una dimostrazione per assurdo di A.

Ad esempio, il fatto che il postulato delle parallele non si possa dedurre dagli altri assiomi della geometria euclidea equivale alla consistenza della geometria (non euclidea) in cui si accetta, insieme agli altri assiomi, la negazione del postulato delle parallele.

Come si vedrà nel seguito, è spesso opportuno considerare un insieme infinito di assiomi. In proposito, vale il seguente

TEOREMA Un insieme Σ di assiomi è coerente se e solo se è coerente ogni sottoinsieme finito di Σ.
Dimostrazione L'implicazione ($\rightarrow$) è ovvia. Per l'implicazione inversa, basta osservare che ogni dimostrazione coinvolge un numero finito di assiomi (anche se la teoria ha infiniti assiomi); per cui, se si può ottenere una contraddizione, lo si può fare a partire da un numero finito di assiomi.

5. Esempi di teorie assiomatiche

Precisiamo che, in tutte le teorie che seguono, compare nel linguaggio il simbolo "=" (da interpretarsi nel modo naturale).

(A) La geometria euclidea è il primo (l'unico?) esempio di teoria assiomatica che viene presentato a Scuola. In realtà, da un punto di vista logico, si tratta di una teoria piuttosto complessa da formalizzare. Vediamo come si può procedere, limitandoci alla geometria del piano.

In primo luogo, per distinguere fra punti e rette, introduciamo due predicati unari P(x) ed R(x), da pensarsi intuitivamente come "x è un punto" ed "x è una retta"; occorrono poi un predicato binario I(x,y) per tradurre la relazione di incidenza ("il punto x sta sulla retta y") e

un predicato a quattro posti U(x,y,z,t) per la congruenza (uguaglianza) fra segmenti ("il segmento di estremi x,y è congruente al segmento di estremi z,t").

Seguendo (più o meno) l'impostazione di Hilbert, introduciamo i seguenti assiomi:

- Assiomi di collegamento o incidenza. Ad esempio: "Per due punti distinti passa una e una sola retta", "ogni retta contiene almeno due punti", "esistono tre punti non allineati".

A titolo di esercizio, traduciamo l'ultimo nel linguaggio formalizzato:

$$\exists x\, \exists y\, \exists z\, [P(x) \wedge P(y) \wedge P(z) \wedge \forall t\, (R(t) \rightarrow \neg\, (I(x,t) \wedge I(y,t) \wedge I(z,t)))]$$

- Assiomi di ordinamento, che permettono di introdurre i concetti di semiretta, segmento, semipiano.
- Assiomi di congruenza, relativi alla congruenza di segmenti (proprietà riflessiva, simmetrica, transitiva, possibilità di trasportare un segmento, ...). Un'osservazione didattica: una trattazione completa degli assiomi di congruenza è difficilmente presentabile in una Scuola Superiore, per cui spesso si fa ricorso all'intuizione. Nelle trattazioni classiche (a partire dagli Elementi di Euclide), i tre criteri di congruenza vengono giustificati con considerazioni intuitive, dopo di che svolgono un ruolo analogo a quello degli assiomi (anche se non sono sufficienti per una trattazione rigorosa dell'argomento).
- Assioma delle parallele. Dati un punto P ed una retta r non passante per P, esiste una e una sola retta per P che non ha punti in comune con r. (In realtà, come era già chiaro ad Euclide, è sufficiente richiedere l'unicità della parallela, in quanto, con gli assiomi di ordinamento e di congruenza, si riesce a dimostrare l'esistenza.)
- Assioma di Archimede. Dati due segmenti (non degeneri) esiste un multiplo del minore che supera il maggiore. Si noti che l'assioma di Archimede non si può formalizzare al prim'ordine, perché il quantificatore esistenziale agisce su una variabile che indica un numero naturale (e non un punto o una retta).
- Assioma di completezza, sul quale ci limitiamo a precisare che le formulazioni usuali non sono esprimibili al prim'ordine.

(B) Nelle geometrie non euclidee (iperbolica ed ellittica) si assume un postulato diverso dall'assioma delle parallele (per la geometria ellittica è necessaria anche una modifica agli assiomi di ordinamento). Non entriamo nei dettagli, ma ci limitiamo a una osservazione generale. Volendo classificare analogie e differenze fra le geometria iperbolica ed ellittica e la geometria euclidea, si possono distinguere:

- risultati comuni alle varie geometrie, come l'uguaglianza degli angoli alla base di un triangolo isoscele (in generale, può capitare che, partendo da ipotesi diverse, si ottengano gli stessi risultati);

- risultati di geometria euclidea che non si ritrovano nelle altre geometrie: ad esempio, il teorema di Pitagora e il suo inverso, i teoremi sulla similitudine, l'usuale introduzione del numero π ;
- risultati delle geometrie non euclidee che non valgono nella geometria euclidea: ad esempio, il "quarto" criterio di uguaglianza dei triangoli, secondo cui se due triangoli hanno gli angoli rispettivamente uguali, allora sono uguali;
- risultati specifici, che cioè riguardano situazioni che non si presentano in geometria euclidea (ad esempio, risultati relativi all'angolo di parallelismo);
- risultati classici che si possono generalizzare a tutte le geometrie a patto di modificare l'enunciato: ad esempio, il teorema dei seni vale in tutte le geometrie se viene enunciato così: in ogni triangolo il rapporto fra la lunghezza della circonferenza avente per raggio un lato ed il seno dell'angolo opposto è costante, cioè è lo stesso per i tre lati (sempre che si adotti una conveniente unità di misura per i segmenti).

(C) La teoria dei gruppi è una teoria assiomatica piuttosto semplice. Il linguaggio comprende una costante "1" (l'elemento neutro), il simbolo "$\cdot$" per l'operazione binaria. Gli assiomi sono quelli ben noti:

$$\forall x\forall y\forall z\,[(x\cdot y)\cdot z = x\cdot(y\cdot z)]\ ;\quad \forall x\,(x\cdot 1 = 1\cdot x = x)\ ;\quad \forall x\,\exists y\,(x\cdot y = y\cdot x = 1)$$

Si può dimostrare che l'insieme degli assiomi, nella forma vista, è dipendente (abbiamo fatto richieste sovrabbondanti) e non completo (ad esempio, dagli assiomi visti non si può dedurre né $\forall x\forall y\,(x\cdot y = y\cdot x)$ né $\neg\forall x\forall y\,(x\cdot y = y\cdot x)$.

Gli assiomi visti (o altri equivalenti) sono spesso presentati come "definizione" di gruppo; del resto, come abbiamo già notato, gli assiomi costituiscono una definizione degli enti primitivi.

Si noti la netta differenza epistemologica fra la teoria dei gruppi e la teoria della geometria euclidea: quest'ultima nasce per descrivere una specifica situazione, mentre la prima ha già in partenza lo scopo di adattarsi a strutture diverse.

Arricchendo opportunamente linguaggio e insieme degli assiomi, si ottiene la teoria dei campi. La cosa ha una rilevanza didattica, perché, entro certi limiti, si può ritenere che si tratti della teoria dell'algebra elementare (gli assiomi dell'algebra, quasi sempre impliciti, sono cioè le proprietà delle operazioni).

(D) I numeri reali costituiscono l'unico campo ordinato completo. Tuttavia, mentre non vi sono difficoltà a descrivere al prim'ordine le proprietà di campo ordinato, non si riesce ad esprimere al prim'ordine la completezza (abbiamo già accennato alle difficoltà poste dall'assioma di Archimede). In altre parole: ci sono assiomi al prim'ordine per i reali (si

aggiungono agli assiomi di campo altri assiomi, ad esempio per richiedere che -1 non sia un quadrato o che $1+1 \neq 0$), ma non è vero che i reali siano l'unica struttura che soddisfa quegli assiomi. Un sistema di assiomi per i numeri naturali verrà discusso nel seguito del corso.

(E) Anche la teoria degli insiemi si può impostare in modo assiomatico (il discorso non è comunque alla portata di studenti delle Superiori). Si potrebbe pensare che, nella teoria assiomatica degli insiemi, vadano intesi come primitivi i concetti di insieme e di elemento; tuttavia, è più semplice assumere come primitiva solo la relazione binaria di appartenenza, senza fare una distinzione a priori fra oggetti che sono elementi ed altri che sono insiemi. Fra gli assiomi, si richiedono in genere i seguenti (ovviamente non sufficienti):

- assioma dell'insieme vuoto ("c'è un insieme privo di elementi"); simbolicamente: $\exists x\ \forall y\ \neg(y\in x)$
- assioma di estensionalità ("se due insiemi hanno gli stessi elementi, allora sono uguali") $\forall x\ \forall y\ [\forall z\ (z\in x \longleftrightarrow z\in y) \rightarrow (x=y)]$

Fra gli altri esempi di teorie assiomatiche che possono avere interesse nella didattica delle Superiori, citiamo:

- impostazione assiomatica per la probabilità;
- altre assiomatiche per la geometria, in particolare: la proposta di Choquet (in cui si assume a priori la conoscenza dei reali) e l'impostazione di Tarski (che riesce ad esprimere al prim'ordine i concetti e le proprietà che usualmente si studiano in geometria, senza parlare esplicitamente di rette, ma solo di punti; più precisamente, si intende che tutte le variabili indichino punti, e si introducono due predicati, uno a tre posti e uno a quattro posti: $\beta(x,y,z)$ da interpretarsi come "i punti x, y, z sono allineati ed y è compreso fra x e z", e $\delta(x,y,u,v)$ da interpretarsi come "il segmento di estremi x, y è congruente al segmento di estremi u, v");
- ancora assumendo la conoscenza dei reali, si possono presentare assiomaticamente gli spazi metrici;
- definendo una relazione d'ordine parziale come una relazione riflessiva, antisimmetrica e transitiva, si introduce (in modo più o meno esplicito) la teoria degli insiemi ordinati.

Alcune proposte didattiche suggeriscono di presentare, specie nel biennio "deduzioni locali": in sostanza, si tratta di limitarsi a mostrare che da certi fatti (accettati perché intuitivi o verificabili sperimentalmente) ne seguono altri (a priori non prevedibili); il tutto senza preoccuparsi dell'indipendenza delle premesse, né di un unico inquadramento per tutti gli argomenti. In un ordine di idee non troppo diverso, si può introdurre la geometria affine (che ha come assiomi solo alcuni degli assiomi della geometria euclidea).

6. Modelli di una teoria - Il problema della coerenza

Per modello di una teoria T si intende una struttura in cui sono soddisfatti gli assiomi di T: ad esempio, ogni gruppo è un modello per la teoria dei gruppi. In ogni modello di T valgono, naturalmente, tutti i teoremi di T (le regole di deduzione sono state scelte proprio in modo da conservare la verità: un ragionamento corretto deve far passare da premesse vere a conclusioni vere).

Si osservi che una dimostrazione matematica non fa riferimento ai modelli, perché è condotta (o dovrebbe essere condotta) per via sintattica. Il discorso non è molto diverso da quello che si incontra anche nella pratica didattica: non si può dimostrare un teorema considerando solo un caso particolare (un teorema può essere verificato, o applicato, in un caso particolare).

L'introduzione e lo studio dei modelli rientra nella semantica di una teoria (contrapposta alla sintassi).

E' facile convincersi che, se un sistema di assiomi ammette un modello, allora non è contraddittorio. Infatti, tenuto conto che in un modello valgono tutti i teoremi della teoria, non può capitare che, in un unica struttura sia verificato un enunciato e, al tempo stesso, la sua negazione.

Come si vedrà, per le teorie al prim'ordine si dimostra anche l'implicazione inversa; in altre parole, una teoria è coerente se e solo se ammette almeno un modello.

L'osservazione precedente costituisce un semplice criterio per stabilire la consistenza di una teoria. Ad esempio, la teoria dei gruppi è consistente perché esiste un gruppo finito (il fatto che il gruppo sia finito permette di eseguire tutti i controlli). Tuttavia, come si vedrà nel seguito del corso, per molte teorie importanti non siamo nelle condizioni di dimostrare la coerenza.

A questo punto siamo in grado di introdurre la "consistenza relativa" con riferimento ai modelli (ci sono anche definizioni sintattiche). Se una teoria T è consistente, allora ammette un modello M; supponiamo che, a partire da M, si riesca a costruire un modello M' di una teoria T'; in queste condizioni, anche T' risulta consistente (perché ammette un modello). Si dice allora che T' è consistente relativamente a T.

ESEMPI

(A) Ricorrendo alla geometria analitica, si dimostra la consistenza relativa della geometria euclidea rispetto alla teoria dei numeri reali: se chiamiamo punti le coppie ordinate di numeri reali e rette gli insiemi dei punti le cui coordinate soddisfano un'equazione lineare in due variabili, gli assiomi risultano soddisfatti, cioè si ottiene un modello per la geometria euclidea.

(B) I modelli (di Klein, di Poincaré, ...) per le geometrie non euclidee sono costruiti all'interno del piano o dello spazio euclideo, nel senso che, proprio sfruttando le proprietà della geometria euclidea, riusciamo a dimostrare che sono soddisfatti gli assiomi delle nuove geometrie. Di conseguenza, se gli assiomi della geometria ellittica (ad esempio) si contraddicessero fra loro, allora troveremmo una contraddizione su una superficie sferica della geometria euclidea. In definitiva: se la geometria euclidea non è contraddittoria, allora anche le geometrie ellittica e iperbolica non sono contraddittorie.

Segnaliamo anche l'importanza didattica dei modelli delle geometrie non euclidee, che permettono di visualizzare i nuovi ambienti.

Alcune citazioni

(LEIBNIZ) «Le definizioni di per sé sono arbitrarie; tuttavia devono adattarsi all'uso e al consenso comune.»

(HOBBES) «Se le definizioni sono arbitrarie, tutta la matematica, che si basa sulle definizioni, è arbitraria.»

(SACCHERI) «Una definizione è figlia di molti teoremi.»

(PEANO) «Le definizioni sono utili ma non necessarie, perché al posto del definito si può sempre sostituire il definiente [...]. Se la nuova definizione non è più lunga e più complicata, quella definizione era poco utile. Se si incontrano difficoltà, la definizione non fu ben data.»

(POINCARÉ, 1900) «Russell distingue tre tesi che egli respinge allo stesso modo: 1° La verità della geometria euclidea ci è nota a priori prima di ogni esperienza. 2° Una geometria è vera e le altre sono false, ma non potremo mai sapere qual è quella vera. 3° Nessuna geometria è vera o falsa. Russell crede che adotti la terza tesi, ma non ne è sicuro; posso rassicurarlo: adotto la terza tesi e rifiuto in modo assoluto le prime due.»

(POINCARÉ, 1902) «Gli assiomi non sono né giudizi sintetici a priori né fatti sperimentali: sono convenzioni. La nostra scelta tra tutte le convenzioni possibili è guidata da fatti sperimentali, ma resta libera ed è condizionata solo dalla necessità di evitare contraddizioni. [...] Gli assiomi della geometria sono semplici definizioni mascherate. Che si deve quindi pensare della questione circa la verità della geometria? Essa non ha alcun senso. Sarebbe come domandare se il sistema metrico sia vero e false le antiche misure; se siano vere le coordinate cartesiane e false quelle polari. Una geometria non può essere più vera di un'altra; può essere soltanto più comoda.»

(EINSTEIN, 1921) «Gli assiomi *definiscono* gli oggetti di cui si occupa la geometria. [...] Io attribuisco speciale importanza all'interpretazione della geometria che ho ora esposto, perché senza di essa non sarei stato capace di formulare la teoria della relatività.»

(FREGE, circa 1900) Gli assiomi «non vengono dimostrati perché la loro conoscenza scaturisce da una fonte conoscitiva di natura extralogica, che possiamo chiamare intuizione spaziale. Il fatto che gli assiomi siano veri ci assicura di per sé che essi non si contraddicono fra loro, e ciò non richiede alcuna ulteriore dimostrazione.»

(HILBERT, in risposta a Frege) «[...] io ho detto esattamente il contrario: se assiomi arbitrariamente stabiliti non sono in contraddizione, con tutte le loro conseguenze, allora essi sono veri. [...] Se voglio intendere un qualunque sistema di enti, per esempio il sistema: amore, legge, spazzacamino, allora basterà che assuma tutti i miei assiomi come relazioni fra questi enti perché le mie proposizioni, ad esempio il teorema di Pitagora, valgano anche per essi. In altre parole, ogni teoria può essere applicata a infiniti sistemi di enti fondamentali. Tale circostanza non rappresenta un difetto della teoria (ne è piuttosto un grandissimo pregio) e in ogni caso è inevitabile.»

(SEVERI, 1952) «Io feci un'esperienza nei miei lontani anni di Padova, nell'insegnamento della geometria proiettiva (1904-05). Per chiudere la porta ad ogni intuizione degli allievi, denotai le idee primitive (punto, retta, piano) con parole ignote ai miei ascoltatori e, enunciati su esse i postulati, dimostrai, col cervello chiuso all'intuizione, i primi teoremi. Presto però mutai rotta, perché compresi che avrei potuto esser seguito da qualche fanatico delle cose nuove e singolari, ma che non avrei insegnato quel che dovevo insegnare.»

Nel 1821 CAUCHY disse che si proponeva di "dare ai metodi dell'analisi lo stesso rigore che si chiede in geometria". L'analisi odierna soddisfa alle esigenze espresse da Cauchy ?

Esercizi

CLAUDIO BERNARDI

ESERCIZIO 1 In ciascuno dei due casi seguenti, scegliere fra (A) e (B) in modo da ottenere una affermazione corretta.

(A) In una definizione ci deve essere una e una sola parola (o locuzione) nuova
(B) Nell'enunciato di un teorema ci deve essere una e una sola parola (o locuzione) nuova

- Se nell'enunciato di un teorema rimpiazziamo ogni parola con la rispettiva definizione e continuiamo questo processo fin che è possibile, troviamo

(A) un enunciato che esprime un legame fra gli enti primitivi
(B) un assioma.

ESERCIZIO 2 Sia Σ un insieme di assiomi indipendente e coerente. Eliminando uno degli assiomi di Σ si ottiene un insieme di assiomi Σ'. Il sistema Σ' è indipendente? E' coerente? Può essere completo?

ESERCIZIO 3
a. Siano Σ e Σ' due diversi sistemi di assiomi equivalenti, cioè che generano lo stesso insieme dei teoremi. Se Σ è consistente, o indipendente, o completo, valgono le analoghe proprietà per Σ' ?

b. Quali delle proprietà considerate al punto a. valgono per un sistema Σ se e solo se valgono per ogni suo sottinsieme finito?

ESERCIZIO 4 Una teoria è completa e non contraddittoria. Dimostrare che se si aggiunge un altro assioma, il sistema di assiomi che si ottiene è o contraddittorio o dipendente.

ESERCIZIO 5 I tre assiomi seguenti descrivono la struttura dell'insieme dei numeri naturali con la sola operazione di successore: "0 non è il successore di alcun numero", "se due numeri hanno successori uguali allora sono uguali", "ogni numero diverso da 0 è il successore di un opportuno numero". Esprimere questi tre assiomi nel linguaggio al prim'ordine che disponga della costante 0, del simbolo ' (operazione unaria per il successore) e del simbolo di uguaglianza. Trovare poi un modello per questa teoria, che non sia isomorfo ai numeri naturali.

ESERCIZIO 6 Scrivere assiomi (in un opportuno linguaggio al prim'ordine) in modo che i modelli siano tutti e soli gli ordini totali densi, privi di minimo e massimo. [Ad esempio, si può assumere fra gli assiomi la formula $\forall x\ \exists y\ (x < y)$. Si può dimostrare che si ottiene una teoria completa.]

ESERCIZIO 7 Sia Σ un insieme di assiomi e sia F una formula chiusa. Dimostrare che se $\Sigma \cup \{F\}$ è indipendente, allora $\Sigma \cup \{\neg F\}$ è non contraddittorio. (Se, accettando ... si arriva ad una contraddizione, allora abbiamo una dimostrazione per assurdo di ...). Dedurre che, se F è una formula tale che né F né $\neg$ F sono dimostrabili in una teoria consistente, allora si può aggiungere come assioma sia F sia $\neg$ F, ottenendo due teorie ancora consistenti. (Un esempio è costituito dal postulato delle parallele e dalla teoria formata dagli altri assiomi della geometria euclidea.)

ESERCIZIO 8 Si consideri la teoria con i seguenti assiomi:

a. assiomi di gruppo (rispetto ad un'operazione indicata con $*$)
b. $\exists x\ \exists y\ [\neg(x=y) \wedge \forall z\ (z=x \vee z=y)]$
c. $\exists x\ \exists y\ \neg(x*x=y*y)$

La teoria è consistente? E delle tre teorie che si ottengono considerando solo due dei tre assiomi, quali sono consistenti? (Si tenga presente che Z_2, cioè le classi di resto modulo 2 con l'addizione, è, a meno di isomorfismi, l'unico gruppo con due elementi.)

ESERCIZIO 9 Sia M un modello per una teoria non contraddittoria T e sia F una formula chiusa. Si può presentare la situazione descritta in ciascuno dei seguenti casi?

a. F è un teorema di T ed F è falsa in M
b. F non è un teorema di T ed F è vera in M
c. $\neg$ F è un teorema di T ed F è vera in M.

Cambiano le risposte se si suppone che T sia completa?

ESERCIZIO 10 Sia M un modello di Σ e si supponga che in M sia soddisfatta una formula chiusa α. Quali delle seguenti affermazioni si possono dedurre da quanto premesso?

a. $\Sigma \vdash \alpha$
b. non è vero che $\Sigma \vdash \neg \alpha$
c. Σ è consistente
d. $\Sigma \cup \{\alpha\}$ è completo

ESERCIZIO 11 Sia Σ un insieme di assiomi e sia α una formula chiusa tale che $\Sigma \cup \{\alpha\}$ è indipendente. Quali delle seguenti affermazioni si possono dedurre da quanto premesso?

a. Σ è consistente
b. $\Sigma \cup \{\neg \alpha\}$ è consistente
c. α non è teorema della teoria generata da Σ
d. $\neg \alpha$ non è teorema della teoria generata da Σ

ESERCIZIO 12 Sia α_n, per ogni n>1, la formula seguente

$$\exists x_1\ \exists x_2 \ldots \exists x_n\ [\neg (x_1 = x_2) \wedge \neg (x_1 = x_3) \wedge \ldots \wedge \neg (x_{n-1} = x_n)\ ].$$

Dopo aver verificato che la formula α_n asserisce che "esistono almeno n elementi distinti", si considerino le tre teorie che hanno come assiomi rispettivamente

(i). $\{ \alpha_2, \alpha_3, \alpha_4, \ldots \}$; (ii) $\{ \alpha_2 , \alpha_5 , \alpha_6 \}$; (iii) $\{ \neg \alpha_6 \}$

Stabilire quali delle teorie citate: (a) ammettono solo modelli finiti; (b) ammettono modelli sia finiti sia infiniti; (c) ammettono come modelli tutti e soli gli insiemi infiniti.

ESERCIZIO 13 Si trovi un insieme T di assiomi tale che, per ogni insieme finito M, M è un modello di T sse (se, e solo se) M ha un numero pari di elementi. (*Suggerimento*: si considerino le formule α_n introdotte negli esercizi precedenti.) La teoria T ammette anche modelli infiniti?

ESERCIZIO 14 L'insieme Σ degli assiomi della teoria (i) del precedente esercizio 12 è infinito. Mostrare che, se si considera un sottoinsieme finito di Σ, si ottiene una teoria non equivalente a quella iniziale (cioè con teoremi diversi). [In effetti, si può dimostrare che la teoria considerata non è "finitamente assiomatizzabile"]. Il sistema di assiomi considerato è indipendente? Si può rendere indipendente eliminando gli assiomi superflui?

ESERCIZIO 15 E' vero che ogni sistema di assiomi si può ridurre ad un sistema equivalente formato da un solo assioma? (Si distingua il caso gli assiomi sono in numero finito da quello in cui vi sono infiniti assiomi.)

Schemi di Deduzione

CARLO MARCHINI
Dipartimento di Matematica
Università di Parma
Via D'Azeglio 85/A
43100 PARMA

La Logica è una sorta di "microscopio" che analizza il discorso matematico per "catturarne" le articolazioni deduttive. Non è stenografia, anche se da esigenze di abbreviazione si sono originati i simboli usati nella Matematica, che oggi la Logica studia e utilizza. La nostra disciplina chiarisce e specifica il concetto di *proprietà*, usato nei postulati di Peano e in teoria degli insiemi; definisce il concetto di *computabile*, ecc. In questi e altri contesti l'analisi deve andare al di là della stenografia. Frege dà inizio alla Logica matematica attuale: egli propone il linguaggio formale come lo strumento (con connotati algebrici) idoneo a descrivere le *proprietà*, intese come sinonimo di *formule*. Il linguaggio formale è inoltre "utensile" privilegiato per dare certezza e chiarezza alle argomentazioni e garanzia della correttezza delle dimostrazioni.

Una prima analisi (pedante) riguarda la forma delle espressioni; è indispensabile per condurre in seguito numerose dimostrazioni basandosi sulla costruzione delle scritture che interessano. Fino a poco fa era di scarsa valenza didattica, oggi la "rigidità" dei calcolatori l'ha rimessa in auge. Un esempio: nella scrittura $3+x \doteq 7$ oltre alla ricerca della soluzione il logico vede aspetti diversi: c'è un predicato binario particolare, l'eguaglianza che per il momento si considera solo come un generico predicato binario [1]. Così si evidenzia che $3+x$ e $3+x \doteq 7$ sono scritture di enti di natura diversa. La seconda esprime un giudizio problema-

[1] Usando la notazione *prefissa* si scrive $=(3+x,7)$. Con un predicato binario sono possibili due scritture: quella che premette il predicato, detta *prefissa*: $=(3+x,7)$ e quella in cui il predicato è posto in mezzo, detta *infissa*: $3+x = 7$. Se si considera un predicato unario o ternario, non c'è più questa possibilità. Così, per uniformità, si preferisce l'uso dei predicati prefissi. Fanno eccezione l'eguaglianza, i simboli usati per relazioni d'ordine e parallelismo, perpendicolarità ed altri, ormai di uso consolidato in Matematica.

tico, la prima ha la stessa "natura" di numero di 3 e 7. In esse compare x, l'*indeterminata*. Da questo esempio si individuano l'indeterminata, x, le costanti (individuali), 3 e 7, la scrittura che assomiglia alle variabile ed alle costanti (individuali), 3+x, a cui si dà il nome generico di *termine* ed infine la frase, *formula*, ottenuta applicando il predicato binario a due "numeri", meglio, termini.

L'analisi delle scritture precedenti e di altro tipo si precisa in modo rigoroso partendo dall'*alfabeto*. La definizione precisa di alfabeto ha notevoli difficoltà di carattere didattico, aridità, scarso coinvolgimento da parte degli studenti, ecc. E' più efficace ottenere una definizione di questo tipo solo con un percorso didattico che preveda la definizione come punto di arrivo, nello stile di Lakatos, così come è stato nella logica, ma si rischia di allungare i tempi. La via migliore è prendere in esame situazioni familiari, ad esempio i polinomi con il loro linguaggio e le loro proprietà. La riflessione didattica dovrebbe staccarsi dagli aspetti algoritmici, predominanti nella prassi, per giungere ad individuare gli oggetti linguistici utilizzati [2]. I polinomi sono l'esempio principale del concetto di termine. Ma questa strategia non è facile ed è poco motivante.

Sempre dall'esperienza scolastica si può introdurre un'operazione grammaticale importante, la *sostituzione*. Il teorema (regola) di Ruffini ne è un esempio: dato un polinomio p(x) si vuole sapere se è divisibile per x - a. Si può procedere in due modi, uno più lungo ed

[2] Per definire i polinomi c'è bisogno di un insieme X di oggetti, non altrimenti specificati, detti *indeterminate* o *variabili*. La scelta tra queste due dizioni merita un poco di attenzione. Molto spesso in Algebra ed in Analisi si adoperano come sinonimi. Ma la prima ha un connotato più grammaticale, la seconda ha un sapore più semantico. Indeterminata è un simbolo che non sta al posto di nessun altro, variabile sottintende un ambito di variazione. La differenza si può cogliere meglio quando si considerano i polinomi, per esempio a coefficienti reali. Dato che $\mathbb{R}$ è un campo, con una costruzione che si presenta solitamente nei corsi universitari di Algebra, si può considerare l'anello commutativo con unità, $\mathbb{R}[x]$ dei polinomi a coefficienti reali. All'università i polinomi sono successioni di numeri reali definitivamente nulle, non la *somma di monomi*! In questo modo l'indeterminata x con le sue potenze è semplicemente un indicatore di posto, in luogo delle virgole con cui solitamente si rappresentano le successioni. Non ha alcun valore numerico. Quando invece si parla della variabile x si considera non più il polinomio, elemento di $\mathbb{R}[x]$, ma una funzione da $\mathbb{R}$ a $\mathbb{R}$ che una volta fissato il valore di x, ad esso fa corrispondere un numero reale. Un esempio di questa "confusione" si può trovare sui testi al momento della presentazione del *Principio di identità dei polinomi*. Per alcuni due polinomi in x sono identici quando, ridotti a forma normale, hanno lo stesso grado ed i coefficienti delle potenze di x rispettivamente eguali. In altri libri si trova che due polinomi sono identici se assumono gli stessi valori per gli stessi valori della variabile x. In una fase strettamente grammaticale è preferibile la dicitura *indeterminata*.

Un polinomio si presenta come scrittura in cui compaiono addizioni e moltiplicazioni tra numeri ed indeterminata(te). Dal punto di vista della struttura delle scritture, ogni costante è un polinomio (di grado 0), x è un polinomio (di grado 1) e se p(x) e q(x) sono polinomi, anche (-p(x)) è un polinomio (con lo stesso grado di p(x)); (p(x) + q(x)) è un polinomio (il cui grado è minore o eguale al massimo dei gradi di p(x) e q(x)) e (p(x)·q(x)) è polinomio (di grado pari alla somma dei gradi di p(x) e q(x)).Questa potrebbe utilmente essere presa come definizione di polinomio a coefficienti in un campo anche nella scuola secondaria. Le parentesi scritte sono superflue, ma possono servire per evidenziare che (-p(x)) oppure (p(x) + q(x)) o (p(x)·q(x)), sono enti nuovi, con una loro individualità, anche se costruiti con "mattoni" già presenti. Invece $\left(\frac{p(x)}{q(x)}\right)$ non è un polinomio. Qui interessa focalizzare il procedimento di costruzione *per ricursione*, che si ripete con opportuni mutamenti in altre parti della logica: dati certi enti, per il momento non meglio precisati, detti polinomi, eseguendo su esse alcune operazioni consentite, si ottiene ancora un ente dello stesso tipo.

uno più veloce: il primo consiste nell'eseguire la divisione e determinare il resto, con uno dei procedimenti noti, il secondo, più rapido: si calcola p(a), la scrittura che si ottiene sostituendo a al posto di x o, più correttamente, calcolando il valore della funzione polinomiale associata a p su a. Si può indicare p(a) con la scrittura più complessa, ma più esplicita, p(x/a) oppure con (x/a)(p(x)) [3]. In essa si specifica che bisogna porre a in luogo di x e non viceversa, anche se a è presente tra i coefficienti del polinomio. Bisogna conferire all'operazione di sostituzione la natura di un omomorfismo, visto che il risultato della sostituzione (x/a)(x^2 - 3x + 1) è a^2 - 3a + 1, omomorfismo per il quale devono essere indicate le strutture tra cui opera. [4]

Come detto più volte, dato il carattere rigoroso della Logica, anche queste operazioni grammaticali vanno correttamente definite, seppure a scapito della semplicità didattica. Gli esempi riportati giustificano, almeno psicologicamente, le definizioni che seguono:

[3] Alcuni autori usano il simbolo di sostituzione scambiando il ruolo dei termini, premettendo il termine da sostituire alla indeterminata: p(a) è ottenuto dalla sostituzione (a|x)(p(x)).

[4] Un altro esempio è dato dai sistemi di equazioni. Per risolvere $\begin{cases} x + 3y - z = 2 \\ x^2 - 3y^2 + xz = 5 \\ x - yz = 0 \end{cases}$ si applica il metodo *per sostituzione*: dalla prima si ricava x = 2 - 3y + z che va sostituito nella rimanenti, con una sostituzione che per analogia a quanto scritto prima può essere indicata con (x/(2 - 3y)). Il risultato è

$$\begin{cases} (2 - 3y + z)^2 - 3y^2 + (2 - 3y + z)z = 5 \\ (2 - 3y + z) - yz = 0 \end{cases}$$

Questo esempio è interessante perché a differenza del precedente, non si opera solo su un polinomio (termine), bensì sulla congiunzione di due eguaglianze (formula) seppure lasciando inalterati i simboli in cui non interviene x. Il fatto che si tratti di una formula vera e propria non è messo bene in luce dalla scrittura: nei sistemi la parentesi graffa aperta che "raduna" le varie equazioni è il connettivo di congiunzione ∧. Quindi ancora una volta alla sostituzione è conferita la natura di omomorfismo, non solo riguardo alle operazioni costitutive dei termini, anche rispetto ai connettivi (e pure dei quantificatori, con qualche cautela). Un esempio più complicato è quello delle formule risolutive. Data la generica equazione di secondo grado $ax^2 + bx + c = 0$, si determinano le soluzioni con la formula ben nota,

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

Spesso si considerano esempi di equazioni di secondo grado in cui i coefficienti a, b e c assumono i valori più svariati. Ad esempio per risolvere l'equazione $(a - b)x^2 - (a + b)x + a = 0$, basta applicare la formula risolutiva ed ottenere

$$x_1 = \frac{(a + b) - \sqrt{(a + b)^2 - 4a(a - b)}}{2(a - b)} \qquad x_2 = \frac{(a + b) + \sqrt{(a + b)^2 - 4a(a - b)}}{2(a - b)}$$

Qui non interessa svolgere i calcoli, ma mettere in luce aspetti logici importanti. Molte volte lo scopo di ripetuti esercizi sull'argomento è quello di far acquisire le formule risolutive e di far apprendere come eseguire le trasformazioni. Bisogna però prestare attenzione: la sostituzione usata è rappresentabile come (a/(a+b) b/(a-b) c/a), ed è una sostituzione simultanea che non può esser effettuata facendo prima la sostituzione di a e poi quella di b e poi quella di c o in qualunque altro ordine si voglia, perché darebbe luogo a risultati ben diversi, a meno di non aggiungere altri simboli.

DEFINIZIONE Sia A l'alfabeto di un calcolo dei predicati e sia X un insieme di indeterminate; si definisce l'*insieme delle sottoformule* di una formula come segue:

1) Per ogni $n \geq 1$ ed ogni costante predicativa $R \in P_n$, presi comunque $t_1, \ldots, t_n \in T_A(X)$; si ha $sub(Rt_1,\ldots,t_n) = \{Rt_1,\ldots,t_n\}$;
2) se φ è una formula, $sub((\neg\varphi)) = sub(\varphi) \cup \{(\neg\varphi)\}$;
3) se φ e ψ sono formule, e $\Diamond \in \{\wedge,\vee,\rightarrow,\leftrightarrow\}$, allora $sub((\varphi\Diamond\psi)) = sub(\varphi) \cup sub(\psi) \cup \{(\varphi\Diamond\psi)\}$;
4) se φ è una formula $x \in X$ e $Q \in \{\forall,\exists\}$, allora $sub((Qx\varphi)) = sub(\varphi) \cup \{(Qx\varphi)\}$;
5) nient'altro [5].

Si dice che ψ è una *sottoformula* di φ se $\psi \in sub(\varphi)$.

Questa è una definizione per ricursione, simile a quella di polinomi data in nota.

Sul concetto di indeterminata *libera* e *vincolata*, agli inizi del secolo c'è stata grande discussione. I logici hanno enucleato il concetto e ne hanno compresa l'importanza, anche se da sempre i matematici e non solo essi ne hanno fatto uso. [6] Una situazione simile a quelle citate negli esempi in nota, si ripete coi quantificatori, che in certo senso generalizzano tali esempi. I quantificatori mutano le indeterminate presenti in una formula da libere in vincolate: se si considera $(\forall x\varphi)$, e $x \neq y$, se y è libera in φ, (nella prossima definizione

[5] In questa definizione, come nelle successive, si considerano tutte le parentesi, anche quelle che nell'uso corrente si eliminano, facendo uso delle consuete convenzioni di scrittura, analoghe a quelle utilizzate per l'ordine delle operazioni algebriche nelle espressioni.

[6] Alcuni esempi matematici: si consideri l'eguaglianza $\sum_{i=0}^{3}\sum_{n=0}^{i} i\cdot n = \sum_{n=0}^{0} i\cdot n + \sum_{n=0}^{1} i\cdot n + \sum_{n=0}^{2} i\cdot n + \sum_{n=0}^{3} i\cdot n = (0\cdot 0) + (1\cdot 0 + 1\cdot 1) + (2\cdot 0 + 2\cdot 1 + 2\cdot 2) + (3\cdot 0 + 3\cdot 1 + 3\cdot 2 + 3\cdot 3) = 0 + 1 + 6 + 18 = 25$. Nella scrittura compaiono sia i che n, variabili qui, perché assumono valori naturali, ma il risultato dipende solo da 0 (due volte) e da 3. Infatti se si considerasse 2 invece di 3, si avrebbe $\sum_{i=0}^{2}\sum_{n=0}^{i} i\cdot n = (0\cdot 0) + (1\cdot 0 + 1\cdot 1) + (2\cdot 0 + 2\cdot 1 + 2\cdot 2) = 0 + 1 + 6 = 7$. In modo analogo se in luogo del primo 0 si ponesse 1 o in luogo del secondo 0 si ponesse 2, il risultato cambierebbe. Il "pezzo" $\sum_{n=0}^{i} i\cdot n$ dipende da i, ma non da n, perché se $i = 2$ esso è dato da $(2\cdot 0 + 2\cdot 1 + 2\cdot 2) = 6$ e se $i = 3$, è dato da $(3\cdot 0 + 3\cdot 1 + 3\cdot 2 + 3\cdot 3) = 18$. Dunque il fatto che n compaia nell'espressione di cui si fa la somma e sotto il simbolo di sommatoria, ha l'effetto di "far sparire" n, cioè di rendere il risultato indipendente da n. In questa scrittura, $\sum_{n=0}^{i} i\cdot n$, il ruolo delle variabili è ben diverso: i è libera, cioè il risultato dipende dal suo valore; n è vincolata, il risultato è indipendente da n. Mettendo poi davanti il simbolo di sommatoria su i, il risultato non ne dipende più: la variabile i diventa vincolata. Situazioni analoghe si incontrano in Matematica anche con i simboli di integrale e di limite: $\int_a^b f(x)dx$ dipende solo da f, da a e b come si può facilmente vedere, non da x, dato che $\int_a^b f(x)dx = \int_a^b f(t)dt$, $\lim_{x\to 0}\frac{\operatorname{sen} x}{x} = \lim_{y\to 0}\frac{\operatorname{sen} y}{y}$. In teoria degli insiemi con la scrittura si presenta la stessa situazione mediante l'operatore di astrazione $\{x \mid P(x)\}$ ed anche con le operazioni infinitarie di unione, intersezione e prodotto cartesiano.

vedremo che cosa ciò significa), allora y resta libera in $(\forall x\varphi)$; se invece si considera $(\forall y\varphi)$, l'indeterminata y non è più libera in $(\forall y\varphi)$. Considerazioni analoghe valgono per il quantificatore esistenziale. I quantificatori sono i soli "autorizzati" a compiere ciò [7]. I connettivi e l'eguaglianza non hanno nessun effetto a questo riguardo. I concetti di indeterminata libera e vincolata e di sostituzione si precisano con le seguenti definizioni:

DEFINIZIONE

a) Per ogni termine di $t \in T_A(X)$ si definisce l'insieme $Lib(t) \subseteq X$, delle indeterminate (libere) presenti in t, con le seguenti clausole:
   - per ogni costante individuale c, $Lib(c) = \emptyset$;
   - per ogni variabile individuale $x \in X$, $Lib(x) = \{x\}$;
   - se $t_1,\ldots,t_n \in T_A(X)$ e $f \in F_n$, $Lib(f(t_1,\ldots,t_n)) = Lib(t_1) \cup \ldots \cup Lib(t_n)$.
b) Sia y una indeterminata e τ un termine, si definisce la sostituzione (y/τ) ponendo
   - per ogni costante individuale c, $(y/\tau)(c) = c$;
   - per ogni variabile individuale $x \in X$, $(y/\tau)(x) = x$, se $x \neq y$; $(y/\tau)(y) = \tau$;
   - se $t_1,\ldots,t_n \in T_A(X)$ e $f \in F_n$, $(y/\tau)(f(t_1,\ldots,t_n)) = f((y/\tau)(t_1), \ldots, (y/\tau)(t_n))$

DEFINIZIONE

a) Per ogni formula $\varphi \in L_A(X)$ si definisce l'insieme $Lib(\varphi) \subseteq X$, delle indeterminate libere presenti in φ, con le seguenti clausole:
   - se $t_1,\ldots,t_n \in T_A(X)$ e $R \in P_n$, $Lib(Rt_1,\ldots,t_n) = Lib(t_1) \cup \ldots \cup Lib(t_n)$;
   - per ogni formula φ, $Lib(\neg\varphi) = Lib(\varphi)$;
   - per ogni formula φ, ψ, $Lib(\varphi\Diamond\psi) = Lib(\varphi) \cup Lib(\psi)$, per $\Diamond \in \{\wedge,\vee,\rightarrow,\leftrightarrow\}$;
   - per ogni formula φ e indeterminata x, $Lib((\forall x\varphi)) = Lib((\exists x\varphi)) = Lib(\varphi) - \{x\}$.
b) Sia x un'indeterminata presente nella formula φ. Si dice che essa è *libera in* φ se $x \in Lib(\varphi)$, altrimenti si dice che l'indeterminata è *vincolata in* φ [8].
c) Si dice *enunciato* una formula φ tale che $Lib(\varphi) = \emptyset$.
d) Data una formula φ tale che $Lib(\varphi) \subseteq \{x_1,\ldots,x_s\}$, la formula $(\forall x_1,\ldots,x_s\varphi)$ è un enunciato [9] che viene detto (*una*) *chiusura universale* di φ.

[7] Nei linguaggi che si considerano qui, vale a dire i cosiddetti linguaggi del primo ordine con eguaglianza. In essi, come detto sopra, la quantificazione è applicabile solo ad indeterminate. Vi sono altri tipi di linguaggi, ad esempio quello in cui si quantifica anche sui predicati. Tali tipi di linguaggi pongono vari ordini di problemi che mi pare esorbitino da una trattazione elementare anche se di essi si fa uso spesso (non sempre corretto) nella comunicazione interpersonale.

[8] In accordo con la definizione precedente tutte le presenze di x nel rango d'azione di un quantificatore Qx sono vincolate, ma un'indeterminata in una formula può avere contemporaneamente presenze libere e vincolate.

[9] In questa scrittura non sono state rispettate le regole sulla formazione delle formule. Secondo la definizione, dopo il primo quantificatore a destra bisogna mettere la parentesi, e questo ripeterlo dopo ogni quantificatore. Ma le convezioni di semplificazione di cui si è detto prima, consentono queste scritture più agevoli, senza dover

e) Sia y una indeterminata e τ un termine, si definisce la sostituzione (y/τ) ponendo

- per ogni formula atomica $Rt_1,\ldots,t_n$,
$$(y/\tau)(Rt_1,\ldots,t_n) = R(y/\tau)(t_1), \ldots,(y/\tau)(t_n);$$
- per ogni formula φ, $(y/\tau)(\neg\varphi) = (\neg(y/\tau)(\varphi))$;
- per ogni formula φ, ψ, $(y/\tau)(\varphi\Diamond\psi) = ((y/\tau)(\varphi)\Diamond(y/\tau)(\psi))$, per $\Diamond\in\{\wedge,\vee,\rightarrow,\leftrightarrow\}$;
- per ogni formula φ e indeterminata x, $(y/\tau)(\forall x\varphi) = (\forall x\varphi)$ e $(y/\tau)(\exists x\varphi) = (\exists x\varphi)$ se $y = x$, altrimenti, $(y/\tau)(\forall x\varphi) = (\forall x(y/\tau)(\varphi))$ e $(y/\tau)(\exists x\varphi) = (\exists x(y/\tau)(\varphi))$.

Il punto d) della definizione precedente insegna a costruire particolari enunciati, le chiusure universali, importanti nel seguito [10].

Diamo ora una definizione di difficile comprensione, la cui utilità diviene chiara quando si introducono i sistemi deduttivi. Sia data una formula φ, in cui compaia eventualmente l'indeterminata x e sia t un termine, si ha

DEFINIZIONE Il termine t si dice *libero per* x *in* φ se non esiste una formula del tipo $(Qy\psi)\in sub(\varphi)$, con $Q\in\{\forall,\exists\}$, tale che $x\in Lib(Qy(\psi))$ e $y\in Lib(t)$.

Gli aspetti grammaticali sono di grande rilevanza didattica e forse sono le uniche parti della Logica indispensabili per contribuire alla precisione del linguaggio. L'addestramento

ripetere i quantificatori, se sono dello stesso nome, e senza interporre parentesi, se non strettamente necessarie per la evitare ambiguità di scrittura.

La definizione di enunciato fa uso di quella di indeterminata libera e vincolata. Talora viene invece utilizzato *enunciato* come sinonimo di *proposizione*. E' solo una scelta linguistica, ma è bene fissare le idee, visto che quando si trattano aspetti grammaticali si considerano solo scritture non interpretate.

[10] Sembra utile aggiungere un altro esempio a chiarificazione del concetto di libero e vincolato. E' necessario per affrontare l'argomento, un "occhio" addestrato che in qualche modo ricorda quello del fisico che vede l'errore in un'eguaglianza tra grandezze fisiche dalla sola analisi delle dimensioni. Sia Antonio il padre di Bruno e Carla. Si descriva questa situazione con un linguaggio adeguato allo scopo. Si può assumere che l'insieme delle costanti individuali C sia dato da {a,b,c} e considerare un solo predicato binario P. L'interpretazione sottintesa di Px,y sia «x è padre di y». Con questi strumenti, la semplice situazione si può rappresentare come Pa,b; Pa,c. Ma questa non esaurisce il potere espressivo del linguaggio scelto. Ad esempio se parlando con una persona che conosce Bruno e Carla, ma non il loro padre, si dicesse che Bruno e Carla sono fratelli, il ricevente comprenderebbe che *esiste* un genitore comune. Ciò che egli conosce non è la congiunzione $Pa,b \wedge Pa,c$, dato che non conosce il nome del padre, ma la formula. $(\exists x(Px,b \wedge Px,c))$. Le due formule esprimono entrambe che Bruno e Carla sono fratelli, ma la prima è più informativa. Tuttavia la seconda esprime correttamente la frase "Bruno e Carla sono fratelli": in essa non compare il nome del loro padre. Così in $(\exists x(Px,b \wedge Px,c))$ compare l'indeterminata x vincolata, quindi il risultato finale non dipende da essa. Sempre col linguaggio di questo esempio, il significato intuitivo dell'enunciato $(\forall x(\exists yPy,x))$ è che ogni individuo (x) ha un padre (y); il significato intuitivo di $(\exists y(\forall xPy,x))$ è che esiste un individuo (y) che è padre di ogni altro (x). La differenza è rimarcata dalla scrittura in cui compaiono i quantificatori in ordine diverso. Ciò che sembra un semplice scambio grafico, ha un significato ben differente. Molti risultati di Matematica vengono letti e interpretati male perché non si presta sufficiente attenzione a scambi analoghi, anche perché spesso si esprimono i quantificatori a parole nei modi più vari. Lo scambio di due quantificatori omonimi (entrambi universali o esistenziali) non provoca questi guai ed è per questo che spesso si scrive un unico quantificatore seguito dai nomi delle indeterminate cui si applica, separate da virgole. E' consigliabile in Matematica un uso il più corretto possibile dei quantificatori per comprendere meglio la struttura delle affermazioni. Basti pensare alla differenza che si fa in algebra, spesso in modo confuso, tra eguaglianza, identità ed equazione.

alla grammatica è una via per favorire il raggiungimento di questo obiettivo, visti anche gli stretti legami tra calcolo letterale, linguaggi di programmazione e grammatica. Spesso la comprensione viene ostacolata dall'incapacità di seguire i passaggi se questi richiedono la modifica delle scritture con vari tipi di sostituzioni. E d'altra parte le sostituzioni sono presenti ed utilizzate come strumento conoscitivo, sia in Matematica, sia in altre discipline (cfr. $[M]_1$ e $[M]_2$).

Una volta si asseriva che con Matematica e Latino si impara a ragionare. Il cambiamento avvenuto nella didattica con l'abbandono sostanziale della geometria sintetica ha dato origine ad un fenomeno di analfabetizzazione logica di ritorno, cfr. [L]. Le aspettative che questo ruolo sia svolto dalla Logica sono destinate a rimanere deluse, almeno se si intende il "ragionare" come la capacità di trovare l'assassino in un libro giallo o compiere la scelta del coniuge. In un qualche senso la Logica è un *a posteriori* (lo diceva Kant), analizza quanto viene fatto. Quindi la Logica non è una "macchina" che dimostra i teoremi di matematica come forse alcuni sperano, senza conoscere profondamente le proprietà degli enti di cui si parla.

E' interessante riguardare i propri testi di Matematica e vedere se sono serviti, con gli esempi, ad insegnare regole corrette, utilizzabili per condurre dimostrazioni [11]. Questa metodologia induttiva permette di scoprire la potenza della logica come strumento utile nelle altre discipline. Cambiando punto di vista, si può dare alla Logica il ruolo di un oggetto di studio in sé. In tale approccio, le regole vanno presentate prima di procedere. Solitamente ciò non avviene, forse c'è paura che mettendo allo scoperto *tutte* le regole indispensabili per la concatenazione dei passaggi che costituiscono le dimostrazioni matematiche, esse (regole) siano in numero così elevato che la loro elencazione porti via troppo tempo [12]. Ciò era vero prima di questo secolo; oggi, grazie a risultati che verranno illustrati in altre lezioni, si possono dissipare tali dubbi.

La presentazione delle regole può essere fatta secondo finalità ben diverse. Si può privilegiare l'aspetto "economico" (alla Hilbert) mostrando che bastano poche regole. In alternativa si può cercare di catturare attraverso le regole la "validità" di certi ragionamenti corretti, che fanno passare dalla verità alla verità (tableaux), o ancora cercare un compromesso (deduzione naturale) che fornisca un sistema agile, non troppo complesso né compresso, in grado di cogliere i ragionamenti che appaiono normalmente sui testi. Questi tre approcci sono qui trattati separatamente. In ciascuno dei sistemi si identificano due parti: un insieme (anche

[11] Solo le regole da seguire tra i vari passaggi di cui sono costituite le dimostrazioni, non a trovare le idee che indicano quali sono i passaggi da fare! Questo è l'aspetto della Logica come strumento, ma la nostra materia è, di per sé, oggetto d'indagine. Alcuni esempi di quanto suggerito vengono indicate negli esercizi.

[12] O forse sarebbero strumento pericoloso in mano agli studenti che potrebbero capire anche gli errori dei docenti!

vuoto) di assiomi detti *logici* ed un insieme di regole di inferenza, cioè di relazioni specificamente assegnate. Prescelto un sistema di derivazione si ha un procedimento che fa passare da (insiemi di) formule a formule. Se le regole vengono rispettate il risultato, una *deduzione*, è un insieme finito di formule, ordinato secondo un ordine lineare oppure no. C'è il problema della correttezza ed adeguatezza dei sistemi. La mentalità corretta per affrontare il tema è quella del gioco: le regole della briscola sono tanto "vere" quanto quelle del tresette. Le varie regole che si presentano in Logica hanno avuto origine dalla ricerca di come determinare un insieme di regole, sufficientemente piccolo, ma sufficiente per le dimostrazioni matematiche e tale che non ci fossero dimostrazioni che non si potessero adeguatamente "tradurre" in esso. Quindi le regole non sono altrettanto "gratuite" di quelle della briscola, ma l'esempio è qui presentato per mettere in evidenza l'aspetto formale delle regole stesse.

Indipendentemente dalle regole prescelte, dato un insieme Γ di formule, sia $C(\Gamma)$ l'insieme delle formule che si possono ottenere, applicando le regole e gli assiomi, a partire da Γ (la chiusura deduttiva di Γ). Le scritture $\varphi \in C(\Gamma)$ e $\Gamma \vdash \varphi$, che si legge φ *è deducibile da* Γ oppure *da* Γ *si deduce* φ, sono equivalenti. In particolare quando $\Gamma = \varnothing$, si scrive $\vdash \varphi$ ed in tal caso φ viene detta *un teorema* (logico). Ovviamente: $\Gamma \subseteq C(\Gamma)$, in quanto ogni elemento di Γ si può pensare ottenuto immediatamente da se stesso. Se poi $\Gamma \subseteq \Delta$, allora $C(\Gamma) \subseteq C(\Delta)$, proprietà di monotonia valida per i sistemi che si considerano qui, non in generale. Poi $C(C(\Gamma)) \subseteq C(\Gamma)$; questa affermazione è spesso richiamata quando in una dimostrazione si utilizza la dizione: *per un teorema precedente*. Infatti essa richiede che se qualche formula φ è ottenibile a partire dalla chiusura deduttiva di Γ, visto che ciascuna formula di $C(\Gamma)$, a sua volta, è ottenibile da Γ con le stesse clausole, allora anche φ è ottenibile da Γ, solo a patto di "allungare" il procedimento deduttivo. Una importante proprietà dei sistemi deduttivi successivi è la cosiddetta *compattezza*. Essa si esprime dicendo che se $\varphi \in C(\Gamma)$, esiste $\Delta \subseteq \Gamma$, con Δ finito, tale che $\varphi \in C(\Delta)$. Queste proprietà, lo si ribadisce, sono dimostrabili in ciascuno dei tre sistemi presentati nel seguito. La prova viene lasciata al lettore.

I sistemi alla Hilbert

Con tale dicitura si indicano diverse proposte. In esse si privilegiano i connettivi di implicazione e negazione ed il quantificatore universale [13]. Ciò ha origine dal trattare l'implicazione come l'analogo formale del procedimento deduttivo: invece di leggere $\varphi \to \psi$ come "φ implica ψ", lo si pensa come "da φ segue ψ"; si potrebbero ricercare più profonde connessioni con l'idea di causalità.

[13] Si può anche fare a meno della negazione, introducendo però un simbolo detto di *falsità*. Gli altri connettivi vengono definiti a partire da questi due (che costituiscono un sistema *adeguato* di connettivi) e pure il quantificatore esistenziale si riottiene a partire dall'universale con la negazione.

Le definizioni per gli altri connettivi ed il quantificatore esistenziale sono le seguenti

$$(\varphi \wedge \psi) \text{ per } (\neg(\varphi \to (\neg\psi))); \qquad (\varphi \vee \psi) \text{ per } ((\neg\varphi) \to \psi);$$
$$(\varphi \leftrightarrow \psi) \text{ per } ((\varphi \to \psi) \wedge (\psi \to \varphi)); \qquad (\exists x \varphi(x)) \text{ per } (\neg(\forall x(\neg\varphi(x))).$$

Un sistema di questo tipo, sostanzialmente tratto da [Me], ha per assiomi:

A1 $\varphi \to (\psi \to \varphi)$;

A2 $(\varphi \to (\psi \to \vartheta)) \to ((\varphi \to \psi) \to (\varphi \to \vartheta))$;

A3 $(\neg \varphi \to \neg \psi) \to ((\neg \varphi \to \psi) \to \varphi)$;

A4 $(\forall x \varphi(x)) \to (x/t)\varphi(x)$, ove t è un termine libero per x in $\varphi(x)$;

A5 $(\forall x(\varphi \to \psi)) \to (\varphi \to (\forall x \psi))$, purché $x \notin \mathrm{Lib}(\varphi)$;

A6 $\forall x(x \doteq x)$

A7 $\forall x \forall y(x \doteq y \to (\varphi(x,x) \to \varphi(x,y)))$, ove x,y sono indeterminate e $\varphi(x,y)$ si ottiene da $\varphi(x,x)$ per sostituzione di x con y [14].

Le regole di inferenza sono due: si tratta di due relazioni sull'insieme delle formule, una detta, *modus ponens* o *regola del taglio*, è una relazione: $\mathsf{MP} = \{\langle\langle\varphi,\varphi\to\psi\rangle,\psi\rangle \mid \varphi,\psi \in \mathbf{L}_A(X)\}$, l'altra, detta *generalizzazione*, è una relazione binaria: $\mathsf{Gen} = \{\langle\varphi,(\forall x\varphi)\rangle \mid \varphi \in \mathbb{L}_A(X),\ x \in X\}$. Spesso le regole (o relazioni), si presentano in scrittura di frazione:

$$\frac{\varphi,\varphi\to\psi}{\psi}\,\mathsf{MP} \qquad \frac{\varphi}{\forall x(\varphi)}\,\mathsf{Gen}.$$

Il "numeratore" viene detto *premessa* (*premesse* nel caso di MP), il "denominatore", *conclusione*. Una rapida verifica convince della correttezza delle regole e della verità degli assiomi.

Per comprendere lo strumento messo a disposizione, si mostrano alcune deduzioni, invitando il lettore a seguirne i passaggi e le rispettive giustificazioni.

Prese comunque le formule φ, ψ e ϑ, si ha

1) $\vdash \varphi \to \varphi$; **2)** $\varphi \to \psi, \vartheta \to \varphi \vdash \vartheta \to \psi$ [15]; **3)** $\varphi \to (\psi \to \vartheta), \psi \vdash \varphi \to \vartheta$; **4)** $\vdash \neg\neg \varphi \to \varphi$; **5)** $\vdash \varphi \to \neg\neg\varphi$; **6)** $\neg\varphi \to \neg\psi \vdash \psi \to \varphi$; **7)** $\psi \to \varphi \vdash \neg\varphi \to \neg\psi$; **8)** $\neg(\varphi \to \psi) \vdash (\varphi \wedge \neg\psi)$; **9)** $(\varphi \wedge \neg\psi) \vdash \neg(\varphi \to \psi)$; **10)** $(\exists x(\varphi(x) \wedge \neg\psi(x)) \vdash \neg(\forall x(\varphi(x) \to \psi(x))$; **11)** $\neg(\forall x(\varphi(x) \to \psi(x)) \vdash (\exists x(\varphi(x) \wedge \neg\psi(x))$; **12)** $(\forall x(\varphi(x) \to \psi(x))), (\forall x(\vartheta(x) \to \varphi(x))) \vdash (\forall x(\vartheta(x) \to \psi(x)))$; **13)** $(\forall x(\varphi(x) \to \psi(x))), (\exists x(\vartheta(x) \wedge \neg\psi(x)) \vdash (\exists x(\vartheta(x) \wedge \neg\varphi(x)))$

[14] Bisognerebbe esser più accurati dicendo che la sostituzione può essere parziale e che y quando viene sostituita a x non cada nel rango d'azione di un quantificatore Qy.

[15] E' tradizione scrivere, nel caso di insiemi finiti $\varphi,\psi \vdash \vartheta$ invece che $\{\varphi,\psi\} \vdash \vartheta$ e $\Gamma,\varphi \vdash \psi$ invece che $\Gamma \cup \{\varphi\} \vdash \psi$.

1)

1. $(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$ A2
2. $\varphi \to ((\varphi \to \varphi) \to \varphi)$ A1
3. $(\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$ MP, 2,1
4. $\varphi \to (\varphi \to \varphi)$ A1
5. $\varphi \to \varphi$ MP 4,5

2)

1. $\varphi \to \psi$ Hp. [16]
2. $(\varphi \to \psi) \to (\vartheta \to (\varphi \to \psi))$ A1
3. $\vartheta \to (\varphi \to \psi)$ MP 1,2
4. $(\vartheta \to (\varphi \to \psi)) \to ((\vartheta \to \varphi) \to (\vartheta \to \psi))$ A2
5. $(\vartheta \to \varphi) \to (\vartheta \to \psi)$ MP 3,4
6. $\vartheta \to \varphi$ Hp,
7. $\vartheta \to \psi$ MP 6,5.

3)

1. $\varphi \to (\psi \to \vartheta)$ Hp
2. ψ Hp
3. $\psi \to (\varphi \to \psi)$ A1
4. $\varphi \to \psi$ MP 2, 3
5. $(\varphi \to (\psi \to \vartheta)) \to ((\varphi \to \psi) \to (\varphi \to \vartheta))$ A2
6. $(\varphi \to \psi) \to (\varphi \to \vartheta)$ MP 1,5
7. $\varphi \to \vartheta$ MP 4,6

4)

1. $(\neg\varphi \to \neg\neg\varphi) \to ((\neg\varphi \to \neg\varphi) \to \varphi)$ A3
2. $\neg\varphi \to \neg\varphi$ Teor prec 1)
3. $(\neg\varphi \to \neg\neg\varphi) \to \varphi$ Teor prec 3), 1, 2
4. $\neg\neg\varphi \to (\neg\varphi \to \neg\neg\varphi)$ A1
5. $\neg\neg\varphi \to \varphi$ Teor prec 2), 4, 3

5)

1. $\neg\neg\neg\varphi \to \neg\varphi$ Teor prec 4)
2. $(\neg\neg\neg\varphi \to \neg\varphi) \to ((\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi)$ A3
3. $(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi$ MP 1,2
4. $\varphi \to (\neg\neg\neg\varphi \to \varphi)$ A1
5. $\varphi \to \neg\neg\varphi$ Teor prec 2) 4, 5

6)

1. $\neg\varphi \to \neg\psi$ Hp
2. $(\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$ A3
3. $(\neg\varphi \to \psi) \to \varphi$ MP 1,2
4. $\psi \to (\neg\varphi \to \psi)$ A1
5. $\psi \to \varphi$ Teor prec 2) 4, 3

7)

1. $\psi \to \varphi$ Hp
2. $\neg\neg\psi \to \psi$ Teor prec 4)
3. $\neg\neg\psi \to \varphi$ Teor prec 2) 2, 1
4. $\varphi \to \neg\neg\varphi$ Teor prec 5)
5. $\neg\neg\psi \to \neg\neg\varphi$ Teor prec 2) 3, 4
6. $(\neg\neg\psi \to \neg\neg\varphi) \to (\neg\varphi \to \neg\psi)$ Teor prec 6) 5
7. $\neg\varphi \to \neg\psi$ MP 5, 6

8)

Per la definizione della congiunzione si prova $\neg(\varphi \to \psi) \vdash \neg(\varphi \to \neg\neg\psi)$.

1. $\neg(\varphi \to \psi)$ Hp
2. $\neg\neg\psi \to \psi$ Teor prec 4)
3. $(\neg\neg\psi \to \psi) \to (\varphi \to (\neg\neg\psi \to \psi))$ A1
4. $\varphi \to (\neg\neg\psi \to \psi)$ MP 2,3
5. $(\varphi \to (\neg\neg\psi \to \psi)) \to ((\varphi \to \neg\neg\psi) \to (\varphi \to \psi))$ A2

[16] Con tale scritta si indica solo che si tratta di un elemento dell'insieme messo a sinistra del simbolo di deduzione.

6. $(\varphi \to \neg\neg\psi) \to (\varphi \to \psi)$ MP 4,5
7. $\neg(\varphi \to \psi) \to \neg(\varphi \to \neg\neg\psi)$ Teor prec 7) 6
8. $\neg(\varphi \to \neg\neg\psi)$ MP 1,7

9)

Come prima si fa prima la traduzione: si deve provare $\neg(\varphi \to \neg\neg\psi) \vdash \neg(\varphi \to \psi)$

1. $\neg(\varphi \to \neg\neg\psi)$ Hp
2. $\psi \to \neg\neg\psi$ Teor prec 5)
3. $(\psi \to \neg\neg\psi) \to (\varphi \to (\psi \to \neg\neg\psi))$ A1
4. $\varphi \to (\psi \to \neg\neg\psi)$ MP 2,3
5. $(\varphi \to (\psi \to \neg\neg\psi)) \to ((\varphi \to \psi) \to (\varphi \to \neg\neg\psi))$ A3
6. $(\varphi \to \psi) \to (\varphi \to \neg\neg\psi)$ MP 4,5
7. $\neg(\varphi \to \neg\neg\psi) \to \neg(\varphi \to \psi)$ Teor prec 7) 6
8. $\neg(\varphi \to \psi)$ MP 1,7

10)

Per le definizioni $\exists$ e $\wedge$, si deve provare $\neg(\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x))) \vdash \neg(\forall x(\varphi(x) \to \psi(x)))$

1. $\psi(x) \to \neg\neg\psi(x)$ Teor prec 5)
2. $(\psi(x) \to \neg\neg\psi(x)) \to ((\varphi(x) \to (\psi(x) \to \neg\neg\psi(x)))$ A1
3. $(\varphi(x) \to (\psi(x) \to \neg\neg\psi(x))$ MP 1,2
4. $(\varphi(x) \to (\psi(x) \to \neg\neg\psi(x))) \to ((\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x)))$ A2
5. $(\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x))$ MP 3,4
6. $((\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x))) \to ((\forall x(\varphi(x) \to \psi(x))) \to ((\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x))))$ A1
7. $(\forall x(\varphi(x) \to \psi(x))) \to ((\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x)))$ MP 5,6
8. $((\forall x(\varphi(x) \to \psi(x))) \to ((\varphi(x) \to \psi(x)) \to (\varphi(x) \to \neg\neg\psi(x)))) \to (((\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \psi(x))) \to ((\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \neg\neg\psi(x)))$ A2
9. $((\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \psi(x))) \to ((\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \neg\neg\psi(x))$ MP 7,8
10. $(\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \psi(x))$ A4 con la sostituzione (x/x)
11. $(\forall x(\varphi(x) \to \psi(x))) \to (\varphi(x) \to \neg\neg\psi(x))$ MP 10,9
12. $\neg(\varphi(x) \to \neg\neg\psi(x)) \to \neg(\forall x(\varphi(x) \to \psi(x)))$ Teor prec 7) 11
13. $\neg\neg(\forall x(\varphi(x) \to \psi(x))) \to \neg\neg(\varphi(x) \to \neg\neg\psi(x))$ Teor prec 7) 12
14. $(\forall x(\varphi(x) \to \psi(x))) \to \neg\neg(\forall x(\varphi(x) \to \psi(x)))$ Teor prec 5)
15. $(\forall x(\varphi(x) \to \psi(x))) \to \neg\neg(\varphi(x) \to \neg\neg\psi(x))$ Teor prec 2) 14, 13
16. $(\forall x(\forall x(\varphi(x) \to \psi(x))) \to \neg\neg(\varphi(x) \to \neg\neg\psi(x))) \to ((\forall x(\varphi(x) \to \psi(x))) \to (\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x))))$ A5, dato che $x \notin Lib((\forall x(\varphi(x) \to \psi(x))))$
17. $(\forall x(\varphi(x) \to \psi(x))) \to (\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x)))$ MP 11, 12
18. $\neg(\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x))) \to \neg(\forall x(\varphi(x) \to \psi(x)))$ Teor prec 7) 17
19. $\neg(\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x)))$ Hp
20. $\neg(\forall x(\varphi(x) \to \psi(x)))$ MP 19,18

11)

Per definizione di $\exists$ e $\wedge$ si prova $\neg(\forall x(\varphi(x) \to \psi(x))) \vdash \neg(\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x)))$

1. $\neg(\forall x(\varphi(x) \to \psi(x)))$ Hp
2. $(\forall x\neg\neg(\varphi(x) \to \neg\neg\psi(x))) \to \neg\neg(\varphi(x) \to \neg\neg\psi(x))$ A4 con la sostituzione (x/x)
3. $\neg\neg(\varphi(x) \to \neg\neg\psi(x)) \to (\varphi(x) \to \neg\neg\psi(x))$ Teor prec 4)
4. $\neg\neg\psi(x) \to \psi(x)$ Teor prec 4)

5. $(\neg\neg \psi(x) \rightarrow \psi(x)) \rightarrow (\varphi(x) \rightarrow (\neg\neg \psi(x) \rightarrow \psi(x)))$ A1 6. $\varphi(x) \rightarrow (\neg\neg \psi(x) \rightarrow \psi(x))$ MP 4,5

7. $(\varphi(x) \rightarrow (\neg\neg \psi(x) \rightarrow \psi(x))) \rightarrow ((\varphi(x) \rightarrow \neg\neg \psi(x)) \rightarrow (\varphi(x) \rightarrow \psi(x)))$ A2

8. $(\varphi(x) \rightarrow \neg\neg \psi(x)) \rightarrow (\varphi(x) \rightarrow \psi(x))$ MP 6,7

9. $(\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))) \rightarrow (\varphi(x) \rightarrow \neg\neg\psi(x))$ Teor prec 2) 3, 2

10. $(\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))) \rightarrow (\varphi(x) \rightarrow \psi(x))$ Teor prec 2) 8,9

11. $(\forall x((\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))) \rightarrow (\varphi(x) \rightarrow \psi(x)))) \rightarrow ((\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))) \rightarrow (\forall x(\varphi(x) \rightarrow \psi(x))))$ A5 dato che $x \notin Lib((\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))))$

12. $(\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x))) \rightarrow (\forall x(\varphi(x) \rightarrow \psi(x)))$ MP 10, 11

13. $\neg(\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow \neg(\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x)))$ Teor prec 7), 12

14. $\neg(\forall x \neg\neg(\varphi(x) \rightarrow \neg\neg\psi(x)))$ MP 1,13

12)

1. $\forall x(\varphi(x) \rightarrow \psi(x))$ Hp 2. $\forall x(\vartheta(x) \rightarrow \varphi(x))$ Hp

Sia y una indeterminata non presente nelle formule 1. e 2.

3. $\forall x(\varphi(x) \rightarrow \psi(x)) \rightarrow (y/x)(\varphi(x) \rightarrow \psi(x))$ A4 4. $\varphi(y) \rightarrow \psi(y)$ MP 1,3 5. $\forall x(\vartheta(x) \rightarrow \psi(x))$ Gen 4.

Questa dimostrazione è una giustificazione del sillogismo universale *Barbara.*

13)

1. $(\forall x(\varphi(x) \rightarrow \psi(x)))$ Hp

2. $(\exists x(\vartheta(x) \wedge \neg\psi(x))$ Hp 3. $(\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow (\varphi(x) \rightarrow \psi(x))$ A4

4. $(\forall x(\vartheta(x) \rightarrow \varphi(x))) \rightarrow (\vartheta(x) \rightarrow \varphi(x))$ A4 5. $\varphi(x) \rightarrow \psi(x)$ MP 1,3

6. $(\varphi(x) \rightarrow \psi(x)) \rightarrow (\vartheta(x) \rightarrow (\varphi(x) \rightarrow \psi(x)))$ A1 7. $\vartheta(x) \rightarrow (\varphi(x) \rightarrow \psi(x)))$ MP 5,6

8. $(\vartheta(x) \rightarrow (\varphi(x) \rightarrow \psi(x))) \rightarrow ((\vartheta(x) \rightarrow \varphi(x)) \rightarrow (\vartheta(x) \rightarrow \psi(x)))$ A2

9. $(\vartheta(x) \rightarrow \varphi(x)) \rightarrow (\vartheta(x) \rightarrow \psi(x))$ MP 7,8

10. $(\forall x(\vartheta(x) \rightarrow \varphi(x))) \rightarrow (\vartheta(x) \rightarrow \psi(x))$ Teor prec 2) 9,

11. $(\forall x((\forall x(\vartheta(x) \rightarrow \varphi(x))) \rightarrow (\vartheta(x) \rightarrow \psi(x)))) \rightarrow ((\forall x(\vartheta(x) \rightarrow \varphi(x))) \rightarrow (\forall x(\vartheta(x) \rightarrow \psi(x))))$ A4 dato che $x \notin Lib((\forall x(\vartheta(x) \rightarrow \varphi(x))))$

12. $(\forall x(\vartheta(x) \rightarrow \varphi(x))) \rightarrow (\forall x(\vartheta(x) \rightarrow \psi(x)))$ MP 10, 11

13. $\neg(\forall x(\vartheta(x) \rightarrow \psi(x))) \rightarrow \neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))$ Teor prec 7, 12

14. $\neg(\forall x(\vartheta(x) \rightarrow \psi(x)))$ Teor prec 10) 2 15. $\neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))$ MP 14, 13

16. $\exists x(\vartheta(x) \wedge \neg\varphi(x))$ Teor prec 11), 15

Questa deduzione può essere vista come la giustificazione di un sillogismo di tipo *Baroco*.

L'esperienza e le ragioni tecniche suggeriscono la scelta degli assiomi e delle regole di inferenza: è una scelta minimale dei tipi di assiomi, dai quali, mediante le regole di inferenza, si ricavano tutte le altre formule valide. Le regole sono comunque tutte qui, meglio, gli schemi, perché le formule che si ottengono sostituendo formule negli schemi in luogo di φ,ψ, ecc.,

fatte salve le limitazioni indicate, sono istanze di assiomi logici o di regole di inferenza. Un ottimo esercizio è mostrare l'esigenza delle limitazioni imposte nelle regole di inferenza.

Da questo esempio di sistema si traggono alcune considerazioni valide anche per gli altri sistemi. Si possono provare le proprietà preannunciate della chiusura deduttiva, in particolare la compattezza: il procedimento deduttivo è finitistico e le ipotesi che servono in una deduzione sono in numero finito. L'ultima formula di una deduzione è detta *deducibile* dall'insieme delle ipotesi. Quando questo è vuoto si dice *teorema logico*. Di una stessa formula si possono avere più deduzioni, partendo sempre dalle stesse ipotesi.

I tableaux (o tavole di confutazione o alberi di confutazione).

Prima di descrivere nei dettagli (abbastanza "pesanti") il sistema, si consideri una generica formula. Si è già visto (e si vedrà in seguito) il concetto di verità della formula. Esso è definito per ricursione sulla costruzione della formula, analizzando le sottoformule di quella data. I tableaux ricostruiscono in modo sintattico i passaggi che vengono compiuti dal punto di vista semantico, con l'analisi del significato delle sottoformule, fino ai costituenti "minimi" della formula stessa. Tra le numerose varianti dei tableaux qui se ne considera una [17] basata su [BM], che ha come primitivi i connettivi di implicazione e negazione e il quantificatore universale, definendo poi gli altri connettivi e il quantificatore esistenziale.

DEFINIZIONE Un *tableau* è costituito da una terna ordinata $\langle G,l,\Phi\rangle$ in cui G, il *grafo* del tableau, è un insieme parzialmente ordinato dotato di minimo m (il *nodo iniziale.*), gli elementi di G sono detti *nodi; l* è un'applicazione del grafo nell'insieme dei numeri naturali, che conserva l'ordine e associa al nodo iniziale il numero 0; se k è un nodo per cui $l(k) = n+1$, allora esiste un unico nodo h tale che $l(h) = n$ e h precede immediatamente k nell'ordine di G. Inoltre Φ è una applicazione del grafo nell'insieme dei sottoinsiemi finiti di $L_A(X)$. Per ogni nodo k, il numero $l(k)$ è detto *livello di* k. Gli elementi massimali del grafo vengono detti*n odi terminali.*

DEFINIZIONE Dato un tableau $\langle G,l,\Phi\rangle$, se esiste un nodo di livello p e non esistono nodi di livello p+1, il numero naturale p viene detto *profondità* del tableau.

Dato un nodo k, il *cammino* di k [18] è una successione $k_0, k_1, \ldots, k_r$ tale che $k_0 = m$, $k_r = k$ e per ogni s, con $0 < s \leq r$, k_s è successore immediato di k_{s-1}. Se k è un nodo terminale, il cammino si dice il *ramo* del tableau, *che termina in* k.

Se h è un nodo del cammino (ramo) di k e $\varphi \in \Phi(h)$, la formula φ si dice *una formula del cammino di* k

[17] Una seconda versione, le cosiddette *tavole semantiche*, verrà illustrata nel contributo di F. Montagna.
[18] Per le condizioni poste sul tableau è univocamente individuato.

Nella pratica vengono confusi i nodi con gli insiemi di formule associati ad essi dalla applicazione Φ. In Logica si usano i tableaux in cui la funzione Φ verifica le seguenti condizioni:

DEFINIZIONE Un *tableau del primo ordine* è un tableau $\langle G,l,\Phi\rangle$ se k è successore immediato di h, $\Phi(k)$ si ottiene in base ad una delle seguenti regole:

(R¬¬) $\Phi(k) = \{\varphi\}$ se $(\neg\neg\varphi)$ è una formula del cammino di h;

(R ¬→) $\Phi(k) = \{\varphi,\neg\psi\}$ se $(\neg(\varphi\rightarrow\psi))$ è una formula del cammino di h;

(R →) $\Phi(k) = \{\neg\varphi\}$ oppure $\Phi(k) = \{\psi\}$ se $(\varphi\rightarrow\psi)$ è una formula del cammino di h.

(R ∀) $\Phi(k) = \{\varphi(x/t)\}$, ove $t\in T_A(X)$, se $(\forall x(\varphi(x)))\in\Phi(h_s)$ è una formula del cammino di h;

(R ¬∀) $\Phi(k) = \{\neg\varphi(x/y)\}$, ove $y\in X$, se $(\neg\forall x(\varphi(x)))$ è una formula del cammino di h e y è un'indeterminata che non compare libera in nessuna delle formule del cammino di h.

(R SI) $\Phi(k) = \{t \doteq t\}$, per ogni $t\in T_A(X)$.

(R SF) $\Phi(k) = \{t_1 \doteq t_{n+1}\wedge \ldots\wedge t_n \doteq t_{2n} \rightarrow f(t_1,\ldots,t_n) \doteq f(t_{n+1},\ldots t_{2n})\}$, comunque presi $t_1,\ldots,t_n,t_{n+1},\ldots t_{2n}\in T_A(X)$ e $f\in F_n$.

(R SP) $\Phi(k) = \{t_1 \doteq t_{n+1}\wedge \ldots\wedge t_n \doteq t_{2n} \rightarrow (Rt_1,\ldots,t_n \rightarrow Rt_{n+1},\ldots t_{2n})\}$, comunque presi $t_1,\ldots,t_n,t_{n+1},\ldots t_{2n}\in T_A(X)$ e ogni predicato R [19].

Nelle regole sull'eguaglianza non ci sono condizioni sul cammino che precede il nodo k, in quanto queste si possono considerare alla stregua degli assiomi di un sistema di Hilbert. Il sistema si può arricchire di regole che trattino gli altri connettivi (e quantificatori), in questo modo la presentazione delle regole si allunga, ma i cammini si accorciano.

Se al nodo iniziale corrisponde un insieme finito di formule, la profondità di un tableau è comunque un numero naturale, anche se il grafo G può avere infiniti nodi. La presentazione scelta ha un difetto: per renderla corretta e breve, viene a mancare l'aspetto dinamico che è così interessante nei tableaux. Dato un insieme "iniziale" di formule, si costruisce il tableau, passo a passo. La costruzione termina se ai nodi terminali corrispondono formule atomiche o negazioni di formule atomiche. Più interessante è la nozione di tableau chiuso:

[19] In questo caso ricade anche l'eguaglianza: $\Phi(k) = \{t_1 \doteq t_3 \wedge t_2 \doteq t_4 \rightarrow (t_1 \doteq t_2 \rightarrow t_3 \doteq t_4))\}$

DEFINIZIONE Un ramo di un tableau si dice *chiuso* se esiste una formula φ tale che tanto φ quanto $\neg\varphi$ siano formule del ramo. Un tableau si dice *chiuso* se ogni suo ramo è chiuso. Se il tableau è chiuso si dice anche che è una *confutazione* di $\Phi(m)$. Se $\Phi(m) = \{\varphi\}$, allora $\vdash\neg\varphi$. Per provare mediante i tableau $\varphi,\vartheta\vdash\psi$, si assume il nodo iniziale dato da $\{\varphi,\vartheta,\neg\psi\}$

In pratica non si giunge alle formule atomiche o alla negazione di formule atomiche, basta ottenere un tableau chiuso. Nella trattazione dei tableaux spesso intervengono indeterminate e termini che a priori possono essere scelti arbitrariamente, in modo opportuno. Il grande merito dei tableaux è quello di fornire, almeno nel caso della struttura proposizionale, un procedimento algoritmico per individuare una dimostrazione sintattica (per assurdo) di una formula, lavorando solo sulle sue sottoformule. Ciò non avviene per i sistemi alla Hilbert, come attestano gli esempi precedenti. Quando intervengono i quantificatori le cose si fanno più delicate, comunque sono strumenti che forniscono molte informazioni.

E.g. si provi coi tableaux una formula analoga a quella vista sopra, la traduzione del sillogismo di tipo *Baroco*: $(\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow ((\exists x(\vartheta(x) \wedge \neg\psi(x))\rightarrow(\exists x(\vartheta(x) \wedge \neg\varphi(x))))$. Poiché si utilizza solo la negazione, l'implicazione ed il quantificatore universale, si prova $(\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow (\neg(\forall x(\vartheta(x) \rightarrow \psi(x)) \rightarrow \neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))))$. Si confuta la negazione di tale formula. La dimostrazione (confutazione) procede così

$\{\neg((\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow (\neg(\forall x(\vartheta(x) \rightarrow \psi(x))) \rightarrow\neg(\forall x(\vartheta(x)\rightarrow \varphi(x)))))\}$

$\{(\forall x(\varphi(x)\rightarrow\psi(x))),\neg(\neg(\forall x(\vartheta(x)\rightarrow\psi(x)))\rightarrow\neg(\forall x(\vartheta(x)\rightarrow\varphi(x))))\}$ **(R¬ →)**

$\{\neg(\forall x(\vartheta(x) \rightarrow \psi(x)),\neg\neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))\}$ **(R ¬ →)**

$\{\neg(\vartheta(y) \rightarrow \psi(y))\}$ **(R ¬ ∀)** purché y non sia presente precedentemente, sostituzione (x/y)

$\{\vartheta(y),\neg\psi(y)\}$ **(R ¬ →)**

$\{\varphi(y) \rightarrow \psi(y)\}$ **(R ∀)** sostituzione (x/y)

$\{(\forall x(\vartheta(x) \rightarrow \varphi(x)))\}$ **(R ¬ ¬)**

$\{\vartheta(y) \rightarrow \varphi(y)\}$ **(R ∀)** sostituzione (x/y)

$\{\neg\varphi(y)\}$ $\{\psi(y)\}$ **(R →)**

$\{\neg\vartheta(y)\}$ $\{\varphi(y)\}$ **(R →)**

I nodi terminali sono individuati dagli insiemi di formule $\{\neg\vartheta(y)\}$, $\{\varphi(y)\}$ e $\{\psi(y)\}$. Il nodo iniziale è dato dal singoletto della formula da confutare. Un cammino è dato da $\{\underline{\neg\vartheta(y)}\}$, $\{\neg\varphi(y)\}$, $\{\vartheta(y) \rightarrow \varphi(y)\}$, $\{(\forall x(\vartheta(x) \rightarrow \varphi(x)))\}$, $\{\varphi(y) \rightarrow \psi(y)\}$, $\{\underline{\vartheta(y),\neg\psi(y)}\}$, $\{\neg(\vartheta(y) \rightarrow \psi(y))\}$, $\{\neg(\forall x(\vartheta(x) \rightarrow \psi(x))), \neg\neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))\}$, $\{(\forall x(\varphi(x) \rightarrow \psi(x))),\neg(\neg(\forall x(\vartheta(x)\rightarrow\psi(x)))\rightarrow\neg(\forall x(\vartheta(x) \rightarrow \varphi(x))))\}$, $\{\neg((\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow (\neg(\forall x(\vartheta(x) \rightarrow \psi(x))) \rightarrow \neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))))\}$. E' un ramo; è chiuso perché compaiono $\neg\vartheta(y)$ e $\vartheta(y)$, come negli insiemi sottolineati. Poiché ogni ramo è chiuso, la formula $\neg((\forall x(\varphi(x) \rightarrow \psi(x)))$

$\rightarrow (\neg(\forall x(\vartheta(x) \rightarrow \psi(x))) \rightarrow \neg(\forall x(\vartheta(x) \rightarrow \varphi(x)))))$ è confutata, dunque è dedotta (anzi provata, trattandosi di un teorema logico) la formula non negata $(\forall x(\varphi(x) \rightarrow \psi(x))) \rightarrow (\neg(\forall x(\vartheta(x) \rightarrow \psi(x))) \rightarrow \neg(\forall x(\vartheta(x) \rightarrow \varphi(x))))$. Si confronti questa confutazione, che sta in mezza pagina, con le tre pagine e mezza necessarie nel sistema alla Hilbert, sistema la cui presentazione occupa una mezza pagina, contro le tre pagine necessarie per definire i tableaux. Come si diceva, la giustificazione delle regole che forniscono i tableaux è data pensando alla valutazione semantica: basta leggere le condizioni come "se ... è vera, allora è vera (oppure sono vere) ....". Nel caso degli insiemi con due formule la virgola va letta come un *et*; nel caso di due insiemi si può leggere come *vel.* Le formule di un cammino vanno viste congiunte da *et.* In questo senso la chiusura di un ramo significa che si considera una formula e la sua negazione, cioè dovrebbe essere vera una formula e la sua negazione, un assurdo, appunto. La chiusura di tutti i rami porta alla confutazione, cioè avere supposto vera la negazione della formula da provare, porta a contraddizioni. E' interessante notare che se un tableau non si chiude, la negazione della formula del nodo iniziale non è provata, anzi il tableau suggerisce la costruzione di un contromodello. Ciò si vede meglio con le tavole semantiche che verranno successivamente illustrate.

Deduzione naturale

Sono possibili varianti più o meno complesse. La deduzione naturale è, in un certo senso, duale al sistema dei tableaux: con essa si cerca una dimostrazione in forma di albero, coi tableaux si cerca una refutazione mediante un albero dall'alto al basso (*top-down*).

Qui si mostra un sistema assai vicino a quello di Prawitz [P], che ha origine dalla formalizzazione del procedere dimostrativo sulla base della sola Logica. Per ciascuno dei connettivi di congiunzione, disgiunzione ed implicazione e i quantificatori sono date regole di introduzione e di eliminazione. La negazione $(\neg\varphi)$ viene definita da $(\varphi\rightarrow\bot)$ ove $\bot$ è il simbolo di *falsità*. Le rispettive regole di introduzione ed eliminazione della negazione sono casi particolari delle regole analoghe dell'implicazione. La presentazione seguita permette di differenziare il tipo di logica usata (minimale, intuizionistica, classica) e come nei tableaux e si opera per sottoformule, stavolta *bottom-up.* In una deduzione intervengono spesso ipotesi ed esse condizionano la deduzione, come verrà precisato in seguito.

DEFINIZIONE La parte *minimale* della sintassi logica è costituita dalle seguenti regole:

$$(i.\wedge)\ \frac{\varphi\quad\psi}{\varphi\wedge\psi} \qquad (e.\wedge)\ \frac{\varphi\wedge\psi}{\varphi},\ \frac{\varphi\wedge\psi}{\psi}$$

$$(i.\vee)\ \frac{\varphi}{\varphi\vee\psi},\ \frac{\psi}{\varphi\vee\psi} \qquad (e.\vee)\ \frac{\varphi\vee\psi \quad \begin{matrix}[\varphi]\\ \vartheta\end{matrix} \quad \begin{matrix}[\psi]\\ \vartheta\end{matrix}}{\vartheta}$$

$$(i.\rightarrow)\ \frac{\begin{matrix}[\varphi]\\ \psi\end{matrix}}{\varphi\rightarrow\psi} \qquad (e.\rightarrow)\ \frac{\varphi\quad\varphi\rightarrow\psi}{\psi}$$

$$(i.\forall)\ \frac{\varphi(a)}{\forall x\varphi(x)} \qquad (e.\forall)\ \frac{\forall x\varphi(x)}{\varphi(t)}$$

$$(i.\exists)\ \frac{\varphi(t)}{\exists x\varphi(x)} \qquad (e.\exists)\ \frac{\exists x\varphi(x) \quad \begin{matrix}[\varphi(a)]\\ \psi\end{matrix}}{\psi}$$

Per la negazione si ha

$$(i.\neg)\ \frac{\begin{matrix}[\varphi]\\ \bot\end{matrix}}{\neg\varphi} \qquad (e.\neg)\ \frac{\varphi\quad\neg\varphi}{\bot}$$

Per l'eguaglianza si ha, con le restrizioni analoghe a quelle dette per **(R ¬ ∀)**:

$$(rifl)\ \frac{}{x \doteq x} \qquad (simm)\ \frac{x \doteq y}{y \doteq x} \qquad (trans)\ \frac{x \doteq y \quad y \doteq z}{x \doteq z} \qquad (sost)\ \frac{x \doteq y \quad \varphi(x)}{\varphi(y)}$$

La prima ha il numeratore "vuoto" perché è un assioma logico, anzi in questa presentazione è l'*unico* assioma logico, che non dipende da premesse. La seconda e la terza sono modi di scrivere le proprietà simmetrica e transitiva dell'eguaglianza e, in un certo senso sono superflue. Il sistema così ottenuto è detto, *minimale*. Si ottiene da esso il sistema *intuizionista* se si aggiunge la regola (che difficilmente viene utilizzata nella prassi scolastica)

$$\bot_i\ \frac{\bot}{\varphi}$$

Il sistema di deduzione *classico* si ottiene aggiungendo al sistema intuizionista la regola

$$\bot C\ \frac{\begin{matrix}[\neg\varphi]\\ \bot\end{matrix}}{\varphi}$$

Il sistema di deduzione naturale è uno strumento per fare le derivazioni proposizionali, considerando solo le regole proposizionali, oppure le derivazioni minimali o intuizionista, considerando solo le corrispondenti regole. Un simile tipo di analisi potrebbe servire a valutare meglio le difficoltà di comprensione da parte degli allievi. Questi aspetti non sono così chiaramente evidenti in altri sistemi formali. Le derivazioni sono rappresentate in forma di albero, in cui i nodi sono le occorrenze delle formule. Le formule che costituiscono i nodi iniziali (le *foglie*) dell'albero sono le *assunzioni* o *ipotesi* da cui parte la derivazione, mentre gli altri nodi dell'albero sono ottenuti dai precedenti mediante l'applicazione di una delle regole di inferenza introdotte sopra. Il nodo finale (la *radice*) dell'albero è la *conclusione* della derivazione.

Non sempre in una derivazione, la conclusione dipende da tutte le formule che costituiscono i nodi iniziali dell'albero. Infatti in un'inferenza ottenuta applicando le regole (e.→), (e.∨), (e.∃) e $\perp_C$ si può *scaricare* un numero qualsiasi, anche nullo, di assunzioni della forma indicata tra parentesi quadre e le conclusioni non dipendono dalle assunzioni che vengono scaricate. In generale si possono scaricare contemporaneamente più assunzioni della stessa forma, sebbene non sia sempre possibile scaricare tutte le assunzioni di quella forma presenti nell'intera derivazione.

Per le regole di inferenza (i.∀) e (e.∃) si aggiungono restrizioni sulla indeterminata libera a, detta *parametro proprio* dell'inferenza: si chiede che per (i.∀) il parametro proprio non compaia nelle assunzioni da cui la formula φ(a) dipende, né nella conclusione. In (e.∃) il parametro proprio non deve comparire in ψ né nelle assunzioni da cui ψ dipende, né nella conclusione. E' ancora un ottimo esercizio comprendere il motivo delle restrizioni, con opportuni esempi.

Se un'occorrenza della formula φ, appartenente alla derivazione, è la conclusione di una regola d'inferenza ottenuta applicando (i.→), (e.∨), (e.∃) o $\perp_C$, allora si scaricano tutte le assunzioni della forma indicata tra le parentesi quadre a cui φ dipende.

Il numero delle assunzioni scaricate può essere nullo. Per indicare in modo chiaro quali assunzioni vengono scaricate applicando una certa regola, si numerano in modo che quelle che non vengono numerate siano esattamente quelle aperte, cioè quelle da cui dipende la conclusione della derivazione. Più assunzioni possono essere contrassegnate dallo stesso numero e tali assunzioni sono trattate allo stesso modo, cioè vengono scaricate contemporaneamente, mentre si possono avere assunzioni della stessa forma, ma segnate con numeri diversi. Per chiarezza in caso di molteplici scarichi si pone accanto alla linea di frazione, il numero corrispondente all'insieme delle assunzioni scaricate. Alla fine restano le assunzioni non scaricate e la conclusione, la radice dell'albero. La presenza di una derivazione dalle assunzioni $\psi_1, \dots \psi_n$, non scaricate, della formula φ si indica con la scrittura $\psi_1, \dots \psi_n \vdash \varphi$.

Il sistema di Prawitz ha finalità logiche di altra natura, ma si presta ad un uso didattico efficace nella fase di riflessione sugli aspetti logici della Matematica. Le regole indispensabili per formalizzare i ragionamenti matematici sono alla portata degli studenti e possono essere oggetto di un insegnamento specifico alla fine del curriculum.

Si presenta, per un confronto, una derivazione in questo sistema del sillogismo *Baroco*

$$\forall x(\varphi(x) \to \psi(x)),\ \exists x(\vartheta(x) \wedge \neg\psi(x)) \vdash \exists x(\vartheta(x) \wedge \neg\varphi(x))$$

$$\dfrac{\dfrac{\exists x(\vartheta(x)\wedge\neg\psi(x)) \quad \dfrac{\dfrac{\dfrac{[\vartheta(a)\wedge\neg\psi(a)]^{2}}{\vartheta(a)}(e.\wedge) \quad \dfrac{\dfrac{\dfrac{\forall x(\varphi(x)\to\psi(x))}{\varphi(a)\to\psi(a)}(e.\forall) \quad [\varphi(a)]^{1}}{\psi(a)}(e.\to) \quad \dfrac{[\vartheta(a)\wedge\neg\psi(a)]^{2}}{\neg\psi(a)}(e.\wedge)}{\perp}(e.\neg)}{\neg\varphi(a)}(i.\neg)(1)}{\vartheta(a)\wedge\neg\varphi(a)}(i.\wedge)}{\exists x(\vartheta(x)\wedge\neg\varphi(x))}(i.\exists)}{\exists x(\vartheta(x)\wedge\neg\varphi(x))}(e.\exists)(2)$$

Riferimenti

[BM] BELL J., MACHOVER M., *A Course in Mathematical Logic,* North Holland, Amsterdam, 1977.

[L] LOLLI G., *Introduzione alla Logica formale,* Il Mulino, Bologna, 1991.

[M] MARCHINI C., [.]1: *Le sostituzioni e la didattica della Matematica,* Bollettino Unione Matematica Italiana, (7) **4** - **A** (1990), 145 - 153; [.]2: *Le sostituzioni e le relazioni,* L'insegnamento della Matematica e delle Scienze Integrate, **13** - n. 7 (Luglio 1990), 732 - 744.

[Me] MENDELSON E., *Introduzione alla Logica Matematica,* Boringhieri, Torino, 1972.

[P] PRAWITZ D., *Ideas and results in Proof Theory,* In: Proceedings of the Second Scandinavian Logic Symposium, (J. Fenstad, editor), North-Holland, Amsterdam, 1971.

Esercizi

CARLO MARCHINI

ESERCIZIO 1 Analizzare i brani seguenti tratti da libri di testo. Studiare gli usi degli enti grammaticali. Aggiungere alla piccola antologia considerata altri brani significativi.[20]

1.1 A. PALATINI, V. FAGGIOLI, *Complementi di matematica per i Licei Scientifici* - Ghisetti e Corvi Ed. - Milano, 1976, p. 81: «Nello studio di una qualsiasi questione si osserva che vi sono quantità il cui valore si mantiene inalterato ed altre che assumono valori diversi: le prime si dicono *costanti* e le seconde *variabili.* ...Una variabile alla quale si possa assegnare un valore a nostro arbitrio si dice *indipendente.* ...Tutte le volte che i valori di una variabile y dipendono da quelli di un'altra variabile x, si dice che y è una *funzione* di x e si scrive $y = f(x)$ (*si legge "y eguale ad effe di x"*). La lettera f può esser sostituita da un'altra lettera qualunque e si scrive, ad esempio, $y = g(x)$ oppure $y = \varphi(x)$, ecc. Il valore che $f(x)$ assume per un particolare valore di x, ad esempio $x = a$, si indica con $f(a)$. Il simbolo $f(x)$ molte volte sta a rappresentare un complesso di operazioni matematiche che si devono eseguire sopra i valori della variabile x per ottenere i valori corrispondenti della y, in tal caso $f(x)$ si dice *funzione matematica*

1.2 A pag. 199 si ha: «Si dice che una *variabile y è funzione della variabile x* quando esiste una legge che faccia corrispondere ad ogni valore della x uno ed un sol valore della y»

1.3 A pag. 204: «Sia x una variabile ed $y = f(x)$ una funzione di x definita in tutti i punti di un intervallo (a,b), eccetto al più in un punto c interno all'intervallo. *Si dice che per x tendente a c la funzione f(x) tende al limite finito l (oppure che ha per limite l) se, fissato un numero positivo ε arbitrariamente piccolo, si può trovare in corrispondenza ad esso, un intorno di c tale che per ogni valore di x di questo intorno (escluso al più il punto x=c) si abbia* $|f(x) - l| < \varepsilon$.»

1.4 In N. DODERO, P. BARONCINI, R. MANFREDI, *Elementi di matematica per gli Istituti Tecnici industriali a indirizzo sperimentale,* Vol. 4, Ghisetti e Corvi, Milano, 1990. a pag. 34 si ha: *«Si dice che, per x tendente a c, la funzione y = f(x) ha per limite l e si scrive*

$$\lim_{x \to c} f(x) = l. \qquad (1)$$

[20] La scelta dei brani non ha intenti critici, ma didattici.

se, fissato un numero positivo ε, arbitrariamente piccolo, si può trovare in corrispondenza ad esso, un intorno completo di c tale che per ogni x di tale intorno (escluso al più x=c) si abbia

$$|f(x) - l| < \varepsilon. \qquad (2)$$

OSSERVAZIONE 3 Per effettuare la verifica di un limite del tipo (1) occorre, fissato ε, risolvere la (2) e constatare che la soluzione determina un intorno completo di c. In caso contrario il limite non è mai verificato»

1.5 A pag. 97: «...2) *la variabile indipendente è sempre continua...*»

1.6 A pag. 147: «TEOREMA *Ogni funzione che ammette derivata finita in un punto, è continua in tale punto.*
... Questo teorema non è invertibile: *vi sono funzioni che sono continue in un punto, ma non sono in esso derivabili.*»
Un uso libero di notazioni per i quantificatori si ha a pag. 204: «Siano x_1 e x_2 due punti qualsiasi di I, con $x_1 < x_2$: per il Teorema di Lagrange si può scrivere $\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c)$ con $c \in (x_1, x_2)$»

1.7 Sulla quantificazione in G. ZWIRNER, *Complementi di Algebra e nozioni di Analisi Matematica per licei scientifici,* Nuova edizione a cura di L. Scaglianti, CEDAM, Padova, 1990, pag. 276: «... Sia E un insieme non vuoto di $\mathbb{R}$. *Si dice che l'insieme E è* limitato superiormente *quando esiste un numero b non minore di tutti i numeri di E.* Cioè quando: $\forall x \in E: x \leq b$.
... Se l'insieme E è limitato sia superiormente che inferiormente, si dice, senz'altro, *limitato;* in tal caso, si può sempre trovare un numero positivo k, tale che $\forall x \in E: |x| \leq k$, cioè tale che ogni numero di E sia, in valore assoluto, minore o uguale a k.»

1.8 Sulla interdefinibilità dei quantificatori si confrontino due brani, rispettivamente a pag. 277 e a pag 293: «... DEFINIZIONE *Sia* E *un insieme di numeri reali* limitato superiormente. *Si chiama* estremo superiore *di* E *il numero* L *che gode delle seguenti proprietà:*

1) Ogni numero di E *è minore od eguale ad* L*;*
2) Comunque si fissi un numero positivo ε*, esiste sempre* almeno *un numero di* E *più grande del numero* L - ε.

Esso risulta, dunque, il più piccolo dei maggioranti di E»

«Si chiama estremo superiore della f in A, l'estremo superiore del codominio f(A) della f, cioè l'estremo superiore dell'insieme dei numeri reali f(x), con $x \in A$, e si indica con: $\sup_{x \in A} f(x)$, oppure $\sup f(A)$ ...Quindi dire, per esempio, che il numero L è l'estremo superiore di f(x) in A, significa che:

1) *nessun valore assunto dalla* f(x) *in* A *supera il numero* L*;*

2) fissato un numero $\varepsilon > 0$, *ad arbitrio, esiste in* A *almeno un punto* x_0 *in cui risulta* $f(x_0) > L - \varepsilon$.»

1.9 L.TONOLINI, F.TONOLINI, *Corso superiore di matematica,* Vol. 2, Minerva Italica, 1991, pag. 43: «Per intorno di un punto P di una retta r intendiamo l'insieme dei punti, diversi da P, di un qualunque segmento contenente P nel suo interno»

1.10 e a pag. 44: «Per intorno di $+\infty$ intendiamo l'insieme di tutti i punti della retta r di ascissa maggiore (o maggiore o uguale) di un numero a»

1.11 A pag. 46: «Prefissato un numero positivo ε dovremo trovare un intorno del punto $x_0 = -2$ tale che per ogni x di questo intorno valga la relazione:\b\bc\|(\f(x - 2,x + 8) - \f(-2,3)) < ε, cioè sia: $\frac{-2}{3} - \varepsilon < \frac{x-2}{x+8} < \frac{-2}{3} + \varepsilon$. Risolviamo dunque il sistema equivalente a questa doppia disuguaglianza:

$$\begin{cases} \frac{x-2}{x+8} > \frac{-2}{3} - \varepsilon \\ \frac{x-2}{x+8} < \frac{-2}{3} + \varepsilon \end{cases}$$

tenendo presente che si può supporre $x + 8 > 0$ [...]. Otteniamo:

$$\begin{cases} x(5 + 3\varepsilon) > -10 - 24\varepsilon \\ x(5 - 3\varepsilon) > -10 + 24\varepsilon \end{cases}$$

e, poiché possiamo senz'altro supporre $5+3\varepsilon > 0$ e $5-3\varepsilon > 0$ ».

1.12 Da G. MELZI, L. TONOLINI, *Geometria per le Scuole Medie Superiori,* Minerva Italica (1991), pag. 49: «*Assioma 22* Esiste ed è unico il sottomultiplo di un segmento secondo un qualunque intero positivo (proprietà della divisibilità indefinita, detta anche postulato di Eudosso - Archimede).»

1.13 Da F. SPERANZA, A. ROSSI DELL'ACQUA, *Matematica per il terzo anno del triennio delle Scuole Medie Superiori,* Zanichelli, Bologna, 1974, pag. 145: «Una mappa f si dice *iniettiva* quando $f(a) = f(b) \Rightarrow a = b$. Una mappa f di *D* in *D'* si dice *suriettiva* allorché $\forall y \in D' \ \exists x \in D \ y = f(x)$.)»

1.14 FONTANA, ROVELLI, *Matematica,* A. Mondadori, Milano 1989, pag 118: «due insiemi finiti sono *uguali* se hanno lo stesso numero di elementi».

ESERCIZIO 2 Formalizzazione di alcune frasi. Tradurle se possibile in termini formali, specificando il vocabolario. Di ciascuna costruire la negazione e l'espressione formale della frase e della sua negazione. Cercare, laddove possibile, di presentare una formalizzazione delle stesse frasi, cambiando il vocabolario.

2.1 «Chi studia è promosso»; 2.2 «Chi non studia non è promosso»; 2.3 «Se non giochi non vinci» (pubblicità del Totip!); 2.4 «Chi tace acconsente»; 2.5 «Can che abbaia non morde»; 2.6 «Lasciate ogni speranza o voi ch'entrate»; 2.7 «Questo e quello per me pari sono»; 2.8 «C'è modo e modo di dire le cose»; 2.9 «Mangia la minestra o salta la finestra»; 2.10 «Non chi dice "Signore, Signore" entra nel Regno dei Cieli»; 2.11 «I contribuenti sono tenuti a versare le imposte entro il 15 ottobre o a versarle maggiorate del 10% dopo il 15 ottobre»; 2.12 «La tessera con 45 punti dà diritto a due tazze di finissima porcellana e con 65 punti ad una zuccheriera finemente decorata»; 2.13 «Disporsi su due file».

ESERCIZIO 3 Individuare presenze libere e vincolate di indeterminate, termini liberi per indeterminate nelle frasi seguenti.

3.1 «Una funzione F(x) si dice una primitiva di f(x) se $\frac{dF(x)}{dx} = f(x)$. Se la funzione f(x) è integrabile, l'integrale indefinito o funzione integrale di f(x), denotata con $\int f(x)dx$ è una primitiva di f(x). Una diversa notazione per la funzione integrale è $F(x) = \int_a^x f(t)dt$»

3.2 A. PALATINI, V. FAGGIOLI, *Complementi di matematica per i licei scientifici* - Ghisetti e Corvi Ed. - Milano, 1976, p. 264 e 265; «TEOREMA DEL VALOR MEDIO (o teorema di Lagrange o di Cavalieri). *Se la funzione f(x), definita nell'intervallo (a,b) è continua e derivabile in tutto l'intervallo, la differenza dei valori della funzione negli estremi a e b è eguale all'ampiezza b - a dell'intervallo moltiplicata per la derivata di f(x) in un punto interno all'intervallo.* [...]

TEOREMA *Se una funzione continua ha derivata nulla in tutti i punti di un intervallo, essa è costante in quell'intervallo.* Infatti se x è un punto qualunque dell'intervallo (a,b) applichiamo il teorema del valor medio all'intervallo (a,x); si ottiene $f(x) - f(a) = (x - a) \cdot f$

'(c) dove c è un punto interno all'intervallo (a,x). Ma per ipotesi, la derivata di f(x) è nulla in ogni punto (a,b) perciò è nulla anche in c, ossia si ha f '(c) = 0; ne segue che f(x) - f(a) = 0 ossia f(x) = f(a). Siccome x è un punto qualunque di (a,b), questo significa che f(x) assume in tutti i punti di (a,b) sempre lo stesso valore, cioè è una costante.»

ESERCIZIO 4 Analizzare i seguenti brani tratti da manuali, mettendo in evidenza i nessi deduttivi utilizzati. Completare l'antologia con altri brani significativi.

4.1 L. CATENI, R. FORTINI, *Il pensiero geometrico,* Vol. II, Le Monnier, Firenze, 1959, p. 19: «Osserviamo che T è prevalente a t_1. ... Analogamente, T è suvvalente a t_2.... Se ne deduce che [dette e l'estensione di t_1, S l'estensione di T ed E l'estensione di t_2], $e < S < E$. ...»

4.2 E. GALLO, *Fare Matematica,* vol. II, S.E.I., Torino, 1988, p. 564: «Risolvi il seguente sistema, secondo le istruzioni: $\begin{cases} 2x + y = 5 \\ 2x^2 - 3xy = 2 \end{cases}$. Esplicita un'incognita da un'equazione da cui compare solo al primo grado, e sostituisci l'espressione trovata nell'altra equazione: $\begin{cases} y = 5 - 2x \\ 2x^2 - 3x(5 - 2x) = 2 \end{cases}$. Risolvi ora l'equazione in una incognita: $\begin{cases} y = 5 - 2x \\ \ldots\ldots \end{cases}$. Trova i corrispondenti valori dell'*altra* incognita: $\begin{cases} x = 3 \\ y = -1 \end{cases}$ *vel* $\begin{cases} x = \ldots \\ y = \ldots \end{cases}$. Affermeresti che questo metodo è sempre utilizzabile per risolvere i sistemi di secondo grado? Motiva la risposta».

4.3 «$A \subseteq A \cup B$. *Dimostrazione*: Se $x \in A$, allora $x \in A \vee x \in B$».

4.4 I. BELLI, A. LUPO PERRICONE, L. PAGNI, S. PALLINI, *Osservare e dedurre,* S.E.I., Torino, 1988, pp. 188-189: «Ma, esistono poligoni con più di un asse di simmetria? e, se sì, qual è il numero massimo di assi per un poligono di n lati? ... Ora, prima di rispondere alla domanda che ci siamo posti per un poligono generico di n lati, vogliamo esaminare lo stesso problema [21]nel caso più semplice dei *poligoni regolari,* ricordando che *un poligono viene detto regolare se ha tutti i lati e tutti gli angoli uguali.* ... Nel caso che il poligono abbia un *numero dispari di lati,* come si verifica per un pentagono, un ettagono, un ennagono regolari, è facile vedere che la *retta bisettrice di uno degli angoli è asse di simmetria per il poligono.* Se r è infatti la bisettrice dell'angolo di vertice A, (fig. 16), è chiaro che $AE \rightleftarrows AB$ e quindi $E \rightleftarrows B$, per l'eguaglianza degli angoli in E e B, si ha poi $ED \rightleftarrows BC$, da cui in E e B, si ha poi $D \rightleftarrows C$. Quanti sono, allora, gli assi di simmetria di un tale poligono. Evidentemente se i

[21] I corsivi sono del testo.

lati sono n, anche gli angoli sono n. Dunque un poligono regolare con un numero dispari di lati ha tanti assi di simmetria quanti sono i suoi lati. Sarà lo stesso se *il numero dei lati è pari*? (esagono, ottagono, ecc.). Questa volta, *non solo le bisettrici degli angoli, ma anche le rette che uniscono i punti medi dei lati opposti sono assi del poligono.* Cerca per esercizio una dimostrazione di questa affermazione. Ti accorgerai che, in ogni caso, si trovano ancora tanti assi di simmetria quanti sono i lati del poligono. In definitiva: un poligono regolare di n lati ha n assi di simmetria.»

Sono tralasciati qui i disegni di un pentagono regolare con un asse di simmetria, la fig. 16, di un quadrato, di un esagono e di ottagono regolare con tutti gli assi di simmetria.

4.5 A. PALATINI, V. REVERBERI FAGGIOLI, *Complementi di Matematica,* Ghisetti e Corvi, Milano, 7ª ed., p. 207: «COR. *Se due funzioni continue* f(x) *e* g(x) *hanno derivate eguali in tutti i punti di un intervallo, esse differiscono per una costante.*

Infatti, posto F(x) = f(x) - g(x), si ha F'(x) = f'(x) - g'(x). Ma per l'ipotesi del teorema, in tutto (a,b) è f'(x) = g'(x), quindi F'(x) = 0, donde F(x) = costante, ossia f(x) - g(x) = costante, come volevasi dimostrare.»

4.6 V. DEL GIUDICE, S. MORINA, *Corso di Matematica,* Petrini, Torino, 1989, p. 409: «L'equazione $(2x+6)x - 12x^2 = -x(10x - 6)$ risulta un'identità perché, applicando al primo membro le proprietà delle operazioni, si può ottenere la catena di identità: $(2x+6)x - 12x^2 = 2x^2 + 6x - 12x^2 = 2x^2 - 12x^2 + 6x = (2 - 12)x^2 + 6x = -10x^2 + 6x$ e applicandole al secondo:$-x(10x - 6) = -10x^2 - x(-6) = -10x^2 + 6x$. Poiché i due membri dell'equazione risultato entrambi identici all'espressione $-10x^2 + 6x$, essi sono identici tra loro.»

4.7 L. SCAGLIANTI, L. VARAGNOLO, G. ZWIRNER, *Lezioni di Matematica: Algebra Informatica 1,* CEDAM, Padova, 1987, pag. 357: «Si dice che i due monomi ax^m e bx^n sono *eguali,* e si scrive: $ax^m = bx^n$, se a valori eguali della x associano lo stesso numero».

4.8 Si analizzi il testo di A. PALATINI, V. FAGGIOLI, *Complementi di matematica per i Licei Scientifici* - Ghisetti e Corvi Ed. - Milano, 1976, p. 264 e 265, riportato come Es. 3.2.

4.9 F. SPERANZA, A. ROSSI DELL'ACQUA, *Il linguaggio della Matematica,* Zanichelli, Bologna, 1981, p. 210: «Una successione di numeri reali diverge a $+\infty$, allorché per ogni $K \in \mathbf{R}_0^+$ esiste un n_0 tale che per ogni $n > n_0$ sia $a_n > K$. ... La successione $a_n = 2n$ diverge a $+\infty$. Infatti fissato K, per avere $a_n > K$, cioè $2n > K$, ovvero $n > \frac{K}{2}$, basta prendere $n_0 > \frac{K}{2}$. ...».

4.10 F. SPERANZA, A. ROSSI DELL'ACQUA, *Il Linguaggio della Matematica*, 2ª ed., Zanichelli, Bologna, 1988, p. 996: «teorema 1. Due rette distinte hanno al massimo un punto in comune (cioè non ne hanno alcuno o ne hanno uno solo).

Naturalmente si ammette che valga l'assioma **A**. Per ricordare questo fatto si potrebbe dire così: se vale l'assioma **A**, due rette distinte hanno al più un punto in comune.

Dimostrazione Immaginiamo che l'affermazione non sia vera, cioè che vi siano due rette distinte che hanno due (o più) punti comuni, diciamo A, B. Allora A, B (che sono distinti) apparterrebbero a due rette distinte, contro l'assioma **A**. Dunque l'ipotesi che due rette distinte abbiano più di un punto in comune contrasta con l'assioma **A**, ed è quindi da respingere.»

4.11 Altro esempio: «Si definisca l'inclusione *propria* tra due insiemi A e B, in simboli $A \subset B$, come l'inclusione di due insiemi diversi, cioè $A \subseteq B \wedge A \neq B$.

PROPOSIZIONE L'inclusione propria è transitiva.
Dimostrazione Si assuma per ipotesi che $A \subset B \wedge B \subset C$. Di qui si ha $A \subseteq B \wedge B \subseteq C$. Per la proprietà transitiva dell'inclusione $A \subseteq C$. Resta da provare $A \neq C$. Questa parte si prova per assurdo: si suppone, ipotesi ausiliaria, che $A = C$. Di qui e da $A \subseteq B \wedge B \subseteq C$ si ottiene $A \subseteq B \wedge B \subseteq A$. Per la proprietà antisimmetrica della inclusione, si ha $A = B$, assurdo.»

4.12 V. DEL GIUDICE, S. MORINA, *Corso di Matematica*, Petrini, Torino, 1989, p. 242: «La forma di ragionamento che, nota la verità di un'implicazione e della sua antecedente, permette di dedurre la verità della conseguente prende il nome di *modus ponens* e può essere rappresentata:

$$\begin{array}{l} p \Rightarrow q \\ \underline{\quad p \quad} \\ q \end{array} \text{ ».}$$

ESERCIZIO 5 Provare in ciascuno dei sistemi deduttivi considerati alcune formule.

5.1 Ripetere le dimostrazioni presentate nelle lezioni col sistema di Hilbert, senza usare il testo;

5.2 Ripetere coi tableaux e con la deduzione naturale le dimostrazioni presentate nelle lezioni col sistema di Hilbert.

5.3 Dimostrare, se possibile, coi vari metodi le seguenti formule. Nel caso non esista la dimostrazione, trovare i contromodelli, sfruttando i tableaux (o le tavole semantiche):

$(((\varphi \rightarrow \psi) \rightarrow \varphi) \rightarrow \varphi)$; $(((\varphi \rightarrow \psi) \rightarrow \varphi) \rightarrow \neg\psi)$;
$\forall x(\varphi(x) \rightarrow \forall y\, \psi(x,y)) \vdash \exists y(\exists x(\varphi(x) \rightarrow \psi(x,y))$;
$\exists y(\exists x(\varphi(x) \rightarrow \psi(x,y)) \vdash \forall x(\varphi(x) \rightarrow \forall y\, \psi(x,y))$
.....

PARTE SECONDA

Conferenze e Percorsi didattici

L'Insegnamento della Logica nella Scuola secondaria superiore

ANNA SGHERRI COSTANTINI
Ministero della Pubblica Istruzione
Viale Trastevere
00153 Roma

Fra gli aspetti più caratteristici della cultura scolastica degli ultimi dieci anni è certamente da considerare l'interesse crescente per la Logica, interesse peraltro dimostrato non solo da parte di insegnanti di Matematica o di Filosofia—discipline in qualche modo privilegiate in questo campo di studio—ma anche, e sempre più diffusamente, da parte di estranei a questi problemi.

In parte ciò deriva dalla convinzione che identifica, tout court, la Logica con l'arte del ragionamento. Tale opinione può essere pericolosa perché crea un'aspettativa che questa disciplina non è in grado di soddisfare circa la capacità—quasi miracolistica—che un suo inserimento fra i programmi di studio secondari possa avere per compensare eventuali carenze degli studenti nel "ragionare".

Com'è noto, infatti, la Logica non insegna—come sua specificità—a ragionare, ma ad analizzare i ragionamenti.

Sarebbe impresa quasi disperata e arbitraria da parte mia tentare una definizione di LOGICA, di logica pura e semplice, senza aggettivi. Forse non esiste altro termine, salvo quello di FILOSOFIA, che abbia conosciuto tanta varietà di discordanti definizioni.

Per taluni la Logica non si propone neppure, almeno in via primaria, di stabilire quali ragionamenti siano corretti, bensì di individuare "calcoli logici" mediante i quali meccanizzare l'attività deduttiva e "dominare" l'insieme delle conseguenze di un insieme di premesse in modo da ragionare sulle teorie nel loro complesso.

Ciò richiede, pertanto, alcune precisazioni circa il significato che s'intende dare a questa disciplina in ragione del suo iter storico e della matrice culturale di chi parla.

Se ogni scienza può essere considerata genericamente come una ricerca della verità, il problema che s'impone è come si consegue la conoscenza della verità, o meglio, come si riconosce che un asserto è vero. Le ipotesi di soluzione possono essere due: o la verità di un asserto può essere stabilita *immediatamente*, oppure la verità risulta per *mediazione,* cioè per implicazione da altri asserti noti come veri.

Queste ipotesi rimandano ai criteri in base ai quali decidere *quando* una proposizione si può accettare come *immediatamente vera* e *quando* una proposizione deve essere considerata *implicata da* certe altre.

Sono problemi non semplici.

La prima via dell'indagine è di pertinenza di specifiche discipline, la filosofia generale, la gnoseologia, la metodologia delle singole scienze. Al compito di definire i requisiti di una buona mediazione cerca di soddisfare la Logica. Si tratta, a parere di chi scrive, di esplicitare il significato del termine *implicazione.*

Quali sono le regole in base alle quali un asserto deve considerarsi *implicato* da un altro o da altri ?

Rispetto al senso comune che lascia imprecisata la nozione o, addirittura, la riporta ad una forma di convincimento psicologico personale, la Logica esplicita e fissa le condizioni dell'implicazione tra asserti.

Sotto questo aspetto essa potrebbe anche essere definita come l'indagine sul "ragionare corretto", quando si voglia caratterizzare l'ambito della Logica servendosi di un discorso sufficientemente preciso, ma non tecnico.

Tuttavia, il termine "ragionamento" ha un significato troppo vasto, comprendendo, infatti, anche il *discorrere per induzione*, per cui conviene usare il termine "implicazione" o "deduzione".

Da qui si passa all'oggetto cui fa riferimento la Logica: il *linguaggio.*

Il pensiero, i ragionamenti si esprimono attraverso il linguaggio. Solo nelle espressioni esplicite, finite ed analizzabili, è possibile praticare un lavoro di catalogazione, di scomposizione e di classificazione dal quale ricavare regole di combinazione esplicite e precise.

Il riconoscimento del diretto rapporto tra *Logica* e *Linguaggio* è prezioso al fine di evitare alcuni equivoci. Noti sono, ad esempio, i pericoli di ricondurre la Logica ad un capitolo della Psicologia, o ad un sottodominio dell'Ontologia. La storia della Logica dal '400 all' '800 è piena di questi casi.

La Logica, quindi, può essere definita come la disciplina che si occupa del corretto dedurre.

Per comprenderne la natura e, di conseguenza, la sua funzione formativa, è necessario tenere presente anche la sua storia, intendendo non i progressi o regressi compiuti nei secoli ma, piuttosto, i diversi aspetti assunti, le diverse forme nelle quali si è realizzata.

Infine è da superare il pregiudizio kantiano circa l'astoricità della Logica. Nella Prefazione della Critica della Ragion Pura, II Edizione, Kant dice:

> "...a cominciare da Aristotele non ha dovuto fare nessun passo indietro...e sino ad oggi la Logica non ha potuto fare un passo innanzi di modo che secondo ogni apparenza essa è da ritenersi come chiusa e completa".

La Logica, in conclusione, è una disciplina autonoma che si è sviluppata in un ampio ventaglio di ricerche e che ha molteplici diramazioni ma in relazione al campo di interesse rappresentato in questa sede, si prefigge l'analisi e la caratterizzazione dei procedimenti deduttivi, l'individuazione e l'esplicitazione dei canoni del corretto dedurre e—non secondariamente—del linguaggio mediante il quale si esprimono i ragionamenti stessi.

Fino dalle sue origini è stata accompagnata da polemiche circa i suoi rapporti di appartenenza o di dipendenza dalla Filosofia e dalla Matematica—e più recentemente dall'Informatica. Certo che la Logica, pur affrontando problemi suoi propri, affonda le sue radici nella tradizione filosofica e scientifica. Si serve di strumenti matematici sempre più sofisticati, ed in questo senso è condivisibile la locuzione largamente praticata di Logica Matematica. D'altra parte la Logica Matematica non esaurisce completamente l'orizzonte della Logica. Gli universi del ragionamento sono molteplici e ciascuno ha caratteristiche peculiari seppure non tutti abbiano raggiunto un soddisfacente grado di formalizzazione.

Passiamo, ora, ad alcune *implicazioni didattiche*, tenendo conto degli spazi offerti dai nuovi curricula sperimentali. Mi riferisco ai nuovi programmi di Filosofia per il Triennio di tutti gli indirizzi previsti dal progetto elaborato dalla Commissione "Brocca". Il fatto che nel programma di matematica siano stati introdotti gli "elementi di Logica", non deve indurre a pensare che la Logica sia da considerare una specie di *premessa* alla Matematica.

Nelle indicazioni metodologiche, molto opportunamente, si dice: "...gli elementi di Logica devono essere visti come una riflessione che si sviluppa a mano a mano che matura l'esperienza matematica dell'allievo..".

Né la Logica deve essere intesa come *premessa* al ragionamento tout court. La maturazione intellettuale del soggetto pone le basi, piuttosto, per fare progressi in questo settore. E per stimolare la maturazione intellettuale entrano in gioco tutte le discipline, in particolare quelle che introducono e addestrano all'astrazione e alla formalizzazione, la Filosofia, appunto.

La fascia scolastica che può trarre il maggior beneficio, a mio parere, è il Triennio superiore; ciò ridimensiona in modo significativo le ipotesi che prevedono serie anticipazioni

in età anche molto precoce. Tuttavia non è da escludere che lo studio dei processi deduttivi—*analisi, esplicitazione, formalizzazione*—abbia una benefica ricaduta anche sull'attività logica naturale, specie per quanto riguarda la capacità espositiva e l'attività definitoria.

Una più stretta collaborazione tra Filosofia e Matematica si potrà avere nel momento in cui si approfondiranno i problemi relativi alla struttura delle teorie matematiche e all'analisi degli aspetti metamatematici. Gli stimoli offerti dalla conoscenza di settori della ricerca filosofica contemporanea come, ad esempio, la filosofia della scienza o la filosofia del linguaggio, sono un utile strumento per capire alcuni aspetti fondamentali della complessa realtà contemporanea, rendendo così più incisivo lo stesso insegnamento filosofico.

Su queste posizioni sono sicuramente convergenti i programmi di Filosofia elaborati dalla Commissione "Brocca".

Lo storicismo di stampo hegeliano che caratterizza l'impostazione del programma tradizionale, ed il manualismo sempre più diffuso, non hanno facilitato l'apertura al contemporaneo anche per coloro che si sono impegnati da tempo nel rinnovamento didattico di questo insegnamento. Per molti anni—o decenni—di Logica si è parlato veramente poco e solo in riferimento ad autori studiati in una visione interpretativa di maniera. Maggiore attenzione è stata riservata, per motivi facili da capire, ad Aristotele, ma anche in questo caso senza alcuna lettura diretta degli scritti di Logica. Lo stesso interesse dimostrato in passato per la logica medioevale, scaturiva dallo stretto legame con le questioni di natura teologica. Valga l'esempio rappresentato dal problema degli universali.

Oltre a tutto questo deve esser considerata anche la formazione iniziale degli insegnanti e il problema, strettamente correlato, della presenza/assenza di corsi specifici nei curricula universitari. Quante sono le cattedre di Logica nelle facoltà di Filosofia in Italia ?

Assumendo, dunque, i programmi "Brocca" come nuovo punto di riferimento, due sono le modalità per individuare le occasioni di collaborazione.

La *prima modalità* è costituita dalla condivisione di alcune finalità, e relativi obiettivi di apprendimento, indicate come specifico dell'insegnamento filosofico e presenti anche per l'insegnamento della Matematica.

Per esempio, le seguenti Finalità dei programmi di Filosofia per gli indirizzi liceali,

5. l'esercizio del controllo del discorso, attraverso l'uso di strategie argomentative e di procedure logiche,
6. la capacità di pensare per modelli diversi e di individuare alternative possibili, anche in rapporto alla richiesta di flessibilità nel pensare, che nasce dalla rapidità delle attuali trasformazioni scientifiche e tecnologiche,

sono da collegare con le Finalità 1,2,4,5 dei programmi di Matematica di seguito riportate:

1. l'acquisizione di conoscenze a livelli più elevati di astrazione e di formalizzazione;
2. la capacità di cogliere i caratteri distintivi dei vari linguaggi (storico-naturali, formali, artificiali);
3. la capacità di utilizzare metodi strumenti e modelli matematici in situazioni diverse;
4. l'attitudine a riesaminare criticamente e a sistemare logicamente le conoscenze via via acquisite;
5. l'interesse sempre più penetrante a cogliere aspetti genetici e momenti storico filosofici del pensiero matematico.

Analogamente, i seguenti Obiettivi di apprendimento 3 e 4 della Filosofia,

3. compiere, nella lettura del testo, le seguenti operazioni:
3.1. definire e comprendere termini e concetti;
3.2. enucleare le idee centrali;
3.3. ricostruire la strategia argomentativa e rintracciarne gli scopi;
3.4. saper valutare la qualità di un'argomentazione sulla base della sua coerenza interna;
3.5. saper distinguere le tesi argomentate e documentate da quelle solo enunciate;
3.6. riassumere, in forma sia orale che scritta, le tesi fondamentali;
3.7. ricondurre le tesi individuate nel testo al pensiero complessivo dell'autore;
3.8. individuare i rapporti che collegano il testo sia al contesto storico di cui è documento, sia alla traduzione storica nel suo complesso
3.9. dati due testi di argomento affine, individuarne analogie e differenze;
4. individuare analogie e differenze tra concetti, modelli e metodi dei diversi campi conoscitivi, a partire dalle discipline che caratterizzano i diversi indirizzi di studio,

sono da collegare con gli Obiettivi 9 e 10 della Matematica:

9. inquadrare storicamente l'evoluzione delle idee matematiche fondamentali;
10. cogliere interazioni tra pensiero filosofico e pensiero matematico.

La *seconda modalità* consiste nell'avviare la trattazione di tematiche e/o di autori afferenti allo sviluppo della Logica in modo contestuale e coordinato. In questo caso gli esempi sono quasi superflui, tale è la ricchezza e la varietà del materiale a disposizione ed utilizzabile didatticamente. Basti ricordare alcuni tra i nuclei tematici indicati per il III, IV, V anno degli indirizzi liceali:

5. Filosofia e scienza nel pensiero antico (III Liceo Classico e Linguistico)
7. Le scienze tra il '700 e l' '800 (IV, Liceo Scientifico)
3. Matematica e Logica nell'800 e nel ' 900
18. L'Intelligenza Artificiale (V, Liceo Scientifico).

E' da aggiungere inoltre che gli insegnanti possono proporre autonomamente altri temi o altri percorsi.

Per passare dal regno dei puri propositi al terreno concreto dell'esperienza, è più opportuno—direi quasi urgente—assicurare ai docenti una preparazione professionale adeguata. Nel campo della Logica gli insegnanti di Filosofia denunciano gravi debolezze anche teoriche. La diversa provenienza culturale (prevalentemente letteraria, o pedagogica, o filosofica) costituisce un ulteriore elemento di complessità nella predisposizione di un piano di formazione in servizio efficace e credibile.

Le iniziative fino ad ora attivate a questo scopo dalla Direzione Generale per l'Istruzione Classica, e le occasioni di incontro e di discussione offerte dai convegni della Società Filosofica Italiana, hanno confermato la necessità di insistere nella sollecitazione a porre le questioni emergenti dalla ricerca filosofica contemporanea, ed in particolare la Logica, al centro di ogni programma volto ad aggiornare e riqualificare gli insegnanti di questa disciplina; e in questa direzione ci stiamo muovendo.

Diversa è la situazione degli insegnanti di Matematica. Il Piano Nazionale per l'Informatica è stato, ad esempio, un'occasione preziosa di cui hanno potuto beneficiare tutti.

Oltre agli interventi specifici di competenza di ogni ambito disciplinare, sarebbe molto utile ipotizzare momenti di formazione comune su alcuni temi particolarmente rilevanti, o progettare insieme segmenti di percorso formativo. Ciò favorirebbe il superamento di quell'isolamento che è rilevabile in ambedue i gruppi di docenti e l'acquisizione di un linguaggio comune; condizioni, queste, preliminari per procedere ad una programmazione *integrata* su quegli aspetti scelti in fase progettuale all'inizio del Triennio.

In conclusione, l'inserimento della Logica deve essere, a mio giudizio, il più possibile autonomo rispetto alle altre discipline, anche se la sua collocazione è presente solo in alcuni programmi—più consistentemente in quelli di Matematica, per evidenti ragioni tecniche e di opportunità. Sia considerata per se stessa che per la natura specifica del suo compito, la Logica richiede competenze e capacità che la candidano come disciplina da inserire nel Triennio superiore.

L'integrazione dei punti di vista matematico e filosofico, seppure questa distinzione continui ad essere solo "accademica", è da considerare, prima, una modalità di approccio metodologico ai problemi, poi, un punto di arrivo, l'esito di un processo formativo a cui concorrono tutte le discipline scolastiche. Ciò richiede, come si è detto, un adeguato lavoro

di preparazione, sia attraverso forme mirate di aggiornamento professionale, sia—e soprattutto—attraverso il coinvolgimento dell'Università nella strutturazione del curriculum destinato ai futuri insegnanti.

In ultimo, un breve cenno al progetto elaborato da Lipman "Philosophy for Children", attualmente in fase di sperimentazione in molti paesi del mondo, ai quali recentemente si è aggiunta anche l'Italia. Lipman parte dall'idea della pervasività didattica della riflessione e della ricerca filosofica. La trasversalità di questo insegnamento sarebbe costituita dalla sua natura *formale,* da cui l'interesse per la Logica.

Ad una lettura superficiale il progetto sembrerebbe identificare la Logica con l' "arte del ragionare bene", e ciò potrebbe convincere ad anticipare elementi di Logica già nella scuola elementare o materna. Ad un esame più approfondito emerge, invece, il senso reale del progetto di Lipman, e la ragione per cui esso è richiamato in questa sede.

E' un progetto che tenta di applicare la convinzione di Bruner che vede nel linguaggio il più potente strumento di formazione dell'uomo e sostiene che una valida teoria dell'istruzione potrà indicare le procedure didattiche capaci di favorire, accelerare, migliorare i processi di crescita. Si tratta di una raffinata operazione di tecnologia educativa, che interviene sulle strutture cognitive che presiedono allo sviluppo del pensiero ipotetico-deduttivo e svolge, pertanto, una funzione propedeutica.

Il libro di Lipman "Il Prisma dei perché" (Edizioni Armando Scuola, 1982) è un'esposizione *introduttiva* delle regole della Logica ricavate dalla vita e applicate in situazioni concrete. La parola "filosofia" è stata una remora alla diffusione del progetto, per le sue molteplici implicazioni e ambigue interpretazioni.

Entro questi limiti, ma non oltre, si può pensare ad "anticipazioni" che appaiono utili in quanto promuovono specifiche attività di pensiero che formano i requisiti di base per i procedimenti discorsivi.

Immagini della Nozione di Dimostrazione

CARLO CELLUCCI
Dipartimento di Studi Filosofici ed Epistemologici
Università di Roma 'La Sapienza'
via Nomentana 118
00161 Roma

1. Attualità della nozione di dimostrazione

La nozione di dimostrazione svolge un ruolo cruciale nell'attività matematica, eppure stranamente molta parte della attuale storiografia della matematica focalizza l'attenzione più sui risultati matematici che sui mezzi attraverso cui essi vengono ottenuti. Così non si tiene conto del fatto che esiste una stretta relazione tra risultati e mezzi, e in particolare si dimentica che la stessa scelta dei risultati è legata all'esistenza di certi mezzi.

Di fatto la nozione di dimostrazione costituisce un tema sottosviluppato della riflessione contemporanea, sia storica che epistemologica, sulla matematica. Questo sarebbe stato giustificato nella prima metà di questo secolo, quando tale nozione veniva studiata, non per comprendere la realtà matematica, ma per realizzare programmi di fondazione della matematica come quelli di Frege o di Hilbert. In tali programmi non si consideravano le dimostrazioni matematiche effettivamente usate dai matematici, ma si introduceva una nozione astratta di dimostrazione, quella di *dimostrazione formale*, solo perché era la più consona a realizzare quei programmi fondazionali, e non perché essa avesse qualcosa a che fare con le dimostrazioni reali. Ma trascurare di dedicare attenzione alla nozione di dimostrazione reale risulta assolutamente ingiustificato oggi, quando nella pratica matematica intervengono alcune grosse trasformazioni che chiamano in causa il concetto stesso di dimostrazione. Ciò è testimoniato da un articolo, apparso alla fine dello scorso anno su *Le Scienze*, in cui si suggerisce che la recente dimostrazione da parte di Wiles della congettura di Fermat potrebbe essere il canto del cigno della nozione tradizionale di dimostrazione.

Senza arrivare a tanto, il logico americano Keith Devlin prevede che nei prossimi cinquant'anni la nozione di dimostrazione perderà molto della sua importanza perché un

numero sempre maggiore di matematici farà matematica senza usare le dimostrazioni. Già adesso esistono centri come il Geometry Center dell'Università del Minnesota dedicato allo sviluppo della geometria senza dimostrazioni, ed esiste una rivista, *Experimental Mathematics*, che ospita lavori dedicati alla nuova matematica senza dimostrazioni. Persino in un universo statico come quello dell'Università italiana si avverte qualche fremito se il Dipartimento di Matematica dell'Università di Pisa ha annunciato che dal prossimo anno accademico, accanto al tradizionale primo biennio del corso di laurea in matematica, verrà istituito un biennio alternativo rivolto a mettere in risalto gli aspetti applicativi della matematica, nell'ambito del quale verrà rivalutata la componente intuitiva rispetto a quella assiomatico-deduttiva della matematica.

All'origine di questa crisi della nozione tradizionale di dimostrazione sta il fatto che sempre più la soluzione di problemi matematici interessanti sembra richiedere dimostrazioni di complessità impressionante. All'inizio del secolo quasi tutte le dimostrazioni erano abbastanza brevi da poter essere lette da cima a fondo in una sola volta ed erano opera di un unico matematico. Oggi le dimostrazioni occupano spesso centinaia di pagine, sono opera di molte persone, in molti casi fanno essenzialmente uso del calcolatore e possono richiedere anni per essere controllate, ammesso che questo sia possibile. Per esempio la classificazione dei gruppi finiti, ultimata all'inizio degli anni Ottanta, è opera di più di cento ricercatori e consta di circa cinquecento articoli, per un totale di circa quindicimila pagine. Si dice che l'unica persona che la comprendesse nella sua totalità fosse Daniel Gorenstein, il coordinatore del progetto, che è morto qualche anno fa portandosi nella tomba questa sua conoscenza. Nel caso, poi, delle dimostrazioni fatte col calcolatore, il controllo diventa quasi impossibile perché coinvolge il problema della correttezza dei programmi, una questione in generale indecidibile e che, anche nei casi in cui è decidibile, non è fattibile in pratica per ragioni di complessità.

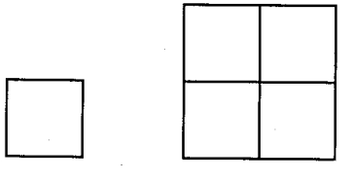
Le difficoltà legate alla complessità hanno fatto nascere l'esigenza di trovare il modo di ottenere risultati matematici superando i limiti della complessità. Le soluzioni correntemente proposte sono di due tipi. La prima consiste nel sostituire la dimostrazione in quanto ragionamento discorsivo con la dimostrazione come vedere. Usando la grafica computazionale si costruiscono modelli visivi che aiutano a cercare la soluzione e la fanno vedere sullo schermo del computer. La seconda soluzione consiste nell'assimilare i risultati della matematica a quelli della fisica: si rinuncia alla certezza assoluta dei risultati, contentandosi di un alto grado di probabilità. Le due soluzioni non sono necessariamente distinte, anzi spesso vengono adottate contemporaneamente.

La dimostrazione, quale intesa tradizionalmente, veniva considerata come diretta a *persuadere* della verità della proposizione dimostrata. Con Frege, a questa funzione della persuasione, si aggiunge anche quella della *fondazione*: scopo della dimostrazione non è più

soltanto quello di persuadere della verità della proposizione dimostrata, ma anche e soprattutto quello di mostrare il fondamento della sua verità. Ma i nuovi sviluppi della matematica sottolineano invece che la dimostrazione, nel nuovo senso empirico, deve far *vedere* perché una proposizione è vera. Inoltre c'è il fatto che spesso nella tradizione filosofica la dimostrazione è stata considerata come uno strumento per *scoprire* verità matematiche. Abbiamo così almeno tre concezioni del dimostrare: 1) dimostrare come vedere; 2) dimostrare come persuadere; 3) dimostrare come scoprire.

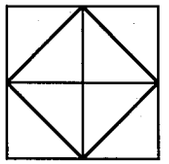
2. Dimostrare come vedere

Il prototipo della concezione del dimostrare come vedere è dato dalla soluzione data da Platone al problema della duplicazione del quadrato.[1] Tale soluzione viene spesso considerata come una tappa verso le dimostrazioni assiomatiche di Euclide, ma in realtà configura una diversa nozione di dimostrazione. Per guidare l'allievo alla soluzione del problema della duplicazione del quadrato, il maestro traccia una serie di figure la cui ispezione conduce l'allievo a vedere la soluzione del problema. In particolare, all'inizio il maestro traccia il quadrato di lato 2 (e quindi di superficie 4) e chiede all'allievo quale dovrebbe essere il lato del quadrato di superficie 8. L'allievo risponde che il lato dovrebbe essere 4. Il maestro allora traccia il quadrato di lato 4 e mostra all'allievo che la sua superficie non è 8 ma 16.



Ma il tentativo, sebbene errato, non è stato vano perché da esso si è imparato qualcosa, cioè che occorre che il lato del quadrato di superficie otto sia maggiore di 2 e minore di 4. L'allievo allora propone come nuova soluzione il lato 3. Ma il maestro ha facile gioco a mostrargli che anche questa risposta è errata perché in tal caso la superficie sarebbe nove. Di fronte a ciò l'allievo si dichiara impotente, perché tutte le soluzioni in numeri interi si sono rivelate errate e così egli crede di aver esaurito tutte le possibilità. A questo punto il maestro compie la mossa risolutiva: traccia di nuovo il quadrato di superficie sedici e lo presenta come la giustapposizione di quattro quadrati, poi traccia le diagonali dei quattro quadrati, e osserva che le diagonali dividono i quattro quadrati a metà.

[1] Cfr. G. CAMBIANO (ed.), *Filosofia e scienza nel mondo antico*, Torino (Loescher) 1986, pp. 88-96.

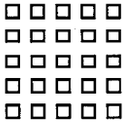


Notando che di tali metà ve ne sono quattro, il maestro mostra all'allievo che il quadrato i cui lati sono costituiti dalle diagonali ha superficie otto (la metà di sedici).

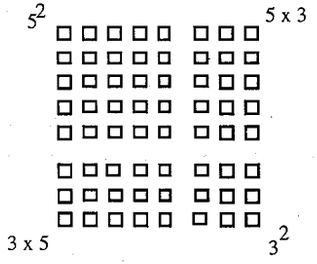
La soluzione del problema della duplicazione del quadrato consta, dunque, dei seguenti tre passi.

1) Si esamina un singolo caso del problema della duplicazione del quadrato, cioè quello della duplicazione del quadrato di lato due.
2) Si procede per tentativi ed errori, considerando vari casi per analogia con casi simili, per esempio considerando il caso del quadrato di lato quattro per analogia con quello del quadrato di lato due.
3) Ci si rende conto che la figura che mostra la soluzione del problema non la mostra solo per il particolare caso considerato, cioè per il quadrato di lato due, ma per qualsiasi quadrato, dal momento che la proprietà mostrata dalla figura non dipende dalla specifica lunghezza del lato, e quindi vale per lati di lunghezza qualsiasi.

Il dimostrare come vedere non riguarda solo risultati geometrici, ma anche risultati che, pur non essendo geometrici, possono essere dimostrati usando figure di qualche tipo. Tale è il caso delle dimostrazioni dei pitagorici di leggi come la formula del quadrato di un binomio $(a+b)^2=a^2+2ab+b^2$. Consideriamo, ad esempio, il caso $a=5$ e $b=3$. Rappresentiamo i numeri quadrati mediante dei quadrati, e.g. il numero quadrato 5^2 tramite la seguente figura.



Per dimostrare la formula del quadrato del binomio nel caso considerato, basta tracciare la seguente figura



Per rendersi conto della verità dell'eguaglianza non è necessario ricorrere ad alcuna deduzione, basta saper osservare la figura.

Questi esempi del dimostrare come vedere costituiscono il prototipo dello sviluppo recente della matematica senza dimostrazioni. In quest'ultima il disegno di figure sulla sabbia è sostituito dalla grafica computazionale.

3. Dimostrare come persuadere

La concezione del dimostrare come persuadere sta alla base del metodo assiomatico. Per primo Aristotele nei *Secondi Analitici* ha introdotto l'idea che la dimostrazione matematica debba essere concepita come una dimostrazione in un sistema assiomatico.[2] Il metodo assiomatico è un metodo che presenta le seguenti caratteristiche fondamentali.

1) Il metodo assiomatico non è un metodo di scoperta bensì un metodo di giustificazione. Scopo della dimostrazione è giustificare la verità della proposizione dimostrata, persuadendo così ad accettarla.
2) Il metodo assiomatico è un procedimento di costruzione della dimostrazione che va dall'alto verso il basso.
3) Sia i principi che la proposizione da dimostrare sono noti all'inizio della ricerca della dimostrazione.
4) In particolare i principi sono dati una volta per tutte: cambiare i principi vuol dire cambiare teoria.
5) Le dimostrazioni hanno un carattere statico e atemporale.
6) Nel metodo assiomatico viene stabilito un ordinamento globale di tutte le proposizioni di una data teoria.
7) I principi sono un punto di partenza assoluto per dimostrare qualsiasi proposizione. In questo senso il metodo assiomatico è un metodo indipendente dalla particolare proposizione che si vuol dimostrare.
8) La strategia della dimostrazione non fa parte del concetto di dimostrazione.
9) Il metodo assiomatico ha un carattere di monologo.
10) In esso la comunicazione non svolge alcun ruolo, quindi non entra a far parte del concetto di dimostrazione.
11) La comunicazione non è necessaria, dal momento che ogni sistema assiomatico deve contenere in sé tutta la conoscenza su un dato dominio.
12) Per il metodo assiomatico i vari campi della matematica sono sistemi chiusi, cioè sistemi i cui principi sono dati una volta per tutte. Anzi, tutta la matematica nel suo complesso è un sistema chiuso.

[2] Cfr. G. CAMBIANO (ed.), *op. cit.*, pp. 229-231.

Nonostante i suoi indubbi meriti, il metodo assiomatico ha alcuni altrettanto indubbi limiti.

1) Il metodo assiomatico non riesce a giustificare i propri principi. Questa è una conseguenza dei due teoremi di incompletezza di Gödel.

2) Il metodo assiomatico non rende conto del diverso ruolo dei principi: per esso gli assiomi sono problematici mentre le definizioni sono banali. Al contrario, spesso introdurre una definizione equivale a introdurre un'ipotesi. Per esempio la definizione euclidea di cerchio presuppone l'esistenza di un cerchio.

3) Il metodo assiomatico non riesce a giustificare la deduzione a partire dai principi. Ciò è mostrato, ad esempio, dal seguente divertente argomento di Lewis Carroll.[3] Consideriamo le proposizioni:

(A) Cose che sono eguali a una stessa cosa sono eguali tra loro.

(B) I due lati di questo triangolo sono cose che sono eguali a una stessa cosa.

(Z) I due lati di questo triangolo sono eguali tra loro.

Accettare (A) e (B) non comporta accettare (Z). Per farlo occorrerebbe aggiungere a (A) e (B), l'ipotesi:

(C) Se (A) e (B), allora (Z).

Ma anche questo non basterebbe e occorrerebbe aggiungere a (A), (B), (C) l'ipotesi:

(D) Se (A), (B) e (C), allora (Z).

E così via all'infinito.

L'impossibilità di dare una giustificazione della deduzione deriva dal fatto che, a differenza degli assiomi logici, le regole di inferenza logiche non possono essere espresse nel linguaggio del sistema ma devono essere espresse nel metalinguaggio, quindi è impossibile darne una giustificazione rimanendo all'interno del sistema. Una giustificazione della deduzione comporterebbe un rimando all'infinito.

4) Il metodo assiomatico lascia inspiegato il processo della scoperta matematica. Per il teorema di indecidibilità di Church-Turing, secondo cui non esiste alcun metodo meccanico per decidere se una proposizione sia dimostrabile a partire da dati assiomi, esso dovrebbe porre il problema della scoperta al centro dell'interesse, eppure non lo fa.

5) Il metodo assiomatico separa nettamente la matematica dalle scienze naturali. Ciò appare chiaro dalla posizione di Polya secondo cui la matematica si basa sul ragionamento dimostrativo mentre le scienze naturali si basano sul ragionamento plausibile, dove il ragionamento dimostrativo è sicuro al di là di ogni possibile dubbio ed è conclusivo, mentre il ragionamento plausibile è rischioso, controverso e provvisorio.[4]

6) Il metodo assiomatico banalizza la dimostrazione, in due sensi differenti. In primo luogo, esso rende pura routine la generazione delle dimostrazioni: basta applicare il cosiddetto *algoritmo del British Museum*, che permette di ottenere l'una dopo l'altra tutte le dimostrazioni di un sistema formale. In secondo luogo, il metodo assiomatico non rende conto del fatto che la dimostrazione dà nuova informazione. Ciò è espresso dal cosiddetto *paradosso dell'inferenza*: Se in un'inferenza la conclusione non è contenuta nelle premesse, allora non può essere valida; e se la conclusione non dice nulla di essenzialmente nuovo rispetto alle premesse, allora è inutile; ma la conclusione non può essere vera e nello stesso tempo dire qualcosa di essenzialmente nuovo rispetto alle premesse; perciò le inferenze o non sono valide oppure sono inutili.

7) Il metodo assiomatico concepisce le dimostrazioni come puramente logiche, quindi come ottenute mediante inferenze (come il *modus ponens* $A,A\rightarrow B/B$) indipendenti dal dominio di oggetti considerato, a differenza delle regole non logiche (come la *regola di induzione* $A(0),\forall x(A(x) \rightarrow A(x'))/\forall xA(x)$) la cui validità dipende strettamente dal dominio.

8) Il metodo assiomatico considera la matematica come un sistema chiuso. Per il primo teorema di incompletezza di Gödel il concetto di sistema assiomatico in quanto sistema chiuso è inadeguato per la matematica e quindi dev'essere sostituito con quello di sistema aperto.

9) Il metodo assiomatico conduce a una degenerazione dell'attività matematica. Come osserva Bourbaki esso dà luogo a una proliferazione teratologica di sistemi privi di prevedibili applicazioni, favorisce la parcellizzazione della ricerca, e scioglie lo stretto rapporto tra la matematica e le sue applicazioni alla fisica.[5]

4. Dimostrare come scoprire

All'origine della concezione del dimostrare come scoprire sta la soluzione data da Ippocrate di Chio al problema della quadratura delle lunule. Essa si basa essenzialmente sul ridurre il problema da risolvere a un altro problema, considerato più facilmente solubile: si suppone vera una proposizione e si cercano ipotesi che la rendono vera, poi si cercano ipotesi che rendono vere tali ipotesi, e così via.

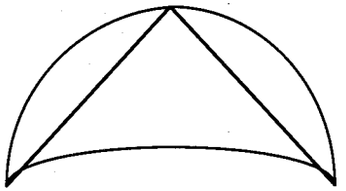
[3] L. Carroll, 'Quello che la Tartaruga disse ad Achille', in D.R. Hofstadter, *Gödel, Escher, Bach*, Milano (Adelphi) 1984, pp. 47-49.

[4] G. POLYA, *Mathematics and plausible reasoning*, vol. I. *Induction and analogy in mathematics*, Princeton (Princeton University Press) 1954, pp. *v-vi*.

[5] N. BOURBAKI, 'L'architecture des mathématiques', in F. Le Lionnais (ed.), *Les grands courants de la pensée mathématique*, Paris (Blanchard) 1962, pp. 35-47.

Per esempio, consideriamo la soluzione data da Ippocrate di Chio al problema della quadratura di una lunula avente la circonferenza esterna eguale a un semicerchio.[6] Essa consiste dei seguenti passi:

1) Si circoscrive un semicerchio intorno a un triangolo rettangolo e isoscele.
2) Intorno alla base del triangolo si circoscrive un segmento di cerchio simile a quelli tagliati fuori dai lati del triangolo. Si ottiene così la seguente figura.



3) Per il teorema di Pitagora il quadrato sulla base del triangolo è eguale alla somma dei quadrati sui lati del triangolo.
4) Si formula l'ipotesi:

   (A) Segmenti di cerchio simili stanno tra loro come i quadrati sulle loro basi.

5) Dall'ipotesi (A) e da 3) segue che il segmento di cerchio circoscritto intorno alla base del triangolo è eguale alla somma dei segmenti di cerchio tagliati fuori dai lati del triangolo.
6) Da 5) segue che, sommando la parte del triangolo che sta sopra il segmento di cerchio circoscritto intorno alla base, con i segmenti di cerchio tagliati fuori dai lati, la lunula sarà eguale al triangolo.
7) A sua volta l'ipotesi (A) dev'essere giustificata da un'altra ipotesi. A tale scopo si formula l'ipotesi:

   (B) Due cerchi stanno tra loro come i quadrati sui loro diametri. E così via.

Il procedimento di Ippocrate di Chio viene elaborato da Platone dando luogo al cosiddetto *metodo delle ipotesi*. Questo consta dei seguenti passi.[7]

1) Per dimostrare una proposizione la si suppone vera.
2) Si formula un'ipotesi che, se vera, implica la verità della proposizione.
3) Se esistono più ipotesi alternative che soddisfano questa condizione, si sceglie tra esse quella che sembra la più salda, cioè quella che meglio si accorda con la proposizione.
4) Si considerano tutte le conseguenze che derivano dall'ipotesi, per vedere se esse si accordano tra di loro, cioè non si contraddicono tra loro, e per vedere se si accordano con i fatti già noti.
5) Se non è così, l'ipotesi dev'essere respinta.
6) Si cerca quindi di giustificare l'ipotesi nello stesso modo, cioè ricorrendo a una nuova ipotesi, e così via.
7) Il processo termina, provvisoriamente, quando si arriva a un'ipotesi giudicata soddisfacente.
8) Tuttavia, questo termine del processo è del tutto provvisorio, dal momento che ogni ipotesi dev'essere giustificata.
9) Per i nostri limiti conoscitivi non possiamo arrivare mai a delle ipotesi ultime, quindi non possiamo esaurire la verità.

Questo lascia aperto il problema della scelta delle ipotesi. Per risolvere tale problema Platone propone due procedimenti.

Il primo mira a sottoporre le ipotesi a un vaglio minuzioso, considerando le loro conseguenze e la loro compatibilità con le conoscenze esistenti. Esso si configura come un dialogo tra due persone, che chiameremo, rispettivamente, *proponente* e *oppositore*. Per dimostrare una data proposizione, il proponente formula un'ipotesi. Il problema di quale ipotesi debba essere scelta inizialmente non è drammatico, perché il procedimento è per tentativi ed errori, e quindi, nella ricerca dell'ipotesi corretta, esso procede per approssimazioni successive. Al limite l'ipotesi iniziale potrebbe essere scelta a caso. L'ipotesi deve soddisfare, però, due requisiti: in primo luogo, dev'essere sufficiente per dimostrare la data proposizione e, in secondo luogo, dev'essere compatibile con le premesse comuni.

Una volta formulata l'ipotesi iniziale, l'oppositore cerca di confutarla, considerando tutte le conseguenze che ne derivano e mostrando che tra esse ve ne sono alcune che si contraddicono tra loro o non si accordano con i fatti già noti. Se il tentativo di confutazione dell'oppositore ha successo, l'ipotesi dovrà essere modificata e, nei casi estremi, abbandonata. Ma il lavoro fatto fino a quel punto, comunque, non sarà stato inutile: infatti il proponente ripercorrerà l'argomentazione che ha portato dalle premesse comuni e dall'ipotesi iniziale fino alla proposizione che contraddice quella che si vuol dimostrare. L'analisi di tale argomentazione rivelerà perché l'ipotesi iniziale è inadeguata, e suggerirà come apportare ad essa modifiche per poter meglio avvicinarsi allo scopo. Come risultato di tale analisi, il proponente formulerà una nuova ipotesi, che sarà un po' meno lontana dall'essere soddisfacente che non l'ipotesi iniziale. Dopodiché il procedimento si ripeterà per la nuova ipotesi, e così via fino ad arrivare a un'ipotesi soddisfacente. Un'ipotesi che voglia essere considerata tale è caratterizzata dal fatto che può essere difficilmente confutata e non può essere eliminata in pratica.

[6] G. CAMBIANO (ed.), *op. cit.*, pp. 85-86.
[7] G. CAMBIANO (ed.), *op. cit.*, pp. 101-108.

Il secondo procedimento mira a trovare le ipotesi. La ricerca si effettua analizzando il problema che si vuole risolvere e mettendolo in relazione con altri fatti. L'analisi scompone il problema nelle sue parti. Per esempio il procedimento attraverso cui si trova una definizione *per suddivisione* comporta un dialogo tra due interlocutori, che si possono ancora chiamare *proponente* e *oppositore*. La funzione del proponente è di proporre via via delle alternative, e quella dell'oppositore è di scegliere tra esse. Le alternative proposte devono esaurire tutti i casi possibili perché si abbia una definizione completa.

5. L'evolutività delle dimostrazioni

Vi sono varie caratteristiche delle dimostrazioni che non sono adeguatamente rispecchiate dalle dimostrazioni costruite col metodo assiomatico e sono meglio modellate da quelle costruite col metodo analitico. Considereremo qui tre di queste caratteristiche, cioè la loro evolutività, plasticità e modularità.

La prima caratteristica delle dimostrazioni matematiche che non è modellata dalle dimostrazioni assiomatiche ma è modellata da quelle basate sul metodo analitico, è data dalla loro evolutività. L'*evolutività* delle dimostrazioni riguarda il fatto che le dimostrazioni non nascono come prodotti finiti, ma si sviluppano per gradi a partire da una rozza ma feconda idea iniziale, che può dirsi il loro *prototipo*. Il prototipo è la matrice dalla quale una dimostrazione prende l'avvio. Le dimostrazioni sono oggetti che si evolvono, passando dai componenti grezzi dell'idea iniziale contenuta nel prototipo, fino ai componenti rifiniti e funzionanti del prodotto finito, attraverso raffinamenti successivi caratterizzati da un grado di astrazione crescente, ciascuno dei quali arricchisce in modo sostanziale l'idea iniziale, e che quindi non sono tutti contenuti implicitamente in essa.

Nel prototipo è inclusa l'idea chiave, la prospettiva illuminante, il *germe* della dimostrazione, ma non nel senso di Frege secondo cui il germe è dato dai principi di un sistema assiomatico (definizioni, assiomi), nei quali la dimostrazione è contenuta come in un seme in cui tutte le varie parti sono ancora inviluppate e nascoste e il cui sviluppo a partire dal seme è dato semplicemente da ripetute applicazioni di inferenze puramente logiche.[8] Il prototipo contiene il germe della dimostrazione piuttosto nel senso che esso ne costituisce il fermento e il lievito, dove questi non vanno confusi col pane e non sono sufficienti per ottenere il pane.

Il principale strumento per trovare un prototipo è dato dall'analisi del problema da risolvere. Il problema può essere visto come un composto che, attraverso l'analisi, viene scomposto nelle sue componenti facendo uso di conoscenze preesistenti. L'analisi chiarisce la reale natura del problema e le sue condizioni di solubilità, e in tal modo fornisce una prima rozza idea della direzione in cui si deve cercare la soluzione, ma l'esistenza di una soluzione

[8] G. FREGE, 'La logica nella matematica', in *Scritti postumi*, Napoli (Bibliopolis) 1986, p. 335.

in generale può essere accertata solo facendo appello a conoscenze disponibili in precedenza. L'analisi, dunque, non è un processo che si limita ad estrarre dal problema l'informazione contenuta più o meno immediatamente in esso. Facendo uso di conoscenze preesistenti, essa può rivelare connessioni che non emergerebbero dalla pura considerazione isolata del problema. Perciò le ipotesi su cui si basa il prototipo della dimostrazione non sono necessariamente scontate, ma possono benissimo risultare moderatamente, o anche fortemente, impreviste.

L'evoluzione successiva della dimostrazione a partire dal prototipo ha luogo per tentativi ed errori. A ogni passo si vagliano le ipotesi formulate e si valutano le ragioni della loro persistente inadeguatezza a risolvere il problema. L'individuazione di tali ragioni in generale è essenziale, e spesso determinante, per ricavarne indicazioni sulle modifiche da apportare alle ipotesi per meglio avvicinarsi allo scopo, cioè per formulare nuove ipotesi che siano un po' meno lontane dall'essere soddisfacenti delle ipotesi esistenti. Dopodiché il procedimento viene ripetuto per le nuove ipotesi, e così via fino a quando si perviene a ipotesi soddisfacenti. Il processo si configura, quindi, come una successione convergente di schemi di dimostrazione che comincia col prototipo e ha come limite la dimostrazione. In questo processo di accostamento per approssimazioni successive alla dimostrazione, l'esame dei tentativi operati in precedenza è essenziale.

Nel risolvere un problema matematico un passo importante, forse addirittura il passo più importante, consiste nell'analizzare tentativi già effettuati in passato, o anche tentativi immaginari, allo scopo di scoprire come mai essi siano andati a vuoto. I tentativi precedenti non si riferiscono necessariamente a quelli fatti da uno stesso matematico, ma possono comprendere anche i tentativi operati da altri matematici. In questo senso si può dire che nessun matematico dovrebbe azzardarsi ad affrontare un serio problema matematico senza prima essersi familiarizzato a fondo con la storia di quel problema, storia che può coincidere con la storia reale o con una storia immaginaria quale un matematico di talento riesce a prefigurarla.

Le dimostrazioni si possono concepire, dunque, come sviluppantesi a partire da un modulo molto semplice, il prototipo, che successivamente si evolve gradualmente diventando sempre più complesso e compiuto, mano a mano che si aggiungono nuove ipotesi o si sottraggono ipotesi che risultino errate. L'evoluzione del modulo iniziale è il risultato di interazioni con altri moduli che stabiliscono connessioni tra il prototipo e altri sistemi di conoscenze matematiche. Tali interazioni modificano il prototipo. Esse, quindi, non sono del tipo di quelle che hanno luogo tra un interruttore elettrico e i suoi utenti. Un interruttore elettrico è un oggetto il cui stato (on/off) è determinato dalle interazioni coi suoi utenti, ma la cui struttura funzionale non è modificata da tali interazioni. Al contrario, le interazioni tra un prototipo e il suo ambiente modificano la stessa struttura funzionale del prototipo, aiutando a

dar corpo all'idea rozza originaria e dandole nuova forma e sostanza.

6. La plasticità delle dimostrazioni

La seconda caratteristica delle dimostrazioni matematiche che non è modellata dalle dimostrazioni assiomatiche ma è modellata dalle dimostrazioni basate sul metodo analitico è data dalla loro plasticità.

La *plasticità* delle dimostrazioni riguarda il fatto che, a differenza delle dimostrazioni assiomatiche, le dimostrazioni che si incontrano nell'effettiva pratica matematica non hanno una struttura rigida ma sono oggetti malleabili, plasmabili e duttili. Il termine del processo di generazione della dimostrazione che comincia col prototipo non consiste in una dimostrazione assiomatica, ma è dato piuttosto da una dimostrazione in cui non si usano né si esplicitano tutti i principi e non si esplicitano tutte le regole di inferenza. Il non esplicitare tutti i principi né tutte le regole di inferenza è essenziale per la fattibilità pratica delle dimostrazioni: richiedere che tutti i principi e tutte le regole siano esplicitati renderebbe la costruzione della dimostrazione così farraginosa e prolissa che essa diverrebbe irrealizzabile già nel caso di teoremi molto elementari.

Se nelle dimostrazioni effettivamente adoperate nella pratica matematica non si usano né si esplicitano tutti i principi né tutte le regole di inferenza, questo è possibile perché ciò che non viene esplicitato in esse viene assunto come conoscenza *implicita*, cioè come conoscenza che già si possiede e che non è necessario esplicitare. Tale conoscenza implicita rappresenta una *conoscenza di sfondo* condivisa. Si può istituire un parallelo tra la comprensione di una dimostrazione e la comprensione di un testo. Quest'ultima ovviamente presuppone un numero più o meno elevato di conoscenze di sfondo condivise. Nel leggere un testo noi facciamo tacitamente certe assunzioni implicite concernenti, l'uso, la notazione e l'interpretazione dei concetti che occorrono in esso, il modo in cui sono formulate le argomentazioni, e l'uso di figure retoriche come mezzo per suffragare le argomentazioni, assunzioni la cui esplicitazione renderebbe il testo così inutilmente complesso da renderlo rapidamente incomprensibile. Analogamente, nel leggere una dimostrazione, noi facciamo tacitamente certe assunzioni implicite concernenti l'uso, la notazione e l'interpretazione dei concetti matematici che occorrono in essa (variabili, funzioni, ecc.), il modo in cui sono formulate le argomentazioni e l'uso di figure sulla carta come mezzo per suffragare le argomentazioni. Queste assunzioni implicite, o conoscenza di sfondo, sono ciò che consente a un testo o a una dimostrazione di essere compresa e comunicata, mantenendosi entro limiti ragionevoli di complessità.

7. La modularità delle dimostrazioni

La terza caratteristica delle dimostrazioni matematiche che non è modellata dalle dimostrazioni assiomatiche ma è modellata dalle dimostrazioni basate sul metodo analitico è data dalla loro modularità. La *modularità* della dimostrazione consiste nel fatto che la dimostrazione può essere concepita come ottenuta per composizione di moduli distinti. Vi sono due tipi di modularità della dimostrazione: la modularità per cooperazione e la modularità per negoziazione.

La *modularità per cooperazione* si riferisce a una situazione in cui tra i moduli esista un rapporto non conflittuale, bensì di cooperazione. La cooperazione è un processo in base a cui più moduli svolgono, in una stessa dimostrazione, ruoli distinti ma mutuamente dipendenti e in armonia tra loro. Questo può essere chiarito più facilmente ricorrendo a un esempio non matematico. Supponiamo che una società fissi una quotazione per un suo prodotto, impegnandosi a fornirlo al prezzo fissato da quella quotazione. Un cliente invia un ordine di acquisto per il prodotto impegnandosi a pagarlo. Il ruolo svolto dal cliente col suo impegno a pagare per il prodotto dipende dal ruolo svolto dalla società col suo impegno a fornirlo, al prezzo fissato dalla quotazione. Il ruolo svolto dalla società col suo impegno a fornire il prodotto dipende dal ruolo svolto dal cliente col suo impegno a pagare per il prodotto il prezzo specificato nell'ordine di acquisto. Dunque la quotazione e l'ordine di acquisto formalizzano la cooperazione tra la società e il cliente.

Questa situazione può essere trasferita alle dimostrazioni. Quando, per dimostrare una proposizione matematica, si usano parecchie ipotesi distinte ma mutuamente dipendenti, si instaura un processo di cooperazione tra esse. Il ruolo svolto da ciascuna ipotesi nell'offrire un contributo per dimostrare la proposizione, dipende dal ruolo svolto dalle ipotesi rimanenti nell'accettarlo. L'offerta di ciascuna ipotesi e la sua accettazione da parte delle rimanenti formalizzano la cooperazione tra le ipotesi. In base a essa, le ipotesi formano un sistema integrato in cui il ruolo di ciascuna di esse dipende da quello delle altre.

La *modularità per negoziazione* si riferisce, invece, a una situazione in cui tra i moduli esiste un rapporto di conflitto, che dev'essere risolto attraverso una negoziazione tra i moduli. Ciò corrisponde alla situazione che si presenta nel metodo analitico quando le nuove ipotesi introdotte sono incompatibili con la conoscenza esistente. Anche qui questo può essere chiarito più facilmente ricorrendo a un esempio non matematico. Supponiamo che una società debba fissare una quotazione per un suo prodotto. Uno dei suoi dirigenti afferma: dobbiamo fissare la quotazione a un livello più elevato di quello della concorrenza perché un profitto maggiore è preferibile a un profitto minore. Un altro dirigente afferma: no, dobbiamo fissare la quotazione a un livello più basso di quello della concorrenza perché una quota di mercato maggiore è preferibile a una quota di mercato minore. Le affermazioni dei due dirigenti forniscono due diverse ipotesi per risolvere il problema, ma tali ipotesi sono in

conflitto. Fissare la quotazione a un livello più elevato di quello della concorrenza aumenterebbe il profitto ma diminuirebbe la quota di mercato. Viceversa, fissare la quotazione a un livello più basso di quello della concorrenza aumenterebbe la quota di mercato ma diminuirebbe il profitto. Problemi di questo tipo vengono di solito risolti attraverso l'algebra lineare, ma ciò non deve occultare il fatto che questo tipo di soluzione è parziale perché non tiene conto di tutti i fattori in gioco. La difficoltà implicita in tali problemi, al di là della loro risolubilità in casi particolari con tecniche tradizionali, è di tipo logico, e una soluzione in generale richiede l'elaborazione di tecniche logiche adeguate.

La situazione ora descritta può essere trasferita alle dimostrazioni. Quando, per dimostrare una proposizione matematica, si considerano nuove ipotesi che risultano in conflitto con le conoscenze esistenti, il conflitto può essere superato solo attraverso una negoziazione. Se questa ha successo, essa dà luogo a un insieme di ipotesi più elevato in cui il conflitto viene composto. Chiaramente il conflitto non potrebbe essere superato usando solo la nozione assiomatica di dimostrazione. Mentre tale nozione può essere idonea per determinare che le ipotesi sono in conflitto con le conoscenze esistenti, essa certamente è inadeguata per formulare e revisionare ipotesi. Quello che realmente si richiede è un dibattito tra scelte alternative e un confronto dell'evidenza a favore di ciascuna di esse.

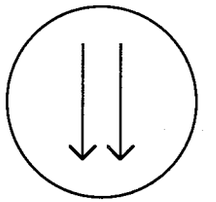
8. I sistemi aperti

Per implementare queste caratteristiche delle dimostrazioni (evolutività, plasticità, modularità) non bastano i sistemi assiomatici, i quali sono *sistemi chiusi* in cui principi e regole di inferenza sono dati una volta per tutte, ma occorre ricorrere a dei *sistemi aperti*.[9] Una fondamentale caratteristica dei sistemi aperti è che in essi non si assume che tutta la conoscenza sia concentrata in un unico sistema, bensì si ammette che essa possa essere distribuita tra più sistemi, e quindi che la comunicazione tra sistemi possa svolgere un ruolo essenziale. Inoltre si ammette che i sistemi possano evolversi, cioè che ad ogni passo si possono aggiungere o eliminare principi o regole di inferenza

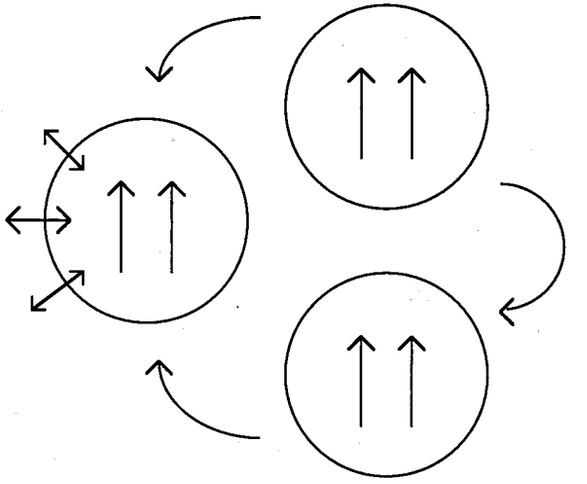
La differenza tra i sistemi assiomatici (chiusi) e i sistemi aperti può essere rappresentata pittoricamente nel modo seguente. I sistemi assiomatici possono essere rappresentati come cerchi aventi un diametro fisso, cioè il cui diametro non può cambiare. Tutti i processi nei sistemi chiusi, che sono costituiti da derivazioni logiche degli assiomi, si svolgono all'interno del cerchio: non si ammette alcun flusso di informazione dall'esterno verso l'interno, né dall'interno verso l'esterno. Inoltre il diametro del cerchio non può cambiare, dove i cambiamenti di diametro rappresentano l'aggiunta o l'eliminazione di

[9] Sulle nozioni di sistema chiuso e sistema aperto cfr. C. CELLUCCI, 'From closed to open systems', in J. Czermak (ed.), *Philosophy of Mathematics: Proceedings of the 15th International Wittgenstein-Symposium*, Wien (Hölder-Pichler-Tempsky) 1993, pp. 206-220.

assiomi o regole di inferenza.



I sistemi aperti, invece, possono essere rappresentati come insiemi di cerchi con diametro variabile: i cambiamenti di diametro corrispondono all'aggiunta o all'eliminazione di ipotesi. Tali cambiamento possono avvenire o dall'interno, come risultato dell'analisi del problema da risolvere, o dall'esterno, come risultato del ricevere informazione da altri sistemi.



I sistemi aperti permettono di modellare, meglio di quelli chiusi, il processo reale di costruzione delle dimostrazioni. Al contrario dei sistemi assiomatici, essi possono trattare in modo naturale conoscenze matematiche in evoluzione perché sono essi stessi dei sistemi in evoluzione le cui ipotesi possono essere modificate ad ogni passo. Essi forniscono l'ambito appropriato per considerare la formazione delle ipotesi perché permettono di riesaminare e di modificare dimostrazioni inadeguate. Inoltre, essi possono trattare in modo appropriato interazioni dinamiche tra sistemi di conoscenze matematiche perché la comunicazione fornisce un mezzo flessibile per stabilire una rete di relazioni dinamiche tra sistemi differenti. Naturalmente i sistemi aperti sono molto più difficili da descrivere dei sistemi assiomatici, ma questa loro complessità è un prezzo che vale la pena di pagare in cambio della migliore approssimazione che essi forniscono alla realtà matematica.

La Deduzione: Esperienze didattiche

CARLO MARCHINI
Dipartimento di Matematica
Università di Parma
Via D'Azeglio 85/A
43100 PARMA

Introduzione

Nella seconda fase del corso di aggiornamento, quella di richiamo, tenutasi presso Otranto, si sono confrontate esperienze condotte in classe da alcuni dei partecipanti al corso al fine di mettere in luce aspetti salienti ed eventualmente giungere ad una proposta didattica sul tema della deduzione e degli schemi di deduzione. E' stata evidenziata la difficoltà di individuare un percorso didattico adeguato e motivante, e nel contempo, non alieno dagli argomenti di Matematica, per non creare una sovrastruttura a sé stante, interessante forse, ma poco significativa.

Quanto viene qui presentato è in gran parte frutto dell'attività svolta dai partecipanti, i professori. Bordoni, Ricci, Paola, Ciceri, Ortolan, dei colloqui avuti con essi e con la prof. Margiotta in quella sede ed in altre occasioni. Ho aggiunto una riflessione personale su argomenti didattici connessi e mi sono ampiamente avvalso della tesina redatta dalla prof. Iaderosa per il corso di perfezionamento in Didattica della Matematica, dell'Università Cattolica di Brescia.

Il materiale che segue è così organizzato: ho iniziato con un breve approccio al problema della valutazione, perché è, a mio parere, un punto essenziale (e spinoso) della didattica della Logica; ho presentato alcune attività finalizzate ad una crescita della competenza grammaticale e sintattica, diverse da quelle presentate dai partecipanti al corso. In seguito si introduce un breve inquadramento di carattere pedagogico per comprendere le difficoltà della dimostrazione, confrontata con l'argomentazione. Sono poi presentate alcune attività proposte dai partecipanti al corso, seguite da una breve conclusione.

1. Valutazione, un problema didattico

In Logica, come in ogni altra innovazione in campo didattico, il tema della valutazione è assai importante. La letteratura inglese distingue tra *assessment* e *evaluation.* [Ci]. Con una traduzione abbastanza infedele, ma suggestiva, si può tradurre la prima come *valutazione formativa*, la seconda come *valutazione sommativa.*

Nella nostra disciplina entrambe le forme di valutazione sembrano essere di difficile individuazione, eppure da sempre gli studenti vengono valutati in Logica, anche quando l'argomento è di altra natura, matematica o no. Spesso, dopo una risposta scorretta, l'insegnante invita lo studente a ragionare, cioè a collegare ed organizzare il materiale che conosce in un quadro coerente. Il voto finale, risultato della valutazione sommativa, tiene conto non solo della capacità di memorizzazione e della quantità di nozioni apprese, ma anche della duttilità e della competenza deduttiva dimostrata. La valutazione formativa è sempre presente nell'azione di un buon insegnante: ella/egli è in grado di cogliere dalle risposte, le eventuali lacune dei suoi allievi, singolarmente o del gruppo-classe, ed appronta nuove strategie per ottenere il fine prefissato.

Ma in questi esempi, la Logica entra in modo induttivo, non esplicitamente, e quindi una valutazione sui suoi aspetti specifici non viene esercitata. Ora invece che il nuovo tema acquista cittadinanza in classe, è bene approntare strumenti valutativi idonei. Quelli che qui vengono proposti non sono stati sperimentati, quindi sono assai discutibili. Solo la sensibilità didattica degli insegnanti potrà dire col tempo se sono adeguati allo scopo.

La mancanza di un indirizzo nella valutazione è però assai dannosa. Molto spesso sono state avanzate proposte anche assai ricche ed articolate nel campo dell'innovazione didattica. Senza un adeguato strumento valutativo esse rischiano di restare inapplicate, in quanto se non forniscono al docente la possibilità di analizzare le prestazioni degli studenti, né di comprendere se il percorso innovativo ha portato ad un effettivo incremento qualitativo della scolaresca, non permettono neppure al docente un'auto-analisi che conforti la sua attività, né gli forniscono adeguati impulsi per un eventuale aggiustamento del tiro.

2. Attività in classe

La scuola superiore ha il compito di portare a compimento la maturazione culturale degli studenti, attingendo al patrimonio delle conoscenze apprese negli ordini scolastici precedenti. Per quanto riguarda la Logica—se non verranno introdotti importanti mutamenti—dopo un tirocinio abbastanza completo e complesso nelle scuole elementari, i programmi delle scuole medie prevedono poco: solo accenni ed anche su temi limitati e discutibili. Poi, ad iniziare dal biennio, si richiedono competenze complesse ed articolate, che però non sono basate su acquisizioni graduali. Tale situazione fa sì che spesso il docente delle scuole superiori si vede

costretto a ricominciare da "zero". Quanto segue può forse sembrare poco idoneo, trattandosi di semplici esercizi, ma la loro utilizzazione si può rivelare preziosa.

2.1. Espressioni letterali

Nella scuola media si introduce il calcolo algebrico facendolo precedere da esercizi il cui scopo è quello di indurre familiarità con le scritture. Ad esempio si fornisce un'espressione algebrica, come la seguente

$$a^2 + (a - 1)(a - 3) + 7$$

si aggiunge la consegna di calcolarne il valore quando si pone $a = -1$, oppure $a = \frac{\sqrt{5}}{3}\pi$, eccetera. Queste attività possono essere iniziate in questa forma semplice per proseguire verso aspetti più complessi, ad esempio invitando a calcolare l'espressione che si ottiene quando si pone $a = a + 1$, oppure $a = \sqrt{\frac{a+b}{a-b}}$.

Tra le due proposte c'è una differenza importante, almeno dal punto di vista logico: nel primo esercizio è evidente l'uso della espressione come funzione polinomiale (o funzione razionale) da valutare su numeri interi o reali, quindi ci si muove in un ambito semantico; nel secondo caso si privilegia l'aspetto grammaticale: la sostituzione applicata su termini fornisce ancora termini. In entrambi i casi il ruolo delle sostituzioni è rilevante.

Un'interessante proposta di lavoro interdisciplinare è indicata in [MP2]. Essa è stata formulata come intervento nella scuola media, tuttavia è utilizzabile anche nel biennio, sfruttando appunto le conoscenze che dovrebbero essere acquisite dall'ordine scolare precedente. In [MP1] la stessa autrice propone alcune schede utilizzabili in classe per verificare il livello di apprendimento e di conoscenza sul tema.

In [M1] si forniscono altri esempi ed un criterio per valutare gli elaborati, basandosi sulla complessità della sostituzione utilizzata. Non è però approfondito a sufficienza il tema della differenza tra sostituzioni "semantiche" e "grammaticali" per verificare se i due aspetti portino a prestazioni diverse.

Attraverso tali attività è possibile mettere in luce eventuali difficoltà degli allievi sui concetti di *costante*, *variabile* e *parametro*. Questo può essere assai importante per le successive applicazioni nella Matematica; un tirocinio sul tema potrebbe essere utilizzato negli argomenti in cui il tema delle sostituzioni è presente, come la risoluzione dei sistemi: sarebbe interessante verificare se esso porti ad un miglioramento delle prestazioni.

2.2. Giochi di carte

E' ben nota la passione degli studenti per il gioco delle carte durante le ore di lezione. Si tratta di trasformare questo passatempo fastidioso (per l'insegnante) in una occasione di approfondimento logico. Il gioco delle carte si presta come *ambiente logico,* nel senso di Arzarello, perché le regole sono abbastanza semplici e note a tutti. Si può iniziare a presentare situazioni di gioco, ad esempio mostrando una o più mani, nello stile dei problemi che compaiono su giornali specializzati ed anche sulla Settimana Enigmistica, però per giochi semplici come la briscola o la scopa, non il bridge. La difficoltà maggiore per l'insegnante credo sia quello di individuare un tipo di rappresentazione delle situazioni, che non sia il disegno. In questa fase ci si muove a livello locale, prendendo in considerazione una mano per volta, non l'intera partita.

Con un'attività di discussione collettiva si può tentare di dare esplicitamente le regole del gioco. In questa fase ci si avvicina moltissimo ai sistemi deduttivi. Vengono infatti a cadere pregiudizi semantici sulla "verità" o "validità" delle regole date. Una fase successiva è quella di fornire le carte di quattro giocatori nelle fasi finali e richiedere una strategia che porti a massimizzare il risultato. In quest'ultima richiesta entrano varie considerazioni: si tratta di fornire un'applicazione globale dei sistemi deduttivi individuati, inoltre sono presenti aspetti semantici, vale a dire come vengono valutate le carte: un fante a briscola vale 2 punti, a tresette $\frac{1}{3}$, a scopa come una carta, se non è di danari. Se giochi "classici" sembrano troppo complessi, si può iniziare dai solitari. Questi giochi permettono variazioni delle regole, senza correre troppi rischi di trovarsi in condizioni contraddittorie. Le situazioni di conflitto sono didatticamente opportune per illustrare il tema della correttezza delle regole, e di conseguenza quello della coerenza. Se ad esempio in un torneo di briscola a quattro si introduce la condizione che chi possiede il sette di danari può decidere di cambiare compagno, dalla mano in cui possiede tale carta, può nascere un problema di conteggio dei punti per decidere quale delle due coppie sia la vincitrice del torneo.

Assieme ai giochi di carte sono possibili variazioni sul tema, con pedine, scacchiere, bastoncini disposti in varie configurazioni, eccetera. Il docente, che a sua volta è stato studente, possiede senza dubbio un campionario di conoscenze sull'argomento che gli permette una scelta opportuna.

In questo ambito la valutazione potrebbe essere effettuata non solo sulla base dell'individuazione della strategia migliore, ma anche sui tempi di reazione, come in pratica avviene in una partita reale.

2.3. Giochi logici

Talvolta su riviste di enigmistica si trovano alcuni giochi etichettati come logici. Alcuni riguardano relazioni, altri vere e proprie inferenze. In quello che qui si riporta, è tratto con qualche modifica inessenziale da [Ia]:

Antonio, Bruno, Carlo e Davide sono medico, avvocato, tabaccaio e ingegnere, ma non necessariamente nell'ordine indicato. Sapendo che

- Davide ha sposato la sorella dell'ingegnere, che a sua volta ha sposato la sorella del medico;
- Carlo è figlio unico, non si è sposato ed è molto amico del tabaccaio;
- Bruno è amico del fratello dell'ingegnere ma non conosce nessuno dei suoi parenti;

accoppiare i nomi con le professioni.

Testi ricchi di esercizi di questo tipo si possono reperire settimanalmente nelle edicole. Il libro del Varga, [V], ha parti da cui possono essere desunti esercizi di questo tipo. I testi di R. Smullyan ricordati nel suo intervento da Arzarello sono una messe quasi inesauribile di problemi logici analoghi ed anche più profondi.

Questi giochi logici spesso mescolano semantica e sintassi, almeno in una fase di gestione "intuitiva". Ad esempio nell'esercizio sopra indicato dalla frase «Davide ha sposato la sorella dell'ingegnere, che a sua volta ha sposato la sorella del medico» si può concludere che Davide non è ingegnere, perché è vietato l'incesto. Ora questa prescrizione che ha origine dalla vita comune non è una condizione di carattere logico, è un assioma extralogico. Ovviamente ci sono altre condizioni che andrebbero esplicitate quali che la relazione «... ha sposato...» non è riflessiva, è simmetrica, ma non ammette (meglio non ammetteva fino a poco tempo fa) come argomenti due individui dello stesso sesso. A questo punto l'esercizio innocente va trasformandosi in una discussione che ha attinenza con ... l'educazione sessuale, argomento per gli studenti forse più interessante della Logica Matematica. Come si vede l'esplicitazione dei sottintesi e delle condizioni implicite nel testo, rischia di trasformare questo esercizio in una esercitazione assai lunga con connotati assai diversi da quelli consueti nelle ore di Matematica.

Iaderosa propone la seguente valutazione in decimi:

- comprensione della soluzione attraverso al compilazione di una tabella: **5**;
- verbalizzazione corretta in una forma che evidenzi i passaggi deduttivi: **2,5**;
- formulazione corretta di ogni proposizione atomica: **0,1**;
- formulazione corretta di ogni proposizione in cui compaia un connettivo: **0,5**.

Il punteggio conseguito potrebbe anche superare **10**, ma è comunque una scala. Di questo schema valutativo è interessante osservare il rilievo dato alla complessità costituita dalla presenza dei connettivi, anche perché questo attiva procedimenti sintattici di riduzione alle pro-

posizioni più semplici, mentre una proposizione atomica viene valutata solo semanticamente.

3. Deduzione ed argomentazione

La distinzione tra argomentazione e dimostrazione è stata oggetto di numerosi convegni ed articoli di ricerca didattica. Una parola chiara e definitiva non è però, a mio parere, ancora stata detta, anche perché è assai difficile definire una tale attività complessa e multiforme. Nel suo intervento di Otranto, Arzarello ha confrontato argomentazione e dimostrazione, mettendone in luce l'impatto psicologico: l'argomentazione è vissuta dallo studente come lo strumento per convincere, la dimostrazione gli appare più come un'esercitazione retorica eventualmente da imparare a ripetere, non lo strumento per stabilire e comunicare risultati.

In Logica si è giunti ad una definizione della dimostrazione, ma la definizione stessa è stata ed è tuttora dibattuta sia sotto l'aspetto tecnico che epistemologico. In un modo grossolano e scorretto, ma suggestivo, l'argomentazione è vicina alla semantica, la dimostrazione ha connotati sintattico. Solitamente si pone scarsa attenzione alla distinzione tra i due livelli e ciò non facilita l'apprendimento della dimostrazione. Alcuni ricercatori francesi, citati in bibliografia, hanno avanzato proposte in tal senso. La loro indagine prende spunto da un'analisi approfondita ed attenta dell'evoluzione del processo dimostrativo nell'allievo, partendo dall'uso della figura geometrica. Essi distinguono in tale utilizzazione diversi livelli di consapevolezza, di rigore e di astrazione, e questo quando ancora l'uso dell'intuizione spaziale legata al disegno è assai forte. In tale situazione l'allievo è più portato ad indurre relazioni generali dalle proprietà della figura, piuttosto che dedurre dagli assiomi le conseguenze volute.

L'interazione sociale, assai bene studiata da [B], è strumento per portare gli studenti, attraverso la discussione collettiva, a convincersi delle proprie argomentazioni—a patto che il docente sappia scegliere argomenti che portino alla discussione, e che i ruoli all'interno del gruppo-classe lo permettano (ad esempio la mancanza di un leader assoluto, al cui parere gli studenti si conformano immediatamente). Quando ha luogo una discussione collettiva su temi di questo tipo è facile che le linee del discorso ripercorrano da vicino i vari passi e momenti storici in cui l'argomento stesso è stato affrontato. Non è detto però che senza sapienti e discreti interventi del docente, si giunga in ogni caso alla utilizzazione del formalismo, così come utilizzato in modo standard dalla matematica contemporanea. Tuttavia l'insegnante deve essere presente nella fase di stipula delle nuove regole sulla base delle quali stabilire, sia pure socialmente, il sapere, per far comprendere il ruolo dell'accettazione coerente delle regole scelte.

Risulta interessante ed importante la distinzione di Balacheff sui seguenti termini, spesso usati come sinonimi: *spiegazione*, *prova*, *dimostrazione*, *ragionamento*, *processo di validazione*. Nella *spiegazione*, il soggetto tenta di chiarire, prima di tutto a sé, la validità di una affermazione utilizzando le proprie conoscenze e seguendo proprie regole di decisione. Quando un discorso che assicuri la validità di un'affermazione viene condiviso da altri, esso (discorso) costituisce una *prova*. Ma tale condivisione può essere erronea, anche se condivisa nel gruppo-classe. Ben diversa è la *dimostrazione*, tipica del procedere matematico. Né è corretto identificare queste fasi col *ragionamento*, che in [B] è visto come il particolare momento in cui il soggetto giunge all'intuizione. Quando il ragionamento viene esplicitato attraverso attività verbali e simboliche, si è in presenza del *processo di validazione*. E' solo in questa forma che il ragionamento può essere condiviso in quanto si presenta come un discorso articolato e, appunto, comunicabile. Con questa griglia di lettura dei vari momenti comunque presenti nell'attività argomentativa-dimostrativa, le prove degli studenti possono essere valutate in modo da metterne in evidenza i diversi gradi di astrazione, distinguendo tra *congetture tratte dall'esame di pochi casi*, in cui è assente il processo di validazione, e *congetture tratte dall'esame di qualche caso, poi messe alla prova* su problemi simili. Questo secondo caso testimonia un maggiore avvicinamento alla generalizzazione. La fase più matura prevede *esperienze mentali* in cui il soggetto passa dalle prove effettive, a prove di carattere intellettuale, testimoni di un'interiorizzazione di esperienze particolari su casi singoli. L'uso consapevole di teorie formalizzate o no dà luogo alle *costruzioni intellettuali* .

La tesi di Balacheff si sofferma in dettaglio sul ruolo del controesempio nell'evoluzione del pensiero. Un controesempio si pone nella fase argomentativa, mediante esso si può giungere a confutare una congettura, esso rappresenta un punto centrale nelle ricerca di una prova o di una dimostrazione; inoltre esso favorisce l'apprendimento di un concetto, mostrandone limiti di applicabilità. Quello della contrapposizione dovrebbe essere un espediente didattico sempre presente, soprattutto quando i concetti sono complessi, in modo da fare acquisire il concetto nei suoi confini più precisi. La presenza di un controesempio può portare, nei casi più sfavorevoli, all'abbandono della congettura e di conseguenza all'abbandono del problema, nel caso più favorevole esso indica la necessità di intraprendere una nuova strada, per giungere alla soluzione per una via alternativa. In questo secondo caso lo studente può modificare la congettura, ritornando al problema che l'aveva ispirata, cercando una diversa formulazione del problema stesso, mutando, ad esempio, l'importanza attribuita da alcuni concetti ed oggetti matematici, oppure modificando il dominio considerato considerando un altro esempio. La congettura può anche venire scomposta in "sottocongetture" per analizzare alla luce del controesempio, quale parte della congettura stessa causa la difficoltà. Lo studente può essere portato a rifiutare il controesempio, ma anche a ricontrollare le definizioni che utilizza verificandole nel caso, operando un controllo sulla confutazione.

Di fronte a questa complessità è chiaro che in fase di apprendimento, la valutazione deve essere più formativa che sommativa: gli errori devono essere analizzati per comprendere il livello di sviluppo del pensiero e dell'acquisizione dell'astrazione. Gli errori più comuni sono, in gran parte, passi obbligati che fanno giungere alla maturazione.

Rouche in [R] soffermandosi sull'uso della dimostrazione in Geometria conferisce un connotato positivo alle prove che l'allievo fornisce a partire dall'evidenza della figura. Sarà poi compito dell'insegnante far mutare il modo di fornire le prove, seguendo l'evoluzione dei significati dei contenuti matematici. Così non ha senso valutare sempre nello stesso modo una stessa prestazione scolastica, perché le accezioni di prova e dimostrazione sono molteplici e sono giustificate dal contesto stesso e soprattutto dalle conoscenze matematiche già acquisite. In tal modo la giustificazione di una proprietà geometrica basata sull'evidenza, reca in sé già alcuni assiomi forse padroneggiati in maniera inconsapevole. La cosiddetta verifica intuitiva spesso scambiata con la dimostrazione, o meglio sostitutiva della dimostrazione, può essere accettata solo come primo gradino del percorso che porterà alla dimostrazione. L'analisi della figura dovrà portare a vedere nella figura stessa le relazioni "generiche" che permetteranno di estendere le proprietà osservate ad una classe più ampia di figure dello stesso tipo. Questa fase di evidenza consapevole, una sorta di ricorso a modelli mentali, viene raggiunta quando le proprietà viene correttamente "realizzata" in un caso particolare e quando è possibile ragionare sull'ente generico, in modo da racchiudere in esso tutti i casi possibili. In fase di valutazione formativa l'insegnante deve saggiare quanto l'alunno abbia interiorizzato le conoscenze per rielaborarle in modo personale. Anche la fase "induttiva" dall'esempio alla classe di figure dello stesso tipo, deve essere compresa e seguita, si possono presentare false "induzioni". Queste sono da valorizzare per sollevare contraddizioni che permettano di fare comprendere appieno l'esigenza della dimostrazione. Senza questo passaggio attraverso l'errore, la dimostrazione rimane, come si diceva prima una fase retorica che può essere solo appresa a memoria. E' comunque assai difficile, didatticamente, fare accettare che la dimostrazione (in Geometria) serva a fornire una validazione di quanto è osservabile dalla figura, che può, per altro, essere scoperto con procedimenti argomentativi. Negli ambiti algebrici, mancando dell'evidenza intuitiva della figura, il ruolo della argomentazione è più sottile, spesso viene confuso con la verifica e la dimostrazione con il calcolo.

Concludo questa parte commentando alcuni punti interessanti tratti dall'articolo [D]. L'autore mette in luce che l'organizzazione del discorso è diversa nelle due fasi qui messe a confronto. Nell'argomentazione si possono cogliere passaggi di natura "inferenziale" da una proposizione all'altra in cui si utilizzano prevalentemente il metalinguaggio ed il contenuto semantico delle proposizioni coinvolte. Lo "stato operatorio" delle proposizioni rimane fisso, in relazione al suo contenuto: opposizione, omonimia, particolarizzazione, ecc. I connettori

presenti, non sempre sostituibili con i connettivi logici, servono ad esplicitare il contenuto della relazione tra due proposizioni (conseguenza, opposizione, giustificazione,...) dato che le proposizioni coinvolte contano solo per il loro contenuto. Nel discorso argomentativo le proposizioni vengono concatenate come in un discorso dialogico, le frasi successive si aggiungono alle precedenti curando una sorta di coerenza tematica globale. Per contro la dimostrazione privilegia lo stato operatorio delle proposizioni, analizzandone la struttura linguistica, sulla base della quale si procede nella inferenza, inoltre le proposizioni di per sé acquistano valori diversi a seconda che siano premesse o conclusioni dei passaggi considerati, quindi il ruolo della proposizione cambia a seconda delle situazioni. La concatenazione di proposizioni assume l'aspetto di un "calcolo".

La proposta di Duval è quella di fare apprendere la dimostrazione insegnando a distinguere tra argomentazione e dimostrazione, mettendone in evidenza le peculiarità e le differenze.

4. Esperienze dei corsisti

I materiali prodotti sono stati provati in IV e V Ginnasio, III, IV Liceo scientifico, in due I Liceo classico. I tempi dedicati ad essi sono compresi tra le 4 e le 10 ore. L'esame di queste proposte, per altro assai diverse, ha evidenziato alcuni aspetti rilevanti.

4.1. Per esperienza comune dei docenti, gli studenti hanno difficoltà a elaborare autonomamente dimostrazioni, sia nel biennio che nel triennio. Un'aspirazione degli allievi, i quali così riprendono senza saperlo idee di alcuni matematici del Rinascimento, è che esista un metodo unitario che permetta di dimostrare ogni teorema. Teoremi come quello di Church garantiscono che procedimenti decidibili per trovare le dimostrazioni non esistono. In realtà la difficoltà permane anche sapendo che un certo enunciato è un teorema, garanzia che per gli studenti è rappresentata dal fatto che gli enunciati da dimostrare sono scritti sul libro o sono presentati dal docente. Questo principio di autorità, cui volentieri gli studenti si conformano, contrasta con l'aspirazione della scuola volta alla formazione di uno spirito critico.

Dalle sperimentazioni effettuate è emerso che c'è negli studenti l'esigenza di verificare la correttezza della regola e la verità delle regole, condizione indispensabile questa, secondo gli allievi, per poter applicare la regola stessa. Forse grazie a una didattica precedente poco attenta alla differenza tra sintassi e semantica, vengono spesso identificate "verità" e "dimostrazione". Si noti che mentre gli studenti accettano che giochi di carte o di scacchi debbano avere regole ben precise, e non necessariamente "vere", quando si passa alla Matematica non viene ritenuto indispensabile, come se il rigore fosse connaturato con la disciplina e le regole di inferenza fossero "naturali"; gli studenti palesano la necessità di trarre conclusioni da teoremi precedenti o da assiomi, non dalle regole, perché troppo ovvie.

Sono poi consuete le confusioni tra tesi e ipotesi e la riduzione della dimostrazione ad esempi, che sostituiscono volentieri il fatto sintattico.

4.2. La presentazione delle attività è avvenuta privilegiando la gestione discorsiva collettiva, piuttosto che la presentazione frontale. E' ovvio che questa strategia, assai usata e proficua in ordini scolastici inferiori, può essere un valido mezzo costruttivo del sapere, sicuramente è quello che necessita di tempi più lunghi. Questa metodologia di presentazione ha poi l'effetto di uno straniamento rispetto alle consuete attività che si svolgono in classe durante le lezioni di Matematica. Gli studenti possono pensare che in queste ore non ci sia da prendere appunti, oppure che quanto viene presentato poi non formerà oggetto di una verifica scritta o orale. In tal modo l'interesse sollevato dalla novità del tema può essere di scarsa durata. In un'esperienza si è ricorso alla lezione frontale per fornire le regole della deduzione naturale.

4.3. I docenti sono abbastanza concordi nel formulare un giudizio complessivamente negativo sulla "didatticità" dell'argomento, o meglio sull'interesse da esso sollevato nella scolaresca, dopo un primo momento di partecipazione attiva. Nell'insegnamento-apprendimento della dimostrazione entrano in conflitto due fattori contraddittori: il matematico lavora spesso nella dimostrazione, concentrando vari passaggi in uno solo, facendo riferimento a risultati di varie teorie e di situazioni che sono ambiguamente poste a cavallo di sintassi e semantica. Egli fa spesso uso di teoremi precedentemente dimostrati. L'analisi logica della dimostrazione è invece, almeno nei suoi primi passi, estremamente dettagliata, tende a presentare tutti i passaggi, essa è incentrata sulle regole utilizzate, non sulle proprietà matematiche che vengono di volta in volta utilizzate nella dimostrazione. I docenti hanno cercato di mettere in luce queste articolazioni ottenendo, quasi unanimemente, una irritata sensazione di fastidio. Prevale infatti negli studenti un atteggiamento "utilitaristico" legato più alle proprietà matematiche degli enti coinvolti (ed ai possibili argomenti di verifiche), che interesse nello strumentario logico che viene spesso sottinteso. Così logica diviene sinonimo di noia. L'atteggiamento di studenti (e docenti) deve mutare assai se si vuole giungere alla riflessione metacognitiva in cui la nostra disciplina ha un ruolo predominante. Si può dire che solo gli importanti risultati della Logica giustificano un tirocinio sull'argomento che ha le stesse "bellezze" del solfeggio, ma questo non basta. Riaffiora qui il tema della valutazione cui è stato dedicato il primo paragrafo.

4.4. Gli studenti che hanno palesato difficoltà nella distinzione tra verità e dimostrabilità, hanno stentato nel riconoscere l'esigenza di regole sintattiche. Questo può essere motivato e giustificato in più modi. Qualche responsabilità l'hanno proprio gli esercizi di logica, almeno a livello iniziale, ad esempio l'insistenza sugli aspetti semantici del calcolo delle proposizioni. Alcune sperimentazioni sono incentrate su questi aspetti, con attività aventi

connotati grammaticali e semantici. E' in esse predominante il problema della negazione e dello scambio che essa opera con connettivi e quantificatori. Qualche esempio esplicito:

- Negare le seguenti frasi:

a) Tutti i ragazzi vanno a scuola .....
b) Nessun ragazzo va a scuola .....
c) Qualche ragazzo va a scuola .....

A detta del docente sperimentatore, 1 o 2 studenti al più danno risposte corrette a quesiti di questo tipo. La constatazione mostra bene la distanza che c'è da compiere per far raggiungere un livello logico accettabile. D'altra parte capita spesso di vedere anche su testi "aggiornati" con l'aggiunta di un capitolo di Logica definizioni come la seguente:

> "Si dice che B è un sottinsieme di A e si scrive $B \subseteq A$, se ogni elemento di B è elemento di A, in simboli $B \subseteq A \Leftrightarrow x \in B \Rightarrow x \in A$",

in cui livelli linguistici e metalinguistici, quantificatori e parentesi sono adeguatamente sconclusionati, quindi l'opera dell'insegnante si può scontrare con le difficoltà causate dall'*ipse dixit* di un testo infelice. Esercizi del tipo precedente sono opportuni, ma nascondono il rischio di essere fine a loro stessi, con la riduzione della didattica della logica ad aspetti linguistici, importanti ma non esclusivi né esaustivi. Tali esercizi poi fanno riferimento ad una semantica intuitiva, perché parlano di situazioni della vita comune. E tuttavia non ha senso, a mio parere, compiere analisi delle proposizioni da dimostrare o di dimostrazioni, se non sono noti alcuni aspetti fondamentali, ad esempio, se non è compreso l'uso dei quantificatori e dei connettivi, che passano dal linguaggio naturale a quello formale, perdendo in significatività ed ambiguità, ma acquistando la chiarezza della funzione che è loro riservata. Ribadisco che sono indicati in una prima fase, ma deve esserci poi un'analisi non semantica dell'uso dei connettivi e quantificatori, bensì il loro ruolo nelle regole di introduzione ed eliminazione.

Un successivo gradino potrebbe essere quello di avvicinare la dimostrazione in Matematica, andando a "saccheggiare" nelle ore di Filosofia i sillogismi categorici. E' un argomento che si presta bene ad attività interdisciplinari. I colleghi di Matematica e Filosofia dovrebbero illustrare i motivi per cui Aristotele e, dopo di lui, altri pensatori si sono serviti di tali forme di ragionamento. C'è però un problema con i tempi. Così facendo si sposta questa attività al terzo anno di scuola superiore, in quelle scuole in cui si tratta la Filosofia, mentre il percorso che porta alla dimostrazione è previsto già nel biennio, esemplificato sulla Geometria.

In alcune proposte in questa ottica si è cercato di evidenziare la differenza tra correttezza del ragionamento e verità degli enunciati che intervengono in esso, nonché la distinzione tra uso di proprietà logiche e uso di proprietà matematiche all'interno di una dimostrazione. Quindi accanto a sillogismi della forma

$$\frac{\begin{array}{c}\text{Ogni multiplo di 4 è pari}\\ \text{Qualche multiplo di 5 non è pari}\end{array}}{\text{Qualche multiplo di 5 non è multiplo di 4}}\ \text{(Ferion)}$$

in cui il contenuto delle due premesse è facilmente verificabile come vero e la conclusione è vera, è bene presentare sillogismi che siano corretti, ma in cui intervengano proposizioni false

$$\frac{\begin{array}{c}\text{tutti i parallelogrammi sono rombi}\\ \text{tutti i rombi sono quadrati}\end{array}}{\text{tutti i parallelogrammi sono quadrati}}\ \text{(Barbara)}$$

oppure altri in cui le premesse sono vere, la conclusione contrasta col senso comune, in quanto viene tirata in ballo l'intensione piuttosto che l'estensione,

$$\frac{\begin{array}{c}\text{chi ruba teme la prigione}\\ \text{chi teme la prigione è onesto}\end{array}}{\text{chi ruba è onesto}}\ \text{(Barbara)}$$

o altri ancora in cui da premesse false si ottiene una conclusione vera:

$$\frac{\begin{array}{c}\text{nessun triangolo rettangolo è un triangolo scaleno}\\ \text{ogni triangolo equilatero è un triangolo rettangolo}\end{array}}{\text{nessun triangolo equilatero è un triangolo scaleno}}\ \text{(Celarent)}$$

Queste attività, per altro raccomandate, anche per la possibilità di scambio tra le competenze di diversi insegnanti, non devono fermarsi a questa fase. Altrimenti si ripercorre il cammino che da Aristotele ha portato a Boole (1854). E' ottima cosa passare dalle frasi della lingua italiana a schematizzazioni del tipo

$$\frac{\begin{array}{c}\text{ogni 'a' è 'm'}\\ \text{qualche 'b' non è 'm'}\end{array}}{\text{qualche 'b' non è 'a'}}\ \text{(Baroco),}$$

e poi alla schematizzazione più logica in cui intervengono simboli per predicati, simboli per connettivi e quantificatori

$$\frac{\begin{array}{c}\forall x(M(x) \rightarrow A(x))\\ \forall y(B(y) \rightarrow M(y))\end{array}}{\forall z(B(z) \rightarrow A(z))}\ \text{(Barbara),}\qquad \frac{\begin{array}{c}\neg\exists x(M(x) \wedge A(x))\\ \forall z(B(z) \rightarrow M(z))\end{array}}{\neg\exists y(B(y) \wedge A(y))}\ \text{(Celarent),}$$

$$\frac{\begin{array}{c}\neg\exists z(M(z) \wedge A(z))\\ \exists y(B(y) \wedge M(y))\end{array}}{\exists z(B(z) \wedge \neg A(z))}\ \text{(Ferion),}\qquad \frac{\begin{array}{c}\forall y(A(y) \rightarrow M(y))\\ \exists y(B(y) \wedge \neg M(y))\end{array}}{\exists y(B(y) \wedge \neg\, A(y))}\ \text{(Baroco)}\ [1]$$

E' importante osservare che i sillogismi categorici spesso vengono presentati con l'uso dei predicati nominali. In essi la copula viene tradotta a volte come implicazione, altre volte come congiunzione. Le motivazioni di ciò vanno illustrate, essendo sostanzialmente causate dalla relativizzazione dei quantificatori.

Da questo esempio, che storicamente è assai importante, si dovrebbe passare a mostrare che anche per semplici deduzioni in cui intervengano predicati non solo monadici, il sillogismo non è sufficiente: un esempio è il cosiddetto (anche se non è un) sillogismo di De Morgan

$$\frac{\text{ogni cavallo è un animale}}{\text{ogni coda di cavallo è coda di un animale}}$$

che si può tradurre in simboli come

$$\frac{\forall x(C(x) \rightarrow A(x))}{\forall y(\exists z(C(z) \wedge D(y,z)) \rightarrow \exists u(A(u) \wedge D(y,u)))}$$

Nelle sperimentazioni viene fatto riferimento anche alla rappresentazione dei sillogismi mediante insiemi, introducendo così quest'altro tema che è strettamente connesso con la Logica.

In altre proposte si è evidenziato il fatto che il ricorso a tavole di verità per l'analisi delle proposizioni viene a nascondere gli aspetti sintattici, che poi sono importanti per il calcolo dei predicati. Per questo si cerca di affrontare il procedere della dimostrazione da semplici passaggi, privilegiando le regole quali il Modus ponens e il Modus tollens:

Se	"Ha svaligiato la Banca degli Onesti Galantuomini"
allora	"E' passato per via Briganti"
e	"Ha svaligiato la Banca degli Onesti Galantuomini"
Si deduce	"E' passato per via Briganti"

[1] Volutamente le indeterminate usate in queste "traduzioni" sono arbitrarie, e questo va spesso fatto osservare per non suggerire involontariamente ipostatizzazioni nell'uso delle lettere.

Se	"Ha svaligiato la Banca degli Onesti Galantuomini"
allora	"E' passato per via Briganti"
e **non**	"E' passato per via Briganti"
Si deduce **non**	"Ha svaligiato la Banca degli Onesti Galantuomini".

Spesso le attività proposte sono incentrate sul "meccanismo" della dimostrazione per assurdo, contrapposte alle dimostrazioni cosiddette "dirette". Le dimostrazioni per assurdo vengono analizzate nelle varie forme, non viene però fatta menzione della possibilità di provare un'implicazione in forma di contronominale.

4.5. Viene posta poca attenzione, nelle dimostrazioni proposte, all'uso delle proprietà matematiche differenziandole da quelle logiche. Ma ciò è importante per convincere gli studenti che gli argomenti logici non costituiscono lo strumento universale di dimostrazione dei teoremi, ma servono per dare "sicurezza" alla dimostrazione e garanzia di correttezza.

Gli esempi di dimostrazioni vengono tratti dalla Geometria, dal Calcolo delle Probabilità e dall'Algebra, e queste ultime sembrano assai interessanti, sia perché si tratta di dimostrazioni semplici, in cui intervengono i soli predicati di eguaglianza e di ordine, sia perché gli assiomi matematici richiesti sono ben noti ed accettati, a tal punto da essere presi come proprietà logiche. Basti pensare al clamore suscitato dalla scoperta dei quaternioni di Hamilton.

Negli esempi riportati sotto, presentati da alcuni docenti, si usano vari libri di testo. La proposta procede in più fasi. Si inizia dal riconoscimento di argomentazioni e dimostrazioni e delle inferenze usate in esse:

Leggi attentamente il seguente teorema e la sua dimostrazione tratta dal libro di testo (LAMBERTI-MEREU-NANI, Nuova Matematica 1, Etas, pag. 124) *e sottolinea quelle parole che, a tuo avviso, indicano che si sta compiendo un'inferenza deduttiva*:

"Teorema. Quali che siano gli eventi E e F, si ha P(E) + P(F) = P(E∪F) + P(E∩F). Dimostrazione. Indichiamo con E_0 = E - E∩F, si ha che E∪F = E_0∪F, inoltre E_0 e F sono incompatibili, quindi P(E) = P(E_0) + P(F). Inoltre E = E_0∪(E∩F), E_0 e E∩F sono incompatibili, quindi P(E) = P(E_0) + P(E∩F), da cui P(E0) = P(E) - P(E∩F) e, quindi P(E∪F) = P(E) - P(E∩F), da cui l'asserto".

Si passa poi all'analisi delle dimostrazioni, in modo abbastanza formale, confrontando passi di dimostrazioni con regole di inferenza presentate in anticipo con una lezione frontale (N.B. nella proposta è utilizzato il sistema di deduzione naturale di Prawitz).

(MARASCHINI-PALMA, Matematica di base 1, Paravia, Pag. 539): "Teorema 8. date due rette *r* e *s* perpendicolari a una stessa retta *t*, sono tra loro parallele. Dimostrazione: se *r* e *s* non sono parallele, allora hanno esattamente un punto in comune. Sono possibili due casi: I) il

punto appartiene anche alla retta *t*. Ciò contrasta con l'assioma 10. II) Il punto non appartiene alla retta *t*. Ciò contrasta col teorema 4. Siamo quindi arrivati ad un assurdo, perché il punto non può né appartenere, né non appartenere a *t*...". *Segna con una croce gli schemi deduttivi fra quelli presenti nel riquadro, che ritieni corretto associare al passo sopra riportato.*

$$(\text{i.}\neg)\ \frac{[\varphi];\bot}{\neg\varphi} \qquad (\text{e.}\neg)\ \frac{\varphi\ \neg\varphi}{\bot} \qquad (\text{Abs}_1)\ \frac{\bot}{\varphi} \qquad (\text{Abs}_2)\ \frac{\overset{[\neg\varphi]}{\bot}}{\varphi}$$

$$(\text{i.}\vee)\ \frac{\varphi}{\varphi\vee\psi}\quad\frac{\psi}{\varphi\vee\psi} \qquad (\text{e.}\vee)\ \frac{\varphi\vee\psi \quad \overset{[\varphi]}{\vartheta} \quad \overset{[\psi]}{\vartheta}}{\vartheta}$$

Così facendo si analizza la struttura locale di una dimostrazione, primo passo indispensabile, a mio avviso, per giungere alla struttura globale.

Viene proposta poi una terza fase, in cui si invitano gli studenti a riconoscere schemi di deduzione utilizzati in passi di dimostrazioni, e non sempre esplicitati, e poi a scrivere lo schema formale esplicito:

Leggi attentamente l'enunciato del teorema e il relativo cenno di dimostrazione che si trova in PALATINI-FAGGIOLI, Elementi di Geometria per Istituti Tecnici, 1992, pag. 60 (con opportuni adattamenti). "Teorema: In un piano, per un punto passa una ed una sola retta perpendicolare ad una retta data. [La dimostrazione si riconduce a dimostrare i seguenti teoremi: 1) In un piano, per un punto appartenente ad una retta *r* passa una ed una sola retta perpendicolare a *r*. 2) In un piano, per un punto esterno ad una retta *r* passa una ed una sola retta perpendicolare a *r*]

Evidenzia gli aspetti logici del ragionamento accennato, esplicitando i passi deduttivi significativi omessi e individuandone lo schema formale.

Queste le proposte, assai interessanti e graduate. E' interessante osservare che a conclusione della attività è prevista una verifica con intenti di valutazione formativa sull'attività degli studenti suddivisi in gruppi.

Da questa sperimentazione si sono messe ben in luce alcune difficoltà: ad esempio alla richiesta di evidenziare gli aspetti logici e deduttivi, molti studenti hanno eseguito una parafrasi della dimostrazione, traducendo anzi le formule in lingua corrente, aiutandosi con disegni e diagrammi. In realtà spesso gli studenti eseguono corrette applicazioni di regole solo dell'eguaglianza. Si avverte poi che l'attenzione posta dagli insegnanti nella spiegazione delle dimostrazioni per assurdo, rischia di far identificare la dimostrazione solo con quella per assurdo.

4.6. Dalle relazioni esaminate non ho rilevato esempi di come una dimostrazione in una teoria, o almeno la parte logica di essa, possa riapplicarsi ad altri contesti, facendo comprendere così che l'ossatura offerta dalla dimostrazione è una impalcatura che può essere rivestita da abiti assai diversi. Forse l'esempio più clamoroso di questa situazione è offerto dalle dimostrazioni per induzione.

5. Conclusioni

Il titolo di questo paragrafo è falso. Conclusioni si possono trarre solo dopo un'approfondita sperimentazione. Le attività proposte dai corsisti, gli itinerari didattici che essi presentano sono "affetti" da un difetto: quello che sono stati approntati e talora sperimentati da docenti assai motivati e capaci. Resta il problema della trasferibilità anche in classi in cui la motivazione di allievi ed insegnanti non è così forte.

La dimostrazione ha un ruolo unificante in tutta la Matematica, dunque si può approfittare della sua ubiquità per fare Logica in ogni occasione, non isolando la nostra disciplina in capitoli o periodi ben precisi. L'educazione alla Logica deve iniziare dal primo anno ed essere condotta fino in quinta, con modalità diverse e con contenuti diversi. E' bene far ripercorrere agli studenti, con discussioni collettive, i passi delineati da Balacheff, di cui si parla nel paragrafo 3. Ma bisogna avere pazienza: dopo aver sollecitato la discussione su concetti ancora vaghi, si possono utilizzare varie occasioni per evidenziare, qua e là, alcuni passaggi. L'idea di un diario di Logica, da tenere aggiornato ogni tanto, in occasione di un'analisi più approfondita, traendo spunto da passaggi di dimostrazioni non chiari alla scolaresca o a qualche alunno, potrebbe far giungere alla "scoperta" che le regole usate poi sono sempre le stesse. Da queste registrazioni, con una sapiente opera del docente, si potrebbe giungere ad esaminare i concetti "forti" quali linguaggio e metalinguaggio, verità e dimostrazione. Su questi temi così generali e complessi, non sarebbe male si svolgessero delle verifiche costituite da temi analoghi a quelli che vengono assegnati dai colleghi di Italiano per vedere se gli alunni hanno ben compreso la poetica di questo o quell'autore. Gli elaborati prodotti possono essere a loro volta argomento di discussione collettiva.

In queste attività scolastiche mi sembra opportuno l'utilizzazione della dinamica di gruppo, come sede di ulteriore elaborazione tra studenti.

Solo se se ne comprende la necessità, potranno essere assimilati e poi esplicitati i motivi che hanno portano i matematici e i logici ad individuare concetti assai complessi. Non si deve trascurare, laddove sia possibile, l'aggancio con l'Informatica.

Riferimenti

[B] BALACHEFF N., *Preuve et démontration en mathématique au college*, Recherches en didactiques des Mathématiques, **3**, n. 3 (1982).

[Ci] GRUGNETTI L. (a cura di), *L'évaluation centrée su l' élève - Assessment focused on the student,* Proceedings of the 45th CIEAEM Meeting, Cagliari 4-10 July, 1993.

[D] DUVAL R., *Structrure du raisonnement deductif et apprentissage de la démonstration*, Educational Studies in Mathematics, **22** (1991).

[Ia] IADEROSA R., *L'avvio alla dimostrazione nella scuola dell'obbligo,* Tesina del Corso di perfezionamento in didattica della Matematica, Università Cattolica del Sacro Cuore - Brescia A.A. 1993/94.

[M1] MARCHINI C., *Le sostituzioni e le relazioni,* L'insegnamento della Matematica e delle Scienze integrate, **13**, n. 7 (Luglio 1990), 732 - 744.

[MP1] MARGIOTTA P., *Un'esperienza con le sostituzioni nella Scuola Media,* La Matematica e la sua Didattica, **5**, n. 2 (Apr.-Giugno 1991), 32 - 36.

[MP2] MARGIOTTA P., *Le sostituzioni in un'ottica interdisciplinare*, L'Educazione Matematica, **XII** (III), vol.2, n. 1 (Aprile 1991), pag. 23 - 44.

[R] ROUCHE N., *Prouver: amener à l'évidence ou controler des implications?* Groupe d'Enseignement Mathématique Louvain-La-Neuve.

[V] VARGA M., *Logica per insegnanti,* Boringhieri, Torino, 1960.

Osservazioni e Spunti per una Proposta didattica: il Concetto matematico di "Infinito"

Resoconto di un'esperienza, a cura di

CLAUDIO BERNARDI
Dipartimento di Matematica
Università Roma "La Sapienza"
Piazzale Aldo Moro 2
00185 Roma

0. Considerazioni generali

E' quasi superfluo soffermarsi sull'importanza del concetto di infinito in matematica. Fin dalle Scuole Elementari, gli studenti imparano che non c'è un numero più grande di tutti, che ogni segmento può essere prolungato perché una retta è "infinitamente lunga", che fra due punti di un segmento ne è sempre compreso un altro, che in certi procedimenti (come nella divisione fra due numeri naturali dei quali il primo non sia multiplo del secondo) si può proseguire quanto si vuole senza arrivare mai alla conclusione. Va aggiunto che la matematica, fra le varie scienze, è l'unica in cui l'infinito entra esplicitamente in gioco.

Ciò nonostante, il concetto di infinito non rientra nei temi di un insegnamento tradizionale e ancor oggi è raramente l'oggetto esplicito di una lezione. I Programmi Brocca, invece, propongono l'argomento alla voce "confronto di insiemi numerici infiniti" nella IV classe del Liceo Scientifico (n. 2.c); nel relativo commento si precisa che tale confronto «dovrà far risaltare la differenza fra la potenza del numerabile e quella del continuo».

Come è stato posto in rilievo nella parte teorica, il tema "l'infinito" è strettamente connesso a concetti di logica matematica. In effetti, l'infinito può essere visto come un tema trasversale, perché presenta ovvi legami anche con argomenti di aritmetica, algebra, geometria. Inoltre, vi sono connessioni con la filosofia e con studio del linguaggio (con un numero finito di simboli e con frasi di lunghezza finita, riusciamo a nominare infiniti oggetti e a descrivere strutture infinite). Più in generale, si tratta di un argomento che da un lato ha

una indubbia portata formativa e culturale, dall'altro è ricco di risultati tecnici non banali e non prevedibili: la trattazione matematica porta a demolire opinioni comuni e spontanee per sostituirle con fatti non intuitivi, quali l'esistenza di una corrispondenza biunivoca fra un insieme e una sua parte propria o la possibilità di costruire una gerarchia di infiniti.

Non vanno naturalmente sottovalutate le difficoltà del tema: l'argomento è decisamente astratto, oggettivamente difficile, presenta grossi nodi concettuali ed ostacoli epistemologici (non per niente, la storia dell'infinito è piena di dubbi e di incertezze). Per questi motivi, può risultare didatticamente più efficace presentare alcuni aspetti dell'argomento, per poi riprenderlo più volte nel seguito, a distanza di mesi o anche di anni: si tratta di concetti che indubbiamente richiedono tempi di maturazione.

Sottolineiamo ancora che affrontare questo tema offre un'occasione per inquadrare storicamente momenti significativi dell'evoluzione del pensiero matematico, cosa che è raccomandata dai Programmi Brocca. Non va nemmeno trascurato il vantaggio di parlare di un momento storicamente recente, al contrario di altri argomenti matematici che danno talvolta l'impressione di una materia ormai "conclusa". Come è detto nella prima lezione, l'opera di Cantor va considerata uno degli eventi più importanti nella storia del pensiero matematico; talvolta, in maniera un po' superficiale, nell'introdurre i primi concetti insiemistici (unione, intersezione, sottoinsieme, ...) si fa riferimento a Cantor, senza porre nel giusto rilievo il suo merito principale: aver trovato una soluzione ai paradossi dell'infinito attuale, fino a renderlo oggetto di studio matematico.

1. Occasioni

E' opportuno ripercorrere le tappe attraverso le quali i ragazzi si avvicinano al concetto di infinito. Cerchiamo di elencare le varie occasioni in cui nell'insegnamento della matematica si parla, in modo più o meno diretto, di infinito.

Sono infiniti i consueti insiemi numerici: l'insieme **N** dei numeri naturali, l'insieme **Z** dei numeri interi, l'insieme **Q** dei numeri razionali, l'insieme **R** dei numeri reali. Sempre nell'ambito aritmetico si incontrano i numeri decimali illimitati (cioè con infinite cifre dopo la virgola): in particolare, abbiamo i numeri periodici, dove alcune cifre si ripetono infinite volte. Va menzionato in proposito il caso del periodo 9: in termini un po' rozzi, si può pensare che la differenza 5,(0) – 4,(9) sia infinitamente piccola (e ci sono così due scritture diverse per uno stesso numero).

Se poi si deve calcolare il risultato della divisione "1:0", la risposta corretta è "la divisione non ammette risultato", o "è impossibile", ma, a livello intuitivo, conviene spesso pensare ad un quoziente infinitamente grande. Il discorso è del tutto analogo nel caso dell'equazione "$0 \cdot x = 1$".

Tutte le figure che si incontrano usualmente in geometria, pensate come insiemi di punti, sono infinite: retta, piano, segmento, cerchio, ... Si noti che retta e piano sono figure non solo infinite, ma illimitate, nel senso che, per ogni numero M, esistono due punti della figura che distano fra loro più di M; invece, segmento e cerchio sono figure infinite, ma limitate (ogni figura illimitata è infinita, mentre non vale l'implicazione inversa).

L'infinità di un segmento è strettamente legata al fatto che si tratta di un insieme ordinato denso (la differenza fra ordini densi e ordini continui è piuttosto complessa e forse, nelle Scuole Superiori, non è il caso di insistere più di tanto). La densità permette poi di dividere un segmento in un qualunque numero finito di parti (in questa circostanza abbiamo a che fare con un infinito potenziale).

Sempre nello studio della geometria si incontrano le grandezze incommensurabili: l'algoritmo delle differenze successive, per determinare un sottomultiplo comune a due segmenti, può non avere termine; in algebra considerazioni analoghe portano all'esistenza dei numeri irrazionali.

Abbiamo poi i procedimenti di approssimazione per la lunghezza della circonferenza e l'area del cerchio (più in generale, per lunghezza di una curva e l'area di una regione piana); negli ultimi anni di alcune Scuole Superiori, il discorso viene precisato con l'introduzione dei concetti di convergenza e di limite.

Anche le principali relazioni geometriche associano a ciascuna delle figure coinvolte infinite figure: si pensi all'uguaglianza (o congruenza), all'equivalenza (o equiestensione), alla similitudine, al parallelismo e alla perpendicolarità fra rette e piani, ... Sull'infinità dei fasci di piani paralleli si basa il principio di Cavalieri per la determinazione dei volumi dei solidi.

Tornando all'aritmetica da un punto di vista più astratto, citiamo il principio di induzione e la sistemazione di Peano per descrivere la successione dei naturali. Abbiamo ancora procedimenti che si possono proseguire indefinitamente, senza termine: dall'algoritmo più semplice "parti da 0, aggiungi 1, poi ancora 1, e così via", alle progressioni aritmetiche e geometriche, alle successioni e alle serie; probabilmente, anche il bambino a cui, nel primo ciclo della Scuola Elementare, si chiede di "numerare per 2 da 20 a 50, si rende conto che l'esercizio pone il termine 50 solo per motivi di tempo, ma, volendo, si potrebbe continuare). Processi più complessi, ma concettualmente non troppo diversi, portano alle frazioni continue, alla curva di Von Kock e ai frattali.

Citiamo, infine, i quantificatori in logica matematica: se facciamo riferimento ad un insieme infinito, un quantificatore universale sostituisce una congiunzione infinita (che non potremmo scrivere direttamente nel nostro linguaggio), mentre un quantificatore universale sostituisce una disgiunzione infinita.

2. L'esperienza

L'esperienza è stata condotta da 9 docenti che hanno partecipato al Corso di Logica *MPI-AILA*. Colgo l'occasione per ringraziare gli insegnanti, che hanno svolto un ruolo attivo ed autonomo sia nella formulazione delle proposte, sia nella realizzazione in classe, sia ancora nell'analisi dei risultati. In effetti, in questo testo mi sono limitato a rielaborare le relazioni scritte dagli insegnanti.

Le classi coinvolte sono state 11: 3 del primo anno delle Superiori (di cui 2 quinte Ginnasio), 4 del terzo anno, 2 del quarto e 2 dell'ultimo anno; in quasi tutte le classi erano adottati i programmi sperimentali *PNI* o Brocca. Va precisato che l'argomento non era stato inserito in modo organico nella programmazione annuale. E' chiaro che il livello scolare ha influito sulle modalità di presentazione: in particolare, nelle classi del triennio sono state viste anche applicazioni all'analisi ed è stata posta maggiore attenzione alla storia e alle implicazioni filosofiche.

L'argomento si presta a discussioni in classe, il che è un fatto indubbiamente positivo. Tuttavia, queste discussioni, se non sapientemente guidate, corrono il rischio di diventare dispersive; si tenga anche presente che successive verifiche hanno talvolta mostrato che idee che sembravano consolidate durante le discussioni in classe, erano invece ancora poco chiare a gran parte degli studenti.

Parlare di infinito richiede negli studenti capacità critiche, anche perché l'argomento viene affrontato in un modo che si discosta da quello consueto nell'insegnamento della matematica. Allo stesso tempo, parlare di infinito può facilitare l'acquisizione di un atteggiamento critico; nonostante le indubbie difficoltà, è capitato che, a posteriori, qualche ragazzo abbia trovato l'argomento "più vicino a noi" di quanto non siano usualmente gli argomenti matematici, e qualcuno ha apprezzato che questa esperienza gli abbia fatto vedere nella matematica "più dubbi che certezze". Paradossalmente, la consapevolezza del sofferto divenire storico delle conoscenze matematiche e del sussistere di limiti ancora non superati, rafforza l'interesse per la disciplina, avvicina alla lettura di libri di matematica, finisce per appassionare gli studenti.

3. Modalità

Quando a Scuola si propone un argomento nuovo, si pongono subito molte domande: Come va affrontato l'argomento? Quali sono i prerequisiti? Quali esercizi vanno assegnati? Quali sono i contenuti essenziali? Che tipi di verifiche si possono eseguire? Queste verifiche entrano nella valutazione globale dello studente?

Nelle esperienze condotte in classe le risposte alle domande precedenti sono state almeno in parte diverse. All'argomento sono state dedicate da un minimo di 8 ad un massimo di 14 ore.

I prerequisiti essenziali per affrontare l'argomento si riducono ai concetti e alle abilità di base della teoria degli insiemi; naturalmente, è utile una certa familiarità con le progressioni aritmetiche e geometriche, e anche un'idea intuitiva dei concetti di limite e di infinitesimo. Per i ragazzi che avevano già svolto i primi capitoli di analisi matematica, si è trattato di un momento di riflessione e sistemazione; per altri ragazzi, l'infinito è stato invece visto come un argomento propedeutico allo studio dell'analisi.

Gli esercizi sono stati assegnati prevalentemente in un contesto numerico, ma alcuni insegnanti hanno privilegiato esempi geometrici, sia legati all'equivalenza dei poligoni, sia con costruzioni grafiche di corrispondenze biunivoche.

Molti dei docenti hanno seguito una metodologia inconsueta: invece che iniziare con lezioni frontali, hanno cominciato distribuendo schede per lavori di gruppo, contenenti questioni di vario tipo. In tal modo sono state favorite discussioni fra studenti: le idee che circolavano erano spesso ingenue, ma i ragazzi si sono trovati direttamente di fronte ai grossi problemi posti dall'infinito, su cui l'umanità ha discusso per secoli. D'altra parte, l'insegnante ha avuto un mezzo per controllare l'intuizione che gli alunni hanno del concetto di infinito ed ha avuto una prima occasione per chiarirlo; uno degli obiettivi iniziali è stata la distinzione fra infinito potenziale e infinito attuale. Con un intento non troppo diverso, altri insegnati hanno preferito assegnare un test per verificare le conoscenze "all'ingresso".

I ragazzi sono stati poi incoraggiati a svolgere un lavoro individuale, in particolare con letture su testi diversi dal manuale adottato; talvolta, sono state eseguite ricerche e consultazioni in biblioteche civiche, fatto anomalo per un argomento di matematica. Le lezioni sono state di frequente integrate con discussioni: la cosa fa "perdere" un certo tempo, ma l'insegnamento può risultare più efficace. Naturalmente, sono stati presentati anche concetti e tecniche più strettamente matematici, con una metodologia più simile a quella usuale.

L'esperienza si è conclusa con un test finale; alcuni insegnanti hanno chiesto agli studenti anche di svolgere un tema o di rispondere alle domande di un questionario. Proprio dall'esame dei temi e dei questionari si rileva che l'argomento è piaciuto alla maggioranza degli studenti (che hanno notato una modalità di approccio diversa dal solito); non sono mancate le difficoltà, e a queste è dedicato il prossimo paragrafo. Secondo alcuni insegnanti, aver parlato in modo esplicito di infinito ha contribuito a rendere più chiare le costruzioni di **Z**, di **Q** e, soprattutto, di **R**.

Un discorso a parte merita la collaborazione con altri colleghi, non di matematica: sono stati coinvolti docenti di storia e filosofia per approfondire gli aspetti storici, critici ed epistemologici, di italiano, di storia dell'arte, e, nel Liceo Classico, perfino di greco. Si sono manifestate differenze di linguaggio con i relativi problemi, ma, nel complesso, la collaborazione è stata fattiva. Proprio per sottolineare quanto l'argomento si presti ad una

didattica trans-disciplinare (credo che oggi si dica così) riporto una citazione che è stata segnalata da un insegnante di storia dell'arte: si tratta di una frase con cui Anassagora, che pare si sia occupato di prospettiva scenografica, intuisce la possibilità di porre in corrispondenza biunivoca due segmenti di lunghezza diversa, o due quadrati di area diversa: «tanto nel grande quanto nel piccolo vi è uno stesso numero di particelle».

4. Alcune difficoltà

Spesso, chi riferisce di una sperimentazione didattica relativa alla presentazione di un nuovo argomento, sottolinea l'interesse degli studenti (anche di quelli in genere più svogliati) e parla di risultati decisamente positivi. Questi giudizi, che naturalmente risentono dell'entusiasmo e delle convinzioni del docente, non sono di grande utilità per chi voglia riprovare a presentare quell'argomento.

Nel nostro caso, anche se il concetto di infinito ha indubbiamente suscitato un certo interesse fra gli studenti (o almeno in alcuni di essi), è più interessante, proprio per le difficoltà del tema, esaminare gli aspetti che sono risultati meno soddisfacenti. Non si tratta di scoraggiare chi voglia ripetere l'esperienza, tutt'altro: si tratta di segnalare alcuni punti delicati o comunque situazioni a cui prestare attenzione.

In primo luogo, noi insegnanti siamo probabilmente portati a pensare che il concetto di infinito sia acquisito (sia pure indirettamente) dagli studenti che hanno concluso la Scuola dell'obbligo. Non pochi ragazzi, all'inizio delle Superiori, hanno invece un'idea piuttosto vaga del concetto di insieme infinito: c'è chi pensa addirittura che un segmento abbia un numero finito di punti (del resto, nel linguaggio corrente "infinito" indica spesso un insieme molto grande, per cui è difficile contare il numero degli elementi).

Le nozioni apprese dagli studenti non sono ancora organizzate in modo coerente: lo stesso ragazzo che crede che un segmento sia un insieme finito, può poi affermare con sicurezza che fra due punti di una retta ne è sempre compreso un altro. Anche il controllo con l'esperienza quotidiana è meno scontato di quanto si possa pensare: pare che, di fronte al paradosso di Achille e la Tartaruga, alcuni ragazzi si convincano davvero del fatto che Achille non raggiungerà la Tartaruga, mentre altri concludono che Achille non può nemmeno muoversi.

Per evitare sorprese, è bene assegnare numerosi esercizi in itinere: probabilmente, la percentuale di risposte corrette sarà diversa dalle attese dell'insegnante. Un inciso di carattere generale: è sempre utile, per un insegnante che assegna un esercizio, fare una previsione sul numero dei suoi studenti che lo risolveranno in modo corretto.

Una grossa difficoltà è legata al linguaggio, che assume un'importanza cruciale nel nostro contesto: la correttezza linguistica è sempre fondamentale in matematica, ma lo studente meno preciso può cavarsela in geometria con una figura e in algebra con un calcolo:

parlando di infinito, invece, si deve appunto parlare. Ad esempio, la parola "paradosso", che pure è stata vista in altre materie, non è chiara per la maggioranza degli studenti. Il problema del linguaggio va affrontato in modo equilibrato: da un lato è essenziale che l'insegnante si soffermi a spiegare tutte le parole di cui i ragazzi non capiscono l'esatto significato, dall'altro l'insegnante non deve preoccuparsi troppo di un eccessivo rigore linguistico. L'importante è trasmettere una visione corretta, ma, specialmente all'inizio, penso sia lecito usare locuzioni come "un numero infinito di punti", o "una retta lunga infiniti metri". Analogamente, non vale la pena di cercare definizioni rigorose di numeri cardinali (né, tanto meno, di ricorrere a questi per introdurre i numeri naturali).

C'è infine una ben nota difficoltà per ogni docente che intenda rinnovare il suo insegnamento, ma non sia disposto ad accettare che i ragazzi deboli rimangano sempre più indietro: il tempo a disposizione è sempre lo stesso, ed è poco se rapportato al programma (inteso come insieme degli argomenti a cui non si può rinunciare).

5. Esempi di esercizi

Un problema che riguarda anche altri settori della logica consiste nella formulazione di esercizi appropriati, non banali né troppo complessi, e in cui si trovi un "risultato" (ad esempio, considerazioni sul linguaggio sono molto utili, ma, talora, l'intera risoluzione si limita alla risposta, e lo studente può non avere la sensazione di avere "fatto" un esercizio). Nel nostro caso, accanto a verifiche di conoscenze specifiche, occorrerà portare gradualmente gli studenti a raggiungere una "autonomia risolutiva" in questioni diverse dal solito.

L'argomento si presta a varie tipologie di esercizi:

domande teoriche, che riguardano anche la storia e la filosofia;

ricerca di esempi (come trovare un insieme numerabile diverso da quelli già visti, oppure costruire, se è possibile, una corrispondenza biunivoca fra due insiemi dati);

classificazione (in particolare, di insiemi a seconda della cardinalità);

ricerca di definizioni adeguate (ad esempio: dati due insiemi A e B, quando è corretto dire che Card A è strettamente minore di Card B ?);

domande più imbarazzanti e riflessione su situazioni strane (due esempi: quanto è lunga in centimetri una retta lunga infiniti metri? si esaminino i prodotti $1{,}6 \cdot 2 = 3{,}2$; $1{,}66 \cdot 2 = 3{,}32$; $1{,}666 \cdot 2 = 3{,}332$; ...: l'ultima cifra del prodotto $1{,}(6) \cdot 2$ è 2 oppure 3?)

Non mancano, naturalmente, esercizi tecnici più simili agli usuali esercizi matematici.

Vediamo, più in concreto, alcuni esercizi, dividendoli per argomento.

Esercizi sugli insiemi ordinati (ordini densi, ...).

- Quanti sono i numeri interi compresi fra 5 e 8? e i numeri decimali (o razionali) compresi fra 0,5 e 0,8?
- In una calcolatrice si possono scrivere numeri con al più 9 cifre, fra parte intera e parte decimale: quanti numeri si possono scrivere su quella calcolatrice compresi fra 5 e 8?
- Quante sono le frazioni comprese fra 3/7 e 5/7 ?
- Si considerino gli insiemi **N**, **Q**, **R**: in quali casi ha senso parlare del successivo del numero 3 ? (per successivo di un numero x si intende il più piccolo numero maggiore di *x*). Nell'insieme dei numeri decimali, c'è il successivo del numero 3,1 ?
- E' giusto dire che i numeri naturali costituiscono un insieme infinito perché per ognuno di essi si può indicare il successivo? si tratta di una caratteristica di tutti gli insiemi infiniti?
- Trovare un numero a (decimale, razionale, o reale) compreso fra 0 e 0,1; trovare poi un numero b compreso fra 0 ed a, e quindi un numero c compreso fra 0 e b. Si può proseguire così fin che si vuole?

Esercizi sulla cardinalità.

Per ciascuno dei seguenti insiemi, dire se si tratta di un insieme finito o infinito; nel primo caso determinare il numero degli elementi dell'insieme, nel secondo caso stabilire se di tratta di un insieme numerabile o non numerabile.

- Insieme dei numeri naturali che sono potenze di 2.
- Insieme dei numeri naturali che sono multipli di 12.
- Insieme dei numeri naturali che sono divisori di 12.
- Insieme dei numeri naturali che non sono multipli di 12.
- Insieme delle radici quadrate dei numeri naturali.
- Insieme delle coppie ordinate di naturali che danno per somma 10.
- Insieme delle coppie ordinate di numeri interi relativi che danno per somma 10.
- Insieme delle frasi scritte in italiano (ciascuna di lunghezza finita).
- Insieme delle sequenze finite di lettere dell'alfabeto (indipendentemente dal fatto che abbiano un significato).
- Insieme dei sottoinsiemi di un insieme con n elementi.
- Insieme dei numeri irrazionali.
- Insieme delle rette di un piano parallele ad una retta data, o passanti per un punto dato.

Figure geometriche.

- Quanti sono i poligoni di area 1 m^2? Quanti i quadrati di area 1 m^2? Quanti i rettangoli di area 1 m^2? (Le domande vanno intese a meno di uguaglianza o congruenza: due figure uguali non vanno contate due volte.)
- Trovare un sottoinsieme proprio di un semipiano che sia uguale (congruente) al semipiano di partenza. (Suggerimento: si disegni una retta parallela al bordo del semipiano.)
- Trovare un sottoinsieme proprio di un angolo retto che sia, a sua volta, un angolo retto. La costruzione si può ripetere nel caso di un angolo non retto?
- Si tracci la retta r tangente ad una circonferenza in un suo punto A. Detto B il punto diametralmente opposto ad A, si considerino le rette passanti per B. Definire una corrispondenza biunivoca fra r e la circonferenza (privata del punto B).
- Sia Γ una semicirconferenza di diametro AB, privata degli estremi A, B; sia r la retta tangente a Γ nel suo punto medio. Costruire, in modo analogo a quanto visto nell'esercizio precedente, una corrispondenza biunivoca fra Γ ed r. (Suggerimento: si consideri il centro di Γ).
- Siano AB e CD due segmenti di lunghezza diversa, posti su rette parallele. Costruire geometricamente una corrispondenza biunivoca fra AB e CD. Supposto poi che AB sia maggiore di CD, costruire una corrispondenza biunivoca fra AB ed un sottoinsieme proprio di CD.

Comportamento diverso di insiemi finiti e infiniti.

Le seguenti affermazioni sono corrette nel caso di insiemi finiti, ma sono sbagliate se si ha a che fare con insiemi infiniti. Trovare un controesempio per ciascuna affermazione.

- Se A e B sono due sottoinsiemi di un insieme C, biiettivi fra loro, allora anche i loro complementari $C-A$ e $C-B$ sono biiettivi fra loro. (In particolare, se A è un sottoinsieme infinito dell'insieme **N**, il complementare di A può essere finito o infinito).
- Se un insieme A è biiettivo ad un sottoinsieme proprio di un insieme B, allora B non può essere biiettivo ad un sottoinsieme proprio di A.
- Dati due insiemi P ed S, se P è biiettivo ad S, allora $P-S$ è biiettivo ad $S-P$ (si ricordi che $A-B$ è l'insieme degli elementi di A che non appartengono a B; l'implicazione inversa vale per tutti gli insiemi, finiti o infiniti).

Gli ultimi tre esercizi sono indubbiamente piuttosto difficili, perché richiedono un buon possesso dell'argomento e molta intuizione (il ragazzo meno sicuro non saprà "da che

parte cominciare"). Si tenga tuttavia presente che la costruzione di esempi e controesempi è una tipica abilità matematica, non sempre posta in rilievo con gli esercizi tradizionali.

Se il tempo lo consente, si potrà proseguire con i primi esercizi di aritmetica cardinale. Con opportuni esempi si illustreranno le proprietà delle operazioni (commutativa ecc.), mentre sarà bene sottolineare che non valgono le leggi di cancellazione: e.g., da $\aleph_0 + 3 = \aleph_0 + 4$ non si può certo dedurre $3 = 4$ (situazioni come queste aiutano a rendere meno scontate le proprietà usuali).

Ancora una osservazione in proposito: il comportamento anomalo dell'infinito è un'occasione per richiamare il comportamento anomalo dello zero.

6. Metafore, paradossi e varianti

Abbiamo già avuto modo di nominare il paradosso di Achille e la Tartaruga. Anche altri fra i celebri ragionamenti di Zenone, come il paradosso della freccia, coinvolgono l'infinito; più in generale è ben noto che il comportamento paradossale dell'infinito portò i Greci al rifiuto dell'infinito attuale.

In questo paragrafo esamineremo altre situazioni più o meno strane. Un'avvertenza importante: mentre in alcuni casi la trattazione matematica offre una spiegazione chiara, almeno nel caso della lampadina non c'è una risposta convincente (la domanda è mal posta).

Due ciclisti A e B partono dagli estremi di una strada rettilinea lunga 30 km e si dirigono ciascuno verso l'altro. A pedala alla velocità costante di 20 km/h, B alla velocità costante di 40 km/h. Una mosca, che vola alla velocità costante di 60 km/h, parte insieme ad A, nella stessa direzione e nello stesso verso; quando incontra B, torna indietro fino a raggiungere A, dopo di che inverte nuovamente la marcia. Se la mosca prosegue così fino a quando A e B si incontrano, quanti km percorre in tutto?

(Il problema è piuttosto complesso se si calcolano le lunghezze degli infiniti tratti percorsi dalla mosca; diviene molto più facile se si fa riferimento ai tempi.)

Una lampada, inizialmente spenta, viene accesa il primo giorno, spenta il secondo giorno, riaccesa il giorno successivo, spenta di nuovo il giorno dopo, e così via. E' facile convincersi che dopo un numero pari di giorni la lampada torna nella situazione iniziale, cioè è spenta, mentre è accesa dopo un numero dispari di giorni. Alla "fine", la lampada è accesa o spenta? Il problema presenta analogie con il calcolo della somma infinita ("serie") $1-1+1-1+1-\ldots$, di cui non si sa dire se il valore sia 0 ovvero 1.

Dimostriamo che il numero periodico 0,(5) è uguale al numero periodico 1,(5):

$$0,(5) = 0,5 + 0,05 + \ldots = (1-0,5) + (1-0,95) + \ldots =$$
$$= 1 + (-0,5 + 1) + (-0,95 + 1) + \ldots = 1 + 0,5 + 0,05 + \ldots = 1,(5)$$

Paradosso di Galileo. E' chiaro che i numeri naturali sono tanti quanti i quadrati perfetti: ogni numero naturale ha un quadrato e numeri diversi hanno quadrati diversi. Tuttavia, è ugualmente chiaro che l'insieme H dei quadrati perfetti è solo una piccola parte dell'insieme **N**, anzi una parte che va diradandosi sempre più.

Galileo non si è accontentato dell'insieme dei numeri pari, che, in un certo senso, è la "metà" di **N**, né dell'insieme M_n dei multipli di n, che è un "ennesimo" di **N**: considerare H rende la situazione ancor più paradossale, perché H è meno "fitto" di M_n per ogni numero n.

L'Albergo di Hilbert. In una certa città sorge un albergo, noto per il fatto di contenere una infinità numerabile di stanze, contrassegnate con i numeri 0, 1, 2, ... L'albergo è al completo; si può sistemare un nuovo cliente? (Facendo traslocare il cliente che occupa la camera 0 nella camera 1, quello che occupa la camera 1 nella 2 e, in generale, quello che occupa la camera n nella camera n+1, resta libera la camera 0 per il nuovo cliente.) In modo analogo, se arriva un numero finito di nuovi clienti, essi riescono a trovare una sistemazione.

Nel caso arrivi una infinità numerabile di nuovi clienti, è sufficiente mandare il cliente che occupava la camera n nella $2n$ per ogni numero n: le camere aventi numero dispari restano così disponibili per i nuovi clienti.

Vogliamo distribuire infinite caramelle (contrassegnate con i numeri naturali 0, 1, 2, ...) a infiniti bambini (anche loro contrassegnati con i numeri 0, 1, 2, ...). La cosa più semplice è dare la caramella con il numero n al bambino con lo stesso numero. Ma i bambini sono molto golosi: riusciamo a fare in modo che ogni bambino riceva due caramelle? E che riceva tre caramelle?

Alcuni riferimenti bibliografici

G. ARRIGO, B. D'AMORE, *Infiniti,* Franco Angeli, Milano 1993

F. ARZARELLO, *Matematica dell'infinito,* CLU, Torino 1980

B. BOLZANO, *I paradossi dell'infinito,* Feltrinelli, Milano

E. FERRARI, G.A. LAGANÀ, E. LUZI, E. TROVINI, *Il concetto di infinito nell'intuizione matematica,* L'insegnamento della Matematica e delle Scienze integrate, vol. 18B, n. 3 (1995), pag. 211-236

L. LOMBARDO RADICE, *L'infinito,* Editori Riuniti 1981

E. MAOR, *All'infinito e oltre,* Mursia, Milano 1993

R. MONDOLFO, *L'infinito nel pensiero dell'antichità classica*

G. PRODI, *Ancora a proposito dell'infinito,* In: Lettera Pristem n. 5, 1992, pag. 19

R. RUCKER, *La mente e l'infinito,* Muzzio, Padova 1991

P. ZELLINI, *Breve storia dell'infinito,* Adelphi, Milano 1980.

F. SPERANZA, (a cura di) "*Epistemologia della Matematica* - Seminari 1989-91", Tecnologie e Innovazioni Didattiche, CNR, quaderno n. 10 (1992). Per averne copia, scrivere al prof. Francesco Speranza, Dipartimento Matematica, via D'Azeglio 85/A, 43100 Parma. Il volume contiene, fra l'altro, le relazioni di un convegno sui concetti di finito e infinito.

PARTE TERZA

Approfondimenti con Complementi ed Esercizi

Verità di una Formula in una Struttura
Conseguenza logica
Modelli di una Teoria
Teorema di Completezza di Gödel

FRANCO MONTAGNA
Dipartimento di Matematica
Università di Siena
via del Capitano 15
53100 Siena

Introduzione

Quando si dice che una formula è vera, oppure che è falsa, in una struttura—ad esempio quando si dice che la formula $\exists x(x \cdot x \approx 2)$ è vera nella struttura dei reali e falsa in quella dei razionali, interpretando i simboli $\approx$ e $\cdot$ rispettivamente nell'uguaglianza e nel prodotto—si ha la sensazione che fra matematici ci si intenda anche senza bisogno di una definizione rigorosa di verità. In effetti, se si eccettua l'interpretazione dell'implicazione, la formalizzazione del concetto di verità di una formula in una struttura è conforme al significato intuitivo dei connettivi e dei quantificatori, e quindi non riserva grosse sorprese; semmai, vale la pena di discutere sul concetto di struttura, o di interpretazione, per un linguaggio al primo ordine.

Prendiamo lo spunto dal linguaggio naturale, anche se l'analogia con quest'ultimo regge solo in parte, sia perché esso è molto più ricco del linguaggio matematico, sia perché dal punto di vista dell'interpretazione del linguaggio naturale non è rilevante solo l'estensione, ma anche l' intensione. Una stessa frase può assumere diversi significati a seconda dei contesti, e in particolare può risultare vera o falsa a seconda della situazione a cui si riferisce. Ad esempio, la verità della frase "Il treno per Bari partirà alle 7,10" dipende da circostanze come il luogo in cui viene pronunciata, il giorno, etc. Questi dati ci consentono di capire cosa si intende dire e, di conseguenza, di concludere che la frase è vera oppure che è falsa. Nella

frase non vengono precisati tutti i dati, dicendo ad esempio: "Il primo treno a partire dalle 7 del mattino che secondo l'orario dovrebbe raggiungere la stazione di Bari partendo dalla stazione di Lecce, partirà dalla medesima alle ore 7,10"; né tantomeno viene specificato cosa si intende per "treno", "partire", etc...; al contrario, la frase resta com'è, e il suo significato viene a dipendere dall'interpretazione dei singoli componenti.

In modo analogo, data ad esempio la formula $\forall x[0 \leq x \rightarrow \exists y(y \cdot y \approx x)]$, in logica matematica è opportuno considerarla in sé, e poi magari dire che la formula è vera nel campo ordinato dei reali e falsa nel campo ordinato dei razionali, piuttosto che tentare l'impresa disperata di includere nella formula la descrizione dell'interpretazione, dicendo $\forall x \in \mathbf{R}\ [0_R \leq_R x \rightarrow \exists y \in \mathbf{R}\ (y \cdot y \approx x)]$, e magari aggiungendo, sempre all'interno della formula, una definizione formale di $\mathbf{R}$, di 0_R, di $\leq_R$ e di $\cdot$ in qualche teoria degli insiemi.

In generale, dare un'interpretazione di un linguaggio significa:

a) fissare il dominio D del discorso; ad esempio D = l'insieme degli esseri umani;
b) individuare su D le relazioni e proprietà rappresentate nel linguaggio naturale attraverso i verbi e nei linguaggi formali attraverso i predicati; così, ad esempio, "amare"indica una relazione binaria in D fra chi ama e chi è amato; così pure, "essere un uomo" indica una proprietà, un sottinsieme di D;
c) individuare le funzioni descritte nel linguaggio naturale dai costrutti descrittivi (ad esempio, essere la madre di) e, nei linguaggi formali, dai simboli funzionali;
d) determinare quali sono gli individui denotati con nomi propri nel linguaggio naturale e con simboli di costante nei linguaggi formali.

Nell'interpretazione dei linguaggi formali per la logica classica, "individuare" una relazione R rappresentata da un predicato P (rispettivamente: individuare una funzione f rappresentata da un simbolo funzionale F) non significa capire il concetto che si vuole esprimere attraverso P (rispettivamente F), bensì determinare l'insieme delle n-uple (ove n è l'arietà, il numero di posti, di P) di elementi del dominio D che stanno nella relazione R (rispettivamente: determinare l'insieme delle n+1 uple di elementi dominio D il cui ultimo elemento è il valore di f sulla n-upla dei primi n). In altre parole, ciò che conta è solo l'estensione degli enti denotati dai simboli.

Nell'esempio relativo alla frase sul treno per Bari, un'interpretazione dal punto di vista della logica classica dovrà associare a "il treno per Bari" un oggetto A, a "7,10" un istante T, e alla costruzione "partire ... alle..." una relazione R(a,t) che lega fra loro un oggetto a, e un istante t se e solo se l'oggetto a inizia a muoversi all'istante t. La frase risulterà vera se e solo se la relazione di cui sopra è soddisfatta per a=A e t=T. E' chiaro che il messaggio contenuto nella frase "Il treno per Bari..." è ben diverso da questo. In effetti la logica classica

è più appropriata per i linguaggi scientifici (e per quello matematico in particolare), di quanto non lo sia per il linguaggio naturale.

Rileviamo anche che nei linguaggi al primo ordine per la logica classica non ci sono simboli di relazione o funzione da interpretarsi in modo speciale. Ad esempio l'addizione o la relazione "essere minore di" non hanno simboli ad hoc nella logica. L'unica eccezione è data—nella logica con identità—dal simbolo predicativo binario $\approx$ che va interpretato costantemente nella relazione di uguaglianza. Per trattare concetti matematici specifici (ad esempio, punto, retta, numero, addizione), l'insieme delle interpretazioni viene ristretto sì, ma solo in modo indiretto: si richiede cioè alle interpretazioni di soddisfare opportuni assiomi che—almeno secondo l'impostazione formalista—definiscono tali concetti.

Struttura soddisfacibilità modello conseguenza. Strutture di Skolem Herbrand

DEFINIZIONE Sia L un linguaggio al primo ordine. Una *struttura per L* (o *interpretazione di L*) è una coppia ordinata $\mathbf{M} = \langle M, {}^M \rangle$ dove M è un insieme non vuoto, detto *dominio di M*, e ^{M} è una funzione che ad ogni simbolo di costante c associa un elemento c^M di M, ad ogni simbolo di funzione n-aria f associa una funzione f^M su M di arietà n (cioè una funzione da M^n a M), e ad ogni simbolo predicativo n-ario P associa una relazione n-aria P^M su M (cioè un sottoinsieme di P^n).

Può succedere che per due simboli di relazione diversi P e Q si abbia $P^M = Q^M$. E dunque ha rilevanza non solo l'insieme degli elementi, delle relazioni e delle funzioni in $\mathbf{M}$ che interpretano rispettivamente le costanti, i predicati e i simboli funzionali, ma anche la relazione fra interpretante e interpretato; in altre parole, è importante mettere in evidenza che P^M è non solo una relazione che fa parte della struttura $\mathbf{M}$, ma è proprio la relazione che interpreta il simbolo predicativo P.

E' chiaro che le strutture usate in algebra sono strutture secondo questa definizione; ad esempio, $\langle Z,+,\cdot,0,1\rangle$ e $\langle Z,+,\cdot,-,0,1\rangle$ sono due diverse presentazioni dell'anello degli interi; la prima è una struttura per un linguaggio L con due simboli di funzioni binarie e due simboli di costante, la seconda è una struttura per un linguaggio con tre simboli di funzioni binarie e due simboli di costante. Per essere precisi, bisognerebbe scrivere $\langle Z, +_Z, \cdot_Z, 0_Z, 1_Z\rangle$, visto che anche la struttura dei razionali viene spesso denotata con $\langle Q,+,\cdot,0,1\rangle$, col risultato che la somma e il prodotto in Z e in Q si trovano ad essere indicati allo stesso modo, mentre non sono la stessa cosa.

Ribadiamo che nulla ci obbliga a interpretare le operazioni binarie + e $\cdot$ del linguaggio L dell'esempio precedente nella somma e nel prodotto, o le costanti 0 e 1 in zero e uno. Analogamente, il dominio dell'interpretazione non deve essere necessariamente Z.

Se prendiamo come dominio un qualsiasi insieme X non vuoto e interpretiamo +, · e 0,1 rispettivamente in due qualsiasi funzioni binarie su X e in due arbitrari elementi di X, otteniamo ancora una struttura per lo stesso linguaggio L.

Passiamo ora alla definizione di verità di una formula in una struttura; per le formule con variabili libere abbiamo il problema di dire quale elemento del dominio denota ciascuna variabile: ad esempio nella struttura $\langle Z, +_Z, \cdot_Z, 0_Z, 1_Z, = \rangle$ la verità della formula $x+y\approx 0$ dipende dai numeri denotati dalle variabili x e y; la formula sarà soddisfatta se e solo se x e y denotano elementi opposti. Introduciamo quindi la seguente definizione (qui e nel seguito, assumiamo che strutture e formule si riferiscano ad un prefissato linguaggio L).

DEFINIZIONI Un'*assegnazione* in una struttura **M** è una funzione σ dall'insieme delle variabili al dominio M di **M**. Dati un'assegnazione σ in **M** e un termine t, l'elemento t^σ *denotato da* t *attraverso* σ si definisce per induzione sul numero di simboli funzionali in t come segue:

- Se t è una variabile, diciamo $t=x$, allora $t^\sigma=\sigma(x)$.
- Se t è una costante, diciamo $t=c$, allora $t^\sigma=c^M$.
- Se $t = f(t_1,...,t_n)$, ove f è un simbolo di funzione n-aria, e $t_1,...,t_n$ sono termini, allora $t^\sigma = f^M(t_1^\sigma,...,t_n^\sigma)$. Si noti che, per l'ipotesi induttiva, gli elementi $t_1^\sigma,...,t_n^\sigma$ sono già stati definiti.

E' abbastanza facile dimostrare che t^σ dipende solo dai valori di σ sulle variabili che hanno occorrenze in t; se non ce ne sono, t^σ non dipende dall'assegnazione σ, (ma dipende ovviamente da **M**), e pertanto è legittimo scrivere t^M anziché t^σ, cosa che faremo d'ora in poi. Date un'assegnazione σ e una variabile x, chiamiamo *x-variante* di σ ogni assegnazione σ' che differisce da σ al più su x. La *complessità* di una formula è il numero di connettivi e quantificatori in essa.

DEFINIZIONE Sia **M** una struttura, e σ un'assegnazione in **M**. Per ogni formula A, definiamo per induzione sulla *complessità* di A, la nozione "A *è soddisfatta in* **M** *sotto l'assegnazione* σ" (in simboli, $\mathbf{M},\sigma\models A$) in questo modo:

- Se A è atomica, diciamo $A=P(t_1,...,t_n)$, con P predicato n-ario, e $t_1,...,t_n$ termini, allora $\mathbf{M},\sigma\models A$ se e solo se $(t_1^\sigma,...,t_n^\sigma) \in P^M$.
- Se $A=B\wedge C$, assumendo per ipotesi induttiva di avere già definito $\mathbf{M},\sigma\models B$ e $\mathbf{M},\sigma\models C$, definiamo:

  $\mathbf{M},\sigma\models B\wedge C$ se e solo se $\mathbf{M},\sigma\models B$ e $\mathbf{M},\sigma\models C$.

Usando in maniera analoga l'ipotesi induttiva, definiamo:

$\mathbf{M},\sigma\models B\vee C$ se e solo se o $\mathbf{M},\sigma\models B$ oppure $\mathbf{M},\sigma\models C$

$\mathbf{M},\sigma\models \neg B$ se e solo se <u>non</u> $\mathbf{M},\sigma\models B$ (abbreviato con $\mathbf{M},\sigma\not\models B$).

$\mathbf{M},\sigma\models B\rightarrow C$ se e solo se o $\mathbf{M},\sigma\not\models B$ oppure $\mathbf{M},\sigma\models C$

$\mathbf{M},\sigma\models \exists xB$ se e solo se esiste una x-variante σ' di σ tale che $\mathbf{M},\sigma'\models B$

$\mathbf{M},\sigma\models \forall xB$ se e solo se per ogni x-variante σ' di σ si ha $\mathbf{M},\sigma'\models B$

Intuitivamente le ultime due ultime stabiliscono che per vedere se $\mathbf{M},\sigma\models\exists xB$, (rispettivamente: per vedere $\mathbf{M},\sigma\models\forall xB$), dobbiamo tenere fissate le denotazioni delle variabili diverse da x mantenendo i valori assegnati loro da σ, e vedere se, per un opportuno valore b (rispettivamente: per ogni valore b) del dominio, la formula B risulti soddisfatta coll'attribuire a x il valore b.

E' facile vedere che la soddisfacibilità di una formula A in una struttura **M** sotto l'assegnazione σ dipende solo dai valori che σ assume sulle variabili che hanno occorrenze libere in A. In particolare, se A è un enunciato la sua soddisfacibilità in una struttura non dipende dall'assegnazione. Se le variabili che hanno occorrenze libere in A sono $x_1,...,x_n$, e se $a_1,...,a_n$ sono elementi del dominio della struttura **M**, si scrive spesso $\mathbf{M}\models A(a_1/x_1,...,a_n/x_n)$ per dire che $\mathbf{M},\sigma\models A$ per tutte le assegnazioni σ tali che $\sigma(x_1)=a_1,...,\sigma(x_n)=a_n$. Con questa convenzione, si ha:

$\mathbf{M}\models\exists xA(x,a_1/x_1,...,a_n/x_n)$ se e solo se esiste un $b\in M$ tale che $\mathbf{M}\models A(b/x,a_1/x_1,...,a_n/x_n)$

$\mathbf{M}\models\forall xA(x,a_1/x_1,...,a_n/x_n)$ se e solo se per ogni $b\in M$ $\mathbf{M}\models A(b/x,a_1/x_1,...,a_n/x_n)$.

Spesso si sente dire che $\mathbf{M}\models A(a_1/x_1,...,a_n/x_n)$ vale se, sostituendo x_1 con a_1,..., x_n con a_n, la formula $A(x_1,...,x_n)$ diventa vera in M, ma si tratta di un modo di esprimersi scorretto: infatti, sostituendo x_1 con a_1,..., x_n con a_n, in $A(x_1,..., x_n)$ non si ha più una formula, poiché generalmente $a_1,...,a_n$ non sono simboli dell'alfabeto del linguaggio.

DEFINIZIONE Diciamo che una formula A è *vera* nella struttura **M** (e scriviamo $\mathbf{M}\models A$) se e solo se per ogni assegnazione σ su **M**, si ha che $\mathbf{M},\sigma\models A$. A è detta *soddisfacibile* se e solo se è vera in almeno una struttura.

Per quanto visto prima, se A è un enunciato, la soddisfazione di A in una struttura **M** non dipende dall'assegnazione σ, e quindi si ha: $\mathbf{M}\models A$ se e solo se per qualche σ, $\mathbf{M},\sigma\models A$ se e solo se per tutte le σ, $\mathbf{M},\sigma\models A$. In particolare si ha l'aut aut: $\mathbf{M}\models A$, oppure $\mathbf{M}\models\neg A$.

ESEMPI La formula $x\cdot y\approx 1$ è soddisfatta nella struttura dei reali sotto tutte e sole le assegnazioni σ tali che $(\sigma x,\sigma y)$ siano le coordinate di un punto appartenente all'iperbole equilatera avente come asintoti gli assi, e passante per il punto di coordinate (1,1). Né questa formula né la sua negazione sono vere nella struttura dei reali. La formula $\prod$ data da $\forall x\exists y(y\cdot y\approx x)$ è chiusa; pertanto, data una qualunque interpretazione, o $\prod$ o $\neg\prod$ risulta vera in essa (indipendentemente dall'assegnazione); ad esempio, $\prod$ è vera nella struttura dei numeri complessi e falsa in quella dei reali.

DEFINIZIONE Una struttura **M** è detta *struttura di Skolem-Herbrand* per un linguaggio L se, a meno di isomorfismi,

(i) il dominio M di **M** è l'insieme dei termini chiusi di L,
(ii) per ogni simbolo c di costante si ha che $c^M=c$,
(iii) per ogni simbolo funzionale n-ario f, si ha che $f^M(t_1,...,t_n) = f(t_1,...,t_n)$, ove $t_1,...,t_n$ sono termini chiusi arbitrari di L.

Nelle strutture di Skolem Herbrand per un linguaggio L, il dominio e l'interpretazione di costanti e simboli funzionali sono univocamente determinati; ciò che può distinguere una interpretazione di Skolem Herbrand da un'altra è solo l'interpretazione dei simboli di predicati. Inoltre, in una struttura di Skolem Herbrand possiamo definire i concetti di soddisfazione e di verità senza far riferimento ad assegnazioni, ma concentrandosi direttamente sugli enunciati (sui quali, come abbiamo detto, i due concetti coincidono). Precisamente, procedendo per induzione sulla complessità degli enunciati, poniamo:

$\mathbf{M}\models P(t_1,...,t_n)$ (P predicato n-ario, $t_1,...,t_n$ termini chiusi) se e solo se $(t_1,...,t_n)\in P^M$
$\mathbf{M}\models B\wedge C$ se e solo se $\mathbf{M}\models B$ e $\mathbf{M}\models C$
$\mathbf{M}\models B\vee C$ se e solo se o $\mathbf{M}\models B$ oppure $\mathbf{M}\models C$
$\mathbf{M}\models \neg B$ se e solo se $\mathbf{M}\not\models B$
$\mathbf{M}\models B\rightarrow C$ se e solo se o $\mathbf{M}\not\models B$ oppure $\mathbf{M}\models C$
$\mathbf{M}\models \exists xB$ se e solo se esiste un termine chiuso t tale che $\mathbf{M}\models B(t/x)$
$\mathbf{M}\models \forall xB$ se e solo se per ogni termine chiuso t si ha $\mathbf{M}\models B(t/x)$

E' importante notare che, essendo $\exists xB$ un enunciato, così pure sarà B(t/x). A differenza dal caso generale, questa volta con la sostituzione nella formula B di ogni occorrenza libera della variabile x con l'elemento t del dominio di **M** si ottiene ancora una formula (addirittura un enunciato) poiché gli elementi del dominio sono termini del linguaggio. In questo caso, e solo in questo caso, è sensato parlare della formula ottenuta sostituendo elementi del dominio alle variabili di una formula. Solo in riferimento a questo tipo di

strutture può aver senso dire che "un elemento è soluzione di un'equazione quando sostituendolo alla variabile nell'equazione si ottiene una formula vera".

DEFINIZIONI Sia Σ un insieme di enunciati nel linguaggio L. Diciamo che una struttura **M** per L è un *modello* di Σ, e scriviamo $\mathbf{M}\models\Sigma$, se e solo se $\mathbf{M}\models A$ per ogni enunciato A di Σ. Un enunciato A di L si dice *conseguenza logica* di un insieme di enunciati Σ di L (e si scrive $\Sigma\models A$) se e solo se in ogni modello **M** di Σ si ha $\mathbf{M}\models A$. A è *logicamente valido* (e si scrive $\models A$) se e solo se è conseguenza logica dell'insieme vuoto, cioè se e solo se per ogni struttura **M** per L è $\mathbf{M}\models A$.

Teoremi di completezza di Gödel

Enunciamo due versioni del Teorema di Completezza (Gödel 1930):

TEOREMA DI COMPLETEZZA DEBOLE Per ogni enunciato A si ha: $\models A$ se e solo se $\vdash A$.

TEOREMA DI COMPLETEZZA FORTE Per ogni insieme di enunciati Σ e per ogni enunciato A (di un linguaggio L), si ha: $\Sigma\models A$ se e solo se $\Sigma\vdash A$.

Quindi: A è logicamente valido se e solo se A è dimostrabile; A è conseguenza logica di Σ se e solo se A è deducibile da Σ.

Prima di affrontare la dimostrazione del Teorema di completezza per la logica classica, discutiamone il significato. Ci sono analoghi teoremi di completezza per la logica intuizionista, per la logica lineare, etc. In generale un teorema di completezza per una logica stabilisce un'equivalenza fra la dimostrabilità (secondo le regole di un calcolo logico per quella logica) e la verità in tutte le possibili interpretazioni del linguaggio per quella logica. Qui "logica" indica sia una classe di linguaggi, sia un insieme di (assiomi e di) regole per dedurre, sia una classe di interpretazioni per quei linguaggi.

Si stabilisce così l'equivalenza fra concetti sostanzialmente diversi: da un lato, per il fatto che le dimostrazioni sono oggetti finiti, la dimostrabilità ha un carattere finitistico e controllabile: si possono scrivere programmi per stabilire, data una configurazione (sequenza, o albero, o grafo) di formule, se essa sia o no una dimostrazione in un dato calcolo logico; dall'altro lato, il concetto di validità è un concetto astratto, infinitistico: secondo la definizione, per verificare la validità di un enunciato, dovremmo completare l'esame di un'infinità di strutture per il linguaggio preso in considerazione—senza dimenticare che le strutture stesse hanno generalmente dominio infinito.

Nel caso della logica classica è abbastanza facile convincersi della correttezza delle regole di uno qualsiasi dei calcoli deduttivi: ciò che è dimostrabile attraverso tali regole è anche valido, e ciò che è deducibile da premesse tramite le regole stesse è conseguenza

logica delle premesse. Ciò che non è a priori ovvio è la completezza del calcolo, ossia il fatto che le regole bastino per ottenere da un insieme di premesse tutte le conseguenze logiche dell'insieme stesso, e dall'insieme vuoto di premesse gli enunciati validi. Ad esempio, gli assiomi e le regole della logica aristotelica sono corretti, ma non completi—tra l'altro, essi non prendono in considerazione i predicati con più di un argomento. Per dimostrare il Teorema di completezza forte basta dimostrare il seguente Teorema, ad esso equivalente:

TEOREMA DEL MODELLO Un insieme Σ di enunciati è coerente (non contraddittorio) se e solo se Σ ha un modello.

Dimostrazione del Teorema di Completezza partendo dal Teorema del Modello. Assumiamo il Teorema del Modello. Allora abbiamo che $\Sigma \models A$ se e solo se non ci sono modelli di $\Sigma \cup \{\neg A\}$ se e solo $\Sigma \cup \{\neg A\}$ è incoerente (per il Teorema del Modello) se e solo se (per la reductio ad absurdum) $\Sigma \vdash A$.

Prima di dimostrare il Teorema del Modello, è opportuno inquadrarlo nell'ambito del Programma di Hilbert. Tale programma prevedeva di fondare la matematica sui sistemi formali; per Hilbert, gli assiomi delle teorie matematiche assumevano il ruolo di definizioni implicite degli enti descritti dalla teoria, e l'unico requisito per l'esistenza di detti enti era la consistenza della teoria che li definisce:

> "Se assiomi arbitrariamente stabiliti non sono in contraddizione con tutte le loro conseguenze, allora essi sono veri, allora esistono gli enti definiti per mezzo di questi assiomi. Questo è per me il criterio della verità e dell'esistenza"
>
> (D. HILBERT).

A prima vista, una frase come questa può sembrare pura follia, ma è chiaro che il Teorema del Modello costituisce un notevole sostegno a questa tesi: se un sistema di assiomi è coerente, esiste un modello in cui detti assiomi sono veri, e quindi gli enti di cui gli assiomi parlano esistono (nel modello stesso). Purtroppo o per fortuna, il Programma di Hilbert era destinato a ricevere dallo stesso Gödel, che con il Teorema di Completezza lo aveva sostenuto, un duro colpo. Un ovvio corollario della frase di Hilbert precedentemente citata è che per credere agli enti di una teoria occorre dimostrarne la consistenza; ebbene lo stesso Gödel avrebbe dimostrato poco dopo l'impossibilità di dimostrazioni di consistenza di teorie forti almeno come l'Aritmetica di Peano, se non con metodi ancora più forti di quelli formalizzabili nella teoria, e quindi difficilmente giustificabili da un punto di vista finitistico.

Dimostrazione informale del teorema del modello

Supponiamo innanzitutto di avere a che fare con un insieme finito di enunciati, $A_1,...,A_n$. Il procedimento che descriveremo ci darà o un modello per $\{A_1,...,A_n\}$, oppure una prova di inconsistenza dell'insieme $\{A_1,...,A_n\}$ con (una lieve variante del) metodo dei tableaux. Poiché il metodo non è del tutto elementare ne daremo prima una presentazione informale.

Tenteremo di costruire un modello di Skolem Herbrand di $\{A_1,...,A_n\}$; ad ogni passo, gli enunciati da rendere veri nel modello che vogliamo costruire verranno ridotti ad enunciati più semplici; spesso introdurremo anche enunciati che dovranno invece essere falsi nel modello finale. Ad un certo punto è possibile che ci si renda conto della manifesta impossibilità di costruire un modello che soddisfi le richieste di cui sopra, e allora l'insieme $\{A_1,...,A_n\}$ verrà dichiarato insoddisfacibile. Per spiegare come vengono effettuate le riduzioni, concentriamoci su uno degli enunciati, diciamo A_i.

—Se A_i ha la forma $B \wedge C$, si richiederà che sia B che C siano veri nel modello finale, e quindi aggiungeremo sia B che C alla lista delle formule da rendere vere; la richiesta che B e C siano vere nel modello finale è equivalente alla richiesta di verità di $B \wedge C$; quindi, una volta aggiunto B e C, contrassegnamo $B \wedge C$ con un asterisco, che vuol dire: "non ci dobbiamo più preoccupare di questo enunciato, il problema della sua verità è stato interamente ricondotto a quello della verità di nuovi, più semplici enunciati".

—Se invece A_i ha la forma $\neg B$, allora si dovrà richiedere che B sia falso nel modello finale. Una volta che abbiamo aggiunto B fra gli enunciati da falsificare, non ci dobbiamo più preoccupare di $\neg B$, e quindi lo asterischiamo.

—Se A_i ha la forma $B \vee C$, ci basta rendere vero o B o C; non avendo alcuna idea su quale dei due convenga rendere vero, salomonicamente ci faremo in due, biforcando il procedimento in due sottoprocedimenti: in uno cerchiamo di rendere vero B, nell'altro di rendere vero C; è chiaro che per ottenere il modello cercato basta che almeno uno dei sottoprocedimenti abbia successo; infine, una volta aggiunto B da una parte e C dall'altra fra gli enunciati da rendere veri, possiamo asteriscare $B \vee C$.

—Supponiamo ora che A_i abbia la forma $\exists xB$; poiché stiamo cercando un modello di Skolem Herbrand, affinché A_i sia vero (nel modello cercato) dovremo trovare un termine chiuso t tale che B(t/x) sia vero; in mancanza di idee migliori, prenderemo come t un nuovo simbolo di costante c, che non sia ancora comparso nel procedimento, e ci proporremo di rendere vero l'enunciato B(c/x). Il ruolo che ha per noi questo c è di denotare un elemento del tutto generico, come Caio, allorché Tizio e Sempronio non sono disponibili; anche nella pratica dimostrativa matematica, quando si assume che esista un elemento che soddisfa la proprietà P ma non si sa quale sia, si dice: "sia c un generico elemento tale che P(c)". E' chiaro che affinché c sia davvero generico, occorre che su c non sia mai stato detto nulla, cioè in altre parole, che c non sia ancora comparso nel procedimento. Questo rende necessario

aggiungere al linguaggio un'infinità di nuove costanti $\{c_1,...,c_n,...\}$ per denotare elementi ciascuno dei quali avente l'unica particolarità di far da testimone a un certo enunciato di tipo $\exists x...$, ma che per il resto, proprio come Tizio, Caio, Sempronio, viene considerato del tutto generico. Ancora una volta, dopo aver aggiunto B(c/x) fra gli enunciati da rendere veri, asterischiamo $\exists xB$.

—Infine, se A_i è della forma $\forall xB$, dovremo rendere vere tutte le istanze B(t/x), per ogni possibile termine chiuso t. Supponiamo di voler prendere di petto il problema, aggiungendo tutti gli enunciati B(t/x) in una volta, e accollandoci un insieme infinito: rischieremo perdite di tempo infinite, pregiudicando irreparabilmente l'effettività del nostro procedere. Scartata dunque questa idea, aggiungiamo B(t*/x) fra gli enunciati da rendere veri, dove t* è il primo termine chiuso tale che B(t*/x) non è ancora stata messa fra le formule da rendere vere (beninteso, avendo avuto cura di prefissare una volta per tutte, ancor prima di partire, una elencazione di tutti i termini chiusi). A questo punto però, non potendo pretendere di esserci tolta la preoccupazione di $\forall xB$, non faremo l'errore di depennarlo con un asterisco: dobbiamo ricordarci di aggiungere anche gli altri enunciati del tipo B(t/x), per ogni possibile termine chiuso t; quello che facciamo è invece di massima acutezza e onestà: asterischiamo sì $\forall xB$, ma al tempo stesso aggiungiamo, non asteriscato, un nuovo $\forall xB$ in fondo alla lista degli enunciati da rendere veri. Benché per il momento questo nuovo arrivato $\forall xB$ non dia nell'occhio, anche per la sua posizione in fondo alla lista, prima o poi dovremo rioccuparcene, e ciò è molto giusto, visto che non abbiamo aggiunto tutte le istanze B(x/t).

Avendo appena terminato di descrivere il trattamento per gli enunciati da rendere veri, i lettori attenti non avranno dubbi sul come—simmetricamente—comportarsi con gli enunciati da rendere falsi: ad esempio, dovendo falsificare $B \vee C$, falsificheremo sia B che C e poi asterischeremo $B \vee C$; dovendo falsificare B, lo asterischeremo, e procederemo con il compito di rendere vero B; dovendo falsificare $B \wedge C$, ci faremo in due, biforcando il procedimento, e ponendoci il compito di falsificare almeno uno tra B e C; e analogamente procederemo per i restanti due casi.

Supponiamo che, proseguendo passo dopo passo ad asteriscare, biforcare, sostituire, collocare, semplificare, secondo la procedura, ad un certo punto succeda che, in tutti i sottoprocedimenti ottenuti per biforcazione, un enunciato compaia sia fra gli enunciati da verificare sia fra quelli da falsificare: allora avremo ottime ragioni per pensare che ogni ulteriore tentativo di trovare un modello non potrà avere successo, visto che un enunciato non può essere sia vero che falso in un modello. Se succede questo, consideriamo il procedimento svolto come una confutazione dell'insieme $\{A_1,...,A_n\}$; se ciò non succede mai, mostreremo come costruire un modello di $\{A_1,...,A_n\}$.

Come abbiamo visto, il nostro insieme di enunciati si evolve nel tempo durante il procedimento, talora sdoppiandosi; per maggiore chiarezza, ne descriveremo l'evoluzione tenendo presente tutti gli stati temporali, disposti ad albero, dove gli insiemi si evolvono dal basso all'alto, e le biforcazioni corrispondono agli sdoppiamenti.

Un'ultima cosa di cui preoccuparci è che, prima o poi, ogni enunciato venga davvero preso in considerazione: ciò non è garantito a priori, visto che ad ogni passo si aggiungono nuovi enunciati. Per evitare omissioni d'atti d'ufficio, come in tutte le code che si rispettino, daremo un numero d'ordine agli enunciati, con l'intesa che quelli che vengono via via aggiunti passino in fondo alla coda: questo garantisce che nessun enunciato debba aspettare in eterno, come succederebbe ad esempio se gli ultimi arrivati continuassero a passare davanti, saltando la coda. Nella realtà della nostra procedura, gli enunciati sono "disposti su due colonne", quella degli enunciati da rendere veri e quella degli enunciati da rendere falsi. Il numero d'ordine sarà tale da alternare un enunciato di una colonna e un enunciato dell'altra—altrimenti, se una colonna avesse la precedenza, agli enunciati dell' altra potrebbe capitare di non essere mai presi in considerazione.

Dimostrazione del teorema: prima parte

a) Introduciamo una sequenza numerabile di nuove costanti, diciamo $\{c_1,...,c_n,...\}$ e chiamiamo L' il linguaggio così esteso; sia $t_1,...,t_n,...$ un'enumerazione dei termini chiusi di L'.

b) Una *tavola* è una coppia ordinata di sequenze finite di enunciati, alcuni dei quali contrassegnati con un asterisco (*asteriscati*); la prima sequenza sarà chiamata *colonna* V (l'idea è che in essa saranno messi gli enunciati da rendere veri), la seconda sequenza verrà chiamata *colonna* F (in essa verranno messi gli enunciati da rendere falsi). Una colonna può anche essere vuota. Una tavola viene indicata con $A_1,...,A_n$; $B_1,...B_m$, dove resta inteso che gli A_i stanno nella colonna V e che i B_j stanno nella colonna F il simbolo ";" separa le due colonne).

Diamo un ordine alle formule della tavola, detto *ordine della tavola*; conveniamo che formule con numero d'ordine inferiore precedano quelle con numero d'ordine superiore; a parità di numero d'ordine, la formula nella colonna V precede quella nella colonna F: ad esempio, B_i precede A_{i+1} che precede B_{i+1}; quindi l'ordine sarà: $A_1, B_1, A_2, B_2,...$. Durante il procedimento una tavola T può essere trasformata in una tavola T' o divisa in due *sottotavole* T_1 e T_2. In questo caso diciamo che T' è il *successore* di T e che T è il *predecessore* di T' (rispettivamente che T_1 e T_2 sono i successori di T e che T è il predecessore di T_1 e T_2) e scriviamo $T \angle T'$, (rispettivamente, $T \angle T_1$, $T \angle T_2$). Diciamo che T' *sta sopra* T (e scriviamo T<T') se esistono $T_1,...T_n$ tali che $T=T_1 \angle ... \angle T_n=T'$.

Ad ogni stadio della costruzione le tavole ottenute formano un albero binario rispetto alla relazione $<$. Allo stadio iniziale, l'insieme delle formule da falsificare è vuoto, e quindi abbiamo la tavola $A_1,...,A_n$; (non c'è nulla dopo il ";"). Tale tavola è detta *radice* dell'albero (è la tavola più in basso). Una tavola T è detta una tavola *foglia* ad un certo stadio della costruzione se a quello stadio non ci sono tavole T' con $T \angle T'$. Intuitivamente, una tavola è una tavola foglia allo stadio n se a quello stadio non c'è nulla sopra di essa.

Si noti che una tavola può essere foglia ad un certo stadio n, ma non esserla allo stadio n+1, in quanto a tale stadio nuove tavole possono essere state aggiunte sopra T. In particolare, allo stadio 0, la tavola iniziale è sia radice che foglia. Una tavola foglia si dice *chiusa* se in essa uno stesso enunciato compare sia nella colonna V che nella colonna F. Una tavola foglia si dice *terminata non chiusa* se non è chiusa e tutti i suoi enunciati sono asteriscati.

c) STADIO n. Consideriamo tutte le tavole che erano foglie allo stadio precedente, in un ordine fissato—ad esempio da sinistra a destra; se n=0, consideriamo la tavola iniziale. Se tutte le tavole foglia sono chiuse, il procedimento termina, e viene considerato una refutazione della tavola iniziale, ossia dell'insieme $\{A_1,...,A_n\}$. Se almeno una tavola foglia è terminata non chiusa, il procedimento termina; edremo in seguito come in questo caso si possa costruire un modello di $\{A_1,...,A_n\}$. Altrimenti sia T la prima tavola foglia non chiusa (e non terminata) non ancora considerata; consideriamo la formule di T nell'ordine della tavola; sia A la prima formula non ancora asteriscata.

Se A è atomica, la asterischiamo (non importa se A è nella colonna V o nella colonna F), ottenendo una nuova tavola T', che mettiamo immediatamente sopra T; poniamo cioè per definizione $T \angle T'$, come nella figura seguente:

T' .....A*,... ; ..... T' .... ; ... A*,...

T .....A,.... ; ..... T ... ; ... A ,....

Se A è della forma $\neg B$ ed è nella colonna V, la asterischiamo, aggiungiamo B alla fine della colonna F, ottenendo una nuova tavola T', e poniamo $T \angle T'$.

Se A è della forma $\neg B$ ed è nella colonna F, la asterischiamo e aggiungiamo B alla fine della colonna V, ottenendo una nuova tavola T', e poniamo $T \angle T'$. I due casi sono illustrati nella figura seguente

T' .....¬B*,... ; .....,B T' ...B ; .... ¬B*,...

T .....¬B,.... ; ..... T ..... ; ... ¬B ,....

Se A è della forma $B \wedge C$ ed è nella colonna V, la asterischiamo e aggiungiamo sia B che C nella colonna V, ottenendo una nuova tavola T', e poniamo $T \angle T'$.

Se A è della forma $B \wedge C$ ed è nella colonna F, costruiamo due nuove tavole T_1 e T_2, entrambe aventi gli stessi elementi della precedente, nelle stesse colonne, con lo stesso ordine e con gli stessi asterischi, con le seguenti eccezioni: (i) $B \wedge C$ è asteriscata sia in T_1 che in T_2; (ii) in fondo alla colonna F di T_1 viene aggiunta B; (iii) in fondo alla colonna F di T_2 viene aggiunta C. Poniamo poi $T \angle T_1$ e $T \angle T_2$. I due casi sono illustrati nella seguente figura

T' ....(B∧C)*...B,C ; ..... T_1 ... ; ...(B∧C)*...B T ... ; ....(B∧C)*...B

T .....B∧C... ; ..... T ..... ; ....B∧C.....

Se A è della forma $B \vee C$ ed è nella colonna V, costruiamo due nuove tavole T_1 e T_2, entrambe aventi gli stessi elementi della precedente, nelle stesse colonne, con lo stesso ordine e con gli stessi asterischi, con le seguenti eccezioni: (i) $B \vee C$ è asteriscata sia in T_1 che in T_2; (ii) in fondo alla colonna V di T_1 viene aggiunta B; (iii) in fondo alla colonna V di T_2 viene aggiunta C. Poniamo poi $T \angle T_1$ e $T \angle T_2$.

Se A è della forma $B \vee C$ ed è nella colonna F, la asterischiamo e aggiungiamo B e C in fondo alla colonna F, ottenendo una nuova tavola T', e poniamo $T \angle T'$. I due casi sono illustrati nella seguente figura

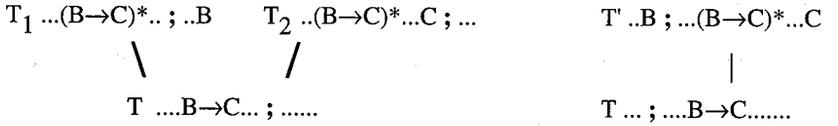
T_1 ...(B∨C)*...B ; ... T_2 ...(B∨C)*...C ; ... T' ... ; ...(B∨C)*...B,C

T ...B∨C... ; ..... T ... ; ... B∨C.....

Se A è della forma $B \rightarrow C$ ed è nella colonna V, costruiamo due nuove tavole T_1 e T_2, entrambe aventi gli stessi elementi della precedente, nelle stesse colonne, con lo stesso ordine e con gli stessi asterischi, con le seguenti eccezioni: (i) $B \rightarrow C$ è asteriscata sia in T_1 che in T_2; (ii) in fondo alla colonna F di T_1 viene aggiunta B; (iii) in fondo alla colonna V di T_2 viene aggiunta C. Poniamo poi $T \angle T_1$ e $T \angle T_2$.

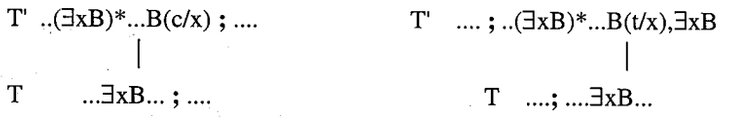
Se A è della forma $B \rightarrow C$ ed è nella colonna F, la asterischiamo e aggiungiamo B in fondo alla colonna V e C in fondo alla colonna F, ottenendo una nuova tavola T'. Poniamo poi $T \angle T'$.

I due casi sono illustrati nella seguente figura



Se A è della forma $\exists xB$ ed è nella colonna V, la asterischiamo ed aggiungiamo B(c/x) in fondo alla colonna V, dove c è la prima fra le costanti $\{c_1,...,c_n,...\}$ che non compare in alcuna delle formule finora introdotte, ottenendo una nuova tavola T'. Poniamo poi T∠T'.

Se A è della forma $\exists xB$ ed è nella colonna F, la asterischiamo, aggiungiamo B(t/x) in fondo alla colonna F, dove t è il primo termine nella enumerazione $t_1,...,t_n,...$ tale che B(t/x) non compare nella colonna F della tavola, e reintroduciamo $\exists xB$ alla fine della colonna F, ottenendo una nuova tavola T'. Poniamo poi T∠T'. I due casi sono illustrati nella figura seguente



Se A è della forma $\forall xB$ ed è nella colonna V, la asterischiamo, aggiungiamo B(t/x) in fondo alla colonna V, dove t è il primo termine nella enumerazione dei termini chiusi $t_1,...,t_n,...$ tale che B(t/x) non compare nella colonna V, e reintroduciamo $\forall xB$ alla fine della colonna V, ottenendo una nuova tavola T'. Poniamo poi T∠T'.

Se A è della forma $\forall xB$ ed è nella colonna F, la asterischiamo e aggiungiamo B(c/x) in fondo alla colonna F, dove c è la prima fra le costanti nella sequenza $\{c_1,...,c_n,...\}$ sopra definita che non compare in alcuna formula nelle tavole finora costruite, ottenendo una nuova tavola T'. Poniamo poi T∠T'. I due casi sono illustrati nella seguente figura

Si passa poi alla tavola foglia immediatamente a destra, e si continua fino ad aver esaurito tutte la tavole foglia. Infine si passa allo stadio successivo ripetendo lo stesso tipo di procedimento rispetto alle nuove tavole foglia. Con questo abbiamo terminato la descrizione del procedimento formale.

Dimostrazione del teorema: seconda parte

Sono possibili tre casi:

Caso a) Il procedimento ha termine perché a un certo stadio tutte le tavole foglia sono chiuse.

Caso b) Il procedimento ha termine perché a un certo stadio almeno una tavola foglia è terminata non chiusa.

Caso c) Il procedimento non termina.

DEFINIZIONE La configurazione prodotta applicando il metodo sopra esposto—finita nei casi a) e b), e infinita nel caso c)—si chiama *albero di riduzione* della tavola iniziale $\{A_1,...,A_n\}$.

DEFINIZIONE Diciamo che la tavola T è *soddisfatta* dal modello **M** (o che **M** soddisfa T) se tutti gli enunciati nella colonna V di T sono veri in **M** e tutti quelli nella colonna F di T sono falsi in **M**.

LEMMA Sia T una tavola nell'albero di riduzione avente almeno un successore; allora, se esiste un modello che soddisfa T, esiste un modello che soddisfa almeno un successore di T.

Dimostrazione Sia **M** un modello che soddisfa T. Poiché T ha un successore, T non può essere chiusa, né terminata non chiusa. Occorre esaminare tutti i casi possibili in cui sono prodotti i successori (o il successore) di T. Vediamone solo alcuni, lasciando gli altri per esercizio.

Supponiamo che la formula di T su cui si agisce sia B→C, e che essa sia nella colonna V; allo stadio successivo vengono introdotti due successori di T, T_1 e T_2; a parte gli asterischi, T_1 differisce da T per la presenza di B nella colonna F, e T_2 differisce da T per la presenza di C nella colonna V; dobbiamo dunque preoccuparci solo di B e di C. Ora, **M** $\models$B→C; quindi, o **M** $\not\models$B o **M** $\models$C; nel primo caso **M** soddisfa T_1; nel secondo caso, **M** soddisfa T_2.

Supponiamo ora che la formula su cui si agisce sia ancora B→C, ma che sia nella colonna F; in questo caso, allo stadio successivo si aggiunge a T la formula B nella colonna V e C nella colonna F, ottenendo una tavola T' che costituisce il successore di T; poiché **M** soddisfa T, e T' differisce da T per la presenza di B nella colonna V e di C nella colonna F, per dimostrare che **M** soddisfa T' basta dimostrare che **M** $\models$B e che **M** $\not\models$C; ma questo segue dal fatto che **M** $\not\models$B→C.

Supponiamo ora che la formula su cui si agisce sia $\exists xB$, e che sia nella colonna V; allo stadio n+1, si aggiunge allora a T una formula B(c/x), c costante non ancora utilizzata, nella colonna V, ottenendo una tavola foglia T' che costituisce il successore di T; sia **M** un modello che soddisfa T; in particolare **M** $\models\exists xB$, onde esiste un a in **M** tale che **M** $\models$B(a/x);

considero un modello **M'** che coincide con **M** salvo al più per l'interpretazione di c, che viene definita ponendo $c^{M'}=a$. Si noti che ogni enunciato che non contiene c è vero in **M'** se e solo se è vero in **M**; quindi **M'** concorda con **M** su tutte le formule della tavola, salvo eventualmente su B(c/x); per dimostrare che **M'** soddisfa T' basta quindi dimostrare che in **M** $\models$B(c/x); ma questo segue dal fatto che c viene interpretata in a, e che **M** $\models$B(a/x), e quindi **M'**$\models$B(a/x).

Supponiamo poi che la formula su cui si agisce sia sempre $\exists$xB, ma che sia nella colonna F; in questo caso, si aggiunge una formula del tipo B(t/x), t termine chiuso, nella colonna F, ottenendo una nuova tavola foglia T' che costituisce il successore di T; sia **M** un modello che soddisfa T; al solito, basta dimostrare che **M**$\models$B(t/x); sia t^M l'interpretazione di t in **M**; se fosse **M** $\models$B(t/x), avremmo **M**$\models$B(t^M/x), onde esisterebbe un elemento b (precisamente $b=t^M$) di **M** tale che **M** $\models$B(b/x); ne seguirebbe **M**$\models\exists$xB, contro quanto supposto.

Gli altri casi sono simili. CVD

COROLLARIO Se tutte le tavole foglia sono chiuse ad un certo stadio n, allora l'insieme $\{A_1,...,A_n\}$ di partenza non ha modelli.

Dimostrazione Chiaramente una tavola chiusa non è soddisfatta in alcun modello; se la tavola iniziale fosse soddisfatta in qualche modello, per il Lemma, qualche suo successore lo sarebbe; lo stesso varrebbe per qualche successore di tale successore; procedendo in questo modo (per dire bene le cose bisognerebbe ragionare per induzione), si troverebbe una tavola foglia soddisfatta in un modello, assurdo, perché le tavole foglia sono tutte chiuse. CVD

Come abbiamo detto, se ad un certo stadio ogni tavola foglia è chiusa, consideriamo (per definizione) la configurazione prodotta a quello stadio come una refutazione di $\{A_1,...,A_n\}$. Se si prescinde dall'interpretazione semantica delle tavole e delle colonne V e F, la configurazione prodotta può essere considerata come entità puramente sintattica, essendo prodotta combinando sequenze di formule e disponendole opportunamente secondo regole puramente meccaniche. Ovviamente, possiamo usare il metodo anche per dimostrare e non solo per refutare: basta considerare una refutazione di $\neg$A come una prova di A, e una refutazione di $\{A_1,...,A_n,\neg A\}$ come una derivazione di A da $\{A_1,...,A_n\}$. Più precisamente, volendo dedurre un enunciato A da un insieme di enunciati $\{A_1,...,A_n\}$, si parte dalla tavola $A_1,...,A_n,\neg$A; e si applica la procedura descritta sopra: se a un certo stadio si arriva a un albero in cui tutte le tavole foglia sono chiuse, la configurazione ottenuta è per definizione una deduzione di A da $\{A_1,...,A_n\}$. Questo metodo di dimostrazione consente di dimostrare esattamente gli enunciati dimostrabili in deduzione naturale o in altri sistemi di dimostrazione.

Il Corollario dimostrato sopra dice che, se $\{A_1,...,A_n\}$ ha un modello, non esiste alcuna refutazione (nel senso precisato sopra) di $\{A_1,...,A_n\}$. Per completare la dimostrazione, occorre far vedere che viceversa, se non si riesce a refutare $\{A_1,...,A_n\}$, cioè se non si verifica mai che tutte le tavole foglia siano chiuse, allora $\{A_1,...,A_n\}$ ha un modello.

DEFINIZIONE Un *ramo finito* nell'albero di riduzione è una successione finita di tavole $T_1,T_2...,T_n$ dove T_1 è la tavola radice, $T_i \angle T_{i+1}$ (i=1,...,n-1), e T_n è una tavola foglia; un *ramo infinito* è una successione infinita $T_1,T_2,...,T_n,...$ dove T_1 è la tavola radice, e, per ogni i, $T_i \angle T_{i+1}$. Se un oggetto è un ramo finito o infinito, lo chiamiamo semplicemente *ramo*.

Osserviamo che nel caso b), quando cioè c'è una tavola terminata non chiusa T, esiste un ramo finito che termina con tale tavola, mentre nel caso c), quando cioè la procedura non ha termine, esiste un ramo infinito (dimostralo!). Pertanto se il caso a) non si verifica, esiste un ramo R che o è infinito oppure termina con una tavola terminata non chiusa.

Sia **F** l'insieme delle formule che si trovano nella colonna F di qualche tavola in R, e sia **V** l'insieme delle formule che si trovano nella colonna V di qualche tavola in R.

LEMMA Gli insiemi **F** e **V** sono disgiunti.

Dimostrazione Se così non fosse, ci sarebbe una formula A che occorre sia nella colonna F di una certa T_i in R, sia nella colonna V di una T_k in R. Supposto senza perdita di generalità $i \leq k$, poiché nessuna formula viene tolta in nessuna colonna nel passaggio da una tavola al successore, la formula A, comparendo nella colonna F di T_i, comparirebbe anche nella colonna F di T_k, e T_k sarebbe chiusa, contro quanto supposto. CVD

Siamo pronti per definire una struttura **M** di Skolem Herbrand che risulterà modello di $\{A_1,...,A_n\}$. Come dominio M di **M**, prendiamo i termini chiusi di L'. Come al solito, per ogni simbolo c di costante poniamo $c^M=c$, e per ogni simbolo funzionale n-ario f, e per ogni n-upla di termini chiusi $t_1,...,t_n$, poniamo $f^M(t_1,...,t_n)=f(t_1,...,t_n)$. Per ogni simbolo predicativo n-ario P, e per ogni n-upla di termini chiusi $t_1,...,t_n$, poniamo $(t_1,...,t_n)\in P^M$ se e solo se $P(t_1,...,t_n)\in$ **V**.

LEMMA Per ogni enunciato A, se A$\in$ **V** allora **M** $\models$A; se A$\in$ **F** allora **M**$\not\models$A.

(Si noti che, essendo $A_1,...,A_n$ in **V**, dal Lemma segue che **M** è un modello di $\{A_1,...,A_n\}$, come volevasi).

Dimostrazione Induzione sulla complessità di A. Ci limitiamo ad esaminare qualche caso.

A atomico, diciamo $A=P(t_1,...,t_n)$. Per definizione, $P(t_1,...,t_n)$ è vero in **M** se e solo se $P(t_1,...,t_n) \in$ **V**; quindi, se A$\in$ **V**, **M** $\models$A; poiché **V** e **F** sono disgiunti, se A$\in$ **F**, A$\notin$ **V**, onde **M**$\not\models$A.

$A=B\vee C$. Supponiamo $A\in \mathbf{V}$; si vede facilmente che esiste una tavola T_i di R tale che A compare in T_i nella colonna V come prima formula non asteriscata; T_i ha due successori, uno dei quali è in R, e contiene nella colonna V o B o C. Quindi o $B\in \mathbf{V}$ o $C\in \mathbf{V}$. Per ipotesi induttiva, $\mathbf{M}\models B$ o $\mathbf{M}\models C$, e quindi $\mathbf{M}\models B\vee C$. Supponiamo ora $A=B\vee C\in \mathbf{F}$. Esiste una tavola T_i di R tale che A compare in T_i nella colonna V come prima formula non asteriscata; T_i ha un unico successore $T_{i+1}\in R$. Inoltre, sia B che C occorrono nella colonna F di T_{i+1}, e quindi sia B che C appartengono a **F**. Per ipotesi induttiva, $\mathbf{M}\not\models B$, $\mathbf{M}\not\models C$, onde $\mathbf{M}\not\models B\vee C$.

$A=\exists xB$. Supponiamo prima $A\in \mathbf{V}$; sia T_i una tavola di R tale che A compaia in T_i nella colonna V come prima formula non asteriscata; T_i ha un unico successore $T_{i+1}\in R$; T_{i+1} avrà una formula del tipo B(c/x) nella colonna V; per ipotesi induttiva, $\mathbf{M}\models B(c/x)$; quindi esiste un $b\in M$ (precisamente b=c) tale che $\mathbf{M}\models B(b/x)$; ne segue $\mathbf{M}\models \exists xB$.

Supponiamo infine che $A=\exists xB\in \mathbf{F}$; è facile convincersi che in questo caso il ramo R è infinito; sia infatti T_i una tavola di R tale che A compare in T_i nella colonna V come prima formula non asteriscata; la formula $\exists xB$ viene asteriscata, ma reintrodotta in fondo alla colonna F di T_{i+1}. Esisterà allora una tavola T_k sopra T_{i+1} tale che $\exists xB$ compaia in T_k nella colonna F come prima formula non asteriscata; ancora una volta $\exists xB$ viene asteriscata ma reintrodotta, per cui il procedimento non termina mai con una tavola di formule tutte asteriscate. Ora, secondo il procedimento, ogni volta che ho $\exists xB$ come prima formula non asteriscata nella colonna F di una tavola T, devo aggiungere B(t/x) in fondo alla colonna F, dove t è il primo termine chiuso nella nostra enumerazione tale che B(t/x) non occorre nella colonna F di T; poiché, come abbiamo visto, questo si verifica infinite volte nelle tavole del ramo R, ne segue che per ogni termine chiuso t la formula B(t/x) occorre nella colonna F di qualche tavola di R. Pertanto ogni formula del tipo B(t/x) è in **F**; per ipotesi induttiva, per ogni termine chiuso t di L', è $\mathbf{M}\not\models B(t/x)$; poiché M è costituito da tutti e soli i termini chiusi di L', se ne deduce $\mathbf{M}\not\models \exists xB$.

Gli altri casi sono simili. CVD

Abbiamo appena concluso la dimostrazione del Teorema del Modello nel caso in cui l'insieme Σ di enunciati sia finito. Se l'insieme Σ di enunciati è infinito ma è numerabile, (cosa che accade inevitabilmente se il linguaggio è numerabile), basta modificare lievemente il procedimento; l'idea è quella di predisporre una enumerazione $A_1,\ldots,A_n,\ldots$ di Σ, e di aggiungere poi al termine di ogni passo un enunciato alla volta in fondo alla colonna V di ogni tavola foglia. Si noti che in questo caso una tavola non risulta mai terminata non chiusa, perché a ogni passo introduciamo un enunciato nuovo non asteriscato. Non entriamo nei dettagli della dimostrazione, che però non differisce di molto dalla precedente.

ESEMPIO Consideriamo l'insieme $\{P(c), P(c)\rightarrow Q(c), \neg Q(d)\}$, P, Q predicati unari, c e d costanti distinte. Applicando il metodo delle tavole semantiche, e indicando per semplicità solo le tavole foglia ad ogni stadio della costruzione, otteniamo successivamente:

$P(c), P(c)\rightarrow Q(c), \neg Q(d)$; (radice), che ha come successore

$P(c)^*, P(c)\rightarrow Q(c), \neg Q(d)$; che si sdoppia in

$P(c)^*, (P(c)\rightarrow Q(c))^*, \neg Q(d)$; $P(c)$ e $P(c)^*, (P(c)\rightarrow Q(c))^*, \neg Q(d), Q(c)$;

La prima tavola è chiusa; la seconda ha come successore

$P(c)^*, (P(c)\rightarrow Q(c))^*, \neg Q(d)^*, Q(c)$; $Q(d)$

che ha come successore

$P(c)^*, (P(c)\rightarrow Q(c))^*, \neg Q(d)^*, Q(c)$; $Q(d)^*$

che ha come successore

$P(c)^*, (P(c)\rightarrow Q(c))^*, \neg Q(d)^*, Q(c)^*; Q(d)^*$, che è una tavola terminata non chiusa. L'insieme di partenza ha come modello $\mathbf{M}=\langle M, P^M, Q^M\rangle$, dove $M=\{c,d\}$, $P^M=Q^M=\{c\}$. (Seguendo pedissequamente la procedura, otterremmo un modello il cui dominio M è l'insieme dei termini chiusi del linguaggio, e in cui $P^M=Q^M=\{c\}$. Tuttavia è abbastanza facile convinversi che solo c e d sono rilevanti per la verità delle formule P(c), $P(c)\rightarrow Q(c)$, $\neg Q(d)$).

Riferimenti bibliografici

Una dimostrazione del teorema di completezza è presente in tutti i testi istituzionali di Logica Matematica. Una dimostrazione simile a quella riportata nel presente testo si trova ad esempio in G. LOLLI, "Introduzione alla Logica Formale", Il Mulino, Bologna, 1991. Dimostrazioni alternative si trovano ad esempio in C.C.CHANG-H.J.KEISLER, Model Theory, North Holland, Amsterdam, London, 1973, oppure in E.MENDELSON, "Introduzione alla Logica Matematica", Boringhieri, Torino, 1972. Tutti questi testi, come pure altri testi di Logica matematica, contengono un'esposizione dettagliata dei concetti di soddisfazione, verità, etc... Commenti critici sull'importanza del Teorema si trovano, oltre che nell'opera citata di Lolli, anche in molti altri testi. Citiamo a titolo d'esempio: E.BALLO, E.CASARI, C.CELLUCCI, M.DALLA CHIARA, G.LOLLI, C.MANGIONE, M.MUGNAI, "9 Lezioni di Logica", Franco Muzzio Ed., Padova, 1990, ed inoltre: F.BELLISSIMA, P.PAGLI, "La verità trasmessa", Biblioteca Universale Sansoni, Firenze, 1993. Per una trattazione storica, ed anche critico-filosofica, degli argomenti esposti in queste note, si veda: C.MANGIONE, S.BOZZI, "Storia della Logica", Garzanti, Milano 1993.

Esercizi

FRANCO MONTAGNA

Verità di una formula in una struttura. Conseguenza logica

ESERCIZIO 1 Sia L un linguaggio avente come simboli non logici un simbolo di operazione binaria * e un simbolo di predicato binario $\approx$; consideriamo l'interpretazione $Z=<Z, \cdot,=>$, dove Z è l'insieme degli interi relativi, $\cdot$ è il prodotto, e = è l'uguaglianza. Per quali assegnazioni σ su Z risulta $Z,\sigma \models x*x \approx y*y$?

ESERCIZIO 2 Un *isomorfismo* da una struttura M a una struttura N dello stesso tipo è una biiezione F da M a N tale che:

(i) per ogni simbolo c di costante sia $F(c^M)=c^N$;

(ii) per ogni simbolo di funzione n-aria f, e per ogni n-upla $a_1,...,a_n$ di elementi di M, sia $Ff^M(a_1,...,a_n)=f^N(Fa_1,...,Fa_n)$;

(iii) per ogni simbolo P di predicato n-ario P, e per ogni n-upla $a_1,...,a_n$ di elementi di M, sia $(a_1,...,a_n) \in P^M$ se e solo se $(Fa_1,...,Fa_n) \in P^N$.

Dimostrare che se F è un isomorfismo da una struttura M a una struttura N, per ogni formula $\alpha(x_1,...,x_n)$ e per ogni $a_1,...,a_n \in M$, è $M \circ \alpha(a_1/x_1,...,a_n/x_n)$ se e solo se $N \circ \alpha(Fa_1/x_1,...,Fa_n/x_n)$.

ESERCIZIO 3 Sia L un linguaggio avente come unico simbolo non logico un simbolo di predicato binario P, e sia **R** l'interpretazione $<R, \geq>$, dove R è l'insieme dei reali, e $\geq$ denota la relazione "essere maggiore o uguale di"; sia poi **R*** l'interpretazione $<R, \leq >$, dove $\leq$ denota la relazione "essere minore o uguale di". Dimostra: per ogni enunciato ϕ di L, è $\mathbf{R} \models \phi$ se e solo se $\mathbf{R}^* \models \phi$. (*Consiglio*: trova un isomorfismo da **R** a **R*** e usa l'esercizio precedente).

ESERCIZIO 4 Riprendi l'esercizio 3; prova ad ampliare il linguaggio e le due interpretazioni in modo da poterle distinguere al primo ordine (ad esempio, aggiungi un'operazione binaria *, da interpretarsi in entrambi i casi come prodotto, e una costante c, da interpretarsi in entrambi i casi come lo zero; dimostra che le due interpretazioni così estese non concordano sulla formula $\forall x P(x*x,c)$).

ESERCIZIO 5 Sia L come nell'esercizio 3); considera le interpretazioni <POL,EQUISC> e <POL,SIM>, dove POL è l'insieme dei poligoni, EQUISC è la relazione di equiscomponibiltà, e SIM è la relazione di similitudine. Prova ad ampliare il linguaggio e le interpretazioni in modo da poter distinguere i due concetti. Ad esempio, introduci nuovi predicati unari $L_3,...,L_n,...$ dove L_n è interpretato in entrambi i casi come "avere n lati", e prova a trovare qualche formula vera in una delle sue strutture e falsa nell'altra.

ESERCIZIO 6 Prova: $\models (A \to B) \leftrightarrow (\neg A \vee B)$; $\models (A \wedge B) \leftrightarrow \neg (\neg A \vee \neg B)$; $\models \exists x A \leftrightarrow \neg \forall x \neg A$; Concludi che nella logica classica si potrebbero usare solo i connettivi $\vee$ e $\neg$ e il quantificatore $\exists$ senza perdere di espressività. E' possibile ridurre i connettivi a uno solo? (*Suggerimento*: poni $A \otimes B = \neg A \wedge \neg B$, ed esprimi $\neg$ e $\vee$ attraverso $\otimes$; ad es puoi esprimere $\neg A$ come $A \otimes A$...).

ESERCIZIO 7 Prova ora ad esprimere i connettivi e i quantificatori attraverso un unico quantificatore binario Qx(A,B), ove x è una variabile e A, B sono formule. (*Consiglio*: poni $Qx(A,B)=\forall x(A \otimes B)$; osserva che $\neg A$ può essere espresso come Qx(A,A) ove x è una variabile che non compare in A...; esprimi ora anche $A \vee B$ e $\forall y A$ tramite Q)

ESERCIZIO 8 Prova a tradurre l'assioma logico $A \to (B \to A)$ in un linguaggio con il solo quantificatore binario Q di cui all'esercizio precedente: ti renderai conto del motivo per cui è sconsigliabile lavorare in un linguaggio simile.

Teorema di Completezza e tavole semantiche

ESERCIZIO 1 Dimostra, sfruttando il metodo delle tavole semantiche, la formula $\exists x \forall y P(x,y) \to \forall y \exists x P(x,y)$. Dimostra ora la dimostrabilità della formula usando il Teorema di Completezza (prova cioè che $\models \exists x \forall y P(x,y) \to \forall y \exists x P(x,y)$).

ESERCIZIO 2 Dimostra l'indimostrabilità nel calcolo dei predicati dell'implicazione inversa $\forall y \exists x P(x,y) \to \exists x \forall y P(x,y)$, P predicato binario. (Puoi considerare ad esempio l'interpretazione il cui dominio è l'insieme degli esseri umani, e in cui P(x,y) è interpretata come "x conosce y"). [*Metaproblema*: l'esercizio è davvero un'applicazione del Teorema di Completezza o piuttosto un'applicazione del Teorema di Validità ?]

ESERCIZIO 3 Sia $A_1,...,A_n$; $B_1,...B_m$ (si intende che le A_i sono nella colonna V e le B_j sono nella colonna F) una tavola semantica in cui tutte le A_i, B_j sono prive di quantificatori. Prova che il metodo delle tavole semantiche applicato alla tavola in oggetto termina in un numero finito di passi. (*Consiglio*: associa ad ogni tavola T un numero n(T) dato dalla somma del numero totale di connettivi e quantificatori nelle formule non asteriscate di T e del numero di

formule atomiche non asteriscate. Prova che se T' è un successore di T allora n(T)>n(T'), e quindi ad un certo momento della costruzione tutte le tavole foglia non chiuse sono...).

Concludi che il problema della dimostrabilità per formule prive di quantificatori è decidibile. Puoi fare lo stesso ragionamento per formule qualsiasi? Perché?

Linguaggi con identità

Un linguaggio con identità è un linguaggio L con un simbolo predicativo binario speciale, diciamo $\approx$. (L'idea è che $\approx$ rappresenta la relazione di uguaglianza). La semantica per i linguaggi con identità è definita come al solito, con la differenza però che $\approx$ deve necessariamente essere interpretato nella relazione di uguaglianza.

ESERCIZIO 1 Prova che le formule seguenti sono vere in tutte le interpretazioni per linguaggi con identità:

$$\forall x(x\approx x);\ \forall x\forall y(x\approx y\rightarrow y\approx x);\ \forall x\forall y\forall z[(x\approx y\wedge y\approx z)\rightarrow x\approx z];$$
$$\forall x_1...\forall x_n\forall y_1...\forall y_n[(x_1\approx y_1\wedge...\wedge x_n\approx y_n)\rightarrow f(x_1,...,x_n)\approx f(y_1,...,y_n)]$$

(f simbolo di funzione n-aria qualsiasi);

$$\forall x_1...\forall x_n\forall y_1...\forall y_n[(x_1\approx y_1\wedge...\wedge x_n\approx y_n)\rightarrow(P(x_1,...,x_n)\rightarrow P(y_1,...,y_n))]$$

(P simbolo di predicato n-ario qualsiasi). Le formule di questo tipo vengono dette "assiomi di identità per il linguaggio con identità L". L'insieme degli assiomi di identità per un linguaggio L verrà denotato con AIL.

ESERCIZIO 2 Nelle logiche con identità il concetto di conseguenza logica è così modificato: A è conseguenza logica di un insieme Σ di enunciati nella logica con identità (in simboli: $\Sigma_{id}\models A$) se e solo se A è vera in tutte le interpretazioni (con identità) in cui sono vere le formule di Σ. Si ha il seguente:

TEOREMA DI COMPLETEZZA PER IL CALCOLO DEI PREDICATI CON IDENTITÀ.
Sia AIL l'insieme degli assiomi di identità per il linguaggio L, Σ un insieme di enunciati di L, e A una formula di L. Si ha: $\Sigma_{id}\models A$ se e solo se $\Sigma\cup AIL\vdash A$.

Prova a dimostrare il Teorema enunciato sopra. *Suggerimento* per il verso più difficile, cioè quello da sinistra a destra: supponi che $\Sigma\cup AIL$ non dimostri A; per il Teorema di Completezza per la logica senza identità, esiste una struttura **M** tale che $\mathbf{M}\models\Sigma\cup AIL$, $\mathbf{M}\models\neg A$; poni, per ogni a,b nel dominio M di **M**, aEb se e solo se $\mathbf{M}\models a\approx b$; E è una relazione di equivalenza; sia [a] la classe di equivalenza di a rispetto alla relazione E, e sia $[M]=\{[a]:a\in M\}$ l'insieme delle classi di equivalenza di elementi di M; forma una struttura

[M] di dominio [M] ponendo, per ogni simbolo c di costante, $c^{[M]}=[c^M]$, per ogni simbolo f di funzione n-aria e per ogni n-upla $([a_1],...,[a_n])$ di elementi di [M], $f^{[M]}([a_1],...,[a_n])=[f^M(a_1,...,a_n)]$, e , per ogni simbolo P di predicato n-ario e per ogni n-upla $([a_1],...,[a_n])$ di elementi di [M], $([a_1],...,[a_n])\in P^{[M]}$ se e solo se $(a_1,...,a_n)\in P^M$. Prova che le definizioni sono lecite, cioè dipendono dalle classi di equivalenza e non dai rappresentanti (usa il fatto che $\mathbf{M}\models AIL$); infine, dimostra per induzione sulla complessità che, per ogni formula $\alpha(x_1,...,x_n)$, e per ogni $a_1,...,a_n\in M$ è $\mathbf{M}\models\alpha(a_1/x_1,...,a_n/x_n)$ se e solo se $[\mathbf{M}]\models\alpha([a_1]/x_1,...,[a_n]/x_n)$. Nota che in particolare $[\mathbf{M}]\models\Sigma\cup AIL$, $[\mathbf{M}]\not\models A$, e che $[\mathbf{M}]\models[a]\approx[b]$ se e solo se $\mathbf{M}\models a\approx b$ se e solo se aEb se e solo se [a]=[b]; quindi, $\approx^{[M]}$ è l'uguaglianza. Concludi che **[M]** è un modello con identità di $\Sigma\cup AIL$ in cui A non è vero, e che pertanto $\Sigma_{id}\not\models A$.

Nel seguito considereremo solo logiche con identità.

ESERCIZIO 3 Sia L un linguaggio con identità avente un simbolo di operazione binaria, *, un simbolo di operazione unaria, $^{-1}$, e un simbolo di costante 1; sia T la teoria che ha come assiomi:

$$\forall x(1*x\approx x);\qquad \forall x(x^{-1}*x\approx 1);\qquad \forall x\forall y\forall z[(x*y)*z\approx x*(y*z)].$$

Usando il teorema di Completezza per la logica con identità, dimostra che, per ogni formula α di L si ha: α è vera in tutti i gruppi se e solo se $T\cup AIL\vdash\alpha$.

ESERCIZIO 4 Con il metodo delle tavole semantiche abbiamo visto che ogni insieme consistente di enunciati ha un modello che ha come dominio l'insieme dei termini chiusi dell'interpretazione, ed è quindi numerabile (in particolare è infinito). La stessa proprietà non vale per la logica con identità: la formula $\forall x\forall y[(x\approx y)$ ha solo modelli con identità di un solo elemento; trova una formula del linguaggio L con $\approx$ come unico simbolo non logico che abbia come modelli con identità tutte e sole le strutture per L di due elementi.

ESERCIZIO 5 In matematica si ha spesso a che fare con funzioni parziali (ad esempio la radice quadrata, il logaritmo); al contrario, in logica, i simboli di funzione vengono interpretati in funzioni totali; prova a rimediare a questo inconveniente aggiungendo al linguaggio un simbolo di costante «, da interpretarsi come "non definito"; naturalmente, avremo fra gli assiomi f(«)=« per ogni simbolo di funzione unaria, e più in generale uno schema di assiomi che esprime che, se almeno uno degli argomenti di una funzione n-aria è non definito, anche il valore della funzione è non definito. Trova alcune formule del linguaggio con identità L=<E, L, M, c,«> (E,L simboli di funzione unaria, M simbolo di predicato binario, c,« simboli di costante) vere nella struttura $<R^{indef}, EXP^{indef}, LOG^{indef}, <^{indef}, 0, indef>$, dove R^{indef} è l'insieme dei reali a cui viene aggiunto un

nuovo elemento "indef", EXP^{indef} e LOG^{indef} sono rispettivamente la funzione esponenziale e la funzione logaritmica estese ponendo $LOG^{indef}(x)$=indef se o $x\leq 0$ oppure x=indef, EXP^{indef}(indef)=indef, e inoltre $x <^{indef} y$ se e solo se $x,y\in R$ e $x<y$. In particolare, nella struttura in questione è vera la formula $\forall x[0 <^{indef} EXP^{indef}(x)]$? E la formula $\forall x LOG^{indef} EXP^{indef}(x)\approx x$?

ESERCIZIO 6 Considera la struttura <R,LOG,<,0>, dove R è l'insieme dei reali, LOG è la funzione logaritmica, < e 0 hanno il solito significato; sostituisci LOG con una relazione binaria GRAFLOG, definita da GRAFLOG(a,b) se e solo se LOG(a)=b, ottenendo una struttura **M**=<R,GRAFLOG,<,0>. Pensa **M** come una struttura per il linguaggio L con identità avente due predicati binari P e Q, e una costante c. Trova alcune formule notevoli di L vere nella struttura **M**.

Cenno sulla logica del secondo ordine

Non sempre un linguaggio matematico è assimilabile a un linguaggio del primo ordine; ad esempio la frase: "ogni funzione reale di variabile reale derivabile è continua" contiene una quantificazione universale su funzioni, mentre al primo ordine possiamo quantificare solo su individui. A questo si rimedia in parte usando il linguaggio al primo ordine della teoria degli insiemi, in cui si può esprimere al primo ordine "essere una funzione", "essere una relazione", "essere un numero reale", "essere una funzione reale di variabile reale"; le funzioni e le relazioni vengono dunque pensate come individui e non come entità di tipo superiore; tuttavia, quando facciamo la semantica per le varie teorie degli insiemi, quantificazioni del tipo "ogni funzione" vengono interpretate sugli oggetti del dominio di interpretazione che sono funzioni (o meglio, che lo sono nel senso dell'interpretazione), e non in tutte le possibili funzioni. Ne viene che la "vera" semantica al second'ordine non può essere assimilata a una semantica per un linguaggio del primo ordine. Purtroppo, come vedremo negli ultimi due esercizi, per il secondo ordine non vale un teorema di completezza analogo a quello del primo ordine.

ESERCIZIO 1 Considera un linguaggio con il simbolo di uguaglianza, $\approx$, due simboli di operazione binaria, + e $\cdot$, e due simboli di costante, 0 e 1. Arricchiamo il linguaggio aggiungendo *simboli di variabile del secondo ordine,* $X_1,...,X_n,...$da interpretarsi sui sottoinsiemi del dominio di interpretazione, e il simbolo di appartenenza, $\in$. Consideriamo la teoria T avente come assiomi:

$\neg(x+1\approx 0)$; $(x+1\approx y+1\rightarrow x\approx y)$; $x+0\approx x$; $x+(y+1)\approx(x+y)+1$; $x\cdot 0\approx 0$; $x\cdot(y+1)\approx(x\cdot y)+x$

(i primi due dicono che l'operazione "aggiungi 1" non dà mai zero ed è iniettiva, le altre danno la definizione induttiva della somma e del prodotto) più l'*assioma di induzione*:

$\{(0\in X)\wedge\forall x[(x\in X)\rightarrow(x+1\in X)]\}\rightarrow\forall x(x\in X)$. Siano M e N due modelli di questa teoria, ove, come già detto, le variabili al secondo ordine possono essere interpretate in arbitrari sottoinsiemi del dominio. Dimostrare che M e N sono isomorfi. (*Suggerimento*: usando l'assioma di induzione, si mostri che $M=\{0^M,1^M,1^M+1^M,...,1^M+...+1^M,...\}$ e $N=\{0^N,1^N,1^N+1^N,...,1^N+...+1^N,...\}$. Si consideri la funzione F che manda 0^M in 0^N e $1^M+...+1^M$, (k volte) in $1^N+...+1^N$, (k volte). Si dimostri che F è un isomorfismo da M a N.

ESERCIZIO 2 Essendo tutti i modelli isomorfi fra loro, per un esercizio precedente si ha che dato un enunciato A, o A è vero in tutti i modelli o è falso in tutti i modelli; se valesse il Teorema di Completezza per il secondo ordine, avremmo che, dato un enunciato A arbitrario, o A o la sua negazione sarebbe dimostrabile in T, e T sarebbe completa. Si dimostri che ciò è impossibile. (*Suggerimento*: se T fosse completa, l'insieme degli enunciati al primo ordine dimostrabili in T sarebbe un'estensione completa e ricorsivamente assiomatizzabile di...).

Compattezza, Categoricità, Paradosso di Skolem[1]

RUGGERO FERRO
Dipartimento di Matematica
Università di Lecce
via per Arnesano
73100 Lecce

0. Introduzione

Il teorema di completezza ha mostrato che si può realizzare un controllo completo della validità e della consequenzialità di una formula, nel senso che se una formula è valida è anche dimostrabile (la sua negazione ha confutazione), e se è conseguenza logica di altre allora si può dedurre dalle altre (le altre e la sua negazione hanno confutazione). Proprio il successo nel trovare un controllo sintattico[2] di validità e consequenzialità, porta a limiti intrinseci delle capacità descrittive di un linguaggio formale del primo ordine. L'obiettivo di questo intervento è precisare alcuni di questi limiti cogliendo anche come hanno origine nel linguaggio formale introdotto.

1. Rafforzamenti del teorema di completezza

Per mostrare il teorema di completezza, si era dimostrato che ogni insieme finito di formule in un linguaggio numerabile del primo ordine è non soddisfacibile se e solo se da quell'insieme di formule si può costruire un albero chiuso. Ci servirà un rafforzamento di

[1] Lavoro parzialmente finanziato con i fondi 40% e 60% del M.U.R.S.T.

[2] La definizione tarskiana di verità di una formula in una interpretazione, e la conseguente nozione di validità, pur essendo del tutto esplicite, non sono di fatto applicabili direttamente per determinare la validità di una formula poiché richiedono di controllare in ogni struttura, adatta al linguaggio usato, la verità della formula data, e le possibili strutture sono di una tale infinità che la loro collezione non può essere un insieme. Inoltre, se la formula contiene dei quantificatori e l'universo della struttura è infinito, anche la determinazione della verità della formula può richiedere infiniti controlli. Tuttavia per alcune formule (ad esempio $\alpha \rightarrow \alpha$) la stessa scrittura della formula ci permette di concludere che è valida (a causa del significato di $\rightarrow$). Allora si pone il problema di cercare un analogo controllo sulla scrittura (sintattico) per determinarne la validità di una qualsiasi formula. Un controllo per essere tale deve essere effettivo, cioè si deve poter determinare in un numero finito (pur senza limiti alla sua grandezza) di passi se l'esito del controllo è positivo. Il teorema di completezza, asserendo che ogni formula valida è dimostrabile, ci permette di effettuare un tale controllo, perché si vede in modo effettivo se una successione di formule è una dimostrazione.

questo risultato che si estende al caso in cui il linguaggio abbia = come simbolo logico (cioè = non va interpretato come una relazione binaria fissata dalla struttura, ma come la relazione binaria costituita da tutte le coppie ordinate di elementi dell'universo il cui primo elemento è lo stesso del secondo) e l'insieme di formule da cui partire possa essere numerabile. In questi casi la costruzione degli alberi a partire da un insieme di formule rimane sostanzialmente la stessa. Anche in questo caso le dimostrazione che se un insieme di formule non ammette albero chiuso allora è soddisfacibile si fa con i metodi già esposti nella lezione del professor Montagna costruendo un modello il cui universo sia costituito da termini, meglio da classi di equivalenza di termini, ora. Pertanto un insieme numerabile soddisfacibile di formule in un linguaggio numerabile avrà un modello al più numerabile.

La limitazione che l'universo del modello sia solo al più numerabile, potrebbe apparire troppo forte. Perché se alcune formule hanno modello devono avere un modello al più numerabile e non soltanto modelli più che numerabili? Di fatto si riesce a dimostrare che è così, e ciò è già un certo tipo di limite del linguaggio nel descrivere sue strutture.

Anche nel caso in cui il linguaggio non sia numerabile si riescono a costruire, con tecniche un po' diverse, alberi a partire da un insieme più che numerabile di formule Φ con rami chiusi se e solo se l'insieme Φ è non soddisfacibile. Ma non insistiamo in questa direzione perché non utilizzeremo tali tecniche: nel caso di linguaggi più che numerabili, penseremo che il controllo sintattico della soddisfacibilità sia stato effettuato non con la tecnica degli alberi, ma con quella delle deduzioni alla Hilbert, per la quale vale ancora il teorema di completezza. La dimostrazione del teorema di completezza per le deduzioni alla Hilbert supera i limiti di questa presentazione, ma vogliamo ricordare che anche per questo risultato si esibisce un modello costruito a partire da classi di equivalenza di termini, sicché la sua cardinalità sarà minore od uguale a quella del linguaggio.

2. I teoremi di compattezza

La nozione di deduzione di una formula da altre (anche infinite altre) richiede l'esistenza di una successione finita di formule, costruita secondo ben precisati criteri sintattici (cioè sul modo di scrittura delle formule), che porti alla formula voluta.

E' conseguenza banale di questa nozione che se una formula si deduce da un insieme di formule allora si deduce anche da un sottinsieme finito dell'insieme di formule dato (il sottoinsieme delle formule che occorrono come elementi della successione che è la deduzione). Il viceversa è ancor più banale[3]. Questo risultato sintattico va sotto il nome di *teorema di compattezza sintattico*.

[3] Ciò corrisponde ad affermare che aumentando le premesse si mantengono i risultati già acquisiti. Si noti che per quanto questa affermazione sia aspettata e vera nei sistemi logici introdotti, ci sono sistemi logici particolari in cui non vale; questi vengono detti non monotoni.

Poiché si dimostra che una formula φ è deducibile da un insieme di formule Φ se e solo se l'insieme di formule $\Phi \cup \{\neg\varphi\}$ è contraddittorio, il teorema di compattezza sintattico può essere riformulato così: un insieme di formule è contraddittorio se e solo se c'è un suo sottinsieme finito che è contraddittorio; o, equivalentemente, così: un insieme di formule è non contraddittorio se e solo se ogni suo sottinsieme finito è non contraddittorio.

Nello sviluppo scelto per questo corso, il controllo sintattico della non soddisfacibilità di un insieme di formule non è stato fatto mediante la deduzione alla Hilbert, ma attraverso alberi di confutazione. Anche gli alberi di confutazione ci portano ad una formulazione del teorema di compattezza sintattico proprio per il fatto che gli alberi chiusi sono finiti (cioè hanno un numero finito di nodi), e così hanno un numero finito di rami ed ogni ramo ha lunghezza finita.

TEOREMA DI COMPATTEZZA SINTATTICO PER GLI ALBERI DI CONFUTAZIONE CON LINGUAGGIO NUMERABILE. *Un insieme di formule in un linguaggio numerabile ha albero di confutazione se e solo se un suo sottinsieme finito ha albero di confutazione.*

Dimostrazione. Una direzione è banale: se un sottinsieme (eventualmente finito) di un insieme di formule dà origine ad un albero con tutti i rami chiusi, allora lo stesso albero chiuso può avere origine dall'intero insieme (si noti che qui stiamo usando la nozione di albero adatta a questa situazione con l'introduzione di un numero numerabile di nuove costanti per l'analisi delle formule quantificate esistenzialmente). Per dimostrare l'altra direzione, si osservi che un albero chiuso ha un numero finito di nodi, ciascuno, esclusa la radice, costituito da un numero finito di formule la cui presenza è giustificata dal fatto che sono state ottenute applicando una delle regole di costruzione degli alberi a una formula in un nodo precedente, sicché solo un numero finito di formule della radice sono state utilizzate per la costruzione dell'albero. Si consideri l'insieme delle formule della radice utilizzate per la costruzione dell'albero. Come detto tale insieme è finito e a partire da esso si può costruire lo stesso albero chiuso ottenuto a partire dall'insieme infinito dato di formule. Si è così ottenuto un sottinsieme finito con albero di confutazione, e la dimostrazione è completata.

Il teorema di completezza visto in questo corso, e qui esteso per linguaggi numerabili, trasforma il teorema di compattezza sintattico per gli alberi di confutazione[4] nella seguente

[4] Anche il teorema di compattezza sintattica visto all'inizio di questa sezione porta al teorema di compattezza semantica disponendo di un teorema di completezza per le deduzioni alla Hilbert. Questo risultato, pur essendo usuale nella letteratura, non è stato dimostrato in questo corso, per una scelta didattica motivata anche dalla limitatezza del tempo disponibile. La non disponibilità del teorema di completezza per le deduzioni alla Hilbert giustifica l'estensione precedentemente tratteggiata del teorema di completezza relativo agli alberi di confutazione per insiemi finiti di formule a quello per insiemi numerabili di formule. Lo stesso fatto giustifica anche la presentazione del teorema di compattezza sintattico per alberi di confutazione, quando il linguaggio è numerabile.

affermazione: *un insieme di formule è soddisfacibile se e solo se ogni suo sottinsieme finito è soddisfacibile*. Questo risultato va sotto il nome di *teorema di compattezza semantico*[5].Il teorema di compattezza semantico vale anche per linguaggi non numerabili, e più avanti lo useremo anche in questa forma, anche se qui non lo abbiamo dimostrato. Di fatto la sua dimostrazione è abbastanza agevole a partire dal teorema di compattezza sintattico per calcoli alla Hilbert e dal teorema di completezza per questi.

Se è ovvio che, se un insieme di formule è soddisfacibile, allora lo è anche ogni suo sottinsieme, in particolare i sottinsiemi finiti, non è assolutamente evidente che, se tutti i sottinsiemi finiti di un insieme di formule sono soddisfacibili, allora anche l'insieme è soddisfacibile. In effetti dire che ogni sottinsieme finito è soddisfacibile equivale a dire che ha almeno un modello (modello che in generale sarà diverso da sottinsieme a sottinsieme), ma ciò non ci suggerisce che possa esserci un unico modello dell'intero insieme di formule, magari ottenuto in qualche modo dai modelli dei sottinsiemi finiti.

Di fatto c'è un modo di costruire un modello per l'intero insieme di formule a partire dai modelli dei sottinsiemi finiti, mediante la tecnica degli ultraprodotti (che qui non affrontiamo), ma questa procedura assolutamente non è banale. Il contrasto tra la banalità della dimostrazione del teorema di compattezza sintattico e la difficoltà dell'ultimo risultato enunciato significa che il teorema di completezza copre le difficoltà per ottenere il teorema di compattezza semantico.

Anche se è stato facilmente dimostrato a partire dal teorema di completezza, il teorema di compattezza semantico è molto importante per il modo con cui può essere utilizzato e per le conseguenze che produce.

3. Possibili modi di usare il teorema di compattezza

Sia T una teoria consistente. Tra i suoi modelli se ne cercano alcuni, se ci sono, che abbiano una particolare proprietà, chiamiamola P. Si cerca di caratterizzare la proprietà mediante enunciati, eventualmente un insieme infinito di enunciati. Cioè si cerca un insieme S, anche infinito, di enunciati che siano veri in una struttura se e solo se quella struttura gode della proprietà P. Se S è un insieme di enunciati che caratterizza la proprietà P dei suoi modelli e T∪S è un insieme soddisfacibile di enunciati, ci sarà un modello di T che gode della proprietà P.

Il teorema di compattezza permette di verificare se l'insieme T∪S è soddisfacibile controllando la soddisfacibilità dei suoi sottinsiemi finiti. Spesso la proprietà P è del tipo: ci sono degli individui che hanno un comportamento ben descritto da certi enunciati. In questo caso la tecnica sopra descritta si precisa ulteriormente. Si amplia[6] il linguaggio con nuovi simboli individuali ciascuno indicante un nuovo individuo, e, per ottenere l'insieme S, si considerano, nel linguaggio ampliato, gli enunciati che descrivono il comportamento dei nuovi individui. Poi, per cercare di far vedere la soddisfacibilità degli enunciati di un qualsiasi sottinsieme finito X di T∪S, spesso si ricorre ad un modello privilegiato **A** di T nel linguaggio iniziale e lo si espande[7] ad un modello **A'** nel linguaggio ampliato interpretando le nuove costanti in modo che siano soddisfatti gli enunciati di S che sono in X.

La tecnica appena esposta può essere utilizzata anche a partire da una struttura **A**, per mostrare che si immerge elementarmente[8] in un'altra con precisate caratteristiche. Si espande la struttura data ad una struttura **A'** in un linguaggio con un nome per ogni elemento dell'universo, nome che si interpreta nell'elemento a cui è associato. Ciò perché **A'** si immerge elementarmente in ogni modello **B'** della teoria di **A'** [9] [10], ed anche **A** si immerge elementarmente nella riduzione[11] **B** di ogni tale **B'** al linguaggio iniziale. Sicché si può applicare la metodologia prima vista alla teoria di **A'**.

4. Immersioni elementari e sottostrutture elementari

Una volta ottenuta una struttura con particolari proprietà in cui la struttura data si immerge elementarmente, non è difficile trovare anche una struttura con quelle proprietà che estenda elementarmente[12] la struttura data (detto altrimenti di cui la struttura data sia una sottostruttura elementare). Cioè vale il seguente

---

[5] Il nome di compattezza dato a questo teorema non è casuale, ma segue dal fatto che il teorema di compattezza implica la compattezza (in senso topologico) di un opportuno spazio topologico costruito a partire da strutture, ma anche questo argomento è fuori dall'ambito di quanto vogliamo perseguire in questa sede.

[6] Per *ampliamento* di un linguaggio si intende un linguaggio che contenga il precedente.

[7] Un'*espansione* di una struttura **A** ad un linguaggio L' che amplia il linguaggio L adatto alla struttura data è una struttura **A'** adatta al linguaggio L' che 1) ha lo stesso universo di **A**; 2) i simboli di L sono interpretati in **A'** come in **A**; 3) sono interpretati anche i simboli di L'-L.

[8] Per *immersione elementare* di una struttura **A** in un'altra **B** si intende una funzione ϕ dall'universo di una nell'universo dell'altra tale che per ogni formula φ del linguaggio adatto ad **A** con variabili tra le prime n+1 e per ogni (n+1)-upla $a_0,...,a_n$ di elementi dell'universo di **A** si ha che $\mathbf{A}\models\varphi[a_1,...,a_n]$ se e solo se $\mathbf{B}\models\varphi[\phi(a_1),...,\phi(a_n)]$. Una tale funzione è anche totale e iniettiva.

[9] La *teoria di una struttura* **A** è l'insieme degli enunciati veri in quella struttura.

[10] Se ad un elemento della prima struttura associamo l'elemento della seconda struttura che ha lo stesso nome (possibile perché tutti gli elementi della prima struttura hanno nome), la funzione così ottenuta è una immersione elementare, sostanzialmente perché il linguaggio è sufficientemente ricco per determinare il comportamento di ogni relazione e di ogni funzione delle strutture. Infatti, in questo linguaggio, per ogni (n+1)-upla di elementi dell'universo e per ogni relazione R della struttura e per ogni funzione F della struttura le formule $P(a_0, \ldots, a_n)$ e $\underline{F}(a_1, \ldots, a_n)=a_0$ descrivono il comportamento della struttura, dove $a_0, \ldots, a_n$ sono i nomi degli elementi dell'(n+1)-upla, P è il nome di R e $\underline{F}$ è il nome di F.

[11] La *riduzione* di una struttura è l'operazione inversa dell'espansione: data una struttura **A** adatta ad un linguaggio L, la sua riduzione al linguaggio L' contenuto in L è la struttura **A'** ottenuta da **A** trascurando di interpretare i simboli in L - L'. Si noti che **A** e **A'** hanno lo stesso universo ed interpretano in modo uguale i simboli in L'.

[12] Si dice che una struttura **B** *estende* una struttura **A**, e che **A** è una *sotto-struttura* di una struttura **B** (e lo si indica così: $\mathbf{B}\supseteq\mathbf{A}$) se 1) **B** e **A** sono strutture dello stesso tipo (cioè adatte allo stesso linguaggio), 2) l'universo B di **B** contiene l'universo A di **A**, $B\supseteq A$, 3) A è chiuso rispetto alle funzioni della struttura **B** (cioè per ogni numero naturale n e per ogni funzione n-aria f di **B** e per ogni n-upla $a_1, \ldots, a_n$ di elementi di A risulta che $f(a_1,..., a_n)\in A$), 4) le relazioni e le funzioni di **A** sono le relazione e le funzioni di **B** ristrette ad A.

TEOREMA. *Se una struttura* **A** *si immerge elementarmente in una struttura* **B**, *allora esiste una estensione elementare* **A'** *di* **A** *che è isomorfa a* **B**.

Dimostrazione. Se **A** si immerge elementarmente in **B** mediante una funzione ϕ, allora ϕ sarà un isomorfismo da **A** nell'immagine ϕ(**A**) di **A** mediante ϕ, e ϕ(**A**) sarà una sottostruttura elementare di **B**. Per convincersi che ϕ è un isomorfismo da **A** in ϕ(**A**), basta notare che ϕ è una biiettività dall'universo di **A** , chiamiamolo A, sull'universo di ϕ(**A**), chiamiamolo ϕ(A), e che la struttura è preservata poiché ciò equivale al conservarsi della verità delle formule atomiche, e ciò segue dall'elementarità dell'immersione. ϕ(A) è un sottinsieme dell'universo di **B**, chiamiamolo B, che è chiuso rispetto alle funzioni n-arie F di **B** poiché per ogni n-upla $b_1,...,b_n$ di ϕ(A), $\phi^{-1}(b_1),...,\phi^{-1}(b_n)$ è una n-upla di elementi di A, e, se F' è la funzione di **A** che ha lo stesso nome f di F, allora ci sarà un elemento a di A tale che

$$\mathbf{A} \models f(x_1,...,x_n)=y[\phi^{-1}(b_1), ...,\phi^{-1}(b_n),a] \quad (\text{o } F'(\phi^{-1}(b_1),...,\phi^{-1}(b_n))=a).$$

Ne segue che $\mathbf{B} \models f(x_1,...,x_n)=y[b_1,...,b_n,\phi(a)]$, e che $F(b_1,...,b_n)$ è l'elemento $\phi(a)$ di ϕ(A). Infine ϕ(**A**) è sottostruttura elementare di **B** poiché per ogni formula φ le cui variabili siano entro le prime n e per ogni (n+1)-upla $b_1,...,b_n$ di elementi di ϕ(A), poiché ϕ è un isomorfismo da **A** su ϕ(**A**) si ha ϕ(**A**) $\models \varphi[b_1,...,b_n]$ se e solo se $\mathbf{A}\models\varphi[\phi^{-1}(b_1),..., \phi^{-1}(b_n)]$; ma l'ultima affermazione equivale a $\mathbf{B} \models \varphi[b_1,...,b_n]$ dal momento che ϕ è una immersione elementare, e la transitività dell'equivalenza permette di concludere.

Si consideri ora un insieme C disgiunto da A e di cardinalità uguale a quella di B-ϕ(A). Sia γ una biettività tra C e B-ϕ(A), esistente per l'equicardinalità dei due insiemi. Poiché C ed A sono disgiunti, la funzione $\delta=\phi\cup\gamma$ è una biiettività da A$\cup$C su B.

Infine si atteggi l'insieme A$\cup$C a struttura attribuendo ai suoi elementi il comportamento indotto da **B** attraverso la funzione δ^{-1}. Cioè si consideri la struttura **A'** il cui universo sia A$\cup$C, le cui relazioni n-arie R siano in corrispondenza delle relazioni n-arie R' di **B** di ugual nome in modo tale che, per ogni n-upla $a_1,...,a_n$ di A$\cup$C, $<a_1,...,a_n> \in R$ se e solo se $<\delta(a_1),...,\delta(a_n)>\in R'$, e le cui funzioni n-arie F siano in corrispondenza delle funzioni n-arie F' di **B** di ugual nome in modo tale che, per ogni n-upla $a_1,...,a_n$ di A$\cup$C, si abbia $\delta(F(a_1,...,a_n))=F'(\delta(a_1),...,\delta(a_n))$.

Si osservi che le restrizioni ad A delle relazioni R e delle funzioni F appena definite su **A'** sono le relazioni R_o e le funzioni F_o di **A** di ugual nome, proprio perché δ ristretta ad A è ϕ e pertanto, per ogni n-upla $a_1,...,a_n$ di A, $<a_1,...,a_n>\in R_o$ se e solo se $<\phi(a_1),...,\phi(a_n)>\in R'$ se e solo se $<\delta(a_1),...,\delta(a_n)>\in R'$ se e solo se $<a_1,...,a_n>\in R$, ed anche $\phi(F_o(a_1,...,a_n))$ $= F'(\phi(a_1),...,\phi(a_n)) = F'(\delta(a_1),...,\delta(a_n)) = \delta(F(a_1,...,a_n))$ e, poiché δ è biiettiva ed estende ϕ, dovrà essere $F_o(a_1,...,a_n) = F(a_1,...,a_n)$. Così si nota anche che A è chiuso rispetto alle funzioni di **A'**.

Si vede subito che δ è un isomorfismo da **A'** su **B** che estende l'isomorfismo ϕ da **A** in ϕ(**A**): infatti è una biiettività che preserva la struttura, proprio per come sono state definite le relazioni e le funzioni in **A'**. Si vede anche che **A** è sottostruttura elementare di **A'**. appunto perché **A** si immerge elementarmente in **B**. Che sia sottostruttura segue dall'osservazione del capoverso precedente. Per dimostrare che è elementare, si osservi che, per ogni formula φ le cui variabili siano entro l'n-esima e per ogni (n+1)-upla $a_o,...,a_n$ di A, $\mathbf{A} \models \varphi[a_o,...,a_n]$ se e solo se $\mathbf{B} \models \varphi[\phi(a_o),...,\phi(a_n)]$, poiché ϕ è una immersione elementare, ma ciò equivale a $\mathbf{B} \models \varphi[\delta(a_o),...,\delta(a_n)]$ che equivale a $\mathbf{A'} \models \varphi[a_o,...,a_n]$ poiché δ è un isomorfismo: l'equivalenza, ottenuta per transitività, tra $\mathbf{A}\models\varphi[a_o,...,a_n]$ e $\mathbf{A'}\models\varphi[a_o,...,a_n]$ è quanto richiede la definizione perché una estensione sia elementare. Così la dimostrazione è completata.

Nella stessa dimostrazione è inserita anche la prova che *se una struttura* **A** *si immerge in una struttura* **B**, *allora esiste una estensione* **A'** *di* **A** *che è isomorfa a* **B**.

Questo è un risultato che serve ad esempio quando si costruiscono i reali a partire dai razionali e si parla di reali razionali da identificarsi (che cosa vuol dire?) con i razionali: si potrebbe introdurre una struttura, ad esempio quella delle sezioni di Dedekind sui razionali, in cui i razionali si immergono, e poi, in base al risultato precedente, affermare che c'è un'estensione dei razionali isomorfa alla struttura introdotta. Applicheremo i risultati ottenuti in situazioni di un certo interesse.

5. Caratterizzabilità sintattica delle strutture finite

Ci possiamo domandare se in un linguaggio del primo ordine si possano caratterizzare le strutture finite[13] mediante un enunciato, cioè se esista un enunciato vero esattamente nella strutture finite, detto altrimenti, se esista un enunciato φ tale che $\{\mathbf{A}: \mathbf{A}\models\varphi\} = \{\mathbf{A}: \mathbf{A}$ è una qualsiasi struttura finita$\}$.

Un modo per dire, usando il linguaggio, che una struttura è infinita è quello di aggiungere al linguaggio infiniti simboli di costanti individuali (ad esempio c_i con i numero naturale) e richiedere che siano veri in quella struttura gli enunciati $c_i\neq c_j$, con i e j numeri naturali diversi tra loro, che asseriscono che quei nomi indicano individui diversi.

Così, se un enunciato φ è vero in una struttura se e solo se la struttura è finita, e se esso viene aggiunto all'insieme di enunciati appena visti, l'insieme Σ di enunciati risultante,

---

Si dice che **B** *estende elementarmente* la struttura **A**, e che **A** è *sottostruttura elementare* della struttura **B** (e si indica così: **B**>**A**) se $\mathbf{B}\supseteq\mathbf{A}$ e se per ogni formula φ le cui variabili sono tra le prime n e per ogni (n+1)-upla a_o, ... , a_n di elementi di A si ha che $\mathbf{B} \models \varphi[a_o, ... , a_n]$ se e solo se $\mathbf{A} \models \varphi[a_o, ... , a_n]$.

[13] Si dice *finita* una struttura il cui universo sia finito. Più in generale, si dice *cardinalità di una struttura* la cardinalità dell'universo di quella struttura.

$\Sigma = \{\varphi\} \cup \{c_i \neq c_j: i \neq j, i \text{ e } j \text{ numeri naturali}\}$, dovrà essere non soddisfacibile, perché ogni suo modello dovrebbe essere simultaneamente finito e infinito.

Invece, sempre nell'ipotesi che ci sia quell'enunciato φ, Σ è soddisfacibile. Infatti faremo vedere che ogni sottinsieme finito Σ_0 di Σ è soddisfacibile, pervenendo così al risultato voluto proprio in virtù del teorema di compattezza.

Sia, dunque, Σ_0 un qualsiasi sottinsieme finito di Σ. Essendo finito il numero degli enunciati occorrenti in Σ_0, ci sarà un indice massimo i_0 dei simboli di costante del tipo c_i occorrenti in enunciati del tipo $c_i \neq c_j$ appartenenti ad Σ_0, se ce ne sono, altrimenti i_0 sia 0. Sia $\mathbf{A}_0$ una struttura, adatta al linguaggio di φ, il cui universo abbia almeno i_0 elementi, che bene ordiniamo. Espandiamo $\mathbf{A}_0$ ad una struttura $\mathbf{A}_0'$, adatta al linguaggio di Σ_0, interpretando ciascun simbolo di costante c_i nell'i-esimo elemento dell'universo nell'ordinamento sopra scelto. Questa interpretazione è stata scelta proprio perché così l'interpretazione di c_i è diversa dall'interpretazione di c_j ogniqualvolta i è diverso da j. Così è ovvio che $\mathbf{A}_0' \models c_i \neq c_j$ per ogni coppia di numeri naturali diversi i e j minori od uguali a i_0. Poiché gli enunciati veri in una struttura continuano ad essere veri in una sua qualunque espansione[14], anche φ sarà vero in $\mathbf{A}_0'$. Per quanto abbiamo visto possiamo concludere che ogni enunciato di Σ_0 è vero in $\mathbf{A}_0'$, fatto che possiamo indicare così: $\mathbf{A}_0' \models \Sigma_0$.

Dunque, come già anticipato, in virtù della compattezza, si può affermare che Σ è soddisfacibile. Ma abbiamo già notato che ciò è impossibile, e pertanto il punto di partenza da cui abbiamo dedotto ciò deve essere falso. Il punto di partenza era l'esistenza dell'enunciato φ vero esattamente nelle strutture finite, enunciato che, perciò, non può esistere.

6. Teoremi di Löwenheim Skolem: parte prima

Una applicazione del teorema di compattezza del tutto analoga alla precedente porta ad affermare l'esistenza di modelli arbitrariamente grandi di teorie con modelli infiniti o di teorie con modelli finiti arbitrariamente grandi. Più precisamente dimostreremo il seguente

TEOREMA DI LÖWENHEIM SKOLEM ASCENDENTE PER TEORIE. *Sia* T *una teoria tale che, fissato un qualsiasi numero naturale* n, *ci siano modelli di* T *di cardinalità maggiore di* n. *Allora, scelta una qualsiasi cardinalità* κ, T *ha modelli di cardinalità maggiore di* κ.

Si noti che l'ipotesi è certamente soddisfatta se T ha modelli di cardinalità infinita.

Dimostrazione. Sia dunque T una tale teoria e κ un arbitrario numero cardinale. Si ampli il linguaggio della teoria T con κ nuovi simboli c_i per costanti individuali, e si consideri il seguente insieme di enunciati $\Sigma = \{c_i \neq c_j: \text{ i e j ordinali minori di } \kappa \text{ e diversi tra loro}\}$. E' evidente che ogni modello di Σ dovrà avere almeno κ elementi nel suo universo. Facciamo

[14] L'espansione non cambia né l'universo della struttura né l'interpretazione dei simboli che erano già interpretati prima dell'espansione.

vedere che $T \cup \Sigma$ ha modello. Per il teorema di compattezza, basta far vedere che ogni sottinsieme finito S_0 di $T \cup \Sigma$ ha modello. In effetti in S_0 occorreranno solo un numero finito di nuovi simboli per costante c_i, e sia n il massimo degli indici di questi simboli, se ce ne sono, 0 altrimenti. Sia **A** un modello di T, nel linguaggio di T, con almeno n elementi nell'universo: c'è un tale modello per le ipotesi fatte. Espandiamo **A** al linguaggio di S_0 interpretando ciascuno dei nuovi simboli per costante c_i, con $i \leq n$, in elementi dell'universo di **A** diversi in corrispondenza di simboli diversi: ciò è possibile per la cardinalità di **A**. Sia **A'** la struttura così ottenuta. Evidentemente $\mathbf{A}' \models S_0$. Come già detto, per il teorema di compattezza, ciò ci garantisce l'esistenza di una struttura **B** modello di $T \cup \Sigma$, e quindi anche di T, che dovrà avere cardinalità almeno κ. Volendo una struttura che risolva il nostro problema nel linguaggio di T, basta considerare la riduzione $\mathbf{B}_0$ di **B** al linguaggio di T; e ciò completa la dimostrazione.

Il risultato appena dimostrato può essere utilizzato per mostrare il seguente

TEOREMA DI LÖWENHEIM SKOLEM ASCENDENTE PER STRUTTURE. *Ogni struttura infinita può essere elementarmente immersa in una struttura di cardinalità maggiore di un cardinale arbitrariamente prefissato.*

Dimostrazione Sia **A** una struttura infinita, e κ un cardinale arbitrario. Sia **A'** l'espansione di **A** ad un linguaggio con un nuovo simbolo di costante individuale per ciascun elemento dell'universo di **A**, in cui ogni tale simbolo sia interpretato nell'elemento dell'universo in corrispondenza del quale è stato introdotto. Sia T la teoria di **A'**, cioè $T = \{\varphi: \mathbf{A}' \models \varphi, \varphi \text{ un enunciato del linguaggio per } \mathbf{A}'\}$. Sappiamo che **A'** è un modello di T infinito, che si immerge elementarmente in ogni modello di T. Così possiamo applicare a T il risultato precedente e concludere che T ha un modello **B** di cardinalità almeno κ. **B** è una struttura di cardinalità almeno κ in cui **A'** si immerge elementarmente. Se $\mathbf{B}_0$ è la riduzione di **B** al linguaggio di **A**, allora $\mathbf{B}_0$ è una struttura di cardinalità almeno κ in cui **A** si immerge elementarmente.

7. Teoremi di Löwenheim Skolem: parte seconda

Più avanti vedremo altre applicazioni del teorema di compattezza. Ora ci domandiamo se è possibile migliorare i risultati della sezione 5, nel senso di poter trovare un modello di cardinalità esattamente κ, e non di cardinalità almeno κ, con le proprietà dei risultati precedenti. Per affrontare questo problema è opportuno richiamare la dimostrazione del teorema di completezza. In essa, ad un certo punto, si doveva costruire una struttura che fosse modello di un opportuno insieme di formule Φ. Quella struttura è stata costruita proprio a partire dal linguaggio, prendendo come elemento dell'universo l'insieme dei nomi di uno stesso elemento, e limitando l'universo stesso agli elementi che hanno nome. Più precisamente, introdotta la relazione di equivalenza tra termini che considera equivalenti due termini t_1 e t_2 se la formula

$t_1=t_2$ è deducibile da Φ, l'universo della struttura modello di Φ era costituito dalle classi di equivalenza di termini rispetto a detta relazione.

L'osservazione centrale, per quanto ci interessa ora, è che la cardinalità della struttura costruita a partire dalle classi di equivalenza di termini dovrà per forza essere minore od uguale alla cardinalità del linguaggio. Combinando questo risultato con i precedenti si ottengono i seguenti ulteriori risultati, che vanno sotto il nome di teoremi di Löwenheim Skolem, il primo per le teorie, il secondo per le strutture.

TEOREMA DI LÖWENHEIM SKOLEM PER LE TEORIE. *Sia* T *una teoria con modelli finiti arbitrariamente grandi.* T *ha modelli di qualsiasi cardinalità infinita* κ *maggiore od uguale alla cardinalità del linguaggio di* T.

Dimostrazione. Si consideri l'insieme di enunciati $\Sigma = \{c_i \neq c_j$: i e j sono ordinali minori di κ tra loro diversi e c_i e c_j sono simboli per costanti non nel linguaggio di T$\}$. Abbiamo già visto che, grazie al teorema di compattezza si può mostrare che $T \cup \Sigma$ è una teoria soddisfacibile; ora possiamo aggiungere che il linguaggio di questa nuova teoria ha cardinalità esattamente κ per l'ipotesi fatta su κ. Per l'osservazione centrale di questa sezione, $T \cup \Sigma$ ha un modello di cardi
nalità minore od uguale alla cardinalità del linguaggio che è κ. D'altra parte non può avere cardinalità minore di κ, altrimenti non potrebbe soddisfare gli enunciati di Σ. Così $T \cup \Sigma$ ha un modello **A** di cardinalità esattamente κ nel linguaggio di $T \cup \Sigma$. La riduzione di **A** al linguaggio di T è un modello di T, nel linguaggio di T, di cardinalità esattamente κ, come si voleva.

TEOREMA DI LÖWENHEIM SKOLEM PER LE STRUTTURE. *Siano* **A** *una struttura infinita e* κ *un qualsiasi cardinale maggiore od uguale al massimo tra la cardinalità di* **A** *e la cardinalità del linguaggio adatto ad* **A** *. Allora* **A** *si immerge elementarmente in una struttura di cardinalità* κ.

Dimostrazione Si consideri, come già fatto altrove, la teoria di **A'**, espansione di **A** ad un linguaggio L' che ha un nome per ogni elemento dell'universo di **A**, cioè l'insieme $T = \{\varphi$: φ è un enunciato di L' tale che $\mathbf{A}' \models \varphi\}$. Ancora la teoria $T \cup \Sigma$ è soddisfacibile, dove Σ è l'insieme di enunciati già usato nel teorema di Löwenheim Skolem per le teorie. Per il risultato precedente $T \cup \Sigma$ ha un modello **B** di cardinalità esattamente κ. Sia **B'** la riduzione di **B** al linguaggio L. Poiché $\mathbf{B}' \models T$, **A'** si immerge elementarmente in **B'**. Sia $\mathbf{B}_0$ la riduzione di **B'** al linguaggio di **A**. **A** si immerge elementarmente in $\mathbf{B}_0$ che ha ancora cardinalità κ. Così anche questa dimostrazione è completa.

Il teorema di Löwenheim Skolem per teorie, in un certo senso, è più forte di quello per strutture. Più precisamente, mentre il primo, con le sue ipotesi, garantisce l'esistenza di modelli di

qualsiasi cardinalità maggiore od uguale a quella del linguaggio, il secondo, con le sue nuove ipotesi, garantisce solo l'esistenza di strutture di cardinalità maggiore od uguale al massimo tra la cardinalità del linguaggio e quella della struttura data, anche se aggiunge che di questi modelli alcuni sono estensioni elementari della struttura di partenza.

8. Teorema di Löwenheim Skolem: parte terza

Evidentemente una struttura non si può immergere elementarmente in un'altra struttura di cardinalità minore. Tuttavia ci possiamo chiedere se, data una struttura di cardinalità κ infinita in un linguaggio di cardinalità λ, con $\lambda<\kappa$, ci siano strutture di una qualsiasi cardinalità compresa tra λ e κ "fortemente legate" alla struttura data. Ovviamente bisognerà anzitutto precisare cosa si intende per "fortemente legate". Se per "fortemente legate" si intende che non siano distinguibili mediante il linguaggio, cioè che in esse siano veri gli stessi enunciati, allora la risposta affermativa è già pronta: basta considerare la teoria T della struttura data, e il teorema di Löwenheim Skolem per teorie applicato a T garantisce l'esistenza di modelli della cardinalità voluta di T, e in questi sono veri esattamente gli enunciati veri nella struttura data. Ma se per "fortemente legate" vogliamo intendere qualcosa di più, ad esempio che la nuova struttura si immerga elementarmente in quella data, allora il teorema di Löwenheim Skolem per teorie non è più sufficiente.

Serve il TEOREMA DI LÖWENHEIM SKOLEM DISCENDENTE che asserisce proprio che *una struttura di cardinalità* κ *maggiore od uguale alla cardinalità* λ *del suo linguaggio ha sottostrutture elementari il cui universo contenga un prefissato sottinsieme di cardinalità* ξ *dell'universo della struttura data, e la cui cardinalità sia arbitrariamente scelta tra* κ *e il massimo tra* λ *e* ξ. Questo teorema ha una dimostrazione molto interessante, che però non riporteremo qui poiché esce dai nostri traguardi e non sarà necessaria per quanto faremo in seguito.

9. Categoricità

Un problema che si affaccia in modo del tutto naturale nello studio della logica è il seguente. Il linguaggio è in grado di caratterizzare in modo univoco una prefissata struttura? Spesso a questo problema viene data una risposta positiva implicita nell'atteggiamento espresso dall'affermazione: se conosci ciò di cui vuoi parlare, e se conosci la lingua, devi essere in grado di descrivere compiutamente il tuo pensiero. Qui "ciò di cui vuoi parlare" può essere inteso come la struttura che si vuol descrivere (che si deve supporre completamente nota, altrimenti non ha neppure senso parlare di verità di un enunciato in quella struttura); e "la capacità di descrivere compiutamente il proprio pensiero" può essere intesa come la capacità di precisare univocamente la struttura che si considera attraverso il linguaggio (naturalmente precisazione univoca a meno di isomorfismi, perché due strutture isomorfe si comportano esattamente

nello stesso modo, e non possono essere distinte, né interessa distinguerle, mediante il linguaggio).

Si dice *categorica* una teoria che ha un solo modello a meno di isomorfismi. Con questa terminologia, il problema iniziale si può formulare così: data una struttura **A** c'è una teoria categorica T di cui è modello?

Poiché la teoria T è sicuramente contenuta nella teoria della struttura **A** (indicata con Th(**A**)), e dal momento che una teoria più ricca non può che precisare meglio una certa struttura, il problema iniziale può essere anche espresso cosi: la teoria di una data struttura è categorica?

Ovviamente la teoria di una struttura si scrive nel linguaggio adatto a quella struttura, che potrebbe anche essere abbastanza povero, sicché un risultato negativo potrebbe essere attribuito a tale povertà di linguaggio piuttosto che ad una impossibilità sancita dall'analisi in corso. Per superare questo sospetto, a partire da una struttura adatta ad un certo linguaggio, consideriamo una sua espansione ad un linguaggio più ricco che abbia almeno un nome per ciascun elemento della struttura. Ciò sicuramente non cambia sostanzialmente la struttura in questione. In questo spirito, riformuliamo in nostro problema nel seguente modo: data una struttura, esiste una sua espansione la cui teoria è categorica?

Il teorema di Löwenheim Skolem per strutture fornisce immediatamente una risposta negativa alla nostra domanda, se la struttura data è infinita. Infatti, proprio per questo teorema, la teoria di una qualsiasi espansione di una struttura infinita ha modelli di qualsiasi cardinalità maggiore od uguale alla cardinalità del linguaggio (si noti che, nel nostro caso, la cardinalità del linguaggio della struttura espansa è almeno uguale a quella della stessa struttura), e due modelli di diversa cardinalità certamente non possono essere isomorfi per l'impossibilità di una biiettività tra gli universi delle due strutture.

Così possiamo affermare che nessun linguaggio potrà mai caratterizzare univocamente (a meno di isomorfismi) una struttura infinita. Ben diversa è la situazione nel caso di strutture finite. Vale infatti il seguente risultato, che non dimostreremo: data una qualsiasi struttura finita, la sua teoria è categorica. Per ottenere questo risultato è essenziale che = sia un simbolo logico del linguaggio[15]. L'analisi condotta mostra un serio limite dei linguaggi formali del primo ordine. Questo limite è strettamente legato all'accettazione della nozione di infinito, nozione, d'altra parte, centrale in tutta la matematica. E' opportuno qui ricordare che sono aspetti essenziali della nozione di linguaggio del primo ordine sia la finitezza di ogni formula, che la effettività della costruzione delle formule, che la restrizione della quantificazione alle sole variabili individuali.

[15] Come si ricorderà, dire che = è un simbolo logico del linguaggio vuol dire che = deve essere interpretato in qualsiasi struttura nella relazione binaria che ad ogni elemento dell'universo associa esattamente lo stesso elemento.

Sono queste caratteristiche del linguaggio che gli impediscono di descrivere compiutamente le strutture infinite. Infatti si dimostra, ma non lo faremo qui, che, se si ammettono formule di lunghezza infinita (ad esempio consentendo la congiunzione di infinite componenti), cadono i teoremi di completezza, di compattezza, di Löwenheim Skolem, e si può caratterizzare, a meno di isomorfismi, ad esempio, l'usuale struttura intesa dei numeri naturali.

Anche se si ammette di poter quantificare su variabili per relazioni, ci si trova in una situazione come la precedente[16], pur di interpretare le variabili per relazioni come un qualsiasi sottinsieme della potenza cartesiana dell'universo delle struttura alla arietà della relazione. Ma la conoscenza di tutti i sottinsiemi di un insieme infinito comporta anche la conoscenza di quell'insieme. Sicché, sia in questo caso, come nel caso precedente, la possibilità di caratterizzare almeno i numeri naturali segue dall'assumere di conoscere già completamente i numeri naturali per poter costruire il linguaggio: così, però, non si risolve il problema di caratterizzare la struttura intesa dei numeri naturali e, più in generale, le strutture infinite.

L'ulteriore caratteristica dei linguaggi formali (che si possa decidere effettivamente quali sono le formule) è irrinunciabile (anche se non sempre rispettata dai linguaggi naturali), perché altrimenti tutto rimarrebbe nella più vaga indeterminazione e non si potrà precisare ciò che si può e ciò che non si può ottenere.

Abbiamo visto che le teorie di strutture infinite non sono categoriche essenzialmente per motivi di cardinalità dei possibili modelli. Sorge allora naturale la domanda se esistono modelli non isomorfi, ma della stessa cardinalità, della teoria di una struttura infinita. Si dice *α-categorica,* con α cardinale infinito, una teoria i cui modelli di cardinalità α sono tra loro isomorfi. Per quanto concerne l'α-categoricità di teorie di strutture la situazione si presenta più articolata. Ci sono esempi di teorie α-categoriche per ogni cardinale infinito α, teorie α-categoriche per certi cardinali infiniti α e non per altri, teorie mai α-categoriche.

L'analisi dettagliata di questa situazione va oltre lo scopo di questa esposizione, ma forse, vale la pena rilevare che l'α-categoricità dipende essenzialmente dalla ricchezza del linguaggio di una teoria, e dalla cardinalità del linguaggio stesso. Ad esempio, si dimostra, ma anche questa volta non lo faremo, che la teoria dell'usuale struttura dei numeri razionali in cui l'unica relazione extralogica a cui si attribuisce un nome sia quella d'ordine è α-categorica esattamente quando α è la cardinalità dei naturali, ma non lo è più se si aggiungono i simboli di funzione per le costanti 0 e 1 e per le funzioni addizione e moltiplicazione.

[16] In questo caso l'unicità, a meno di isomorfismi, dell'usuale struttura intesa dei numeri naturali è il classico risultato di Peano ottenuto sfruttando la sua assiomatizzazione dei numeri naturali.

10. L'aritmetica non è aleph-zero-categorica

Vogliamo invece provare il seguente

TEOREMA. *L'aritmetica (cioè la teoria dell'usuale struttura dei numeri naturali relativamente al linguaggio i cui simboli propri siano* $+ \times ' 0$, *da interpretarsi rispettivamente nelle funzioni addizione moltiplicazione, passaggio al successore, e costante* 0) *non è aleph-zero*[17]*-categorica.*

Dimostrazione. Per mostrare ciò, costruiremo un modello della teoria T, di cardinalità aleph-zero, non isomorfo ad **N** (indicheremo con **N** l'usuale struttura intesa dei numeri naturali) e di cui **N** sia sottostruttura elementare. Un modello di una teoria non isomorfo al modello usualmente inteso di quella teoria viene detto modello non standard, mentre viene chiamato standard il modello usuale. Mostreremo l'esistenza di un tale modello numerabile non standard dell'aritmetica, immergendo elementarmente **N** in una struttura numerabile con un elemento maggiore di tutti quelli corrispondenti agli elementi di **N**. Si noti che la relazione di maggiore ($x>y$) può essere espressa nel linguaggio indicato mediante la formula $x\neq y \wedge \exists z\ x=y+z$. Si noti anche che ciascun elemento n dell'universo di **N** ha nome, e precisamente $0''\cdots'$, in cui il nome dell'operazione successore viene ripetuto esattamente n volte: il termine $0''\cdots'$, indicato da $\underline{n}$, viene chiamato numerale di n.

Anzitutto mostriamo l'esistenza di un tale modello, poi faremo vedere che non è isomorfo ad **N**. Sfrutteremo ancora il teorema di compattezza. Sia T la teoria di **N**. **N** si immergerà elementarmente in ogni modello di T poiché nel linguaggio di T ci sono i numerali, cioè i nomi di ogni elemento dell'universo di **N**. Arricchiamo il linguaggio adatto ad **N** con un nome per una costante individuale, c, e consideriamo il seguente insieme Σ di enunciati: $\Sigma = \{c\neq\underline{n}: \underline{n} \text{ è un numerale}\}$. Si noti che il nuovo linguaggio L' è ancora numerabile.

Mostriamo ora che la teoria $T\cup\Sigma$ è soddisfacibile. Infatti ogni sottinsieme finito S di $T\cup\Sigma$ è soddisfacibile. Per mostrare ciò, iniziamo col notare che ci sarà un massimo numero n tale che il numerale $\underline{n}$ occorra in S (S è finito), se un qualche numerale occorre in S, altrimenti n sia 0. Espandiamo ora **N** a **N'**, struttura adatta al linguaggio di $T\cup\Sigma$, interpretando c nel numero n+1. Evidentemente tutti gli enunciati di S sono veri in **N'**. Allora, per il teorema di compattezza, anche la teoria $T\cup\Sigma$ è soddisfacibile.

Per il teorema di Löwenheim Skolem, $T\cup\Sigma$ ha un modello numerabile, chiamiamolo **A'**. Chiaramente l'interpretazione di c in **A'** deve essere un elemento maggiore di tutti quelli che interpretano i numerali (poiché $\mathbf{A}'\models\Sigma$). Sia **A** la riduzione di **A'** al linguaggio di T. **A** è numerabile, **N** si immerge elementarmente in **A** e i corrispondenti degli elementi dell'universo di **N** saranno le interpretazioni dei loro nomi.

[17] Aleph-zero è la cardinalità dell'insieme dei numeri naturali.

Mostriamo ora che **A** è la struttura che cercavamo facendo vedere che **A** non è isomorfa ad **N**. In effetti l'elemento, chiamiamo a, dell'universo di **A** in cui è interpretato il simbolo c, non può essere il corrispondente nell'isomorfismo ϕ di alcun elemento n dell'universo di **N**. Infatti, se fosse $a=\phi(n)$ allora sarebbe anche $a < \phi(n+1)$ (poiché $n< n+1$ e ϕ è un isomorfismo), contraddicendo il fatto che a è maggiore di tutti i corrispondenti nell'isomorfismo di elementi dell'universo di **N**. Così abbiamo mostrato, per assurdo, che la struttura **A** non può essere isomorfa ad **N**.

11. Archimedeità

In modo del tutto analogo a quanto appena visto si dimostra anche la seguente proposizione

TEOREMA. *Un campo archimedeo si immerge elementarmente in un campo non archimedeo della stessa cardinalità.*

Ci sono vari modi tra loro equivalenti di caratterizzare un campo archimedeo, che deve comunque essere ordinato. Qui adottiamo la seguente formulazione. Un campo è archimedeo se è ordinato e se per ogni suo elemento x maggiore di zero ha anche un elemento del tipo $1/n$ che sia minore di x, dove n è l'elemento del campo che si ottiene, a partire da 0 (l'elemento neutro rispetto all'addizione nel campo), aggiungendo 1 (l'elemento neutro rispetto alla moltiplicazione nel campo) n volte, con n numero naturale maggiore di 0.

Si noti la differenza tra n e n: n è un elemento del campo ottenuto in un certo modo, mentre n, in generale, non è un elemento del campo, ma un numero naturale che indica quante volte si è ripetuta l'operazione di aggiungere 1 a partire da 0 per ottenere n.

Poiché il fatto che una struttura sia un campo si può esprimere con un numero finito di enunciati[18], e l'ordine del campo è una sua relazione primitiva della struttura caratterizzabile mediante pochi[19] enunciati, allora anche ogni struttura in cui un campo ordinato si immerga elementarmente è un campo ordinato. Così per esprimere la non archimedeità di una struttura in cui si immerga elementarmente un campo archimedeo, bisogna affermare l'esistenza di un elemento positivo minore di tutti gli elementi del tipo $1/n$, con $n>0$. Sicché, per dimostrare la proposizione enunciata, si dovrà esibire una struttura in cui il campo dato si immerga elementarmente e avente un elemento positivo minore di tutti quelli del tipo $1/n$, con $n>0$.

Dimostrazione Al solito, sia **A** la struttura di campo ordinato di caratteristica zero di partenza, ed L il linguaggio adatto ad essa. Espandiamo **A** ad una struttura **A'** adatta ad un linguaggio L' in cui ci siano i nomi di ciascun elemento dell'universo di **A**. Si noti che la car-

[18] I soliti assiomi di campo ben noti dall'algebra.

[19] Si possono utilizzare le cosiddette proprietà antiriflessiva, transitiva, tricotomica, e le note proprietà che legano l'ordine alla operazioni di addizione e moltiplicazione.

dinalità di L' è uguale sia a quella di **A** che a quella di **A'**. Sia T la teoria di **A'**, cosicché **A** si immergerà elementarmente in ogni riduzione al linguaggio L di un modello di T.

Ampliamo ulteriormente L' a L" mediante l'aggiunta di un simbolo per costante, c, che vorrà indicare un elemento positivo minore di tutti quelli del tipo *1/n*, *n>0*. Anche la cardinalità di L" è uguale a quella di **A**.

Sia Σ l'insieme di enunciati che esprime questa caratteristica di c, e precisamente Σ = {c>0}∪{c< 1/n: n è un numero naturale e n>0} (0, 1, n sono i nomi degli elementi *0*, *1*, *n* rispettivamente). Dimostriamo che la teoria T∪Σ è soddisfacibile.

Anche questa volta sfrutteremo il teorema di compattezza, e cominciamo col far vedere che ogni sottinsieme finito S di T∪Σ è soddisfacibile. Sia m il più grande numero naturale tale che un enunciato del tipo c<1/m occorre in S, 1 se in S non ci sono enunciati di tale tipo. Si espanda **A'** ad una struttura **A"**, adatta al linguaggio di T∪Σ, interpretando c in *1/(m +1*). E' immediato che **A"** è modello di S.

Così ogni sottinsieme finito di T∪Σ è soddisfacibile, e, per il teorema di compattezza, anche T∪Σ è soddisfacibile. Per il teorema di Löwenheim Skolem, c'è un modello **B"** di T∪Σ di cardinalità uguale alla cardinalità di **A**. Per come abbiamo proceduto **A** si immerge elementarmente nella riduzione di **B"** a **B**, struttura adatta al linguaggio L. Pertanto, per quanto già osservato, **B** è un campo ordinato. Resta solo da far vedere che **B** non è archimedeo. Di fatto, l'interpretazione di c in **B"**, chiamiamola C, è un elemento positivo minore di tutti gli elementi del tipo *1/n*, n numero naturale maggiore di 0, sicché **B"** sarà non archimedeo proprio perché **B"** è modello di Σ. L'elemento C resta nell'universo di **B** con le stesse caratteristiche, anche se ora non ha più un nome specifico. Pertanto anche **B** sarà non archimedeo, come si voleva far vedere.

Il risultato di cui abbiamo appena completato la dimostrazione può essere formulato anche nel modo seguente. Non c'è alcun enunciato φ, né alcun insieme di enunciati Φ, veri in un campo ordinato se e solo se il campo è archimedeo. Infatti, con riferimento alla simbologia precedente, Φ dovrebbe essere contenuto in T, e T∪Σ dovrebbe essere non soddisfacibile, mentre abbiamo visto che lo è.

12. Relazioni concorrenti

Le situazioni studiate nelle due sezioni precedenti hanno delle caratteristiche comuni. In entrambi i casi, data una struttura, si cercava una sua estensione elementare che avesse un elemento in una certa relazione binaria R con ciascun elemento dell'universo della struttura data; ed il problema aveva soluzione, grazie al teorema di compattezza, poiché, scelto un qualsiasi sottinsieme finito S dell'universo della struttura data, c'era sempre un elemento in quell'universo in relazione R con ciascun elemento di S.

Una relazione binaria R si dice *concorrente* se è tale che, per ciascun insieme finito S, ci sia sempre un elemento in relazione con ogni elemento di S (espresso altrimenti: per ogni insieme finito S esiste b tale che per ogni a∈S si ha R(a,b)).

Vogliamo far vedere che la caratteristica essenziale che permette di ottenere i risultati delle due precedenti sezioni è proprio la concorrenza, mostrando che quei risultati valgono anche quando al posto delle relazioni considerate si prenda una qualsiasi relazione concorrente.

TEOREMA. *Siano* **A** *una struttura infinita*, R *un insieme di sue relazioni binarie concorrenti* R_i, i∈I, *con* I *un insieme di indici, e* κ *un cardinale maggiore od uguale al massimo tra la cardinalità di* **A**, *e quella del linguaggio* L *di* **A** *(si noti che* κ *è maggiore od uguale alla cardinalità di* I *poiché ciascuna delle relazioni* R_i *avrà un nome). In queste ipotesi esiste una struttura* **B** *di cardinalità* κ *in cui* **A** *si immerge elementarmente, tale che per ogni relazione* R_i *esiste un elemento* b_i *tale che per ogni a nell'universo di* **A** *si ha* $R_i(a,b_i)$.

Dimostrazione Al solito iniziamo considerando un'espansione **A'** di **A** ad un linguaggio L' con un nome a per ogni elemento a dell'universo A di **A**, nome da interpretarsi nell'elemento a. Sia T la teoria di **A'**. Ampliamo ulteriormente il linguaggio L' ad un linguaggio L" aggiungendo nuovi simboli di costante individuale c_i per ciascun i∈I (cioè in corrispondenza di ciascuna relazione concorrente considerata) e d_j per ogni j<κ. Si noti che la cardinalità di L" è esattamente κ.

Sia Σ l'insieme degli enunciati $P_i(a,c_i)$ al variare di a in A e di i in I (dove P_i è il predicato che è nome della relazione R_i) e degli enunciati $d_{j_1}\neq d_{j_2}$ per ogni coppia j_1,j_2 di elementi di κ tra loro diversi. Abbiamo ormai imparato che **A** si immerge elementarmente nella riduzione **B** ad L di un modello **B'** di cardinalità κ di T∪Σ, se T∪Σ, è consistente, che anche **B** avrà cardinalità κ, e che nell'universo B di **B'**, che è anche l'universo di **B**, ci saranno elementi b_i, che interpreteranno i simboli c_i, tali che per ogni a in A risulterà $R_i(a,b_i)$. Sicché, per ottenere il risultato voluto, basta far vedere che T∪Σ ha un modello di cardinalità κ. Ma la presenza in Σ degli enunciati del tipo $d_{j_1}\neq d_{j_2}$, grazie al teorema di Löwenheim Skolem, riduce il problema alla soddisfacibilità di T∪Σ.

Per dimostrare quest'ultimo punto ricorriamo ancora una volta al teorema di compattezza. Sia S un sottinsieme finito di T∪Σ. Facciamo vedere che S è soddisfacibile. Si espanda **A** alla struttura $\mathbf{A}_o$ adatta ad un linguaggio per S interpretando i simboli c_i in S in elementi α_i di A tali che $R_i(a,\alpha_i)$ per ogni a tale che $P_i(a,c_i)\in S$ e i simboli del tipo d_j in S in elementi tra loro diversi. Tutto ciò è possibile perché le relazioni R_i sono concorrenti e perché **A** è una struttura infinita, mentre S è un insieme finito. Con queste scelte è evidente che $\mathbf{A}_o|=S$, e ciò completa la dimostrazione.

13. Paradosso di Skolem

Vogliamo ora considerare una situazione particolarmente interessante per la matematica e, da certi punti di vista, sorprendente: essa riguarda l'usuale collezione degli insiemi con l'appartenenza. In senso stretto, ciò che stiamo considerando non è neppure una struttura, poiché il suo universo non è un insieme, ma una classe propria[20]. Anche l'appartenenza non è una relazione, nel senso che non è un insieme di coppie ordinate, ma una collezione propria di coppie ordinate.

Ma non tutto è perduto. Possiamo cercare di estendere la nozione di struttura per includere cose del tipo presentato. Per una maggior chiarezza espositiva, conveniamo di chiamare *realizzazione* il nuovo concetto, simile a quello di struttura, che stiamo per introdurre.

Una realizzazione non sarà più una terna, perché una terna deve avere tre elementi, mentre le classi proprie non sono elementi. Consideriamo, allora, alcune collezioni (in generale non una collezione di collezioni, perché ad una collezione appartengono solo elementi): una collezione da chiamarsi universo; una o più collezioni i cui elementi sono n-uple ordinate di elementi dell'universo, con n ben fissato per ciascuna di queste collezioni, (ciascuna di queste collezioni sarà detta una relazione); ed ancora delle collezioni (eventualmente nessuna) di (n+1)-uple ordinate di elementi dell'universo, con n ben fissato per ciascuna di queste collezioni, con l'ulteriore proprietà che, scelti i primi n elementi dell'(n+1)-upla, è unico l'(n+1)-esimo elemento tale che l'(n+1)-upla appartenga alla collezione (ciascuna di queste collezioni sarà detta una funzione).

Le realizzazioni possono non essere insiemi. Pertanto non si potrà più parlare di proprietà di certe realizzazioni, o di relazioni tra realizzazioni, o di collezioni di realizzazioni. Ma si potranno fare considerazioni che dipendono dagli elementi delle varie collezioni. In particolare si potrà ancora dire come interpretare un linguaggio in una realizzazione, ed anche dire quando una formula è vera in una realizzazione con certe attribuzioni di valori alle variabili libere: si può ancora considerare la teoria di quella realizzazione.

L'usuale collezione degli insiemi con l'appartenenza è una realizzazione e possiamo considerare la sua teoria T, che chiameremo teoria completa degli insiemi[21]. Se l'usuale "struttura" degli insiemi non è un'assurdità la sua teoria T sarà consistente, e, come tale, avrà un modello nel vecchio senso. Anzi, poiché il linguaggio è numerabile, avendo il solo simbolo extralogico $\in$, per il teorema di Löwenheim Skolem, ci saranno modelli numerabili di T. Ma tra gli enunciati di T ci sarà il cosiddetto teorema di Cantor, quello che afferma che

[20] Stiamo pensando ad un concetto di insieme per cui valga l'assioma di regolarità.

[21] Si sarebbe potuto far a meno di accennare alle realizzazioni, e pervenire ugualmente ad una teoria degli insiemi, non quella completa (se ha senso parlare di essa), ma una qualsiasi costruita a partire da opportuni assiomi. Ancora ci sarebbe stato il problema della sua consistenza, ma accettata questa, e il fatto che in essa ci sia il teorema di Cantor, l'esposizione potrebbe proseguire in modo del tutto analogo.

esistono insiemi strettamente più che numerabili, ad esempio l'insieme dei sottinsiemi dei naturali. Indichiamo con $\exists z \neg \psi(z)$ questo enunciato[22]. Esso dovrà essere vero in ogni modello di T, in particolare in un modello numerabile la cui esistenza era stata affermata nel capoverso precedente.

Non è strano che *in una struttura numerabile ci sia un elemento "più che numerabile* "? Detta in questo modo l'affermazione ha tutta l'aria di una assurdità!

Questa stranezza va sotto il nome di *paradosso di Skolem.* Si noti che, se il paradosso di Skolem fosse davvero un'assurdità, metterebbe in crisi la teoria degli insiemi e tutta la matematica moderna che si fonda su di essa: infatti per giungere al paradosso non si sono introdotte altre ipotesi se non quelle usuali della teoria degli insiemi e gli sviluppi di logica basati su di essa.

Ma la frase in corsivo va letta con attenzione. A ragion veduta la locuzione "più che numerabile" è stata messa tra virgolette poiché è l'interpretazione dell'enunciato $\exists z \neg \psi(z)$ in un modello numerabile di T in cui il predicato $\in$ sarà interpretato in una relazione di "appartenenza", chiamiamola E, che può non avere nulla a che vedere con l'usuale relazione di appartenenza (essere uno degli elementi di una certa collezione)[23]. Così, il termine "più che numerabile" non è detto che indichi ciò a cui siamo abituati.

Sostanzialmente, se d è l'elemento di un modello M voluto dall'enunciato $\exists z \neg \psi(z)$ vero in M, allora l'insieme che deve essere "più che numerabile" sarà l'insieme $D=\{y: yEd\}$. D'altra parte D è un sottinsieme dell'universo numerabile, e, pertanto, D sarà al più numerabile, pur dovendo essere "più che numerabile": ecco il paradosso. Niente dice che D debba appartenere all'universo del modello (pur appartenendovi d), né che d debba essere un sottinsieme dello stesso universo (pur essendolo D).

Osserviamo che il paradosso si gioca tutto sul contrasto tra l'affermazione di numerabilità dell'universo della struttura, e l'interpretazione dell'enunciato che proclama l'esistenza di un elemento "più che numerabile". L'affermazione di numerabilità della struttura è, evidentemente, effettuata nel metalinguaggio, dal di fuori della stessa struttura; mentre l'enunciato che proclama l'esistenza di un elemento "più che numerabile" è nella teoria della struttura e fa una affermazione dal di dentro della struttura.

Questa osservazione non porterebbe alla soluzione del paradosso se la sottile distinzione tra affermazioni dal di dentro e dal di fuori della struttura fosse priva di senso,

[22] $\psi(z)$ sarà una formula della teoria degli insiemi, nella sola variabile libera z, che dice che c'è una funzione iniettiva da z in ω (l'ordinale dei numeri naturali).

[23] Si noti che il modello può essere il più stravagante possibile pur di rendere veri gli enunciati di T; magari può essere quello costruito a partire dalle classi di equivalenza dei termini, e quindi la relazione che interpreta $\in$ in questa struttura può essere ben lontana dal significato che usiamo dare alla parola appartenenza: certamente, se l'enunciato $t_1 \in t_2$ sta in T, non è vero che la classe di equivalenza del termine t_1 sia un elemento della classe di equivalenza del termine t_2.

cioè se le affermazioni dal di dentro e dal di fuori della struttura dessero le stesse informazioni sulla struttura.In effetti per certe affermazioni è proprio così. Si pensi ad esempio all'enunciato $\exists v_1 \exists v_2 (v_1 \neq v_2)$, chiamiamolo σ: dire che σ è vero in una struttura equivale a dire che in quella struttura ci sono almeno due elementi. Detto altrimenti, $\mathbf{A} \models \sigma$ se solo se l'universo della struttura ha almeno due elementi. Ma l'ultima affermazione, che indicheremo con $<\sigma>$ non è altro che l'enunciato σ letto nel metalinguaggio e relativizzato all'universo di **A**.

Cioè un enunciato è vero in una struttura se e solo se quella struttura ha le caratteristiche espresse dall'enunciato letto da fuori, nel metalinguaggio, ovvero se quell'enunciato dice le stesse cose dal di dentro della struttura che dal di fuori. Quando sia ben chiaro in quale struttura vadano fatte le interpretazioni, quanto detto diviene: σ e vero in **A** se e solo se $<\sigma>$; che è come dire, in modo più colorito, un enunciato è vero se e solo se vale quanto affermato da quell'enunciato nel metalinguaggio, ossia un enunciato è vero dal di dentro di una struttura se e solo se è vero dal di fuori (cioè se e solo se la situazione è proprio come descritta dall'enunciato). Così, se il paradosso di Skolem non è un'assurdità (e guai se lo fosse, come abbiamo osservato), ci dovranno essere degli enunciati la cui verità non equivale a quanto essi affermano nel metalinguaggio, enunciati che non dicono dal di dentro esattamente quanto dicono dal di fuori.

In effetti le cose stanno proprio così per la formula $\psi(z)$, che dice che z ha cardinalità minore o eguale a quella di ω. Essa è del tipo $\exists f \ \eta(f,z)$ dove $\eta(f,z)$ è una formula che dice che f è una funzione iniettiva da z in ω, l'ordinale dei numeri naturali. Ora affinché la formula $\psi(z)$ non sia vera in un modello numerabile di T, diciamo $\mathbf{A}=(A,E)$, quando z viene valutata nell'elemento d, bisogna che in A non ci sia alcuna funzione iniettiva f_0 da d in ω. Detto in altre parole, affinché d sia numerabile, dal di dentro c'è l'ulteriore *condicio sine qua non* che la funzione f_0 sia elemento di A.

Così è possibile che ci siano funzioni iniettive da d in ω (e ciò fa sì che d sia numerabile dal di fuori), ma tali funzioni non saranno in A (per cui d sarà non numerabile dal di dentro). Ci siamo così resi conto che non è equivalente il guardare una affermazione dal di dentro e dal di fuori.

14. Conseguenze didattiche

Finora abbiamo visto, grazie al teorema di compattezza e al teorema di Löwenheim Skolem, la non categoricità delle teorie con modelli infiniti, l'esistenza di modelli non standard, l'esistenza di formule non assolute con la conseguente necessità di distinguere tra affermazioni dal di dentro e dal di fuori. Ora vorremmo vedere che conseguenze hanno questi risultati sulla didattica della matematica. Ovviamente non propongo di ripetere la presentazione di questi argomenti agli studenti delle scuole superiori, ma penso che debbano

essere conosciuti dagli insegnanti perché portano ad una visione della matematica che può modificare l'intero percorso dell'insegnamento di questa materia.

Pensiamo a come avviene oggi, ad esempio, l'insegnamento dei numeri naturali. Direi che nella scuola elementare si insegna l'uso di certe parole per contare. Poi i numeri naturali sono dati per noti. Che concetto di numero naturale possono essersi fatti gli studenti dopo una tale presentazione? Sono consapevoli delle scelte fatte per costruire questo sistema numerico, oppure pensano che i numeri esistano indipendentemente dalle razionalizzazioni umane? Dei numeri si devono cogliere il senso ed il comportamento? Sono enti perfettamente noti (forse non allo studente, ma allo studioso), o ci sono ancora aspetti che rimangono oscuri?

Come possiamo pretendere che uno studente sappia rispondere a queste domande dopo un usuale corso di studi offerto dalla nostra scuola? E magari gli chiediamo di usare con sicurezza il principio d'induzione perché serve in varie circostanze.

La situazione non cambia molto se passiamo dall'aritmetica alla geometria, anche se su questa si insiste molto più diffusamente anche negli anni successivi alle elementari. Che cosa studia la geometria? Lo spazio reale, o uno spazio immaginario, o le figure geometriche, e cosa sono queste? E, se adottiamo una presentazione assiomatica di questa disciplina, che valore e significato hanno gli assiomi, cosa sono?

Per non parlare poi dei numeri reali, il cui stesso nome suggerisce che siano quelli veri, quelli che sono effettivamente presenti nella realtà. Cos'è la continuità, perché è importante, in che rapporti è con l'approssimazione? La continuità si coglie solo attraverso i reali, o ci sono altri sistemi numerici ispirati ad essa?

Forse non è neppure giusto fare tutte queste domande ad uno studente, ma l'insegnante in qualche modo deve aver dato le sue risposte, magari inconsciamente, per poter impostare la propria azione didattica. E' auspicabile poi che le risposte date non siano ingiustificate posizioni personali, ma tengano conto delle conoscenze acquisite dalla ricerca.

Già la scoperta delle geometrie non euclidee aveva fatto capire che gli assiomi non sono affermazioni evidentemente vere, ma scelte tra varie possibili posizioni. I risultati di non categoricità e di non α-categoricità, ottenuti grazie al teorema di compattezza, si collocano, in qualche modo, nello stesso filone dei risultati geometrici citati, ma, direi, ad un maggior livello di profondità. Ora viene messa in discussione la stessa possibilità di descrivere con precisione enti che scegliamo di costruire, ma, evidentemente, senza poter cogliere appieno la portata di quanto abbiamo scelto di fare.

La consapevolezza dei limiti del linguaggio ci deve rendere cauti su quanto richiediamo agli studenti (forse la spiegazione fatta agli studenti non escludeva certe interpretazioni che non volevamo consentire) e guardinghi nella esposizione, coscienti che certe nozioni non possono essere trasmesse con precisione attraverso il linguaggio. Che fare in questi casi? La mia posizione è la seguente:

Anzitutto non nascondere la difficoltà agli studenti, non imbrogliarli facendo credere che il concetto trasmesso sia perfettamente individuato da una pretesa definizione.

Poi far ricorso ai dati dell'esperienza, alla problematica da risolvere, e a quant'altro per far intuire il senso di ciò che proponiamo, pur non essendo in grado di dire esattamente cos'è.

Quindi dichiarare che, pur nell'imprecisione, riteniamo che almeno certi comportamenti (esplicitati negli assiomi, anche informali) debbano essere accettati per approssimare la nozione che vogliamo trasmettere. Ciò comporta che si debbano accettare anche le conseguenze degli assiomi.

Infine, difendere le scelte fatte dalla matematica (ma bisogna conoscerle), non tanto perché sono le uniche giuste e possibili, ma per la fecondità e la convenienza dimostrata nel tempo, sempre disponibili ad accettare altre impostazioni che si dimostrassero più opportune, ma con il sano scetticismo di chi sa quanto è stato duro arrivare alle posizioni proposte, e quanto sarà difficile migliorarle.

Abbiamo notato precedentemente, e non bisogna dimenticarlo a questo punto, il ruolo dell'infinito. Quando tutto si svolge in un mondo limitato da un ben precisato confine finito, allora il linguaggio è sufficiente per descrivere compiutamente la struttura che si vuol proporre. Ma, non appena si permettono quantità anche finite, ma arbitrariamente grandi, si sono già ritrovati tutti i possibili problemi dovuti ai limiti espressivi del linguaggio.

Finora abbiamo considerato le conseguenze didattiche della non categoricità delle strutture infinite e della collegata non esprimibilità attraverso il linguaggio di certe proprietà delle strutture, ma non abbiamo scomodato il paradosso di Skolem. Il motivo per cui ritengo importante conoscere il paradosso di Skolem ai fini della didattica è perché mette in luce l'esistenza di nozioni non assolute che hanno significati diversi dal di dentro e dal di fuori del mondo che si vuole considerare.

Ciò si collega direttamente alle difficoltà didattiche di iniziare un argomento: da dove cominciare la presentazione? Forse per prima cosa si dovrebbe decidere se presentarlo dal di dentro o dal di fuori, quando tali opzioni sono possibili. A priori vedo varie vie per introdurre un nuovo concetto.

1) Attraverso esempi mostrare come si usa una parola, senza assolutamente impegnarsi sul suo significato, lasciando che lo studente si costruisca lui una sua immagine mentale del concetto che vogliamo comunicare. Può essere il caso della presentazione dei numeri naturali alle elementari come vocabolario per contare. D'altra parte questo è il modo naturale con cui si comincia ad apprendere la lingua madre quando non c'è altra via comunicativa disponibile.

2) Lasciare il concetto che interessa incognito, da scoprire. Si cercherà di determinare solo alcune caratteristiche del concetto su cui si concorda affinché quel concetto sia utile nell'affrontare i problemi per i quali si ha interesse a formularlo; poi il concetto stesso sarà

una qualunque cosa che goda delle caratteristiche stipulate. Certe presentazioni del concetto di dimostrazione seguono questa via: non si sa cosa vuol dire dimostrare, ma, qualunque cosa voglia dire, dovrà avere almeno certe proprietà; sicché si propone di chiamare dimostrazione una qualsiasi cosa che goda delle proprietà stabilite.

3) Partire da una idea del concetto che si vuol comunicare, cercando di illustrarla al meglio dal di dentro della struttura che si vuol costruire. Così si stabiliranno gli aspetti chiave mediante enunciati (assiomi) nel linguaggio della struttura, eventualmente motivati da osservazioni metalinguistiche su ciò che si vuol fare. Gli assiomi dovranno essere sufficienti per poter dedurre da essi tutti gli altri aspetti di cui si intende far uso. Anche se l'ascoltatore si farà un concetto diverso da quello che intendiamo proporre, sarà sufficientemente analogo per gli scopi che ci proponiamo, forse anche più di quanto lo imporrebbero i soli assiomi. Un esempio abbastanza tipico di uso di questo metodo può essere il seguente. Si può avere un'idea del concetto di insieme come una collezione, quasi un elenco di elementi, che può essere pensato come cosa singola e gli assiomi cercheranno di cogliere gli aspetti salienti di questa proposta. E' ben difficile immaginare di poter descrivere un tale concetto da un mondo esterno a quello degli insiemi: che possiamo inventare di più generale di tanto? Sicché siamo quasi costretti a presentare il concetto di insieme dal di dentro mediante enunciati nel linguaggio degli insiemi.

4) Ancora si parte da un'idea del concetto che si vuol illustrare, ma si considerano anche le esigenze che spingono a introdurre il concetto in questione, i problemi che deve risolvere e ci si mette in un ambiente esterno a quello da costruire in cui si dispone di un linguaggio per descrivere i problemi che il nuovo concetto dovrà risolvere e la struttura in cui vive il nuovo concetto dall'esterno. Forse un esempio ben noto può chiarire meglio questa via. Penso alla presentazione dei numeri naturali all'interno della teoria degli insiemi: essi sono particolari insiemi adatti a contare gli elementi di un qualsiasi insieme finito. Gli oggetti da contare, gli insiemi finiti sono all'interno della teoria degli insiemi. In essa si può anche descrivere l'operazione di contare, e si possono trovare degli elementi che risolvono il problema. Infine si può considerare la struttura costituita dagli elementi che risolvono il problema con le loro proprietà caratteristiche.

Le vie 3) e 4) si prestano più facilmente a fornire delle motivazioni allo studio da parte dello studente, mentre per le altre è forte il rischio di fraintendimenti. Può essere che ci siano anche altre vie ma queste mi paiono almeno le principali. In certe situazioni la scelta di una via da utilizzare per la presentazione può essere obbligata o almeno limitata (la quarta via non sembra praticabile per presentare gli insiemi). A volte, invece, si tratta di fare una scelta didattica su quale via seguire essendo tutte possibili. E' il caso dei numeri naturali o anche dei numeri reali.

Non voglio qui addentrarmi in una discussione su quale sia la via più opportuna in tali ed in altri casi. Penso che sia essenzialmente un problema che ciascun insegnante deve risolvere per proprio conto in funzione del livello scolare e degli obiettivi che si prefigge con il proprio insegnamento, ed anche della facilità di esplicazione e di comprensione del concetto usando il metodo scelto, ed infine della sensibilità e del punto di vista personale. Tuttavia mi pare importante che un insegnante sia cosciente della scelta che fa e che si attenga ad essa, perché mescolando i metodi si casca facilmente in circoli viziosi e si disorienta lo studente.

Concludendo, anche se gli argomenti di logica toccati sono ben lontani da quanto si può presentare in una scuola superiore, appaiono bagaglio indispensabile per un docente che voglia affrontare consapevolmente i più comuni problemi dell'insegnamento della matematica.

Bibliografia

I risultati tecnici menzionati (fuorché la particolare traccia di dimostrazione del teorema di completezza per un linguaggio con una quantità più che numerabile di simboli) si trovano in genere su un qualsiasi manuale di logica. Di sicuro sono tutti trattati esaustivamente nel manuale

J.L.BELL, M.MACHOVER *A course in mathematical logic*, North Holland, Amsterdam, 1977.

Esercizi

RUGGERO FERRO

ESERCIZIO 1 Completare le tracce di dimostrazioni sparse nel testo, in particolare nella prima sezione.

ESERCIZIO 2 I seguenti punti faranno vedere che la teoria di una struttura di una certa cardinalità finita assegnata è categorica, mettendo così ancor più in risalto il ruolo dell'infinito.

a) Esiste un enunciato che dice quanti sono gli elementi di un suo modello finito, cioè un enunciato vero in tutte e sole se strutture che abbiano un certo ben precisato numero finito come cardinalità. Per esercizio, esibire un tale enunciato.

b) Siano **A** e **B** due strutture di cardinalità n, un certo numero naturale, adatte ad un linguaggio L. Si supponga che le due strutture $(\mathbf{A}, a_0, \ldots, a_i)$ e $(\mathbf{B}, b_0, \ldots, b_i)$, con $i < n-1$, siano elementarmente equivalenti relativamente al linguaggio L' ottenuto da L aggiungendo le nuove costanti $c_0, \ldots, c_i$. Allora, comunque scelto un elemento a dell'universo di **A** (un elemento b dell'universo di **B**) esiste un elemento b dell'universo di **B** (un elemento a dell'universo di **A**) tale che le due strutture $(\mathbf{A}, a_0, \ldots, a_i, a)$ e $(\mathbf{B}, b_0, \ldots, b_i, b)$ siano elementarmente equivalenti rispetto al linguaggio L" ottenuto da L aggiungendo le nuove costanti $c_0, \ldots, c_i, c$.

Infatti, sia Φ l'insieme degli enunciati di L" veri in $(\mathbf{A}, a_0, \ldots, a_i, a)$. Sia v una variabile; si sostituisca v a c in ogni enunciato di Φ ottenendo un insieme Φ' di formule nella variabile libera v (per evitare cattura di variabili si eseguano i dovuti cambi alfabetici). Si quantifichi ogni formula di Φ' esistenzialmente rispetto a v. Si ottiene così un nuovo insieme $\Phi"$ di enunciati di L' veri in $(\mathbf{A}, a_0, \ldots, a_i)$, e dunque anche in $(\mathbf{B}, b_0, \ldots, b_i)$, per l'ipotesi fatta.

Ma ciascuna formula di $\Phi"$ e' esistenziale, e sarà vera in $(\mathbf{B}, b_0, \ldots, b_i)$ se e solo se c'è una attribuzione di valore alla variabile v che renda vera la formula privata della quantificazione iniziale. A priori detta attribuzione di valore alla variabile v può cambiare da formula a formula, e potrebbe succedere che non ci sia una attribuzione di valore alla variabile v che renda vere in $(\mathbf{B}, b_0, \ldots, b_i)$ tutte le formule di Φ', ma di fatto non è così.

Infatti, se per assurdo fosse così, i sottinsiemi Φ_x di Φ' delle formule non vere in $(\mathbf{B}, b_0,\ldots,b_i)$ quando alla v viene attribuito il valore x dell'universo di **B** non sono vuoti al variare di x. Sia φ_x una formula di Φ_x, e si consideri la congiunzione φ di queste formule, possibile perché x varia in un insieme finito. Per nessuna attribuzione di valore alla variabile v, può succedere che φ sia vera in $(\mathbf{B}, b_0,\ldots,b_i)$. Si consideri ora la formula $\varphi(v/c)$: è un enuncia-

to di L" tale che $(\mathbf{A},a_0,...,a_i,a)\models\varphi(v/c)$, sicché $(\mathbf{A},a_0,...,a_i) \models \exists v\varphi(v)$, ed anche, per l'ipotesi di elementare equivalenza, $(\mathbf{B}, b_0,...,b_i)\models\exists v\varphi(v)$. Ciò contraddice l'affermazione che φ non può essere vera in $(\mathbf{B}, b_0,...,b_i)$ per nessuna attribuzione di valore alla variabile v. Così ci deve essere un elemento b in B che attribuito alla variabile v rende vere in $(\mathbf{B},b_0,..., b_i)$ tutte le formule di Φ'. Lo stesso b può essere l'interpretazione di c ed allora nella struttura $(\mathbf{B},b_0,...,b_i,b)$ sono veri gli enunciati di Φ. Ne segue l'elementare equivalenza rispetto al linguaggio L" tra $(\mathbf{A}, a_0, \dots, a_i, a)$ e $(\mathbf{B}, b_0, \dots, b_i, b)$.

c) Applicando ripetutamente b), dimostrare che, se **A** e **B** sono elementarmente equivalenti, allora ci sono indicizzazioni $a_0, \dots, a_{n-1}$ di **A** e $b_0, \dots, b_{n-1}$ di **B** tali che $(\mathbf{A}, a_0, \dots, a_{n-1})$ e $(\mathbf{B}, b_0, \dots, b_{n-1})$ sono elementarmente equivalenti relativamente al linguaggio L* ottenuto da L aggiungendo le nuove costanti $c_0, \dots, c_{n-1}$. Cosi' **A** e **B** sono strutture isomorfe rispetto sia al linguaggio L che al linguaggio L*, per un risultato già incontrato.

d) Sia **A** una struttura di cardinalità n, un numero finito. Sia **B** un qualsiasi modello della teoria di **A**. Allora **B** è isomorfo ad **A** e la teoria di **A** è categorica.

Il Teorema di Incompletezza di Gödel

FERDINANDO ARZARELLO
Dipartimento di Matematica
Università di Torino
via Carlo Alberto 10
10123 Torino

1. Introduzione storico-critica

Kurt Gödel (1906-1978), laureatosi nel 1929 (nella sua tesi dimostrò la completezza del calcolo dei predicati), scrisse il suo lavoro più importante nell'estate del 1930 a Vienna. Esso sarebbe apparso l'anno successivo nei "Monatshefte für Mathematik und Physik" (il titolo era: *Sulle proposizioni formalmente indecidibili dei* Principia Mathematica *e sistemi affini I*). Il suo contenuto è schizzato nelle prime righe: dopo avere ricordato la crescente formalizzazione della matematica, per cui le "dimostrazioni possono essere svolte seguendo solo alcune regole meccaniche", a partire da alcuni assiomi, come avviene nei *Principia Mathematica* di Whitehead e Russell e nella teoria degli insiemi di Zermelo e Fraenkel, egli così prosegue:

"Si potrebbe perciò pensare che questi assiomi e regole di inferenza siano sufficienti per decidere *tutte* le questioni matematiche che possono essere formalmente espresse in questi sistemi. Sarà mostrato oltre che non è questo il caso, che al contrario esistono nei due sistemi menzionati problemi relativamente semplici della teoria dei numeri naturali che non possono essere decisi sulla base degli assiomi" (trad. di G.Lolli, Incompletezza, Il Mulino, 1992, p.8).

In linguaggio più moderno, i sistemi assiomatici dei Principia Mathematica e di Zermelo Fraenkel sono incompleti; esistono al loro interno enunciati indecidibili; anzi la natura di tali enunciati e le ipotesi necessarie per sviluppare la dimostrazione di non decidibilità sono così elementari che anche sistemi meno potenti di questi presentano lo stesso 'inconveniente': ciò avviene per esempio per il sistema dell'aritmetica di Peano

formulato nel linguaggio della logica del primo ordine (assiomi logici + regole di inferenza), cioè in un linguaggio come quello usato da Whitehead e Russell nel loro lavoro.

I risultati di incompletezza e indecidibilità si inquadrano classicamente nella problematica tipica delle scuole fondazionali dei primi tre decenni del secolo ventesimo, in particolare quella di matrice hilbertiana. Il (primo) teorema di Gödel pone fine in questo senso al cosiddetto programma fondazionale di Hilbert.

In questo paragrafo introduttivo accennerò, schematicamente, ad alcuni punti essenziali della problematica, rimandando al citato libro di Lolli per approfondimenti e commenti ulteriori.

Ci sono due aspetti almeno da evidenziare nel lavoro fondazionale di Hilbert e della sua scuola, che sono rilevanti per il risultato di Gödel, cioè:

a) L'Entscheidungsproblem, il problema della decidibilità di ogni questione matematica mediante un numero finito di operazioni;
b) La formalizzazione della matematica.

In tali problematiche troviamo mescolati, a volte in modo confuso, vari concetti che solo il lavoro di Gödel avrebbe messo a fuoco in tutta la loro chiarezza:

(i) la *categoricità* di un sistema di assiomi (cioè il fatto che essi abbiano un unico modello, a meno di isomorfismi);
(ii) la *completezza* di un sistema di assiomi (cioè il fatto che ogni enunciato sia dimostrabile o refutabile): questa proprietà veniva spesso vista come non ramificazione del sistema (impossibilità che sia un enunciato sia la sua negazione fossero compatibili con gli assiomi);
(iii) la *decidibilità* del sistema (ogni questione che rientra nel sistema viene appunto decisa con un procedimento meccanico).

L'intrico di tali concetti nei due filoni di problemi costituisce il groviglio entro cui nasce e sboccia il risultato di Gödel, che d'ora in avanti indicheremo con TG1.

Vediamo una breve sintesi della faccenda.

Il punto di partenza (e di arrivo!) è ben rappresentato dalla lista dei 23 problemi posti da Hilbert alla comunità matematica nella sua conferenza al secondo Congresso Internazionale dei Matematici (Parigi, agosto 1900). Nell'introduzione, la filosofia della matematica elaborata da Hilbert in quel periodo è ben sintetizzata. Il punto della sua discussione riguarda i problemi in matematica e per i matematici: "Fin tanto che una branca della scienza offre abbondanza di problemi, essa è viva; la mancanza di problemi preannuncia l'estinzione o la cessazione di uno sviluppo indipendente".

Dopo avere esaminato le diverse fonti di problemi per la matematica, egli esamina la questione di che cosa costituisca una soluzione: "...sarà possibile stabilire la correttezza della soluzione in un numero finito di passi basati su di un numero finito di ipotesi, che sono implicate dall'enunciato del problema e che devono essere correttamente formulate. Tale richiesta di deduzione logica per mezzo di un numero finito di processi è semplicemente la richiesta di rigore nel ragionamento." E oltre aggiunge: "Quando siamo coinvolti nello studio dei fondamenti di una scienza, dobbiamo elaborare un sistema di assiomi che contiene una descrizione esatta e completa delle relazioni che sussistono tra le idee elementari di quella scienza... nessun enunciato nell'ambito di quella scienza sulle cui fondamenta stiamo investigando si deve assumere come corretto a meno che possa essere derivato dagli assiomi in un numero finito di passi".

Particolarmente interessante il decimo problema, che suona così: "Data un'equazione diofantea in un numero qualsiasi di incognite e a coefficienti numerici interi razionali: individuare un processo secondo il quale si possa determinare in un numero finito di operazioni se l'equazione ha soluzioni intere razionali".

[Come ben noto, nel 1970 si chiudeva grazie al russo Matiasievich l'ultimo anello di una lunga storia, in cui si dimostrava che nessun algoritmo può fare quanto richiesto da Hilbert nel suo decimo problema: anche questo problema risulta indecidibile!].

In questo il decimo problema è singolare: è l'unico caso in cui non si chiede di dimostrare o confutare un'asserzione ma un algoritmo che esibisca direttamente le risposte positive o negative fornite da tali dimostrazioni e confutazioni (dell'esistenza di una soluzione all'equazione diofantea).

Nella ricerca della soluzione di un problema (che può consistere in una dimostrazione o in una confutazione dell'asserzione del problema stesso) si possono incontrare difficoltà varie: tipicamente si può trovare che il problema è mal posto, ad esempio le ipotesi possono risultare insufficienti, ecc.

In tal caso il problema va riformulato in modo corretto e quindi si passa alla soluzione del nuovo problema: infatti secondo Hilbert la soluzione è sempre possibile: "Prendete un qualunque preciso problema non risolto, come l'irrazionalità della costante C di Eulero-Mascheroni, o l'esistenza di un numero infinito di primi della forma 2^n+1. Per quanto intrattabili tali problemi ci possano sembrare e per quanto noi possiamo restare senza speranza di fronte ad essi, ciononostante abbiamo la ferma convinzione che la loro soluzione deve seguire da un numero finito di processi puramente logici.

È questo assioma della risolubilità di ogni problema una caratteristica peculiarità del solo pensiero matematico, o è una legge generale che riguarda la natura della mente, per cui tutte le questioni che si pone debbano perciò essere risolubili?"

"La convinzione della risolubilità di ogni problema matematico è un potente incentivo a chi ci lavora. Noi sentiamo di continuo il richiamo: Ecco il problema. Cerca la soluzione. Tu lo puoi trovare con la pura ragione, in quanto in matematica non esistono *ignorabimus*".

Hilbert formulava questo principio in sintonia con la mentalità del tempo: la sua discussione è un po' il ballo Excelsior della matematica, in un certo senso.

Egli avrebbe precisato successivamente in forma più precisa il senso del suo principio, chiarendo il ruolo fondamentale giocato in ciò dalla formalizzazione. Nel 1904 infatti sottolinea "la rilevanza teorica della traduzione dei testi matematici in linguaggi simbolici, anche prescindendo dalla utilità o meno per la comunicazione tra matematici; la traduzione permette di spostare le questioni relative al senso e alla coerenza dei discorsi matematici...a questioni relative a manipolazioni sintattiche, combinatorie, di un ben definito insieme di segni" (Lolli, op. cit., p.12).

Nel 1917 "formula un vero e proprio programma, che mira a provare la non contraddittorietà delle teorie matematiche dimostrando la impossibilità che mai compaiano due figure fisicamente contraddittorie nella generazione sistematica di tutte le formule, secondo le regole prefissate"(ibid.).

Hilbert avrebbe trovato fieri oppositori al suo credo: in particolare intorno agli anni venti si sarebbe scontrato aspramente con L.E.J.Brouwer, il fondatore dell'intuizionismo, che criticava il principio di risolubilità di ogni problema e le regole logiche che ne conseguivano (terzo escluso, doppia negazione). Ciò nonostante, egli sarebbe rimasto fedele a questo principio fino alla fine della sua carriera accademica: ad esempio, l'otto settembre 1930, in occasione del suo pensionamento tenne un discorso alla radio di Könisberg in cui prendeva veementemente posizione contro lo scetticismo in matematica:

> "Invece dello sciocco ignorabimus, noi ci rifacciamo al nostro motto: Noi dobbiamo sapere, noi vogliamo sapere"

(Sulla sua pietra tombale è riportato proprio questo motto: "WIR MUSSEN KENNEN, WIR WILLEN KENNEN").

In effetti, la non contraddittorietà di una teoria è da considerarsi per gran parte dei matematici contemporanei di Hilbert (ad esempio Poincaré) come condizione sufficiente per l'esistenza degli enti dei quali la teoria stessa parla. Quindi la non contraddittorietà assicura l'esistenza di un modello (almeno) per la teoria.

Inoltre, se la teoria è anche categorica, essa ha un solo modello; ne segue subito la sua completezza: tutto ciò che è vero si dimostra. Ciò implica la risolubilità di ogni problema matematico relativo alla teoria: infatti se così non fosse, un certo problema potrebbe avere

coerentemente due risposte diverse, una positiva e una negativa; ne seguirebbe l'esistenza di due modelli non isomorfi, contro la categoricità della teoria.

Affrontare la categoricità in termini di completezza ha il vantaggio di evitare il riferimento a strutture infinite; permette di affrontare la questione fondazionale secondi i rigidi canoni del finitismo, così cari a Hilbert. Nel 1928 al congresso internazionale dei matematici a Bologna, Hilbert formula così la completezza:

"se agli assiomi della teoria dei numeri viene aggiunta una formula appartenente alla teoria dei numeri ma non dimostrabile, allora dal sistema di assiomi esteso può essere derivata una contraddizione" (trad. di Lolli, cit., p. 26).

In sostanza, il programma di Hilbert è di provare la completezza delle teorie matematiche (ad esempio, la teoria dei numeri), in ultima analisi di tutta la matematica, usando strumenti finitisti, in quanto ci si riduce ogni volta alla non contraddittorietà di certi sistemi.

Il programma di Hilbert può essere riassunto nel seguente asserto: "dare una descrizione assiomatica completa della matematica da cui ogni asserzione matematica possa essere decisa in un numero finito di passi logici".

Oggi, con tutti i risultati di indecidibilità da cui siamo invasi, la cosa non appare sensata: ci circondano molti ignorabimus. Ma ciò si chiarisce appunto con i due lavori di Gödel (tesi del 29 e TG1 del 30), che mettono nella giusta luce anche il paradosso di Löwenheim-Skolem.

Da questi segue che i risultati di categoricità visti richiedono tutti una logica diversa da quella del primo ordine, esplicitata nei Principia Mathematica e studiata dalla scuola di Hilbert.

Oggi sappiamo infatti che gran parte dei risultati di categoricità significativi richiedono la cosiddetta logica del secondo ordine, per la quale non vale un risultato di completezza analogo a quello provato da Gödel nel 1929 per la logica del primo ordine: l'insieme degli enunciati validi non è ricorsivamente enumerabile, perde cioè una delle caratteristiche tipiche dei sistemi deduttivi (al primo ordine). Non solo: ma la logica del secondo ordine non soddisfa oltre al teorema di Löwenheim-Skolem nemmeno il teorema di compattezza. La logica del primo ordine, per la quale valgono tutti i teoremi piacevoli (come la enumerabilità ricorsiva dei teoremi, il teorema di compattezza nonché il teorema di Löwenheim-Skolem), non è più categorica e risulta indecidibile.

Il lavoro di Gödel mette in luce un doppio piano, rispetto al quale leggere i risultati sui fondamenti del primo trentennio del secolo: è come se si accendesse una lampada su di una scena che prima sembrava piattamente bidimensionale e che ora con la nuova luce appare avere una dimensione in più.

2. Considerazioni euristiche

Un avvicinamento euristico al TG1 è quello proposto da Smullyan: si tratta di un'elaborazione del paradosso del bugiardo (che così come è noto, risulta invece più vicino al teorema di Tarski). Un ulteriore approccio sarà accennato in un capitolo successivo.

Gli indovinelli di Smullyan sono una palestra formidabile per introdurre la logica (si vedano per esempio i volumi "Qual è il titolo di questo libro?", "Donna o tigre", presso Zanichelli) e presentano un indubbio interesse anche per un eventuale avvicinamento ai teoremi di Gödel. L'isola dei furfanti e dei cavalieri potrebbe costituire un ingrediente fondamentale in ogni approccio didattico ai problemi dell'incompletezza e alla logica in generale. Il discorso dovrebbe essere intrecciato con altre situazioni, ad esempio i paradossi e l'autoreferenzialità in contesti più intuitivi (macchine, geometria, punti fissi). Qui mi limiterò ad introdurre il tema dell'incompletezza sulla falsariga degli indovinelli di Smullyan: per molti più dettagli ed esempi, si legga il volumetto di Smullyan, *Forever undecided: A Puzzle Guide to Gödel, Alfred A. Knopf, 1987*, oppure il cap. XI del suo bellissimo trattato *Gödel's Incompleteness Theorems, Oxford University Press, 1992*.

Veniamo ora ad un primo esempio. Immaginiamo un'isola, chiamiamola Gödelandia, abitata da furfanti (che dicono sempre il falso) e da cavalieri (che dicono sempre il vero). Alcuni abitanti dell'isola sono ateniesi, altri sono cretesi. I primi sono tutti cavalieri i secondi tutti furfanti; esistono (eventualmente) furfanti e cavalieri provenienti da altre parti del mondo. Risolviamo il seguente

PROBLEMA 1 Quale frase può dire un abitante di Gödelandia per convincerti che è un cavaliere ma non è ateniese?

[Soluzione. Il nativo può dire: "Io non sono un ateniese". Infatti un furfante non può dire la frase, perché sarebbe la verità. Segue che chi parla è un cavaliere e dunque non è ateniese in quanto le sue affermazioni sono vere].

Nostro obiettivo è di sviluppare l'analogia di questa situazione, passando dall'isola dei furfanti e dei cavalieri agli enunciati dimostrabili oppure veri in un sistema di assiomi (sistema formale). Il ruolo degli ateniesi sarà quello degli enunciati non solo veri ma anche provabili nel sistema. Il nostro cavaliere non ateniese personifica un esempio di enunciato vero ma non provabile. Il ruolo dei cretesi (poco interessante nell'aneddoto) è quello degli enunciati refutabili. Il TG1 prova appunto che gran parte dei sistemi formali è come Gödelandia: nei sistemi formali per l'aritmetica esiste l'analogo dei cavalieri non ateniesi.

Passiamo ora ad un altro esempio, più articolato. Esso presenta un qualche interesse per certi sistemi assiomatici delle logiche modali, nonché per gli studiosi di intelligenza artificiale, in quanto permette una riformulazione "epistemica" dei teoremi di incompletezza.

[Nota per il lettore: più volte nel testo, ho dovuto premettere questi avvertimenti in cui preannuncio "tradimenti" più o meno spinti dei teoremi di Gödel, per un avvicinamento didattico a questi. Non sono infatti riuscito a trovare situazioni di apprendimento che non snaturino, quale più, quale meno, l'essenza di tali teoremi. Nell'esposizione orale cercherò di illustrare questo punto con ulteriori considerazioni].

Il teorema di Gödel tratta di sistemi (formali) di assiomi e di enunciati provabili o meno nell'ambito di tali sistemi. In questo approccio euristico considereremo un logico (che quindi è maestro nei ragionamenti) e le proposizioni che lui crede vere. Mentre in un sistema formale si hanno dunque formule dimostrate (teoremi); qui considereremo le proposizioni credute (vere, in base a ragionamenti) dal nostro logico: la dimostrabilità nel sistema cede il passo alle credenze del logico; altra cosa da queste sono le proposizioni vere indipendentemente dalle credenze razionali del logico (vedremo che un'ipotesi consiste proprio nel supporre vere le credenze del logico, ovvero nel supporre che il logico creda solo in cose vere).

In tutte le situazioni si avrà a che fare con un indigeno dell'isola e con il logico.

SITUAZIONE 1

Supponiamo che le credenze del logico non siano mai false: egli è cioè perfettamente accurato nei suoi ragionamenti.

PROBLEMA 2 Il nativo dice una frase X. Il risultato è che il logico non crederà mai che il nativo sia un cavaliere né che egli sia un furfante. Esplicitare un enunciato opportuno per X.

[Soluzione. Una soluzione è: X = "Non crederai mai che io sia un cavaliere"

Supponiamo che il nativo sia un furfante. Dunque X è falso. Quindi il logico giungerà a credere che il nativo è un cavaliere. Ciò è assurdo, in quanto il logico non crede mai proposizioni false.

Segue che il nativo è un cavaliere.

Siccome dice il vero, il logico non crederà mai che lui è un cavaliere.

D'altra parte il logico non crederà mai proposizioni false, quindi non crederà mai che il nativo sia un furfante.

Segue che il logico rimarrà indeciso per sempre sulla natura del nativo].

Notazioni. Useremo la lettera c per rappresentare la proposizione "il nativo è un cavaliere"; useremo la lettera L _ per rappresentare il predicato "il logico crederà _ , prima o poi". Ad esempio, Lc sta per "il logico crederà prima o poi che il nativo sia un cavaliere" (crederà prima o poi può volere dire anche che già lo sta credendo).

Si osservi che se il nativo asserisce la proposizione p, essa risulta logicamente equivalente alla proposizione c: $p \leftrightarrow c$. Infatti se il nativo è un cavaliere p è vera, mentre se è

un furfante (cioè vale ¬c) p è falsa. Chiameremo questo principio che regola la vita in Gödelandia: legge dell'isola.

L'ipotesi di correttezza è traducibile così: Lp → p (per ogni proposizione p); cioè, se il logico crederà p prima o poi, p è vera.

Discussione "formalizzata" della soluzione alla situazione 1.

1. Il nativo asserisce ¬Lc; quindi vale: c ↔ ¬Lc. Abbiamo visto che ne segue sia ¬Lc sia ¬L(¬c). Infatti: ¬c → Lc, Lc → c, assurdo, dunque c; segue ¬Lc. Inoltre: se fosse L(¬c), seguirebbe ¬c; ma vale c; dunque ¬L(¬c).

Ciò vale per ogni proposizione p: se vale p ↔ ¬Lp e se il logico non crede mai proposizioni false, segue ¬Lp e ¬L(¬p) e dunque il logico non crederà mai né p né ¬p, mentre p risulta vera (equivale all'enunciato vero ¬Lp).

SITUAZIONE 2

Nella discussione fatta nella precedente situazione, non occorre che davvero i cavalieri dicano il vero e che i furfanti dicano il falso. Ciò che importa è che il logico creda che ciò si verifichi, cioè creda che valga la legge dell'isola.

Esplicitiamo le ipotesi sulle credenze del logico necessarie per condurre il ragionamento fatto nella situazione 1:

1. *Credenze del logico in virtù di assiomi e regole logiche.*

Supponiamo che il logico possieda in atto un sistema di regole logiche corrette e complete (calcolo dei predicati del primo ordine). Esempi di formule credute dal logico: le tautologie, i teoremi logici; il mondo delle sue credenze sarà chiuso per modus ponens (se crede p e p→q, allora crede anche q); inoltre, nel caso deduca un enunciato q usando la premessa p, allora egli crede anche p→q (teor. di deduzione). Chiameremo rispettabile un logico che soddisfi alla proprietà 1. Ma la rispettabilità è troppo poco; occorrono altre virtù.

In primo luogo il logico deve avere una sua ipotesi sulla natura del mondo in cui è sbarcato, per potervi ragionare su. Si tratta di ipotesi extra-logiche; enunciamo quindi come punto

2. *La legge dell'isola: il logico crede che valga la seguente legge:* c ↔ ¬Lc.

Chiameremo avveduto un logico che crede nella legge dell'isola.

3. *Credenze del logico in virtù di principi ad hoc: normalità e stabilità.*

Infine, facciamo le due seguenti ipotesi sulle regole cui soddisfano le credenze del logico. Si tratta di ipotesi ragionevoli, anche se discutibili (pensate a che cosa ne potrebbe tirare fuori un Pirandello):

Supponiamo che il logico sia normale: se crede una proposizione p, allora crede di crederla. In formule: Lp → LLp. L'inverso della normalità costituisce ciò che chiamiamo stabilità: LLp→Lp.

Diremo infine che il logico è incoerente se crede sia p che ¬p, per almeno una certa proposizione p. Un logico incoerente crede ogni proposizione (cfr. il teorema logico che corrisponde al principio "ex absurdo quod libet"): è un credulone! (si osservi che però un logico incoerente è automaticamente normale e stabile, anche se credulone; per fortuna non vale il viceversa). Diremo invece che il logico è coerente se non è incoerente (cioè non crede a tutto).

PROBLEMA 3 Supponiamo che un logico rispettabile, avveduto e normale visiti la nostra isola; incontra un indigeno che gli dice: "*Tu non crederai mai che io sono un cavaliere*".
Provare che: a) Se il logico crederà mai che il nativo è un cavaliere allora egli è incoerente.
b) Se il logico crederà mai che il nativo è un furfante egli è incoerente o non stabile.
Dunque, se il logico è anche coerente e stabile, non deciderà mai se l'indigeno è furfante o cavaliere.

[Soluzione. Il nativo ha asserito ¬Lc.

a. Supponiamo che il logico creda c a un certo punto; siccome crede c → ¬Lc (legge dell'isola), usando il modus ponens crederà anche ¬Lc. D'altra parte, il logico crede c, dunque essendo normale crede anche Lc. Segue la sua incoerenza.

b. Supponiamo che il logico creda invece ¬c a un certo punto. Allora egli crederà anche Lc, in virtù dell'ipotesi sulla legge dell'isola e delle seguenti equivalenze logiche: [c ↔ ¬Lc] ↔ [¬c ↔ Lc]. Se il logico è stabile, dalla credenza di Lc passerà alla credenza diretta di c. Dunque egli risulta incoerente, in quanto crederà sia ¬c che c].

Le situazioni 1 e 2 introducono a due forme del primo teorema di Gödel (TG1), la prima di tipo semantico (infatti fa riferimento alla correttezza del nostro logico), mentre la seconda rimane chiusa nell'ambito delle credenze del logico e la nozione di verità non è più necessaria.

Noi spingeremo il più possibile l'analogia tra Gödelandia e i sistemi formali nell'avvicinamento al TG1: le proposizioni analizzate nei nostri indovinelli diventano gli enunciati dei sistemi di assiomi per l'aritmetica che considereremo; la loro verità e falsità non riguarderà più furfanti e cavalieri ma le proprietà dei numeri (alcune false altre vere); le credenze del logico, conseguenza dei suoi ragionamenti (condotti usando le regole logiche a partire dalla legge dell'isola, assunta come unico assioma non logico), saranno rimpiazzate dai teoremi dei sistemi considerati, conseguenza delle dimostrazioni fatte nel sistema stesso (a partire dai suoi assiomi non logici usando le opportune regole logiche). Se

riusciremo a scoprire anche per i nostri sistemi formali un analogo della legge dell'isola, allora avremo la possibilità di spingere l'analogia fino alle estreme conseguenze, cioè provare l'esistenza di enunciati aritmetici veri ma non dimostrabili né refutabili nel sistema formale considerato. Si noti che occorrerà trovare l'analogo della proprietà $c \leftrightarrow \neg Lc$, cioè una proposizione come c, che risulti equivalente alla di lei indimostrabilità (essendo la nozione di credenza rimpiazzata da quella di dimostrabilità). Ciò sarà fatto nel prossimo paragrafo.

3. Assiomi per l'aritmetica

Considereremo ora un sistema di assiomi per l'aritmetica, che indicheremo con **ASSAR**; si tratta di un sistema di assiomi adatto a parlare della struttura

$$\mathbf{N} = (\omega; \leq, ', +, \times, 0)$$

cioè del monoide ordinato dei numeri naturali, dotati della funzione "successivo di" e delle operazioni di somma e prodotto, con l'elemento 'distinto' zero. Esso si suppone formulato in un linguaggio con i seguenti simboli non logici (oltre al simbolo di uguaglianza =):

(i) variabili numeriche: x, x', x",.....(useremo liberamente le lettere x, y, z);
(ii) una costante numerica: **0** (nome del numero 0);
(iii) funzioni: ' (successivo di), + (somma), · (prodotto);
(iv) la relazione: < .

Si è detto che è nostro obiettivo estendere il ragionamento euristico fatto nel paragrafo precedente per ottenere in modo via via più rigoroso il teorema di Gödel. In questo lavoro di avvicinamento cercheremo anche di capire progressivamente quali proprietà dovrà soddisfare il nostro sistema di assiomi, per potere sviluppare in forma precisa al suo interno l'analogo del ragionamento del logico sbarcato sull'isola dei furfanti e dei cavalieri.

Ricordiamo: si è provato che egli non avrebbe mai potuto decidere la natura dell'indigeno, sempre che fosse un logico rispettabile, avveduto, normale, coerente e stabile (cfr. situazione 2). Tutto dipendeva dalla legge dell'isola: $c \leftrightarrow \neg Lc$ (base delle credenze dei logici avveduti).

Ora è nostro obiettivo sviluppare il ragionamento sostituendo il predicato epistemico L_ che riguarda le credenze del logico con quello di dimostrabilità nel sistema **ASSAR**, che indicheremo con Prov(_). Ad esempio se F è un assioma del nostro sistema è chiaro che vale Prov(F). La cosa vale anche per tutti i teoremi di **ASSAR**: ad esempio, se Prov(F) e $Prov(F \rightarrow G)$ segue che anche Prov(G).

Occorrerà quindi trovare un enunciato G tale che $G \leftrightarrow \neg Prov(G)$, cioè un enunciato del linguaggio di **ASSAR** che asserisca la propria indimostrabilità in **ASSAR**. Questo è il

problema di fondo, molto difficile. Occorrerà esprimere nel linguaggio di **ASSAR** enunciati come Prov(_), cioè enunciati che non riguardano solo le proprietà matematiche dei numeri ma proprietà metamatematiche, sul sistema formale stesso.

Supposto di avere fatto ciò, rimane da chiarire che cos'è l'analogo del logico rispettabile, avveduto, coerente, normale e stabile per un sistema di assiomi. Questo è un problema importante, ma meno difficile, una volta che sia risolto il problema di fondo.

Ragioniamo:

- che un sistema di assiomi sia rispettabile è molto facile da tradurre: significa che **ASSAR** è formulato con un calcolo logico del primo ordine, cioè **ASSAR** contiene gli assiomi logici del calcolo dei predicati del primo ordine;
- che sia avveduto significa che esso assume alcune ipotesi ragionevoli sul mondo che considera (assiomi non logici sui numeri);
- coerente significa ipotizzare che esiste una formula F di **ASSAR** tale che né lei né la sua negazione ¬F sono teoremi di **ASSAR**;
- la normalità significa che vale l'implicazione $Prov(E) \rightarrow Prov(Prov(E))$; cioè, se E è di fatto un teorema di **ASSAR**, ciò è provabile nel sistema stesso: la normalità esprime quindi una forma di completezza del sistema **ASSAR** rispetto alle formule del tipo Prov(E);
- che sia stabile significa che vale la relazione inversa della normalità $Prov(Prov(E)) \rightarrow Prov(E)$; è una forma di riflessione rispetto agli enunciati del tipo Prov(E): se si prova formalmente in **ASSAR** la provabilità di un enunciato E, allora tale enunciato risulta di fatto provabile.

Per ottenere un sistema che goda di tutte le proprietà in elenco, si rivelerà ampiamente sufficiente il seguente sistema di assiomi, denotato da **N**, i cui assiomi non logici sono:

N1. $Sx \neq \mathbf{0}$.
N2. $Sx = Sy \rightarrow x = y$.
N3. $x + \mathbf{0} = x$.
N4. $x + Sy = S(x + y)$.
N5. $x \cdot \mathbf{0} = \mathbf{0}$.
N6. $x \cdot Sy = (x \cdot y) + x$.
N7. $\neg(x < \mathbf{0})$.
N8. $(x < Sy) \leftrightarrow (x < y) \vee (x = y)$.
N9. $(x < y) \vee (x = y) \vee (y < x)$.

Gli assiomi **N1-N2** asseriscono che S è una funzione 1-1, il cui codominio non contiene lo zero.

Gli assiomi **N3-N4** sono le equazioni che definiscono la somma per ricorrenza.

Gli assiomi **N5-N6** sono le equazioni che definiscono il prodotto per ricorrenza.

Gli assiomi **N7-N9** riguardano le proprietà fondamentali della relazione "essere minore di".

Con tali assiomi si riescono a dimostrare le proprietà numeriche necessarie per condurre in porto l'analogia tra Gödelandia ed enunciati aritmetici. Il sistema è preso dal libro dello Shoenfield, *Mathematical Logic*, p.22 ed è una variante di un sistema molto famoso di assiomi per l'aritmetica, noto come Aritmetica di Raphael Robinson; il suo maggiore pregio è di essere costituito da un sistema finito di assiomi. Il sistema in questione è interessante in quanto aggiungendo un unico schema di assiomi, cioè l'induzione matematica, si ottiene un sistema molto più potente di aritmetica, cioè l'aritmetica di Peano **PA** al primo ordine; **PA** è il sistema **N1...N9 + $\mathbf{Ind_A}$**, dove è il seguente schema di assiomi (uno per ogni formula **A** del linguaggio dell'aritmetica sopra schizzato)

$$\mathbf{Ind_A}: \quad \mathbf{A[0]} \,\&\, \forall x(\mathbf{A} \rightarrow \mathbf{A[Sx]}) \rightarrow \mathbf{A}$$

In **PA** si dimostrano tutti gli usuali teoremi propri della teoria elementare dei numeri, nonché moltissimi risultati non elementari di teoria dei numeri.

4. Il primo teorema di incompletezza di Gödel: una dimostrazione semantica

La versione qui esposta del TG1 è antistorica, in quanto basata sulla nozione semantica di verità di una formula, che non compare nella trattazione originale di Gödel. Essa è fatta col 'senno del poi', in quanto tiene conto anche delle riflessioni di Tarski, ispirate da Gödel stesso, ma successive al 1931, che portarono il logico polacco a dimostrare il suo celebre teorema sulla non rappresentabilità dell'insieme degli enunciati veri dell'aritmetica (si veda la discussione in Lolli, p.30). Essa ha però il pregio di apparire più intuitiva e quindi didatticamente più efficace, perlomeno in un primo approccio al problema. Inoltre presenta la risoluzione al problema in modo che la dimostrazione sintattica (che qui non vedremo) appare come un affinamento delle tecniche usate, ma l'idea rimane sostanzialmente la stessa. Essa corrisponde alle situazione 1 dell'isola dei furfanti e dei cavalieri, che abbiamo visto nel §2, e si basa su tre ingredienti, necessari per risolvere sia il problema di fondo, sia quelli collegati, discussi nel §2; tali ingredienti permetteranno di tradurre rigorosamente le situazioni di Gödelandia ai nostri sistemi formali per l'aritmetica:

a) la nozione di esprimibilità di un insieme (o relazione) in un sistema formale, ad esempio **PA**;
b) l'aritmetizzazione delle formule;
c) la diagonalizzazione.

Quanto diremo vale per molti sistemi formali, oltre **N**, **PA**, anche se concretamente ci riferiremo solo ad uno di questi, diciamo **PA**, tanto per fissare le idee.

Cominciamo con l'esprimibilità; un esempio chiarirà la questione: l'insieme dei numeri primi è esprimibile in **PA** dalla formula, diciamo P(x), che esplicita nel linguaggio dell'aritmetica la proprietà di essere primo. Tale formula risulterà vera per tutti e soli i numeri primi. E' possibile verificare che gran parte degli insiemi e delle relazioni riguardanti i numeri che si incontrano usualmente in matematica sono esprimibili in **PA**. Diamo perciò la

DEFINIZIONE Un insieme numerico A è *esprimibile* in un sistema formale, ad esempio nell'aritmetica di Peano, con la formula H(x) nel caso che, per ogni numero naturale n: H(n) è vero sse (se, e solo se) $n \in A$.

La definizione è estendibile anche alle relazioni a più argomenti. Si noti che non tutti gli insiemi numerici sono esprimibili in **PA** (né in un qualunque altro sistema formale). Infatti **PA** dispone solo di un'infinità numerabile di formule, mentre l'insieme dei sottoinsiemi di ω ha cardinalità più che numerabile, per il teorema di Cantor.

Il secondo punto è l'aritmetizzazione. Per non rendere troppo macchinosa l'esposizione, ci limiteremo a supporre che esista una biiezione g tra l'insieme delle formule del sistema formale e i numeri naturali; se E è una formula, g(E) è detto il suo numero di Gödel. Scriveremo anche E_n per indicare la formula il cui numero di Gödel è n.

Il terzo punto è la diagonalizzazione. Espliciteremo unicamente il punto saliente sorvolando su alcuni aspetti tecnici, irrilevanti nella sostanza, ma che rischiano di appesantire troppo ad una prima lettura.

Usiamo la parola *numerale* per indicare il termine canonico ed ufficiale con cui rappresentiamo i numeri nel linguaggio di PA: la costante **0**, i termini **0'**, **0''**, ..., sono numerali (che rappresentano i numeri 0, 1, 2,...) ; per i numerali useremo lettere in grassetto.

Supponiamo che la formula E_n abbia un'unica variabile libera v: sostituiamo alla variabile v nella formula il numerale **n**; otteniamo un enunciato (senza variabili libere): ad esempio se la formula $(\exists v')(v < v')$ ha un certo numero di Gödel, diciamo c, la sostituzione produrrà $(\exists v')(\mathbf{c} < v')$. Tale formula avrà un certo numero di Gödel, che dipende ovviamente dal numero n della formula di partenza; lo indicheremo con d(n).

La funzione d associa ad ogni numero n (ossia, ad ogni formula E_n) il numero della formula che si ottiene con la sostituzione vista (qui sorvolo sul caso in cui E_n non abbia esattamente una variabile libera: comunque le cose si aggiustano facilmente, come vedremo).

Il significato intuitivo della diagonalizzazione è così esprimibile: sia P(v) la formula con una variabile libera (predicato) costruita nell'esercizio per esprimere nell'aritmetica

l'insieme dei numeri primi; sia p il suo numero di Gödel d(p) è il numero di Gödel della formula P(**p**): essa risulta vera sse p è primo. Analogamente, ¬P(v) indicherà i numeri composti, avrà un certo numero, sia q: è facile pensare a ¬P(**q**), che risulta vera sse q è composto.

Supponiamo ora di considerare sistemi formali corretti: significa che tutti gli enunciati (formule chiuse) dimostrabili nel sistema sono veri, mentre quelli refutabili (cioè quelli la cui negazione è dimostrabile) sono falsi (cioè non veri). Tutti i sistemi assiomatici per l'aritmetica esistenti sul mercato sono corretti; esempi: **PA**, **N**.

Ci avvicineremo a TG1 con la fig.1: nel disegno, in alto è illustrato il mondo delle rappresentazioni numeriche: lì vivono i numeri che aritmetizzano (rappresentano, sono il numero di Gödel del-)le componenti sintattiche dei nostri sistemi formali. In basso sono illustrate le formule vere e proprie; la linea orizzontale può essere pensata come uno specchio che fa passare dal mondo 'reale' delle formule a quello delle loro rappresentazioni come numeri, tramite la funzione g. Per comodità, in alto si considerano due copie dello stesso oggetto: tra le due copie opera la funzione d. In effetti il teorema di Gödel è basato su un continuo gioco di specchi: useremo la metafora "al di qua, al di là dello specchio" per non ripetere continuamente la pesante distinzione tra enti sintattici e loro gödeliani.

Il disegno illustra (tra l'altro) la diagonalizzazione di una formula E_n: gli ovali di sinistra illustrano (al di qua e al di là dello specchio) l'insieme *P* delle formule dimostrabili nel sistema (i teoremi). L'ovale rigato rappresenta l'insieme *P** degli enunciati che si ottengono diagonalizzando *P*: si tratta cioè degli enunciati $E_n(\mathbf{n})$ che risultano dimostrabili.

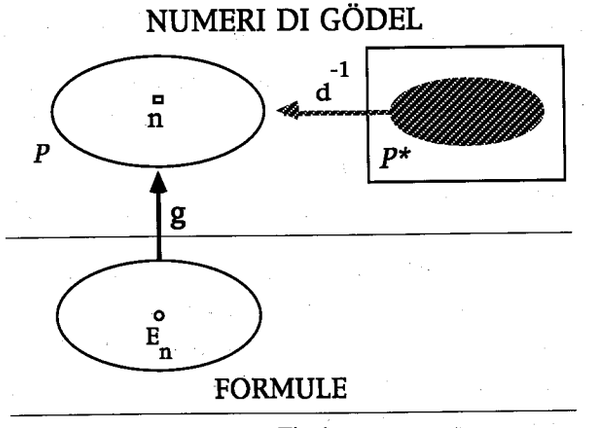


Fig.1

L'insieme *P* è semplice da immaginare: si tratta dei teoremi al di là dello specchio; l'insieme *P** è più barocco: bisogna considerare solo gli enunciati diagonali, come quello che riguarda i numeri primi visto poco fa e tra questi considerare quelli che sono teoremi. Ad esempio, siccome usualmente ¬ P(**q**) è un teorema, lo troveremo al di là dello specchio in *P**. Invece, siccome P(**p**) non è un teorema (nelle usuali aritmetizzazioni), esso non sta in *P**. In altre parole, *P** è l'insieme degli enunciati diagonali che sono teoremi.

Formuliamo ora le seguenti ipotesi sul nostro sistema formale, ad esempio **PA**:

1. Supponiamo che sia corretto: significa che l'insieme *P* è contenuto nell'insieme *V* degli enunciati veri del sistema. Ne segue che anche *P** è contenuto in *V*:
2. Supponiamo che l'insieme ¬*P** (cioè l'insieme complemento di *P** in N) sia esprimibile; esiste quindi una formula H(v) tale che, per ogni numero naturale n, H(**n**) risulta vero sse n non appartiene all'insieme *P**.

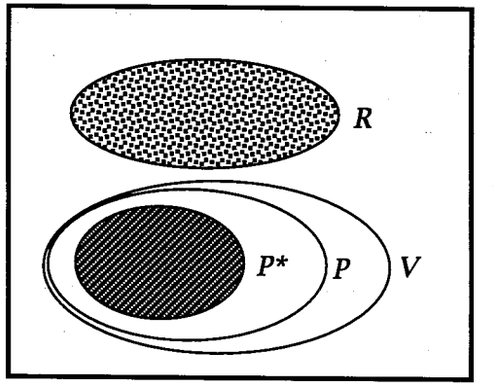


Fig.2

Quindi H(**n**) risulta vero sse $E_n(\mathbf{n})$ non è dimostrabile. Ora H(v) è una formula del linguaggio del nostro sistema formale, e.g. è una formula ufficiale di **PA**. Vive cioè nel nostro mondo al di qua dello specchio. Ha quindi un suo numero diciamo h, al di là dello specchio.

Possiamo ora diagonalizzare la formula H(v): si ottiene l'enunciato H(**h**); si noti che è H(v) = E_h e quindi $E_h(\mathbf{h})$ = H(**h**). Intuitivamente, H(**n**) significa: "$E_n(\mathbf{n})$ non è provabile", cioè l'ennesimo enunciato, diagonalizzato, non è provabile, mentre H(**h**) asserisce esattamente: "l' h-simo enunciato, diagonalizzato, non è provabile", vale a dire "H(**h**) non è provabile". In altre parole H(**h**) asserisce la propria indimostrabilità.

In forma più rigorosa, possiamo asserire anche per questo enunciato quanto sappiamo degli enunciati ottenuti per diagonalizzazione:

H(**h**) è vero sse H(**h**) non è dimostrabile.

Da un punto di vista logico risulta che l'equivalenza vale esattamente in due casi:

(i) H(**h**) è vero e H(**h**) non è dimostrabile, oppure
(ii) H(**h**) non è vero e H(**h**) è dimostrabile.

Il secondo caso non può sussistere in quanto se H(**h**) è dimostrabile, per la correttezza del sistema (ipotesi 1) risulta anche vero, il che non è. Rimane solo il caso (i): cioè H(**h**) è un esempio concreto di enunciato vero non dimostrabile. In onore di Gödel, indichiamolo con G. Risulta anche che la negazione di G non è dimostrabile; se lo fosse, per la correttezza del sistema, G sarebbe non vero, il che non è. Abbiamo quindi dimostrato in forma semantica il

TEOREMA DI INCOMPLETEZZA DI GÖDEL Se **ASSAR** è un sistema formale corretto, in cui $\neg P^*$ è esprimibile, esiste un enunciato G di **ASSAR** che è vero, ma né G né la sua negazione sono dimostrabili in **ASSAR**.

Nota Il teorema è enunciato in forma generale, sorvolando sulle specifiche caratteristiche di **ASSAR**, in quanto sistema formale. Esplicitandole, occorre che **ASSAR** sia un linguaggio del primo ordine, in cui abbia senso parlare del seguente elenco di cose (ognuno pensi al significato delle parole in elenco nel caso dei sistemi visti per l'aritmetica e immagini di generalizzare la nozione):

1. Un insieme E di Espressioni, in quantità numerabile;
2. Un insieme S di Enunciati, $E \supset S$;
3. Un insieme P di enunciati provabili, $S \supset P$;
4. Un insieme R di enunciati refutabili, $S \supset R$;
5. Un insieme H di predicati, $E \supset H$ (intuitivamente, un predicato è il nome di un insieme di numeri naturali);
6. Un nome **n** (detto numerale) per ogni naturale n;
7. Una funzione che assegni ad ogni espressione E e ad ogni numero naturale n un'espressione, indicata con E(**n**); nel caso che E sia un predicato H, H(n) risulta un enunciato (intuitivamente, H(n) esprime il fatto che il numero n appartiene all'insieme nominato da H).
8. Un insieme V di enunciati veri, $S \supset V$.

Nei sistemi per l'aritmetica che abbiamo considerato, le cose in elenco hanno un significato preciso; è comunque possibile definirle in generale per sistemi formali qualsiasi e conseguentemente enunciare il teorema di Gödel in forma molto generale.

Osservazioni

1. Delle due ipotesi fatte (correttezza, esprimibilità di $\neg P^*$), la prima è facile da provare per sistemi come **PA**, la seconda è molto più ardua da dimostrare. E' proprio l'aritmetizzazione a garantirne la possibilità.
2. L'argomento sviluppato per P e $\neg P^*$ vale in realtà per ogni insieme A, tale che A* è esprimibile nel sistema. Si dimostra cioè che in tal caso esiste un enunciato $E_h(\mathbf{h})$ tale che l'enunciato è vero sse $h \in A$ (se l'enunciato è vero, h appartiene ad A, mentre non vi appartiene se l'enunciato è falso). L'osservazione è dovuta a Carnap (1934).
3. Come conseguenza del punto 2, risulta che l'insieme V degli enunciati veri dell'aritmetica non è esprimibile in **PA**, tanto meno in **N** (teorema di Tarski, 1934; in realtà già Gödel si era accorto del risultato nell'autunno del 1930). Infatti in tal caso si dimostra che sarebbero esprimibili anche $\neg V$ e $\neg V^*$; ma allora esisterebbe un enunciato $E_h(\mathbf{h})$ tale che l'enunciato è vero sse $h \notin V$; cioè è vero sse è falso, il che è assurdo. Il teorema di Tarski marca in modo drammatico la differenza tra provabilità e verità. Mentre la nozione di provabilità è esprimibile negli usuali sistemi logici, non così è per la verità: l'insieme degli enunciati aritmetici veri (nel modello usuale dei numeri naturali) va oltre le capacità espressive dell'aritmetica.

5. Il teorema di Gödel-Chaitin

Negli ultimi decenni, il teorema di Gödel ha ricevuto una nuova e affascinante interpretazione nell'ambito della cosiddetta teoria della complessità algoritmica, sviluppata soprattutto grazie ai lavori di due studiosi, che lavoravano ai poli opposti del mondo di allora, Andrei Kolmogorov, accademico in URSS, e Gregory Chaitin, della Divisione Ricerca dell'IBM in USA.

(Per un'introduzione, si legga la sezione intitolata Ordine e Caso in: *Numeri, Caso e Sequenze*, a cura di R. Magari, Quaderno n. 45, Le Scienze 1982).

Essi si proposero di analizzare matematicamente la differenza tra successioni generate seguendo una regola, come la (1) qui sotto e le cosiddette successioni senza regola (derivanti dalla nozione di collettivo di von Mises), come la (2),

(1) 10101010101010101010101010101010101010

(2) 10000100111011011001111000100011101101 00

La loro idea fu molto semplice, ma profonda: una successione è senza regole se è incompressibile da un punto di vista algoritmico, se cioè non si può descrivere con un metodo più compatto che non sia il semplice elenco dei suoi termini. In altre parole, una successione come la (1) è regolare in quanto posso scrivere un breve programma per un

computer, che genererà tutti i suoi termini e la lunghezza del programma sarà decisamente più breve che non l'elenco dei termini della successione (la cosa sarebbe più appariscente se le successioni (1) e (2) contenessero 1000 termini). La successione (2) invece è senza regola, in quanto l'unico programma sensato per farla generare da un computer consiste nel dare al computer l'elenco stesso dei suoi termini: non esiste nulla di più compatto.

La loro teoria fu messa a punto negli anni settanta; in tal modo, nozioni importanti, proprie della teoria dell'informazione, furono estese in forma nuova e suggestiva ai concetti matematici analizzati dal nuovo punto di vista della complessità algoritmica.

Secondo la teoria, Kolmogorov e Chaitin definiscono come "casuali" quelle successioni per le quali non esistono algoritmi che le generino, la cui lunghezza sia decisamente più corta rispetto alla sequenza stessa. Quindi la loro nozione di "successione casuale" coincide in realtà con quella di "sequenza generata senza alcuna regola", ovvero di "sequenza disordinata" (nel caso di sequenze infinite occorre una modifica tecnica importante su cui non possiamo qui entrare): è però concettualmente preferibile tenere distinti i due concetti di "casuale" e di "disordinato"; quindi preferiamo chiamare "senza regola" (SR) tali successioni.

Le successioni SR godono di interessanti proprietà.

Ad esempio, esse sono molto frequenti, relativamente a quelle generate secondo una regola: un semplice conteggio mostra che già il 99.999 % delle successioni di 0 e 1, lunghe 20, sono SR. La regolarità è cioè un'eccezione nel mondo delle successioni aritmetiche. Un'altra proprietà interessante delle successioni SR è che esse superano tutti i principali test di casualità elaborati dagli specialisti di statistica. In altre parole, anche se differenti dalle successioni casuali (ad esempio quali risultano dai lanci di una moneta) sul piano teorico, le successioni SR risultano indistinguibili da quelle casuali da un punto di vista esterno e comunque a livello concreto.

La teoria della complessità algoritmica ha messo in luce interessanti connessioni con il teorema di Gödel. Il risultato di Gödel, alla luce della definizione di successione SR data da Kolmogorov e Chaitin, asserisce che tipici esempi di queste proprietà indimostrabili sono forniti dalle successioni SR: pur essendocene moltissime (come abbiamo visto, gran parte delle successioni sono SR), e quindi pur essendo quasi certo che se si prende una successione generica di 100 cifre, questa risulta SR, ciononostante non è possibile (tranne che per poche eccezioni) dare una dimostrazione formale del fatto che essa sia senza regola. E' un fatto vero che gran parte delle successioni siano SR e si comportino casualmente, ma proprio per questo, tale verità risulta inattingibile dai nostri strumenti dimostrativi.

Vediamo di intuire un pochino il significato di questo risultato, che offre un'altra via di attacco al teorema di Gödel. Il punto di partenza può essere un celebre paradosso, attribuito a Berry, bibliotecario alla biblioteca Bodleiana ai tempi di Russell. Come è noto, il

mondo dei paradossi è un'utile palestra per acquisire metafore e intuizioni relative a profonde verità matematiche, in particolare al teorema di Gödel (per una discussione più approfondita sui paradossi si legga il cap. XI del volume di Lolli citato).

Tradizionalmente, i paradossi interessanti nello studio dei fondamenti della matematica sono tre :

1. Paradosso di RUSSELL: *Sia R l'insieme di tutti gli insiemi che non contengono se stessi come elemento.* R risulta contraddittorio, in quanto se contiene se stesso significa che soddisfa la proprietà che lo caratterizza, cioè non contiene se stesso come elemento. E viceversa.

2. Paradosso di EPIMENIDE cretese: Si consideri l'enunciato F: *"questo enunciato è falso"*. F non può essere né vero né falso: se fosse vero allora sarebbe falso. E viceversa.

3. Paradosso di BERRY: *Sia N il più piccolo numero naturale che per essere descritto in italiano richiede più di diciotto parole.* [N.B.: l'enunciato in corsivo è esattamente di 18 parole]. L'enunciato non definisce alcun numero, in quanto se così fosse, il numero così definito, in quanto definito esattamente con 18 parole (cioè quelle del nostro enunciato), non cadrebbe sotto l'enunciato stesso.

Ora, i primi due paradossi sono buone metafore per avvicinarsi al teorema di Gödel, in quanto mettono in luce i problemi autoreferenziali tipici della dimostrazione di tale teorema; il paradosso di Berry invece ha una natura diversa, di carattere più marcatamente linguistico (osservazione già fatta a suo tempo da Peano) ed è legato in modo forse più trasparente ai problemi della complessità delle successioni numeriche discussi in questo paragrafo, oltre che al teorema di Gödel.

Vediamo meglio. Identifichiamo al solito numeri e successioni finite di numeri: il paradosso di Berry mette in forse la definibiltà di successioni SR. Infatti concentriamoci sulle successioni lunghe 98; è chiaro che una successione lunga 98 e che per essere descritta richieda più di 98 segni è SR secondo la definizione di Kolmogorov e Chaitin. Ma il paradosso mette in dubbio che se ne possa considerare concretamente una! Analizzato per bene, esso esprime il fatto che non si può dimostrare per una sequenza SR che essa è tale.

Il risultato ammette eccezioni; per le sequenze non troppo lunghe, ciò può essere possibile: infatti, è chiaro che non si riesce a compattare l'enunciato del paradosso in una frase con meno di un certo numero di lettere (diciamo 50 lettere), ma da un certo punto in poi (oltre le 50 lettere, ad esempio) avremo a che fare solo con successioni lunghe e SR, per le quali non disporremo di una prova formale che sono tali: avremo il sospetto, l'evidenza sperimentale per la loro complessità; quel che volete ma non una dimostrazione matematica.

Vediamo di valutare la portata di tali risultati. Si può dire in prima approssimazione che l'analisi di successioni di dati costituisce il lavoro dello scienziato; inoltre egli si serve del linguaggio matematico per fare tale analisi. Lo scienziato va a caccia di regolarità, ma spesso

queste sono difficili da scoprire; a volte sembra proprio che non ci siano. Il punto è che se queste regolarità ci sono, allora non tutto è perduto: non è impossibile che con gli strumenti matematici a sua disposizione (o altri ancora che si forgerà via via) gli riesca di scoprire e provare matematicamente che tali regolarità sono soddisfatte dalle sequenze di dati frutto dei suoi esperimenti e delle sue osservazioni. Ma se per caso, tali regolarità non ci fossero o ci fossero solo in parte? Allora, la battaglia rischia di essere perduta in partenza perché il linguaggio matematico sarà probabilmente impotente a spiegarne la massima parte. C'è quasi la certezza che sia così: solo vincendo una scommessa incredibile, basata su scarsissime probabilità di successo (anzi con la quasi certezza di perdere), potrebbe riuscire a descrivere in forma matematica la complessità dell'universo.

Detto in altro modo: gran parte delle realtà matematiche (e forse fisiche) sono così complesse che non può esistere nemmeno una prova rigorosa di tale complessità; solo forme particolarmente efficienti di compattamento dell' informazione riescono a gettare una qualche luce su aspetti specifici di tale complessità. Se la spiegazione c'è, deve passare necessariamente attraverso la cruna di un ago, e non sarà mai globale e complessiva, ma frammentaria e specifica. Perlomeno il mondo dei numeri, frequentato dai matematici, ma su cui si modellano anche i fenomeni fisici, è fatto così; la regola è che esistono infinite verità indimostrabili (e molte non sono banali), ma d'altra parte esistono molte verità dimostrabili (e alcune non sono banali), la cui dimostrazione può essere ottenuta solo in modo estremamente creativo e significativo. Quale parte della matematica corrisponda al mondo fisico non sappiamo; si può essere pessimisti (e il freddo raziocinio dovrebbe spingerci a esserlo) e supporre che i modelli relativi al mondo fisico stiano tutti dalla parte complessa indimostrabilmente tale; oppure si può essere ottimisti e accettare la scommessa, fidando da un lato nell'intelligenza dell'uomo e dall'altro nel presupposto che la divinità sia sottile ma non maliziosa (forse era la posizione di Einstein): in tal caso si ammetterà che almeno una parte dell'universo sia modellata su quella parte dell'universo matematico su cui è possibile mettersi in gara per trovare le spiegazioni più compatte e ingegnose. L'impressione è che molti scienziati si comportino di fatto così.

Non vale più la speranza positivistica conclamata da Hilbert all'inizio del secolo, con la quale abbiamo iniziato la nostra chiacchierata: "in scienza non esistono ignorabimus"; ma non sarebbe nemmeno corretto sopravvalutare acriticamente i teoremi limitativi, tipo quelli di Gödel-Chaitin. L'ottimismo è molto meno diffuso; l'immagine della matematica oggi è molto meno monolitica di quanto non fosse un secolo fa, prima dei teoremi limitativi. Ma è anche meno compatta di quanto non pensassero i bourbakisti intorno agli anni cinquanta. L'immagine non è più quella di un grande edificio collegato da comode e potenti vie di accesso, ma di tante vette scoscese ed aguzze, ognuna diversa dalle altre e con le sue proprie

vie di accesso. Ognuna va costruita con fatica e con metodi specifici: questo è forse il senso del teorema di Gödel alle soglie del terzo millennio.

Bibliografia

LOLLI G., *Incompletezza*, Saggio su Kurt Gödel, il Mulino, Bologna, 1992.

MAGARI R. (curatore), *Numeri, Caso e Sequenze*, Le Scienze, Quaderno n. 45, 1982.

SMORINSKY C., *The incompleteness theorems*, In: Barwise J. (ed.), Handbook of Mathematical Logic, North Holland, Amsterdam, 1977, p.821.

SHOENFIELD J.R., *Mathematical Logic*, Addison Wesley, 1967; trad. it. presso Boringhieri, Torino.

SMULLYAN R. M., *Qual è il titolo di questo libro?*, Zanichelli, Bologna.

SMULLYAN R. M., *Donna o tigre?*, Zanichelli, Bologna.

SMULLYAN R. M., *Forever undecided: A Puzzle Guide to Gödel*, Alfred A.Knopf, New York, 1987.

SMULLYAN R. M., *Gödel's Incompleteness Theorems*, Oxford University Press, New York, Oxford, 1992.

Algoritmi e Calcolabilità La Tesi di Church Applicazioni della Logica all'Informatica

DANIELE MUNDICI
Dipartimento Scienze Informazione
Università Studi Milano
via Comelico 39-41
20135 Milano

PARTE I: ALGORITMI E CALCOLABILITÀ. LA TESI DI CHURCH

Anche se da molti secoli l'umanità sa costruire figure geometriche, solo nel 1801 Gauss trovò una condizione necessaria e sufficiente su n affinché l'n-gono regolare sia costruibile con riga e compasso. Per questo risultato non basta avere familiarità con le costruzioni geometriche: occorre invece una lenta evoluzione intellettuale che renda concepibile la stessa formulazione del problema di esistenza della costruzione.

Analogamente, benché fin dal principio fosse chiaro che funzioni come l'addizione e la moltiplicazione sono effettivamente calcolabili, solo a partire dagli anni trenta del ventesimo secolo fu fatto il salto concettuale necessario per circoscrivere la nozione di calcolabilità. In pochi decenni questa nozione ha avuto uno sviluppo rapido e si è concretizzata in strumenti di calcolo sempre più flessibili e potenti. Perché questa accelerazione ?

Per rispondere dobbiamo partire dalla logica. Infatti la nozione di calcolabilità è figlia del problema di come dedurre verità riposte da verità evidenti, mediante un unico e globale *calculus ratiocinator* che, invece di numeri, manipoli espressioni linguistiche formalizzate. La formalizzazione logica ha comportato un lavoro millenario di distillazione e selezione di parti del discorso, modi verbali, simboli, regole. Fonte ispiratrice è stata la geometria, attraverso l'esperienza acquisita nel raccogliere in assiomi e teoremi una congerie di osservazioni frammentarie sulle figure geometriche concrete.

Ogni insegnante sa quale crescita intellettuale intercorra dalla constatazione che un triangolo di lati 3, 4, 5 è rettangolo, alla comprensione del teorema di Pitagora. Altrettanta ne occorre per passare dall'osservazione che se Pico è uomo, visto che gli uomini sono mortali, anche Pico è mortale, alla comprensione dei teoremi logici fondamentali.

Frege e Hilbert

Benché il metodo assiomatico e la formalizzazione del ragionamento sillogistico risalgano entrambi ai Greci, e nonostante le anticipazioni e le intuizioni di Leibniz, Boole, Peano e altri, la prima costruzione di un linguaggio logico formale aderente ai moderni requisiti di espressività e precisione risale a Frege. Egli nella sua opera di trascrizione concettuale, la "Begriffsschrift" del 1879, assimila il procedimento dimostrativo a un calcolo, le cui regole sono in numero finito e tutte esplicitate a priori, e la correttezza dei cui risultati prescinde dal contenuto delle formule. Ciò non significa che la realtà di cui parlano le formule non abbia alcun ruolo; basti dire che le regole sono dettate dall' esigenza di buon funzionamento del calcolo in riferimento a tale realtà. Tuttavia questo ruolo è come quello di un'assemblea costituente che, incaricata di scrivere la costituzione di un paese ideale, ne produce una così perfetta che per il resto della storia di quel paese rimane inconcepibile la stessa idea di revisione costituzionale.

Tutti noi, inclusi i nostri colleghi nati prima di Frege, abbiamo imparato a manipolare formule logiche—per esempio nella risoluzione di sistemi di equazioni lineari. Un' equazione lineare con n incognite è un tipo molto semplice di formula logica con n variabili. Un sistema di equazioni è una congiunzione di tali formule. E la loro risoluzione è data da un pacchetto finito e definitivo di regole, garantite da un teorema che ne certifica la perfetta rispondenza alla finalità di risolvere sistemi lineari: nessun ingegno brillante aggiungendo nuove regole riuscirà a trovare nuove soluzioni; lo stesso insegnamento di queste regole appare standardizzato in tutti i paesi del mondo; la loro meccanicità è tale che la risoluzione di sistemi di equazioni lineari è ormai principalmente affidata alle macchine, come la precompilazione delle bollette della luce.

Uno dei programmi della logica è sempre stato quello di ricondurre *tutto* il processo di deduzione a un vasto sistema di equazioni risolvibili in base a regole costituzionali definitive. Negli esercizi 1-17 tale programma viene esemplificato in un contesto di logica giocattolo: quello del calcolo proposizionale booleano. Come voleva Frege, il calcolo deduttivo viene proposto senza curarsi del significato delle formule ma solo del loro aspetto esteriore; e siccome l'occhio vuole la sua parte, con qualche trucco le formule vengono smagrite per renderle più agili nei calcoli: come nel caso dei sistemi di equazioni, gli elementi linguistici più importanti, come i verbi e le congiunzioni sono messi in sordina, tanto che chi risolve questi esercizi non si accorgerà di stare facendo apprendistato logico.

Non è detto che questo sia il modo migliore per introdurre gli studenti all'idea di calculus ratiocinator: dal punto di vista della didattica della logica, presentare un calcolo che manipola formule senza preoccuparsi del loro significato è un' impresa difficile. Ogni insegnante è giustamente prevenuto verso tutti i meccanicismi perché ne conosce bene gli effetti diseducativi.

Come se non bastasse il problema didattico, dal punto di vista storico e culturale si stenta a credere che uno dei risultati di più di duemila anni di ricerca sul vero e il falso possa essere un criterio di verità basato sull' esame esteriore degli enunciati, senza entrare nel merito del loro contenuto: che verità significative si potranno mai accertare esaminando gli enunciati, ad esempio, del codice della strada, solo in quanto successioni di simboli ? a chi dare più credito: ai logici che—dermatologicamente—scrutano i segni calligrafici degli enunciati, o agli aruspici che scrutano le interiora degli animali ? e come potrà non ridere un logico che incontra un altro logico ?

Per questa caratteristica di non sporcarsi mai le mani coi contenuti, nei secoli passati la logica veniva talora vista come cavillosa, o scientificamente vacua, o inutilmente formalistica, o poco istruttiva. Nonostante ciò l' approfondimento autocritico le ha dato un ruolo sempre più importante, mostrando che molti problemi, oltre che insolubili, erano anche più gravi del previsto. Forse c'è qui un'analogia con l'autoalimentarsi del debito pubblico.

Varie ragioni hanno spinto i logici alla politica delle mani contenutisticamente pulite: nelle intenzioni di Frege la "scrittura concettuale" doveva servire a far trasparire i concetti e le relazioni tra di loro, come il denaro serve ad esprimere beni e risorse e la loro scambiabilità; può essere successo che il simbolismo abbia preso la mano, visto che i riferimenti oggettivi talora divengono incerti, o vorticosamente interscambiabili, mentre le manipolazioni formali hanno sempre una loro realtà intrinseca, indipendente da ogni riferimento a strutture rigide o a beni rifugio e alle loro vicissitudini, come—per tornare all'economia—ben sanno i movimentatori di capitali.

Fatto sta che si è definitivamente affermato il principio secondo cui le regole di inferenza debbano far riferimento solamente alla struttura esteriore delle formule e non al loro significato, così da poter essere applicabili anche da chi è totalmente digiuno di matematica, e persino dalle macchine calcolatrici. Riprendendo l'analogia coi sistemi di equazioni lineari, e pensando alle regole di inferenza come metodi di risoluzione di sistemi di equazioni universali, si pone subito il problema fondamentale, l'*Entscheidungsproblem*, o problema della decisione, che enunciamo seguendo Hilbert: "L'*Entscheidungsproblem* è risolto quando si conosca una procedura per decidere la validità o la soddisfacibilità di una data espressione logica, mediante un numero finito di operazioni." Qui la parola chiave è "procedura mediante un numero finito di operazioni", che useremo come sinonimo di algoritmo, procedura effettiva, procedura meccanica.

La prospettiva di una matematica interamente meccanizzata portava von Neumann nel 1927 a prefigurare una risposta negativa al problema della decisione; ma come definire "procedura meccanica" ? Ripensiamo a un risultato come quello citato di Gauss, sugli n- goni regolari non costruibili con riga e compasso. Pur riferendosi a metodologie di costruzione assai circoscritte, questo risultato limitativo presuppone la definizione rigorosa di "figura non costruibile con riga e compasso"—definizione che deve catturare l'essenza di ogni possibile e immaginabile costruzione con riga e compasso. Su scala molto più vasta, per rispondere negativamente all' *Entscheidungsproblem* occorreva predisporre una nozione di metodo di calcolo così ampia da includere ogni ragionevole procedura meccanica passata, presente e futura.

Una tale nozione fu data da Turing nel 1936.

La macchina di Turing

Prima di Turing molti autori avevano ideato ed anche costruito strumenti di calcolo; ricordiamo i progetti di calcolatori meccanici di Pascal e Leibniz, e la macchina analitica costruita da Babbage. Le macchine introdotte da Turing si differenziano da quelle precedenti per la loro universalità: almeno sulla carta esse sono in grado di eseguire ogni algoritmo immaginabile. Turing pervenne a una nozione universale di algoritmo analizzando e stilizzando il modo di procedere di uno scolaro mentre fa i compiti di matematica sul suo quaderno a quadretti. Infatti una *macchina di Turing* T è costituita da:

(i) un insieme finito A, detto *alfabeto,* i cui elementi sono chiamati *simboli.* Ogni alfabeto contiene un simbolo speciale ø, detto *blank,* e almeno un simbolo *nonblank* |. Fino ad avviso contrario, potremo supporre $A = \{ø, |\}$.

(ii) un insieme finito S i cui elementi sono chiamati *stati* (mentali) di T; per bandire ogni psicologismo, S viene identificato con l'insieme dei primi r+1 numeri naturali $0,...,r$; il numero 0 viene chiamato lo *stato iniziale* di T.

(iii) un *nastro* costituito da una fila di quadretti, o caselle, finita ma estendibile a piacere in entrambe le direzioni; ogni quadretto può contenere uno e un solo simbolo di A. Per non interrompere il nostro filo rigoroso identifichiamo il nastro con l'insieme degli interi $\mathbf{Z} = \{..., -2, -1, 0, 1, 2,...\}$.

(iv) un *pennino ottico* che può leggere e scrivere su un quadretto alla volta sul nastro, seguendo il *programma* di T, ossia un elenco di istruzioni.

Matematicamente, ogni *istruzione* è una quintupla del tipo $[a, s, a', D, s']$ (oppure: $[a, s, a', S, s']$) da leggersi in questo modo:

a	se nel quadretto in esame leggi il simbolo a
s	e sei nello stato s
a'	allora scrivi a' al posto di a
D	e passa al quadretto a destra (oppure, quando c'è "S" invece di "D", passa al quadretto a sinistra).
s'	e finalmente vai nello stato s'.

Mentre la bidimensionalità del quaderno a quadretti non è ritenuta da Turing elemento decisivo per la calcolabilità, decisive sono invece la finitezza dell'alfabeto e dell'insieme di stati mentali di ogni macchina T; visto che il nastro e l'alfabeto sono eguali per tutti, possiamo identificare T con l'elenco delle sue quintuple. Ogni casella del nastro vergine contiene il simbolo ø. Noi comunichiamo l'input a T scrivendo una successione di simboli |, un simbolo per casella. Poi posizioniamo il pennino ottico sul primo di tali simboli e mettiamo T in stato 0. A questo punto il pennino ottico di T non può prendersi molte licenze: la computazione avviene attraverso una successione di piccole modifiche della posizione del pennino, del simbolo stampato, dello stato di T, sempre secondo quanto prescritto dalle quintuple di T. Per evitare che il pennino ottico faccia la fine dell'asino di Buridano si stipula che, per ogni simbolo $a = ø, |$ e per ogni stato $s = 0,...,r$ tra le quintuple di T ce ne sia *al più una* che comincia con $[a, s,...]$. La computazione si interrompe, per mancanza d'istruzioni, quando il pennino si trova a leggere un simbolo a, in uno stato s, e nessuna quintupla di T comincia con $[a, s,...]$.

Per capire come una macchina di Turing possa calcolare una funzione, occorre definire la semplice nozione di "configurazione successiva". L'esercizio 27 è dedicato a questo scopo. Un *passo di calcolo* della macchina di Turing T è una coppia di configurazioni di T, di cui la seconda successiva alla prima.

Una funzione $f: \mathbf{N} \to \mathbf{N}$ è *Turing-computabile* se esiste una macchina di Turing T che avendo in input n simboli nonblank consecutivi, posta in stato 0 col pennino ottico sul primo di tale simboli, si ferma dopo un numero finito di passi, lasciando scritti sul nastro $f(n)$ simboli nonblank consecutivi. Un po' prosaico, no ? Eppure su questa definizione è passata una sorta di rivoluzione intellettuale. La generalizzazione a funzioni a due o più variabili è una rielaborazione di questa, e non aggiunge nulla di sostanziale; e così è il passaggio a macchine di Turing che lavorano con alfabeti meno striminziti del nostro $\{ø, |\}$, per esempio, scrivendo i numeri in notazione binaria o decimale, invece delle nostra trita notazione "unaria".

E siccome per le macchine di Turing non fa differenza che l'input rappresenti cifre numeriche o lettere alfabetiche, con la stessa tecnica si definisce la calcolabilità di funzioni f:{insiemi di parole} → {insiemi di parole}; per esempio esiste una macchina di Turing K che avendo in input una parola $l_1...l_u$ scrive in output la parola rovesciata $l_u ...l_1$. Se fossimo Turing-programmatori ci potrebbe capitare un esercizio che chiede di scrivere le quintuple per K. Un metodo pratico per convincerci dell'esistenza di K, benché estraneo ai canoni del rigore matematico, verrà descritto nella prossima sezione.

Tesi di Church

Gödel, che aveva ripetutamente cercato una nozione onnicomprensiva di "passo di calcolo deduttivo", scrisse che con le macchine di Turing si era ottenuta una definizione precisa e indiscutibilmente adeguata del concetto generale di sistema formale. Come giustificare questa affermazione ? Che portata ha la nozione di Turing-computabilità ?

Non c'è dubbio che una funzione Turing-computabile sia effettivamente calcolabile. Semmai ci può venire il dubbio che, stante la modestia delle operazioni del pennino ottico, certe funzioni complicate, benché computabili in un numero finito di passi, non siano alla portata di nessuna macchina di Turing.

Chi intraprenda lo studio della Turing-computabilità passa inizialmente per una fase di apprendistato di Turing-programmatore (esercizio 28) scrivendo le quintuple di macchine che calcolano funzioni come il successore $n+1$ (per $n \geq 1$), o l'addizione $x+y$ (per $x,y \geq 1$).

Uno degli effetti collaterali di questo tirocinio è di mostrare la macchinosità della Turing-programmazione. Tuttavia (esercizio 29) riusciremo a scrivere programmi di macchine di Turing per funzioni basilari come il successore, la costante zero, le varie funzioni identità $I(x_1,...,x_n) = x_i$. *Messa in guardia:* Gli studenti simultaneamente impegnati nello studio di una funzione complicata a variabile reale, e nella Turing-programmazione di una funzione identità, non avranno esitazione a ritenere la seconda attività meno gratificante della prima, anche perché non vedono a cosa possa servire una funzione identità.

In una fase successiva cercheremo tecniche che permettano di vivere di rendita, garantendo l'esistenza di nuovi programmi a partire da programmi già scritti. Queste tecniche sono scimmiottature delle operazioni che permettono di passare da funzioni ad altre funzioni. Familiare a tutti è l'operazione di *composizione,* come in $(x+y)^2$; meno familiare è la ricorsione primitiva, che ci limitiamo a definire per funzioni a una variabile:

DEFINIZIONE Sia c un numero e $g(x,y)$ una funzione. Sia f data da:
$f(0) = c$;
$f(n+1) = g(n, f(n))$
Allora f è ottenuta per *ricorsione primitiva* a partire da c e dalla funzione g.

ESEMPIO La funzione $f(n) = n!$ è definita da
$0! = 1$;
$(n+1)! = (n+1)\, n!$
e dunque, è ottenuta per ricorsione primitiva dal numero 1 e dalla funzione $g(x,y) = (x+1)y$.

Quando una funzione è ottenibile dalle funzioni basilari mediante composizioni o ricorsioni primitive, allora è detta *primitiva ricorsiva.* Abbiamo appena abbozzato il pedigree della funzione fattoriale; il pedigree completo chiede di far risalire la funzione g agli antenati basilari. Questo è il tema dell' esercizio 30. Data qualsiasi funzione primitiva ricorsiva f, accompagnata dal suo pedigree P, ogni Turing-programmatore esperto sa ricavare da P le quintuple di una macchina di Turing che calcola f.

Mentre la Turing-programmazione è un'attività macchinosa, pochi esercizi standard mostrano che primitive ricorsive sono tutte le funzioni dell'aritmetica elementare. Dunque tutte le funzioni dell'aritmetica elementare sono Turing-calcolabili. Trovare funzioni effettivamente calcolabili che non siano primitive ricorsive è così difficile che si arrivò a congetturare che non ce ne fossero, finché Ackermann e, indipendentemente, Sudan, ne costruirono una. Ackermann usò una procedura di *diagonalizzazione,* di cui daremo una versione semplificata nelle prossime righe.

Notiamo subito che, per definizione, le funzioni primitive ricorsive *unarie* (ossia, quelle a una sola variabile) possono essere elencate, cominciando da semplicissime funzioni di base (successore, identità x e costante zero), poi procedendo con funzioni di complessità crescente, in base al numero di composizioni o ricorsioni necessarie per scriverle: tale ordine si chiama *lessicografico.* Pensiamo ora a una serie di fascicoli, da raccogliere in un' enciclopedia-pacco, a cui pagina 1 contiene la prima funzione, pagina 2 la seconda funzione, ..., pagina x la x-ma funzione p_x,. Chiamiamo ora r^* la funzione che, dato x, va a pagina x, calcola il valore di $p_x(x)$ e gli aggiunge 1, in simboli:

$$r^*(x) = p_x(x) + 1.$$

La nozione di "procedura meccanica" che abbiamo innata ci porta a dire che la stampa delle prime x pagine dell'enciclopedia, e il calcolo di $p_x(x) + 1$ sono meccanicamente eseguibili. Dunque r^* è effettivamente calcolabile.

TEOREMA *r^* non è primitiva ricorsiva.*

La dimostrazione del teorema (esercizio 31) si basa su un *argomento diagonale,* assai simile all'argomento usato da Cantor per mostrare che l'insieme dei numeri reali non è numerabile.

Questo è un caso in cui una nozione imprecisa—quella di "effettivamente calcolabile"—risulta non equivalente a una nozione precisa, al di là di ogni ragionevole dubbio. Prontamente si dimostrò che la funzione di Ackermann è Turing-computabile. La

scoperta di r^* poneva un delicato problema di completamento, che ottimisticamente poteva essere formulato così: per catturare tutte le funzioni Turing-computabili basterà aggiungere alla composizione e alla ricorsione primitiva qualche operazione che ci è momentaneamente sfuggita ?

Fortunatamente la risposta è positiva: basta aggiungere l'operazione di minimizzazione, che è definita in questo modo: data una funzione Turing-computabile $g(x,y) : \mathbf{N}\times\mathbf{N} \to \mathbf{N}$ sia ZER(x) l'insieme dei numeri y tali che $g(x,y) = 0$. Si supponga che *per ogni* x, ZER(x) *sia non vuoto.* Allora per *minimizzazione* di g si ottiene la funzione f: $\mathbf{N} \to \mathbf{N}$ data da $f(x)$ = minimo elemento di ZER(x). La generalizzazione al caso di più variabili è ovvia.

Kleene nel 1938 definì le funzioni *parziali ricorsive* abrogando la condizione che ZER(x) non sia mai vuoto, e ammettendo così anche funzioni definite non per tutti gli x. Giustificò l'introduzione di queste funzioni nel miglior modo possibile, provando che *una funzione* f *è Turing-calcolabile se e solo se è parziale ricorsiva.* (Il numero x sta nel dominio di f se e solo se la macchina di Turing che calcola f avendo in input x si ferma in un numero finito di passi.)

Dunque con la minimizzazione si completa il quadro delle operazioni necessarie e sufficienti per ottenere le funzioni Turing-computabili. Molti approcci alla nozione di computabilità effettiva risultarono equivalenti alla Turing-computabilità. A lungo andare si affermò l'aspettativa che, per quanto imprecisa sia la nozione di calcolabilità, nessuno più sarebbe stato capace di mostrarne la differenza rispetto alla Turing-computabilità, come invece Ackermann ne aveva mostrato la differenza dalla calcolabilità primitiva ricorsiva.

La *TESI DI CHURCH* sostiene che questa aspettativa non andrà mai delusa: *per ogni funzione* f *effettivamente computabile c'è una macchina di Turing che calcola* f. L'imprecisione della nozione di "effettivamente computabile" produce strani effetti: da un lato pochi hanno dubbi sull'effettiva calcolabilità delle funzioni Turing-computabili. Più opinabile è il viceversa: non potrebbe l'evoluzione umana portare a metodi di computabilità effettiva così potenti da far apparire le macchine di Turing come oggetti d'antiquariato ? Per la riga e il compasso nessuno ha avanzato la tesi che essi esauriscano tutte le possibili e immaginabili metodologie di costruzione geometrica. Invece la tesi di Church asserisce che i passi delle macchine di Turing sono il non plus ultra delle procedure effettive.

Se la tesi di Church è vera, l'*Entscheidungsproblem*, che, come vedremo, Turing mostrò essere insolubile rispetto alla Turing-calcolabilità, risulta insolubile anche rispetto alla formulazione data da Hilbert, che fa riferimento a procedure con un numero finito di passi.

È sempre possibile dare versioni suggestive dei teoremi sacrificandone la precisione dell'enunciato: lo faremo una volta tanto, solo per amor d'analogia, dicendo che il teorema di completezza di Gödel ci induce a pensare che nessuna nuova "tecnica dimostrativa" varrà

ad aumentare il patrimonio di teoremi dimostrabili con le regole codificate nella logica dei predicati; su questa falsariga potremo allora dire che la tesi di Church ci induce a pensare che nessuna nuova "tecnica algoritmica" riuscirà a sopravanzare il potere di calcolo delle macchine di Turing.

Nella trattazione della Turing-computabilità i riferimenti alla tesi di Church sono sempre eliminabili. Quando la si trova menzionata in contesti matematici per asserire la Turing-computabilità di una funzione, come osservò Roberto Magari, la sua valenza è simile a quella dell'avverbio "ovviamente", soprattutto quando maschera il non desiderio dell'autore di fornire troppi noiosi dettagli, confidando nella comprensione del lettore a cui viene risparmiato di controllare tali dettagli.

Così per esempio, vale il seguente teorema: *la funzione che fornisce l'n-ma cifra decimale di* π *è Turing-computabile.* La dimostrazione che c'è una macchina di Turing che avendo in input un numero n, dopo un numero finito di passi scrive sul nastro l'n-ma cifra di π non è tra le più difficili, ma non è neanche tra le più istruttive. Molti di coloro che non hanno mai seguito in tutti i dettagli tale dimostrazione si convincono pienamente della validità del teorema, e di molti altri del genere, utilizzando una scorciatoia non rigorosa ma molto spiccia: la tesi di Church.

Risultati fondamentali della calcolabilità

Utilizzando l'equivalenza tra funzioni parziali ricorsive e funzioni Turing-computabili, e procedendo per analogia con quanto fatto per le funzioni primitive ricorsive, si ottiene (esercizio 32) un' enumerazione effettiva delle funzioni parziali ricorsive (unarie).

Siano dunque f_1, f_2, ... le funzioni così enumerate. Il teorema di enumerazione di Kleene dice che la funzione $U(x,y) = f_x(y)$ è parziale ricorsiva. Siccome le funzioni parziali ricorsive coincidono con le funzioni Turing-computabili, possiamo realizzare U come una lista finita di istruzioni di macchina di Turing, la *macchina di Turing universale.* La macchina U, avendo in input un arbitrario programma X di macchina di Turing e un numero y, simula il calcolo di X su input y e produce in output il valore $X(y)$. Qui nessun argomento di diagonalizzazione riesce a minare alla radice la macchina U: infatti, quando il valore di $U(y,y)$ non è definito, evapora ogni rischio di contraddizione nello scrivere $U(y,y) = U(y,y) + 1$.

Il programma X altro non è che un elenco finito di quintuple: pur contenendo tutta l' essenza di una macchina di Turing, X è una successione di simboli. La concezione pitagorica secondo cui tutto è numero ha qui un indubbio riscontro, anche se la riduzione della macchina X a numero x è chiamata gödelizzazione invece che pitagorizzazione. Una volta ridotta alla stregua di un numero, la macchina può essere messa in input ad un' altra macchina—oppure anche a se stessa. Questo processo prefigura la metamorfosi

dell'hardware nel software: mentre la programmazione assomigliava all'attività di un centralinista, von Neumann e Turing capirono che i programmi possono essere trascritti e conservati come successioni di simboli, esattamente come qualsiasi dato. Nacque così, ad opera di Eckart ed altri, lo *stored program computer,* incarnazione della macchina universale di Turing.

Accanto ai risultati positivi, assai importanti sono anche certi risultati negativi. Ad esempio, non esiste una macchina di Turing H con questa proprietà: avendo in input una coppia ordinata (X, y) ove X è un elenco di quintuple ed y è un numero, H si ferma se e solo se la macchina X si ferma avendo in input y (vedi l'esercizio 33). Anche questo risultato si ottiene con un semplice argomento di diagonalizzazione. Abbiamo appena mostrato *l'insolubilità del problema della fermata.* Ebbene, questo risultato ha una portata così grande da risolvere l'*Entscheidungsproblem.* Infatti, trascrivendo i passi di calcolo in espressioni della logica dei predicati, Turing ridusse il problema della fermata al problema della decisione. E concluse che se il secondo fosse risolvibile da macchine di Turing, anche il primo lo sarebbe. Dunque l'*Entscheidungsproblem è insolubile.* Questo è il teorema di Turing-Church sull'indecidibilità della logica dei predicati.

Notiamo qui la differenza profonda tra "ars decidendi" e "ars inveniendi": il teorema di completezza di Gödel fornisce le quintuple di una macchina di Turing R raziocinante: avendo in input una formula F della logica dei predicati, se F è valida R dopo un certo numero di passi *trova,* e produce sul nastro, una dimostrazione di F. Ricordando la prassi del silenzio-assenso siamo tentati di utilizzare R per *decidere* se F sia valida: se dopo un certo numero di passi di calcolo R non ha trovato una dimostrazione di F vorrà dire che F non è valida.

Già, ma quanto dobbiamo aspettare ? Intuitivamente, il teorema di indecidibilità dell' *Entscheidungsproblem* dice che i tempi d'attesa tipici del metodo del silenzio-assenso sono quelli della peggior geologia: rigorosamente parlando, non c'è una regola meccanica (Turing-computabile) che, noto F, delimiti il numero di passi di calcolo entro cui R trova, se esiste, una dimostrazione di F.

Se il teorema di completezza è sibillino di fronte a formule F non valide, esso fornisce importanti informazioni sull' insieme V delle formule valide. Siccome le dimostrazioni sono elenchi finiti di simboli presi da un alfabeto finito prefissato, elencando alfabeticamente tutte le dimostrazioni D_1, D_2, D_3,... in una seconda enciclopedia-pacco, e chiamando F_j la formula dimostrata in D_j, abbiamo una procedura meccanica, una macchina di Turing, per costruire un elenco completo F_1, F_2, F_3,... delle formule valide. In termini matematici, l' insieme V delle formule valide forma un insieme *ricorsivamente enumerabile:* in altre parole, V coincide con l'insieme dei valori di una funzione Turing-computabile.

L'intuizione suggerisce, e una dimostrazione rigorosa conferma, che la decidibilità è una condizione più forte dell'enumerabilità ricorsiva: un esercizio standard nella teoria della calcolabilità chiede di mostrare che un insieme W (di formule o di numeri) è decidibile se e solo se sia W che il suo complementare sono ricorsivamente enumerabili.

Noi abbiamo appena incontrato un insieme di formule indecidibile e ricorsivamente enumerabile: il Teorema di Turing-Church di indecidibilità della logica dei predicati si ottiene dimostrando che non è ricorsivamente enumerabile l'insieme NV delle formule non valide —mentre il complementare di NV è ricorsivamente enumerabile, visto che l'insieme V delle formule valide è ricorsivamente enumerabile.

Ricapitolazione

Ripercorriamo il circolo di idee attorno alla calcolabilità:

Punto di partenza è il problema della decisione per la logica dei predicati, l'*Entscheidungsproblem* di Hilbert: per poter divenire teorema, l'intuizione che il problema ha soluzione negativa richiede ben di più che la solita familiarità con procedure algoritmiche. Richiede una definizione rigorosa e convincente del confine tra la calcolabilità e la non calcolabilità.

Turing scopre una nozione soddisfacente di calcolabilità puntando il microscopio sui passi di calcolo di uno scolaro, fino alla messa a fuoco della loro ultima struttura, al di là della quale il processo di semplificazione non può andare.

Ogni macchina di Turing altro non è che una lista finita x di quintuple, e per pitagorizzazione-gödelizzazione, x è identificabile con un numero. Guardando ora col telescopio la totalità delle macchine di Turing, l'hardware si trasmuta in software: Turing scrive, una volta per tutte, il programma di una macchina universale U capace di simulare ogni macchina di Turing x avente un input arbitrario y.

Questa macchina universale U da un lato è inattaccabile da introspezioni diagonali alla Cantor-Ackermann, perché le funzioni parziali ricorsive non sempre rispondono ad ogni input. Dall'altro lato, U si sobbarca i problemi di fermata di tutte le macchine di Turing: quando scriviamo in input i due numeri x e y sorge il problema $U(x,y)$ se U si fermerà, o proseguirà indefinitamente a passeggiare sul nastro. Si tratta di un problema sostanziale, come il problema di sapere se un sistema di equazioni ha soluzioni.

Passando ora al problema *diagonale* $U(x,x)$, per costruzione si vede che nessuna macchina di Turing c (ove c sta per "cretese") riesce a decidere il problema $U(x,x)$ per ogni x: infatti, se così fosse, che cosa dovremmo pensare del problema $U(c,c)$?

Il problema $U(x,y)$ è una goccia nel mare di problemi che si possono trascrivere nella logica dei predicati. Se una macchina e risolvesse il problema della decisione, da e riusciremmo a costruire una macchina c che decide $U(x,x)$ per ogni x. Segue che nessuna

macchina di Turing riesce a risolvere l'*Entscheidungsproblem* . Questo è il teorema di Turing-Church.

Per affrontare ad armi pari l'*Entscheidungsproblem*, che Hilbert formulò servendosi della nozione imprecisa di "procedura mediante un numero finito di operazioni", ci può servire la tesi di Church, che identifica questa nozione con la Turing-computabilità.

Mentre l'umanità (matematica) aveva meditato per millenni sulle idee di spazio e numero, le sue elaborazioni sulla nozione di calcolabilità rimasero rudimentali fino agli anni trenta del ventesimo secolo: tutti riconoscevano l'effettiva calcolabilità dell'addizione o della moltiplicazione, ma nessuno aveva pensato a definire la non calcolabilità. Negli anni trenta emersero gli aspetti più sottili e riposti di granularità, finitezza, incompletezza, universalità delle macchine di Turing. Poi, a partire da quegli anni la teoria della calcolabilità ha avuto un'accelerazione.

PARTE II: APPLICAZIONI DELLA LOGICA ALL'INFORMATICA

Anno dopo anno nel secolo ventesimo le relazioni tra logica e informatica si sono andate sviluppando come le relazioni tra analisi matematica e fisica nel secolo diciannovesimo. Molti settori dell'informatica non avrebbero potuto esistere senza le idee e gli strumenti della logica: la dimostrazione automatica di teoremi, la verifica della correttezza dei programmi, la programmazione logica, le forme di rappresentazione ed elaborazione della conoscenza e del linguaggio naturale fornite da svariate logiche su misura. È quasi una tautologia dire che non si possono applicare concetti logici senza avere studiato la logica. Come corollario, in Italia e in altri paesi di consolidate tradizioni scientifiche la logica matematica ha assunto un ruolo fondamentale nell'insegnamento e nella ricerca.

Nelle righe successive discuteremo, per cenni necessariamente brevi, alcune applicazioni della logica all' informatica.

Linguaggi e programmi: il problema della certificazione

Abbiamo già incontrato un esempio di linguaggio, la Turing-programmazione, scoprendone pregi e difetti. Chi non ha ancora fatto gli esercizi 28-29 può comunque fare questo esperimento con due Turing-programmatori, che chiameremo A e B: presa una funzione computabile non banalissima f, si chieda ad A di scrivere un programma P per f; poi si mostri P a B, chiedendogli che funzione è calcolata da P.

Per aumentare il grado di mutua comprensibilità tra programmatori, negli anni intorno al 1950 si introdussero linguaggi più comunicativi, come il Fortran, il Cobol e l'Algol 60. Anche qui però rimanevano carenze di vario tipo, che rendevano difficile scrivere programmi molto complessi e molto affidabili nello stesso tempo. Quasi tutti sanno che una virgola mancante in un programma può far perdere un satellite artificiale.

Nacquero così linguaggi di programmazione modulari come il Pascal, i cui programmi sono fatti di sottoprogrammi, ogni sottoprogramma essendo fatto di sotto-sottoprogrammi, e così via, con ovvi vantaggi per quanto riguarda l'assemblaggio, il collaudo e la ricambiabilità dei pezzi difettosi.

Purtroppo non basta collaudare e ricollaudare un programma per certificarne la correttezza: il software commerciale di buona qualità è stato ampiamente collaudato, ma spesso contiene errori. Quando i primi programmi cominciavano a girare nessuno si preoccupava di questioni di correttezza. Col diffondersi del fenomeno di cattivo funzionamento dei programmi si ebbe la "crisi del software" accompagnata dalla nascita di una nuova disciplina: la verifica (della correttezza) dei programmi. La correttezza di un programma, come la correttezza di un enunciato matematico, va provata in maniera formale (esercizio 34).

La logica dei predicati è lo strumento principale per intervenire nel merito dei programmi. Già abbiamo parlato del prototipo di tali interventi, quello attuato da Turing per descrivere il comportamento della macchina universale *U*.

Come mostrato dall' esercizio 34, la certificazione di una macchina di Turing *T* si fa in due tempi: nel primo tempo si scrive un enunciato *E* che dice che fissato l'input e fissata l'inizializzazione di *T*, dopo un certo numero di passi *T* si ferma su un certo output; nel secondo tempo si dimostra *E*, utilizzando le definizioni della Turing-computabilità, assieme a tutti gli strumenti matematici disponibili: l'induzione, la teoria dei numeri, l'analisi combinatoria. Ecco una delle ragioni per cui un informatico non può non avere anche una buona formazione matematica.

Visto che le computazioni di *T* sono concetti matematici proprio come le rotazioni di un triangolo, non vi è altro modo di descrivere le proprietà input-output di *T* se non mediante affermazioni matematiche. E non vi è altro modo per certificare la veridicità di questi comportamenti se non dimostrando queste affermazioni. Il resto è gesticolazione, tanto meno convincente quanto più altisonanti sono i nomi di cui si fregia.

Ritornando all' esercizio 34, notiamo che l'affermazione in esso contenuta non si impegna a dirci che cosa *T* farà se all'inizio il nastro contiene un input strano (del tipo | | | ø ø ø | |) , oppure se *T* non è messa nello stato 0, o se il pennino non è posizionato a dovere. Né ci dà tutti i dettagli sulla configurazione finale o su quelle intermedie. Se dovessimo utilizzare *T* come sottomacchina di una macchina *S* avente compiti molto più vasti, questi *distinguo* cavillosi su situazioni critiche potrebbero divenire decisivi, e dunque richiedere un supplemento di indagine per ricavare le dovute certificazioni più generali. Da tale indagine quasi sempre risulterà che sono necessarie sostanziali revisioni di *T*, prima di poterla incorporare in *S*.

Dunque vediamo come sia utile la proprietà di modularità dei linguaggi come il Pascal: come è più facile dimostrare un teorema smontandolo in lemmi, così è più facile certificare un programma lungo smontandolo nei suoi sottoprogrammi corti.

L'analogia tra teorema/dimostrazione e programma/certificazione non finisce qui, benché percentualmente siano assai di più i teoremi corredati da dimostrazione corretta che i pacchetti software corredati da certificazione valida. Ancor oggi il rigore nella certificazione del software è paragonabile al rigore monetario. Quasi si trattasse di certificazione *DOC* per vini provenienti da uve non ancora maturate, certuni ritengono che la certificazione dei programmi quando serve è impossibile, e quando è possibile o non serve o è fasulla; questo atteggiamento è simile a quello che alcuni principianti hanno verso le dimostrazioni in geometria.

Logica e programmazione logica

In questo capitolo ci interessiamo del *modo* dei verbi usati per fare i calcoli. Mentre le frasi principali degli enunciati dei teoremi e delle dimostrazioni usano quasi solamente l'indicativo (con un pizzico di congiuntivo nelle ipotesi, e di condizionale nelle argomentazioni per assurdo), nelle quintuple delle macchine di Turing le istruzioni parlano con verbi all'imperativo. I verbi all'imperativo corrispondono alla percezione dello svolgimento di un calcolo o di una costruzione geometrica come successione di passi in ciascuno dei quali viene dato ed eseguito un ordine: l'imperativo è il modo tipico dell' insegnante quando cerca di aiutare uno studente incerto alla lavagna. Che altro modo ci dovrebbe essere per suggerire un calcolo o una costruzione ?

Eppure negli assiomi della geometria elementare i verbi sono all'indicativo: per due punti distinti *passa* una sola retta. Nella definizione di qualsiasi funzione, per esempio, il fattoriale, i verbi sono all'indicativo: 0! *è* eguale a 1, e $(n+1)!$ *è* eguale a $(n+1)$ moltiplicato per $n!$ Se poi dovessimo calcolare il valore di 2!, lo faremmo applicando successivamente le definizioni, continuando a leggere i passaggi all'indicativo. Quanto poi alla moltiplicazione, ogni casella delle tabelline stampate nelle pagine di copertina dei quaderni d'aritmetica si legge, e si può calcolare (esercizio 26) utilizzando verbi all'indicativo. Questo capitolo parlerà del ruolo dell'indicativo per calcolare logicamente le funzioni. Non pretenderà invece di dire che l'indicativo può essere usato per aiutare uno studente incerto alla lavagna.

La prima idea di ogni linguaggio di programmazione basato sulla logica è che ogni funzione Turing-computabile *f* può essere trattata allo stesso modo del fattoriale: scrivere un programma per il calcolo di *f* significa assiomatizzare *f*, con verbi all'indicativo. Una macchina di Turing *R* preventivamente istruita con le regole del calculus ratiocinator si incaricherà di calcolare $f(x)$, sulla base della definizione, da noi fornita, della funzione *f*, e del numero *x*. *R* procederà per deduzioni successive, al modo di una dimostrazione matematica diretta. Se per caso *R* sbagliasse a dedurre il valore di $f(x)$ dovremo prendercela solo con noi stessi, per aver sbagliato qualcosa nel testo dettato del problema—il che non è mai esaltante, perché trasforma un problema che credevamo altrui in un problema nostro. Certo non possiamo dubitare delle regole deduttive della logica dei predicati.

Se l'*Entscheidungsproblem* fosse fulmineamente decidibile, il problema della correttezza di un programma logico per calcolare $f(x)$ sparirebbe per incanto—basta saper formulare la domanda con precisione. Come sappiamo invece, l'*Entscheidungsproblem* è indecidibile; tutto quello che abbiamo è il teorema di completezza di Gödel che certifica che la macchina *R* riesce infallibilmente a dimostrare ogni formula valida.

Benché *R* sia sibillina con le formule non valide, e geologicamente lenta con quelle valide, ciò non ha scoraggiato il progetto di una deduzione automatizzata. E così il

calculus ratiocinator è stato rivoltato in ogni sua piega, tradotto e ritradotto nelle forme più adatte ai linguisti, ai logici, agli informatici, alle macchine—ovviamente, senza alterare il patrimonio di teoremi dimostrabili—proprio come l'algoritmo del gioco delle parole intersecanti (esercizi 16,17) è una rielaborazione dall'algoritmo delle tavole di verità.

Come dimostrato da Alan Robinson nel 1965, una sola regola deduttiva, chiamata *risoluzione*, è sufficiente per ottenere tutte le deduzioni possibili nella logica dei predicati, una volta adottata la notazione "clausale" (esercizi 18-24, introduttivi ai problemi della programmazione logica). Svolgendo questi esercizi vedremo che le macchine a cui viene affidata la deduzione logica del calcolo di $f(x)$ non lavorano gratis.

Per cominciare, la dettatura di f e di x, ed anche la lettura del risultato $f(x)$ sono meno semplici del previsto. Proporzionalmente all'efficienza che chiediamo alle macchine, esse ci chiedono di stringare al massimo il nostro modo di fare matematica, abolendo tutte quelle invenzioni simboliche, immagini geometriche, figurazioni retoriche che caratterizzano il linguaggio dei matematici, nell'insegnamento e nella ricerca. Le macchine di Turing non apprezzano la nostra capacità di ragionare per analogie—non capiscono che la loro stessa esistenza deriva dal respiro matematico di un'analogia felice.

La stessa pregiudiziale dell' oratore verso l' attore, e del logico verso l'oratore, ora viene lanciata dalla macchina di Turing R al logico: il *mio* linguaggio ha meno gesticolazione e più efficacia sulla realtà. E così per le esigenze dell' efficacia deduttiva, e di una maggiore applicabilità a problemi concreti, la logica dei predicati autoriduce le sue potenzialità espressive. Nelle versioni più spintamente "machine oriented" essa è resa quasi irriconoscibile, spezzettata in polpette clausali d'informazione tutta positiva, quasi come il linguaggio della pubblicità. Si veda l'esercizio 25. A partire dagli anni 70 il frammento della logica dei predicati costituito dalle clausole Horn ha acquisito lo status di linguaggio di programmazione.

Chi è passato attraverso il lungo esercizio 26 si convincerà che la programmazione logica riesce a riempire correttamente ogni casella della tabellina dell'addizione, della moltiplicazione, e di ogni funzione Turing-computabile. Ma chi ha continuato a pensare all' esercizio anche dopo la soluzione (un insieme certamente non vuoto), continuerà a farsi delle domande.

Una domanda, suggerita dall' antiquariato delle tavole dei logaritmi e dei regoli calcolatori a scorrimento, è se i metodi dell'insegnamento tradizionale della matematica cederanno il passo a metodi deduttivi interamente vigilati da calcolatore—puntando a demandare tutta l'attività raziocinante umana alle due protesi offerte dal calcolatore e dal televisore.

Un'altra domanda potrebbe essere: perché questo mordersi la lingua per un simbolo in più o in meno ? Che cosa succede se aggiungiamo un simbolo di relazione Uxy a cui far

dire "x è uguale ad y" ? *Risposta*: L'eguaglianza non è una relazione d'equivalenza qualsiasi. Per trattare logicamente il nuovo simbolo U non basterà scortarlo con tre clausole che ne attestino la riflessività, simmetria e transitività, come nell' esercizio 23. La logica si fa carico della specificità del simbolo U aggiungendo regole deduttive ad hoc per espressioni contenenti U. Tali regole supplementari, ben spiegate nella letteratura, complicano e rendono meno efficiente il calculus ratiocinator Hornicus,come prezzo per l'accresciuto potere espressivo.

Passando ora dal livello della programmazione logica a quello della logica dei predicati, l' esercizio 26 è istruttivo anche per chi pensa alla logica come a una fiera della vanità simbologica, in cui, simbolo più simbolo meno, tutto rimane sempre allo stato superficiale. Al contrario, certi simboli, ad esempio quello di infinito ∞, sono così potenzialmente carichi di significato che le regole deduttive per trattarlo a livello logico ... semplicemente non possono esistere. Questa conseguenza del teorema di incompletezza di Gödel impone al logico di rinunciare a un simbolo più importante per lui che per l'oratore.

Forse perché passata in pochi decenni attraverso molte prove, la logica dei predicati appare ben stagionata e ben lontana dall'invecchiamento. Essa è una creazione dell' immaginazione umana perennemente impegnata a (i) esprimere simbolicamente porzioni sempre più significative della conoscenza, mai perdendo di vista la possibilità di (ii) elaborare ed espandere tale conoscenza manipolandone le espressioni simboliche. Per quello che abbiamo visto in questo capitolo, a tutti i livelli appare una sorta di proporzionalità inversa tra grado di espressività di una logica, e grado di efficacia del suo calcolo deduttivo.

L'equilibrio tra (i) e (ii) realizzatosi nella logica dei predicati ha comportato drastiche limitazioni espressive rispetto al sogno di una *ars magna* di molti logici dei secoli passati. Ma, come e di più che per la scala musicale finalmente temperata in dodici semitoni, l'equilibrio raggiunto è stato un punto di partenza per la costruzione di strumenti meglio intonati, precisi e potenti—su cui si è copiosamente riversata l'intuizione creatrice.

Complessità e calcolabilità polinomiale

Mentre la Turing-programmazione ha lasciato il posto a linguaggi più comunicativi, le macchine di Turing hanno ancora un ruolo insostituibile nella teoria e nella didattica della calcolabilità. Oltre che per i teoremi fondamentali citati precedentemente, il passo di macchina di Turing è ancora adottato come unità di conto per misurare i costi dei calcoli. In questo capitolo, tutto dedicato ai problemi dell'efficienza, diviene necessario arricchire l'alfabeto delle macchine di Turing, in modo da poter scrivere i numeri in notazione binaria o decimale. Altrimenti molti problemi divengono falsamente facili (relativamente alla lunghezza dell'input) solo perché l'input è stato scritto in maniera irresponsabilmente lunga.

In una lettera a von Neumann del 1956, Gödel si sofferma a considerare una macchina di Turing T che, avendo in input una formula F della logica dei predicati ed un numero x, decide se F possegga una dimostrazione di lunghezza $\leq x$. Denotiamo con passi(x) il massimo numero di passi occorrenti a T per rispondere al problema al variare di F tra tutte le formule, necessariamente di lunghezza $\leq x$. Gödel chiede a von Neumann una stima sulla velocità di crescita della funzione passi(x) al crescere di x. Gödel chiede anche se von Neumann abbia in mente qualche algoritmo veloce per riconoscere, ad esempio, i numeri primi.

Non si sa se von Neumann abbia risposto a questa lettera: egli era molto ammalato e morì nel febbraio del 1957. In ogni caso, Gödel in questa lettera poneva il problema di quantificare il grado di facilità di problemi computabili, in termini di numero di passi di macchine di Turing.

Oggi sappiamo che passi(x) $\leq$ qualche polinomio in x se e solo se la soddisfacibilità di formule nella logica proposizionale è decidibile con un numero di passi polinomiale: e questo è uno dei problemi fondamentali e insoluti dell'informatica.

Nella teoria della complessità l'accento è posto sui costi degli algoritmi. Ciò ha senso, visto che il tempo di cui disponiamo è poco e decresce rapidamente. Ma come definire rigorosamente la "facilità" di un problema ? che cos'è un "problema" ? Risposta all'ultima domanda: il problema dei numeri primi è l'insieme dei numeri primi; il problema di soddisfacibilità è l'insieme delle formule soddisfacibili, ... ogni problema viene identificato con l'insieme delle sue soluzioni—come se dicessimo che un problema è il "luogo" delle sue soluzioni. Ciò non toglie che nel presentare un problema usiamo le consueta terminologia, con i verbi all'infinito-imperativo. Tra i problemi più interessanti citiamo:

—il problema di *soddisfacibilità*: data una formula in forma normale congiuntiva, riconoscere se essa è soddisfacibile;

—il problema dell'*intersecabilità*: dato un elenco di parole con lettere maiuscole e minuscole, riconoscere se esista una parola normale che le intersechi tutte;

—il problema dei *francobolli*: dato un numero A e una collezione di francobolli di valore $b_1, b_2, \ldots, b_n$ decidere se con un'oculata scelta si riesca ad ottenere esattamente l'affrancatura A.

Per formulare matematicamente il fatto che un problema è facile, si usa la seguente definizione:

DEFINIZIONE Un problema Q è *decidibile in tempo polinomiale,* (in breve, Q appartiene alla classe P) se c'è un polinomio r e una macchina di Turing T con queste proprietà: avendo qualsiasi input x di lunghezza n, T riesce a decidere se x è soluzione di Q entro al più $r(n)$ passi.

Mentre per l'omologazione computabilità = Turing-computabilità ci fu la tesi di Church, nessuno ha avuto il coraggio di dire che la nozione di "problema facile" è pienamente catturata dalla classe P. Al massimo si è disposti a concedere che quando un problema non è nella classe P allora è assai ostico. La classe P comprende tutti i problemi più elementari, come il riconoscere se un numero è pari o dispari, multiplo di 7, quadrato perfetto, o riconoscere se due numeri sono primi tra loro.

Forse l'unico algoritmo difficile insegnato nella scuola elementare è quello che fornisce il minimo comun denominatore di due frazioni, utilizzando la scomposizione in fattori primi. Per un capriccio della sorte scolastica, Euclide descrisse un algoritmo che oltre a dare i minimi comun denominatori in tempo fulmineo, fornisce spunti culturali non inferiori a quelli della scomposizione in fattori primi, in quanto direttamente collegato alla commensurabilità di segmenti, e alla teoria delle grandezze. Invece il popolarissimo algoritmo di scomposizione in fattori primi resiste ad ogni tentativo di soluzione in tempo polinomiale.

Impermeabile ad ogni tentativo di semplificazione è, dopo anni di assedio, un manipolo di problemi importanti in tutti i campi dove si eseguono calcoli, tra cui quello di soddisfacibilità e quello dei francobolli. Se analizziamo la struttura generale di questi problemi vediamo una bipartizione:

(i) indovinare magicamente un'assegnazione o un' affrancatura,
(ii) banalmente controllare che tale assegnazione o affrancatura è soluzione del problema.

Come ulteriore esempio dei due tempi (i) e (ii), ricordiamo la congettura di Fermat secondo cui ogni numero della forma $1 + 2^{2^n}$ è primo. Per $n = 1, 2, 3, 4$ tutto fila via liscio, ma Euler coprì—e questa è la fase (i)—che $1 + 2^{2^5} = 4294967297 = 641 \times 6700417$. Per scrupolo—e questa è la fase (ii)—possiamo prenderci la briga di controllare la validità di questa fattorizzazione con un calcolatore tascabile, o con carta e penna. Come avrà fatto Euler a scoprire questa fattorizzazione ?

Vogliamo approfondire i distinguo tra la fase (i) e la fase (ii). Con una certa vaghezza diciamo che un problema è di classe NP (*polinomiale nondeterministico*) se la sua soluzione passa per queste due fasi. Rimandiamo all' esercizio 35 per una definizione rigorosa. Benché a prima vista la differenza tra (i) e (ii) sia paragonabile alla differenza tra scrivere una sinfonia ed ascoltarla, nessuno ha dimostrato che NP contiene problemi non di classe P. Invece molti hanno scoperto che spesso i problemi in NP hanno casi facili, di classe P: per esempio, come abbiamo visto nell' esercizio 11, il problema di soddisfacibilità diviene facilissimo se ci restringiamo a formule Horn: per queste clausole la fase inventiva (i) è un pro-forma. Anche il problema dei francobolli ha casi facili, come si vedrà dall'esercizio 36.

Siccome nella pratica e nella teoria contano soprattutto i casi difficili, proprio come al tempo di Euler non mancano le opportunità di impiego per scopritori di fattorizzazioni, assegnazioni, affrancature. Il lettore che abbia svolto gli esercizi 15-17 ricorderà che non c'è molta differenza tra il problema di intersecabilità e il problema di soddisfacibilità: l'uno è facilmente *riducibile* all'altro. Nell'esercizio 37 viene descritto un semplice esempio di riduzione. Ebbene, lo stesso vale per molti problemi in NP. Cook dimostrò nel 1971 che il problema di soddisfacibilità ha un ruolo emblematico in questa classe: ogni problema in NP è velocemente riducibile al problema di soddisfacibilità. Dunque chi avesse talento di scopritore di parole intersecanti sarebbe automaticamente in buona posizione nelle liste di collocamento per scopritori di affrancature, di fattorizzazioni (come e più di Euler...).

Il problema delle parole intersecanti è sostanzialmente l'unico problema in NP: tutti gli altri sono sue varianti e possono essere velocemente ridotti ad esso. Chi scoprisse un algoritmo veloce per questo problema avrà d'un sol colpo risolto in tempo polinomiale tutti i problemi in NP, ridimensionando l'importanza della fase inventiva (i) che li caratterizza, e rispondendo positivamente alla domanda fatta da Gödel nella lettera a von Neumann.

Ringraziamento I commenti di Gabriele Lolli e di Dario Palladino mi hanno aiutato a migliorare questo testo.

Bibliografia

F. BELLISSIMA, P. PAGLI, *La Verità Trasmessa,* Sansoni, Firenze, 1993.

S. BOZZI, C. MANGIONE, *Storia della Logica,* Garzanti, Milano, 1993.

E. CASARI (a cura di), *La Logica del Novecento,* Loescher, Torino, 1981.

M. DAVIS (a cura di), *The Undecidable,* Raven Press, New York, 1965.

M. FRANCHELLA, C. MANGIONE, (a cura di), *Letture di Logica,* LED, Edizioni Universitarie di Lettere Economia Diritto, Milano, 1993.

G. LOLLI, *Introduzione alla logica formale,* Il Mulino, Bologna, 1991.

M. MUGNAI (a cura di), *La Logica da Leibniz a Frege,* Loescher, Torino, 1982.

D. MUNDICI, W. SIEG, *La Matematica Studia le Macchine,* in: Le Scienze della Mente, a cura di C. Mangione, Enciclopedia "Le Scienze e le Tecnologie, ieri, oggi e domani", Grandi Opere, Milano, 1995.

D. MUNDICI (a cura di), *La Scienza dei Calcolatori,* Le Scienze Quaderni, Edizione italiana di Scientific American, Vol. 56, ottobre 1990.

M.NEGRI, *Elementi di Logica,* LED, Edizioni Universitarie di Lettere Economia Diritto, Milano, 1994.

J. VAN HEIJENOORT (a cura di), *From Frege to Gödel,* Harvard University Press, Cambridge, Massachusetts, 1967.

Esercizi

DANIELE MUNDICI

UN CALCOLO DEDUTTIVO SENZA LOGICA

Parole che si intersecano

Il nostro punto di partenza è un insieme finito di lettere minuscole e maiuscole, a, A, b, B, c, C, ... Successioni di lettere, anche senza senso compiuto, sono chiamate *parole;* per esempio, AAA, Anna, ottO, a, OttO, aliDdOSO, caNE, GATto, sono parole. Diciamo che due parole p_1 e p_2 *si intersecano* se hanno almeno un simbolo in comune; per esempio gAtTo e cAnE si intersecano (hanno in comune il simbolo A) mentre GaTtO e cAnE non si intersecano. Diciamo che una parola p interseca un insieme di parole $p_1,...,p_n$ se p interseca ciascuna parola p_i. Ad esempio, l'insieme di parole Ab, AB, aB, ab è intersecato dalla parola Aa; invece la parola AB non lo interseca.

ESERCIZIO 1 Scrivi alcune parole $p_1,...,p_n$ con maiuscole e minuscole; per ogni i, j = 1,...,n guarda se p_i interseca p_j. La relazione di intersezione tra parole è riflessiva, simmetrica, transitiva ? Osserva che se q interseca $p_1,...,p_n$ e la parola r è un anagramma di q, anche r interseca $p_1,...,p_n$. Dunque l'ordine con cui si scrivono le lettere è irrilevante ai fini dell'intersezione di parole. Quali di queste operazioni sono altrettanto irrilevanti ?

(i) elidere da una parola eventuali doppioni di una lettera maiuscola lasciandone una copia sola; stessa cosa per copie multiple di una lettera minuscola (esempio, passare da PIOgGereLLIna a PIOgGerLna)

(ii) aggiungere per ogni lettera maiuscola che compare nella parola, anche la corrispondente minuscola (esempio, passare da CaNe a CaNecAnE)

(iii) scrivere la parola al contrario (esempio, trasformare StrAdA in AdArtS)

(iv) scambiare maiuscole con minuscole (esempio, trasformare Roma in rOMA)

Una parola è detta anormale se contiene sia la maiuscola che la minuscola di una lettera; altrimenti la parola è detta *normale.* Per esempio, Oslo, aNDREA, ALla, SisiFo, inNO, elLA, AnNa sono parole anormali, mentre AnnA, sIsIfO, ELLA, ellA, cinque, b, ANDREA, InnO sono parole normali. Ci interessa il seguente gioco, della cui importanza parleremo il più tardi possibile:

Date le parole $p_1,...,p_n$, cercare una parola normale p che le intersechi, oppure dimostrare che tale p non esiste.

Prendere le lettere iniziali di $p_1,...,p_n$ ci aiuterà a trovare p solo nei casi fortunati: in generale le iniziali formano una parola anormale. Anche l'idea di pescare astutamente un

simbolo da ciascuna parola evitando doppioni maiuscolo/minuscolo non ha finora prodotto risultati apprezzabili, benché su quest'idea siano già stati spesi moltissimi anni-uomo.

ESERCIZIO 2 Prendi tre versi della tua poesia preferita. Riscrivili alternando lettere maiuscole e minuscole, per esempio trasformando "La donzelletta vien dalla campagna" in "La DoNzElLeTtA vIeN dAlLa CaMpAgNa". Decidi se le parole così ottenute sono intersecabili da una parola normale.

La ricerca esaustiva

In attesa di risultati più brillanti, ecco una burocratica, ma infallibile procedura per decidere se $p_1,...,p_n$ è *intersecabile,* ossia decidere se esista una parola *normale* che interseca ciascuna parola p_i. Cominciamo col definire *ridondante* una parola in cui occorra due o più volte lo stesso simbolo: ad esempio ciascuna delle parole gaTTo, PaRigi, auTomobili, dissimilissiMi, MaMmA è ridondante, mentre le parole gaTto, PaRiGI, automObiLI, MamA non sono ridondanti. Per decidere se l'insieme di parole $p_1,...,p_n$ è intersecabile procediamo così:

PASSO 1 Intanto restringiamo il campo di ricerca a parole che sono simultaneamente (i) normali, (ii) non ridondanti e (iii) contenenti esclusivamente i simboli che si trovano in $p_1,...,p_n$. Questo ci permetterà di esaminare solo un campione finito (perché il campione è finito ?) di candidati $q_1,...,q_m$, senza paura di sbagliare il verdetto finale.

ESERCIZIO 3 Se l'insieme $p_1,...,p_n$ è intersecato da una parola q normale allora esso è intersecato anche da una parola che soddisfa (i)-(iii).

PASSO 2 Per ciascuna parola q_j dell'elenco preparato al passo 1 controlliamo se essa interseca $p_1,...,p_n$. Se ne troviamo una, dichiariamo $p_1,...,p_n$ intersecabile; se invece ci troviamo ad aver percorso tutto l'elenco senza aver trovato una parola intersecante, dichiariamo $p_1,...,p_n$ non intersecabile. L'esercizio appena svolto assicura che nessuno ci smentirà.

Questo procedimento, noto con il nome di *ricerca esaustiva,* va bene finché i simboli delle parole $p_1,...,p_n$ si contano sulle dita della mano. Invece, già per una dozzina di simboli è difficile trovare volontari disposti a fare il passo 1. Quando i simboli sono un centinaio, anche il più volenteroso non avrebbe abbastanza tempo per stilare l'elenco $q_1,...,q_m$.

ESERCIZIO 4 Utilizzando il metodo di ricerca esaustiva, dimostra che l'insieme di parole ONU, Uno, uno, oNu, Onu, UNo, UnO, uNO non è intersecabile.

Se non altro per mancanza di tempo e d'inchiostro dovremo cercare un metodo più sbrigativo per decidere l'intersecabilità di $p_1,...,p_n$. Il nostro ideale di "metodo sbrigativo" ci è

suggerito da quello che capita in certi casi fortunati dell'aritmetica: ad esempio, dati i numeri naturali $P_1, P_2, ..., P_n$, con $1 < P_1 < P_2 < ... < P_n$, c'è un metodo fulmineo per decidere se esista un numero $Q > 1$ che li divida tutti: tale metodo, associato al nome di Euclide, risolve il problema calcolando direttamente il massimo comun divisore con un piccolo numero di divisioni, senza scomporre i numeri P_i in fattori primi. La tartaruga esaustiva, che per ogni numero $d = 2, 3,..., P_1$ controlla se d divide tutti i P_i, vede sfrecciare il razzo euclideo che in un baleno risolve il problema. Infatti, mentre la scomposizione in fattori primi di un numero di 500 cifre sfida ancora oggi i più potenti calcolatori, un personal computer calcola il massimo comun divisore di due numeri di 5000 cifre in pochi secondi utilizzando l'algoritmo euclideo.

ESERCIZIO FACOLTATIVO Trovare un metodo sbrigativo per decidere se esista o no una parola p normale che intersechi un insieme arbitrario di parole $p_1,...,p_n$.
Suggerimento Questo esercizio è uno dei problemi fondamentali dell'informatica. Il problema, a tutt' oggi irrisolto, fu posto da Gödel in una lettera a von Neumann, di cui parliamo nelle lezioni.

Una partita fra due squadre: la completezza

Un principio condiviso da tutti coloro che debbono risolvere equazioni è di non mettersi a cercare soluzioni se non ne esistono. Al problema dell'esistenza di soluzioni va data precedenza assoluta; per rispondere al problema non occorrono molti numeri o parole, ma un semplice bit 1 = sì, oppure 0 = no. Anche per il problema delle parole intersecanti considereremo prima il problema d'esistenza.

Immaginiamo una partita tra due squadre di giocatori, la squadra del sì e quella del no. Al fischio di inizio appare su un tabellone dello stadio un elenco di parole $p_1,...,p_n$. Vincerà la squadra del sì se $p_1,...,p_n$ è intersecabile, altrimenti vincerà la squadra del no. Ogni squadra si impegna per vincere e convincere. Come in certi sport ultradifensivisti, può passare molto tempo senza che succeda niente; al punto che l'arbitro, per conto proprio, si mette a fare una ricerca esaustiva del vincitore, non tanto per spirito di protagonismo, ma per un ragionamento di questo tipo: se la squadra vincente fosse così inetta da non riuscire a dimostrare di aver vinto, o se la squadra perdente cercasse di barare con false dimostrazioni di vittoria, con la mia ricerca esaustiva io sarò comunque in grado di decretare infallibilmente la vittoria a tavolino.

Ma allora che cosa ci stanno a fare le squadre se—come nel peggior calcio scommesse—la vittoria/sconfitta è decisa in una sede diversa dal campo di gioco ? Risposta: qui, come in tanti altri casi della vita, benché vittoria e sconfitta dipendano da un giudizio ineluttabile, i tempi del giudizio sono esasperantemente lenti. E così ogni squadra

cerca di bruciare le tappe, escogitando tutte le possibili scorciatoie alla ricerca esaustiva. E in effetti le emozioni non mancano nemmeno nel gioco delle parole: ogni tifoso della squadra del sì grida goal quando la sua squadra riesce a trovare una parola normale q che interseca $p_1,...,p_n$.

ESEMPIO Le parole: I, cIpReSsI, cHe, A, boLgHeRi, AlTi, E, sChIeTtI, vAn, Da, SaN, gUiDo, In, duplice, FILAR, quaSI, In, CORsa, GIGANTI, giovinetti
sono intersecate dalla parola normale: AcDEgINOtu

Questa parola, da sola, costituisce una prova di intersecabilità così decisiva che la stessa squadra del no, arrendendosi all'evidenza, smette di cercare prove di non intersecabilità e ammette di aver perso la partita. E l'arbitro, senza paura di contestazioni, decreta la vittoria della squadra del sì, e manda tutti negli spogliatoi ben contento di poter interrompere la ricerca esaustiva.

Invece, per esperienza, ogni tifoso della squadra del no sa che le vittorie della sua squadra sono più sofferte, al punto da far rimpiangere le vittorie di Pirro: le regole del gioco dicono che la squadra del no vince dimostrando che $p_1,...,p_n$ *non* è intersecabile; il copione è sempre lo stesso: la squadra, pazientemente, intraprende la ricerca esaustiva di tutte le possibili parole normali, senza farsi illusioni di vittoria fino all'ultimissimo minuto, quando l'ultimo candidato è stato esaminato e risulta non intersecare $p_1,...,p_n$. Solo allora la squadra vince. Unico conforto, la certezza che nel frattempo la squadra del sì non è riuscita ad imbrogliare l'arbitro e portarsi via una vittoria fasulla. Per questa ragione ogni tifoso della squadra del no ha la sensazione che la sua squadra sia svantaggiata rispetto alla squadra del sì. Possibile che quelle fortunate invenzioni che permettono vittorie fulminee alla squadra avversaria siano del tutto precluse alla sua squadra ? Questo è un altro problema importante e insoluto dell'informatica.

Ecco un esempio di piccolo colpo di fortuna che può capitare alla squadra del no: chiamiamo *singoletto* una parola costituita da una sola lettera; quando due singoletti S_1 ed S_2 sono costituiti dalla stessa lettera che occorre minuscola in S_1 e maiuscola in S_2; (o viceversa) diciamo che sono due singoletti *opposti.* Supponiamo che nell'elenco di parole $p_1,...,p_n$ da intersecare appaiano due singoletti opposti S_1 e S_2. Allora anche il tifoso della squadra del no potrà gridare al goal, perché la squadra del no esibirà S_1 ed S_2 come una prova di non intersecabilità: infatti se una parola p interseca $p_1,...,p_n$, in particolare conterrà sia la lettera di S_1 che quella di S_2 e dunque, per definizione, p non può essere normale.

ESERCIZIO 5 Considera le seguenti parole: I, cIpReSsI, cHe, A, boLgHeRi, AlTi, E, sChIeTtI, vAn, Da, SaN, gUiDo, In, duplice, FILAR, quaSI, In, CORsa, GIGANTI, giovinetti, MI, balzarono, INCONTRO, e, mi, PARLAR. Esse non sono intersecabili da nessuna parola normale. La squadra del no ha fatto goal. E se correggessimo il secondo "e" rendendolo maiuscolo ?

Benché l'esistenza di una coppia di singoletti opposti sia sufficiente per concludere che $p_1,...,p_n$ non è intersecabile, il viceversa non vale necessariamente. Per esempio, l'insieme di parole AB, Ab, aB, ab non è intersecabile (perché ?) pur non possedendo nessuna coppia di singoletti opposti. Con grande delusione dei tifosi della squadra del no, la ricerca di coppie di singoletti opposti è un metodo di prova incompleto. Da dove partiremo per descrivere un metodo completo che permetta anche alla squadra del no di vincere senza fare la ricerca esaustiva ?

Elementi dinamici: la risoluzione

Per rispondere a questa domanda, procedendo per analogia con le tecniche di risoluzione di sistemi di equazioni lineari, cercheremo di trasformare l'insieme di parole dato $p_1,...,p_n$ in un altro, diciamo $q_1,...,q_m$, equivalente al primo, ma più facile da trattare. Questa idea di semplificare un problema senza snaturarlo è una delle più feconde in tutta la scienza, e si trova frequentemente in matematica.

DEFINIZIONE Un insieme di parole $p_1,...,p_n$ è *equivalente* a un altro insieme $q_1,...,q_m$ se ogni parola normale che interseca il primo insieme interseca anche il secondo—e viceversa.

ESERCIZIO 6 Questa relazione è davvero riflessiva, simmetrica, transitiva ?

REGOLA DI RISOLUZIONE Date due parole s e t, supponiamo che una lettera minuscola x sia contenuta in s e la corrispondente maiuscola X sia contenuta in t. Allora il *risolvente* r = R(s,t; x,X) si ottiene cancellando in s tutte le x, cancellando in t tutte le X, e attaccando le due parole così ottenute. In simboli,

R(s,t; x,X) = (s senza le x) + (t senza le X).

ESEMPIO
R(cane, GATTO; a, A) = cneGTTO
R(USUFRuttuario, Ululato; u, U) = USUFRttariolulato
R(dissiMILISSIMI, ASSestasSero; s, S) = diiMILISSIMIAestasero

ESERCIZIO 7 Risolvi coppie di parole scritte con maiuscole e minuscole, scelte a piacere; può una coppia avere più di un risolvente ? può una coppia non avere nessun risolvente ?

ESEMPIO CRITICO Applicando la regola di risoluzione ai due singoletti opposti a, A si ottiene la *parola vuota*, in simboli

R(a, A; a, A) = parola vuota = ø

Come lo zero fu adottato tardivamente nella famiglia dei numeri, per meriti di servizio, anche noi allargheremo la definizione di "parola" accogliendo la parola vuota. Non ci sono problemi ad aggiustare le nozioni precedenti: per esempio, la parola vuota è normale, e non interseca nessuna parola. Vedremo ora che aggiungendo risolventi a un insieme di parole si ottiene un insieme equivalente:

ESERCIZIO 8 Sia dato un insieme di parole $p_1,...,p_n$; sia x una lettera minuscola in p_i, e si supponga che la corrispondente maiuscola X occorra in p_j ($i \neq j$). Allora aggiungendo a $p_1,...,p_n$ il risolvente $r = R(p_i, p_j; x, X)$ si ottiene un insieme equivalente a quello dato.

Risoluzione Indubbiamente se p interseca l'insieme $p_1,...,p_n$, r allora p interseca $p_1,...,p_n$. Viceversa, supponiamo che una parola (normale, non vuota) q intersechi $p_1,...,p_n$. Per costruzione, q non contiene entrambe x ed X.

Caso 1: q non contiene X.

Allora q ha in comune con p_j un simbolo diverso da X; per definizione di $R(p_i, p_j ; x, X)$ questo simbolo è contenuto anche in r. Dunque, q interseca anche r.

Caso 2: q non contiene x.

Si ragiona come sopra, con p_i nel ruolo di p_j.

QED

Aggiungere a $p_1,...,p_n$ un risolvente $r = R(p_i, p_j ; x, X)$ non solo non altera la sostanza del problema dell'intersecabilità, ma spesso sblocca una dimostrazione di non intersecabilità, così facilitando il compito alla squadra del no. Questo succede, per esempio, se $r = ø$. In questo caso la squadra del no presenterà all'arbitro l'identità $ø = R(p_i, p_j ; x, X)$, e l'arbitro accetterà tale identità come prova inconfutabile che $p_1,...,p_n$ non è intersecabile, visto che l'elenco equivalente $p_1,...,p_n$, ø non lo è. Se poi la squadra del sì sporgesse reclamo, l'arbitro ha una moviola infallibile per difendere il suo operato: l'esercizio 8.

Anche senza aver la fortuna di trovare la parola vuota nella prima generazione di risolventi, come figlia di p_i e p_j, la squadra del no può sperare che la parola vuota sia nipote di parole $p_1,...,p_n$, ossia abbia un *albero genealogico* con due livelli di risolventi: risolventi diretti di $p_1,...,p_n$ e risolventi di risolventi. Ad esempio, date le parole AB, Ab, aB, ab, si ottiene la parola vuota alla seconda generazione, in questo modo:

aa = R(ab, aB; b, B),

AA = R(Ab, AB; b,B),

ø = R(aa, AA; a, A) = R(R(ab, aB; b, B), R(Ab, AB; b, B); a, A).

L' identità ø = R(R(ab, aB; b, B), R(Ab, AB; b, B); a, A) può essere pensata come albero genealogico della parola vuota, attestante che essa discende dai nonni AB, Ab, aB, ab attraverso due generazioni-risoluzioni; tale albero genealogico costituisce una prova inconfutabile del fatto che AB, Ab, aB, ab non è intersecabile.

ESERCIZIO 9 Costruisci tre o quattro generazioni nell'albero genealogico di questo insieme di parole:

abc, Abc, aBc, abC, ABc, AbC, aBC,

e vedi se riesci a trovare la parola vuota; aggiungendo ora la parola ABC alle precedenti sette parole, riusciremo a trovare la parola vuota ?

Come ricorderemo, il metodo dei singoletti opposti non riusciva a dare la vittoria alla squadra del no, con gran delusione dei suoi tifosi. Essi ora si chiedono se il metodo risolutivo sia *completo:* in altre parole, per ogni elenco non intersecabile $p_1,...,p_n$ siamo sicuri che costruendo l'albero genealogico di $p_1,...,p_n$ apparirà in qualche generazione successiva, la parola vuota ? La risposta positiva è il seguente teorema di completezza :

TEOREMA Un insieme di parole non è intersecabile se e solo se la parola vuota è ottenibile da esso mediante un numero finito di applicazioni della regola di risoluzione.

Dimostrazione Una direzione è già stata sistemata; l'altra prende solo una ventina di righe, ma forse è meglio ometterla per non turbare l'atmosfera sportiva. La lasciamo come esercizio alla migliore tra la squadra del sì e quella del no.

Appena le parole $p_1,...,p_n$ vengono scritte sul tabellone e la partita comincia, mentre la squadra del sì si lancia alla ricerca di una parola normale q che intersechi $p_1,...,p_n$, la squadra del no costruisce generazioni successive dell'albero genealogico di $p_1,...,p_n$, sperando che tra i discendenti a un certo punto appaia la parola vuota.

In generale non si sa quanto il metodo degli alberi genealogici sia più efficace del metodo di ricerca esaustiva. Tuttavia esistono casi interessanti in cui un forte miglioramento c'è. Questo è l'oggetto dei seguenti due esercizi:

ESERCIZIO SPECIALE 10 (parole di Krom) Quando una parola ha meno di tre lettere si chiama una parola di Krom. Considera questo insieme di parole di Krom:

Ab, cB, Ad, DE, Gh, aD, Eb, bc, h, Ef, FG, gh, Ge.

Valuta se questo insieme sia intersecabile, utilizzando il metodo degli alberi genealogici. Ripeti l'esercizio per altri insiemi di parole di Krom. Osserva come, pur essendoci tante parole e tante lettere in gioco, sia fulmineo decidere se queste parole sono intersecabili. Per spiegarsene la ragione bastano queste osservazioni: (i) il risolvente di due parole Krom è ancora Krom; (ii) la risoluzione non aggiunge nuove lettere; (iii) il numero di parole di Krom con due lettere è eguale al quadrato del numero di lettere disponibili; ancora minore è il numero di parole di Krom con una lettera. Inutile dire che il metodo di ricerca esaustiva procederebbe bovinamente ad esaminare tutti i possibili candidati, con le solite lungaggini.

ESERCIZIO SPECIALE 11 (parole di Horn) Quando una parola ha meno di due lettere maiuscole si chiama parola di Horn. Considera questo insieme di parole di Horn:
cane, A, gatto, Cine, Giottto, O, N, E, neI, cAn, cenE, Oce, iNe, cinE.
Valuta se questo insieme sia intersecabile, utilizzando il metodo degli alberi genealogici. Ripeti l'esercizio per altri insiemi di parole di Horn. Osserva come, pur essendoci tante parole e tante lettere in gioco, anche in questo caso è fulmineo decidere l'intersecabilità. Nota come ogni risolvente di due parole Horn sia ancora una parola Horn. Invece il metodo di ricerca esaustiva tratta un elenco di parole Horn con la solita flemma burocratica.

SIGNIFICATO LOGICO

Le parole cominciano a balbettare

Fedeli al principio di parsimonia, oltre al simbolo di negazione $\neg$, che sta per "non" ci limiteremo a introdurre il simbolo $\wedge$ per la *congiunzione* "e", ed il simbolo $\vee$ per la *disgiunzione* "oppure". Chiamiamo i simboli $\neg$, $\wedge$, $\vee$ *connettivi.*

ESERCIZIO 12 Siano P, Q, R, S abbreviazioni delle seguenti proposizioni:

P = "Jones era in casa Smith alle 21 ieri",
Q = "Smith ha perso le chiavi di casa",
R = "il teste Ross ha visto bene",
S = "vi è una seconda copia della chiavi di casa Smith".

Usando P, Q, R, S e con il solo aiuto dei connettivi $\neg$, $\vee$, $\wedge$, esprimi queste proposizioni composte:

(i) se il teste Ross ha visto bene allora Jones era in casa Smith alle 21 ieri;
(ii) Smith ha perso le chiavi di casa, a meno che il teste Ross non abbia visto male;
(iii) Jones era in casa Smith alle 21 ieri solo se Smith ha perso le chiavi di casa;
(iv) è impossibile che il teste Ross abbia visto bene e che Jones fosse in casa Smith alle 21 ieri;

(v) se Smith non ha perso le chiavi di casa e Jones era in casa Smith alle 21 ieri, allora vi è una seconda copia delle chiavi di casa Smith;
(vi) delle ipotesi che Smith abbia perso le chiavi di casa e che il teste Ross abbia visto bene una e una sola vale.

(*Soluzione parziale*: (i): $\neg R \vee P$; (ii):$\neg R \vee Q$; (iii)$\neg P \vee Q$; (vi): $(Q \wedge \neg R) \vee (\neg Q \wedge R)$)

ESERCIZIO 13 Scrivi tre proposizioni P, Q, R di cui non si sappia se sono vere o false. Forma poi alcune proposizioni composte, come $(P \wedge \neg Q)$, $(P \vee \neg P)$, $(\neg\neg R \vee \neg P)$, $(P \wedge \neg(P \vee Q \vee R))$, $(P \vee \neg P) \vee (R \wedge \neg R)$ e vedi se di qualcuna di esse si possa, nondimeno, dire se è vera o falsa. Ad esempio, se P è un'abbreviazione di "il pianeta Marte è abitato da organismi unicellulari", la proposizione $P \vee \neg P$ è vera.

Non è il caso di scomodare la bio-astronomia per convincersi della verità dell' affermazione "il pianeta Marte è abitato da organismi unicellulari, oppure il pianeta Marte non è abitato da organismi unicellulari". Analogamente, non è il caso di scomodarsi a fare una divisione per convincersi che il numero 555555 è divisibile per 33. Non basterebbe un colpo d'occhio se usassimo la notazione romana. La scelta del modo di scrivere i numeri è tanto importante per fare i calcoli quanto la scelta del modo di scrivere le formule è importante per ragionarvi sopra.

Il punto di partenza sono le "variabili proposizionali"—ma nessuno soffrirà su di esse, come invece molti studenti soffrono quando cominciano a fare calcoli letterali, perché non hanno risposte chiare alle domande "che cosa è una variabile, una variabile indipendente, una variabile dipendente, un'incognita, una costante letterale, un parametro ?".

Le nostre variabili sono invece concrete, inossidabili, alla mano, e quindi facilissime da capire. Infatti, fedeli alla politica delle mani contenutisticamente pulite, chiameremo *variabili proposizionali* nient'altro che le lettere maiuscole dell'alfabeto A, B, C,...Z . Quando avremo bisogno di una 22-ma, 23-ma,... variabile proposizionale, daremo le indicazioni necessarie, ma forse allora non sarà necessario, perché avremo capito tutto.

Le proposizioni, o formule, composte si ottengono applicando i connettivi di negazione $\neg$, congiunzione $\wedge$, disgiunzione $\vee$, secondo le ferree istruzioni che stiamo per dare. I più semplici esempi di formule sono le variabili proposizionali e le loro negazioni $\neg A$, $\neg B$, $\neg C$,...:

♦ Chiamiamo *letterale* una variabile proposizionale oppure la sua negazione.

Utilizzando i letterali come mattoni da costruzione, semplici formule si ottengono inserendo tra due letterali consecutivi il simbolo $\vee$, e scrivendo per esempio $A \vee \neg R \vee S \vee \neg M \vee A \vee \neg M$:

♦♦ Chiamiamo *clausola* una disgiunzione di letterali.

Utilizzando infine le clausole come mattoni da costruzione, altre formule si ottengono inserendo tra due clausole consecutive il simbolo $\wedge$, e scrivendo ad esempio $(A \vee B) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$; benché non necessarie,le parentesi sono messe per una forma di rispetto ortografico verso l'occhio:

♦♦♦ Chiamiamo *formula CNF* una congiunzione di clausole.
CNF sta per "forma normale congiuntiva".

Qui c'è un abbozzo di analogia:

letterale	variabile
clausola	monomio
formula CNF	polinomio

Introdurremo subito un'analogia ben più utile, che ci permetterà di progredire nella trattazione della verità logica,rivedendo sotto nuova luce il gioco delle parole intersecanti.

Quid est veritas ?

Verità, falsità, conseguenza—ecco alcuni temi tradizionali della logica. Già abbiamo visto che, pur senza sapere se su Marte ci sia vita, una proposizione come "il pianeta Marte è abitato da organismi unicellulari oppure il pianeta Marte non è abitato da organismi unicellulari" è vera. Questa apparente banalità è conseguenza di alcune importanti definizioni, che costituiscono la *logica proposizionale booleana*:

1. I possibili valori di verità di ogni proposizione sono due: 1=vero e 0=falso. (Questa è la *definizione di valore di verità*)
2. Se una proposizione è vera, automaticamente la sua negazione è falsa, e viceversa. (Questa è la *definizione della negazione*);
3. Il valore di verità di una disgiunzione è il massimo dei valori di verità dei disgiunti (Questa è la *definizione di disgiunzione*).
4. Il valore di verità di una congiunzione è il minimo dei valori di verità dei congiunti (Questa è la *definizione di congiunzione*).

Da 1 e 2 segue che, indipendentemente da ogni considerazione bio-astronomica, sarà vera una e una sola delle due proposizioni "il pianeta Marte è abitato da organismi unicellulari " oppure "il pianeta Marte non è abitato da organismi unicellulari ". Da 3 segue allora che la proposizione "il pianeta Marte è abitato da organismi unicellulari oppure il pianeta Marte non è abitato da organismi unicellulari " è vera.

Le stipulazioni 1-4 hanno carattere di massima generalità: per qualsiasi proposizione Z esse sanciscono d'ufficio la verità della proposizione $Z \vee \neg Z$. E' utile notare che variabili, letterali, clausole, formule CNF rimangono, e rimarranno sempre, successioni di simboli. All'inizio delle lezioni abbiamo notato come spesso ci si lasci prendere la mano dai simboli e dalla loro paradossale e familiare concretezza, soprattutto per gli effetti perversi dell' inflazione di parole come "contenuto espressivo", "semantica", "significato oggettivo", che invece richiedono una drastica amministrazione matematica. E in effetti per rendere totalmente asettiche le stipulazioni 1-4 e dare un chiaro segnale di rigore nel trattare concetti così importanti, la logica matematica ha prima di tutto abolito ogni terminologia altisonante, sostituendola con le nozioni di assegnazione, soddisfazione, tautologia:

DEFINIZIONE Un'*assegnazione* è una funzione μ definita su un insieme di variabili proposizionali e avente valori nell'insieme $\{0,1\}$. Diciamo poi che

♣ μ *soddisfa la variabile proposizionale* P se $\mu(P)=1$
♣♣ μ *soddisfa il letterale negato* $\neg P$ se $\mu(P)=0$
♣♣♣ μ *soddisfa la clausola* C se μ soddisfa almeno uno dei letterali in C
♣♣♣♣ μ *soddisfa la formula CNF* F se μ soddisfa tutte le clausole di F.

NOTE Questa è una tipica definizione a (per fortuna, non più di quattro) scatole cinesi, in cui prima di interiorizzare la riga n+1 il lettore deve aver appreso perfettamente le righe 1,2,...,n.

In questa definizione si ipotizza tacitamente che il dominio di definizione di μ sia abbastanza vasto da comprendere tutte le variabili che occorrono nella formula F. La definizione appena data fissa, una volta per tutte, il significato, la semantica delle variabili proposizionali e dei connettivi nella logica booleana. Per completare il discorso sulla verità si dà ora la seguente

DEFINIZIONE Una formula CNF F è *universalmente vera* (F è una *tautologia*) se ogni assegnazione π la soddisfa. F è *soddisfacibile* se almeno una assegnazione μ la soddisfa; altrimenti F è *insoddisfacibile.*

ESERCIZIO 14

(a) Ci sono sedici possibili assegnazioni per le proposizioni P,Q,R,S dell'esercizio 12. Scrivi queste assegnazioni. Per ciascuna di esse controlla la verità o falsità delle affermazioni (v) e (vi).

(b) Sia π la seguente assegnazione: $\pi(A)=1$, $\pi(B)=1$, $\pi(C)=1$, $\pi(D)=\pi(E)=\ldots=\pi(Q)=0$. Quali delle seguenti sei formule CNF sono soddisfatte da π ?

$(C \vee D \vee \neg B) \wedge (\neg C \vee \neg G \vee B) \wedge (\neg D \vee \neg A \vee B)$

$(\neg C \vee A \vee \neg B) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$

$(\neg A \vee B \vee \neg G \vee \neg H \vee \neg M \vee \neg K) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$

$(\neg A \vee \neg B) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$

$(A \vee H) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg G)$

$(P \vee \neg Q) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg E \vee \neg B) \wedge (A \vee B) \wedge (\neg C \vee \neg G \vee B)$

NOTA L' esercizio 14 (b) contiene un errore, perché promette sei, ma scrive solo cinque formule CNF. In una formula infatti appare una variabile proposizionale clandestina. Ma chi se ne è accorto ? E allora, promossi per disattenzione, abbiamo capito che tutto quanto definito finora vige immutato se l'insieme delle variabili proposizionali viene arricchito da un 22-mo, 23-mo,... simbolo. Anche se questo ampliamento concedesse la cittadinanza a un'infinità di nuove variabili proposizionali, non ci troveremmo di fronte a una rivoluzione concettuale—nemmeno i progettisti di tastiere di computer farebbero una piega, asserendo anzi che le con le tastiere in commercio si scrive benissimo ogni nuova variabile proposizionale passata, presente e futura. I domini di definizione delle assegnazioni avranno ora più di 21 elementi—o addirittura un'infinità di elementi. Ciò diverrà importante, per esempio, per dispiegare un quantificatore universale come una congiunzione infinita. Ma in queste nostre pagine le assegnazioni manterranno sempre dominio finito.

Intersecabilità = soddisfacibilità

Metteremo ora a frutto l'esperienza accumulata con il gioco delle parole intersecanti, utilizzando il metodo degli alberi genealogici come un potente metodo di calcolo deduttivo. Visto che, per definizione, le variabili proposizionali sono state identificate con le lettere maiuscole A,B,C,... procediamo così:

(i) Per sbarazzarci del connettivo di negazione, scriveremo a,b,c,... invece di $\neg A$, $\neg B, \neg C, \ldots$

(ii) Per analogia col trattamento riservato al simbolo di moltiplicazione, eliminiamo il connettivo $\vee$, identificando per esempio la clausola $A \vee \neg R \vee S \vee \neg M \vee A \vee \neg M$ con la parola ArSmAm.

(iii) A questo punto, scrivendo una virgola al posto del connettivo $\wedge$, una congiunzione di clausole diviene la stessa cosa che un elenco di parole.

Ecco un esempio di questa metamorfosi (i)-(iii):

—data la formula CNF $(A \vee B) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$

—per eliminazione del connettivo $\neg$ abbiamo : $(A \vee B) \wedge (c \vee a \vee B) \wedge (d \vee a \vee B) \wedge (c \vee b \vee b)$

—per eliminazione dei connettivi $\vee$ e $\wedge$ otteniamo le parole: AB, caB, daB, cbb

Visto che sulle formule CNF dobbiamo fare dei calcoli senza distrarci a considerare il loro contenuto semantico, la scelta del sistema di notazione diviene importante; ben sappiamo come sia diverso insegnare a fare l' addizione o la moltiplicazione usando la notazione posizionale oppure quella romana. Anche ora, potendo scegliere tra la farfalla AB, caB, daB, cbb e il bruco $(A \vee B) \wedge (\neg C \vee \neg A \vee B) \wedge (\neg D \vee \neg A \vee B) \wedge (\neg C \vee \neg B \vee \neg B)$, non avremo esitazioni.

ESERCIZIO 15 Siano P,Q,R,S le stesse proposizioni dell'esercizio 12. Considera l'elenco di parole Pqr, pQS, Qrs; scrivi la formula CNF corrispondente; rileggi per esteso la formula nell' italiano parlato.

Prendiamo ora un' assegnazione arbitraria μ: facciamo *corrispondere* a μ quella parola m che ha per maiuscole le variabili proposizionali a cui μ dà valore 1, e per minuscole quelle a cui μ dà valore 0. Per pignoleria o per amor di univocità, mettiamo in ordine alfabetico le lettere di m. Visto che μ è una funzione avente dominio finito, segue che m è una parola normale. Viceversa, data una parola normale m, è facile costruire l' assegnazione μ corrispondente a m.

ESEMPIO L'assegnazione $\mu(A)=1$, $\mu(B)=0$, $\mu(C)=1$, $\mu(D)=\mu(E)=\mu(F)=0$ corrisponde alla parola normale AbCdef. La parola normale AcDE corrisponde all'assegnazione $\pi(A)=1$, $\pi(C) = 0$, $\pi(D) = \pi(E) = 1$.

ESERCIZIO 16 A coronamento della metamorfosi costruiamo una perfetta analogia tra il gioco delle parole intersecanti e il problema di soddisfacibilità di formule CNF:
sia F una tale formula e siano $C_1,\ldots,C_n$ le sue clausole, trascritte come parole $p_1,\ldots,p_n$; sia μ un' assegnazione e sia m la parola normale corrispondente a μ. Dimostra che:

"μ soddisfa C_i" è equivalente a "m interseca p_i"
"μ soddisfa F" è equivalente a "m interseca $p_1,\ldots,p_n$ ".

ESERCIZIO 17 Per analogia con quanto fatto per il gioco delle parole intersecanti:
Definisci la nozione di equivalenza logica di due formule CNF.
Enuncia la regola deduttiva corrispondente alla regola di risoluzione.
Formula e dimostra la proposizione che afferma la correttezza della regola di risoluzione.
Formula il teorema di completezza, che garantisce il monopolio della soddisfacibilità di formule CNF ad una sola regola deduttiva, la risoluzione.

Forse alcuni dei lettori più familiarizzati con il metodo delle tavole di verità si chiederanno a che cosa possa servire il metodo alternativo costituito dagli alberi genealogici—metodo che,

per di più è applicabile solo a formule preconfezionate. *Risposta*: il metodo delle tavole di verità corrisponde alla ricerca esaustiva, il più burocratico metodo che si conosca per accertare se una formula è soddisfacibile. Nessuno sa in generale quanto il metodo risolutivo sopravanzi per efficacia il metodo esaustivo: ma come mostrato dagli esercizi 10 e 11 vi sono casi in cui la superiorità è sicura. In particolare, gli insiemi di clausole Horn sono importanti nella logica dei predicati, dove stanno alla base della programmazione logica.

Ecce calculus ratiocinator

Le nozioni di clausola, forma normale congiuntiva, risoluzione, soddisfacibilità, completezza da noi introdotte solo per le formule CNF del calcolo proposizionale booleano sono estendibili alla logica dei predicati, come ampiamente discusso in altre lezioni di questo testo, a cui rimandiamo. Anche la regola di risoluzione, con opportuni adattamenti che non ne intaccano la semplicità, continua ad essere completa nella logica dei predicati. Dunque imparare a fare alberi genealogici con insiemi di clausole predicative significa dotarsi di un armamentario dimostrativo infallibile e semplice per ottenere tutti i teoremi. Così semplice che anche le macchine possono padroneggiarlo. Anziché appesantirci con dettagli tecnici, proviamo a svolgere alcuni esercizi autoesplicativi:

ESERCIZIO 18 Ogni uomo è mortale;

(ossia, parlando in clausole: per ogni x, o x non è uomo, oppure x è mortale)

Socrate è uomo;

dunque:

Socrate è mortale.

Soluzione Utilizzando due simboli di relazione U ed M, e un simbolo di costante s, scriviamo le due clausole predicative:

{¬Ux,Mx}

{Us}

Seguendo il *metodo refutazionale,* scriviamo ora la negazione della tesi che si vuol dimostrare

{¬Ms} (ipotesi assurda)

Puntiamo ad ottenere una contraddizione, la clausola vuota, utilizzando l'unica regola deduttiva a nostra disposizione—la risoluzione. Dando valore s alla variabile x nella clausola {¬Ux,Mx} otteniamo, come caso particolare, {¬Us, Ms}. Per risoluzione di questa clausola con {Us} otteniamo la clausola {Ms}. Risolvendo {Ms} con l'ipotesi assurda {¬Ms} otteniamo la desiderata contraddizione. In simboli,

R({¬Ux,Mx},{Us}; ¬Us, Us) = {Ms}

R({¬Ms},{Ms};{¬Ms},{Ms}) = ø

Commenti: in questa deduzione nullo è stato il ruolo di Socrate, della mortalità e dell'umanità; infatti abbiamo dimostrato anche che "se ogni unionista è musicale, e Smith è unionista, allora Smith è musicale", e inoltre che "se tutti gli ungulati sono monocordi, e Sincrate è ungulato, allora Sincrate è monocorde". Chissà che non abbiano qualche fondamento le critiche malevole contro la logica, per il suo distacco formalistico. E, rimanendo formalistici, e individuati i *letterali* Ux, Mx, Ms, Us, ¬Ux, ¬Mx, ¬Ms, ¬Us, notiamo come ogni clausola in questo esercizio è una *clausola Horn*, nel senso che in essa non vi sono due o più letterali privi di negazione.

ESERCIZIO 19 Pico è monocellulare;

la mamma di ogni essere monocellulare è monocellulare;

dunque:

la nonna materna di Pico è monocellulare.

Soluzione Introduciamo un simbolo di funzione m(z) per la mamma di z; ovviamente, m(m(z)) sarà la mamma della mamma di z.

{Mp}

{¬Mx, Mm(x)}

{¬Mm(m(x))}

Fedeli al metodo refutazionale, abbiamo appena negato che la nonna materna di Pico sia monocellulare; speriamo di ricavare contraddizione dalle tre clausole a nostra disposizione; mettendo p al posto di x abbiamo:

R({Mp},{¬Mx, Mm(x)};{Mp},{¬Mp}) = {Mm(p)}

R({Mm(p)},{¬Mx, Mm(x)};{Mm(p)},{¬Mm(p)}) = {Mm(m(p))}

R({Mm(m(p))}, {¬Mm(m(x))};{Mm(m(p))}, {¬Mm(m(p))}) = ø

Individua i letterali di questo esercizio. Individua le clausole. Sono tutte clausole Horn ?

ESERCIZIO 20 (trascrivere in clausole, e dedurre mediante risoluzione)

Ogni avvocato nell'isola di Utopos difende coloro che non si difendono da sé, e non difende chi si difende da sé;

dunque:

non ci sono avvocati nell'isola di Utopos.

ESERCIZIO 21 (trascrivere in clausole, e dedurre mediante risoluzione)
ogni marziano è verde;
Tico è blu;
nessuno è verde e blu;
dunque:
Tico non è marziano.

ESERCIZIO 22 (trascrivere in clausole, e dedurre mediante risoluzione)
ogni docimologo è stonato;
ogni ornitologo è intonato;
dunque
non esistono docimologi ornitologi.

ESERCIZIO 23 (più difficile; trascrivere in clausole, e dedurre mediante risoluzione)
la relazione binaria U è simmetrica (se Uxy allora Uyx);
la relazione U è transitiva (se Uxy e Uyz allora Uxz);
la relazione U è piena (per ogni x c'è un y—che con un po' d'arbitrarietà chiameremo f(x), non nascondendoci la sua dipendenza da x—tale che Uxf(x));
dunque:
la relazione U è riflessiva (per ogni x, vale Uxx).
Nota che tutte le clausole anche di questo esercizio sono Horn.

ESERCIZIO 24 (trascrivere in clausole, e dedurre mediante risoluzione)
ogni marziano ha un figlio verde;
ogni figlio di un marziano è marziano;
dunque:
se ci sono marziani, allora ci sono marziani verdi.

ESERCIZIO 25 (non solo dedurre, ma anche risolvere equazioni a incognite greche)
Premesse
Platone disprezza chi disprezza l'aritmetica (detto altrimenti, chiunque sia x, se x disprezza l'aritmetica, allora Platone disprezza x; ancora, in altre parole, o x non disprezza l'aritmetica, o Platone disprezza x);
Gorgia disprezza la retorica;
Gorgia disprezza l'aritmetica.
Problema
ammesso e non concesso che Platone disprezzi qualcosa o qualcuno, trovarlo;

Soluzione
{¬Dxa, Dpx}
{Dgr}
{Dga}

prima di tutto vediamo se la soluzione c'è; come prescritto dal metodo refutazionale, aggiungiamo la clausola

{¬Dpy}

la quale dice che Platone non disprezza alcuno; speriamo ora di ricavare una contraddizione; con alcune sostituzioni appropriate di costanti al posto di variabili otteniamo:

R({¬Dpy} , {¬Dxa,Dpx}, {¬Dpy}, {Dpy}) = {¬Dya}
R({¬Dya}, {Dga}; {¬Dga}, {Dga}) = ø

Abbiamo ricavato una contraddizione (la potevamo ricavare anche se non avessimo saputo che Gorgia disprezza la retorica); dunque esiste qualcuno che Platone disprezza. Abbiamo avuto il via libera per ricercare una soluzione. A questo scopo, prepariamo un nuovo simbolo di predicato E; leggiamo Ey come "È y (una soluzione del nostro problema)!"; anziché scrivere la clausola {¬Dpy} scriviamo la clausola più realistica {¬Dpy, Ey}; essa prende atto che chiunque sia y, se Platone disprezza y, allora y è una soluzione al nostro problema. Procedendo con il metodo risolutivo succede che y, il quale attualmente ha la natura di variabile, viene ad incarnarsi in una...costante. Più precisamente, con il solito metodo risolutivo otteniamo:

R({¬Dpy, Ey} , {¬Dxa,Dpx}; {¬Dpy}, {Dpy}) = {¬Dya, Ey}
R({¬Dya, Ey},{Dga}; ¬{Dga},{Dga}) = {Eg}

ossia è Gorgia una soluzione al problema. La sintassi della logica predicativa tiene ben distinti i simboli di variabile x,y,z,... da quelli costante a,b,c,s,g,...

ESERCIZIO 26 (non solo incognite greche, ma anche le solite incognite)
Vediamo come il metodo risolutivo possa servire anche per calcolare funzioni numeriche. Ci concentriamo su una casella della tabellina dell'addizione. Con le tecniche della programmazione logica, mostreremo che tre più due fa cinque (da cosa nasce cosa).
Soluzione Introduciamo un predicato Axyz che dice "x più y fa z"; introduciamo un simbolo di funzione s(x) per il successore di x; introduciamo un simbolo di costante o (per lo zero). E così abbiamo il modo (alquanto bovino) di scrivere qualunque

numero naturale: per comodità scriveremo ssso, invece di s(s(s(o))), pensando al numero tre. Per fare le somme basteranno queste due clausole:

{Axox}
{¬Axyz, Axs(y)s(z)}

Non è difficile capire che cosa dicono.
Neghiamo che esista soluzione al problema:

{¬Asssossoz}

Cerchiamo di ricavare una contraddizione dalle tre clausole

{Axox}, {¬Axyz, Axs(y)s(z)}, {¬Asssossoz}
Con un po' di pazienza, fai saltare fuori tale contraddizione applicando la regola di risoluzione.

Avendo avuto semaforo verde, andiamo ora alla ricerca di una soluzione; preparato un nuovo predicato Eu a cui far dire che u è soluzione del problema, scriviamo la clausola

{¬Asssossoz, Ez},

la quale afferma che ogni z che faccia tre più due è soluzione al problema.
Con alcune applicazioni della regola risolutiva, a partire dalle clausole

{Axox}, {¬Axyz, Axs(y)s(z)}, {¬Asssossoz, Ez}

ricaviamo la clausola

{Essssso},

che ci informa che cinque è soluzione al nostro problema.
Commento Questo esercizio ci può lasciare un po' di amaro in bocca. Con qualche giro di parole abbiamo aggirato i nostri handicap espressivi allorché abbiamo dovuto esprimere le proprietà fondamentali dell'addizione: x+0 = x, e x + successore(y) = successore(x+y), senza avere il simbolo di eguaglianza. Ma come fare per dedurre che cinque è l'unica soluzione al nostro problema ?

Siamo arrivati al punto giusto per apprezzare il problema dell' eguaglianza nella logica dei predicati: consultando le altre lezioni di questo corso, i lettori vedranno che il problema ha una soluzione soddisfacente, al punto che oggi il simbolo = è incorporato nel

calculus ratiocinator—un calculus ratiocinator con qualche regola deduttiva in più rispetto al prototipo senza eguaglianza.

Nello stesso tempo, stante la semplicità di questo esercizio, tutto costituito da risoluzioni su clausole Horn, a partire da {Axox}, {¬Axyz, Axs(y)s(z)}, {¬Asssossoz} viene da chiedersi: è mai possibile che per accertare matematicamente che 3+2 = sssssssso *non* è ottenibile da queste clausole (mentre abbiamo appena ottenuto 3+2 = ssssso) occorra scomodare niente meno che una revisione costituzionale del calculus ratiocinator in senso egualitario ?

Questo è un problema importante della programmazione logica. Ai lettori più attenti, e a quelli più sensibili alle analogie potrà ritornare in mente la squadra del no, con le sue difficoltà a trovare dimostrazioni di non esistenza. Qui la partita è ancor più dura, continua dopo il goal {Essssso} della squadra del sì, senza la garanzia di un arbitro che, finita la sua ricerca esaustiva, mandi tutti a casa. Ma questo esercizio è come un provino: il film è molto più interessante.

MODELLI PER LA COMPUTAZIONE

Stilizziamo il nastro delle macchine di Turing identificandolo con l'insieme **Z** = {..., –2, –1, 0, 1, 2,...} degli interi. Economizziamo sulla tastiera-alfabeto, supponendo che contenga solo due simboli, il "blank" ø e il "non blank" |. Allora una *foto di nastro* altro non è che una funzione f : **Z**→ {ø, |}; naturalmente, f(y) è il simbolo stampato nella casella y. Una *configurazione* di macchina di Turing deve contenere le seguenti informazioni: una foto di nastro f, lo stato s in cui si trova la macchina, la posizione x del pennino ottico sul nastro. Dunque una configurazione è una terna ordinata (f, s, x). Note le quintuple della macchina e nota una configurazione, c'è un unico modo di definire ragionevolmente la configurazione successiva.

ESERCIZIO 27 (configurazioni)
Diamo nome T1 alla macchina le cui quintuple sono

[ø,0,ø,S,1] e [|,0,ø,D,0].

Supponi che la foto di nastro f riveli sette simboli nonblank consecutivi, in un nastro altrimenti pieno di blank. Poni il pennino ottico sul primo nonblank a sinistra, e poni T1 nello stato 0. Qual è la configurazione successiva ? Descrivi le configurazioni successive fino all'arresto della macchina.

ESERCIZIO 28 (Turing-programmazione, upgrade)
Scrivi le quintuple di una macchina di Turing T2 che computi la funzione successore, in questo senso: se sul nastro vengono scritti n simboli nonblank consecutivi ($n \geq 1$), e il pennino ottico di T2 è messo sul primo di essi, in stato 0, allora T2 passa in rassegna questi simboli senza cancellare nulla, poi ne aggiunge uno immediatamente alla destra dell'ultimo, e poi si ferma. Dopo aver programmato T2, supponi di dover calcolare anche il successore di zero. Devi fare un "upgrade" (un potenziamento) della tua T2, o basta la versione precedente ? Supponi ora di dover garantire che la macchina ritorni all'inizio dell'input. Fai l'upgrade.

ESERCIZIO 29 (Turing-programmazione delle funzioni identità e della funzione somma)
Dimostra la Turing-calcolabilità della seconda funzione identità a tre posti, ossia la funzione $I^2_3(x, y, z) = y$; in dettaglio, programma una macchina di Turing che avendo in input
x simboli nonblank | seguiti da
un blank ø, seguito da
y simboli |, seguiti da
un blank ø, seguito da
z simboli |
dà in output solamente gli y simboli |. Supponi $x, y, z \geq 1$. Programma una macchina di Turing T3 che avendo in input m nonblank seguiti da un simbolo blank di separazione ø e poi da n nonblank ($m,n \geq 1$), dà in output m+n nonblank consecutivi. Che upgrade si possono fare per estendere l'azione di T3 anche al caso $m, n = 0$?
Suggerimento La scelta di estrema povertà del nostro alfabeto ci può metter in condizione di dover scrivere i numeri in maniere un po' strane; anche se decidessimo di scrivere lo zero come |, l'uno come ||, il due come |||, eccetera, non cambierà la nozione di Turing-calcolabilità.

ESERCIZIO 30 (pedigree)
Esprimi funzioni come x!, xy, x^y ,... partendo dalle funzioni di base (addizione, successore, la funzione identità x, la costante zero) utilizzando solamente le operazioni di composizione e ricorsione primitiva

ESERCIZIO 31 (diagonale)
L'insieme delle funzioni primitive ricorsive non è enumerabile da una funzione primitiva ricorsiva.
Soluzione Ricordiamo le caratteristiche e le notazioni della nostra enciclopedia-pacco. Se, per ipotesi assurda, r* fosse primitiva ricorsiva, la troveremmo diciamo a pagina x*; dunque le funzioni r* e p_x* sono eguali; per tutti gli x abbiamo

$p_x^*(x) = r^*(x) = p_x(x) + 1$,

e in particolare, per $x = x^*$

$p_x^*(x^*) = r^*(x^*) = p_x^*(x^*) + 1$,

contraddizione.

ESERCIZIO 32 Enumera le funzioni parziali ricorsive.
Soluzione Visto che l'alfabeto consiste dei due soli simboli ø e |, se il numero di stati è q, segue che l'insieme Q delle quintuple possibili ha esattamente $8q^2$ elementi. Siccome ogni programma di macchina di Turing è un sottoinsieme di Q, esiste solo un numero finito di macchine di Turing con q stati; se necessario, non avremmo problemi ad elencarle tutte in ordine alfabetico. Se poi accresciamo q aggiungendo un elemento alla volta, potremo facilmente scrivere *tutte* le possibili macchine di Turing, senza ometterne alcuna.

ESERCIZIO 33 (l' halting non è Turing-computabile)
In questo esercizio converrà allargare l'alfabeto delle macchine di Turing in modo che contenga tutti i simboli con cui si scrivono le quintuple. Un tipico alfabeto adatto a questo scopo è A = {|, ø,], [, S, D, 0, 1}. Usiamo il blank ø come sostituto della virgola; scriviamo gli stati come numeri in notazione binaria. Chiamiamo una macchina M "autostoppista" se M ha questo comportamento, per qualsiasi input x:

—nel caso (eccezionalissimo) che x coincida con l'elenco delle sue proprie quintuple, M ne prende atto fermandosi dopo un certo numero di passi;
—in caso contrario, M prosegue indefinitamente senza mai fermarsi.

Dimostrare che non esiste una macchina di Turing T che avendo in input le quintuple di una qualsiasi macchina N, si ferma se e solo se N *non* è autostoppista.
Suggerimento Usa un argomento di diagonalizzazione. Come si comporterà T avendo in input l'elenco delle sue proprie quintuple ?

ESERCIZIO 34 (certificazione)
Sia T4 la seguente macchina di Turing

[|,0,|,D,0]
[ø,0,|,D,1]
[ø,1,ø,S,2]

Supponi che T4 sia accompagnata dal seguente certificato: se posizionata in stato 0 all'inizio dell'input, costituito da n simboli nonblank consecutivi, ($n \geq 0$), in un nastro per il resto pieno di blank, T4 aggiunge un simbolo nonblank alla fine, e si ferma in stato 2 su quest'ultimo simbolo.
È corretta questa certificazione ?
Dimostra, per induzione su $n \geq 1$, che: se l'input contiene almeno n simboli nonblank consecutivi, allora nell'n-ma configurazione il pennino ottico è posizionato sull'n-mo di essi, con la macchina T sempre in stato 0.

Sia T5 questa macchina di Turing

[|,0,|,D,0]
[ø,0,ø,S,1]
[|,1,ø,S,2]

Certifica che questa macchina computa la funzione predecessore (qui per definizione, il predecessore di 0 è eguale a 0).

ESERCIZIO 35 (la classe NP)
Per render capace una macchina di Turing T di inventare soluzioni di problemi, permetteremo a T di scegliere "nondeterministicamente" quale azione intraprendere e in quale stato successivo andare. Abroghiamo la restrizione secondo cui al più una quintupla di T comincia con [a, s, ...]. In questo modo, ogni configurazione potrà avere un numero finito di configurazioni successive. La computazione di T su input x non procede più lungo una successione di configurazioni, ma procede invece biforcandosi: le configurazioni si collocano lungo un albero: alla radice sta la configurazione iniziale, che si dirama in configurazioni successive, ciascuna delle quali si dirama, e così via. Diremo che T accetta x se nell'albero c'è almeno un cammino di configurazioni successive, che partendo dalla configurazione iniziale termina su una configurazione di fermata. Chiamiamo tale cammino *cammino d'accettazione per x*. Diciamo che *un problema L è di classe NP* se esistono un polinomio r e una macchina nondeterministica T che avendo in input una stringa x di lunghezza n si comporta così:

—se x è soluzione di L, T ha un cammino d'accettazione per x con meno di r(n) passi
—se x non è soluzione di L, T non ha nessun cammino d'accettazione per x.

Il seguente esercizio, più che la scrittura delle quintuple di macchine nondeterministiche, richiede una presa di coscienza che tali quintuple possono essere scritte:

Dimostrare che i seguenti problemi sono in NP: soddisfacibilità di formule CNF, riconoscimento di numeri composti, problema dei francobolli, problema delle parole intersecanti.
Suggerimento Prendiamo il problema del riconoscimento dei numeri composti. Ecco una veloce procedura nondeterministica per riconoscere se un numero è composto:

(i) indovinare un fattore, come fece Euler, e
(ii) velocemente controllare, facendo la divisione, che esso è fattore.

Avvertimento: Al lettore convinto che, con questo tipo di argomentazione, si riesca a dimostrare che ogni problema è in NP. Cercare di applicare l'argomento ai seguenti problemi: riconoscimento delle formule *in*soddisfacibili; *Entscheidungsproblem.*

ESERCIZIO 36 (un caso facile di un problema difficile)
Supponi, come nel problema dei francobolli, di avere un numero A (l' affrancatura) e un insieme C di francobolli $b_1 \leq b_2 \leq ... \leq b_n$, con la seguente domanda: *È possibile ottenere l'affrancatura A scegliendo oculatamente tra i francobolli a disposizione ?* In generale questo problema è così complicato come il problema dell'intersecabilità (si conoscono riduzioni fulminee dell'uno all'altro). Come il caso Horn per il problema di soddisfacibilità, anche il problema dei francobolli ha dei casi facili, che, opportunamente mascherati in modo da sembrare difficili, sono usati nella crittografia.

Un caso facile si ha, per esempio, supponendo che la collezione C dei francobolli sia *supercrescente,* ossia

$b_2 > b_1$
$b_3 > b_2 + b_1$
$b_4 > b_3 + b_2 + b_1$
... ...
$b_n > b_1 + b_2 + ... + b_{(n-1)}$.

Trova un metodo fulmineo per decidere se l'affrancatura A è realizzabile.
Abbozzo di soluzione Come se stessimo scrivendo un galateo durevole di comportamento con il Ministro delle Poste, che non vada al macero dopo la prima crisi di governo, cominciamo con chiamare A l'affrancatura *pro tempore*. Questo è un modo elegante per dire che non ci aspettiamo che A resti in vigore per molto tempo, ma non vogliamo nemmeno scomodarci a scrivere A_1, A_2, A_3,... per la sfilza delle affrancature pro tempore che seguiranno A, e preferiamo usare il singolare "l'affrancatura pro tempore", come se questa o quella affrancatura per noi pari fossero.

Disponiamo ora i francobolli della collezione C sul tavolo, in fila, in ordine di valore crescente $b_1 < b_2 < ... < b_n$. Arrivati in fondo alla fila, ritorniamo indietro, esaminando i francobolli uno dopo l'altro a partire da quello di valore massimo.

L'esame consiste in questo: se un francobollo ha valore superiore a quello dell'affrancatura pro tempore, mettiamolo nel frigorifero, e proseguiamo; se un francobollo ha valore minore o eguale a quello dell'affrancatura pro tempore, facciamo una piccola pausa per eseguire queste due operazioni:

(uno) mettiamo questo francobollo in uno scrigno;

(due) sottraiamo all' affrancatura pro tempore il valore di questo francobollo, e chiamiamo il risultato affrancatura pro tempore. Nello stesso istante, la vecchia affrancatura pro tempore va nel dimenticatoio, e perde ogni titolo e ufficio postale.

Per fissare le idee, pensiamo a una serie di crisi di governo in un paese ideale in cui ogni neoministro (pro tempore) delle poste fa abbassare le tariffe postali. Ebbene, procedendo in questo modo da affrancatura a affrancatura, in pochi passi si arriva a un aut aut:

—o l'affrancatura pro tempore diviene eguale a zero, e allora i francobolli nello scrigno risultano formare l'affrancatura A di partenza, mentre gli eventuali francobolli rimasti da esaminare finiranno tutti nel frigorifero in quanto hanno valore superiore all'affrancatura pro tempore,

—oppure l'affrancatura pro tempore rimane sempre maggiore di zero, anche quando tutti i francobolli di C sono ormai stati ripartiti tra frigorifero e scrigno.

Allora potremo tranquillamente dichiarare che l'affrancatura A è irrealizzabile. Infatti valgono queste due affermazioni:

(1) una collezione C supercrescente ammette al più un sottoinsieme S di francobolli che realizzano l'affrancatura A

(2) il nostro metodo uno-due è infallibile a trovare tale S, quando esiste.

Ora tocca a te, lettrice-lettore resistente. Fai alcuni esempi concreti di insiemi C di francobolli supercrescenti.

Esempio Supponi che $b_1 = 10$ lire, $b_2 = 100$ lire, $b_3 = 1000$ lire, $b_4 = 10000$ lire e $b_5 = 100000$ lire. Verifica che questa collezione è supercrescente. Quali di queste affrancature sono realizzabili ?

$A = 10010$ lire

$A' = 10011$ lire

$A'' = 101010$ lire

$A^{\wedge} = 110010$ lire

$A^{\circ} = 1111001$ lire

$A^{*} = 20000$ lire

Per ciascun esempio esegui la procedura uno-due descritta sopra, e verifica l'aut aut finale. Nota una volta per tutte che se l'affrancatura pro tempore diviene zero, allora i francobolli nello scrigno realizzano precisamente l'affrancatura A. Dimostra le affermazioni (1) e (2).

Suggerimento L'affermazione (1) è una conseguenza diretta delle definizione di supercrescenza, e si dimostra per assurdo. Per dimostrare l'affermazione (2), supponi che esista un insieme C supercrescente e un'affrancatura A realizzabile con un sottoinsieme S di C. Mettiamo i francobolli di C e di S in ordine decrescente e, passo dopo passo, usando l'induzione matematica, controlliamo che davvero il metodo uno-due metta nello scrigno i francobolli di S e solo questi.

ESERCIZIO 37 (riduzioni)

Supponiamo di aver bisogno di una macchina di Turing che riconosca i numeri primi, scritti nella solita notazione decimale. Supponiamo che sul mercato siamo riusciti a trovare solamente un riconoscitore R di *successori di numeri primi.* In dettaglio, R si comporta così: se sul nastro scriviamo 3, 4, 6, 8, 12,... la macchina R pulisce il nastro e scrive il simbolo | in output, mentre scrive il simbolo ø quando in input abbiamo scritto 0, 1, 2, 5, 9, 10, 11,..., che non sono successori di primi. Notiamo che la macchina R può essere facilmente utilizzata per riconoscere numeri primi in questo modo: dato in input il numero y trasformiamolo in $y' = y+1$ (chiunque è disposto a pagare le risorse di calcolo necessarie per questa trasformazione); diamo y' in input ad R; trascriviamo la risposta di R. Ovviamente, y sarà primo se e solo se R risponde | su input y'. La trasformazione da y a y' *riduce* con grande efficienza il problema di riconoscimento dei primi al problema di riconoscimento dei successori di primi.

Riduci il problema di soddisfacibilità di formule CNF al problema dell'intersecabilità delle parole. Concludi che se un giorno si troverà un algoritmo fulmineo per decidere l'intersecabilità, allora si sarà trovato anche un algoritmo fulmineo per il problema di soddisfacibilità .