UNIVERSITÀ DI PISA

Facoltà di Lettere e Filosofia

CORSO DI DOTTORATO IN DISCIPLINE FILOSOFICHE

Tesi finale di dottorato

Connections between default-assumption and preferential approaches to defeasible reasoning

Candidato:
Giovanni Casini

Relatore:
Prof. Mauro Mariani

Anno Accademico 2007/2008

Discussa il 12/09/2008

Contents

1 Introduction **11**

1.1 Logic and reasoning. 11
1.2 AI and logic. 13
1.3 Normative vs. descriptive approach 18
1.4 What are we going to talk about? 20
1.5 Basic notation. 22
1.6 Main results. 23

2 Defeasible reasoning. Main approaches and desiderata. **25**

2.1 What is defeasible reasoning? 25
2.2 Skeptical vs. credulous approaches 29
2.3 Main proposals 31
2.3.1 Default-assumption approach 32
2.3.2 Default-Rule Approach 34
2.3.3 Preferential Approach 37
2.4 A consequentialist perspective. 40
2.4.1 Tarskian consequence relation 41
2.4.2 Defeasible inference relations 45

3 Connections between the basic approaches **55**

3.1 Default-assumption inference relations and preferential relations 56
3.2 Default-assumption and default-rule inference relations 63

4 The preferential approach **67**
4.1 The KLM models . . . 67
4.1.1 Preferential entailment. . . . 75
4.1.2 Other Horn-conditions . . . 78
4.2 Non-Horn Conditions . . . 80
4.2.1 Rational monotony . . . 81
4.2.2 Derived properties . . . 84
4.2.3 Rational closure . . . 88

5 Weakly rational inference relations **95**
5.1 Representation theorem . . . 97
5.2 Relations with Injectivity . . . 105

6 Using default-assumptions in closure operations **109**
6.1 Generated preferential models . . . 110
6.1.1 Injectivity . . . 110
6.1.2 Consistent default-assumption sets . . . 113
6.1.3 Injective weak rationality . . . 116
6.2 Default-assumption closures . . . 118
6.2.1 Desiderata . . . 118
6.2.2 Knowledge and beliefs . . . 120
6.2.3 Weakly rational closure . . . 126
6.2.4 Rational closure . . . 139
6.2.5 Conclusions . . . 143

7 Working with default-assumptions: reasoning about normality **145**
7.1 Characterizing normality . . . 146
7.1.1 The normality operator $\rhd$ . . . 152
7.2 What are default-assumptions? . . . 169
7.2.1 Reasoning with stereotypes . . . 171

8 Default updating **185**

8.1 Theory revision . . . 185

8.1.1 Expansion . . . 191

8.1.2 Contraction . . . 193

8.1.3 Revision . . . 197

8.1.4 Meet contractions and entrenchment relations . . . 200

8.2 Base revision . . . 205

8.3 Default revision . . . 208

8.3.1 Expansion . . . 209

8.3.2 Contraction . . . 211

8.3.3 Revision . . . 215

8.4 Default-base revision . . . 223

8.4.1 Default-base expansion . . . 223

8.4.2 Default-base contraction . . . 225

8.5 Conclusion . . . 231

9 Conclusions **233**

Acknowledgments

I would like to express my gratitude to the people who have contributed to this thesis.
First of all, of course, my supervisor, Prof. Mauro Mariani, who has always been helpful and has given me precious advice. I surely need also to thank Prof. Guglielmo Tamburrini, to whom I owe a great deal, Prof. Massimo Mugnai, for his interest and help, and Edoardo Datteri, for his advice, his friendship, and because he has always been an example of trustworthiness and devotion to work.
A particular thanks to Prof. David Makinson, who during his visit to Pisa in 2007 dedicated much time and attention to my work, and who gave me essential advice.

A special thanks to all my friends: they always make the little everyday troubles disappear, and this is so important, especially as I was writing my thesis.

In the end, I should thank Hykel Hosni, but it is a great challenge to find an appropriate way. His contribution has been essential, both on the human and on the professional side: he has taught me really a lot and has always been extremely helpful, and working with him has been a true pleasure.

Abstract

Owing to the research program of Artificial Intelligence, in the last decades a big effort has been undertaken in order to develop interesting models of human reasoning by means of logical tools, receiving contributions from various fields, as Philosophy, Mathematics and Computer Science. One of the main problems has been the characterization of defeasible inference, i.e. that kind of inference, modeling common-sense reasoning, in which an agent draws tentative conclusions, using as supplementary information what he maintains as holding in most normal situations. Such conclusions are open to revision in case more complete information about the actual situation becomes available to the agent.

This thesis focusses on defeasible logics (or nonmonotonic logics). In particular we analyze the connection between two of the main approaches to the formalization of defeasible reasoning: the default-assumption and the preferential formalizations. On the basis of such connection we can have a deeper understanding of both approaches, and use the tools provided by each approach to work in the other one.

In the first two chapters the thesis presents the main problems and the main formal approaches to the development of logical models for defeasible reasoning. We briefly present the main proposals in the field of nonmonotonic logics and delineate the consequentialist view to the study of defeasible reasoning, i.e. an approach focused on the analysis of the behaviour of the inference relations generated by the different types of logical systems. In particular, following the recent literature, we delineate three main views, the default-assumption, preferential and default-rule approaches, distinguished by the kind of formalization used to represent default information (i.e. information about what normally holds).

In the third chapter we show that there is a correspondence between the basic formulations of the three different approaches, in particular stressing a

strong connection between the preferential and the default-assumption ones, the former referring to a preference order defined over the set of the semantic valuations of the language, the latter using a set of formulae as background information, to be added to actual information as extra-premises. We shall refer to such a connection all along the thesis.
The fourth chapter is dedicated to a brief presentation of the main results in the study of defeasible reasoning from a consequentialist point of view, presenting the main representation theorems, relating the satisfaction of desirable properties of the inference relations to particular classes of preferential models.

In the fifth chapter we isolate an interesting class of inference relations, weakly rational inference relations, that we shall use in the following chapters, and prove a representation theorem connecting such inference relations to the class of optimal preferential models.
The content of the sixth chapter is directly connected to the correspondence between the default-assumption and the preferential approach: we show how it is possible to use the default-assumption approach in order to build interesting preferential models, defining well-behaved inference relations.
In the seventh chapter we use the correspondence between default-assumption and preferential approaches in order to define in a precise way the behaviour of default formulae, by means of a normality operator. In the end of the chapter we present a generalization of a model of stereotypical reasoning proposed by Lehmann.
In the last chapter we move into the field of belief revision, defining a possible approach to the revision of default information, referring as a starting point to the main results of the AGM approach, one of the cornerstones in the field.

Chapter 1

Introduction

Abstact. We briefly depict the theoretical framework in which the problem of defeasible reasoning has arisen.

1.1 Logic and reasoning.

Philosophy has always dedicated many efforts in the characterization of thought and mental activities in general. In particular, one of the main targets of this investigation has been the individuation of the properties of *arguments*, i.e. the move from initial information (premises) to a new piece of information (conclusion). Historically, philosophers have been interested in identifying the *correct* way of thinking, i.e. which constraints have to be respected in order to be justified in moving from an assertion to another.

Aristotle, establishing logic as an autonomous discipline, depicted it as the study of the elements and the structures of reasoning, identifying how we can denote an argument as a correct one; his *theory of syllogisms* has been configured as the study of the correct argument schemes, delineating when we are justified in claiming the truth of the conclusion on the basis of the truth of the premises.

The theory of syllogisms was presented as a general theory of inference, and the syllogism as a *valid inference scheme* of the form:

$$\text{If } \alpha \text{ and } \beta, \text{ then } \gamma$$

where α, β and γ are a particular kind of propositions, corresponding to a fragment of monadic predicate logic.

So, from its beginnings, the interest of logic has been centered on the characterization of the correct forms of reasoning by means of an analysis of the '*If..., then...*' patterns, independently of the particular dominion of the discourse. This position toward the role of logic has continued after Aristotle and all along the medieval period, and the characterization of the valid derivations, by means of the clarification of the relation between premises and conclusions, has always been one of the tasks qualifying the logical endeavour.

Between the XIX and XX centuries, logic has been re-established, both by means of a strong formalization effort and with a rethinking of the aims of the discipline, delineating the research interests and the formal tools we know and use today. From this period we have received the exact definition of logic's foundational concepts, as formal languages, axiomatization and derivation rules, coherence, semantics, consequence relations, and so on. However, we have to keep in mind that this re-definition of logic has been carried forward within a research program that is distinct from the aristotelian one. Frege's main interest was to investigate the possibility of a foundation of arithmetic by means of logic; also after the failure of such a program, the main achievements in logic have sprung investigating the relations between logic and mathematics. So, the interests generating modern logic have been different from Aristotle's ones, and the development of a theory of generic argumentation and reasoning has not played a real role in the formulation of classical logic (referring by "classical" to propositional and first-order logic). Nevertheless, in the second half of the past century we have seen a recovery of the research program pointing toward the investigation of human reasoning and arguments by means of logical formalisms. In philosophical logic, the development of modal logic probably represents a main turning point, since

it unfolded the possibility to modify mathematical logic in order to formalize and investigate extra-mathematical notions: the foundational work of Hintikka [22], aimed at the formalization of the notions of knowledge and belief by means of specific modal operators, combined with the subsequent work of Kripke [26], that has defined a complete semantical characterization to the main modal systems, has allowed the growth of modal formalisms apt to the description of intensional states, formalisms which have become popular and powerful tools.
Still, the main pressures towards a recovery of the original program, pointing to a characterization of human reasoning by means of formal logic, came in the fifties from the field of Artificial Intelligence (AI).

1.2 AI and logic.

The traditional picture, describing classical logic as a successful theory regulating reasoning and argument in general, has been put under pressure especially by the problems placed by Artificial Intelligence (AI).
AI's research program is aimed at the development of computational models of intelligent behavior, i.e. programs able to solve problems typically deemed as faceable only by humans. In the evolution of the discipline, it became manifest that one of the priorities was the correct formalization of intensional notions (as beliefs, desires, etc.), that is, the correct form of representation of knowledge and reasoning.
Classical logic was introduced as the most obvious candidate for this task, and the use of logical formalisms allowed the development of *expert systems*, one of the first successes of AI. This kind of programs is based on an explicit, logical formalization of knowledge (see [40]), on which the agent operates by means of logical rules.
Expert systems have probably been the major success in the first stage of the AI endeavour, characterized by the use of formalisms directly reducible to classical logic.

However, by the eighties, a theoretical debate about what we have to define as an intelligent task changed the priorities of the research agenda, stressing the necessity of moving from problems of an "argumentative", "theoretical" kind, such as chess-playing and theorem proving, to more "practical" problems, centered on the interaction between agents and their environment. This move has given priority to the investigation of notions linked directly to the interaction with the environment, such as "action" and "deliberation", in spite of the study of more abstract and theoretical dominions.
This drove toward a disqualification of the standard picture, which saw mathematical logic as a valid theory of reasoning and arguing.

In general, we could say that logic, historically, has been interested in the formalization of the correct kinds of argument in three main dominions:

- Mathematics.
- Philosophical/scientific argumentations.
- Common-sense reasoning.

AI is mainly interested in the production of models aimed at the investigation of the third point, while classical logic is a tool developed in order to investigate the first kind of argumentation.
Formalisms based on classical logic resulted in not being suitable for modeling everyday situations for various reasons. First of all, logical formalizations need all relevant information to be explicitly present in the knowledge base, while many forms of intelligent behavior often make use of common-sense implicit knowledge. Moreover, computational costs of logical reasoning are not usually apt to model real-time deliberation in complex environments. Finally, by means of classical logic we can model agent-environment interaction only with respect to a strict class of situations: every minimal departure from what the designer has contemplated can disqualify an agent's performance. In general, the strict idealizations made by classical logic have turned out to be unsuitable for the development of interesting AI models.

To overcome such problems, AI researchers found themselves in front of a crossroad: they could give up the symbolic approach, or they could look for new logical formalizations of reasoning.
Many of them have preferred to recede from the use of logical formalisms, moving towards characterizations of cognition as a sub-symbolic enterprise. Other researchers have maintained a symbolic approach to cognition, being conscious of the exigency to move from classical logic and to develop new formalisms appropriate to the management of the new problems.
We don't want to enter into such a debate: both the approaches, the symbolic one and the subsymbolic one (considered in its different versions, from *Connectionism* to *Behaviour Based Robotics*), have merits and demerits, and neither of them has arrived at a dead-end. Probably, each of them is still going to give great contributions to the development of naturalistic models of cognition, and there are possibilities of fruitful interplays between them.
However, we would like to stress that a total abandonment of the symbolic approach seems impractical. It is possible, perhaps probable, that in the future we will conclude that a symbolic approach to cognition is unrealistic, that deliberation, and reasoning and behaviour in general, are sublinguistic tasks and can be successfully modeled only by means of nonsymbolic tools. However, such an outcome would not mine the role of the symbolic approach in the interpretation of such systems: in everyday life, the only way we have to interpret the cognitive behaviour of other agents is by means of common-sense psychological notions, articulated in symbolic argumentation structures. Such an approach is the standard way we interpret not only other agents, but also our own behaviour.
Hence, if one of the main aims of AI is to model the interplay between agents and their environments, the role of a symbolic theory of reasoning seems necessary, since both introspection and the interpretation of other agents' behaviour by means of symbolic models seem to play a fundamental role in our reasoning paths.

The revision of the research priorities, focusing on the analysis of the interplay between agents and complex environments, has stressed the need for a revision of the desiderata of the logical formalisms. The new theoretical issues have introduced a characterization of cognition and rationality dismissed from the strongly idealizations characterizing the first period of AI, aimed at the identification of *optimal* deliberation patterns, and have taken under consideration the uncertainty and the scarcity of resources, inevitably associated with real situations.
The endeavour of modeling a rational agent points now to the definition of a proficient 'management' of the scarcity of cognitive resources, aiming at the delineation of rationality bounds of both logical and economical nature.
Agent's scarcity of cognitive resources can be distinguished into three types:

- Information: the amount of information the agent has at its disposal about the environment and its own status.
- Time: the amount of time the agent has at its disposal in order to undertake a deliberation.
- Computational power: the computational bounds of agent's reasoning capabilities.

We can consider such resources as parameters: varying them we can define different characterizations of rationality. On one hand, we have the strongly *ideal* characterization, which refers to 'super-natural' agents with unbounded resources (complete information, no time and computational limits, and consequently the possibility of an optimal behavior) and which is useful in the development of theoretical models, aimed at the explication of ideal constraints an agent should respect. On the other hand, we have a *bounded* characterization: the agent's behavior is fruitful (referring to its own tasks) with respect to its knowledge (typically a partial one), its time resources (probably the agent is obliged to undertake a decision as soon as possible) and its reasoning capabilities. We have to refer to this last kind of rationality in the characterization and valuation of real computational models of cognitive functions.

Such limits are extraneous to classical logic, and imposing them influences directly the kind of formalisms we are looking for. We need to abandon a dogmatic clutch towards classical logic, and instead swing toward new formalisms apt to the characterization of bounded reasoning processes. The new problems logicians have to face are questions of epistemic and temporal kinds that the idealizations of classical logics are not apt to manage. Some of the most notorious problems are:

- Frame problem: every external event perturbs the state of the environment. Using classical formalisms, in order to be aware of the new environmental state, we should declare exactly which external properties are changed by the occurred event and which are not.

- Temporal problems: given dynamic environments, or many agents that are interacting, we need formalizations appropriate for modeling the evolution in time of the states of both agents and environments.

- Lack of information: the agent could have only partial information about its environment, and nonetheless it should be able to deliberate the same.

- Inconsistent beliefs: acquiring information from the environment and from other agents, an agent's database could result in inconsistent information. We have to implement means apt to manage such inconsistent information.

- Goal-orientation, deliberation and planning: we have to define the different epistemic attitudes an agent could have toward information with respect to the functional role every such piece of information plays in the agent's reasoning processes (beliefs, goals, intentions, duties...), and how such different kinds of information interact, defining the agent's deliberation processes.

- Interaction between agents: if we want more agents to interact in the same environment, we have to develop models apt to the formalization of the epistemic states of multiple agents.

- Vagueness and ambiguity: classical logic is not appropriate for the treatment of vague or ambiguous concepts, and typically it produces paradoxes. Notwithstanding, such kinds of concepts are strongly present in our everyday life, and we need appropriate formalisms to handle them.

On the theoretical level, the above problems have been mainly fronted referring to the philosophy of mind (e.g., [10, 7] about deliberation, or [33] about belief revision) or to classical economic models (game theory, maximization of utility...). Contrariwise, fields such as psychology or biology are not usually taken under consideration as inspirations for the development of logical models.

1.3 Normative vs. descriptive approach

These new problems have determined a turn in the priorities of logical research, a *practical turn*, as Gabbay and Woods have recently designated it ([16]), which has determined a turn also in the role of logical models. First of all, there is a visible tension between a normative and a descriptive value of such models. However, we think we have to distinguish between two senses of *normative*.

We have a very strong sense of normativity, seen as a form of *idealization*. In this sense we abstract from the limits imposed by the perspective that our agent is a real one immersed in a real environment, and consequently we abstract from the limitations imposed on the cognitive resources (information, time, computational power) of a real agent. In this sense, classical logic is a perfect model of deduction, but, as we have seen, such a level of idealization is too strong to provide useful models of rationality. However, it is still an efficient ideal system, to which we have to refer in the development of less idealized models. In general, the appeals to ideal notions such as logical consistency or inferential correctness (from the point of view of logic), or maximization of utility (on the economical side) have such an intuitive charm, that, notwithstanding their direct inapplicability in real situations, it is not conceivable not to refer to them as ultimate points of reference.

On a lower level we could define a weaker sense of a normative model, intending it as *prescriptive*. In such a sense, our formal systems consider the cognitive limits of the agent, and aim at the definition of the correct behaviour, given agent's resources. Here we abstract from a real agent, in the sense of not considering the fact that real agents often behave in wrong ways. The more our models point toward the description of the behaviour of a real, actual agent, the more they can be considered as *descriptive* ones.

Of course, these three notions are vague, with no clear boundaries, and the nature of a particular model will be a problem of degree, also with respect to our research interests (and every model, in its being a model, will always have a certain degree of idealization).

Both AI and epistemology, in general, have been interested both in prescriptive and descriptive models, but logicians have always been more centered on the prescriptive dimension, given the origins and the nature of the discipline. A prescriptive model has to manage in an efficient manner the problems listed at the end of the preceding paragraph. All such problems have been faced by logicians. An optimal logical system for agent-modeling should be able to manage all these dimensions of bounded rationality; obviously, we cannot point directly to the development of such a system. The general attitude has been towards a modular approach, facing each problem individually, defining the desiderata with respect to each one and proposing formalizations

appropriate to the satisfaction of such desiderata. The coordination of the various problems by means of their integration into further logical systems is a matter to be faced later. So logic is moving, step by step, from the extreme idealizations of classical logic towards the realization of prescriptive models for real rational agents.
The various problems seen above have been addressed by means of new logical proposals as:

- Temporal logics and dynamic logics.
- Defeasible logics.
- Belief-revision theory and paraconsistent logics.
- Agent-oriented logics.
- Multi-agent logics.
- Nonomniscient epistemic logics.

Each of these kind of systems moves from classical logic, the maximally idealized reference point, modifying it in order to obtain logical systems apt to the formalization of the particular problems under exam.

1.4 What are we going to talk about?

We are going to concentrate on the formalisms developed for dealing with lack of information, i.e. *defeasible reasoning* (or *nonmonotonic reasoning*). In particular, we are going to investigate some interesting properties derivable from the correspondences between the preferential approach and the default-assumption approach. These two systems, which will be presented in the following chapter, are two cornerstones in the study of defeasible reasoning. On one hand, the default-assumption approach can be considered as a very simple but expressive and powerful tool, appropriate also for modeling abductive processes (see [45]) and learning behaviours (see [9]), without considering the strict relation with classical models of belief-revision (see [1]). On the other hand, the preferential approach has been the most fruitful formalization in order to delineate from a logical point of view the core structural desiderata for a satisfying characterization of defeasible reasoning. By means of a strict correspondence between the two approaches, we will deepen the analysis of the behaviour of both of them.

In chapter 2 there is a brief presentation of the main systems proposed in the field: default-assumption, default-rule and preferential systems. We will also introduce the theoretical approach we will assume, centered on the notion of logical consequence.

In chapter 3, we shall investigate the correspondences between the different systems presented in Chapter 2, focusing in particular on the strict relations between the default-assumption approach and the preferential one.

In chapter 4, we shall present Kraus, Lehmann and Magidor's preferential system, a cornerstone in the study of defeasible reasoning, and the main systems obtained on the basis of their results.

Chapter 5 will be dedicated to the analysis of the condition of *weak rationality* for defeasible inference relations.

In chapter 6, we shall see how some systems presented in chapter 4 can be analyzed by means of the Default-assumption approach, in particular how the construction of their semantical models can be simplified.

In chapter 7, we shall use the correspondence between the preferential approach and the default-assumption approach, presented in chapter 3, to deepen the behaviour of default-assumption consequence relations, in particular how such an approach characterizes normality.

Finally, in chapter 8, we shall briefly present the main points of the classical belief-revision theory ,and we shall develop a theory of default revision in the default-assumption approach.

1.5 Basic notation.

We are going to work with classical propositional languages, since it suffices for the instantiation of the main problems of the field. In particular, we are going to use propositional languages built from a finite set of propositional letters, i.e. logically finite languages.

Contrary to the classical approach of mathematical logics, the use of finite languages does not need to be seen as a limit: since the final aim of the field is to model the reasoning capabilities of real agents, which obviously have a finite vocabulary, the results obtained with logically finite languages are fully relevant. So, $P = \{p_1, \ldots, p_n\}$ will be a finite set of propositional letters, and ℓ is the propositional language generated from P in the usual, recursive way by means of the classical truth-functional connectives $(\wedge, \vee, \neg, \rightarrow)$.

The sentences of ℓ will be denoted by lowercase Greek letters $\alpha, \beta, \gamma, \ldots$, and subsets of ℓ will be denoted by capital Roman letters $A, B, C, \ldots$.

We will denote by means of $\vDash$, $\vdash$, $\mid\!\sim$ (Cl, Cn, C) different kinds of consequence relations (operations). In particular, we will denote by $\vDash$ the classical consequence relation, by $\vdash$ a generic monotonic consequence relation, and by $\mid\!\sim$ a defeasible inference relation (see the next chapters for elucidations).

An element of a consequence relation $A \mid\!\sim \beta$ will be in general called a *sequent*. We assume that the set of premises A is a *finite* set of propositions.

On the semantical side, we shall use kripkian possible-worlds models. We are going to deal with sets of classical propositional valuations $W = \{w, v, \ldots\}$, that we will call worlds, and sets of states $S = \{s, r, t, \ldots\}$.

As usual, the symbol $\vDash$ will be used also as a satisfaction relation between valuations and formulae: $w \vDash \alpha$ is to be read as 'The valuation w satisfies the formula α'.

Given a set of formulae $A \subseteq \ell$ and a set of valuations W, we shall write $[A]_W$ for indicating the set of the valuations in W satisfying all sentences in A:

$$[A]_W = \{w \in W \mid w \vDash \phi \quad \text{for every} \quad \phi \in A\}.$$

Finally, given a set of sentences A, we will denote by A_w the subset of A which is satisfied by the valuation w, that is to say, $A_w = \{\phi \in A \mid w \vDash \phi\}$.

1.6 Main results.

We now briefly list the main results presented here. Two results, specifically Theorems 3.1.4 and 6.2.22, have been proven independently, finding only later that analogous results had already been presented by Freund in [12].

In Chapter 3 we prove the correspondence between the basic formulations of default-assumption and preferential inference relations.

> **Theorem 3.1.4.** *Let ℓ be a logically finite propositional language. Given an arbitrary default-assumption consequence relation $\mid\!\sim_{\Delta}$ defined over ℓ, we can define a preference consequence relation $\mid\!\sim_{\delta_{\Delta}}$ s.t. $A \mid\!\sim_{\Delta} \phi$ iff $A \mid\!\sim_{\delta_{\Delta}} \phi$, and, conversely, given an arbitrary preference consequence relation $\mid\!\sim_{\delta}$ defined over ℓ, we can define a default-assumption consequence relation $\mid\!\sim_{\Delta^{\delta}}$ s.t. $A \mid\!\sim_{\Delta^{\delta}} \phi$ iff $A \mid\!\sim_{\delta} \phi$.*

Such a result proves a strong connection between two of the main families of nonmonotonic inference relations, and will be used all along the thesis.

In Chapter 5 there is a representation theorem for the class of weakly rational inference relations.

> **Theorem 5.1.20.** [Representation Theorem for Weakly Rational Inference Relations] *A consequence relation is a weakly rational inference relation iff it is defined by some optimal model.*

In Chapter 6 we define two interesting semantical constructions, $\mathbb{W}$-closure and $\mathbb{R}$-closure, which are based on the default-assumption approach, and we prove the correspondence between $\mathbb{R}$-closure and Lehmann and Magidor's rational closure (see [31]).

Theorem 6.2.22. *The preferential model* $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\Delta_{R\mathcal{B}}} \rangle$ *is a canonical model of the rational closure* $\mathbb{R}(\mathcal{B})$ *defined in [31]*

In Chapter 7 we introduce an operator $\triangleright$ which results appropriate for describing the behaviour of default formulae. We also define a semantical notion of distance in order to define a model of stereotypical reasoning, and we prove that the model we propose defines cumulative inference relations.

Theorem 7.2.6. *Given a set of stereotypes* S *and a notion of distance satisfying (d1)-(d3'), the consequence relation* $\mid\sim_S$ *is cumulative.*

Finally, in Chapter 8 we propose some functions appropriate for the revision of sets of defaults.

Chapter 2

Defeasible reasoning. Main approaches and desiderata.

Abstract. We briefly present the main proposals in the field of nonmonotonic logics and delineate the consequentialist approach to the study of defeasible reasoning.

2.1 What is defeasible reasoning?

As we have seen, there are new directions in logic aiming at a formal characterization of reasoning of a real agent placed in its environment. This change in the dominion of analysis is coupled with different demands and desiderata from the ones of classical logic. The study of defeasible reasoning is connected to the representation of an agent that, in the absence of sufficient information, has of necessity to *presume* what's happening in its own environment. In particular, we can think of an agent that has to take a decision quickly, with insufficient information at its disposal about what is actually happening or what it is going to happen. We assume that the agent is supplied with an amount of background knowledge, expressing facts normally holding in its environment. By means of such information, the agent *completes* its knowledge base in order to derive what is presumably holding.

For example, if our agent sees that the ground in front of him is wet, and it is known that typically wet pavements are slippery, it will deliberate to proceed cautiously on the basis of this background information. In such a case the agent has to display *flexible* reasoning capabilities, referring by 'flexible' to the ability to draw conclusions on the ground of 'poor' epistemic situations, and, if necessary, to 're-adapt' such conclusions when confronted with new evidence.

Our agent must be capable of deriving plausible conclusions about the exact situation of its own environment, and to behave properly on the grounds of such presumptions.

Until now, this problem has been undertaken by means of a single type of approach: the agent completes its partial information by means of an implicit reservoir of background information. This background information, which is formalized by various means (formulae, rules, particular semantic formalizations...), is interpreted as indicating what the agent maintains as normally holding.

The core mechanism for the formalization of this kind of reasoning is the same for all the main approaches to defeasible reasoning: we add to what the agent knows the portion of background knowledge *compatible* with it, and then the agent treats this completed information in a classical way.

The way we formalize background information, and how we determine the compatibility of default information with what we know, are the main discriminants used to distinguish between the various approaches to the formalization of defeasible reasoning.

So, the term 'defeasible' in our characterization of common sense reasoning has an obvious reason: if we utilize information holding only presumptively, deeper investigations can later reveal that pieces of such information are false; such epistemic change would also falsify every conclusion drawn on the basis of such premises. In the example above, the agent could discover that the pavement it is proceeding on is very rough, and so it is not slippery. The negation of the agent's presumptions *defeats* also the conclusions and the

decisions based on it.
Aside from providing a logical basis appropriate to the formalization of deliberation, more generally we can use defeasible reasoning to deal with every dominion characterized by rules which admit exceptions. For example, let us take the most "popular" case of defeasible reasoning used in the literature: we know that Tweety is a bird, and that normally birds fly; so we can presume that Tweety flies. However, later we are informed that Tweety is a penguin; obviously, this is an exception to the rule of flying birds, and we are forced to revise our conclusions about Tweety in the face of the new evidence, however retaining the general rule that, typically, birds are flying creatures.

This is not possible in a classical frame. For example, assume we have a vocabulary P composed by three elementary propositions ($P = \{b, p, f\}$), where b is interpreted as "Tweety is a bird", p as "Tweety is a penguin" and f as "Tweety flies". We could formalize our default information by means of a set of formulae $\Delta = \{b \rightarrow f, p \rightarrow \neg f, p \rightarrow b\}$.
Using a classical formalization, such formulae have to be interpreted as strict implications, and they enter in conflict with each other. In fact, we obtain $\{b \rightarrow f, p \rightarrow \neg f, p \rightarrow b\} \vDash \neg p$, negating to Tweety the possibility of being a penguin; our knowledge tells us that birds fly, and, on the ground of such a law, it is not conceivable to be a bird and not to fly.

In classical logic, given a set of premises, we deduce conclusions such that they *necessarily* inherit truthfulness from the truth of the premises. Such a position is reflected in the property of monotonicity (see below). This property stipulates that every conclusion derived from a set of premises remains valid also augmenting the initial set of premises by means of new information, i.e. every conclusion is not defeasible in front of new information.
As we have seen, reasoning about presumptions is characterized by the flexibility of our reasoning, i.e. the possibility to retract previous conclusions in front of new information. So, the property of monotonicity is not compatible with the behaviour we want to model. Hence, on formal grounds, defeasibil-

ity of conclusions results in logical systems with a *nonmonotonic* trend.
The use of background information to complete what we really know is a phenomenon we can observe in most of our everyday reasoning, in a more or less conscious manner. Reasoning in the face of uncertainty is ubiquitous, and it is manifest that, practically in every decision, we reason on the ground of an enormous amount of *background*, *frame* information, most of which is certainly unconscious. We treat such background information as a body of 'hazardous' generalizations, not behaving as a classical sequent $A \vDash \beta$. For the latter to be true, exceptions are not allowed; a necessary tie between premises and conclusions is stipulated, and the discovery of an exception forces the negation of the validity of such a rule. On the other hand, the aim of default information is to model what we maintain as normally holding, and to use this kind of information to augment our inferential power in order to deal with our environment in an efficient manner. This kind of information holds in typical situations, and we don't want our defaults to be falsified by the discovery of exceptions; such exceptions have generally to be perceived as a signal that we are in an exceptional situation, not that our default information is incorrect.

There are many domains that look at nonmonotonic logic as a powerful tool. First of all, we have everyday reasoning, where we refer to "everyday" as the normal interaction between an agent and its environment. If we aim at developing a cognitive model of such an interaction, there are many functions which seem to need a defeasible management of information, from the interpretation of sensory information, to the deliberation processes.
However, in general we can refer to defeasible reasoning in every field in which we reason on the basis of defeasible rules, i.e. rules admitting exceptions. Also, disciplines like physics, which refers to its rules as necessary laws, have to deal with defeasibility, for example as in the management of *ceteris-paribus* conditions, i.e. conditions stating something like "law x is necessarily valid but only if our initial situation satisfies a series of unex-

pressed (and probably inexpressible in practice) frame conditions, holding in normal experimental situations".

Many have addressed nonmonotonic logics as an useless field, referring to the universal value of classical logic: anything that can be expressed in any formal language can also be expressed in classical logic. This is surely possible in principle, but it is totally unpractical in most of the situations. If we want to translate the use of a defeasible generalization in classical logic, we should transform it into a necessary generalization, making explicit every condition invalidating the rule. This is an unmanageable task also for very simple situations, both for the development of the knowledge base of the agent, and for its computational tractability.

2.2 Skeptical vs. credulous approaches

Generally, the birth date of nonmonotonic logics is indicated as 1980 when a special issue of *Artificial Intelligence* on the subject was published. In that issue some of the main foundational works have been proposed, such as McCarthy's Circumscription [41], Reiter's Default Logic [50] and McDermott and Doyle's Modal Nonmonotonic Logic [42]. From that year on, there has been a flourishing of publications on the subject, most of which propose new nonmonotonic systems or modifications of existing systems, in both cases pointing toward the characterization of logical systems appropriate to the formalization of particular dominions.
However, it is possible to conceive of a very general and quite informal structure representing the way these systems generally work.
We assume the epistemic state S of an agent to be characterized by a couple of elements: $S = \langle A, \Delta \rangle$, where A is a set of formulae representing the agent's beliefs about what is actually holding, while Δ represents default information, i.e. what the agent holds as normally holding that can be formalized in different possible ways (formulas, rules, semantical models...).

The agent treats A as 'hard' information, surely holding, while Δ represents 'soft' regularities, information about normal situations, which we assume as true until they are negated by evidence. So, the agent completes its actual beliefs by adding background information.
The use of Δ to complete the set of actual beliefs is regulated by a *consistency check*, i.e. we have to check if the default information we want to use is logically consistent with the set of premises.
If the set A is consistent with the default information, the agent uses all its default information to derive its presumptive conclusions. Otherwise, if the agent's beliefs are not consistent with the entire background information, the agent looks for the 'biggest' portions of the information formalized by Δ consistent with A, and derives its conclusions. The agent assumes defaults as much as possible, until it reaches inconsistency, i.e. the representation of a situation that is not possible.
This allows to define an extremely general notion of *extension* $E(A)$, that is, a set of conclusions generated from A and a set of information contained in Δ which is maximally compatible with A, i.e. to which no more default information can be added without obtaining a contradiction.
It is possible for an epistemic state $S = \langle A, \Delta \rangle$ to have more than a single extension.
To make a trivial example, assume that $A = \{\alpha\}$ and that we have encoded our default information by means of a set of formulae (see the following section); in particular we have $\Delta = \{\alpha \to \beta, \alpha \to \neg\beta\}$. We cannot complete the information in A by adding the entire default set because we would obtain a contradiction. So we have two possibilities, depending on which default formula we decide to add to A. Both of these alternatives are possible extensions of A by means of Δ.
The way we decide to use extensions in deriving defeasible conclusions from a set of premises is one of the main discriminants between the different proposals. The are two possibilities, positioned at opposite extremes:

- *Skeptical/cautious approach*: Given a family $\mathfrak{E}$ of extensions of a set of premises A, we consider as defeasible consequences of A only the formulas true in *every* extension in $\mathfrak{E}$.

- *Credulous/bold approach*: Given a family $\mathfrak{E}$ of extensions of a set of premises A, we consider as defeasible consequences of A all the formulas true in *at least one* extension in $\mathfrak{E}$.

The skeptical approach moves from A to the intersection of its extensions, $\bigcap \mathfrak{E}$, while the credulous approach takes to the union, $\bigcup \mathfrak{E}$. A third in-between possibility is the *choice* approach, that consists in choosing only a single extension of A.
In general, the credulous approach is not very interesting, especially for logicians, since it often derives contradictory information from a consistent set of premises. The skeptical approach is surely the most solid from a logical point of view. Also, the choice approach is interesting, but we need an extralogical apparatus for the choosing procedure of the extension.
Hence, the skeptical approach, on which we will focus our attention, is the one most investigated by logicians, and generally its results represent also a solid base if we want to move into a choice perspective.

2.3 Main proposals

Makinson's recent work, [39], actually offers a simple but efficacious reordering of the entire field. Makinson proposes to categorize all the various approaches to defeasible reasoning under three main families, determined by the kind of formalization we give to default information with respect to classical logic: propositions, rules or semantic restrictions.
We propose a brief survey of such approaches, both for a general characteri-

zation of the field, and in function of the following chapters.

2.3.1 Default-assumption approach

The simplest approach to the formalization of default information is by means of a set of background formulae. Such formulae are interpreted as information about what normally holds, and they work as the implicit assumptions the agent uses to complete its premises. The original idea is due to Poole (see [45])[1], but other approaches to defeasible reasoning, like the well known *closed-world assumption* (see [49]), can be seen as variations of this general approach.

The idea behind the default-assumption consequence relation is that we have a set of propositions Δ which indicate what normally holds (our default-knowledge). Given another set of propositions A, interpreted as premises and representing what we know about actual facts, we find which portion of default-information is consistent with what we actually know and, considering both default-information and actual information, we derive what presumably holds.

The interaction between the two kinds of initial information (actual and default) is determined by the following consistency check.

Definition 2.3.1 (maximally A-consistent sets)**.** *Given two sets of formulae, Δ and A, we say that a set Δ' is a* maximally A-consistent *subset of Δ iff Δ' is consistent with A and for no Δ'' s.t. $\Delta' \subset \Delta'' \subseteq \Delta$, Δ'' is consistent with A.*

Given a premise set A and a default set Δ, the maximally A-consistent subsets of Δ represent all the default-information which is compatible with our knowledge. In order to represent the set of every A-maximal consistent

[1]Actually, Poole's proposal was aimed to the characterization of abductive reasoning, but his technique has always been considered also a very interesting approach to defeasible deduction.

subsets of Δ, we can use the notion of *remainder set*, that is used in the field of belief revision. A remainder set of Δ modulo α is a maximal subset of Δ that does not imply α.

Definition 2.3.2. (Remainder Sets).
Let A, B be two sets of formulae. The set $A \perp B$ ('A less B') is the set such that $C \in A \perp B$ if and only if:

1. $C \subseteq A$
2. $\beta \notin Cn(C)$ *for every* $\beta \in B$
3. *There is no set C' such that $C \subseteq C' \subseteq A$, and $\beta \notin Cn(C')$ for every $\beta \in B$*

Obviously, there is a perfect correspondence between maxiconsistent sets and remainder sets: given a default assumption set Δ and a formula α, a set A is a α-maxiconsistent subset of Δ iff $A \in \Delta \perp \neg\alpha$.
So, every set of formulae $Cl(A \cup \Delta')$, with Δ' maximally A-consistent, is a possible *extension* of A with respect to Δ.
Intuitively, then, we take under consideration such sets in order to determine what the agent might expect or presume to be true in a situation in which A holds.

Definition 2.3.3 (Default-assumption consequence relation)**.** *β is a default-assumption consequence of the set of premises A, given a set of default-assumptions Δ, (written $A \mid\sim_\Delta \beta$) if and only if β is a classical consequence of the union of A and every maximally A-consistent subset of Δ.*

$$A \mid\sim_\Delta \beta \quad \textit{iff} \quad A \cup \Delta' \vDash \beta \textit{ for every maximally } A\textit{-consistent } \Delta' \subseteq \Delta$$

Theoretically, this is a very simple system for the formalization of defeasible reasoning, and we will see below that moreover its behaviour satisfies interesting properties.

2.3.2 Default-Rule Approach

The idea behind the default-rule approach is to formalize defeasible generalizations (background information) by means of derivation rules. *Default rules* are meta-rules whose role is to further complete an underlying incomplete informational state.

Definition 2.3.4 (Default Rule)**.** *A default rule is a derivation rule of the form*

$$\frac{\alpha : \beta_1, \ldots, \beta_m}{\gamma}$$

where:

- α *is the* precondition.
- γ *is the* consequent.
- $\{\beta_i | 1 \leq i \leq m\}$ *is the set of the* justifications.

The informal interpretation of the default rule above is that the consequent γ can be derived (the rule can be *triggered*) if the the precondition α is believed and the justifications $\{\beta_i | 1 \leq i \leq m\}$ are consistent with everything believed, i.e. there is a successful consistency check between the belief set of the agent and the justifications. Note that the consistency check is not with respect to the consequent; since such a possibility would be a more intuitive consistency requirement, we can isolate a particular class of default rules, the class of *normal rules*, in which the consequent and the justification are identical.

Definition 2.3.5 (Normality)**.** *A default rule is* normal *if it is of the form*

$$\frac{\alpha : \beta}{\beta}$$

The simplest forms of defaults are normal defaults without preconditions, i.e. rules of the form $\frac{\top:\beta}{\beta}$. The intuitive appeal and the utility of non-normal

defaults have often been discussed (see e.g. [29]), and we can refer just to normal defaults in this brief presentation.
The epistemic state of an agent is defined by means of a *default theory*, that is formed by a couple of sets: a set of formulae defining the belief set of the agent, and a set of default rules describing the default information.

Definition 2.3.6 (Default Theory). *A default theory is a pair* $S = \langle A, \Delta \rangle$ *where:*

A is a set of formulae, the 'factual' information;

Δ is a set of default rules.

Formally, the set of conclusions which can be added to our knowledge base is defined in terms of a *default extension*.

Definition 2.3.7 (Default Extension). *Given a default theory* $S = \langle \Delta, A \rangle$, $E_\Delta(A)$ *is an* extension *of A under Δ iff there is a sequence* $E_0, E_1, \ldots$ *s.t.:*

- $E_0 = A$
- $E_\Delta(A) = \bigcup\{E_i | i \geq 0\}$
- *for all* $i \geq 0$, $E_{i+1} = Cl(E_i) \cup \{\beta\}$, *with β s.t.*
  - $\frac{\alpha:\beta}{\beta}$ *is an instance of a default in* Δ
  - $E_i \vDash \alpha$
  - *β is consistent with* $E_\Delta(A)$

As can be seen, in this definition of the construction of an extension $E_\Delta(A)$, $E_\Delta(A)$ itself is mentioned for the consistency check (last point). So, the construction is considered as *semi-inductive*, and it is equivalent to the following *fixed-point* construction.

Theorem 2.3.1. *([50], Theorem 2.1) An extension of a set of formulae A is the smallest set $E_\Delta(A)$ satisfying*

1. $A \subseteq E_\Delta(A)$

2. $Cl(E_\Delta(A)) = E_\Delta(A)$

3. *if* $\frac{\alpha:\beta}{\beta}$ *is an instance of a default in* Δ, *and* $\alpha \in E_\Delta(A)$, *and* $\neg\beta \notin E_\Delta(A)$, *then* $\beta \in E_\Delta(A)$.

So, an extension is a set of beliefs which are in some sense 'reasonable' in the light of what is known about the world.

This is the classical definition of extension, typically used in Computer Science because it has some computational advantages, but it is also possible to give an equivalent definition of extension in an inductive form (see [8]), which is surely more intuitive.
First of all we can define an inductive way to construct a default extension of $S = \langle A, \Delta \rangle$; this can be done by imposing an ordering over the elements of the set Δ, which can be interpreted as a possible priority order over the default rules: it states in which order we have to check if a rule can be triggered.

Definition 2.3.8 (Default Extension - inductive construction)**.** .

- *Define a finite or* ω*-ordering* $\langle \Delta \rangle$ *of the given set* Δ *of rules.*
  We define $E_{\langle \Delta \rangle}(A) = \cup\{A_n | n < \omega\}$, *and set* $A_0 = A$.

- *We break the definition of* A_{n+1} *into two cases, invoking a consistency check.*
  Case 1. Suppose there is a rule $\frac{\alpha:\beta}{\beta} \in \Delta$ *such that* $\alpha \in A_n$ *but* $\beta \notin A_n$ *and* β *is consistent with* A_n. *Then choose the first such rule and put* $A_{n+1} = Cn(A_n \cup \beta)$.
  Case 2. Suppose there is no such rule. Then put $A_{n+1} = A_n$.

Every ordering imposed over the set of rules generates a different extension $E_{\langle\Delta\rangle}(A)$.
To obtain the skeptical consequences of a set A, we have to intersect all the extension of A.

Definition 2.3.9 (Sceptical Default-Rule Consequence relation)**.** .
We define a default-rule consequence operator C_Δ by putting

$$C_\Delta(A) = \cap\{E_{\langle\Delta\rangle}(A) | \langle\Delta\rangle \textit{ is a finite or } \omega\textit{-ordering of } \Delta\}.$$

In relational notation:

$$A \mid\!\sim_\Delta \beta \textit{ iff } \beta \in \cap\{E_{\langle\Delta\rangle}(A)\}.$$

The default-rule approach is probably the most popular in the field of nonmonotonic logics, mainly in the field of Computer Science. However, the analysis of its behaviour has always been problematic and, as we are going to see below, it does not satisfy some of the main desiderata for defeasible logics.
We have to deem as variations of this kind of approach also the classical modal approaches to nonmonotonic reasoning, as nonmonotonic modal logic [42] and autoepistemic logic [24], since they use fixed point constructions as well and, in specific forms, behave exactly in the same way (see e.g. [23]).

2.3.3 Preferential Approach

The last kind of consequence relation on which we are going to focus on is the family of *preferential* consequence relations. It is a strictly semantical approach, built in a possible-worlds frame. Its first formalization was presented by Shoham in [55] as a generalization of *Circumscription*, one of the most popular proposals in nonmonotonic reasoning (see [41]).
Given a set of classical propositional valuations $W = \{w, v, \ldots\}$, each representing a possible situation, we define an order $\prec$ over it, s.t. $w \prec v$ is interpreted as meaning that the situation described by the valuation w has to be considered preferred to (more normal than) the situation associated to

the valuation v.
Hence, our background information is formalized by means of such an ordering over the possible situations.
Given a set A of factual information, the agent assumes to find itself in one of the most normal situations in which the facts in A hold, i.e. in one of the $\prec$-minimal situations in $[A]_W$ (the subset of W composed by those worlds satisfying the formulae in A). Each of those worlds can be considered as a semantical extension of the set of premises A.
We define a model $\mathfrak{M} = \langle W, \delta \rangle$, where W is a set of classical propositional valuations and δ is an irreflexive and transitive relation over W (δ is a strict order over W, and we will write $w \prec_\delta v$ for $(w, v) \in \delta$).
Given a set of worlds $V \subseteq W$, we call $min_\delta(V)$ the set of the minimal worlds in V with respect to the order δ.

$$min_\delta(V) = \{w \in V | \text{ there is no } v \in V, \text{ s.t. } v \prec_\delta w\}$$

If a world w is s.t. $w \in min_\delta([\alpha]_W)$, i.e. it is between the minimal models in $\langle W, \delta \rangle$ satisfying α, we will write $w \vDash_\prec \alpha$.

A well-behaved preferential model should also satisfy the *smoothness condition*.

Definition 2.3.10. (Smooth sets).
Let W a set of elements, $\prec$ a binary relation on W, and $V \subseteq W$. V is smooth *iff for every $w \in V$, $w \in min_\delta(V)$, or there is a $v \in V$ s.t. $v \prec w$ and $v \in min_\delta(V)$.*

Definition 2.3.11. (Smoothness condition).
Let W be a set of worlds and $\prec$ an asymmetric binary relation on W. The model $\langle W, \prec \rangle$ satisfies the smoothness condition *iff, for every $\alpha \in \ell$, $[\alpha]_W$ is a smooth set.*

This condition allows to assume that, given an order δ and taken a non-empty subset of W, we always have a non-empty set of minimal worlds.

Since δ satisfies irreflexivity and transitivity, in our model there are no cycles (sequences of valuations of the form $\alpha \prec \beta_1 \prec \ldots \prec \beta_n \prec \alpha$), that, by transitivity, would imply $\alpha \prec \alpha$; the absence of cycles prevents the presence of the same valuation twice in a $\prec$-chain of worlds. Consequently, working with a finite set of worlds, if we take a valuation w we will have only finite $\prec$-chains of valuations starting from w (only finite sequences of worlds preferred to w), since otherwise we would have a cycle. It is immediate to see that the finiteness of the $\prec$-chains of worlds implies the satisfaction of the smoothness condition. Hence, working with finite sets of valuations (i.e. with logically finite propositional languages), the smoothness condition is automatically satisfied.
To define the defeasible consequences of a set of premises A in $\mathfrak{M}$ we simply consider the formulas satisfied by every semantical extension of A in $\mathfrak{M}$, i.e. by every minimal world in $[A]_W$.

Definition 2.3.12 (Preferential consequence relation)**.** *Given a preferential model $\mathfrak{M} = \langle W, \delta \rangle$, β is a preferential consequence of the set of premises A (written $A \mid\sim_\delta \beta$) if and only if β is satisfied by every world that is δ-minimal in the set $[A]_W$.*

$$A \mid\sim_\delta \beta \quad \textit{iff} \quad w \vDash \beta \textit{ for every } w \in min_\delta([A]_W)$$

This is interpreted as stating that, given A, we consider as presumably holding every formula satisfied in every most normal situation compatible with the truth of A.
As we will see below, such a semantical approach to defeasible reasoning has been extremely fruitful in the logical analysis of nonmonotonic reasoning.

2.4 A consequentialist perspective.

We have presented different forms of nonmonotonic reasoning: Poole's maxi-consistent constructions, Reiter's default systems, Shoham's preferential models. Makinson treats them as the basic systems representing the three main approaches to nonmonotonic reasoning, but there is a multitude of other systems which can be seen as their specializations.

Today, most researchers admit that there is not a single, correct nonmonotonic logic, and that different domains need to be treated by means of different kinds of defeasible reasoning. Nevertheless, human nonmonotonic reasoning has always appeared as a quite disciplined field, and the reference to a core of basic structures, constraining the development of new particular systems, seems a necessity.

Moreover, a theoretical reordering of the field has been felt as a necessity, in order to define a set of desired properties that our systems have to satisfy.

Such a reordering has been firstly proposed by Gabbay [15], and developed ten years later by Makinson [38]. Their approach focuses on the characterization of the premises-conclusions relations holding in defeasible logics, i.e. on the properties of the consequence relation defining the behaviour of our logical systems. So, we place ourselves into a *consequentialist perspective*, assuming the identification of the reasoning capabilities of an agent with the inference relation arising from the particular "machinery" modeling its reasoning processes. We could see this as a *black box* approach, focusing on the conditions defining input-output dependencies of the observed systems, independently of the inner structure of the system. The analysis of the structure of the premises-consequences dependencies, abstracting from the way we derive the conclusions, becomes the central target of this field.

From their point of view, the analysis of the relation of logical consequence becomes the ground for the construction of logical systems for defeasible reasoning.

We briefly present the classical definition, due to Tarski, of such a notion.

In the following we will use $\vDash$ and Cl as symbols of the classical propositional consequence relation and of the classical propositional consequence operator respectively, while $\vdash$ and Cn will be used to refer to a generic monotonic consequence relation and the associated operator. To refer to nonmonotonic consequence relations and operators, we will use $\mid\!\sim$ and C. It is common use to distinguish such relations from monotonic ones calling them *inference relations* and *inference operators* instead of consequence relations and operators.

2.4.1 Tarskian consequence relation

Consequence relations $\vdash$ are the traditional means for the characterization of argumentation patterns; they are relations $\vdash: \wp(\ell) \times \ell$, i.e. holding between sets of formulae and single formulae (or between sets of formulae and sets of formulae, but we won't use such a formalization here).
$A \vdash \alpha$ has to be interpreted as 'α is a logical consequence of the set of formulae A', i.e. that if we reason assuming that the formulae in A are true, we can conclude that α is also true. A consequence relation can also be described by means of a *consequence operator*, a function $Cn : \wp(\ell) \mapsto \wp(\ell)$, associating a set of premises to the set of their conclusions:

$$Cn(A) = \{\alpha | A \vdash \alpha\}$$

and conversely:

$$A \vdash \alpha \text{ iff } \alpha \in Cn(A)$$

In the following, we shall switch freely between symbolisms using consequence relations and the associated consequence operators.
The standard approach to consequence relations is the account given by Tarski in his 1936 classical paper [58].
Tarski gives the conditions of acceptability of a notion of 'following logically'. The notion of 'truth preservation' works as the reference point for the adequacy of the definition of logical consequence, i.e. a formula α is a logical

consequence of a set A iff the truth of the set of premises A implies the truth of the formula α: it can't be that A is true without α being true.
Formally, Tarski interpreted such desiderata on the semantical side, grounding his definition of following logically on the notion of semantical model:

> "We say that the sentence X *follows logically* from the sentences of the class $\mathfrak{K}$ if and only if every model of the class $\mathfrak{K}$ is at the same time a model of the sentence X."([58], p.186)

This can be characterized as an 'inclusion' definition: given a class of logical models, to tell that α is a logical consequence of A we need every model of A to be also a model of α, i.e. we need the set of models satisfying all the formulae in A to be a subset of the set of models satisfying α. That corresponds to saying that it is not conceivable a situation s.t. A is true and α is not true.

Consequence relations can also be treated apart from the semantical characterization by means of axioms and rules defining the behaviour of the consequence relation. For example, the behaviour of a tarskian consequence relation can be defined by means of the following three rules/conditions:

- Reflexivity (REFL):

$$K \vdash \alpha \text{ for every } \alpha \in K$$

- Cut (CT):

$$\frac{K \cup \{\alpha\} \vdash \beta \qquad K \vdash \alpha}{K \vdash \beta}$$

- Monotony (MON):

$$\frac{K \vdash \beta}{K \cup \{\alpha\} \vdash \beta}$$

Briefly, their intuitive meaning is:

- *Reflexivity*: we can derive every formula contained in our premises. It is really an intuitive and elementary condition, and generally is considered as a minimal requirement for reckoning a relation as a consequence one[2].

- *Cut* (or *Cumulative Transitivity*): this condition is a cautious version of plain transitivity (if $A \vdash \beta$ for every $\beta \in B$ and $B \vdash \gamma$, then $A \vdash \gamma$).It states that we can accumulate our conclusions into our premises without an amplification of inferential power, so it characterizes our consequence relation as a closure one. Moreover, CT implies that the conclusions have the same 'status' independently of the number of steps needed to infer them from the premises: once inferred, a proposition can be added to the original set of premises, and every conclusion obtained will be implied also by the original set. CT is important in the dynamics of proofs because it allows one to add to a set of premises other information in order to reach the desired conclusion, and then to justify the added premises on the ground of the original set.

- *Monotony*: Monotony is obviously our 'cause of the scandal'. It tells us that augmenting the information in the premises, whatever we had concluded before remains true. In the dominion of mathematical logic, obviously, it is a desirable property, since it guarantees that we are dealing with 'sure' information; the main problem of defeasible logic is to weaken such a property in an 'educated' way.

We can state a corresponding characterization in terms of the consequence operator Cn:

[2]The only case known to the author of a non-reflexive consequence relation comes from the field of modal dynamic logics (see [3], pp.19-21), but it can be justified by the particular interpretation of the semantical structure proposed.

- Inclusion (INC):

$$A \subseteq Cn(A)$$

- Cut (CT):

$$A \subseteq B \subseteq Cn(A) \quad \Rightarrow \quad Cn(B) \subseteq Cn(A)$$

- Monotony (MON):

$$Cn(B) \subseteq Cn(A) \text{ whenever } B \subseteq A$$

Given inclusion, we can derive from cut an important property, idempotence, which defines our operator as a closure one.

- Idempotence (IDP):

$$Cn(A) = Cn(Cn(A))$$

If a set of formulae is closed under an operator Cn (i.e. $Cn(A) = A$), we will call such a set A a *theory* of Cn.
Logical consequence is one of the foundational concepts of logic. For many authors, a formal machinery for modeling reasoning processes can be considered a logic if it can be formalized by means of a formal language and a consequence relation. That is, a logical system can be identified by a pair $\langle \ell, C \rangle$, where ℓ is a formal language and C a consequence operator defined over ℓ. From a formal point of view, such a characterization is satisfactory, and it allows one to analyze the behaviour of a logical system from the standpoint of its meta-properties, such as consistency, compactness, and so on.
The Tarskian characterization of the concept of logical consequence stands as a point of reference, and it is absolutely solid from the point of view of mathematical logic. However, as we have said in the previous chapter, the tools of classical logic are not generally appropriate in the formalization of common-sense reasoning.

2.4.2 Defeasible inference relations

Tarski himself has stressed that a notion of consequence relation, to be successful, first of all has to account for the everyday, intuitive usage of the notion of 'following', i.e. our notion has to be *materially adequate* with respect to the dominion formalized (in particular, in the case of Tarski, with respect to the mathematical context).
The classical paper *'On the Concept of Following Logically'* begins with the following period:

> "The concept of *following logically* belongs to the category of those concepts whose introduction into the domain of exact formal investigations was not only an act of arbitrary decision on the side of this or that researcher: in making precise the content of this concept, efforts were made to conform to the everyday 'pre-existing' way it is used. [...] the way it is used is unstable, the task of capturing and reconciling all the murky, sometimes contradictory intuitions connected with that concept has to be acknowledged a priori as unrealizable, and one has to reconcile oneself in advance to the fact that every precise definition of the concept under consideration will to a greater or lesser degree bear the mark of arbitrariness."([58], p.176)

Our target is the formalization of common-sense reasoning, and, as we have said, it is characterized by the property of nonmonotony. So, tarskian consequence relation is not adequate for our aims.
This does not absolutely put under discussion the value of tarskian consequence relation for the field it was developed for, i.e. mathematical logic. Simply, if we agree with the practical turn of logic, depicted in the previous chapter, we have to reformulate our desiderata in order to deal with our targets.
The property of monotony, at least referring to the tarskian formulation, is strictly tied to the notion of 'truth-preservation', a natural desiderata in the

field of mathematics. 'Truth-preservation' is too strong a notion with respect to everyday reasoning and has to be substituted with a 'softer' claim, more epistemic and also more vague, such as 'acceptability', or 'take-as-true preservation', i.e. α is a consequence of a set of premises A iff is acceptable, reasonable, to take α as true on the basis of the knowledge of A.
We have to move from an interpretation of 'if α, then β' as 'if α is true, then β is true' to a new interpretation as 'if α is true, then, normally/presumably, β is true'.
This move justifies the change in the desiderata of our consequence relation. Our approach to the characterization of logical consequence will start from the definition of the patterns that we maintain as characterizing logically valid arguments with respect to the nature of the kind of argumentation we are looking for, i.e. with respect to the target we want our consequence relation to model.
So, we start from the identification of the argument schemes we want our formal system to satisfy, tempting to define the desirable logical properties in terms of such patterns.

Pure conditions

We consider as 'pure' the proprieties which refer to the inference relation alone, without regard to its interaction with the classical consequence operation and the connectives. For example, Monotony, Cut and Reflexivity are all pure conditions.
The inference relations $\mid\sim \subseteq \wp(\ell) \times \ell$ or operations $C : \wp(\ell) \to \wp(\ell)$ we are investigating may happen to fail monotony.
So, if we eliminate monotony from the tarskian properties, what remains?

As we have said, reflexivity appears as a necessary condition in the definition of a consequence relation, and it is obvious that, if we know that α, we presume that α.
Cut occupies a strategic point: formulated in terms of $\mid\sim$ and C, cut is

$$\frac{A \cup \{\alpha\} \mid\!\sim \beta \qquad A \mid\!\sim \alpha}{A \mid\!\sim \beta}$$

$$A \subseteq B \subseteq C(A) \quad \Rightarrow \quad C(B) \subseteq C(A)$$

Cut is considered a desirable property, because it does not imply monotonicity and it is a really powerful logical property. It can be interpreted as saying that a plausible conclusion is as secure as the assumptions it is based on, and, since we are not looking for logics modeling also degrees of belief (as probabilistic logics), we find it intuitive.
Moreover, from Cut and Reflexivity we can derive idempotence, that characterizes C as a closure operator.

Despite the fact that defeasible logics are firstly characterized as 'non-monotonic', their main task is not to eliminate the property of monotony: such a property remains intuitive and is often satisfied by common-sense reasoning. What we aim to is just a weakening of such a property in order to deal with presumptive reasoning and the use of background information. Notwithstanding that classical logic is not appropriate in modeling defeasible reasoning, we continue to look at it as our ideal reference point, and we aim to preserve monotony and classical properties as much as possible.
The converse of cut is an important restricted form of monotony, *Cautious Monotony*:

- Cautious monotony (CM):

$$\frac{A \mid\!\sim \gamma \qquad A \mid\!\sim \beta}{A \cup \{\beta\} \mid\!\sim \gamma}$$

$$A \subseteq B \subseteq C(A) \quad \Rightarrow \quad C(A) \subseteq C(B)$$

Cautious monotony says that if β is a plausible conclusion from a set of premises A, we can safely add it to A without invalidating the other plausible

conclusions of A.
It is a very natural and intuitive property, and it is the minimum requirement for a good monotonic behaviour in common-sense reasoning.
Cut and Cautious Monotony can be expressed together by the propriety of *Cumulativity*:

- Cumulativity (CUM):

$$A \subseteq B \subseteq C(A) \quad \Rightarrow \quad C(A) = C(B)$$

These conditions correspond to natural ways of organizing our reasoning. They tell us that we accumulate our conclusions into our premises without loss of inferential power (cautious monotony) or amplification of it (cut), describing the reasoning process as 'stable'. Cumulativity can be seen as a principle for organizing our conclusions coherently.
Cut can also be seen as expressing the claim that we do not want to allow the length of a derivation to affect the value of the conclusion. On the other hand, Cautious Monotony can be seen as expressing a form of 'irreversibility' in the drawing of conclusions: once inferred, a proposition may be retained irrespective of what other inferred propositions are added to the stock of usable information.

Non-pure conditions

Non-pure conditions regulate the behaviour of our inference relation with respect to other consequence relations and truth-functional connectives.
As was said above, despite the fact that classical logic is not apt to the characterization of common-sense reasoning, it is important to hold it theoretically as a point of reference of the ideal reasoning, and technically as the most solid and well-behaved formalization of deduction we have at our disposal.
So it is of extreme importance to relate the behaviour of our consequence

relation to the properties of classical consequence relation.
Moreover, we are going to work with a propositional language, and so it is important to investigate what kind of conditions a defeasible inference relation should satisfy with respect to classical truth-functional connectives.
The main property relating a defeasible operator C to the classical operator Cl is supraclassicality, stating that C is an extension of Cl.

- Supraclassicality (SCL):

$$\frac{A \vDash \alpha}{A \mid\!\sim \alpha}$$

$$Cl(A) \subseteq C(A)$$

This property means that every classical consequence of a set of premises has to be included between its defeasible consequences. It simply states that 'sure' conclusions have to be included in what is presumed, that by reasoning in a defeasible way we extend our derivative capability beyond the 'safeness' of classical deduction.
Supraclassicality has a series of interesting consequences, as full absorption.

Proposition 2.4.1 (Absorption)**.** *([38], Observation 2.2.1) Let C be a* supraclassical *inference operation.*
If C satisfies Idempotence ($C(A) = C(C(A))$)*, then C satisfies* left absorption*:*

$$Cl(C(A)) = C(A) \text{ for every premise set } A$$

If C is cumulative ($A \subseteq B \subseteq C(A) \Rightarrow C(A) = C(B)$)*, then it satisfies* full absorption*:*

$$Cl(C(A)) = C(A) = C(Cl(A)) \text{ for every premise set } A$$

Full absorption states that our operator C is really well behaved with respect to classical consequence operation, since it is an extension of Cl, but

theories of C are also closed under classical consequence.
Makinson holds that an approach to nonmonotonic reasoning has to be considered as *logical* only if the inference operation C satisfies the full absorption principle, since, given a set of premises A, the propositions we're allowed to infer from A form a classical theory ($C(A) = Cl(C(A))$, and depend only upon the logical content of A rather than upon its manner of presentation ($C(A) = C(Cl(A))$).
The absorption property implies other important non-pure conditions.

Proposition 2.4.2. *([38], Observation 2.2.2) Let $C : \wp(\ell) \to \wp(\ell)$ be any operation.*
If C satisfies left absorption, then it satisfies:

- *Right conjunction (AND):*

$$\frac{\alpha \mid\!\sim \beta \qquad \alpha \mid\!\sim \gamma}{\alpha \mid\!\sim \beta \wedge \gamma}$$

- *Right weakening (RW):*

$$\frac{\alpha \mid\!\sim \beta \qquad \beta \models \gamma}{\alpha \mid\!\sim \gamma}$$

If C satisfies right absorption, then it satisfies:

- *Left logical equivalence (LLE):*

$$\frac{\alpha \mid\!\sim \gamma \qquad \models \alpha \leftrightarrow \beta}{\beta \mid\!\sim \gamma}$$

$$Cl(A) = Cl(B) \;\Rightarrow\; C(A) = C(B)$$

◦ *Subclassical cumulativity (SUB):*

$$A \subseteq B \subseteq Cn(A) \ \Rightarrow \ C(A) = C(B)$$

These are all desirable properties.
AND states that, if I believe that, given α, presumably β holds, and presumably γ holds, then I will believe that presumably $\beta \wedge \gamma$ would hold.
RW states that if I presume β, I will suppose also all its classical consequences.
The intuitiveness of such claims is manifest, and they are considered the core conditions for the interaction of C with truth-functional connectives.
LLE and SUB define our consequence operator as a typical logical one, i.e. sensible only to the logical structure of its premises, and independent of their syntactical forms.
From LLE we can derive another intuitive and technically important property, i.e. conjunction in the premises.

◦ Conjunction in the premises (L$\bigwedge$):

$$A \mid\!\sim \alpha \ \Leftrightarrow \ \bigwedge(A) \mid\!\sim \alpha,$$

where $\bigwedge$ is a function that takes as argument a set of formulae and gives back a single formula composed by their conjunction.
This property, which is very intuitive, allows to treat our inference relations as single-premise relations (assumed we are working with finite sets of premises).
Supraclassicality is immediately verified in any $\mid\!\sim$ satisfying REF and RW:

$$\frac{\vDash \alpha \rightarrow \beta \qquad \alpha \mid\!\sim \alpha}{\alpha \mid\!\sim \beta}$$

Another fundamental non-pure condition for C is the condition of *distribution*:

- Distribution (DIS):

$$C(A) \cap C(B) \subseteq C(Cn(A) \cap Cn(B)) \text{ for all } A, B \subseteq \ell$$

Any inference operation satisfying distribution will be *distributive*. The condition of distribution does not have an intuitive appeal; its justification lies in its logical power, because it implies other important non-pure conditions.

Proposition 2.4.3. *([38], Observation 2.2.3) If C is an inference operation satisfying distribution, supraclassicality, and absorption, then C satisfies the following proprieties:*

- *Disjunction in the Premises (OR):*

$$\frac{\alpha \mid\!\sim \gamma \qquad \beta \mid\!\sim \gamma}{\alpha \vee \beta \mid\!\sim \gamma}$$

$$C(A \cup \{\alpha\}) \cap C(A \cup \{\beta\}) \subseteq C(A \cup \{\alpha \vee \beta\})$$

- *Conditionalization (S):*

$$\frac{\alpha \mid\!\sim \beta}{\mid\!\sim \alpha \rightarrow \beta}$$

$$\beta \in C(A \cup \{\alpha\}), \textit{ then } \alpha \rightarrow \beta \in C(A)$$

OR has an intuitive appeal: if a formula γ is presumable assuming α or assuming β, it is natural to think that it is presumable also assuming $\alpha \vee \beta$. Moreover, OR allows to beneficiate of the *proof by cases* rule:

$$\frac{A, \alpha \mid\!\sim \beta \qquad A, \neg\alpha \mid\!\sim \beta}{A \mid\!\sim \beta}$$

Conditionalization is also called the 'hard half of the deduction theorem'. It is intuitive (If I maintain that 'given α, normally β' holds, then I think that normally α implies β), and is very useful.
The converse (the 'easy half': if $A \mid\!\sim \alpha \rightarrow \beta$, then $A, \alpha \mid\!\sim \beta$), does not hold for any interesting nonmonotonic relation, since generally it implies monotonicity.
Since the centrality of their role in the study of defeasible inference, *cumulativity*, *supraclassicality* and *distribution* are addressed by Makinson as the *core* conditions of nonmonotonic logic.

Another important property, not derivable by those just mentioned, is *consistency preservation*:

- Consistency Preservation (CP):

$$\frac{\alpha \mid\!\sim \bot}{\alpha \vDash \bot}$$

$$Cl(A) \neq \ell \;\Rightarrow\; C(A) \neq \ell$$

Also consistency preservation relates the consistency notion in C to classical consistency in Cl. It can be seen as a condition limiting the power of our inferences: C takes us to contradictions only if they were already classically implicit in our premises. This is again a principle imposing on our reasoning a good logical behaviour, indicating the strong classical notion of contradiction as a reference point.
Intuitively (and semantically), the principle could be interpreted as saying: if a set of premises is classically consistent, then it is conceivable at least a situation confirming it, and so (at least in a skeptical or choice approach to

defeasibility) we are not justified in concluding a contradiction.

We can briefly indicate if the systems described above generate inference relations satisfying the structural properties just seen.
The default-assumption approach manifest a very good behaviour, satisfying all the desired properties: cumulativity and supraclassicality ([38], Observation 3.3.1), distribution ([38], Observation 3.3.2) and consistency preservation ([38], Observation 3.3.3).
The same can be said for the simple preferential systems defined above ([38], Observations 3.4.2, 3.4.3, and 3.4.6).
A different situation comes out for default-rule approach: while it satisfies cut ([38], Observation 3.2.1), full absorption ([38], Observation 3.2.2), and consistency preservation (but just in the case of normal rules), there is the possibility of the failure of both distribution ([38], Observation 3.2.3), and cautious monotony ([38], Observation 3.2.4).
So, if we place ourselves from a consequentialist point of view, default-assumption and preferential approaches manifest a very regular behaviour, but the same cannot be said about Reiter's systems.
One of the main open problems in nonmonotonic logics has been indicated as the definition of the relations between the behaviours of the various approaches. In the following chapter we are going to see that there is a connection between the basic formulations of the three systems presented, that is particularly strong between the default-assumption and the preferential ones.

Chapter 3

Connections between the basic approaches

Abstract. We show that there is a correspondence between the basic formulations of default-assumption, default-rule and preferential inference relations.

The precise definition of how the different approaches to defeasible reasoning are related between them is still an open problem . However, it is possible to show easily that, working with a finite language, the basic formulations of the systems defined in the previous chapter manifest the same behaviour. In particular the connections between default-assumption and preferential approaches are quite strong.

In the following, we assume a propositional language ℓ generated from a finite set P of elementary letters, and the finite set W of all the possible classical valuations of ℓ.

3.1 Default-assumption inference relations and preferential relations

The purpose of this section is to establish a perfect correspondence, in the finite case, between the family of the preferential inference relations and the family of the default-assumption inference relations. The main result in this section, Theorem 3.1.4, has been developed independently, finding only later that it had already been proved By Freund in [12].
Default assumption extensions are defined by means of maxiconsistent subsets of defaults; this means that we consider a situation as more normal than another if it satisfies a larger set of defaults. From this consideration it is immediate, given a default set Δ and a set of valuations W, to define a preference order between them, a *generated strict order*, with respect to the amount of default formulae satisfied by every valuation.

Definition 3.1.1 (Generated strict order)**.** *Given a set Δ of formulae, a relation δ is generated by Δ (and we write δ_Δ) iff $\delta = \{(w,v) \in W \times W \mid \Delta_w \supset \Delta_v\}$.*

Obviously, δ_Δ is irreflexive ($\Delta_w \not\supset \Delta_w$ for any $w \in W$) and transitive (for the transitivity property of '$\supset$').
We want to show that, if an order δ is generated from a set of formulae Δ ($\delta = \delta_\Delta$), then the default-assumption consequence relation defined by the assumption set Δ corresponds exactly to the preferential consequence relation defined by the order δ_Δ. That is

$$A \mid\!\sim_\Delta \phi \quad \text{iff} \quad A \mid\!\sim_{\delta_\Delta} \phi \quad \text{for every premise set } A$$

Lemma 3.1.1. *Let δ be a strict order generated by a set of assumptions Δ. That is, $\delta = \delta_\Delta$.*
For every set of premises A, $w \in min_\delta([A]_W)$ iff Δ_w is a maximally A-consistent subset of Δ.

Proof.
(⇒): $w \in min_\delta([A]_W) \;\Rightarrow\; \Delta_w$ is a maximally A-consistent subset of Δ.

Assume that Δ_w is not a maximally A-consistent subset of Δ, that is that there is a $\Delta' \subseteq \Delta$ s.t. $\Delta' \supset \Delta_w$ and Δ' is A-consistent.
Given that W contains all the valuations of our language, there will be a valuation v s.t. $v \vDash \Delta' \cup A$. That means that $\Delta_v \supseteq \Delta'$, and so $\Delta_v \supset \Delta_w$.
Then we have that $v \prec_\delta w$ with $v \in ([A]_W)$.
Therefore $w \notin min_\delta([A]_W)$.

(⇐): Δ_w is a maximally A-consistent subset of $\Delta \Rightarrow\; w \in min_\delta([A]_W)$.

Assume that $w \notin min_\delta([A]_W)$. Then there is a $v \in [A]_W$ s.t. $v \prec_\delta w$. So $\Delta_v \supset \Delta_w$ and Δ_v is A-consistent. Δ_w is not a maximally A-consistent subset of K.
■

Proposition 3.1.2. *For every set of premises A, we have that*

$$A \mid\!\sim_\Delta \phi \quad \textit{iff} \quad A \mid\!\sim_{\delta_\Delta} \phi$$

Proof.
We know that $A \mid\!\sim_\Delta \phi$ iff $A \cup \Delta' \vDash \phi$ for every $\Delta' \subseteq \Delta$ that is maximally A-consistent. This means that for every $w \in W$, if $w \vDash A \cup \Delta'$ for some $\Delta' \subseteq \Delta$ that is maximally A-consistent, then $w \vDash \phi$.
By lemma 3.1.1, a valuation w satisfies the set A and a set Δ', s.t. $\Delta' \subseteq \Delta$ is maximally A-consistent, iff $w \in min_{\delta_\Delta}([A]_w)$. Therefore $A \mid\!\sim_\Delta \phi$ iff $w \vDash \phi$ for every $w \in min_\delta([A]_w)$, as required.
■

We now show that, given any strict order δ over W, we can find a set of default assumptions Δ^δ from which it can be generated.

We define the set of formulae $\{\alpha_w | \ w \in W\}$, which is the set of the distinctive formulae of the worlds in W, called *atoms*, i.e. α_w is the formula that describes completely the valuation w, s.t. it is satisfied only by the world w.

$$\alpha_w := \bigwedge\{p \mid p \in P \text{ and } w \vDash p\} \wedge \bigwedge\{\neg q \mid q \in P \text{ and } w \vDash \neg q\}$$

Proposition 3.1.3. *Given a strict order δ over W, there exists a set of formulae Δ^δ from which we can generate δ.*

Proof.
Assume a strict order δ over W.
For every valuation $w \in W$ we define a formula β_w such that:

$$\beta_w := \alpha_w \vee \bigvee_{v \prec_\delta w} \{\alpha_v\}$$

β_w is the formula that characterizes the subset of W composed by w and all the worlds below it, i.e. it is satisfied only by those worlds.
Now we define the set of default-assumptions Δ^δ as the set of all the formulae β_w:

$$\Delta^\delta = \{\beta_w | w \in W\}$$

Let δ_{Δ^δ} be the order generated by Δ^δ: $\delta_{\Delta^\delta} = \{(w, v) \in W \times W | \Delta^\delta_w \supset \Delta^\delta_v\}$.
We need to show that $\delta = \delta_{\Delta^\delta}$.

($\Rightarrow$): $(w, v) \in \delta \ \Rightarrow \ (w, v) \in \delta_{\Delta^\delta}$
Suppose $(w, v) \in \delta$. We have to show that $\Delta^\delta_w \supset \Delta^\delta_v$, i.e. if $\phi \in \Delta^\delta_v$, then $\phi \in \Delta^\delta_w$, and that there is a formula ψ s.t. $\psi \in \Delta^\delta_w$ and $\psi \notin \Delta^\delta_v$.
Suppose that $\beta_u \in \Delta^\delta_v$ ($v \vDash \beta_u$) for some $\beta_u \in \Delta$. Then , by the definition of β_u, $v = u$ or $v \prec_\delta u$.
In both cases, since $w \prec_\delta v$, we have that $w \prec_\delta u$, therefore $w \vDash \beta_u$, i.e. $\beta_u \in \Delta^\delta_w$.
So for every $\beta_u \in \Delta^\delta$, we have that if $\beta_u \in \Delta^\delta_v$, then $\beta_u \in \Delta^\delta_w$.
Since $w \vDash \beta_w$ and $w \prec_\delta v$, we have that $v \neq w$ and $v \not\prec_\delta w$. But this means

that $v \nvDash \beta_w$ ($\beta_w \notin \Delta^\delta_v$).
Therefore $\Delta^\delta_w \supset \Delta^\delta_v$, and $(w, v) \in \delta_{\Delta^\delta}$.

($\Leftarrow$): $(w, v) \in \delta_{\Delta^\delta} \;\Rightarrow\; (w, v) \in \delta$
Suppose $(w, v) \notin \delta$. We can have either $w = v$ or $w \neq v$.
In the former case $\Delta^\delta_w = \Delta^\delta_v$, and so $(w, v) \notin \delta_{\Delta^\delta}$.
in the latter case, since $w \neq v$ and $w \not<_\delta v$, then $w \nvDash \beta_v$, while $v \vDash \beta_v$. Hence $\Delta^\delta_w \not\supset \Delta^\delta_v$ and $(w, v) \notin \delta_{\Delta^\delta}$.
■

We will refer to the set Δ^δ defined in the proof as the *characteristic set* of an ordering δ, and every formulae β_w as a *characteristic default* of δ.

So, we have shown that for every default-assumption system we can generate a corresponding preferential model and conversely. Now, we can show the correspondence between default-assumption and preferential inference relations, for the finite case.

Theorem 3.1.4. *Let ℓ be a logically finite propositional language. Given an arbitrary default-assumption consequence relation $\mid\!\sim_\Delta$ defined over ℓ, we can define a preference consequence relation $\mid\!\sim_{\delta_\Delta}$ s.t. $A \mid\!\sim_\Delta \phi$ iff $A \mid\!\sim_{\delta_\Delta} \phi$, and, conversely, given an arbitrary preference consequence relation $\mid\!\sim_\delta$ defined over ℓ, we can define a default-assumption consequence relation $\mid\!\sim_{\Delta^\delta}$ s.t. $A \mid\!\sim_{\Delta^\delta} \phi$ iff $A \mid\!\sim_\delta \phi$.*

Proof.
That's obvious from the previous propositions. By Proposition 3.1.2, we know that a preferential consequence relation defined by an order δ and the default-assumption consequence relation defined by a set Δ from which δ is generated behaves identically.
By Definition 3.1.1, we know that, given a default-assumption set, we can generate a strict order.

By Proposition 3.1.3, we know that, given a strict order, we can find a default-assumption set from which it is generated.
■

All these results can be generalized to a preorder (i.e. reflexive, transitive relations) ε.
As usual, we shall write $w \prec_\varepsilon v$ if $w \preceq_\varepsilon v$ and $v \not\preceq_\varepsilon w$.
In order to generalize previous results it will be enough to restate the definition of generated orders.

Definition 3.1.2 (Generated preorder)**.** *Given a set* Δ *of formulae, we say that a relation* ε *is generated by* Δ *(written* ε_Δ*) iff*
$\varepsilon = \{(w, v) \in W \times W \mid \Delta_w \supseteq \Delta_v\}$
or equally
$\varepsilon = \{(w, v) \in W \times W \mid v \models \psi \;\Rightarrow\; w \models \psi \textit{ for every } \psi \in \Delta\}$.

One can immediatly see that ε_Δ is an extension of δ_Δ:

$$\varepsilon_\Delta = \delta_\Delta \cup \{(w, v) \mid \Delta_w = \Delta_v\}$$

Keeping all the other definitions fixed, we can restate the previous results for ε_Δ.

Lemma 3.1.5. *Let* ε *be a preorder generated by a set of assumption* Δ*, that is,* $\varepsilon = \varepsilon_\Delta$*.*
Then, for every set of premises A*,* $w \in min_\varepsilon([A]_W)$ *iff* Δ_w *is a maximally* A*-consistent subset of* Δ*.*

Proposition 3.1.6. *For every premise set* A*, we have that*

$$A \mid\!\sim_\Delta \phi \quad \textit{iff} \quad A \mid\!\sim_{\varepsilon_\Delta} \phi$$

Proposition 3.1.7. *Given a preorder ε over W, there is a set of formulae Δ^{ε} from which we can generate ε ($\varepsilon = \varepsilon_{\Delta^{\varepsilon}}$).*

For this proposition it's sufficient to restate β_w-formulae as

$$\beta_w := \bigvee_{v \preceq_{\varepsilon} w} \{\alpha_v\}$$

Theorem 3.1.8. *Given an arbitrary default-assumption consequence relation $\mid\!\sim_{\Delta}$, we can define a preferential consequence relation $\mid\!\sim_{\varepsilon_{\Delta}}$ s.t. $A \mid\!\sim_{\Delta} \phi$ iff $A \mid\!\sim_{\varepsilon_{\Delta}} \phi$, and, conversely, given an arbitrary preference consequence relation $\mid\!\sim_{\varepsilon}$, we can define a default-assumption consequence relation $\mid\!\sim_{\Delta^{\varepsilon}}$ s.t. $A \mid\!\sim_{\Delta^{\varepsilon}} \phi$ iff $A \mid\!\sim_{\varepsilon} \phi$*

We omit the proofs of these propositions, because they are analogous to the previous ones.

We can generalize these results, in order to comprehend also systems equipped with a background knowledge, i.e. a set of indefeasible formulae K, representing background 'hard' information of the agent, laws that it maintains as undoubtedly true.
So, we could model a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, s.t. K participates both in the consistency check and in the derivation of presumptive conclusions[1]. It is sufficient to restate some definitions.

Definition 3.1.3 (maximally A-consistent sets)**.** *Given a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$ and a set of premises A, we say that a set Δ' is a* maximally A-consistent *subset of Δ iff $\Delta' \cup K$ is consistent with A and for no Δ'' s.t. $\Delta' \subset \Delta'' \subseteq \Delta$, $\Delta'' \cup K$ is consistent with A.*

Definition 3.1.4 (Default-assumption inference relation)**.** *Given a default-assumptions system $\mathfrak{S} = \langle K, \Delta \rangle$, β is a default-assumption consequence of*

[1] Such a possibility was contemplated also in [45], but in a different way: since the interest of the author was in modeling abduction, in his model, K participates to the consistency check, but not in the definition of the consequences.

the set of premises A, (written $A \mid\!\sim_{\langle K,\Delta\rangle} \beta$) if and only if β is a classical consequence of the union of A and every maximally A-consistent subset of Δ and the knowledge set K.

$$A \mid\!\sim_{\langle K,\Delta\rangle} \beta \quad \textit{iff} \quad A \cup K \cup \Delta' \vDash \beta \textit{ for every maximally } A\textit{-consistent } \Delta' \subseteq \Delta$$

Obviously, the basic formulation of default-assumption consequence relation of Definition 2.3.3 corresponds to the consequence relation defined by the system $\mathfrak{S} = \langle \emptyset, \Delta \rangle$, i.e. the default-assumption system with no background knowledge.
It is easy to show that every inference relation $\mid\!\sim_{\langle K,\Delta\rangle}$ corresponds exactly to the inference relation $\mid\!\sim_{\Delta}$, putting K between the premises.

Lemma 3.1.9. *Given a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, $A \mid\!\sim_{\langle K,\Delta\rangle} \beta$ iff $A \cup K \mid\!\sim_{\Delta} \beta$ for every set of premises A.*

Proof. It is immediate from Definition3.1.3 that Δ' is a maximally A-consistent subset of Δ in $\mathfrak{S} = \langle K, \Delta \rangle$ iff it is a maximally $A \cup K$-consistent subset of Δ in the system $\mathfrak{S}' = \langle \emptyset, \Delta \rangle$.
So, the set of defeasible consequences of A in $\mathfrak{S}$ corresponds exactly to the set of defeasible consequences of $A \cup K$ in $\mathfrak{S}'$, since they are defined exactly in the same way:

$$A \cup K \cup \Delta' \vDash \beta \text{ for every maximally } A\text{−consistent } \Delta' \subseteq \Delta$$

■

In the preferential systems, we can assume the background knowledge K restricting the set of valuation available in the model to the set $U_K = \{w \in W | w \vDash \phi \text{ for every } \phi \in K\}$. Correspondingly, given a model $\mathfrak{M} = \langle U, \delta \rangle$, with $U \subseteq W$, we can define the knowledge set of the agent by means of the disjunction of all the characteristic formulae on the worlds in U: $K_U = \bigvee\{\alpha_w | w \in U\}$.

As above, we can show that every such inference relation $\mid\!\sim_{U,\prec}$ behaves exactly as the previous ones, putting K_U between the premises.

Lemma 3.1.10. *Given a preferential system* $\mathfrak{M} = \langle U, \delta \rangle$, $A \mid\!\sim_{\langle U,\delta\rangle} \beta$ *iff* $A \cup K_U \mid\!\sim_{\langle W,\delta\rangle} \beta$ *(where* W *is the set of all the possible valuations of our language), for every set of premises* A.

Proof. We simply have to show that $min_{\langle U,\delta\rangle}(A) = min_{\langle W,\delta\rangle}(A \cup K_U)$. Since a world w satisfies K_U iff $w \in U$, it is immediate to see that $w \in min_{\langle W,\delta\rangle}(A \cup K_U)$ iff it is a minimal world in U satisfying A, i.e. $w \in min_{\langle U,\delta\rangle}(A)$. ■

From these results, we can generalize the Theorem 3.1.4.

Theorem 3.1.11. *Given an arbitrary default-assumption inference relation* $\mid\!\sim_{\langle K,\Delta\rangle}$, *we can define a preferential inference relation* $\mid\!\sim_{\langle U_K,\delta_\Delta\rangle}$ *s.t.* $A \mid\!\sim_{\langle K,\Delta\rangle} \phi$ *iff* $A \mid\!\sim_{\langle U_K,\delta_\Delta\rangle} \phi$, *and, conversely, given an arbitrary preferential inference relation* $\mid\!\sim_{\langle U,\delta\rangle}$, *we can define a default-assumption inference relation* $\mid\!\sim_{\langle K_U,\Delta^\delta\rangle}$ *s.t.* $A \mid\!\sim_{\langle K_U,\Delta^\delta\rangle} \phi$ *iff* $A \mid\!\sim_{\langle U,\delta\rangle} \phi$

So, given a logically finite language, there is a perfect correspondence between Shoham's preferential systems and Default-assumption systems.

3.2 Default-assumption and default-rule inference relations

The connection between default-assumption consequence relation and Reiter's approach is much more feeble, holding only between default-assumption systems and Reiter's systems with elementary rules of form $\frac{\top:\omega}{\omega}$ (normally, if it is consistent to assume ω, assume ω).
The following is an alternative proof, based on the inductive definition of

default-rule consequence, to the one given in [45].
Assume we have a default-assumption system defined by a finite set of default formulae $\Delta = \{\alpha_1, \ldots, \alpha_n\}$, and a default-rule system defined by a set of rules $R = \{\frac{\top:\alpha_1}{\alpha_1}, \ldots, \frac{\top:\alpha_n}{\alpha_n}\}$, s.t. the justifications/consequences of the rules in R are the same formulae contained in Δ.
It is sufficient to show that the set of extensions of these two systems correspond to each other.

Lemma 3.2.1. *Assume a set of premises A, a default-assumption system $\mathfrak{S}_\Delta$ with $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ and a default-rule system S_R with $R = \{\frac{\top:\alpha_1}{\alpha_1}, \ldots, \frac{\top:\alpha_n}{\alpha_n}\}$. Then a set of formulae E is an extension of A in $\mathfrak{S}_\Delta$ iff it is an extension of A in S_R.*

Proof.

($\Leftarrow$): Following the inductive definition 2.3.8, every extension of S_R is generated by an ordering $\langle R \rangle$ over the set of rules. Working with rules of form $\frac{\top:\omega}{\omega}$, the precondition $\top$ is always satisfied, and the inductive condition for building of the extension E is just:

- set $A_0 = A$.
- Break the definition of A_{n+1} into two cases.
  Case 1. Suppose there is a rule $\frac{\top:\omega}{\omega} \in \langle R \rangle$ such that $\omega \notin A_n$ and ω is consistent with A_n. Then choose the first such rule and put $A_{n+1} = Cn(A_n \cup \omega)$.

  Case 2. Suppose there is no such rule. Then put $A_{n+1} = A_n$.
- Put $E = \bigcup\{A_i\}$

Designate by $Cons_R$ the set of the consequents of the rules in R. Since the procedure iterates over the elements of R, it triggers a set of rules R' s.t. $R' \subseteq R$. It is immediate to see that we obtain an extension $E =$

$Cn(A \cup Cons_{R'})$, where $Cons_{R'}$ is an A-maxiconsistent subset of $Cons_R$, i.e., since $Cons_R = \Delta$, we obtain a default-assumption extension of A, given a set of defaults Δ.

($\Rightarrow$): We have an A-maxiconsistent subset Δ' of Δ, defining an extension E of A in $\mathfrak{S}_\Delta$ ($E = Cn(A \cup \Delta')$). To see that it is also an extension of A in S_R, it is sufficient to take whichever ordering $\langle R \rangle$ s.t. it begins the sequence of rules with those having as justifications/consequents the elements of Δ', followed by the rules having as justifications/consequents the elements of Δ/Δ'.
■

From the lemma above we immediately obtain the correspondence of the two inference operators.

Theorem 3.2.2. *Given a default-assumption system* $\mathfrak{S}_\Delta$ *with* $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ *and a default-rule system* S_R *with* $R = \{\frac{\top:\alpha_1}{\alpha_1}, \ldots, \frac{\top:\alpha_n}{\alpha_n}\}$, $C_\Delta(A) = C_R(A)$ *for every set of formulae* A.

Proof.
It is immediate, given that, for every A, they have the same extensions and that $C_\Delta(A)$ and $C_R(A)$ are defined by the intersection of such extensions.
■

So there is a (very) basic correspondence also between default-assumption and default-rule approaches. However, such correspondence seems not to be extendable to more complex structures. The behaviour of default rules is particular mainly because we are using meta-linguistic rules instead of material implications inside the language: while the formers do not participate in the reasoning processes, unless their precondition is satisfied and they can be triggered, a material implication does not have to be 'triggered' to participate to the derivations, and we can reason freely with it by means of a lot

of classical operations, like contraposition, left strengthening and so on...
In this particular case, we have been able to face default-assumption with Reiter's approach just because the precondition of every rule was $\top$. $\top$ is always satisfied and the check for the activation of the rule was just of a consistency kind, as in the other two approaches.

Chapter 4

The preferential approach

Abstract. We briefly present the main results in the study of defeasible reasoning from a consequentialist point of view.

4.1 The KLM models

As we have seen in Chapter 2, Makinson has isolated a set of conditions that should be satisfied by an inference relation modeling defeasible reasoning. Kraus, Lehmann and Magidor [25], starting from Shoham's preferential models (see Sect. 2.3.3), have shown that possible-worlds settings are an optimal tool for formalizing defeasible reasoning starting from the desired properties of the inference relations, and they have investigated which classes of semantical models can represent particular classes of defeasible inference relations. The move from the syntactical level to the semantical one is made by interpreting an agent's assetion:

$$\text{`If } \alpha, \text{ then normally/usually } \beta\text{'}$$

as meaning

'In all the most expected/typical situations in which α is true, β is also true.'

Kraus, Lehmann and Magidor have keyed out the class of inference relations satisfying a series of minimal required properties:

REF	$\alpha \mathrel{\mid\!\sim} \alpha$	Reflexivity
LLE	$\dfrac{\alpha \mathrel{\mid\!\sim} \gamma \qquad \models \alpha \leftrightarrow \beta}{\beta \mathrel{\mid\!\sim} \gamma}$	Left Logical Equivalence
RW	$\dfrac{\alpha \mathrel{\mid\!\sim} \beta \qquad \beta \models \gamma}{\alpha \mathrel{\mid\!\sim} \gamma}$	Right Weakening
CT	$\dfrac{\alpha \mathrel{\mid\!\sim} \beta \qquad \alpha \wedge \beta \mathrel{\mid\!\sim} \gamma}{\alpha \mathrel{\mid\!\sim} \gamma}$	Cut (Cumulative Transitivity)
CM	$\dfrac{\alpha \mathrel{\mid\!\sim} \beta \qquad \alpha \mathrel{\mid\!\sim} \gamma}{\alpha \wedge \beta \mathrel{\mid\!\sim} \gamma}$	Cautious Monotony

This class of inference relations represents the class satisfying *inclusion*, *cumulativity* and *supraclassicality* (see section 2.4.2), all core properties of defeasible reasoning.

Definition 4.1.1 (Cumulative Inference Relations)**.** *An inference relation* $\mathrel{\mid\!\sim}$ *is* cumulative *iff it is closed under REF, LLE, RW, CM, CT.*

At the same time, they have defined a class of possible-worlds models, and a semantical notion of consequence relation related to them.

Definition 4.1.2 (Cumulative Models)**.** *A* cumulative *model* $\mathfrak{M}$ *is a triple* $\langle S, l, \prec \rangle$*, where*
- S *is a set, the elements of which are called states,*
- $l : S \mapsto W$ *is a labeling function assigning a propositional valuation (a world) to each state, and*
- $\prec$ *is a binary relation satisfying the* smoothness condition *(see Def. 2.3.11).*

As in the case of Shoham's model, $\prec$ is interpreted as a normality order over the set of states, i.e. $s \prec t$ means that 'The situation described by the state s is more normal than the situation described by the state t'. Also the consequence relation is defined analogously to Shoham's.

Definition 4.1.3. *Let $\mathfrak{M} = \langle S, l, \prec \rangle$ be a* cumulative *model. The consequence relation defined by $\mathfrak{M}$ (we will write $\mid\!\sim_{\mathfrak{M}}$) is defined by:*

$$\alpha \mid\!\sim_{\mathfrak{M}} \beta \text{ iff for every state } s \in \ S, \text{ if } s \vDash_{\prec} \alpha, \text{then } s \vDash \beta.$$

That is, every minimal (preferred/most normal) state in $[\alpha]_S$ satisfies β, i.e. β is normally true given α.

Kraus, Lehmann and Magidor have shown that the class of cumulative inference relations is completely represented by the class of cumulative models.

Theorem 4.1.1 (Representation Theorem for Cumulative Relations)**.** *[[25], Theorem 3.25]*
An inference relation $\mid\!\sim$ is cumulative iff it is defined by some cumulative model (i.e. it corresponds to a consequence operation $\mid\!\sim_{\mathfrak{M}}$ generated by some cumulative model $\mathfrak{M}$).

In the cumulative framework, we can derive easily also other intuitive structural rules (see [25], Lemma 3.3):

- Right conjunction (AND):

$$\frac{\alpha \mid\!\sim \beta \qquad \alpha \mid\!\sim \gamma}{\alpha \mid\!\sim \beta \wedge \gamma}$$

- $\mid\!\sim$-equivalence:

$$\frac{\alpha \mid\!\sim \beta \quad \beta \mid\!\sim \alpha \quad \alpha \mid\!\sim \gamma}{\beta \mid\!\sim \gamma}$$

◦ Modus Ponens (MP):

$$\frac{\alpha \mid\!\sim \beta \rightarrow \gamma \qquad \alpha \mid\!\sim \beta}{\alpha \mid\!\sim \gamma}$$

We can obtain a new fundamental inference relation if we add the OR property to cumulative $\mid\!\sim$:

$$\text{OR} \quad \frac{\alpha \mid\!\sim \gamma \qquad \beta \mid\!\sim \gamma}{\alpha \vee \beta \mid\!\sim \gamma} \quad \text{Disjunction in the Premises}$$

Definition 4.1.4 (KML-Preferential Inference Relations)**.** *An inference relation* $\mid\!\sim$ *is* KML-preferential *iff it is closed under REF, LLE, RW, CM, CT, and OR.*

As we have seen (see Sect.2.4.2), OR is an intuitive property, associated with the distribution rule, that validates the 'proof by cases' technique. To represent KML-preferential relations, it is sufficient to restrict the orders between the states, imposing transitivity and irreflexivity.

Definition 4.1.5 (KLM-Preferential models)**.** *A KLM-preferential model* $\mathfrak{M}$ *is a triple* $\langle S, l, \prec \rangle$ *where*
- S *is a set, the elements of which are called states,*
- $l : S \mapsto W$ *assigns a world to each state, and*
- $\prec$ *is a strict partial order on* S *(i.e. an irreflexive, transitive relation), satisfying the* smoothness condition.

Such a semantical bond over $\prec$ sounds very intuitive, corresponding to our standard use of the relation 'to be more normal than'. If a situation x is considered more normal than a situation y and, in turn, y is considered more normal than a situation z, the application of transitivity is spontaneous: x is more normal than z.
Given a definition of the inference relation $\mid\!\sim_{\mathfrak{M}}$, generated by a KML model,

identical to the definition of the inference relation in the cumulative case, Kraus, Lehmann and Magidor have proved a representation result between such classes.

Theorem 4.1.2 (Representation Theorem for KML-Preferential Relations)**.** *[[25], Theorem 5.18]*
An inference relation $\mid\!\sim$ is preferential iff it is defined by some KLM-preferential model (i.e. it corresponds to a consequence operation $\mid\!\sim_{\mathfrak{M}}$ generated by some KML-preferential model $\mathfrak{M}$).

As can be seen, Shoham and KLM's semantical approach to defeasible reasoning readapt for our interests the classical semantical definition of logical consequence by Tarski: we move from an 'all-or-nothing' notion, in which *every* model satisfying the set of premises has to satisfy also the consequent, to a more specific notion, in which only some particular models of the premises (the 'most normal' ones) are called into question.
As can be noted by the definitions above, the models proposed by Kraus, Lehmann and Magidor generalize Shoham's original proposal (see Section 2.3.3), distinguishing between 'states' (elements of the dominion S, and arguments of the relation $\prec$) and 'worlds' (propositional valuation). Such a move allows for the repetition of the same valuations in different positions in the model, since the labeling function l permits to associate the same classical valuation to different states in S. This is needed in order to obtain the representation results, since there are preferential relations not representable by models with a unique copy for each valuation (see [25], p.193, for an example). However, it is quite counterintuitive to place the same situation in two different positions in a normality ordering, and this sounds a bit like a technical trick to guarantee the representation result. Theoretically, this move has been justified by declaring that two copies of the same valuation do not represent the same situation, but they have to be assumed as differentiated by 'inexpressible' information; such justification does not sound as a very solid one.

However, models without multiple copies of the same valuations are identified as *injective* models, since the function l is injective. We will meet them in the following chapters.
From the six KML conditions we can derive also the Hard Half of the Deduction Theorem (see [25], Lemma 5.2):

$$\text{S} \quad \frac{\alpha \mid\!\sim \beta}{\mid\!\sim \alpha \rightarrow \beta} \quad \text{Conditionalization}$$

A word apart is needed for *consistency preservation*,

$$\text{CP} \quad \frac{\alpha \mid\!\sim \bot}{\alpha \models \bot} \quad \text{Consistency preservation}$$

which states that our notion of consistency is the same as that of classical consequence. Our model is consistency preserving only if there is at least one copy of each valuation of our language.

Lemma 4.1.3. *Assume a finite language ℓ. A preferential relation $\mid\!\sim$ is consistency preserving iff in its canonical model $\mathfrak{M}$ there is at least one copy of every valuation of our language ℓ.*

Proof.
($\Rightarrow$): Let $\alpha \not\models \bot$. Then there is at least a valuation w s.t. $w \models \alpha$. Since at least a copy of w is present in $\mathfrak{M}$, and, since $\mathfrak{M}$ respects the smoothness condition, we have $min_{\prec}(\alpha) \neq \emptyset$, i.e. $\alpha \not\mid\!\sim \bot$.
($\Leftarrow$): Assume there is a classical valuation w not present in $\mathfrak{M}$ and consider its characteristic formula α^w. We have $\alpha^w \mid\!\sim \bot$, since there is not a state in $\mathfrak{M}$ satisfying it, but $\alpha^w \not\models \bot$, since it is satisfied by the classical valuation w. ■

However, we can define a kind of consistency preservation with respect to the 'monotonic core' of our inference relation $\mid\!\sim$. First of all we can define the notion of knowledge for preferential models, i.e. a fixed and indefeasible set of formulas which the agent maintains as a set of background laws it has to respect. So, knowledge is the 'hard', indefeasible portion of the agent's information, s.t. the agent takes under consideration only situations satisfying it. On the semantical side, we can identify the set of known formulas as the set of formulas satisfied by every valuation in the model (that is to say, we determine the known formulas treating the model as a classical universal epistemic model, where a formula is known by an agent if and only if it is satisfied by every valuation in the model).

Definition 4.1.6 (Knowledge set **K**)**.** *Given a preferential model $\mathfrak{M}$, its* knowledge set $\boldsymbol{K}_{\mathfrak{M}}$ *is composed by every formula that is satisfied by every state s of $\mathfrak{M}$.*

We will call a *knowledge base* a finite set of formulae A_K s.t. $Cl(A_K) = \mathbf{K}$. Given a preferential relation $\mid\sim$, and its canonical model $\mathfrak{M}_{\mid\sim}$, it is easy to prove that a formula α is in the knowledge set of $\mathfrak{M}_{\mid\sim}$ ($\alpha \in \mathbf{K}_{\mid\sim}$) iff $\neg\alpha \mid\sim \bot$.

Lemma 4.1.4. $\alpha \in \boldsymbol{K}_{\mid\sim}$ *iff* $\neg\alpha \mid\sim \bot$.

Proof.
Assume a preferential relation $\mid\sim$ and its canonical model $\mathfrak{M}_{\mid\sim}$.
$\alpha \in \mathbf{K}_{\mid\sim}$ iff every valuations in $\mathfrak{M}_{\mid\sim}$ satisfies α, i.e. there are no valuations satisfying $\neg\alpha$. This corresponds to $min_{\prec}([\neg\alpha]_S) = \emptyset$, i.e. $\neg\alpha \mid\sim \bot$.
■

We can define the *monotonic core* of an inference relation $\mid\sim$.

Definition 4.1.7 (Monotonic core of a preferential relation $\mid\sim$)**.** *The* monotonic core *of an inference relation $\mid\sim$ is the monotonic consequence operation $Cn_{\mid\sim}$ (relation $\vdash_{\mid\sim}$) obtained from Cl and the addition of $\boldsymbol{K}_{\mid\sim}$ as a set of extra-axioms:*

$$Cn_{\mid\sim}(A) = Cl(\boldsymbol{K}_{\mid\sim} \cup A) \text{ for every } A \subseteq \ell$$

$$A \vdash_{\mid\sim} \alpha \text{ iff } A \cup \boldsymbol{K}_{\mid\sim} \vDash \alpha \text{ for every } A \subseteq \ell$$

It is easy to prove that every preferential $\mid\sim$ preserves consistency with respect to its own monotonic core, as already informally shown in [38], p.81.

Theorem 4.1.5. *Given an inference relation* $\mid\sim$ *and its monotonic core* $\vdash_{\mid\sim}$*, we have:*

$$\alpha \mid\sim \bot \;\Rightarrow\; \alpha \vdash_{\mid\sim} \bot$$

Proof.
If $\alpha \mid\sim \bot$, then $\neg\alpha \in \mathbf{K}_{\mid\sim}$. Since $Cn_{\mid\sim}(\alpha) = Cl(\mathbf{K}_{\mid\sim}, \alpha)$, we have $\alpha \vdash_{\mid\sim} \bot$. ■

Moreover, it is possible to strengthen the supraclassicality of $\mid\sim$ with respect to the new monotonic consequence relation:

Theorem 4.1.6. *Given an inference relation* $\mid\sim$ *and its monotonic core* $\vdash_{\mid\sim}$*, we have:*

$$\alpha \vdash_{\mid\sim} \beta \;\Rightarrow\; \alpha \mid\sim \beta$$

Proof.
If $\alpha \vdash_{\mid\sim} \beta$, then every world in $[\alpha]_{\mathfrak{M}_{\mid\sim}}$ satisfies also β. Since $min_{\prec}([\alpha]_{\mathfrak{M}_{\mid\sim}}) \subseteq [\alpha]_{\mathfrak{M}_{\mid\sim}}$, we have $\alpha \mid\sim \beta$ ■

On the other hand, we can do an analogous characterization of the beliefs of the agent, as the formulas the agent presumes as holding in the most normal situations, in every minimal state s of the entire domain S (we could also say all the states s s.t. $s \in min_{\prec}([\top]_S)$). Obviously such set of formulas corresponds exactly to the set $\{\alpha | \top \mid\sim \alpha\}$.

Definition 4.1.8 (Belief set **B**)**.** *Given an inference relation* $\mid\sim$*, its* belief set $\boldsymbol{B}_{\mid\sim}$ *is composed by every formula that is satisfied by every minimal state* s *of* $\mathfrak{M}$*, which corresponds to say* $\boldsymbol{B}_{\mid\sim} = \{\alpha | \top \mid\sim \alpha\}$

We will call a *belief base* a finite set of formulae A_B s.t. $Cn_{\mid\sim}(A_B) = \mathbf{B}$. One can immediately see that $\mathbf{K}_{\mid\sim} \subseteq \mathbf{B}_{\mid\sim}$ for every inference relation $\mid\sim$, since if a formula α is true in every state of a model, it is obviously true in every preferred state.

4.1.1 Preferential entailment.

As we have seen, KLM-preferential relations are described completely by a structural axiom (REF) and five structural conditions (LLE, RW, CM, CT, OR).
Such conditions have the form of Horn-rules, i.e. 'If $\alpha_1 \mid\sim \beta_1, \dots \alpha_n \mid\sim \beta_n$, then $\alpha_{n+1} \mid\sim \beta_{n+1}$', with only positive instances in the premises and a single positive consequence.
It is well known that the satisfaction of Horn-conditions is closed under intersection. That is, given a family I of relations $\{\mid\sim_i \mid i \in I\}$ satisfying a set R of Horn-conditions, the relation $\mid\sim_{\bigcap I}$ obtained by the intersection of the relations in I ($\mid\sim_{\bigcap I} = \bigcap\{\mid\sim_i \mid i \in I\}$) continues to satisfy the conditions in R.
Just to make an example, take a relation $\mid\sim_1$ satisfying a rule of the form

$$\alpha_1 \mid\sim_1 \beta_1, \dots \alpha_n \mid\sim_1 \beta_n / \alpha_{n+1} \mid\sim_1 \beta_{n+1}$$

and a relation $\mid\sim_2$ satisfying the corresponding rule of the form

$$\alpha_1 \mid\sim_2 \beta_1, \dots \alpha_n \mid\sim_2 \beta_n / \alpha_{n+1} \mid\sim_2 \beta_{n+1}$$

and define a new relation $\mid\sim_3 = \mid\sim_1 \cap \mid\sim_2$.
$\mid\sim_3$ too will satisfy

$$\alpha_1 \mid\sim_3 \beta_1, \dots \alpha_n \mid\sim_3 \beta_n / \alpha_{n+1} \mid\sim_3 \beta_{n+1}$$

For every instance of the scheme, if both $\mid\sim_1$ and $\mid\sim_2$ satisfy the set of premises, they will both satisfy the conclusion, and so will $\mid\sim_3$, verifying the rule; if either $\mid\sim_1$ or $\mid\sim_2$ does not satisfy the set of premises, the same $\mid\sim_3$, again verifying the rule.

So, the intersection of a set of preferential inference relations will be a preferential relation too. Then, assume a finite set of sequents $\mathcal{B} = \{\alpha \mid\!\sim \beta, \ldots, \gamma \mid\!\sim \rho\}$, and take under consideration every preferential model satisfying the sequents in $\mathcal{B}$; every such model defines a preferential inference relation $\mid\!\sim$ s.t. $\mathcal{B} \subseteq \mid\!\sim$. The inference relation $\mid\!\sim_{\mathcal{B}}$, defined by the intersection of every preferential $\mid\!\sim$ satisfying $\mathcal{B}$, will be a preferential relation too; in particular, it will be the smallest preferential relation satisfying $\mathcal{B}$.
This allows to develop a notion of semantical entailment, with respect to KLM-preferential models, between sequents: given a set of sequents $\mathcal{B}$, we can consider as its preferential consequences the set of every sequent satisfied by every preferential model of $\mathcal{B}$. Moreover, we can reinterpret preferential structural properties as rules of proof between sequents: we can start from a finite set of sequents as premises and calculate their preferential consequences by means of REF, LLE, RW, CM, CT, OR, having a complete semantical characterization by means of the class of KLM-preferential models.
Now we are going to use $\mid\!\sim$ not as a meta-linguistic symbol, but as a connective inside the language: we define a new language ℓ' composed by all the formulas of form $\alpha \mid\!\sim \beta$, s.t. α and β are propositional formulas. We can define a consequence relation $\Vdash\!\vdash_P$, having as arguments sequents, and the associated closure operation.

Definition 4.1.9 (Preferential entailment $\Vdash\!\vdash_P$)**.** *Let $\mathcal{B}$ be a set of sequents. $\alpha \mid\!\sim \beta$ is preferentially entailed by $\mathcal{B}$, written '$\mathcal{B} \Vdash\!\vdash_P \alpha \mid\!\sim \beta$', iff it is satisfied by all preferential models of $\mathcal{B}$.*

Definition 4.1.10 (Preferential closure)**.** *The set of all sequents that are preferentially entailed by $\mathcal{B}$ will be denoted by $\mathbb{P}(\mathcal{B})$ ($\mathbb{P}(\mathcal{B}) = \{\alpha \mid\!\sim \beta | \mathcal{B} \Vdash\!\vdash_P \alpha \mid\!\sim \beta\}$).*
The preferential inference operation $\mathbb{P}(\mathcal{B})$ is called the preferential closure of $\mathcal{B}$.

From the behaviour of Horn rules under intersection and our representation result (Theorem4.1.2), we can easily obtain:

Theorem 4.1.7. *([25], Corollary 3.26, reinterpreted for preferential* $\mid\!\sim$*) Let* $\mathcal{B}$ *be a set of sequents. The following conditions are equivalent:*

1. $\mathcal{B} \Vdash\!\vdash_P \alpha \mid\!\sim \beta$
2. $\alpha \mid\!\sim \beta$ *is provable from* $\mathcal{B}$ *by means of REF, LLE, RW, CM, CT, OR.*

This means that the conditions defining the preferential consequence relations can be used as a complete axiomatic system to derive new valid sequents from a set of assumed sequents. Moreover, since the proofs are always finite, and therefore use only a finite number of premises, preferential entailment is *compact.*
We can also explicitly model the knowledge and the beliefs of an agent, using $K(\alpha)$ and $B(\alpha)$ as abbreviations respectively of $\neg\alpha \mid\!\sim \bot$ and $\top \mid\!\sim \alpha$.

It is easy to see that $\Vdash\!\vdash_P$ is a tarskian (and hence monotonic) consequence operator, since this semantic consequence operator is of 'all-or-nothing' kind, i.e. it depends upon the satisfaction of the consequent in *all* the models of the premises.

Proposition 4.1.8. $\Vdash\!\vdash_P$ *satisfies Reflexivity, Monotony and Cut.*

proof We check the three properties in turn.

- Reflexivity (REF):

$$\mathcal{B} \Vdash\!\vdash_P \alpha \mid\!\sim \beta \text{ for every } \alpha \mid\!\sim \beta \in \mathcal{B}$$

  This property is obviously satisfied.

◦ Monotony (MON):

$$\frac{\mathcal{B} \Vvdash_P \alpha \mathrel{|\!\sim} \beta}{\mathcal{B}, \gamma \mathrel{|\!\sim} \delta \Vvdash_P \alpha \mathrel{|\!\sim} \beta}$$

If $\alpha \mathrel{|\!\sim} \beta$ is satisfied by every model satisfying $\mathcal{B}$, it will be obviously satisfied in every model satisfying $\mathcal{B} \cup \{\gamma \mathrel{|\!\sim} \delta\}$, since they form a subset of the former ones.

◦ Cut (CT):

$$\frac{\mathcal{B}, \gamma \mathrel{|\!\sim} \delta \Vvdash_P \alpha \mathrel{|\!\sim} \beta \qquad \mathcal{B} \Vvdash_P \gamma \mathrel{|\!\sim} \delta}{\mathcal{B} \Vvdash_P \alpha \mathrel{|\!\sim} \beta}$$

If $\mathcal{B} \Vvdash_P \gamma \mathrel{|\!\sim} \delta$, then the set of models satisfying $\mathcal{B}$ and the set of models satisfying $\mathcal{B} \cup \{\gamma \mathrel{|\!\sim} \delta\}$ are identical.

■

KML's results have been felt as a main turn in the study of defeasible reasoning: they have shown that it is possible to develop nonmonotonic systems on the basis of theoretical desiderata, and moreover they have developed a complete calculus for defeasible reasoning. Their preferential system is considered the main reference point in the consequentialist approach.

4.1.2 Other Horn-conditions

The general feeling is that the KML conditions exhaust the set of interesting Horn-conditions for the characterization of defeasible reasoning.
There are other classical horn conditions to be considered:

◦ Easy half of the deduction theorem (EHD):

$$\frac{\alpha \mid\sim \beta \rightarrow \gamma}{\alpha \wedge \beta \mid\sim \gamma}$$

- Transitivity (T):

$$\frac{\alpha \mid\sim \beta \qquad \beta \mid\sim \gamma}{\alpha \mid\sim \gamma}$$

- Contraposition (CONTR):

$$\frac{\alpha \mid\sim \beta}{\neg\beta \mid\sim \neg\alpha}$$

These three conditions cannot be taken under consideration, since they are logically connected with monotonicity.

Proposition 4.1.9. *([25], Lemma 3.4) Given cumulativity and supraclassicality, the conditions of Monotonicity, EHD, and Transitivity are all equivalent.*

Proposition 4.1.10. *([25], Lemma 3.5) Given LLE and RW, the condition of Contraposition implies the condition of Monotonicity.*

So, we have to avoid consequence relations satisfying these three rules. There are other possible Horn-conditions between cautious monotony and monotony (i.e. implied by MON and not by CM), but these are more of a technical than a conceptual interest (see [4], sect.8).
The only one that could be interesting is Conjunctive Insistence.

- Conjunctive Insistence (CI):

$$\frac{\alpha \mid\sim \gamma \qquad \beta \mid\sim \gamma}{\alpha \wedge \beta \mid\sim \gamma}$$

However, it is easy to find counterexamples to the rationality of such a rule.
Assume Steve has to study some texts and to do a little lab project to pass his biology class, but he knows that the teacher is not very demanding; so he believes that if he studies the texts just superficially, he should presumably pass the exam the same, and that concluding his project hastily, he should presumably pass the exam the same. However, he would be irresponsible in thinking that studying superficially the texts, and concluding hastily his project, he should presumably pass the exam.
So, generally, we consider as exhausted the set of interesting Horn-conditions in the field of defeasible reasoning.

4.2 Non-Horn Conditions

A good reasoner should always support the KLM-preferential rules, and there does not seem to be other Horn-conditions to be included between our desiderata. However, such rules often appear insufficient for the characterization of an efficient rational behaviour, and we can identify a series of other desirable conditions which do not have a Horn-form, i.e. which have one or more premises in negative form:

$$\alpha_1 \mid\sim \beta_1, \ldots \alpha_i \mid\sim \beta_i, \alpha_j \not\mid\sim \beta_j, \ldots \alpha_n \not\mid\sim \beta_n / \alpha_{n+1} \mid\sim \beta_{n+1}$$

Or, equivalently, they have multiple conclusions, to be interpreted disjunctively:

$$\alpha_1 \mid\sim \beta_1, \ldots \alpha_i \mid\sim \beta_i / \alpha_j \mid\sim \beta_j, \ldots \alpha_n \mid\sim \beta_n, \alpha_{n+1} \mid\sim \beta_{n+1}$$

The meaning of such kind of rules is that we can derive that some relations do hold not only from the presence of some other sequents, but also from the absence of them. With non-Horn rules, ignorance plays a direct role in defeasible reasoning.
It is important to note that such non-Horn conditions are not closed under

intersection as the Horn ones. For example, take the following instance of a rule:

$$\alpha \mid\sim \beta / \gamma \mid\sim \delta, \pi \mid\sim \rho$$

Assume a relation $\mid\sim_1$ s.t.

$$\alpha \mid\sim_1 \beta, \gamma \mid\sim_1 \delta, \text{ and } \pi \not\mid\sim_1 \rho$$

and a relation $\mid\sim_2$ s.t.

$$\alpha \mid\sim_2 \beta, \gamma \not\mid\sim_2 \delta, \text{ and } \pi \mid\sim_2 \rho$$

They both validate the rule. Intersecting the two, we obtain a new relation $\mid\sim_3 = \mid\sim_1 \cap \mid\sim_2$, s.t.

$$\alpha \mid\sim_3 \beta, \gamma \not\mid\sim_3 \delta, \text{ and } \pi \not\mid\sim_3 \rho$$

$\mid\sim_3$ does not satisfy the above rule anymore.

We are going to present briefly some desirable Non-horn rules, stating the known representation results and the dependencies between them. They are all weakened forms of the classical structural rule (monotony, transitivity, contraposition, deduction,...).

4.2.1 Rational monotony

The 'patrician' rule between Non-horn rules is surely *Rational Monotony*:

- Rational Monotony (RM):

$$\frac{\alpha \mid\sim \gamma \qquad \alpha \not\mid\sim \neg\beta}{\alpha \wedge \beta \mid\sim \gamma}$$

It is a very strong rule: it allows to add whichever new premise β assuming monotonicity ($\alpha \mid\sim \gamma \;\Rightarrow\; \alpha \wedge \beta \mid\sim \gamma$), unless we are explicitly aware that such a formula would be exceptional in the actual situation ($\alpha \mid\!\sim \neg\beta$). Probably, it sounds also more intuitive in the form

$$\frac{\alpha \mid\!\sim \gamma \qquad \alpha \wedge \beta \not\mid\!\sim \gamma}{\alpha \mid\!\sim \neg\beta}$$

If we think that when Susan sings, normally she is very good, but that if she sings and she has a cold, normally she is awful, it is natural to conclude that when Susan sings, normally she does not have a cold.
We define a *Rational Inference Relation* as a KLM-preferential relation satisfying also RM. In [31], the authors have proved that the class of rational relations is semantically characterized by the class of *modular* preferential models, i.e. models characterized by a ranked preference relation.

Definition 4.2.1. *A partial order $\prec$ over a set S is* ranked *iff there is a totally ordered set Ω (the strict order on Ω will be denoted by $<$) and a function $r : S \mapsto \Omega$ (the ranking function) s.t. $s \prec t$ iff $r(s) < r(t)$.*

So, a ranked $\prec$ generates a model partitioned in a hierarchy of sets of states. Informally, this has to be interpreted as saying that our notion of normality is organized in a linear 'scale', and every situation finds its place in such a scale of normality. Whether such a characterization of normality, ordered on a linear scale, sounds an intuitive model of our everyday notion of normality is an open question, left to the intuition of everyone. However, Lehmann and Magidor have shown that every rational $\mid\!\sim$ is generated by a preferential model of such a kind.

Theorem 4.2.1 (Representation Theorem for Rational Relations)**.** *[[31], Theorem 5]*
An inference relation $\mid\!\sim$ is rational iff it is defined by some modular preferential model (i.e. it corresponds to a consequence relation $\mid\!\sim_{\mathfrak{M}}$ generated by some ranked preferential model $\mathfrak{M}$).

As we have seen above, transitivity is not a desirable property, since it implies monotonicity. However, Freund, Lehmann and Morris (see [14], Theorem 2.1) have proved that, given the KLM-preferential conditions, Rational monotony is equivalent to a weakened form of transitivity:

- Weak Transitivity (WT):

$$\frac{\alpha \mid\!\sim \beta \quad \beta \mid\!\sim \gamma \quad \beta \not\mid\!\sim \neg\alpha}{\alpha \mid\!\sim \gamma}$$

That is, transitivity between α, β, and γ holds just if α does not represent an out-of-the-ordinary situation in β. Let us recall the penguin example: we know that penguins are birds ($\alpha \mid\!\sim \beta$), and that birds normally fly ($\beta \mid\!\sim \gamma$): we are forced to derive that penguins normally fly ($\alpha \mid\!\sim \gamma$), unless we are informed that birds are not normally penguins ($\beta \mid\!\sim \neg\alpha$), excluding penguins from an acritical attribution of the properties of normal birds.

From rational monotony we can derive every interesting Non-horn condition for defeasible reasoning. Lehmann and Magidor have argued in favour of conceiving rational monotony as a universally valid rationality principle, so maintaining rational inference relations as the 'core' system for defeasible reasoning. Others, as Makinson, think that rational monotony is too strong a rule, not apt to be assumed in every model of defeasible reasoning.

> "Rational monotony is too strong to insist upon, at least for relations of strong support. For we may have A strongly supporting z, and although A does not go so far as to strongly support $\neg x$, it may still suggest the possibility of $\neg x$ sufficiently to undermine the inference of z from $A \cup \{x\}$." ([38], p.97)

Sometimes it is better to add to preferential rules only some of the weaker rules derivable from RM.

Notwithstanding, rational relations remain a reference point, because RM is an intuitive rule and is generally applicable, and, moreover, the modular order of the semantical structures allows to reason on models in a simple way, as we are going to see below.

4.2.2 Derived properties

RM strictly implies some structural properties which confirm the intuitiveness of rational monotony.

- Weak Contraposition (WC):

$$\frac{\gamma \wedge \alpha \mathrel{|\!\sim} \beta \qquad \gamma \mathrel{|\!\not\sim} \beta}{\gamma \wedge \neg\beta \mathrel{|\!\sim} \neg\alpha}$$

- Weak Rational Monotony (WRM):

$$\frac{\top \mathrel{|\!\sim} \alpha \rightarrow \beta \qquad \top \mathrel{|\!\not\sim} \neg\alpha}{\alpha \mathrel{|\!\sim} \beta}$$

- Disjunctive Rationality (DR):

$$\frac{\alpha \vee \beta \mathrel{|\!\sim} \gamma \qquad \alpha \mathrel{|\!\not\sim} \gamma}{\beta \mathrel{|\!\sim} \gamma}$$

- Negation Rationality (NR):

$$\frac{\alpha \mathrel{|\!\sim} \beta \qquad \alpha \wedge \gamma \mathrel{|\!\not\sim} \beta}{\alpha \wedge \neg\gamma \mathrel{|\!\sim} \beta}$$

Weak contraposition allows the contraposition between two formulae α and β, given a background γ, just if the conclusion that β normally holds is dependent from the satisfaction of α, i.e. if we have γ, and β is negated,

presumably also α does not hold.
For example, we know that it is not normal for Bob to work on Sundays ($\gamma \not\mid\!\sim \beta$), but that, if it is Sunday and he has to finish an urgent work, he normally goes on working ($\gamma \wedge \alpha \mid\!\sim \beta$); then we can conclude that, if it is Sunday and Bob is not working, he presumably does not have any urgent work to finish ($\gamma \wedge \neg\beta \mid\!\sim \neg\alpha$).

Weak rational monotony is a weakened form of RM, and we will analyze its properties in the next chapter.
A word apart is needed for Disjunctive rationality.

Disjunctive rationality and Injectivity

The behaviour of disjunctive rationality (DR) has been analyzed deeply by Freund in [11].
Probably, its most intuitive form is

$$\frac{\beta \not\mid\!\sim \gamma \qquad \alpha \not\mid\!\sim \gamma}{\alpha \vee \beta \not\mid\!\sim \gamma}$$

That is, for example: I do not think that, given that the day is rainy, it is normally a good time for a walk ($\beta \not\mid\!\sim \gamma$), and that, given that the day is very cold, it is a good time to go for a walk ($\alpha \not\mid\!\sim \gamma$). Then I cannot think that, if the day is rainy or very cold, it is normally a good time to go for a walk ($\alpha \vee \beta \not\mid\!\sim \gamma$).
Some authors argue against the general validity of DR:

> "[...]Suppose a crime has been committed in a house where two persons x and y live. Let γ stand for 'Sherlock Holmes is interested in finding the murderer', α stand for 'y is the murderer' and β for 'x is the murderer'. Then we have $\alpha \vee \beta \mid\!\sim \gamma$, but we do not have necessarily $\alpha \mid\!\sim \gamma$ or $\beta \mid\!\sim \gamma$"([11], p.234)

However, this kind of examples refers to the defeasible management of information in a particular context, that is, in an autoepistemic multi-agent context; such auto-referentiality has been traditionally recognized as problematic in the development of logical epistemic models, and has to be often treated apart. The fact that DR raises problems in such a context does not weaken its general intuitive value in more standard situations, where the auto-epistemic dimension of reasoning does not play any role.
If we limit ourselves to the specification of defeasible reasoning about factual information, then we are not aware of any counterexample to disjunctive rationality. Also Makinson seems to agree with such an advice.

> "[...] disjunctive rationality does seem to be a *sine qua non* for a reasonable inference relation on a language without indexical connectives." ([38], p.97)

In the language of consequence operators, DR can be rewritten as

$$C(\alpha \vee \beta) \subseteq (C(\alpha) \cup C(\beta))$$

In analyzing DR, Freund has also focused on a weakened version of such property:

$$C(\alpha \vee \beta) \subseteq Cn(C(\alpha) \cup C(\beta))$$

He has proven that, given a finite language, such a condition characterizes exactly the class of the inference operations defined by *injective* models, i.e. preferential models s.t. every state is labeled by a distinct propositional valuation (see above in Sect.4.1).

Theorem 4.2.2. *([11], Theorem 4.13)*
Let ℓ be a logically finite language and $\mid\sim$ a preferential inference relation on ℓ. Then $\mid\sim$ is represented by an injective model iff $C(\alpha \vee \beta) \subseteq Cn(C(\alpha) \cup C(\beta))$ for every pair of formulas α and β.

If an inference relation $\mid\sim$ satisfies such a property, we will call $\mid\sim$ an *injective* inference relation.
Between injective models, we can identify a particular class of models.

Definition 4.2.2. *A preferential model* $\mathfrak{M} = \langle S, l, \prec \rangle$ *is said to be* filtered *if whenever two states* s *and* t *of* S *satisfy a formula* α *without being minimal in* $[\alpha]_S$*, there exists a state* r *s.t.* $r \prec s$*,* $r \prec t$ *and* $r \vDash \alpha$*.*

Injective filtered preferential models represent disjunctive inference relations.

Theorem 4.2.3. *([11], Theorem 5.2) A preferential relation* $\mid\!\sim$ *is disjunctive if and only if it can be defined by a standard filtered model.*

DR has as a special case negation rationality (NR),

$$\frac{\alpha \wedge \neg\gamma \not\mid\!\sim \beta \qquad \alpha \wedge \gamma \not\mid\!\sim \beta}{\alpha \not\mid\!\sim \beta}$$

For example, assume Bob does not believe that if tomorrow is a sunny day, he will presumably travel to Rome, and, contemporarily, he does not believe that if tomorrow is not a sunny day, he will presumably travel to Rome. Bob cannot believe that tomorrow he will presumably travel to Rome. As DR, negation rationality has a strong intuitive appeal as long as we do not have a situation with autoepistemic or other self-referential components. Makinson proposes this example:

> "Suppose, for example, that x and z are both propositions of the language, but z says that the truth value of x is undermined by the information available to us. Then we may indeed have $A \not\mid\!\sim x$, $A \not\mid\!\sim \neg x$, so that $A \mid\!\sim z$, but neither $A, x \mid\!\sim z$, nor $A, \neg x \mid\!\sim z$." ([38], p.92)

In [13] it has been proven that we can obtain DR from NR and injectivity. As we have seen, RM implies DR, which in turn implies Injectivity. So every rational relation $\mid\!\sim$ has an injective canonical model. Hence, we know that every rational relation has a ranked canonical model (not necessarily injective) and an injective canonical model. We could ask whether such an injective model is also ranked. The answer is affirmative.

Theorem 4.2.4. *([4], Corollary 4.7) Every rational inference relation* $\mid\sim$ *is generated by some injective ranked preferential model.*

4.2.3 Rational closure

It would be interesting to consider the question of building a notion of entailment also for these stronger inference relations. For example, given a set $\mathcal{B}$ of sequents, we would like to define the set of its rational consequences intersecting all the sets of sequents satisfied by every rational models of $\mathcal{B}$, analogously to the preferential entailment. Unfortunately, as we have seen above, non-Horn conditions are not preserved under intersection, and the intersection of a set of rational relations (or, in general, relations characterized by non-Horn conditions) could be not a rational relation.
In particular, Lehmann and Magidor proved that, given a set of sequents $\mathcal{B}$, and the corresponding set of those ranked models satisfying it, the notion of entailment defined by the intersection of the consequence relations satisfied by such models corresponds to the preferential entailment of $\mathcal{B}$, i.e. the only conditions preserved are the Horn ones.

Theorem 4.2.5. *([31], Theorem 6) If the sequent* $\alpha \mid\sim \beta$ *is satisfied by all* ranked *models that satisfy all the sequents in* $\mathcal{B}$*, then it is satisfied by all* preferential *such models.*

A result of this kind can be easily validated for every relation characterized by a non-Horn condition. Such results emphasize the impossibility of modeling a 'classical' notion of entailment for this kind of consequence relations, i.e. a notion of entailment defined semantically by the intersection of *all* the models satisfying the set of premises.
Lehamnn and Magidor ([31]) have proposed to determine the rational consequences of a set $\mathcal{B}$ by selecting a single rational model satisfying $\mathcal{B}$. We could say that, moving on the meta-linguistic level, we move from a *skeptical* approach (as in preferential entailment) to a *choice* approach (see Sect.2.2), focusing on a specific rational model of $\mathcal{B}$, and determining $\mathcal{B}$'s rational con-

sequences only referring to it and ignoring the other models.
So, given a set of sequents $\mathcal{B}$, we move from the problem of determining a form of closure of $\mathcal{B}$ by means of a notion of logical entailment ($\mathbb{P}(\mathcal{B})$), to the definition of a form of *rational closure* $\mathbb{R}(\mathcal{B})$ by choosing a single rational model of the set $\mathcal{B}$.
The main problem is to define a principle for selecting which of the rational expansions of $\mathcal{B}$ has to be chosen, i.e. which one contains the sequents we intuitively expect to follow from $\mathcal{B}$.
Lehmann and Magidor suggest that, if we are in a situation s.t., given $\alpha \mid\sim \beta$, we don't know which new sequent to add, $\alpha \mid\sim \neg\gamma$ or $\alpha \wedge \gamma \mid\sim \beta$, we should tend, whenever it is possible, to assume monotonicity and not to add arbitrarily new information, i.e. we should tend to add $\alpha \wedge \gamma \mid\sim \beta$ and ignore $\alpha \mid\sim \neg\gamma$.
Moreover, we should tend not to modify the belief set and the knowledge set determined by the preferential closure of $\mathcal{B}$, i.e.

$$\top \mid\sim \alpha \in \mathbb{R}(\mathcal{B}) \text{ iff } \top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$$

$$\neg\alpha \mid\sim \bot \in \mathbb{R}(\mathcal{B}) \text{ iff } \neg\alpha \mid\sim \bot \in \mathbb{P}(\mathcal{B})$$

Exceptionality ranking

Lehmann and Magidor define a notion of rational closure of a set of sequents $\mathcal{B}$ apt to respect such desiderata. First of all, they build an ordering of exceptionality for formulas and sequents.

Definition 4.2.3. *(Exceptionality) Let $\mathcal{B}$ be a set of sequents and α a formula. The formula α is said to be* exceptional *for $\mathcal{B}$ iff $\mathcal{B}$ preferentially entails the sequent $\top \mid\sim \neg\alpha$. The sequent $\alpha \mid\sim \beta$ is said to be exceptional for $\mathcal{B}$ iff its antecedent α is exceptional for $\mathcal{B}$.*

We call $E(\mathcal{B})$ the set of all the sequents of $\mathcal{B}$ exceptional for $\mathcal{B}$ itself ($E(\mathcal{B}) \subseteq \mathcal{B}$). We can define a hierarchy of decreasing subsets of $\mathcal{B}$ $R = \{C_0, \ldots, C_n\}$:

- $C_0 = \mathcal{B}$
- $C_{i+1} = E(C_i)$

If $\mathcal{B}$ is a finite set, the sequence of C_i will be finite, i.e., after some point, all Cs will be equal and *completely exceptional* (maybe empty).

Definition 4.2.4 (Rank of formulas)**.** *We shall say that a formula α has rank i for $\mathcal{B}$ iff C_i is the least set s.t. α is not exceptional for C_i (i.e. $C_i \not\Vdash_P \top \mid\sim \neg\alpha$ and, if $i \neq 0$, $C_{i-1} \Vdash_P \top \mid\sim \neg\alpha$).*
If α has rank i, then every sequent $\alpha \mid\sim \beta$ will have rank i for every β

A formula α has no rank iff it is completely exceptional, that is iff $\emptyset \Vdash_P \alpha \mid\sim \bot$ (α is $\mid\sim$-inconsistent).

On the basis of such a ranking, Lehmann and Magidor define a rational closure of a set $\mathcal{B}$ satisfying the desiderata seen above.

Definition 4.2.5 (Rational closure)**.** *Let $\mathcal{B}$ be a set of sequents. The rational closure $\mathbb{R}(\mathcal{B})$ of $\mathcal{B}$ exists and is the set of all the sequents $\alpha \mid\sim \beta$ s.t. either*

- *The rank of α is strictly less than the rank of $\alpha \wedge \neg\beta$, or*
- *α has no rank.*

So, given a set of sequents $\mathcal{B}$, it is possible to define a rational closure operation $\mathbb{R}(\mathcal{B})$ s.t. it observes some intuitive desiderata with respect to $\mathcal{B}$. We can construct a modular canonical model of $\mathbb{R}(\mathcal{B})$ by means of the complex technique used in [31], Appendix A, to prove the representation theorem for RM.
In Chapter 6 we will see how the connections with the default-assumption approach allows for a simple method to build a canonical model for the rational closure of a set $\mathcal{B}$.

Valuation of the rational closure

The quality of rational closure has been evaluated by means of a series of examples. Lehamann and Magidor in [31] have tested the behaviour of rational closure with respect to some classical example for the valuation of nonmonotonic logics.

∘ Nixon diamond.

Assume the following conditional base:

1. $\rho \mid\sim \neg\pi$
2. $\mu \mid\sim \pi$

We interpret ρ, π, μ respectively as 'being a republican', 'being a pacifist', and 'being a quaker'.
With rational closure, we obtain intuitive results, as $\rho\wedge\mu \not\mid\sim \pi$ and $\rho\wedge\mu \not\mid\sim \neg\pi$.

∘ Penguin triangle.

Assume the following conditional base:

1. $\pi \mid\sim \beta$
2. $\pi \mid\sim \neg\phi$
3. $\beta \mid\sim \phi$

Here π, β, ϕ are respectively interpreted as 'being a penguin', 'being a bird', and 'being able to fly'.
The following intuitive sequents result valid:
$\phi \mid\sim \neg\pi$, $\neg\phi \mid\sim \neg\beta$, $\neg\phi \mid\sim \neg\pi$, $\beta \mid\sim \neg\pi$, $\neg\beta \mid\sim \neg\pi$, $\beta\wedge\pi \mid\sim \neg\phi$, $\beta\wedge green \mid\sim \phi$, $\pi \wedge black \mid\sim \neg\phi$.
Instead, the following sequents are counterintuitive, and the rational closure does not endorse them:
$\beta \wedge \neg\phi \mid\sim \pi$, $\beta \wedge \neg\phi \mid\sim \neg\pi$, $\pi \mid\sim \phi$.

However, notwithstanding a generally satisfying behaviour, rational closure can return counter-intuitive results. In [44], Paris presents three examples of strange behaviour of the rational closure.
The first two examples are related to the failure of property inheritance from a class to an exceptional subclass.

(i) Consider the conditional base $\mathcal{B} = \{\beta \mid\sim \phi, \pi \mid\sim \neg\phi, \pi \wedge \neg\beta \mid\sim \bot, \beta \mid\sim \omega\}$, where β, π, ϕ, ω stand for 'bird', 'penguin', 'flies', and 'has wings' respectively. In $\mathbb{R}(\mathcal{B})$, the rational closure of $\mathcal{B}$, we have $\pi \not\mid\sim \omega$, while we would like that birds' property of 'wingedness' should be inherited by the exceptional classes, like penguins, if we are not informed of the contrary.

(ii) Consider the conditional base $\mathcal{B} = \{\sigma \mid\sim \tau, \sigma \mid\sim \phi\}$, where σ, τ, ϕ stand for 'Swedish', 'tall', and 'fair' respectively. In this case $\mathbb{R}(\mathcal{B})$ does not satisfy 'Short Swedes are usually fair' ($\neg\tau \wedge \varrho \not\mid\sim \phi$), despite the fact that one could feel that the property of fairness is sufficiently independent from the property of tallness to be inherited by the class of short Swedes.

The third example is more significative, since it shows the drawing of unjustified conclusions; while the non-derivability of desirable conclusions, as in (i) and (ii), can be justified by the absence of relevant information in the premises, the derivation of undesired conclusion emphasize a sure problem in the structure.

(iii) Take $\mathcal{B} = \{\phi \wedge \pi \mid\sim \omega, \gamma \mid\sim \neg\omega\}$, where $\phi, \pi, \omega, \gamma$ stand for 'factory worker', 'manager', 'well-off', and 'drives an economy car' respectively.
In $\mathbb{R}(\mathcal{B})$ we do not have $\phi \wedge \pi \wedge \gamma \mid\sim \omega$. However , if we expand $\mathcal{B}$ adding $\phi \mid\sim \neg\omega$, surprisingly, in $\mathbb{R}(\mathcal{B} \cup \{\phi \mid\sim \neg\omega\})$ we have $\phi \wedge \pi \wedge \gamma \mid\sim \omega$, despite the fact that the added sequent would seem to support the conclusion that managers working in factories and driving economy cars are usually not well-off.

Hence, rational closure is generally well-behaved, but shows some problems in the property-heritage from more normal to more exceptional situations. We will see this better at the end of Chapter 6.

Chapter 5

Weakly rational inference relations

Abstract. We prove a representation theorem for weakly rational inference relations.

We are going to present a representation theorem for the weakly rational inference relations. *Weak rational monotony* (WRM) is a property firstly presented in [19], and it has not received a lot of attention. We are going to investigate its properties since it will be useful in the following chapters.
We define a weakly rational inference relation $\mid\!\sim$ as a preferential relation closed under weak rational monotony, i.e. satisfying the following rules:

REF	$\alpha \mid\!\sim \alpha$	Reflexivity
LLE	$\dfrac{\alpha \mid\!\sim \gamma \qquad \models \alpha \leftrightarrow \beta}{\beta \mid\!\sim \gamma}$	Left Logical Equivalence
RW	$\dfrac{\alpha \mid\!\sim \beta \qquad \beta \models \gamma}{\alpha \mid\!\sim \gamma}$	Right Weakening
CT	$\dfrac{\alpha \mid\!\sim \beta \qquad \alpha \wedge \beta \mid\!\sim \gamma}{\alpha \mid\!\sim \gamma}$	Cautious Cut (Cautious Transitivity)
CM	$\dfrac{\alpha \mid\!\sim \beta \qquad \alpha \mid\!\sim \gamma}{\alpha \wedge \beta \mid\!\sim \gamma}$	Cautious Monotony
OR	$\dfrac{\alpha \mid\!\sim \gamma \qquad \beta \mid\!\sim \gamma}{\alpha \vee \beta \mid\!\sim \gamma}$	Disjunction in the Premises
WRM	$\dfrac{\top \mid\!\sim \alpha \rightarrow \beta \qquad \top \not\mid\!\sim \neg\alpha}{\alpha \mid\!\sim \beta}$	Weak Rational Monotony

WRM is an important property, because, combined with conditionalization, it gives a weakened form of the deduction theorem:

$$\top \not\mid\!\sim \neg\alpha \;\Rightarrow\; (\top \mid\!\sim \alpha \rightarrow \beta \;\Leftrightarrow\; \alpha \mid\!\sim \beta)$$

It is also intuitive, being interpreted as: "If in any normal situation α implies β, and there are normal situations in which α holds, then, given α, β normally holds".
We want to show that this class of consequence relations is represented by a particular class of preferential models.

Definition 5.0.6. (Optimal models)
Given a preferential model $\mathfrak{M} = \langle S, l, \prec \rangle$*, we call it optimal iff its minimal states (i.e. the set* $min_{\prec}(S)$*) is composed by states* s *s.t.* s *is preferred to every other nonminimal world in* S*.*

$$s \in min_{\prec}(S), t \in (S \setminus min_{\prec}(S)) \;\Rightarrow\; s \prec t$$

This means that a normal situation (a situation in $min_{\prec}(S)$) is considered more normal than any exceptional situation, and we will call every normal situation an *optimal* situation. This seems a very intuitive claim.

5.1 Representation theorem

We want to show that weakly rational inference relations are represented by the class of optimal preferential models. We will show a representation theorem with the same approach used in [25] for their representation theorem (Theorem 5.18).
First of all, we have to prove the soundness of our rules in every optimal model.

Lemma 5.1.1 (Soundness)**.** *For any optimal model* $\mathfrak{M}$*, the inference relation* $\mid\!\sim_{\mathfrak{M}}$ *it defines is a weakly rational inference relation.*

Proof. Since we already know that every preferential rule is sound with respect to every KLM preferential model ([25], Lemma 5.8), we need to simply show that every inference relation defined by an optimal model satisfies WRM.
Suppose $\top \mid\!\sim_{\mathfrak{M}} \alpha \rightarrow \beta$ and $\top \not\mid\!\sim_{\mathfrak{M}} \neg\alpha$. $\top \not\mid\!\sim_{\mathfrak{M}} \neg\alpha$ implies that there is at least one optimal state s s.t. $s \vDash \alpha$. Since every such state is preferred to every non-optimal world, we have that $min_{\prec}([\alpha]_S) \subseteq min_{\prec}(S)$. Since $\top \mid\!\sim_{\mathfrak{M}} \alpha \rightarrow \beta$, we have that, for every state s, if $s \vDash_{\prec} \alpha$, then $s \vDash \alpha \rightarrow \beta$, i.e.

$s \vDash \beta$. So, we have that $\alpha \mid\!\sim_{\mathfrak{M}} \beta$.

■

Now we have to show that every weakly rational inference relation $\mid\!\sim$ corresponds exactly to the inference relation $\mid\!\sim_{\mathfrak{M}}$ defined by an optimal model $\mathfrak{M}$, its canonical model.
We assume a weakly rational inference relation $\mid\!\sim$, and we introduce an ordering between the formulas of our language ℓ:

$$\alpha \leq \beta \text{ iff } \alpha \vee \beta \mid\!\sim \alpha.$$

This means that normally, if we know that $\alpha \vee \beta$ holds, we assume that α holds, i.e. that α is at least as normal as β.
It can be shown that $\leq$ is a reflexive and transitive relation ([25], Lemma 5.10).
Since from REF we have $\alpha \vee \beta \mid\!\sim \alpha \vee \beta$, and by LLE $(\alpha \vee \beta) \vee \alpha \mid\!\sim \alpha \vee \beta$, we obtain $\alpha \vee \beta \leq \alpha$ for every $\alpha, \beta \in \ell$.
We can also see that the beliefs of the agent are exactly the minimal (preferred) elements of such an ordering, i.e. that a formula α is at least as normal as β for every $\beta \in \ell$ iff $\top \mid\!\sim \alpha$.

Proposition 5.1.2. *$\alpha \leq \beta$ for every $\beta \in \ell$ iff $\top \mid\!\sim \alpha$.*

Proof.
$\Leftarrow$: If $\top \mid\!\sim \alpha$, then, by RW, $\top \mid\!\sim \alpha \vee \beta$ for every formula β. This implies, by CM, that $\alpha \vee \beta \mid\!\sim \alpha$, i.e. $\alpha \leq \beta$ for every β.
$\Rightarrow$: If $\alpha \leq \beta$ for every formula β, then we also have $\alpha \leq \top$, that implies $\alpha \vee \top \mid\!\sim \alpha$, i.e. $\top \mid\!\sim \alpha$ by LLE.

■

There are other properties of the beliefs of a preferential agent that will be useful in the following:

Lemma 5.1.3. *If $\top \mid\sim \alpha$ and $\top \mid\sim \beta$, then $C(\alpha) = C(\beta)$.*

Proof.
If $\top \mid\sim \alpha$ and $\alpha \mid\sim \gamma$, then, by CT, $\top \mid\sim \gamma$, and, by CM, $\beta \mid\sim \gamma$, hence $C(\alpha) \subseteq C(\beta)$. The same in the other direction.
■

Lemma 5.1.4. *If $\top \mid\sim \beta$ and $\alpha \leq \beta$, then $\top \mid\sim \alpha$.*

Proof.
$\top \mid\sim \beta$ implies $\top \mid\sim \alpha \vee \beta$ (by RW).
If $\alpha \leq \beta$, we have $\alpha \vee \beta \mid\sim \alpha$, which implies, by CT, $\top \mid\sim \alpha$.
■

To build the canonical model for $\mid\sim$, we have to introduce the notion of *normal world* with respect to a formula, that is, a valuation satisfying all the defeasible consequences of such formula.

Definition 5.1.1 (Normal worlds)**.** *We call the world w a* normal world *for α iff $w \vDash \beta$ for every $\beta \in \ell$ s.t. $\alpha \mid\sim \beta$.*

Kraus, Lehmann and Magidor have shown that the intersection of the formulas satisfied by every normal world of a formula α returns exactly the set of the defeasible consequences of such a formula.

Lemma 5.1.5 ([25], Lemma 3.18)**.** *Suppose a consequence relation $\mid\sim$ satisfies REF, RW and AND, and let $\alpha, \beta \in \ell$. All normal worlds for α satisfy β iff $\alpha \mid\sim \beta$.*

We point toward the construction of an optimal model s.t., for every formula α, the set of α-preferred worlds corresponds to the set of α-normal worlds.
Since a weakly rational $\mid\sim$ is also a preferential inference relation, we can use

some results from [25].

Lemma 5.1.6 ([25], Lemma 5.11)**.** *If $\alpha \leq \beta$ and w is a normal world for α that satisfies β, then w is a normal world for β.*

Lemma 5.1.7 ([25], Lemma 5.12)**.** *If $\alpha \leq \beta \leq \gamma$ and w is a normal world for α that satisfies γ, then w is a normal world for β.*

Now we are going to define a model $\mathfrak{M}$ on the basis of the relation $\leq$. We will have to prove that it is an optimal model and that it is the canonical model for our weakly rational inference relation $\mid\!\sim$.

$\mathfrak{M} = \langle S, l, \prec \rangle$, where:

(i) $S = \{\langle w, \alpha \rangle |\ w \text{ is a normal world for } \alpha\}$

(ii) $l(\langle w, \alpha \rangle) = w$

(iii) $\langle w, \alpha \rangle \prec \langle v, \beta \rangle$ iff ($\alpha \leq \beta$ and $w \not\models \beta$) or ($\top \mid\!\sim \alpha$ and $\top \not\mid\!\sim \beta$)

First of all, we have to show that $\mathfrak{M}$ is an optimal model, i.e. that $\prec$ is a strict order and that $\mathfrak{M}$ satisfies optimality and smoothness.

Lemma 5.1.8. *The relation $\prec$ is a strict partial order, i.e., it is irreflexive and transitive.*

Proof.
$\prec$ is irreflexive, since $\langle w, \alpha \rangle \prec \langle w, \alpha \rangle$ would imply $w \not\models \alpha$ and $w \models \alpha$, or $\top \mid\!\sim \alpha$ and $\top \not\mid\!\sim \alpha$.
For transitivity, assume $\langle w, \alpha \rangle \prec \langle v, \beta \rangle$ and $\langle v, \beta \rangle \prec \langle s, \gamma \rangle$. Fist of all we can see that $\top \not\mid\!\sim \beta$ and $\top \not\mid\!\sim \gamma$, otherwise $\langle w, \alpha \rangle \not\prec \langle v, \beta \rangle$ and $\langle v, \beta \rangle \not\prec \langle s, \gamma \rangle$. So $\beta \leq \gamma$ and $v \not\models \gamma$.
If $\top \mid\!\sim \alpha$, $\langle w, \alpha \rangle \prec \langle s, \gamma \rangle$ immediately.
If $\top \not\mid\!\sim \alpha$, we have $\alpha \leq \beta$ and $w \not\models \beta$.
From the transitivity of $\leq$, we obtain $\alpha \leq \gamma$.

Since $\alpha \leq \beta \leq \gamma$ and $w \not\models \beta$, we have $w \not\models \gamma$ by Lemma 5.1.7.
So, $\langle w, \alpha \rangle \prec \langle s, \gamma \rangle$.
■

We have to show that $\mathfrak{M}$ satisfies optimality. First, we characterize the set $min_{\prec}(S)$ by the following.

Lemma 5.1.9. *$\langle w, \alpha \rangle \in min_{\prec}(S)$ iff $\top \mid\!\sim \alpha$.*

Proof.

$\Rightarrow$: Assume $\top \not\mid\!\sim \alpha$. Then for every world $\langle v, \beta \rangle$ s.t. $\top \mid\!\sim \beta$ we have $\langle v, \beta \rangle \prec \langle w, \alpha \rangle$, hence $\langle w, \alpha \rangle \notin min_{\prec}(S)$.
$\Leftarrow$: Assume $\langle w, \alpha \rangle \notin min_{\prec}(S)$, i.e. there is a world $\langle v, \beta \rangle$ s.t. $\langle v, \beta \rangle \prec \langle w, \alpha \rangle$. Since $\top \mid\!\sim \alpha$, the second claim in the definition of $\prec$ cannot be satisfied, and so $\beta \leq \alpha$ and $v \not\models \alpha$.
By Lemma 5.1.4, we wave that if $\beta \leq \alpha$ and $\top \mid\!\sim \alpha$, we obtain $\top \mid\!\sim \beta$. So, by lemma 5.1.3, we have that in $\langle v, \beta \rangle$, $v \models \alpha$. So $\langle v, \beta \rangle \prec \langle w, \alpha \rangle$ cannot hold.
■

Lemma 5.1.10. *If $\top \mid\!\sim \alpha$, then $\langle w, \alpha \rangle \prec \langle v, \beta \rangle$ for every $\langle v, \beta \rangle \notin min_{\prec}(S)$*

Proof.
If $\langle v, \beta \rangle \notin min_{\prec}(S)$, then $\top \not\mid\!\sim \beta$. So, by the definition of $\prec$, $\langle w, \alpha \rangle \prec \langle v, \beta \rangle$.
■

This shows that $\mathfrak{M}$ is an optimal model, since the states defined as normal worlds for the beliefs of the agent are preferred to any other state.
From optimality it is easy to prove smoothness.

Lemma 5.1.11. $\mathfrak{M}$ *satisfies smoothness.*

Proof.
Optimality implies that $[\top]_S$ ($[\top]_S = S$) is a smooth set, i.e. $S \neq \emptyset$ implies $min_{\prec}(S) \neq \emptyset$.
Assume that $\mathfrak{M}$ does not satisfy smoothness, i.e. there is a formula $\alpha \in \ell$ s.t. we have an infinite descending chain of worlds satisfying α. Since such worlds would also satisfy $\top$, we would have an infinite descending chain of worlds satisfying $\top$, and that would contradict the smoothness condition of the set S.
■

So, $\mathfrak{M}$ is a preferential model satisfying optimality.
Now we have to show that, for every formula α, the set of minimal worlds of α corresponds to the set of its normal worlds.

We have to distinguish two cases: $\top \mid\sim \neg\alpha$ and $\top \not\mid\sim \neg\alpha$.

First of all, let us consider the case that $\top \not\mid\sim \neg\alpha$.

Lemma 5.1.12. *Let* $\top \not\mid\sim \neg\alpha$.
If w *is a normal world for* α*, then there is a copy of the valuation* w*,* $\langle w, \beta \rangle$*, for every* β *s.t.* $\top \mid\sim \beta$.

Proof.
It is sufficient to show that, if w is normal for α and $\top \not\mid\sim \neg\alpha$, w is a normal world also for any belief β.
Given $\top \mid\sim \beta$, we have, by CT, that $\top \mid\sim \gamma$ for every γ s.t. $\beta \mid\sim \gamma$.
From $\top \mid\sim \gamma$ and $\top \not\mid\sim \neg\alpha$ we obtain, by RW and WRM, $\alpha \mid\sim \gamma$.
So, since w is a normal world for α, if $\beta \mid\sim \gamma$, then $\alpha \mid\sim \gamma$ and so $w \vDash \gamma$, i.e. w is a normal world for β and $\langle w, \beta \rangle \in S$.

■

Lemma 5.1.13. *If $\langle w, \beta \rangle$ is an optimal world and $w \vDash \alpha$, then w is a normal world for α.*

Proof.
Assume $\alpha \mathrel{|\!\sim} \gamma$. By conditionalization, we have $\top \mathrel{|\!\sim} \alpha \rightarrow \gamma$. So, since $w \in min_{\prec}(S)$, $w \vDash \alpha \rightarrow \gamma$, that, together with $w \vDash \alpha$, gives $w \vDash \gamma$. So w is a normal world for α.
■

From the above lemmas and optimality we have:

Lemma 5.1.14. *Let $\top \mathrel{|\!\not\sim} \neg\alpha$.*
$min_{\prec}([\alpha]_S) \subseteq min_{\prec}(S)$, and there is a state $\langle w, \beta \rangle \in min_{\prec}([\alpha]_S)$ iff w is a normal world for α.

Now, let us move to the case that $\top \mathrel{|\!\sim} \neg\alpha$.

Lemma 5.1.15. *Let $\top \mathrel{|\!\sim} \neg\alpha$.*
In the model $\mathfrak{M}$, $\langle w, \beta \rangle \in min_{\prec}([\alpha]_s)$ iff $w \vDash \alpha$ and $\beta \leq \alpha$.

Proof.
$\Rightarrow$: Assume $\langle w, \beta \rangle \in min_{\prec}([\alpha]_s)$. Then $w \vDash \alpha$.
If $\top \mathrel{|\!\sim} \beta$, then $\beta \leq \alpha$ for every α.
If $\top \mathrel{|\!\not\sim} \beta$, we have to prove that $\alpha \vee \beta \mathrel{|\!\sim} \beta$. Suppose there is a normal world v for $\alpha \vee \beta$ s.t. $v \nvDash \beta$. Since $\alpha \vee \beta \leq \beta$ for every formula, we have $\langle v, \alpha \vee \beta \rangle \prec \langle w, \beta \rangle$. But since $v \vDash \alpha \vee \beta$ and $v \nvDash \beta$, we have $v \vDash \alpha$, contrary to the α-minimality of $\langle w, \beta \rangle$.
So every normal world for $\alpha \vee \beta$ satisfies β, and, by Lemma 5.1.5, we have that $\alpha \vee \beta \mathrel{|\!\sim} \beta$, i.e. $\beta \leq \alpha$.

$\Leftarrow$: Assume $w \vDash \alpha$ and $\beta \leq \alpha$ and that $\langle w, \beta \rangle \notin min_{\prec}([\alpha]_S)$, i.e. there is a world $\langle v, \gamma \rangle$ s.t. $\langle v, \gamma \rangle \prec \langle w, \beta \rangle$ and $v \vDash \alpha$. Since $\top \mathrel{|\!\sim} \neg\alpha$, v cannot be an

optimal world and, by Lemma 5.1.9, $\top \not\mid\!\sim \gamma$.
So $\langle v, \gamma \rangle \prec \langle w, \beta \rangle$ because $\gamma \leq \beta$ and $v \not\models \beta$.
Hence, $\gamma \leq \beta \leq \alpha$, v is a normal world for γ, $v \not\models \beta$ and $v \models \alpha$, which is in contradiction with Lemma 5.1.7.
■

From this Lemma and Lemma 5.1.6, we obtain:

Lemma 5.1.16. *Let* $\top \mid\!\sim \neg\alpha$.
If $\langle w, \beta \rangle \in min_{\prec}([\alpha]_s)$, *then* w *is a normal world for* α.

Proof.
If $\langle w, \beta \rangle \in min_{\prec}([\alpha]_s)$, then , by Lemma 5.1.15, $\langle w, \beta \rangle \models \alpha$ and $\beta \leq \alpha$. So, by Lemma 5.1.6, w is a normal world for α.
■

Lemma 5.1.17. *Let* $\top \mid\!\sim \neg\alpha$.
If w *is a normal world for* α, $\langle w, \alpha \rangle \in min_{\prec}([\alpha]_S)$.

Proof.
$\langle w, \alpha \rangle \models \alpha$ and $\alpha \leq \alpha$ (since $\alpha \vee \alpha \mid\!\sim \alpha$ is preferentially valid). So, by Lemma 5.1.15, $\langle w, \alpha \rangle \in min_{\prec}([\alpha]_s)$.
■

From Lemma 5.1.14, Lemma 5.1.16 and Lemma 5.1.17, we obtain:

Lemma 5.1.18. *In the model* $\mathfrak{M}$, *for every formula* α, *the set* $min_{\prec}([\alpha]_s)$ *is composed exactly by at least one copy of every normal world* w *for* α.

This result and Lemma 5.1.5 allow to derive that $\mathfrak{M}$ is a canonical model for the inference relation $\mid\!\sim$.

Lemma 5.1.19. *For every* α *and* β, $\alpha \mid\!\sim \beta$ *iff* $\alpha \mid\!\sim_{\mathfrak{M}} \beta$

Proof.
$\alpha \mid\!\sim_{\mathfrak{M}} \beta$ iff $w \vDash \beta$ for every $w \in min_{\prec}([\alpha]_s)$; such condition, by Lemma 5.1.18 and Lemma 5.1.5, corresponds to say that $\alpha \mid\!\sim \beta$.
■

Now we know that every optimal model generates a weakly rational inference relation (Lemma 5.1.1), and that for every weakly rational inference relation there is an optimal model representing it (Lemma 5.1.19).
This can be summed up in the following.

Theorem 5.1.20 (Representation Theorem for Weakly Rational Inference Relations)**.** *A consequence relation is a weakly rational inference relation iff it is defined by some optimal model.*

5.2 Relations with Injectivity

Weak rational monotony is a special case of rational monotony. In [13], the authors have proven that weak rational monotony and negation rationality are independent, and so it also does not imply disjunctive rationality.
Since RM implies injectivity, we want to investigate if also WRM implies injectivity, since in the next chapter we will mainly deal with optimal injective models.
To prove that WRM does not imply injectivity it is sufficient to show that there is a non-injective model defining a weakly rational consequence relation that cannot be generated by means of an injective model.

Proposition 5.2.1. *Weak rational monotony does not imply injectivity.*

Proof. Take a language ℓ generated from two elementary letters p, q. We define the optimal non-injective model $\mathfrak{M}$ described in Fig. 5.1. We have to prove that there is not an injective model defining the same consequence relation.

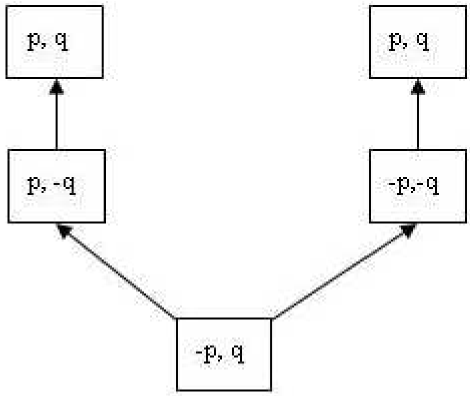

Figure 5.1: Non-injective optimal model

An injective model of the language ℓ cannot have more than four states, since four are the possible valuations for a two-letters language. Observe that in $\mathfrak{M}$ the following sequents hold: $\top \mid\!\sim \neg p \wedge q$; $p \wedge q \not\mid\!\sim \bot$; $p \leftrightarrow q \not\mid\!\sim p \wedge q$; $p \leftrightarrow q \not\mid\!\sim \neg p \wedge \neg q$; $p \not\mid\!\sim q$; $p \not\mid\!\sim \neg q$; $p \vee \neg q \mid\!\sim \neg q$.
We can try to build a corresponding injective model (see Fig. 5.2).
(1) $\top \mid\!\sim \neg p \wedge q$ implies that the valuation $(\neg p, q)$ has to be the preferred one.
(2) Since $p \wedge q \not\mid\!\sim \bot$, the valuation (p, q) has to be in the model, above $(\neg p, q)$.
(3) Since $p \leftrightarrow q \not\mid\!\sim p \wedge q$ and $p \leftrightarrow q \not\mid\!\sim \neg p \wedge \neg q$, in the model there must be also $(\neg p, \neg q)$. It has to be above $(\neg p, q)$, but it cannot be neither preferred to (p, q) (otherwise $p \leftrightarrow q \mid\!\sim \neg p \wedge \neg q$), nor above it (otherwise $p \leftrightarrow q \mid\!\sim p \wedge q$).
(4) Since $p \not\mid\!\sim q$ and $p \not\mid\!\sim \neg q$, with an argument similar to that in (3), we are forced to position $(p, \neg q)$ above $(\neg p, q)$ and disjoint from the other valuation. However, this model does not satisfy $p \vee \neg q \mid\!\sim \neg q$.

■

So, this counterexample shows that there can be non-injective weakly rational consequence relations.

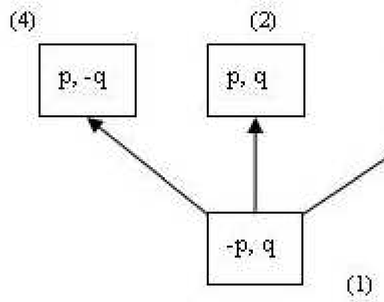

Figure 5.2: Injective model

Chapter 6

Using default-assumptions in closure operations

Abstract. We shall show how it is possible to use the default-assumption approach in order to build in a simple way semantical models for non-Horn closures of conditional bases.

In Chapter 3 we have seen that there is a correspondence between Shoham's preferential models and default-assumption models. In this chapter we are going to see how such a correspondence can be used in the construction of semantic models for some non-Horn inference relations.
The chapter is organized as follows.
In Section 6.1 we are going to analyze more deeply the behaviour of preferential models generated from sets of default-formulas. In particular, we shall see that such generation methods can be used only in dealing with injective preferential models, and we will check, assuming consistent default-assumption sets, which kind of preferential models we can obtain.
In Section 6.2 we shall define the generation of preferential models representing weakly rational closures and rational closures of a given set of sequents $\mathcal{B}$. We shall begin by stating the desiderata of such closure operations, then

we shall investigate more deeply the role of knowledge, beliefs and inconsistencies in such closures, and how they are related to the construction of a default-assumption set. Finally, we shall present our closure operations.

6.1 Generated preferential models

First of all, we have to analyze the behaviour of the preferential models generated by sets of default formulae.

6.1.1 Injectivity

We can note from the constructions in chapter 3 that the correspondence between the default-assumption approach and the preferential one is limited only to injective models, and cannot be extended to models with multiple copies. The reason is simple: in generating an order from a set of defaults Δ (see Definition 3.1.1), we refer only to the valuations associated with the states, and it becomes impossible to differentiate the relative positions in the ordering of different states associated with the same valuation. We say that two states have the same *relative position* in a preferential model iff they have the same sets of states above and below them.

Lemma 6.1.1. *Assume a set of states S, a non-injective labeling function l and a set of default formulae Δ. If two states $s, t \in S$ are associated with the same propositional valuation w, then they have the same relative position in the ordering δ_Δ.*

Proof.
Recall that $\delta_\Delta = \{(r, u) \in S \times S \mid \Delta_r \supset \Delta_u\}$. Since s and t are associated with the same valuation w, we have that $\Delta_s = \Delta_t$, both equal to Δ_w.
Obviously from the definition of δ_Δ, for every state r:

$$r \prec s \text{ iff } r \prec t$$

$$s \prec r \text{ iff } t \prec r$$

■

We say that two preferential models are *inferentially equivalent* if they define the same inference relation $\mid\!\sim$.
We can prove that a non-injective model generated by a set of default-assumptions, as the one described in the lemma above, is inferentially equivalent to its injective restriction (i.e. the model obtained by eliminating every state labeled with the same valuation as another state).

Lemma 6.1.2. *Assume a model $\mathfrak{M} = \langle S, l, \prec \rangle$ s.t. two states $s, t \in S$ are associated with the same valuation w and have the same relative position in $\prec$. Let $\mathfrak{M}'$ be a model s.t. it corresponds to $\mathfrak{M}$ restricted to the set of states $S' = S - t$. Then $\mid\!\sim_{\mathfrak{M}} = \mid\!\sim_{\mathfrak{M}'}$*

Proof.
We have to show that for every pair of formulas α and β, $\alpha \mid\!\sim_{\mathfrak{M}} \beta$ iff $\alpha \mid\!\sim_{\mathfrak{M}'} \beta$.
Given the same relative position and the same associated valuation, for every formula α, $s \in min_{\prec}([\alpha]_S)$ iff $t \in min_{\prec}([\alpha]_S)$.
Take a state r s.t. $r \neq t$. We have to show that, for every formula α, $r \in min_{\prec}([\alpha]_S)$ in $\mathfrak{M}$ iff $r \in min_{\prec}([\alpha]_{S'})$ in $\mathfrak{M}'$. If $r \in min_{\prec}([\alpha]_S)$, then for every state q s.t. $q \prec r$, $q \nvDash \alpha$. Since S' is a restriction of S, then, also in S', for every state q s.t. $q \prec r$, $q \nvDash \alpha$, and $r \in min_{\prec}([\alpha]_{S'})$. Assume $r \in min_{\prec}([\alpha]_{S'})$; then, moving to S, the only possibility for $r \notin min_{\prec}([\alpha]_S)$ is that $t \prec r$ and $t \vDash \alpha$. But, since s is in the same relative position of t and is associated to the same valuation, in S' we would have $s \prec r$, $s \vDash \alpha$, and $r \notin min_{\prec}([\alpha]_{S'})$.
So, for every formula α s.t. $s, t \notin min_{\prec}([\alpha]_S)$, we have $min_{\prec}([\alpha]_S) = min_{\prec}([\alpha]_{S'})$.
Take a formula α s.t. $s, t \in min_{\prec}([\alpha]_S)$. Since s and t contribute with the same valuation to the determination of the defeasible consequences of α, if we eliminate t from the set $min_{\prec}([\alpha]_S)$, $min_{\prec}([\alpha]_{S'})$ contains the same set

of valuations, since the contribution of t is preserved by s. Conversely, if $s \in min_{\prec}([\alpha]_{S'})$, we shall have that $s, t \in min_{\prec}([\alpha]_S)$, but, since the contribution of t is guaranteed in S' by s, its addition is redundant and $min_{\prec}([\alpha]_S)$ contains the same valuations as $min_{\prec}([\alpha]_{S'})$.
Hence, for every formula α, $min_{\prec}([\alpha]_S)$ and $min_{\prec}([\alpha]_{S'})$ contain the same valuations, i.e. α has the same defeasible consequences. ■

From Lemmas 6.1.1 and 6.1.2, we can easily derive the following:

Proposition 6.1.3. *Given a Δ-generated non-injective model $\mathfrak{M}$, it defines the same inference relation $\mid\!\sim$ generated by the corresponding Δ-generated injective model $\mathfrak{M}'$.*

Analogously, such problems with the use of default assumptions in analyzing non-injective preferential models also arise if we try to extract the characteristic set of defaults (see Section 3.1) from a non-injective preferential model. Here is an example.

Assume the non-injective model depicted in Fig. 6.1.
Using the method delineated in Proposition 3.1.3, we obtain the character-

Figure 6.1: Non-injective model

istic set of defaults $\Delta = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$, where φ_1 is $(p \wedge q) \vee (p \wedge \neg q)$, φ_2 is $(p \wedge q) \vee (\neg p \wedge \neg q)$, φ_3 is $(p \wedge \neg q)$, and φ_4 is $(\neg p \wedge \neg q)$.

If we try to re-generate the ordering of such set of worlds by means of Δ, we obtain a totally disjoint model, depicted in Fig. 6.2, which obviously does not generate the same inference relation as the previous one.

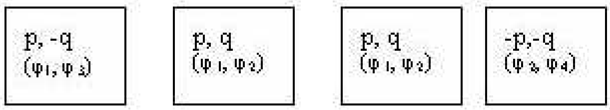

Figure 6.2: Model generated by Δ (every state has in parenthesis the defaults it satisfies)

Hence, the method to extract the characteristic set of an ordering, presented in Section 3.1, does not work for non-injective models. The correspondence between the preferential and the default-assumption approaches is restricted to injective inference relations.

6.1.2 Consistent default-assumption sets

We would also like to analyze the behaviour of Δ-generated preferential models, assuming we are working with a consistent default-set Δ. In fact, as we will see in more detail in the next chapter, the consistency of Δ is not necessary for the construction of well-behaved injective preferential models, but the consistency of our set of default-assumptions is obviously an intuitive desideratum, and corresponds to an interesting class of preferential models. At the end of Section 3.1, we have generalized our default-assumption consequence relations in order to also admit 'hard' background information, i.e. an agent with a knowledge set. So, we need a notion of consistency relative to such knowledge set.

Definition 6.1.1. (K-consistency)

*Assume a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$. We say that Δ is K-*consistent *iff* $\Delta \cup K \nvDash \bot$

Remember that, given a system $\mathfrak{S} = \langle K, \Delta \rangle$, we can construct the corresponding preferential model $\mathfrak{M} = \langle U_K, \delta_\Delta \rangle$, where U_K is the set of all the valuations satisfying all the formulae in K, and δ_Δ the preferential order generated by Δ over U_K (see Theorem 3.1.11).

Lemma 6.1.4. *Assume a system $\mathfrak{S} = \langle K, \Delta \rangle$, and the corresponding preferential model $\mathfrak{M} = \langle U_K, \delta_\Delta \rangle$. If Δ is K-consistent , then $\mathfrak{M}$ is an optimal model (see Definition 5.0.6).*

Proof.
If Δ is K-consistent, then there is at least a world w in U_K s.t. $w \vDash \gamma$ for every $\gamma \in \Delta$ ($\Delta_w = \Delta$). $[\Delta]_U$ is the set of such worlds in U_K satisfying every formula in Δ. From the definition of the generated order δ_Δ (see Definition 3.1.1), it is immediate to see that the worlds in $[\Delta]_U$ are preferred to any other world in U_K. Hence $\mathfrak{M}$ is an optimal model.
■

We can also prove that every optimal injective model can be generated by a consistent set of default-assumptions. Recall from Section 3.1 the formula $K_U = \bigvee\{\alpha_w | w \in U\}$, characterizing the knowledge contained in a set of valuations U (α_w is the formula characterizing univocally a valuation w).

Lemma 6.1.5. *Given an optimal model $\mathfrak{M} = \langle U, \delta \rangle$ there is a K_U-consistent default-assumption set generating it.*

Proof.
Assume an optimal model $\mathfrak{M} = \langle U, \delta \rangle$. Define the set $min_\delta(U)$ of its optimal valuations. Recall from Section 3.1 the notion of the characteristic set Δ^δ of the order δ:

$$\Delta^\delta = \{\beta_w | w \in U\}$$

where

$$\beta_w := \alpha_w \vee \bigvee_{v \prec_\delta w} \{\alpha_v\}$$

We know, generalizing Proposition 3.1.3, that $\mathfrak{S} = \langle K_U, \Delta^\delta \rangle$ generates $\mathfrak{M} = \langle U, \delta \rangle$.

Transform Δ^δ in Δ', substituting β'_w to every β_w in the following way:

$$\beta'_w = \begin{cases} \beta_w & \text{If } w \notin min(U) \\ \bigvee\{\alpha_v | v \in min(U)\} & \text{If } w \in min(U) \end{cases}$$

Since $w \vDash \bigvee\{\alpha_v | v \in min(U)\}$ iff $w \in min(U)$, it is easy to see by Definition 3.1.1 that Δ' generates the same order as Δ^δ, so $\mathfrak{S}' = \langle K_U, \Delta' \rangle$ generates $\mathfrak{M} = \langle U, \delta \rangle$.

We have to show that Δ' is K_U-consistent, i.e. $\bigwedge \Delta' \nvdash \bot$ (where $\vdash$ is the supraclassical monotonic consequence relation generated from the classical $\vDash$ adding K_U as extra-axioms).

Since, by optimality, $w \in min_\delta(U)$ iff $w \prec v$ for every $v \notin min_\delta(U)$, we have that $\bigvee\{\alpha_v | v \in min_\delta(U)\}$ is part of the disjunction forming every β'_w, i.e. every β'_w has the form $\gamma \vee \bigvee\{\alpha_v | v \in min_\delta(U)\}$ for some γ, with $\vdash \gamma \leftrightarrow \top$ only in case $w \in min(U)$.

Let π be short for $\bigvee\{\alpha_v | v \in min_\delta(U)\}$. Then we have that $\bigwedge \Delta'$ is equivalent to $\pi \wedge (\gamma_1 \vee \pi) \wedge \ldots \wedge (\gamma_n \vee \pi)$, with each $\gamma_i \vee \pi$ characterizing a β_w for $w \notin min(U)$.

Since we have that, for every formulas ϕ and ψ, $(\phi \vee \psi) \wedge \psi$ is logically equivalent to ψ, we have that $\bigwedge \Delta'$ is logically equivalent to π, i.e. to $\bigvee\{\alpha_v | v \in min_\delta(U)\}$.

Since every α_v with $v \in U$ is necessarily K_U-consistent, we have that $\bigvee\{\alpha_v | v \in min_\delta(U)\}$ is K_U-consistent.

■

From Lemmas 6.1.4 and 6.1.5, we obtain:

Theorem 6.1.6. *An injective model $\mathfrak{M} = \langle U, \delta \rangle$ is optimal iff it can be generated by a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, s.t. Δ is K-consistent.*

Analogously to the preferential definition (see Definition 4.1.8), we define the belief set $\mathbf{B}_{\mathfrak{S}}$ of a default system $\mathfrak{S} = \langle K, \Delta \rangle$ as the set of formulas holding in most normal situations: $\mathbf{B}_{\mathfrak{S}} = \{\alpha | \top \mid\!\sim_{\mathfrak{S}} \alpha\}$.
Given a system $\mathfrak{S} = \langle K, \Delta \rangle$ with a K-consistent Δ, it is easy to define the belief set.

Proposition 6.1.7. *Assume a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$ with a K-consistent Δ. Then $\boldsymbol{B}_{\mathfrak{S}} = Cl(K \cup \Delta)$*

Proof.
If Δ is K-consistent, then $\top \mid\!\sim_{\mathfrak{S}} \alpha$ iff $K \cup \Delta \vDash \alpha$. Hence $\mathbf{B}_{\mathfrak{S}} = Cl(K \cup \Delta)$.
■

6.1.3 Injective weak rationality

We know from the previous chapter that every optimal preferential model generates a weakly rational inference relation $\mid\!\sim$, and from chapter 4 that every injective preferential model generates an injective inference relation $\mid\!\sim$. Hence, from Theorem 6.1.6, we can derive the following proposition.

Proposition 6.1.8. *Every default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, with a K-consistent Δ, defines an injective weakly rational inference relation $\mid\!\sim$.*

We want to show that (assuming that ℓ is a finite propositional language, as usual) every injective weakly rational inference relation $\mid\!\sim$ has an optimal injective canonical model. Up to this point, in fact, we can derive that every injective weakly rational inference relation $\mid\!\sim$, being injective, has an injective canonical model, and, being weakly rational, has an optimal canonical model; this does not imply that $\mid\!\sim$ has a canonical model that is both injective and optimal.
To show this, it is sufficient to prove that every non-optimal injective model generates an inference relation that is not weakly rational.

Lemma 6.1.9. *If $\mathfrak{M} = \langle U, \delta \rangle$ is an injective non-optimal model, then the inference relation $\mid\!\sim_{\mathfrak{M}}$ defined by $\mathfrak{M}$ is not weakly rational.*

Proof.

If $\mathfrak{M}$ is injective and non-optimal, then there are two worlds w and v s.t. $w \in min_{\prec}(U)$, $v \notin min_{\prec}(U)$, and $w \not\prec v$.
We want to show that there is a formula α s.t. $\top \mid\!\sim \alpha \rightarrow \beta$, $\top \not\mid\!\sim \neg\alpha$ and $\alpha \not\mid\!\sim \beta$.
Define $\alpha_{v,w}$ as the characteristic formula of the set $\{w, v\}$, i.e. $\alpha_{v,w} = \alpha_w \vee \alpha_v$. Take a formula β as the characteristic formula of $min_{\prec}(U)$, i.e. $\beta = \bigvee\{\alpha_w | w \in min_{\prec}(U)\}$. The definition of β implies that, given a world u, $u \vDash \beta$ iff $u \in min_{\prec}(U)$.
We have $\top \not\mid\!\sim \neg\alpha_{v,w}$, since $w \in min_{\prec}(U)$ and $w \vDash \alpha_{v,w}$
From the definition of β, we obtain $\top \mid\!\sim \beta$, and hence $\top \mid\!\sim \alpha_{v,w} \rightarrow \beta$ by RW.
But we also have $\alpha_{v,w} \not\mid\!\sim \beta$, since $v \in min_{\prec}([\alpha_{v,w}]_U)$ and, given $v \notin min_{\prec}(\top)$, $v \nvDash \beta$.
So WRM fails.
■

Hence, if $\mid\!\sim$ is an injective weakly rational inference relation, its injective canonical model has to be also optimal.
From Theorems 4.2.2, 5.1.20, and Proposition 6.1.9, we can state:

Theorem 6.1.10. (Representation Theorem for Injective Weakly Rational Inference Relations)
A consequence relation is an injective weakly rational inference relation iff it is defined by some injective optimal model.

Finally, from Theorems 6.1.10 and 6.1.6, we obtain the following result.

Theorem 6.1.11. *A consequence relation is an injective weakly rational inference relation iff there is a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, with a K-consistent Δ generating it.*

6.2 Default-assumption closures

We have seen that every injective inference relation can be generated by a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$ (Theorem 3.1.11), and that every injective weakly rational inference relation can be generated by a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, with a K-consistent Δ (Theorem 6.1.11). We want to see how such connections can be used to easily create models for interesting closures of finite sets of sequents. In particular, we will see how to generate interesting models for weakly rational and rational closures of a set of sequents.

6.2.1 Desiderata

Assume we have a finite set of sequents $\mathcal{B} = \{\alpha \mid\!\sim \beta, \ldots, \gamma \mid\!\sim \phi\}$; we will call such a set a *conditional base*. KLM-rules allow to determine its preferential closure (see Section 4.1.1). However, as we have seen, if we want to go beyond preferential closure by means of the satisfaction of non-Horn rules, we have to renounce to the preferential axiomatic characterization and to the classical notion of entailment, and to choose a semantical model representing an interesting closure of $\mathcal{B}$ of the desired kind (see Section 4.2).
We will call $\mathbb{C}(\mathcal{B})$ an arbitrary closure operation of $\mathcal{B}$.

There is a set of minimal requirements that we want to be satisfied by $\mathbb{C}(\mathcal{B})$ in order to consider it a reasonable closure of $\mathcal{B}$:

1. $\mathcal{B} \subseteq \mathbb{C}(\mathcal{B})$.

2. $\mathbb{P}(\mathcal{B}) \subseteq \mathbb{C}(\mathcal{B})$ (where $\mathbb{P}(\mathcal{B})$ is the preferential closure of $\mathcal{B}$, see Section 4.1.1).

3. The knowledge and belief sets of $\mathbb{C}(\mathcal{B})$ have to be the same as those of $\mathbb{P}(\mathcal{B})$.

The first desideratum is an obvious requirement, i.e. our closure has to be reflexive. However, it is not always satisfied by default-assumption preferential models: given a set of sequents $\mathcal{B}$, it could not be easy to find a set of formulae Δ generating a preferential model satisfying $\mathcal{B}$.
The most intuitive potential candidate to represent $\mathcal{B}$ by means of a set of default-formulae is surely the set $\overrightarrow{\mathcal{B}}$ of the *materializations* of the sequents in $\mathcal{B}$, i.e. the material implications corresponding to the sequents:

$$\overrightarrow{\mathcal{B}} = \{\alpha \rightarrow \beta | \alpha \mid\sim \beta \in \mathcal{B}\}$$

However, it is easy to see that stating $\Delta = \overrightarrow{\mathcal{B}}$ we do not guarantee the generation of a preferential model satisfying $\mathcal{B}$.
Let us refer to the 'penguin' example. We have $\mathcal{B} = \{\beta \mid\sim \gamma, \alpha \mid\sim \neg\gamma, \alpha \mid\sim \beta\}$. Using $\overrightarrow{\mathcal{B}}$ as the default-assumption set, we have $\Delta = \{\beta \rightarrow \gamma, \alpha \rightarrow \neg\gamma, \alpha \rightarrow \beta\}$. Such a default set is consistent, so there are some worlds satisfying it, and such worlds are the preferred ones. Since Δ is consistent with β, our model validates $\beta \mid\sim \gamma$. However, $\Delta \vDash \neg\alpha$, so we have to identify the α-maximal consistent subsets of Δ to determine the plausible consequences of α. There are three such α-maximal consistent sets: $\Delta' = \{\alpha \rightarrow \neg\gamma, \alpha \rightarrow \beta\}$, $\Delta'' = \{\beta \rightarrow \gamma, \alpha \rightarrow \beta\}$, $\Delta''' = \{\beta \rightarrow \gamma, \alpha \rightarrow \neg\gamma\}$.
As can be easily seen, $\alpha, \Delta' \vDash \beta$, $\alpha, \Delta'' \vDash \beta$, but $\alpha, \Delta''' \nvDash \beta$, so $\alpha \not\mid\sim \beta$. Analogously for $\alpha \mid\sim \neg\gamma$: $\alpha, \Delta' \vDash \neg\gamma$, $\alpha, \Delta''' \vDash \neg\gamma$, but $\alpha, \Delta'' \nvDash \neg\gamma$, so $\alpha \not\mid\sim \neg\gamma$.

The second point is a desideratum since, as we have seen in Chapter 4, the preferential closure is considered as the minimal requirement for the characterization of a rational agent. However, the satisfaction of the first desideratum implies automatically the satisfaction of the second: if $\mathfrak{M}$ is a preferential model of $\mathcal{B}$, automatically it is also a model of $\mathbb{P}(\mathcal{B})$.

For the third point, we prefer not to modify the belief and knowledge sets generated by the preferential closure: in defining a closure operation amplifying the preferential one, we want to amplify the agent's reasoning capabilities with respect to sequents; the epistemic situation of the agent is primarily characterized by what it maintains as hard background information, its knowledge, and what it holds as normally holding, its beliefs. Such kinds of information are defined in a solid logical way by the preferential closure (see Chapter 4), and we do not feel any need to change them. Hence, every modification of knowledge and belief sets would sound as an unjustified move. Moreover, as we are going to see below, the knowledge and belief sets preferentially defined from a set of sequents can be used to define a notion of preferential consistency.

6.2.2 Knowledge and beliefs

In Chapter 4 we have defined the knowledge set of a preferential $\mid\sim$ as the set $\mathbf{K}_{\mid\sim} = \{\alpha | \neg\alpha \mid\sim \bot\}$, and every set A_K s.t. $Cl(A_K) = \mathbf{K}_{\mid\sim}$ as a knowledge base of $\mathbf{K}_{\mid\sim}$. So, we want the knowledge set associated to a set of sequents $\mathcal{B}$ ($\mathbf{K}_{\mathcal{B}}$) to be the knowledge set defined by its preferential closure:

$$\mathbf{K}_{\mathcal{B}} = \{\alpha | \neg\alpha \mid\sim \bot \in \mathbb{P}(\mathcal{B})\}$$

Moreover, we would like, given a set $\mathcal{B}$, to identify a finite knowledge base for $\mathbf{K}_{\mathcal{B}}$. Bochman, in [5], has defined a simple method to identify the knowledge base associated to a set of sequents.

Assume a set of sequents $\mathcal{D}$, and define $\mathcal{D}^{\vee}$ the disjunction of all the antecedents of the sequents in $\mathcal{D}$, i.e.

$$\mathcal{D}^{\vee} = \bigvee\{\alpha | \alpha \mid\sim \beta \in \mathcal{D}\}$$

Bochman defines $\mathcal{D}$ a *clash* iff $\overrightarrow{\mathcal{D}} \models \neg\mathcal{D}^{\vee}$, i.e. a set of sequents which, in its materialized form, classically implies the negation of *all* its antecedents.

He proves the following lemma.

Lemma 6.2.1. *([5], Lemma 7.5.4)*
$\alpha \in \boldsymbol{K}_{\mathcal{B}}$ iff $\neg\mathcal{D}^{\vee} \vDash \alpha$, for some clash $\mathcal{D} \subseteq \mathcal{B}$.

Since the union of every clash is itself a clash, for every set $\mathcal{B}$ we can identify a unique greatest clash $\mathcal{C}_B$ of $\mathcal{B}$, given by the union of all the clashes in $\mathcal{B}$.
So, we can define a knowledge base $Q_{\mathcal{B}}$, which closure is the knowledge set of $\mathbb{P}(\mathcal{B})$.

Definition 6.2.1. *(Knowledge base of $\mathcal{B}$)*
Given a set of sequents $\mathcal{B}$, and identified its greatest clash $\mathcal{C}_B$, we call the knowledge base *of $\mathcal{B}$ the set $Q_{\mathcal{B}}$ s.t.*

$$Q_{\mathcal{B}} = \{\neg\alpha | \alpha \mid\sim \beta \in \mathcal{C}_B\}$$

Lemma 6.2.2. *([5], Lemma 7.5.5)*
$Q_{\mathcal{B}}$ is a knowledge base for $\boldsymbol{K}_{\mathcal{B}}$,i.e.

$$\boldsymbol{K}_{\mathcal{B}} = Cl(Q_{\mathcal{B}})$$

So, given a set of sequents $\mathcal{B}$, we can identify its greatest clash, and consequently, we can define the knowledge base of $\mathcal{B}$. This, in turn, defines the *monotonic core* (see Definition 4.1.7) of $\mathcal{B}$.

Definition 6.2.2. *(Monotonic core of $\mathcal{B}$)*
The monotonic core *of set of sequents $\mathcal{B}$ corresponds to the monotonic core of its preferential closure $\mathbb{P}(\mathcal{B})$, i.e. the monotonic consequence operation $Cn_{\mathcal{B}}$ (relation $\vdash_{\mathcal{B}}$) obtained from Cl and the addition of $Q_{\mathcal{B}}$ as a set of extra-axioms:*

$$Cn_{\mathcal{B}}(A) = Cl(Q_{\mathcal{B}} \cup A) \text{ for every } A \subseteq \ell$$

$$A \vdash_{\mathcal{B}} \alpha \text{ iff } A \cup Q_{\mathcal{B}} \vDash \alpha \text{ for every } A \subseteq \ell$$

As expressed in the desideratum, we look for a closure of $\mathcal{B}$ that preserves $\mathbf{K}_{\mathcal{B}}$ as its knowledge set.
So, as Bochman suggests ([5], p.180), we can partition a conditional base $\mathcal{B}$ in the part defining its knowledge core, its greatest clash $\mathcal{C}_B$, and the part defining defeasible information, $\mathcal{B}\backslash\mathcal{C}_B$.

Definition 6.2.3. *(knowledge and defeasible portion of $\mathcal{B}$)*
Given a conditional base $\mathcal{B}$, we define:

- $\mathcal{B}_{\mathbf{K}} = \mathcal{C}_B$*: the* knowledge portion *of $\mathcal{B}$ (where $\mathcal{C}_B$ is the greatest clash in $\mathcal{B}$).*
- $\mathcal{B}_{\mathbf{D}} = \mathcal{B} \backslash \mathcal{C}_B$*: the* defeasible portion *of $\mathcal{B}$.*

Analogously, we want our closure operation $\mathbb{C}(\mathcal{B})$ to preserve the same belief set of $\mathbb{P}(\mathcal{B})$, which is (see Definition 4.1.8):

$$\mathbf{B}_{\mathcal{B}} = \{\alpha | \top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})\}$$

We shall show that, given a set $\mathcal{B}$, with its monotonic core $Cn_{\mathcal{B}}$, the belief set of $\mathcal{B}$ is equivalent to the closure under $\mathcal{B}$'s monotonic core of the materialization of the conditional part of $\mathcal{B}$, i.e.:

$$\mathbf{B}_{\mathcal{B}} = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_{\mathbf{D}}})$$

As a first step, we want to show that $\mathbf{B}_{\mathcal{B}} = Cl(\overrightarrow{\mathcal{B}})$.

Lemma 6.2.3. *If $\alpha \in Cl(\overrightarrow{\mathcal{B}})$, then $\alpha \in \boldsymbol{B}_{\mathcal{B}}$*

Proof.
Recall that $\alpha \in \mathbf{B}_{\mathcal{B}}$ iff $\top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$.
So, we have to show that if $\alpha \in Cl(\overrightarrow{\mathcal{B}})$, then $\top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$.
From conditionalization, we have that $\phi \mid\sim \psi \in \mathcal{B}$ implies $\top \mid\sim \phi \rightarrow \psi \in \mathbb{P}(\mathcal{B})$. So we have that if $\alpha \in \overrightarrow{\mathcal{B}}$, then $\top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$. Assume $\beta \in Cl(\overrightarrow{\mathcal{B}})$. Then, by compactness, there is a finite set $\{\gamma_1, \dots, \gamma_n\}$ s.t. $\{\gamma_1, \dots, \gamma_n\} \subseteq \overrightarrow{\mathcal{B}}$ and $\{\gamma_1, \dots, \gamma_n\} \models \beta$.
Correspondingly, we have $\{\top \mid\sim \gamma_1, \dots, \top \mid\sim \gamma_n\} \subseteq \mathbb{P}(\mathcal{B})$; since $\{\gamma_1, \dots, \gamma_n\} \models \beta$, we have that $\{\top \mid\sim \gamma_1, \dots, \top \mid\sim \gamma_n\}$ implies $\top \mid\sim \beta$ by AND and RW, and $\top \mid\sim \beta \in \mathbb{P}(\mathcal{B})$, i.e. $\beta \in \mathbf{B}_{\mathcal{B}}$.
■

Lemma 6.2.4. *If* $\alpha \mid\sim \beta \in \mathbb{P}(\mathcal{B})$*, then* $\alpha \rightarrow \beta \in Cl(\overrightarrow{\mathcal{B}})$

Proof.

We prove it by cases w.r.t. the derivation steps of $\alpha \mid\sim \beta$ from $\mathcal{B}$ by means of the preferential rules.

If $\alpha \mid\sim \beta \in \mathcal{B}$, then, by definition, $\alpha \rightarrow \beta \in \overrightarrow{\mathcal{B}}$.

By reflexivity, $\alpha \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$ for every α, and, by tautology, $\alpha \rightarrow \alpha \in Cl(\overrightarrow{\mathcal{B}})$ for every α.

-LLE:

If $\alpha \mid\sim \beta \in \mathbb{P}(\mathcal{B})$ and $\vDash \alpha \leftrightarrow \gamma$, then $\gamma \mid\sim \beta \in \mathbb{P}(\mathcal{B})$.

If $\alpha \rightarrow \beta \in Cl(\overrightarrow{\mathcal{B}})$ and $\vDash \alpha \leftrightarrow \gamma$, then $\gamma \rightarrow \beta \in Cl(\overrightarrow{\mathcal{B}})$.

-RW:

If $\alpha \mid\sim \beta \in \mathbb{P}(\mathcal{B})$ and $\vDash \beta \rightarrow \gamma$, then $\alpha \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$.

If $\alpha \rightarrow \beta \in Cl(\overrightarrow{\mathcal{B}})$ and $\vDash \beta \rightarrow \gamma$, then $\alpha \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$.

-CM:

If $\alpha \mid\sim \beta, \alpha \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$, then $\alpha \wedge \beta \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$.

If $\alpha \rightarrow \beta, \alpha \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$, then $(\alpha \wedge \beta) \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$.

-CT:

If $\alpha \wedge \gamma \mid\sim \beta, \alpha \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$, then $\alpha \mid\sim \beta \in \mathbb{P}(\mathcal{B})$.

If $(\alpha \wedge \gamma) \rightarrow \beta, \alpha \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$, then $\alpha \rightarrow \beta \in Cl(\overrightarrow{\mathcal{B}})$.

-OR:

If $\alpha \mid\sim \gamma, \beta \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$, then $\alpha \vee \beta \mid\sim \gamma \in \mathbb{P}(\mathcal{B})$.

If $\alpha \rightarrow \gamma, \beta \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$, then $(\alpha \vee \beta) \rightarrow \gamma \in Cl(\overrightarrow{\mathcal{B}})$.

■

Corollary 6.2.5. *If* $\top \mid\sim \alpha \in \mathbb{P}(\mathcal{B})$*, then* $\alpha \in Cl(\overrightarrow{\mathcal{B}})$

Proof.

It is just a special case of the lemma above.

■

Hence, by Lemma 6.2.3 and Corollary 6.2.5, we have:

Proposition 6.2.6. $\alpha \in \boldsymbol{B}_{\mathcal{B}}$ *iff* $\alpha \in Cl(\overrightarrow{\mathcal{B}})$

Now we shall prove that $Cl(\overrightarrow{\mathcal{B}}) = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_{\mathbf{D}}})$.

Proposition 6.2.7. $Cl(\overrightarrow{\mathcal{B}}) = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_{\mathbf{D}}})$

Proof.
It is sufficient to prove that $\overrightarrow{\mathcal{B}_{\mathbf{K}}}$ is logically equivalent to $Q_{\mathcal{B}}$ (recall that $Q_{\mathcal{B}} = \{\neg\alpha | \alpha \mid\!\sim \beta \in \mathcal{B}_{\mathbf{K}}\}$).
Given that $\mathcal{B}_{\mathbf{K}}$ is the greatest clash in $\mathcal{B}$, we have that $\overrightarrow{\mathcal{B}_{\mathbf{K}}} \vDash \neg\alpha$ for every $\neg\alpha \in Q_{\mathcal{B}}$.
By RW we have that $Q_{\mathcal{B}} \vDash \alpha \rightarrow \beta$ for every $\neg\alpha \in Q_{\mathcal{B}}$, i.e. for every $\alpha \rightarrow \beta \in \overrightarrow{\mathcal{B}_{\mathbf{K}}}$.
Obviously, $Cl(\overrightarrow{\mathcal{B}}) = Cl(\overrightarrow{\mathcal{B}_{\mathbf{D}}} \cup \overrightarrow{\mathcal{B}_{\mathbf{K}}})$. Given the logical equivalence between $Q_{\mathcal{B}}$ and $\overrightarrow{\mathcal{B}_{\mathbf{K}}}$, we have that $Cl(\overrightarrow{\mathcal{B}_{\mathbf{D}}} \cup \overrightarrow{\mathcal{B}_{\mathbf{K}}}) = Cl(\overrightarrow{\mathcal{B}_{\mathbf{D}}} \cup Q_{\mathcal{B}}) = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_{\mathbf{D}}})$.
■

Hence, by Propositions 6.2.6 and 6.2.7, we have proven the following:

Theorem 6.2.8. $\alpha \in \boldsymbol{B}_{\mathcal{B}}$ *iff* $\alpha \in Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_D})$

From Proposition 6.2.7 we also have the following result.

Corollary 6.2.9. $\overrightarrow{\mathcal{B}}$ *is classically consistent iff* $\overrightarrow{\mathcal{B}_D}$ *is* $Cn_{\mathcal{B}}$*-consistent.*

Preferential inconsistency

From Proposition 6.2.6, we can also derive an interesting corollary about the presence of inconsistencies in a conditional base $\mathcal{B}$. As we have seen, there is nothing wrong if a sequent $\alpha \mid\!\sim \bot$ is preferentially derivable from $\mathcal{B}$, since it simply means that, assuming $\mathcal{B}$, the agent takes for sure the truth of $\neg\alpha$. That is, $\alpha \mid\!\sim \bot$ means that α is a $\mid\!\sim$-inconsistent formula, but now we are on the meta-level: our object language does not contain formulae, but sequents, so we would like to see when a set of sequents $\mathcal{B}$ is to be considered as inconsistent.

Obviously, the problem of inconsistency would arise if we had in our knowledge both α and $\neg\alpha$, i.e., by OR, $\mathcal{B} \Vdash_P \alpha \vee \neg\alpha \mid\sim \bot$, or, equivalently, $\mathcal{B} \Vdash_P \top \mid\sim \bot$.
Equivalently, if we have both α and $\neg\alpha$ between our beliefs, which, by AND, gives again $\mathcal{B} \Vdash_P \top \mid\sim \alpha \wedge \neg\alpha$, i.e. $\mathcal{B} \Vdash_P \top \mid\sim \bot$.
If our conditional base preferentially entails $\top \mid\sim \bot$, then our agent will believe everything, since $\bot$ will be in its knowledge set and in its belief set.
Moreover, from $\top \mid\sim \bot$, we can derive every possible sequent by means of CM and RW.
Take two formulae ϕ and ψ: From $\top \mid\sim \bot$, by RW, we have $\top \mid\sim \phi$, and, by CM, $\phi \mid\sim \bot$. Apply again RW and we obtain $\phi \mid\sim \psi$.

Definition 6.2.4. *(preferential inconsistency) A conditional base $\mathcal{B}$ is* preferentially inconsistent *iff* $\mathcal{B} \Vdash_P \top \mid\sim \bot$

We can prove that the inconsistency of a conditional base is directly connected to the inconsistency of its materialization.

Theorem 6.2.10. $\mathcal{B} \Vdash_P \top \mid\sim \bot$ *iff* $\bot \in Cl(\overrightarrow{\mathcal{B}})$

Proof.
$\mathcal{B} \Vdash_P \top \mid\sim \bot$ is equivalent to $\bot \in \mathbf{B}_{\mathcal{B}}$. By Proposition 6.2.6 such an occurrence is possible iff $\bot \in Cl(\overrightarrow{\mathcal{B}})$ ■

Hence, we can check the inconsistency of a set of sequents $\mathcal{B}$ simply by checking the inconsistency of its materialization.
Obviously, if $\mathcal{B}$ is preferentially inconsistent, $\mathcal{B}$ will be the clash of itself.

Proposition 6.2.11. *If $\mathcal{B}$ is preferentially inconsistent, then its knowledge base $Q_{\mathcal{B}}$ will be inconsistent, and its defeasible portion $\mathcal{B}_D$ will be empty.*

Proof.
If $\mathcal{B} \Vdash_P \top \mid\sim \bot$, then, by Corollary 6.2.10, $\overrightarrow{\mathcal{B}} \vDash \bot$. Hence $\overrightarrow{\mathcal{B}} \vDash \neg\alpha$ for every $\alpha \mid\sim \beta \in \mathcal{B}$, and consequently $\mathcal{B}$ is a clash of itself, and $\mathcal{B}_D$, its defeasible portion, will be empty.

Given that $\mathcal{C}_{\mathcal{B}} = \mathcal{B}$, we have that $Q_{\mathcal{B}} = \{\neg\alpha | \alpha \mid\!\sim \beta \in \mathcal{B}\}$. Since $\neg\alpha \vDash \alpha \rightarrow \beta$ for every α, we have that $Q_{\mathcal{B}} \vDash \bigwedge \overrightarrow{\mathcal{B}}$, i.e. $Q_{\mathcal{B}} \vDash \bot$.
■

Hence, the preferential closure of an inconsistent set of sequents simply determines an inconsistent knowledge set.
These results stress the importance of keeping fixed the knowledge and belief sets determined by the preferential closure of a conditional base in defining a new closure operation (the third desideratum in Section 6.2.1), since, assuming that the conditional base is preferentially consistent, we guarantee that such consistency is preserved, impeding the validity of $\top \mid\!\sim \bot$ by means of our closure operation.

6.2.3 Weakly rational closure

We want, given a conditional base $\mathcal{B}$, to use the default-assumption approach to build in a simple way a satisfying weakly rational closure $\mathbb{W}(\mathcal{B})$.
As we have said, the general desiderata are

- $\mathbb{P}(\mathcal{B}) \subseteq \mathbb{W}(\mathcal{B})$
- The knowledge and belief sets of $\mathbb{W}(\mathcal{B})$ have to be the same as $\mathbb{P}(\mathcal{B})$

From the second desideratum, we can derive the following rule:

- If we have $\top \mid\!\sim \alpha \rightarrow \beta$ in $\mathbb{P}(\mathcal{B})$, and $\top \mid\!\sim \neg\alpha \notin \mathbb{P}(\mathcal{B}), \alpha \mid\!\sim \beta \notin \mathbb{P}(\mathcal{B})$, we prefer a closure operation including $\alpha \mid\!\sim \beta$ instead of $\top \mid\!\sim \neg\alpha$.

Recalling that WRM can be written also as

$$\frac{\top \mid\!\sim \alpha \rightarrow \beta}{\top \mid\!\sim \neg\alpha \qquad \alpha \mid\!\sim \beta}$$

So the request above can be interpreted as saying that WRM is an 'educated' application of the easy half of the deduction theorem ($\top \mid\sim \alpha \to \beta \Rightarrow \alpha \mid\sim \beta$), not a way to amplify our belief set.

So, given $\mathcal{B}$, we want to build an optimal preferential model $\mathfrak{M}$ satisfying $\mathcal{B}$ and s.t. $\mathbf{B}_{\mathfrak{M}} = \mathbf{B}_{\mathcal{B}}$ and $\mathbf{K}_{\mathfrak{M}} = \mathbf{K}_{\mathcal{B}}$, where $\mathbf{B}_{\mathfrak{M}}$ and $\mathbf{K}_{\mathfrak{M}}$ are, respectively, the belief set and the knowledge set defined by the model $\mathfrak{M}$.
Assume an agent characterized by $\mathcal{B}$; we can define the knowledge set by looking for the biggest clash, identifying the knowledge base $Q_{\mathcal{B}}$.
Once we have found $Q_{\mathcal{B}}$, we have to restrict the set of possible valuations of our model to those respecting the knowledge of the agent, that is, we have to construct our model over $U = \{w \in W | w \vDash \phi \textit{ for every } \phi \in Q_{\mathcal{B}}\}$, where W is the set of all the valuations of our language.
If we want a weakly rational closure, we need to define an optimal model, which can be done by means of a consistent set of defaults Δ.
Moreover, given Theorem 6.2.8, if we want our belief set to be the same as the preferential one, we need to have $Cn_{\mathcal{B}}(\Delta) = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_D}) = Cl(\overrightarrow{\mathcal{B}})$, i.e. our set of defaults has to be logically equivalent (with respect to $\vdash_{\mathcal{B}}$) to the materialization of the defeasible part of our conditional base.
However, in Section 6.2.1, we have seen that taking as default set simply the materialization $\overrightarrow{\mathcal{B}}$ of the conditional base $\mathcal{B}$ is not sufficient for the first desideratum, i.e. for the sequents in $\mathcal{B}$ to be satisfied in the model.
So we have to define a set of defaults such that its conjunction is $Cn_{\mathcal{B}}$-equivalent to $\overrightarrow{\mathcal{B}_D}$, and it guarantees the validity of every sequent in $\mathcal{B}$.
To reach such an aim, we can use the exceptionality order defined in [31] for the rational closure (see Section 4.2.3).

Construction of the model.

Let $\mathcal{B}$ be a conditional base.
Lehmann and Magidor ([31]) define a formula α *exceptional* for $\mathcal{B}$ iff $\mathcal{B}$ preferentially entails the sequent $\top \mid\sim \neg\alpha$. The sequent $\alpha \mid\sim \beta$ is said to be

exceptional for $\mathcal{B}$ iff its antecedent α is exceptional for $\mathcal{B}$.
The exceptionality of a formula α simply means that α is not satisfied in most normal worlds, i.e. the agent presumes that α is false (i.e. $\top \mid\sim \neg\alpha$). So, to verify a sequent $\alpha \mid\sim \beta$, we have to move to exceptional situations.
By definition, we have that $\top \mid\sim \neg\alpha$ iff $\neg\alpha \in \mathbf{B}_{\mathcal{B}}$, and we know, by Theorem 6.2.8, that $\neg\alpha \in \mathbf{B}_{\mathcal{B}}$ iff $\neg\alpha \in Cl(\overrightarrow{\mathcal{B}})$, that is, iff $\neg\alpha \in Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_D})$.
So we can construct the exceptionality ordering presented in [31] by checking the materializations of our sequents.

Proposition 6.2.12. *α is exceptional for a base $\mathcal{B}$ iff $\overrightarrow{\mathcal{B}} \vDash \neg\alpha$.*

Proof.
It is immediate from Definition 4.2.3 and Theorem 6.2.8.■

It is easy to see that, if every sequent in $\mathcal{B}$ is not exceptional, i.e. its antecedent is consistent with the agent's belief set, then we can build a model just stating $\Delta = \overrightarrow{\mathcal{B}_D}$.

Lemma 6.2.13. *Assume a base $\mathcal{B}$, partitioned in a knowledge portion $\mathcal{B}_K$ (with $Q_{\mathcal{B}}$ as the associated knowledge base) and a defeasible portion $\mathcal{B}_D$. If $\mathcal{B}_D$ is non-empty and does not contain exceptional sequents, then the default-assumption system $\mathfrak{S}_{\mathcal{B}} = \langle Q_{\mathcal{B}}, \overrightarrow{\mathcal{B}_D}\rangle$ is an injective weakly rational closure of $\mathcal{B}$.*

Proof.
If $\mathcal{B}_D$ is non-empty, then, by Proposition 6.2.11, $\mathcal{B}$ is preferentially consistent, and, by Corollary 6.2.10, $\overrightarrow{\mathcal{B}}$ is consistent. By Corollary 6.2.9, we have also that $\overrightarrow{\mathcal{B}_D}$ is $Cn_{\mathcal{B}}$-consistent.
This implies, by Theorem 6.1.11, that $\mathfrak{S}_{\mathcal{B}} = \langle Q_{\mathcal{B}}, \overrightarrow{\mathcal{B}_D}\rangle$ generates an injective weakly rational inference relation $\mid\sim_{\mathfrak{S}_{\mathcal{B}}}$.
The knowledge set of $\mathfrak{S}_{\mathcal{B}}$ is obviously the one defined by the preferential closure of $\mathcal{B}$, since $Q_{\mathcal{B}}$ is its knowledge base.

The belief base of $\mathbb{P}(\mathcal{B})$ is defined by the set $Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_D})$ (see Theorem 6.2.8); the belief set of $\mathfrak{S}_{\mathcal{B}} = \langle Q_{\mathcal{B}}, \overrightarrow{\mathcal{B}_D} \rangle$ is $\mathbf{B}_{\mathfrak{S}_{\mathcal{B}}} = Cl(Q_{\mathcal{B}} \cup \overrightarrow{\mathcal{B}_D})$ (see Proposition 6.1.7), that corresponds to $\mathbf{B}_{\mathfrak{S}_{\mathcal{B}}} = Cn_{\mathcal{B}}(\overrightarrow{\mathcal{B}_D})$. Hence, also the belief set is preserved.
Given WRM and the non-exceptionality of the sequents, it is immediate to check that $\mathcal{B}$ is satisfied by $\mathfrak{S}_{\mathcal{B}}$: if $\alpha \mid\sim \beta \in \mathcal{B}$, then $\top \mid\sim_{\mathfrak{S}_{\mathcal{B}}} \alpha \rightarrow \beta$ is valid. Since $\alpha \mid\sim \beta$ is not exceptional for $\mathcal{B}$, and $\mathfrak{S}_{\mathcal{B}}$ has the same belief set as $\mathcal{B}$, we have $\top \not\mid\sim_{\mathfrak{S}_{\mathcal{B}}} \neg\alpha$, and, since $\mid\sim_{\mathfrak{S}_{\mathcal{B}}}$ satisfies WRM, $\alpha \mid\sim_{\mathfrak{S}_{\mathcal{B}}} \beta$.
■

This result is also valid for preferential models $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\mathcal{B}_D} \rangle$, generated by $\mathfrak{S}_{\mathcal{B}} = \langle Q_{\mathcal{B}}, \overrightarrow{\mathcal{B}_D} \rangle$, i.e. where $U_{\mathcal{B}}$ is the set of worlds satisfying $Q_{\mathcal{B}}$, and $\delta_{\mathcal{B}_D}$ is the preferential order generated from the default set $\overrightarrow{\mathcal{B}_D}$.

Proposition 6.2.14. *Assume a base $\mathcal{B}$, partitioned in a knowledge portion $\mathcal{B}_K$ (with $Q_{\mathcal{B}}$ as the associated knowledge base) and a defeasible portion $\mathcal{B}_D$. If $\mathcal{B}_D$ is non-empty and does not contain exceptional sequents, then the preferential model $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\mathcal{B}_D} \rangle$ generates an injective weakly rational closure of $\mathcal{B}$.*

Otherwise, if we have in $\mathcal{B}$ some exceptional sequent $\alpha \mid\sim \beta$, we have to guarantee the validity of such a sequent in the model, since it cannot be derived by means of its material implication and WRM.
The use of the exceptionality ranking of [31] (see Section 4.2.3) seems the right tool.
Take a finite $\mathcal{B}_D$ and define the rank of every sequent (see Definition 4.2.4), in such a way that it is directly proportional to the exceptionality of the sequent: $rank(\alpha \mid\sim \beta) = 0$ if $\alpha \mid\sim \beta$ is not exceptional, while $rank(\gamma \mid\sim \rho) \geq rank(\phi \mid\sim \psi)$ means that $\gamma \mid\sim \rho$ is no less exceptional than $\phi \mid\sim \psi$.
Since $\mathcal{B}_D$ is clash-free, there will be a last rank $n+1$ which will result empty.

Lemma 6.2.15. *If a finite set of sequents $\mathcal{B}$ is clash-free, then its ranking is upper-bounded by $\emptyset$.*

Proof.
If $\mathcal{B}$ is clash-free, then there is no subset $\mathcal{D} \subseteq \mathcal{B}$ s.t. $\overrightarrow{\mathcal{D}} \models \neg\mathcal{D}^{\vee}$, i.e. s.t. every sequent in $\mathcal{D}$ is exceptional for $\mathcal{D}$ itself. This implies that every set C_{i+1} in the construction of the ranking (see Section 4.2.3) is a strict subset of the set C_i.
As said in Section 4.2.3, if $\mathcal{B}$ is a finite set, the sequence of C_i will stabilize to ∞, i.e., after some point, all Cs will be equal and completely exceptional. But, since $C_{i+1} \subset C_i$ for every C_{i+1} and $\mathcal{B}$ is finite, the stable set will necessarily be the empty set.
■

Let $\mathcal{A}_i$ indicate the set of the antecedents of the sequents in $\mathcal{B}_D$ of rank i.

$$\mathcal{A}_i = \{\alpha | \alpha \mid\sim \beta \in \mathcal{B}_D \text{ and } rank(\alpha) = i\}.$$

The intuition behind the construction of the default base is the following. Assume as default set the materializations of the sequents $\overrightarrow{\mathcal{B}_D}$. If a sequent $\gamma \mid\sim \rho$ is exceptional, its antecedent is not consistent with the entire default set $\overrightarrow{\mathcal{B}_D}$, in particular it is negated by the conjunction of its own materialization with the materializations of more normal sequents. So, given $\overrightarrow{\mathcal{B}_D} \vdash_{\mathcal{B}} \neg\gamma$, in looking for γ-maxiconsistent subsets of $\overrightarrow{\mathcal{B}_D}$, we could find some set in which $\gamma \rightarrow \rho$ is not present (see the 'penguin' example in Section 6.2.1, with respect to the exceptional sequents $\alpha \mid\sim \beta$ and $\alpha \mid\sim \neg\gamma$); in such a case it is possible that the exceptional sequent $\gamma \mid\sim \rho$ is not satisfied by the model.
We want to modify our default set in such a way that $\gamma \rightarrow \rho$ is present in every γ-maxiconsistent set. We could change the default formulae associated with a more normal sequent, i.e. generated by its materialization, in such a way that it will be γ-inconsistent by itself; in this way, we eliminate the defaults associated to more normal situations from the construction of γ-maxiconsistent sets. In particular, we shall obtain a single γ-maxiconsistent

set, composed by the defaults associated to the sequents in $\mathcal{B}$ at least as exceptional than γ, since, by the construction of the ranking, the conjunction of their materializations is consistent with γ, and all the other sequents, more normal than $\gamma \mid\sim \rho$, have been made γ-inconsistent by themselves.
We can obtain the γ-inconsistency of more normal sequents in a direct way: If γ is an exceptional formula, while $\alpha \mid\sim \beta$ is sequent of lower ranking, we can think of the normal situations for $\alpha \mid\sim \beta$, i.e. the minimal valuations for the antecedent α, to be such that γ is negated. To do this is sufficient to conjunct $\neg\gamma$ to the materialization $\alpha \to \beta$.
An example. Assume a conditional base $\mathcal{B}_D = \{\alpha \mid\sim \beta, \gamma \mid\sim \rho\}$ and suppose $\alpha \mid\sim \beta > \gamma \mid\sim \rho$, because $\alpha \to \beta, \gamma \to \rho \vDash \neg\gamma$, but $\alpha \to \beta \nvDash \neg\gamma$, $\alpha \to \beta \nvDash \gamma \to \rho$ and $\gamma \to \rho \nvDash \neg\gamma$. Assume as our defaults the materializations $\alpha \to \beta$ and $\gamma \to \rho$. In this case, when we look for γ-maxiconsistent sets, we have $\{\alpha \to \beta\}$ and $\{\gamma \to \rho\}$, and, as a result, $\gamma \not\mid\sim \rho$.
Add $\neg\gamma$ to the materialization $\alpha \to \beta$, so that a world satisfying $(\alpha \to \beta) \wedge \neg\gamma$ becomes more normal than (preferred to) every world satisfying γ.
Imposing this, in looking for the minimal worlds satisfying γ, we do not take into consideration the defaults generated by more normal sequents than $\gamma \mid\sim \rho$.

On the basis of this intuition, we can define a default formula $\sigma_{\alpha\mid\sim\beta}$ from every sequent $\alpha \mid\sim \beta$ in $\mathcal{B}_D$:
given $\alpha \mid\sim \beta \in \mathcal{B}_D$ and $rank(\alpha \mid\sim \beta) = i$,

$$\sigma_{\alpha\mid\sim\beta} = (\alpha \to \beta) \wedge \bigwedge\{\neg\gamma | \gamma \in \mathcal{A}_j, \text{ with } j > i\}$$

and we define the default-assumption set $\Delta_{WR\mathcal{B}}$ for the weakly rational closure as

$$\Delta_{WR\mathcal{B}} = \{\sigma_{\alpha\mid\sim\beta} | \alpha \mid\sim \beta \in \mathcal{B}_D\}$$

The form of the formula $\sigma_{\alpha\mid\sim\beta}$ may look a bit cumbersome, but it simply states that we put in our default set the materialization of every sequent together with the negation of the antecedent of every more exceptional sequent.

First of all, we can prove that $\Delta_{WR\mathcal{B}}$ is $Cn_{\mathcal{B}}$-equivalent to $\overrightarrow{\mathcal{B}_D}$ and so preserves the preferential belief set.

Lemma 6.2.16. *Assume a conditional base $\mathcal{B}$ s.t. $\mathcal{B}_D$ is its defeasible portion and $Cn_{\mathcal{B}}$ the consequence operation defining its knowledge. $\Delta_{WR\mathcal{B}}$ is $Cn_{\mathcal{B}}$-equivalent to $\overrightarrow{\mathcal{B}_D}$.*

Proof.
It is sufficient to show that $\Delta_{WR\mathcal{B}}$ and $\overrightarrow{\mathcal{B}_D}$ are logically equivalent, since the supraclassical monotonic operator $Cn_{\mathcal{B}}$ satisfies Left Logical Equivalence.
Since $\overrightarrow{\mathcal{B}_D} \vDash \neg\gamma$ for every exceptional γ, i.e. every γ s.t. $rank(\gamma) > 0$, and since Cl satisfies idempotence, the adding of such formulae $\neg\gamma$ to $\overrightarrow{\mathcal{B}_D}$ does not change its classical closure. Hence,

$$Cl(\Delta_{WR\mathcal{B}}) = Cl(\overrightarrow{\mathcal{B}_D})$$

■

This guarantees that, substituting $\Delta_{WR\mathcal{B}}$ to $\overrightarrow{\mathcal{B}_D}$ as default set, the belief set does not change.
We can prove that the default base $\Delta_{WR\mathcal{B}}$ validates every sequent in $\mathcal{B}$.

Lemma 6.2.17. *Assume a base $\mathcal{B}$, partitioned in a knowledge portion $\mathcal{B}_K$ (with $Q_{\mathcal{B}}$ as the associated knowledge base) and a defeasible portion $\mathcal{B}_D$. The default-assumption system $\mathfrak{S}_{\mathfrak{B}} = \langle Q_{\mathcal{B}}, \Delta_{WR\mathcal{B}} \rangle$ is an injective weakly rational closure of $\mathcal{B}$.*

Proof.
Since $\Delta_{WR\mathcal{B}}$ and $\overrightarrow{\mathcal{B}_D}$ are logically equivalent, we have that $\Delta_{WR\mathcal{B}}$ is $Cn_{\mathcal{B}}$-consistent, and so $\mathfrak{S}_{\mathfrak{B}}$ defines an injective weakly rational inference relation

$\mid\!\sim_{\mathfrak{S}_{\mathcal{B}}}$.
From the definition of $Q_{\mathcal{B}}$ and the lemma above we are guaranteed that the belief and the knowledge sets are the preferential ones.
We have to show that if $\alpha \mid\!\sim \beta \in \mathcal{B}$, then $\alpha \mid\!\sim_{\mathfrak{S}_{\mathcal{B}}} \beta$.
Assume $rank(\alpha) = 0$; then α is consistent with the entire set $\Delta_{WR\mathcal{B}}$. Since $\alpha \rightarrow \beta$ is a conjunct in $\sigma_{\alpha \mid\sim \beta}$, then $\Delta_{WR\mathcal{B}} \vDash \alpha \rightarrow \beta$, and $\alpha, \Delta_{WR\mathcal{B}} \vDash \beta$.
So, $\alpha \mid\!\sim_{\mathfrak{S}_{\mathcal{B}}} \beta$.
Assume, $rank(\alpha) = i$, with $i > 0$. Then α is inconsistent with every $\sigma_{\gamma \mid\sim \rho}$ with $rank(\gamma \mid\!\sim \rho) < i$, since in every such formula is present $\neg\alpha$. On the other hand, from the Definition 4.2.4, we can see that α is consistent with the set $\Delta_{\alpha} \subseteq \Delta_{WR\mathcal{B}}$, s.t. $\Delta_{\alpha} = \{\sigma_{\gamma \mid\sim \rho} | rank(\gamma \mid\!\sim \rho) \geq i\}$, and that Δ_{α} is the α-maximal consistent subset of $\Delta_{WR\mathcal{B}}$.
Since $\sigma_{\alpha \mid\sim \beta} \in \Delta_{\alpha}$, we have $\Delta_{\alpha} \vDash \alpha \rightarrow \beta$, hence $\alpha, \Delta_{\alpha} \vDash \beta$ and $\alpha \mid\!\sim_{\mathfrak{S}_{\mathcal{B}}} \beta$.
■

As above, we can state the same result referring to the generated preferential model.

Proposition 6.2.18. *Assume a base $\mathcal{B}$, partitioned in a knowledge portion $\mathcal{B}_K$ (with $Q_{\mathcal{B}}$ as the associated knowledge base) and a defeasible portion $\mathcal{B}_D$. The preferential model $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\Delta_{WR\mathcal{B}}} \rangle$ generates an injective weakly rational closure of $\mathcal{B}$.*

Hence, given whatever conditional base $\mathcal{B}$, we can find in an easy way a default-assumption set generating a satisfying weakly rational closure of $\mathcal{B}$.

Clash identification

As mentioned above, Bochman defined the notion of a clash of sequents, but he did not define a method to identify the "biggest" clash of a set of sequents. The use of materializations and Lehmann and Magidor exceptionality ranking can help us in finding such a clash in an easy way, suggested

by the Lemma 6.2.15.
We want to show that, given a finite conditional base $\mathcal{B}$, it has a largest clash iff its ranking ends up with a non-empty stable set.

Proposition 6.2.19. *Assume a finite conditional base $\mathcal{B}$ and rank its sequents. $\mathcal{B}$ has a largest clash $\mathcal{B}_K$ iff its ranking has a non-empty stable set of exceptional sequents ST, and $\mathcal{B}_K = ST$*

Proof.
Recall that a clash is a set of sequents $\mathcal{D}$ s.t. $\overrightarrow{\mathcal{D}} \models \neg \mathcal{D}^{\vee}$. As said in Section 4.2.3, if $\mathcal{B}$ is a finite set, the sequence of C_i will end with a stable set, i.e., after some point, all Cs will be equal and completely exceptional (maybe empty); we will call such final set ST.
Lemma 6.2.15 says us that if $ST \neq \emptyset$, then $\mathcal{B}$ has a clash. Equally, the definition of the notion of clash tells us that the ranking construction stops in a non-empty ST exactly if it finds a clash.
By the construction of the ranking, since we start from $\mathcal{B}$ eliminating step-by-step non-exceptional formulas, ST has to be the largest clash of $\mathcal{B}$.
■

Construction steps

So, given a conditional base $\mathcal{B}$, the construction of its injective weakly rational closure can be defined by means of the following steps:

- Materialize $\mathcal{B}$ in $\overrightarrow{\mathcal{B}}$.
- Define the ranking of the sequents in $\mathcal{B}$ by means of $\overrightarrow{\mathcal{B}}$.
- Identify the knowledge portion $\mathcal{B}_K$ of $\mathcal{B}$ by means of the stable set of the ranking operation. Automatically we have defined also $Q_{\mathcal{B}}$ and $\mathcal{B}_D$.
- Use the ranking and $\overrightarrow{\mathcal{B}_D}$ to define $\Delta_{WR\mathcal{B}}$.

- Construct $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\Delta_{WR\mathcal{B}}} \rangle$.

Valuation of the $\mathbb{W}$-closure

For some authoritative authors, as Makinson, Rational Monotony is too strong a condition for the definition of a satisfying closure operation over sequents (see Section 4.2.1), while we deem both Weak Rational Monotony and Disjunctive Rationality as desirable conditions. Hence, our original aim was the definition of a closure operation over sequents satisfying both WR and DR in an intuitive way, but not RM.
First of all, note that the satisfaction of WR and DR does not imply the satisfaction of RM.

Proposition 6.2.20. *The satisfaction of Weak Rational Monotony and of Disjunctive Rationality does not imply the satisfaction of Rational Monotony*

Proof.
It is sufficient to build a counter-example. Consider the model in Fig. 6.3.
Such a model is optimal and filtered, and consequently, by Theorems 4.2.3 and 5.1.20, it defines an inference relation that is both weakly rational and disjunctive. Notwithstanding, it does not satisfy Rational Monotony: we have $\neg p \vee \neg q \mid\!\sim \neg(p \leftrightarrow q)$ and $\neg p \vee \neg q \not\mid\!\sim q$, but $(\neg p \vee \neg q) \wedge \neg q \not\mid\!\sim \neg(p \leftrightarrow q)$.
■

Notwithstanding, we have not been able by now to identify a syntactical constraint over a default set Δ implying the satisfaction of the filteredness condition by the generated model. Hence, our closure operation does not imply the satisfaction of DR.
Our model does not satisfy DR.
For example, assume a preferentially consistent conditional base $\mathcal{B}$ composed of four sequents of rank 0, s.t. we can define a default set for the $\mathbb{W}$-closure with four default formulas $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$. They are mutually consistent and

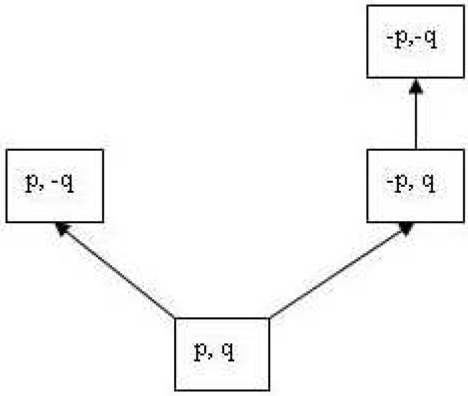

Figure 6.3: A model defining an inference relation satisfying WR, DR, but not RM.

the formula $\phi := \neg(\sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4)$ is of rank 1, i.e. an exceptional formula. The formula ϕ cannot be satisfied by the optimal worlds of the model, i.e. the worlds satisfying $(\sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4)$. However, assume we have six valuations $w_1 - w_6$ such that:
$w_1 \models \sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \neg\sigma_4$
$w_2 \models \sigma_1 \wedge \sigma_2 \wedge \neg\sigma_3 \wedge \sigma_4$
$w_3 \models \sigma_1 \wedge \neg\sigma_2 \wedge \sigma_3 \wedge \sigma_4$
$w_4 \models \neg\sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4$
$w_5 \models \sigma_1 \wedge \sigma_2 \wedge \neg\sigma_3 \wedge \neg\sigma_4$
$w_6 \models \neg\sigma_1 \wedge \neg\sigma_2 \wedge \sigma_3 \wedge \sigma_4$
We have that $w_1, w_2, w_3, w_4 \in min_{\prec}([\phi])$ and $w_5, w_6 \notin min_{\prec}([\phi])$. However, we have that only w_1 and w_2 are preferred to w_5, while only w_3 and w_4 are preferred to w_6. Hence, our model is not filtered.
Since RM implies DR, this example automatically shows also that $\mathbb{W}$-closure does not satisfy Rational Monotony, as desired.
Hence, $\mathbb{W}$-closure satisfies WR and Injectivity, that can be seen as a weakened form of Disjunctive Rationality (see Section 4.2.2).

However, does $\mathbb{W}$-closure define a satisfying inference relation?
We can evaluate it with respect to the examples presented in Section 4.2.3 for the evaluation of the behaviour of the rational closure.

∘ Nixon diamond:

1. $\rho \mid\sim \neg\pi$

2. $\mu \mid\sim \pi$

We interpret ρ, π, μ respectively as 'being a republican', 'being a pacifist', and 'being a quaker'.
The ranking of both the sequents is 0, since $\rho \to \neg\pi, \mu \to \pi \not\models \neg\rho$ and $\rho \to \neg\pi, \mu \mid\sim \pi \not\models \neg\mu$. Hence, by Theorem 6.2.13, we have $\Delta = \{\rho \to \neg\pi, \mu \to \pi\}$.
As in the rational case, we have $\rho \wedge \mu \not\mid\sim_{\Delta} \pi$ and $\rho \wedge \mu \not\mid\sim_{\Delta} \neg\pi$, as desired.

∘ Penguin triangle:

1. $\pi \mid\sim \beta$

2. $\pi \mid\sim \neg\phi$

3. $\beta \mid\sim \phi$

Here π, β, ϕ are respectively interpreted as 'being a penguin', 'being a bird', and 'being able to fly'.
(1) and (2) are exceptional sequents, while (3) is not. Hence, the default set is $\Delta = \{\pi \to \neg\beta, \pi \to \neg\phi, (\beta \to \phi) \wedge \neg\pi\}$.
Again, we have the same behaviour as rational closure. The following sequents are valid, since the premise is consistent with the entire Δ:
$\phi \mid\sim_{\Delta} \neg\pi$, $\neg\phi \mid\sim_{\Delta} \neg\beta$, $\neg\phi \mid\sim_{\Delta} \neg\pi$, $\beta \mid\sim_{\Delta} \neg\pi$, $\neg\beta \mid\sim_{\Delta} \neg\pi$, $\beta \wedge \pi \mid\sim_{\Delta} \neg\phi$, $\beta \wedge green \mid\sim_{\Delta} \phi$.
$penguin \wedge black \mid\sim_{\Delta} \neg fly$ is valid too, since the only subset of Δ maxiconsistent with $\pi \wedge black \mid\sim_{\Delta} \neg\phi$ is $\Delta' = \{\pi \to \neg\beta, \pi \to \neg\phi\}$.

As in the rational case, the following sequents are not endorsed:
$\beta \wedge \neg\phi \mid\!\sim_{\Delta} \pi$, $\beta \wedge \neg\phi \mid\!\sim_{\Delta} \neg\pi$, $\pi \mid\!\sim_{\Delta} \phi$.

Also, for the examples that result problematic for rational closure, the behaviour of our $\mathbb{W}$-closure is quite similar to the rational one. Only in the case of example (ii) we solve the problem.

(ii) Assume a conditional base $\mathcal{B} = \{\sigma \mid\!\sim \tau, \sigma \mid\!\sim \phi\}$, where σ, τ, ϕ stand for 'Swedish', 'tall', and 'fair' respectively. The sequents are not exceptional, and so the default set is simply $\Delta = \{\sigma \rightarrow \tau, \sigma \rightarrow \phi\}$. Contrary to $\mathbb{R}(\mathcal{B})$, $\mathbb{W}(\mathcal{B})$ satisfies 'Short Swedes are usually fair' ($\neg\tau \wedge \sigma \mid\!\sim \phi$). The reason of such a difference, as will be clarified in the next session, is due to the fact that rational closure treats the satisfaction of the sequents on the same level conjunctively, and if one of these sequents is not valid, then also the other ones do not hold; on the other hand, one advantage of $\mathbb{W}$-closure is just that the validity of each sequent on the same level is treated independently.

However, with examples (i) and (iii), $\mathbb{W}$-closure manifests exactly the same behaviour as $\mathbb{R}$-closure, showing the same problems of property-heredity from a situation to a more exceptional one. As the treatment by means of default-assumptions shows, the problem is mainly linked to the fact that we associate with the materialization of every sequent the negation of more exceptional situations, preventing, for example, normal properties to be inherited by exceptional items, even if they could be consistently added, as in the 'winged penguin' of example (i) in Section 4.2.3.

Hence, $\mathbb{W}$-closure seems to have a slightly better behaviour than rational closure, but they are still strongly similar.

6.2.4 Rational closure

We can extend this method to deal with closure under Rational Monotony. We will define a method to generate a preferential model corresponding to the rational closure presented in [31] (see Section 4.2.3), $\mathbb{R}(\mathcal{B})$. Such a result has been developed independently, finding only later that an analogous result was presented in [12].
Again, we start from a conditional base $\mathcal{B}$, and we define the ranking of its sequents working with the set of materializations $\overrightarrow{\mathcal{B}}$.
By means of such ranking we define the knowledge portion $\mathcal{B}_K$, the associated knowledge base $Q_{\mathcal{B}}$, and the defeasible portion $\mathcal{B}_D$.
Given such 'parameters', it is sufficient to change the definition of the set of default-assumptions to obtain a rational closure.
Recall that RM has the form

$$\frac{\alpha \mid\sim \gamma \qquad \alpha \not\mid\sim \neg\beta}{\alpha \wedge \beta \mid\sim \gamma}$$

Since WRM can be equivalently rewritten as

$$\frac{\top \mid\sim \beta \qquad \top \not\mid\sim \neg\alpha}{\alpha \mid\sim \beta}$$

intuitively, RM can be seen as a generalization of WRM from the premise $\top$ to every kind of antecedent formula α. This is also reflected by the characteristic models: weak rational inference relations are characterized by preferential models with a 'ground' level of optimal worlds, while rational inference relations are characterized by modular preferential models.
We will call our rational closure operation $\mathbb{C}(\mathcal{B})$, proving only at the end that it is equivalent to the rational closure $\mathbb{R}(\mathcal{B})$ by Lehmann and Magidor.
The main desiderata are the same as before:

- $\mathbb{P}(\mathcal{B}) \subseteq \mathbb{C}(\mathcal{B})$
- The knowledge and belief sets of $\mathbb{C}(\mathcal{B})$ have to be the same as $\mathbb{P}(\mathcal{B})$

And we have to readapt the third desiderata to deal with RM:

- If we have $\alpha \mid\sim \gamma$ in $\mathbb{P}(\mathcal{B})$, and $\alpha \mid\sim \neg\beta, \alpha \wedge \beta \mid\sim \gamma \notin \mathbb{P}(\mathcal{B})$, we prefer a closure operation adding $\alpha \wedge \beta \mid\sim \gamma$ instead of $\alpha \mid\sim \neg\beta$.

This is to stress that RM is a principle finalized to regulate the use of the monotony principle, i.e. it is functional to the strengthening of the premises, and not to the adding of unjustified information, as the adding of $\alpha \mid\sim \neg\beta$ would be.

Essentially, we have to modify the construction of the set of default-assumptions (that we will call $\Delta_{R\mathcal{B}}$) in order to generate a modular model $\mathfrak{M}$, but we have to preserve the validity of $\mathcal{B}$ in $\mathfrak{M}$ and the logical equivalence between $\Delta_{R\mathcal{B}}$ and $\overrightarrow{\mathcal{B}}$, in order to leave the belief set $B_{\mathcal{B}}$ unchanged.

The intuition is the following.

We want a modular model respecting the ranking of the sequents in $\mathcal{B}_D$.

Assume the sequents in $\mathcal{B}_D$ have been divided into n ranks.

Let us call $\mathcal{B}_{Di}$ the set of the sequents in $\mathcal{B}_D$ of rank i.

$$\mathcal{B}_{Di} = \{\alpha \mid\sim \beta \in \mathcal{B}_D | rank(\alpha \mid\sim \beta) = i\}$$

Such sets form a partition of $\mathcal{B}_D$ ($\mathcal{B}_D = \bigcup\{\mathcal{B}_{Di} | 0 \leq i \leq n\}$).

We point to the construction of a modular model of n levels, s.t. the worlds in level i are functional to the satisfaction of the sequents of rank i.

For example, assume that the domain of our model is the set of valuations $U_{\mathcal{B}}$, i.e. the valuation satisfying the knowledge base $Q_{\mathcal{B}}$. We want to partition $U_{\mathcal{B}}$ in levels $U_{\mathcal{B}i}$ ($0 \leq i \leq n$) in such a way that, given a sequent $\alpha \mid\sim \beta$ of rank i, the minimal worlds for α are between the worlds of level i ($min([\alpha]_{U_{\mathcal{B}}}) \subseteq U_{\mathcal{B}i}$), and $\alpha \rightarrow \beta$ is satisfied in every world in level i.

To obtain such a model, we can use the defaults generated for the weakly rational closure of the preceding paragraph. It is sufficient to join in a single default all the defaults generated for the sequents of the same rank.

That is, we want to generate a set $\Delta_{R\mathcal{B}}$ of default formulas composed by n defaults, one for each level of the ranking.

$$\Delta_{R\mathcal{B}} = \{\sigma_i | 0 \leq i \leq n\}$$

where

$$\sigma_i = \bigwedge\{\sigma_{\alpha \mid\sim \beta} | rank(\alpha \mid\sim \beta) = i\}.$$

$\sigma_{\alpha \mid\sim \beta}$ is the default generated by the sequent $\alpha \mid\sim \beta \in \mathcal{B}$, defined in the previous section for the weakly rational closure operation.
By logical equivalence, we obtain that

$$\sigma_i = \bigwedge\{\alpha \to \beta | rank(\alpha \mid\sim \beta) = i\} \wedge \bigwedge\{\neg\gamma | \gamma \in \mathcal{A}_j, \text{ with } j > i\}$$

By means of such defaults, we can obtain what we were looking for: for every sequent $\alpha \mid\sim \beta$ of rank i, the antecedent α is negated exactly by every σ_j with $j < i$, while σ_i is α-consistent ($min([\alpha]_{U_{K_\mathcal{B}}} \subseteq U_{\mathcal{B}i}$) and $\alpha \to \beta$ is implied by σ_i.
Since from the construction of the exceptionality ranking (see Section 4.2.3), we have that

$$\bigwedge\{\alpha \to \beta | rank(\alpha \mid\sim \beta) \geq i\} \vDash \bigwedge\{\neg\gamma | \gamma \in \mathcal{A}_j, \text{ with } j > i\},$$

and we have that $\neg\gamma \vDash \gamma \to \rho$ for every γ and ρ, by logical equivalence we obtain

$$\sigma_i = \bigwedge\{\alpha \to \beta | rank(\alpha \mid\sim \beta) \geq i\}$$

We have to show that $\mathfrak{S} = \langle Q_\mathcal{B}, \Delta_{R\mathcal{B}} \rangle$ generates a model $\mathfrak{M} = \langle U_\mathcal{B}, \delta_{\Delta_{R\mathcal{B}}} \rangle$ that is modular.

Lemma 6.2.21. *$\mathfrak{M} = \langle U_\mathcal{B}, \delta_{\Delta_{R\mathcal{B}}} \rangle$ is a modular model.*

Proof.
Recall by Definition 4.2.1 that an order δ over a set S is *ranked* iff there is a totally ordered set Ω (the strict order on Ω will be denoted by $<$) and a function $r : S \mapsto \Omega$ (the ranking function) s.t. $s \prec_\delta t$ iff $r(s) < r(t)$.
From our last formulation of the σ_is in $\Delta_{R\mathcal{B}}$ ($\sigma_i = \bigwedge\{\alpha \to \beta | rank(\alpha \mid\sim \beta) \geq i\}$) it is immediate to see that, if $i < j$, then

$$\sigma_i = \bigwedge\{\alpha \to \beta | j > rank(\alpha \mid\sim \beta) \geq i\} \wedge \sigma_j$$

That is, if $i < j$, $\sigma_i \vDash \sigma_j$. So, for every valuation $w \in U_{\mathcal{B}}$, if $w \vDash \sigma_i$, then $w \vDash \sigma_j$, for every $j \geq i$.
Hence, $(\Delta_{R\mathcal{B}})_w \supset (\Delta_{R\mathcal{B}})_v$ iff $|(\Delta_{R\mathcal{B}})_w| \supset |(\Delta_{R\mathcal{B}})_v|$, i.e. the superset ordering corresponds exactly to the size ordering.
Hence, we can define a totally ordered set $\Omega = \{0, \ldots, n\}$, and a function $r : U_{\mathcal{B}} \mapsto \Omega$ based on the cardinality of the set of defaults satisfied by every valuation:

$$r(w) = i \text{ iff } |(\Delta_{R\mathcal{B}})_w| = i$$

■

So, given $|(\Delta_{R\mathcal{B}})| = n$, we have $n+1$ levels of valuations, where the most normal worlds satisfy n defaults, and the most exceptional one satisfies no defaults at all.
Since the model is modular, it defines a rational inference relation. Obviously, we preserve also the preferential belief set, since $\Delta_{R\mathcal{B}}$ is logically equivalent to $\bigwedge\{\overrightarrow{\mathcal{B}}\}$.
We have to show that the sequents in $\mathcal{B}$ are valid in $\mathfrak{M}$
That is obvious, since, as we have seen before, if $rank(\alpha \mid\sim \beta) = i$, then $min([\alpha]_{U_{\mathcal{B}}}) \subseteq U_{\mathcal{B}i}$ and $\alpha \rightarrow \beta$ is satisfied in every world in $U_{\mathcal{B}i}$.
Finally, we can show that the inference operator defined by $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\Delta_{R\mathcal{B}}} \rangle$ corresponds exactly to the rational closure operation defined by Lehmann and Magidor.

Theorem 6.2.22. *The preferential model $\mathfrak{M} = \langle U_{\mathcal{B}}, \delta_{\Delta_{R\mathcal{B}}} \rangle$ is a canonical model of the rational closure $\mathbb{R}(\mathcal{B})$ defined in [31]*

Proof.
Assume a finite set of sequents $\mathcal{B}$.
Recall from Definition 4.2.5 that the rational closure $\mathbb{R}(\mathcal{B})$ of $\mathcal{B}$ is the set of all the sequents $\alpha \mid\sim \beta$ s.t. either

- The rank of α is strictly less than the rank of $\alpha \wedge \neg\beta$, or

- α has no rank.

By definition, α has no rank iff it is completely exceptional, i.e. iff there is a sequent $\alpha \mid\sim \beta$ in the clash of $\mathcal{B}$. This corresponds to say that $\neg\alpha$ is known by the agent, and that $\alpha \mid\sim \bot$ is valid in $\mathfrak{M}$. Hence, if α has no rank, $\alpha \mid\sim \beta$ is valid in $\mathfrak{M}$ for every β, and, conversely, if $\alpha \mid\sim \bot$, then α has no rank. $rank(\alpha) < rank(\alpha \wedge \neg\beta)$ means that, assumed $rank(\alpha) = i$, we have $min([\alpha]_{U_\mathcal{B}}) \subseteq U_{\mathcal{B}i}$, and $U_{\mathcal{B}i} \vDash \neg(\alpha \wedge \neg\beta)$, i.e. $U_{\mathcal{B}i} \vDash \alpha \rightarrow \beta$. Hence $\alpha \mid\sim \beta$ is valid in $\mathfrak{M}$. Conversely, if $\alpha \mid\sim \beta$ is valid in $\mathfrak{M}$, and $rank(\alpha) = i$, then $\sigma_i \vDash \alpha \rightarrow \beta$, and, by monotony, for every σ_j s.t. $j < i$, $\sigma_j \vDash \alpha \rightarrow \beta$. Hence $rank(\neg(\alpha \rightarrow \beta)) > i$.
■

So, we have defined a simple way to build a model of the rational closure of a finite set of sequents $\mathcal{B}$.

Construction steps

Given a conditional base $\mathcal{B}$, the construction of its rational closure can be defined by means of the following steps:

- Materialize $\mathcal{B}$ in $\overrightarrow{\mathcal{B}}$.
- Define the ranking of the sequents in $\mathcal{B}$ by means of $\overrightarrow{\mathcal{B}}$.
- Identify the knowledge portion $\mathcal{B}_K$ of $\mathcal{B}$ by means of the stable set of the ranking operation. Automatically we have defined also $Q_\mathcal{B}$ and $\mathcal{B}_D$.
- Use the ranking and $\overrightarrow{\mathcal{B}_D}$ to define $\Delta_{R\mathcal{B}}$.
- Construct $\mathfrak{M} = \langle U_\mathcal{B}, \delta_{\Delta_{R\mathcal{B}}} \rangle$.

6.2.5 Conclusions

As we have just seen, the default-assumption approach can be used to define injective preferential models appropriate for the definition of interesting non-Horn closures of sets of conditionals.

Lehmann and Magidor's rational closure can be built by means of a default set which turns out very easy to define. Moreover, we have defined a new closure operation which, referring to the examples, seems to have a slightly better behaviour than rational closure. However, both the operations manifest a problem with respect to the heredity of normal properties in exceptional situations; this is a problem that, by now, can be solved only in a trivial way, declaring explicitly in the conditional base which normal properties remain valid in the exceptional cases.

Chapter 7

Working with default-assumptions: reasoning about normality

Abstract. We define in a precise way the behaviour of default formulae by means of the correspondence between default-assumption and preferential approaches. In the end of the chapter we present a generalization of a model of stereotypical reasoning by Lehmann.

In the previous chapter, we have seen how the default-assumption approach can be used in the preferential framework to build up in a simple way interesting models of non-Horn closures. Instead, in the present chapter, we shall use the correspondence with injective preferential models to investigate more deeply the behaviour of the default-assumption approach.
The chapter is organized as follows.
In Section 7.1, we shall analyze, given a default-assumption set, which kinds of changes we can do to it without changing the generated consequence relation. Such behaviour shall be formalized by means of a *normality operator* $\rhd$.

In Section 7.2, we shall try to define the role of default-assumptions in common-sense reasoning on the basis of the results of the previous section, and finally, we propose a model for reasoning with stereotypes.

7.1 Characterizing normality

It is immediate from Definition 2.3.3 that distinct choices of a default information set Δ might give rise to indistinguishable consequence relations $\mid\!\sim_\Delta$. For example, it is evident that $\Delta = \{\alpha, \beta\}$ and $\Delta' = \{\alpha, \beta, \alpha \wedge \beta\}$ give rise to the same default-assumption consequence relations, since Δ and Δ' give rise to logically equivalent maxiconsistent sets for every formula γ. Notwithstanding, it is a notorious fact (see [39], p.34) that the behaviour of default-assumption systems is dependent upon the syntactical form of the set Δ, i.e. logically equivalent sets of formulae can give rise to different default-assumption inference relations.
On the other hand, on the side of injective preferential models, each preferential order δ determines uniquely the preferential inference relation $\mid\!\sim_\delta$.

Proposition 7.1.1. *Assume a set of valuations U. Given two strict orders δ, δ' over U,*

$$\delta = \delta' \text{ iff } \mid\!\sim_\delta = \mid\!\sim_{\delta'}$$

Proof. The direction from left to right follows directly from Definition 2.3.12. As to the converse, assume that $\mid\!\sim_\delta = \mid\!\sim_{\delta'}$ but $\delta \neq \delta'$. Then there is at least a pair (w, v) such that either $(w, v) \in \delta$ and $(w, v) \notin \delta'$, or $(w, v) \in \delta'$ and $(w, v) \notin \delta$. Assume, without loss of generality, that the first case holds. Let γ be a sentence satisfied only by w and v. Since $(w, v) \in \delta$ and $(w, v) \notin \delta'$ we have that $min_\delta(\{w, v\}) = \{w\}$, while

$$min_{\delta'}(\{w, v\}) = \begin{cases} \{w, v\}, & \text{if } (v, w) \notin \delta' \\ \{v\}, & \text{if} (v, w) \in \delta'. \end{cases}$$

Either way, by the injectivity of the preferential model, we have $\{\phi \mid \gamma \mid\!\sim_\delta \phi\} \neq \{\phi \mid \gamma \mid\!\sim_{\delta'} \phi\}$, contradicting the hypothesis that $\mid\!\sim_\delta = \mid\!\sim_{\delta'}$.
■ □

Note that by Theorem 3.1.3 this extends immediately to default-assumption consequence relations, that is to say, each ordering determines uniquely the corresponding default-assumption consequence relation.

If we think of a consequence relation as an agent, this can be intuitively interpreted as saying that, given a set of default information Δ, there is a certain amount of "change" that we can operate on a set Δ itself *while keeping its generated ordering fixed*, that is to say, according to our discussion of preferential reasoning, without altering the *normality* of the situation at hand. Roughly speaking then, our characterization of normality could be viewed as identifying the "epistemic changes" that a default-assumption consequence relation is capable of tolerating before "disgregating".
That is, we want to see which changes can be done to our default-assumption set without modifying the inference relation it determines, i.e. without modifying the preferential order it generates.

Definition 7.1.1. (Order stability)
Given a default assumption set Δ and a sentence ϕ, we shall say that the generated strict ordering δ_Δ is stable with respect to ϕ *just if $\delta_\Delta = \delta_{\Delta \cup \{\phi\}}$.*

It so happens that the statement and the proof of many of the following results is greatly simplified if we take reflexive orders as primitives instead of strict orders. This, however, does not make any conceptual difference, since, as Theorem 7.1.1 below shows, a strict order δ_Δ is stable with respect to a formula ϕ exactly when the correspondent preorder ε_Δ (see Definition 3.1.2) is stable with respect to ϕ.

Lemma 7.1.2. $\delta_\Delta = \delta_{\Delta \cup \{\phi\}}$ *if and only if for every $(w, v) \in \delta_\Delta$ and every w, v s.t. $\Delta_w = \Delta_v$, $v \vDash \phi$ implies $w \vDash \phi$.*

Proof.

($\Rightarrow$): $\delta_\Delta = \delta_{\Delta\cup\{\phi\}} \Rightarrow v \vDash \phi$ implies $w \vDash \phi$ for every $(w, v) \in \delta_\Delta$ and every w, v s.t. $\Delta_w = \Delta_v$

This is obvious by Definition 3.1.1.

($\Leftarrow$): $v \vDash \phi$ implies $w \vDash \phi$ for every $(w, v) \in \delta_\Delta$ and every w, v s.t. $\Delta_w = \Delta_v \Rightarrow \delta_\Delta = \delta_{\Delta\cup\{\phi\}}$

If $\phi \in \Delta$ the result is obvious, because the implicated part ($\delta_\Delta = \delta_{\Delta\cup\{\phi\}}$) is always true.
So we are going to analyze only the case in which $\phi \notin \Delta$.
Suppose that $\delta_\Delta \neq \delta_{\Delta\cup\{\phi\}}$. We want to show that this implies that there is a $(w, v) \in \delta_\Delta$ or there is a couple of worlds w,v with $\Delta_w = \Delta_v$ s.t. $v \vDash \phi$ does not imply $w \vDash \phi$, i.e. $v \vDash \phi$ and $w \nvDash \phi$.
If $\delta_\Delta \neq \delta_{\Delta\cup\{\phi\}}$, then there is at least a couple (u, t) s.t. $(u, t) \in \delta_\Delta$ and $(u, t) \notin \delta_{\Delta\cup\{\phi\}}$ or s.t. $(u, t) \in \delta_{\Delta\cup\{\phi\}}$ and $(u, t) \notin \delta_\Delta$. We'll treat each option separately.
First, some obvious facts.
The content of $(\Delta \cup \{\phi\})_w$,i.e. the subset of $\Delta \cup \{\phi\}$ satisfied by w, can be defined in relation to Δ_w and the valuation of ϕ in w:

If $w \vDash \phi$, $(\Delta \cup \{\phi\})_w = \Delta_w \cup \{\phi\}$
If $w \nvDash \phi$, $(\Delta \cup \{\phi\})_w = \Delta_w$

So, we have that $(\Delta \cup \{\phi\})_w \supseteq \Delta_w$.
Let's see the first option:

(a). If there is a couple (u, t) s.t. $(u, t) \in \delta_\Delta$ and $(u, t) \notin \delta_{\Delta\cup\{\phi\}}$, then there is a $(w, v) \in \delta_\Delta$ or there is a couple of worlds w,v with $\Delta_w = \Delta_v$ s.t.

$v \vDash \phi$ doesn't imply $w \vDash \phi$, i.e. $v \vDash \phi$ and $w \nvDash \phi$.

Suppose $(u,t) \in \delta_\Delta$ and $(u,t) \notin \delta_{\Delta\cup\{\phi\}}$. Then we have that $\Delta_u \supset \Delta_t$ and not $(\Delta \cup \{\phi\})_u \supset (\Delta \cup \{\phi\})_t$.
The values of $(\Delta \cup \{\phi\})_u$ and $(\Delta \cup \{\phi\})_t$ depend on the valuations of ϕ in u and t. So we have four possible situations:

1a) $u \vDash \phi \quad t \vDash \phi$
2a) $u \vDash \phi \quad t \nvDash \phi$
3a) $u \nvDash \phi \quad t \vDash \phi$
4a) $u \nvDash \phi \quad t \nvDash \phi$

1a) We have $(\Delta \cup \{\phi\})_u = \Delta_u \cup \{\phi\}$ and $(\Delta \cup \{\phi\})_t = \Delta_t \cup \{\phi\}$.
$\Delta_u \supset \Delta_t$, $\phi \notin \Delta_u$ and $\phi \notin \Delta_t$ hold, so we have $(\Delta\cup\{\phi\})_u \supset (\Delta\cup\{\phi\})_t$, and (u,t) must belong to $\delta_{\Delta\cup\{\phi\}}$ too. We are in contradiction with the hypothesis.

2a) We have $(\Delta \cup \{\phi\})_u = \Delta_u \cup \{\phi\}$ and $(\Delta \cup \{\phi\})_t = \Delta_t$.
$\Delta_u \supset \Delta_t$, and $\Delta_u \cup \{\phi\} \supset \Delta_u$ hold, so we also have $\Delta_u \cup \{\phi\} \supset \Delta_t$, and (u,t) must belong to $\delta_{\Delta\cup\{\phi\}}$ too. We are in contradiction with the hypothesis.

3a) We have $(\Delta \cup \{\phi\})_u = \Delta_u$ and $(\Delta \cup \{\phi\})_t = \Delta_t \cup \{\phi\}$.
We have $\phi \notin (\Delta \cup \{\phi\})_u$ and $\phi \in (\Delta \cup \{\phi\})_t$, so obviously $(\Delta \cup \{\phi\})_u \supset (\Delta \cup \{\phi\})_t$ is not true.

4a) We have $(\Delta \cup \{\phi\})_u = \Delta_u$ and $(\Delta \cup \{\phi\})_t = \Delta_t$.
We have $\Delta_u \supset \Delta_t$, so we have $(\Delta \cup \phi)_u \supset (\Delta \cup \phi)_t$, and (u,t) must belong to $\delta_{\Delta\cup\{\phi\}}$ too. We are in contradiction with the hypothesis.

The only case in which we don't have contradiction with the premise is (3a); that is, if (3a) is not satisfied, then it doesn't hold that $(u,t) \in \delta_\Delta$ and $(u,t) \notin \delta_{\Delta\cup\phi}$. So, if there is a pair (u,t) s.t. $(u,t) \in \delta_\Delta$ and $(u,t) \notin \delta_{\Delta\cup\phi}$, then there is a $(w,v) \in \delta_\Delta$, i.e. (u,t) itself, s.t. $v \vDash \phi$ and $w \nvDash \phi$.

(b). If there is a pair (u,t) s.t. $(u,t) \in \delta_{\Delta \cup \{\phi\}}$ and $(u,t) \notin \delta_\Delta$, then there is a $(w,v) \in \delta_\Delta$ or there are two worlds w,v with $\Delta_w = \Delta_v$ s.t. $v \models \phi$ doesn't imply $w \models \phi$, i.e. $v \models \phi$ and $w \not\models \phi$.

Suppose $(u,t) \notin \delta_\Delta$ and $(u,t) \in \delta_{\Delta \cup \{\phi\}}$. Then we have that $(\Delta \cup \{\phi\})_u \supset (\Delta \cup \{\phi\})_t$ and not $\Delta_u \supset \Delta_t$.
We define again our possible four possible situations:

1b)	$u \models \phi$	$t \models \phi$
2b)	$u \models \phi$	$t \not\models \phi$
3b)	$u \not\models \phi$	$t \models \phi$
4b)	$u \not\models \phi$	$t \not\models \phi$

1b) We have $(\Delta \cup \{\phi\})_u = \Delta_u \cup \{\phi\}$ and $(\Delta \cup \{\phi\})_t = \Delta_t \cup \{\phi\}$.
$\Delta_u \cup \{\phi\} \supset \Delta_t \cup \{\phi\}$, $\phi \notin \Delta_u$ and $\phi \notin \Delta_t$ hold, so we have $(\Delta)_u \supset (\Delta)_t$, and (u,t) must belong to δ_Δ too. We are in contradiction with the hypothesis.

2b) We have $(\Delta \cup \{\phi\})_u = \Delta_u \cup \{\phi\}$ and $(\Delta \cup \{\phi\})_t = \Delta_t$.
$\Delta_u \cup \{\phi\} \supset \Delta_t$, $\phi \notin \Delta_u$ and $\phi \notin \Delta_t$ hold. So we have that $\Delta_u \supseteq \Delta_t$.
In this case, the only possibility in which we have that $\Delta_u \supset \Delta_t$ is not true is when we have $\Delta_u = \Delta_t$.

3b) We have $(\Delta \cup \{\phi\})_u = \Delta_u$ and $(\Delta \cup \{\phi\})_t = \Delta_t \cup \{\phi\}$.
We have $\Delta_u \supset \Delta_t \cup \{\phi\}$ and $\Delta_t \cup \{\phi\} \supset \Delta_t$, so $\Delta_u \supset \Delta_t$ and (u,t) must belong to δ_Δ as well. We are in contradiction with the hypothesis.

4b) We have $(\Delta \cup \{\phi\})_u = \Delta_u$ and $(\Delta \cup \{\phi\})_t = \Delta_t$.
We have $(\Delta \cup \phi)_u \supset (\Delta \cup \phi)_t$, so we have $\Delta_u \supset \Delta_t$, and (u,t) must belong to δ_Δ as well. We are in contradiction with the hypothesis.

The only case in which we don't have contradiction with the premise is (2b) with $\Delta_u = \Delta_t$; that is, if (2b) with $\Delta_u = \Delta_t$ is not satisfied, then it doesn't hold that $(u, t) \notin \delta_\Delta$ and $(u, t) \in \delta_{\Delta \cup \phi}$. So, if there is a pair (u, t) s.t. $(u, t) \notin \delta_\Delta$ and $(u, t) \in \delta_{\Delta \cup \phi}$, then there are two worlds w and v, i.e. t and u respectively, s.t. $\Delta_w = \Delta_v$ and $v \vDash \phi$ doesn't imply $w \vDash \phi$ ($v \vDash \phi$ and $w \nvDash \phi$).

Then our results are:
(a) If there is a pair (u, t) s.t. $(u, t) \in \delta_\Delta$ and $(u, t) \notin \delta_{\Delta \cup \phi}$, then there is a $(w, v) \in \delta_\Delta$, i.e. (u, t) itself, s.t. $v \vDash \phi$ doesn't imply $w \vDash \phi$.
(b) If there is a couple (u, t) s.t. $(u, t) \notin \delta_\Delta$ and $(u, t) \in \delta_{\Delta \cup \phi}$, then there are two worlds w and v, i.e. t and u respectively, s.t. $\Delta_w = \Delta_v$ and $v \vDash \phi$ doesn't imply $w \vDash \phi$.

From these we have that:
If $\delta_\Delta \neq \delta_{\Delta \cup \phi}$, then there is a $(w, v) \in \delta_\Delta$ s.t. $v \vDash \phi$ doesn't imply $w \vDash \phi$ or there are two worlds w and v s.t. $\Delta_w = \Delta_v$ and $v \vDash \phi$ doesn't imply $w \vDash \phi$.

From this we obtain:
If $v \vDash \phi$ implies $w \vDash \phi$ for every $(w, v) \in \delta_\Delta$ and every w, v s.t. $\Delta_w = \Delta_v$, then $\delta_\Delta = \delta_{\Delta \cup \phi}$.
■

On the other hand, the stability of a generated preorder ε_Δ is regulated by the following simple rule.

Lemma 7.1.3. $\varepsilon_\Delta = \varepsilon_{\Delta \cup \{\phi\}}$ *if and only if* $v \vDash \phi$ *implies* $w \vDash \phi$ *for every* $(w, v) \in \varepsilon_\Delta$

Proof. The implication from left to right follows directly from the definition of $\varepsilon_{\Delta \cup \{\phi\}}$.

As to the other direction note that

$$\varepsilon_\Delta = \{(w,v) \mid v \vDash \psi \;\Rightarrow\; w \vDash \psi \quad \text{for every}\, \psi \in \Delta\}. \tag{7.1}$$

$$\varepsilon_{\Delta\cup\{\phi\}} = \{(w,v) \mid v \vDash \psi \;\Rightarrow\; w \vDash \psi \quad \text{for every}\, \psi \in \Delta\cup\{\phi\}\}. \tag{7.2}$$

Now, since $v \vDash \phi$ implies $w \vDash \phi$ for every $(w, v) \in \varepsilon_K$, then equations (7.1) and (7.2) define exactly the same pairs. Thus $\varepsilon_\Delta = \varepsilon_{\Delta\cup\phi}$. □

The upshot of Lemma 7.1.2 and Lemma 7.1.3 is the following:

Theorem 7.1.1. $\delta_\Delta = \delta_{\Delta\cup\{\phi\}}$ *if and only if* $\varepsilon_\Delta = \varepsilon_{\Delta\cup\{\phi\}}$, *that is, if and only if* $v \vDash \phi$ *implies* $w \vDash \phi$, *for every* $(w, v) \in \varepsilon_\Delta$.

As a consequence of Theorem 7.1.1, we shall be freely swapping between δ_Δ and ε_Δ in what follows.

7.1.1 The normality operator $\triangleright$

Recall from Proposition 7.1.1 that every distinct default-assumption consequence relation is semantically represented by a distinct strict preferential order. We now define a preferential model and a corresponding notion of satisfiability, with the desideratum that only those sentences which, if added to Δ keep $\mid\!\sim_\Delta$ fixed, should be satisfied. This satisfiability relation gives us the building block to construct our normality operator.

Let $\mathcal{C}_U$ be the class of models of the form $\mathfrak{M} = (U, \varepsilon)$, with U a fixed set of valuations and ε a preorder over U.
Recall that the dominion U reflects the 'hard' information of the agent, its knowledge core (with K_U naming the associated knowledge set and Cn_U the associated monotonic consequence operator). To characterize the behaviour of defeasible information, we will assume that our agent has a fixed knowledge set, i.e. the models in the class $\mathcal{C}_U$ vary only with respect to the preorder ε, while the knowledge of the agent, i.e. the valuations in the dominion U of the model, is fixed.

We say that $\mathfrak{M}$ satisfies ϕ, written $\mathfrak{M} \Vdash \phi$, just if ϕ is *compatible* with ε, that is

$$\mathfrak{M} \Vdash \phi \quad \text{iff} \quad v \vDash \phi \Rightarrow w \vDash \phi, \forall (w, v) \in \varepsilon. \tag{7.3}$$

We can now define our normality operator $\rhd$ by putting $\Delta \rhd \phi$ only if ϕ is satisfied by every model $\mathfrak{M} \in \mathcal{C}_U$ that satisfies Δ, in the sense of formula (7.3):

Definition 7.1.2.

$$\Delta \rhd \phi \quad \textit{iff} \quad \mathfrak{M} \Vdash \psi, \forall \psi \in \Delta, \textit{ implies} \, \mathfrak{M} \Vdash \phi.$$

The next Proposition justifies the intuitive reading of $\rhd$ as a *normality operator* in the light of the above discussion.

Proposition 7.1.4.

$$\Delta \rhd \phi \quad \text{iff} \quad \varepsilon_\Delta = \varepsilon_{\Delta \cup \{\phi\}}.$$

Proof. ($\Rightarrow$): suppose that $\Delta \rhd \phi$. This amounts to say that $v \vDash \phi \Rightarrow w \vDash \phi$ holds in every preorder ε such that $v \vDash \psi \Rightarrow w \vDash \psi$, for every $\psi \in \Delta$ and every $(w, v) \in \varepsilon$. Since ε_Δ is one of those preorders, we have $v \vDash \phi \Rightarrow w \vDash \phi$ for every $(w, v) \in \varepsilon_\Delta$. So $\varepsilon_\Delta = \varepsilon_{\Delta \cup \{\phi\}}$.

($\Leftarrow$): Suppose that $\varepsilon_\Delta = \{(w, v) \in U \times U \mid v \vDash \psi \;\Rightarrow\; w \vDash \psi$ for every $\psi \in \Delta\} = \varepsilon_{\Delta \cup \{\phi\}}$. Let $\mathfrak{M} = (U, \varepsilon)$ be an arbitrary model in $\mathcal{C}_U$. If $\mathfrak{M} \Vdash \psi$ for every $\psi \in \Delta$, then every pair $(w, v) \in \varepsilon$ satisfies $v \vDash \psi \;\Rightarrow\; w \vDash \psi$ for every $\psi \in \Delta$. But since all those pairs of valuations are in ε_Δ, it follows that $\varepsilon \subseteq \varepsilon_\Delta$. Since $\varepsilon_\Delta = \varepsilon_{\Delta \cup \{\phi\}}$, then $v \vDash \phi \;\Rightarrow\; w \vDash \phi$ for every $(w, v) \in \varepsilon_\Delta$. It therefore holds that $v \vDash \phi \;\Rightarrow\; w \vDash \phi$ for every $(w, v) \in \varepsilon$, that is to say, $\mathfrak{M} \Vdash \phi$. But $\mathfrak{M}$ was an arbitrary model, so we conclude that $\Delta \rhd \phi$, as required.

■ □

It is natural to ask, at this point, which kind of object is the operator $\rhd$. We shall begin by observing that $\rhd$ is a Tarskian operator.

Proposition 7.1.5. $\rhd$ *satisfies Reflexivity, Monotony and Cut.*

Proof.

- Reflexivity (REF):

$$\Delta \rhd \phi \text{ for every } \phi \in \Delta$$

This property is obviously satisfied: if $\mathfrak{M} \Vdash \psi$ for every $\psi \in \Delta$, then $\mathfrak{M} \Vdash \psi$ for every $\psi \in \Delta$.

- Monotony (MON):

$$\frac{\Delta \rhd \phi}{\Delta \cup \{\psi\} \rhd \phi}$$

Let Δ' be a superset of Δ. If $\mathfrak{M} \Vdash \gamma$ for every $\gamma \in \Delta'$, then, since $\Delta \subseteq \Delta'$, it holds that $\mathfrak{M} \Vdash \psi$ for every $\psi \in \Delta$. Therefore, since $\Delta \rhd \phi$, we obtain $\mathfrak{M} \Vdash \phi$.

- Cut (CT):

$$\frac{\Delta \cup \{\psi\} \rhd \phi \qquad \Delta \rhd \psi}{\Delta \rhd \phi}$$

Suppose that $\Delta \cup \{\psi\} \rhd \phi$ and $\Delta \rhd \psi$. Then, Proposition 7.1.4, the equations $\varepsilon_{\Delta\cup\{\psi\}} = \varepsilon_{\Delta\cup\{\psi\}\cup\{\phi\}}$ and $\varepsilon_{\Delta} = \varepsilon_{\Delta\cup\{\psi\}}$ hold. Thus, by substitution we obtain $\varepsilon_{\Delta} = \varepsilon_{\Delta\cup\{\phi\}}$. And from this we get, again by Proposition 7.1.4, that $\Delta \rhd \phi$, as required.

■

The following proposition relates $\rhd$ to the classical consequence relation $\models$.

Proposition 7.1.6. $\rhd$ *satisfies*

- *Tautology ($\top$):*

$$\Delta \rhd \top, \quad \textit{for any tautology } \top \textit{ of } \ell.$$

- *Contradiction* ($\bot$):

$$\Delta \rhd \bot, \quad \textit{for any contradiction } \bot \textit{ of } \ell.$$

- *Singleton Left Logical Equivalence (sLLE):*

$$\frac{\vDash \phi \leftrightarrow \psi \qquad \Delta \cup \{\phi\} \rhd \gamma}{\Delta \cup \{\psi\} \rhd \gamma}$$

- *Right Logical Equivalence (RLE):*

$$\frac{\vDash \phi \leftrightarrow \psi \qquad \Delta \rhd \phi}{\Delta \rhd \psi}$$

Proof. ($\top$) is straightforward: every world satisfies $\top$, so for every couple of worlds (w, v) in every ordering is valid $v \vDash \top \;\Rightarrow\; w \vDash \top$.
This means that we can add whichever tautology to our assumption set without changing the consequence relation.
Analogously, ($\bot$) is valid since no world satisfies a contradiction, and so for every couple of worlds (w, v) in every ordering is valid $v \vDash \bot \;\Rightarrow\; w \vDash \bot$.
Hence, adding a contradiction to Δ does not affect the maximally A-consistent subsets of Δ and therefore leaves the generated order unchanged.
Both (sLLE) and (RLE) follow from the fact that $\vDash \phi \leftrightarrow \psi$ implies $\mathfrak{M} \Vdash \phi$ if and only if $\mathfrak{M} \Vdash \psi$. ■

If we assume an agent equipped with a knowledge set K_U (i.e. the dominion of our preferential models is a set of valuation $U \subset W$, with W the set of all the valuation of our language), then we can strengthen the above rules with respect to K_U:

- Cn_U-Tautology ($K - \top$):
  If $\top \vdash_U \alpha$ (i.e., if $\alpha \in K_U$), then $\Delta \rhd \alpha$

- Cn_U-Contradiction $(K - \bot)$:
  If $\alpha \vdash_U \bot$ (i.e., if $\neg\alpha \in K_U$), then $\Delta \rhd \alpha$

- Singleton Left Cn_U-Equivalence (sLKE):

$$\frac{\vdash_U \phi \leftrightarrow \psi \qquad \Delta \cup \{\phi\} \rhd \gamma}{\Delta \cup \{\psi\} \rhd \gamma}$$

- Right Cn_U-Equivalence (RKE):

$$\frac{\vdash_U \phi \leftrightarrow \psi \qquad \Delta \rhd \phi}{\Delta \rhd \psi}$$

Again, they are proved by the fact that Cn_U-tautologies are satisfied by every world in U, while Cn_U-contradiction by no world in U.
Both (sLKE) and (RKE) follow from the fact that $\vdash_U \phi \leftrightarrow \psi$ implies $\mathfrak{M} \Vdash \phi$ if and only if $\mathfrak{M} \Vdash \psi$, for every $\mathfrak{M}$ with U as dominion.

Note that since $\rhd$ aims at characterizing invariance under any "normal refinement" of a default information set, it is only sensitive to *contingent facts* and therefore disregards as uninformative both tautologies (as we remarked above) and contradictions. This latter case can be illustrated by taking $\Delta = \{p, q\}$ and $A = \{\neg q\}$. Clearly there is only one maximally A-consistent subset of Δ, namely $\Delta_1 = \{p\}$. Let us now add a contradiction to Δ, so $\Delta' = \{p, q, \alpha \wedge \neg\alpha\}$. Again, there is only one maximally A-consistent subset of Δ', which is still $\Delta_1 = \{p\}$.

Note also that although default-assumption consequence relations are not closed under substitution of logically equivalent default information sets (see below, about Right Weakening), (sLLE) ensures that $\rhd$ is closed under singleton substitution.

We now move on to the behaviour of $\rhd$ with respect to the standard propositional connectives. Let us begin with the properties which $\rhd$ satisfies.

- Disjunction in the premises (OR):

$$\frac{\Delta \cup \{\phi\} \rhd \gamma \qquad \Delta \cup \{\psi\} \rhd \gamma}{\Delta \cup \{\phi \vee \psi\} \rhd \gamma}$$

Let $\mathfrak{M} = (U, \varepsilon)$ be a model. Assume $\Delta \cup \{\phi\} \rhd \gamma$, $\Delta \cup \{\psi\} \rhd \gamma$ and $\mathfrak{M} \Vdash \rho$ for every $\rho \in \Delta \cup \{\phi \vee \psi\}$, which means that for every $(w, v) \in \varepsilon$, if $v \vDash \rho$, then $w \vDash \rho$. Take one of those pairs $(w, v) \in \varepsilon$. We need to check three cases:

1) $v \vDash \phi \vee \psi$ and $w \vDash \phi \vee \psi$.
   Since $w \vDash \phi \vee \psi$, then either $w \vDash \phi$ or $w \vDash \psi$. Hence at least one of $v \vDash \phi \Rightarrow w \vDash \phi$ and $v \vDash \psi \Rightarrow w \vDash \psi$ is satisfied. Either way, $v \vDash \gamma$ implies $w \vDash \gamma$

2) $v \nvDash \phi \vee \psi$ and $w \vDash \phi \vee \psi$.
   The same argument as (1) applies.

3) $v \nvDash \phi \vee \psi$ and $w \nvDash \phi \vee \psi$.
   We have $v \nvDash \phi$, $v \nvDash \psi$, $w \nvDash \phi$ and $w \nvDash \psi$. Then $v \vDash \phi$ implies $w \vDash \phi$ and $v \vDash \psi$ implies $w \vDash \psi$. Hence $v \vDash \gamma$ implies $w \vDash \gamma$.

Summing up, if we assume that $\Delta \cup \{\phi\} \rhd \gamma$, $\Delta \cup \{\psi\} \rhd \gamma$, and $\mathfrak{M} \Vdash \rho$ for every $\rho \in \Delta \cup \{\phi \vee \psi\}$, then, for every pair $(w, v) \in \varepsilon$, $v \vDash \gamma \Rightarrow w \vDash \gamma$ holds, that is $\mathfrak{M} \Vdash \gamma$. So $\Delta \cup \{\phi \vee \psi\} \rhd \gamma$, as required.

- Introduction of conjunction ($I\wedge$):

$$\{\phi\} \cup \{\psi\} \rhd \phi \wedge \psi$$

If an ordering ε satisfies ϕ and ψ, then for every $(w, v) \in \varepsilon$, if $v \vDash \phi$, then $w \vDash \phi$, and if $v \vDash \psi$, then $w \vDash \psi$. Therefore, for such (w, v), if $v \vDash \phi \wedge \psi$, we have that $v \vDash \phi$ and $v \vDash \psi$, so also $w \vDash \phi$ and $w \vDash \psi$, i.e. $w \vDash \phi \wedge \psi$.

This property, together with MON and CT, gives us the AND rule:

$$\frac{\Delta \rhd \phi \qquad \Delta \rhd \psi}{\Delta \rhd \phi \wedge \psi}$$

- Cautious Introduction of disjunction ($CI\vee$):

$$\{\phi\} \cup \{\psi\} \rhd \phi \vee \psi$$

If an ordering ε satisfies ϕ and ψ, that means that for every $(w, v) \in \varepsilon$, if $v \vDash \phi$, then $w \vDash \phi$, and if $v \vDash \psi$, then $w \vDash \psi$. Then, for such (w, v), if $v \vDash \phi \vee \psi$, we have that $v \vDash \phi$ or $v \vDash \psi$, so also $w \vDash \phi$ or $w \vDash \psi$, i.e. $w \vDash \phi \vee \psi$.
Note that we need both premises to derive the disjunction. In particular the classical Introduction of disjunction ($\phi \rhd \phi \vee \psi$) is not valid. To see this, take $(w, v) \in \varepsilon$ s.t. $v \nvDash \phi$, $v \vDash \psi$, $w \nvDash \phi$ and $w \nvDash \psi$, so $v \vDash \phi \vee \psi$ and $w \nvDash \phi \vee \psi$. For this pair $v \vDash \phi \;\Rightarrow\; w \vDash \phi$, but $v \vDash \phi \vee \psi \;\not\Rightarrow\; w \vDash \phi \vee \psi$.

Finally, let us look at some of the properties which $\rhd$ does not satisfy.

- Right Weakening (RW):

$$\frac{\vDash \phi \rightarrow \psi \qquad \Delta \rhd \phi}{\Delta \rhd \psi}$$

- Modus Ponens (MP):

$$\frac{\Delta \rhd \phi \rightarrow \psi \qquad \Delta \rhd \phi}{\Delta \rhd \psi}$$

To see that $\rhd$ satisfies neither of the above, let (w, v) be $v \nvDash \phi$, $v \vDash \psi$, $w \nvDash \phi$, and $w \nvDash \psi$. This pair satisfies $v \vDash \phi \;\Rightarrow\; w \vDash \phi$, both v and w satisfy $\phi \rightarrow \psi$, but $v \vDash \psi \;\not\Rightarrow\; w \vDash \psi$.

◦ Contraposition (CONTR):

$$\frac{\Delta \cup \{\psi\} \rhd \phi}{\Delta \cup \{\neg\phi\} \rhd \neg\psi}$$

Given $\Delta \cup \{\psi\} \rhd \phi$, it is sufficient to show that there is a model $\mathfrak{M}$ s.t. $\mathfrak{M} \Vdash \Delta \cup \{\neg\phi\}$ and $\mathfrak{M} \nVdash \neg\psi$, i.e. there is a pair (w, v) s.t. (w, v) satisfies $v \vDash \rho \;\Rightarrow\; w \vDash \rho$ for every $\rho \in \Delta$, $v \vDash \neg\phi \;\Rightarrow\; w \vDash \neg\phi$ and $v \vDash \neg\psi \;\not\Rightarrow\; w \vDash \neg\psi$. Take a pair (w, v) s.t. $v \vDash \rho$ for every $\rho \in \Delta$, $v \vDash \phi$, $v \nvDash \psi$, $w \vDash \rho$ for every $\rho \in \Delta$, $w \vDash \phi$, and $w \vDash \psi$. This pair satisfies $v \vDash \rho \;\Rightarrow\; w \vDash \rho$ for every $\rho \in \Delta$, $v \vDash \phi \;\Rightarrow\; w \vDash \phi$, $v \vDash \psi \;\Rightarrow\; w \vDash \psi$, and $v \vDash \neg\phi \;\Rightarrow\; w \vDash \neg\phi$, but it does not satisfy $v \vDash \neg\psi \;\Rightarrow\; w \vDash \neg\psi$.

The failure of LLE is connected with the failure of another rule of classical consequence relation:

◦ Left Conjunction (L$\bigwedge$):

$$\frac{\Delta \rhd \phi}{\bigwedge \Delta \rhd \phi}$$

Such a rule is not valid, and thus we cannot substitute a set of premises with a single formula obtained by their conjunction.
Let us look at a simple example. We always have $\phi, \psi \rhd \phi$ by REF, but it is possible that $\phi \wedge \psi \rhd \phi$ results not valid. Assume a pair of worlds (w, v) s.t. $v \vDash \phi$, $v \nvDash \psi$, $w \nvDash \phi$ and $w \nvDash \psi$. They satisfy $v \vDash \phi \wedge \psi \;\Rightarrow\; w \vDash \phi \wedge \psi$, since $\phi \wedge \psi$ is false in both of them, but $v \vDash \phi \;\not\Rightarrow\; w \vDash \phi$.
Obviously, this is also a counter-example for a special case of RW, the elimination-of-conjunction rule ($E\wedge$):

$$\phi \wedge \psi \rhd \phi.$$

Completeness

Let $C_{\triangleright}$ be the normality operator defined by the relation $\triangleright$:

$$C_{\triangleright}(A) = \{\alpha | A \triangleright \alpha\}.$$

Given a set of default-assumptions Δ, $C_{\triangleright}(\Delta)$ defines all the formulas we can add to Δ without changing the generated consequence relation $\mid\!\sim_{\Delta}$. We can prove that $C_{\triangleright}(\Delta)$ corresponds to the closure under conjunction and disjunction of the set Δ (modulo Cn_U-equivalence).

Note: in this paragraph and the following one we shall not consider the presence in the default set of logical equivalent formulae. That is, given a default system $\langle K, \Delta \rangle$, by '$\alpha$', for example, we do not refer properly to a formula $\alpha \in \Delta$, but the class of the Cn_K-equivalent formulae to α. Obviously, given RKE, if $\alpha \in \Delta$, every Cn_K-equivalent formula to α is in $C_{\triangleright}(\Delta)$.
Hence, from now to the end of the chapter, the constraints and results in the proofs have to be considered *modulo Cn_K-equivalence*, i.e. considering as if in our default sets we have only a single witness for every set of Cn_K-equivalent formulae.

First of all, some notation.
Given a set of formulae A, $A^{\wedge}$ is the set of formulae obtained by closing A under conjunction.

$$A^{\wedge} = \{\bigwedge B | B \subseteq A\}$$

Analogously for disjunction.

$$A^{\vee} = \{\bigvee B | B \subseteq A\}$$

We can immediately see that these operations are monotonic, that is, if $A \subseteq B$, then $A^{\wedge} \subseteq B^{\wedge}$ and $A^{\vee} \subseteq B^{\vee}$.
We also have to introduce the notion of *maximal* worlds for a formula α, i.e. the most exceptional worlds satisfying α.

Given a formula α, we call $max_\varepsilon([\alpha]_U)$ the set of the maximal worlds in $[\alpha]_U$ with respect to the preorder ε.

$$max_\varepsilon([\alpha]_U) = \{w \in [\alpha]_U | \text{ there is no } v \in [\alpha]_U, \text{ s.t. } w \prec_\varepsilon v\}$$

A world w is *maximal* for α iff $w \in max_\varepsilon([\alpha]_U)$.
Given that we work with finite models, the smoothness condition (see Definition 2.3.11) is valid also with respect to maximal worlds, i.e. $[\alpha]_U \neq \emptyset$ implies $max_\varepsilon([\alpha]_U) \neq \emptyset$ for every formula α.
Recall also that, given a preferential model $\mathfrak{M} = \langle U, \varepsilon\rangle$, we can define its characteristic default set Δ^ε (see Proposition 3.1.7):

$$\Delta^\varepsilon = \{\beta_w | w \in U\}$$

with

$$\beta_w := \bigvee_{v \preceq_\varepsilon w} \{\alpha_v\}$$

where α_v is the formula characterizing the valuation v.

We can use a generalization of Observation 15 in [12] to show that the closure under disjunction of the characteristic default set Δ^ε defines every default formula in ε.

Lemma 7.1.7. *Given a model $\mathfrak{M} = \langle U, \varepsilon\rangle$, and its characteristic default-assumption set Δ^ε, then a formula ϕ is a default formula in $\mathfrak{M}$ ($\Delta^\varepsilon \triangleright \phi$) iff it is Cn_U-equivalent to a formula $\rho \in (\Delta^\varepsilon)^\vee$.*

Proof.
From the validity of the CI$\vee$ rule, we automatically have that $\phi \in (\Delta^\varepsilon)^\vee$ implies that ϕ is a default in $\mathfrak{M}$.
We have to prove the converse. Assume ϕ is a default formula in $\mathfrak{M}$. If ϕ is Cn_U-inconsistent, then it is a disjunction over an empty set.
Suppose ϕ is consistent. Denote by β_ϕ the disjunction of every β_w, s.t. $w \in max_\varepsilon([\phi]_U)$. We want to prove that ϕ is Cn_U-equivalent to β_ϕ.

Suppose that a world v satisfies ϕ. If $v \in max_\varepsilon([\phi]_U)$, then $v \vDash \beta_v$, and hence $v \vDash \beta_\phi$. If $v \notin max_\varepsilon([\phi]_U)$, then there is a world u s.t. $v \prec u$ and $u \in max_\varepsilon([\phi]_U)$; then $v \vDash \beta_u$, and hence $u \vDash \beta_\phi$. So every world in U satisfying ϕ satisfies β_ϕ. $\phi \vdash_U \beta_\phi$.
For the converse, take a world v s.t. $v \vDash \beta_\phi$. There is a world u s.t. $u \in max_\varepsilon([\phi]_U)$ and $v \vDash \beta_u$.
We can have that $v = u$ or $v \preceq u$. If $v = u$, then obviously $v \vDash \phi$.
If $v \preceq u$ we have the same that $v \vDash \phi$, since $u \vDash \phi$ and ϕ is a default formula in $\mathfrak{M}$ (hence downward persistent).
So every world in U satisfying β_ϕ satisfies ϕ. $\beta_\phi \vdash_U \phi$.
Hence, every default formula ϕ in $\mathfrak{M} = \langle U, \varepsilon \rangle$ is Cn_U-equivalent to the disjunction of some subset of Δ^ε.
■

Given a default-assumption set Δ, we can build the generated preorder ε_Δ, and from this we can identify its characteristic default set $\Delta^{\varepsilon_\Delta}$. We want to show that $\Delta^{\varepsilon_\Delta}$ is equal to the closure of Δ under conjunction, i.e. the set $\Delta^\wedge$.
Assume, as usual, a set of worlds U as dominion of our models, defining our background knowledge K_U. First, let us prove that every consistent subset of Δ is Cn_U-equivalent to some Δ_w for some $w \in U$ (recall that $\Delta_w = \{\alpha \in \Delta | w \vDash \alpha\}$).

Lemma 7.1.8. *Assume that $\Psi \subseteq \Delta$ and that Ψ is Cn_U-consistent; then Ψ is Cn_U-equivalent to Δ_w for some $w \in U$.*

Proof.
Assume $\Psi \subseteq \Delta$. Define a set Δ_Ψ as

$$\Delta_\Psi = \Delta \cap Cn_U(\Psi)$$

Obviously, $\Delta_\Psi \subseteq \Delta$ and , since $\Psi \subseteq \Delta$, we have $\Psi \subseteq \Delta_\Psi$, and Ψ and Δ_Ψ are Cn_U-equivalent.
It is sufficient to prove that Δ_Ψ is the maximally consistent subset of Δ for

some formula.
Given a set Δ, define the set $\Delta^{\neg}$ as the set composed by the negated formulas in Δ:

$$\Delta^{\neg} = \{\neg\phi | \phi \in \Delta\}$$

Take the set $\Delta \setminus \Delta_{\Psi}$. If $\Delta \setminus \Delta_{\Psi} = \emptyset$, then $\Delta_{\Psi} = \Delta$, and, given the consistency of Ψ, Δ_{Ψ} is the set of defaults satisfied in every optimal world. If $\Delta \setminus \Delta_{\Psi} \neq \emptyset$, then take the set of the negated defaults that are not in Δ_{Ψ}, i.e. $(\Delta \setminus \Delta_{\Psi})^{\neg}$. There is at least one nonempty subset $(\Delta \setminus \Delta_{\Psi})^{\neg'}$ of $(\Delta \setminus \Delta_{\Psi})^{\neg}$ s.t. $(\Delta \setminus \Delta_{\Psi})^{\neg'}$ is Cn_U-maxiconsistent with Δ_{Ψ}. Otherwise, we have that $\Delta_{\Psi}, \neg\phi \vdash_U \bot$ for every ϕ in $\Delta \setminus \Delta_{\Psi}$, that is, $\Delta_{\Psi} \vdash_U \phi$. In such a case, by the definition of Δ_{Ψ} and since $\phi \in \Delta$, we would have $\phi \in \Delta_{\Psi}$, which contradicts $\phi \in \Delta \setminus \Delta_{\Psi}$.
Hence, there is a set $(\Delta \setminus \Delta_{\Psi})^{\neg'}$ s.t. $(\Delta_{\Psi} \cup (\Delta \setminus \Delta_{\Psi})^{\neg'})$ is satisfied by a world w, and Δ_{Ψ} is the maximal subset of Δ consistent with $\bigwedge\{(\Delta \setminus \Delta_{\Psi})^{\neg'}\}$, that is, $\Delta_{\Psi} = \Delta_w$.
■

Assumed a background knowledge K, let $\bot_K$ be the symbol for Cn_K-contradictions.

Lemma 7.1.9. *Given a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$, every ϕ in $\Delta^{\varepsilon_\Delta}$ is Cn_K-equivalent to a formula in $\Delta^{\wedge} - \bot_K$, and conversely.*

Proof.
Assume a default-assumption system $\mathfrak{S} = \langle K, \Delta \rangle$. From $\mathfrak{S}$ generate the corresponding preferential model $\mathfrak{M} = \langle U_K, \varepsilon_\Delta \rangle$. We can define its characteristic default set $\Delta^{\varepsilon_\Delta}$:

$$\Delta^{\varepsilon_\Delta} = \{\beta_w | w \in U\}$$

with

$$\beta_w := \bigvee_{v \preceq_\varepsilon w} \{\alpha_v\}$$

Obviously, every β_w is consistent, since it is a disjunction of consistent formulas.

As a first step, we want to show that, if Δ is Cn_K-consistent, then $\Delta^{\wedge} = \Delta^{\varepsilon_\Delta}$, i.e. that every β_w is Cn_K-equivalent to a consistent conjunction of defaults in Δ; otherwise, if Δ is not Cn_K-consistent, $\Delta^{\varepsilon_\Delta} = \Delta^{\wedge} - \bot_K$.
Name $\bigwedge \Psi$ the conjunction of the elements of some $\Psi \subseteq \Delta$.

Assume Δ is Cn_K-consistent; then every $\bigwedge \Psi$ is Cn_K-consistent as well. It is sufficient to prove that every β_w in $\Delta^{\varepsilon_\Delta}$is Cn_K-equivalent to Δ_w, the set of every default satisfied in w.
It is easy to prove that, for every v in U_K, $v \vDash \beta_w$ iff $v \vDash \phi$ for every $\phi \in \Delta_w$: if $v \vDash \phi$ for every $\phi \in \Delta_w$, then $\Delta_v \supseteq \Delta_w$ and $v \preceq w$, hence $v \vDash \beta_w$; otherwise, if $v \nvDash \phi$ for some $\phi \in \Delta_w$, then $\Delta_v \not\supseteq \Delta_w$, $v \not\preceq w$ and $v \nvDash \beta_w$.
This implies the Cn_K-equivalence between β_w and Δ_w.
Given that every $\bigwedge \Psi$ is Cn_K-consistent, by the lemma above we have that $\bigwedge \Psi$ is Cn_K-equivalent to some Δ_w.
Hence, every formula $\bigwedge \Psi$, with $\Psi \subseteq \Delta$, is Cn_K-equivalent to some β_w, and, conversely, every β_w is Cn_K-equivalent to some $\bigwedge \Psi$, with $\Psi \subseteq \Delta$ (in particular, it is Cn_K-equivalent to $\bigwedge \Delta_w$).
So, given the validity of the substitution of equivalent single formulas (sLKE and RKE), if Δ is Cn_K-consistent, then $\Delta^{\wedge} = \Delta^{\varepsilon_\Delta}$.

If Δ is not Cn_K-consistent, then $\bigwedge \Psi$ could be inconsistent too.
If $\bigwedge \Psi$ is inconsistent, then $\bigwedge \Psi$ is logically equivalent to $\bot_K$.
If $\bigwedge \Psi$ is consistent, then the argument above applies here too.
Hence, $\Delta^{\varepsilon_\Delta} = \Delta^{\wedge} - \bot_K$.
■

Finally, by Lemmas 7.1.7 and 7.1.9, we can obtain

Proposition 7.1.10. $C_{\triangleright}(\Delta)$ *is composed by the formulas* Cn_U*-equivalent to the formulas in* $(\Delta^{\wedge})^{\vee}$.

Proof.

By Lemmas 7.1.7 and 7.1.9 we have

$$C_{\triangleright}(\Delta) = (\Delta^{\varepsilon_\Delta})^{\vee} = (\Delta^{\wedge} - \bot_K)^{\vee}$$

Since $\emptyset \subseteq \Delta^{\varepsilon_\Delta}$ and $(\emptyset)^{\vee} = \bot_K$, we have $\bot_K \in (\Delta^{\wedge} - \bot_K)^{\vee}$, and we can simplify the above equation to:

$$C_{\triangleright}(\Delta) = (\Delta^{\wedge})^{\vee}$$

■

We can simplify the proposition above, showing that the closure under conjunction and under disjunction of Δ corresponds exactly to $(\Delta^{\wedge})^{\vee}$, modulo Cn_U-equivalence.

Theorem 7.1.11. *ϕ is a default formula in a generated preferential model $\mathfrak{M} = \langle U_K, \varepsilon_\Delta \rangle$ iff ϕ is Cn_U-equivalent to a formula in the disjunctive and conjunctive closure of Δ.*

Proof.
Since every formula in $(\Delta^{\wedge})^{\vee}$ is obviously also a formula in the conjunctive and disjunctive closure of a set of formulae Δ, it is sufficient to prove that every formula in the conjunctive and disjunctive closure of Δ is Cn_U-equivalent to a formula in $(\Delta^{\wedge})^{\vee}$. We can prove this by induction in the construction of the conjunctive and disjunctive closure of Δ.
Let γ be in such closure of Δ.
If $\gamma \in \Delta$, obviously we have that $\gamma \in (\Delta^{\wedge})^{\vee}$. Otherwise, γ have been obtained by the conjunction or the disjunction of two formulae α, β.
By induction step, we assume $\alpha, \beta \in (\Delta^{\wedge})^{\vee}$, or better, that they are Cn_U-equivalent to two formulae in $(\Delta^{\wedge})^{\vee}$. This corresponds to saying that α is Cn_U-equivalent to a formula $\alpha_1 \vee \ldots \vee \alpha_n$, with $\alpha_1, \ldots, \alpha_n \in \Delta^{\wedge}$, and that β is Cn_U-equivalent to a formula $\beta_1 \vee \ldots \vee \beta_m$, with $\beta_1, \ldots, \beta_m \in \Delta^{\wedge}$.
If γ is Cn_U-equivalent to $\alpha \vee \beta$, then γ is Cn_U-equivalent to $\alpha_1 \vee \ldots \vee \alpha_n \vee \beta_1 \vee \ldots \vee \beta_m$, which is a formula in $(\Delta^{\wedge})^{\vee}$.

If γ is Cn_U-equivalent to $\alpha \wedge \beta$, then γ is Cn_U-equivalent to $(\alpha_1 \vee \ldots \vee \alpha_n) \wedge (\beta_1 \vee \ldots \vee \beta_m)$, which, in turn, is equivalent to $(\alpha_1 \wedge \beta_1) \vee \ldots \vee (\alpha_1 \wedge \beta_m) \vee \ldots \vee (\alpha_n \wedge \beta_1) \vee \ldots \vee (\alpha_n \wedge \beta_m)$. Since every subformula $(\alpha_i \wedge \beta_j)$ is in $\Delta^\wedge$, we have that the formula is in $(\Delta^\wedge)^\vee$, i.e. γ is Cn_U-equivalent to a formula in $(\Delta^\wedge)^\vee$.
■

Hence, logical equivalence, closure under disjunction and closure under conjunction allow to reach every default formula determined by a preorder ε over a set of worlds U; this corresponds to saying that, given a set of default formulae Δ and a knowledge set K, the structural rules

- $(K - \top)$
- $(K - \bot)$
- (RKE)
- $(I\wedge)$
- $(CI\vee)$

represent a complete characterization of the derivation of default formulae from a set Δ, given a background knowledge K.

Minimal bases

Since we work with finite sets of formulae, given a finite default set Δ, it is easy to see that also its closures under conjunction ($\Delta^\wedge$) and under disjunction ($\Delta^\vee$) are finite: if $|\Delta| = n$, then $|\Delta^\wedge| = |\Delta^\vee| = 2^n$; hence, also $(\Delta^\wedge)^\vee$ is finite ($|(\Delta^\wedge)^\vee)| = 2^{2^n}$).
It is possible to show that every $\triangleright$-closed set D has got, again modulo Cn-equivalence, a minimal base, i.e. a smaller set of formulae that generates D under $\triangleright$-closure. This will be especially useful in the next chapter, when we

shall analyze default revision.

Theorem 7.1.12. *If a finite set D is closed under conjunction ($D = D^{\wedge}$), then there is a minimal set of formulae Δ s.t. $\Delta^{\wedge} = D$.*

Proof.
Assume a background knowledge K (corresponding to a semantical dominion U) and a finite set D, closed under conjunction. Given three formulae α, β, γ, we know that γ is Cn_U-equivalent to $\alpha \wedge \beta$ iff $[\gamma]_U = [\alpha]_U \cap [\beta]_U$.
Define an enumeration $\langle D \rangle = \langle \alpha_1, \ldots, \alpha_n \rangle$ of the formulae in D.
We construct Δ in the following way:

$$\alpha_i \in \Delta \text{ iff } [\alpha_i]_U \neq [\alpha_j]_U \cap [\alpha_k]_U \text{ for every} j, k \neq i$$

This simply states that we put a formula α in Δ iff we cannot obtain α from the conjunction of two other formulas in D. We have to prove that $\Delta^{\wedge} = D$. Since $\Delta \subseteq D$, we obviously have that $\Delta^{\wedge} \subseteq D$. Assume that $D \nsubseteq \Delta^{\wedge}$, i.e. there is a formula α_i s.t. $\alpha_i \in D$ and $\alpha_i \notin \Delta^{\wedge}$. This implies that $[\alpha_i]_U = [\alpha_j]_U \cap [\alpha_k]_U$ for some $\alpha_j, \alpha_k \in D$. Define as $\alpha_i^{\uparrow}$ the finite set of the formulae in D implied by α_i, that is, the set of the formulae s.t. the set of the worlds satisfying one of them includes the set of the worlds satisfying α_i; obviously, $\alpha_j, \alpha_k \in \alpha^{\uparrow}$. Necessarily, at least one between α_j and α_k is not in $\Delta^{\wedge}$, otherwise $\alpha_i \in \Delta^{\wedge}$.
Assume, without loosing generality, that $\alpha_j \notin \Delta^{\wedge}$. We can apply the same argument used for α_i, identifying two formulae in the set $\alpha_j^{\uparrow}$ s.t. α_j corresponds to their conjunction. Obviously $\alpha_j^{\uparrow} \subset \alpha_i^{\uparrow}$.
Since, given the finiteness of D, for every formula γ, $\gamma^{\uparrow}$ is a finite set, the iteration of the procedure above will end up with a formula β s.t. $\beta^{\uparrow} = \emptyset$. That means that β does not correspond to the conjunction of any other formula in D. This implies that β is in Δ. Contradiction.
Hence, $\Delta^{\wedge} = D$.
Finally, we have to show that Δ is a minimal base for D, i.e. that for every $\Gamma \subseteq D$ s.t. $\Gamma^{\wedge} = D$, $\Delta \subseteq \Gamma$.

Assume there is a Γ s.t. $\Delta \not\subseteq \Gamma$. Then there is a formula α s.t. $\alpha \in \Delta$ and $\alpha \notin \Gamma$.
$\alpha \in \Delta$ implies that $[\alpha] \neq [\beta] \cap [\gamma]$ for every $\beta, \gamma \in D$, i.e. for every $\beta, \gamma \in \Gamma^{\wedge}$.
Hence $\alpha \notin \Gamma^{\wedge}$ and $\Gamma^{\wedge} \neq D$. Contradiction.
Hence $\Delta \subseteq \Gamma$ for every set Γ s.t. $\Gamma^{\wedge} = D$.
■

Theorem 7.1.13. *If a finite set D is closed under disjunction ($D = D^{\vee}$), then there is a minimal set of formulae Δ s.t. $\Delta^{\vee} = D$.*

Proof.
The proof is analogous to the previous one. It is sufficient to refer to union instead of disjunction ($[\gamma]_U = [\alpha]_U \cup [\beta]_U$ instead of $[\gamma]_U = [\alpha]_U \cap [\beta]_U$), and to define, instead of $\gamma^{\uparrow}$, the set $\gamma^{\downarrow}$ of the formulae s.t. the set of the worlds satisfying them is included in the set of the worlds satisfying γ.
■

Theorem 7.1.14. *Given a default set Δ, its closure $(\Delta^{\wedge})^{\vee}$ has got a minimal default base Δ_m, that is, $(\Delta_m^{\wedge})^{\vee} = (\Delta^{\wedge})^{\vee}$ and $\Delta_m \subseteq \Gamma$ for every Γ s.t. $(\Gamma^{\wedge})^{\vee} = (\Delta^{\wedge})^{\vee}$.*

Proof.
By Theorem 7.1.13, $(\Delta^{\wedge})^{\vee}$ has got a minimal base Φ s.t. $\Phi^{\vee} = (\Delta^{\wedge})^{\vee}$ and $\Phi \subseteq \Delta^{\wedge}$.
We have that $\Phi \subseteq \Phi^{\wedge} \subseteq \Delta^{\wedge}$; by the monotonicity of $\vee$-closure and the fact that $\Phi^{\vee} = (\Delta^{\wedge})^{\vee}$, we have that $\Phi^{\vee} = (\Phi^{\wedge})^{\vee} = (\Delta^{\wedge})^{\vee}$.
By Theorem 7.1.12, $\Phi^{\wedge}$ has too a minimal base Δ_m, s.t. $\Delta_m \subseteq \Phi \subseteq \Delta$ and $\Delta_m^{\wedge} = \Phi^{\wedge}$.
Hence, $(\Delta_m^{\wedge})^{\vee} = (\Phi^{\wedge})^{\vee} = (\Delta^{\wedge})^{\vee}$.
■

7.2 What are default-assumptions?

Default-assumptions are generally deemed as referring to formulae holding in normal situations. However, if we defend such an interpretation, the analysis above highlights a strange behaviour. In particular, the failure of Right Weakening and Modus Ponens sounds strange.

Let us briefly comment on these two negative results. The failure of MP means that the normality closure of a sentence given a set of default information does not obey the laws of material implication. For example, in the default-assumption model, if an agent has as default information that the weather is sunny, and that if the weather is sunny the streets are not wet, such an agent is not allowed to deem as a default that the streets are not wet.

In particular, note that failure of RW implies the failure of Supraclassicality, that is, our operator $\rhd$ does not extend the classical consequence operator

If we want our default-formulae to describe a normal situation, it would be intuitive for such a characterization to be classically closed. Moreover, there is a gap between the formulae $\rhd$-derivable from a set Δ and the set of sequents of form $\top \mid\sim_{\Delta} \alpha$ generated by Δ: for example, given a $\Delta = \{\alpha, \alpha \rightarrow \beta\}$, we cannot generally add β to Δ without modifying the associated inference relation $\mid\sim_{\Delta}$, but we will have $\top \mid\sim_{\Delta} \beta$ as valid, i.e. we will deem β as normally holding in the model generated from Δ.

So, we cannot interpret default-assumption sets simply as sets containing formulae we deem as holding in normal situations, as generic *expectations*, in the way Gärdenfors and Makinson have characterized them (see [19]): the kind of closure defined by $\rhd$ does not account for such an interpretation.

If we interpret our defaults as simple expectations, we feel the need to characterize the default information of an ideal agent by means of the classical closure of its default set $Cl(\Delta)$. However, it is well known that if we close the set of default-assumptions Δ under classical consequence, not only we radically change the generated inference relation $\mid\sim_{\Delta}$, but, moreover, such change results in a trivial behaviour.

Theorem 7.2.1. *([39], Theorem 2.7) When $\Delta = Cl(\Delta)$, then $C_\Delta(A) = Cl(A)$ whenever A is inconsistent with Δ.*

Such a behaviour has been felt as one of the main flaws of default-assumption approach.
Hence, there is the risk to be forced to hold the default-assumption approach just as an useful tool to be combined with the classical preferential approach in order to make model-construction more simple, as in the previous chapter, without giving to it any role in the epistemological interpretation of our agents.
We do not think it is the case. Simply, it is necessary to differentiate between formulae simply holding in normal situations and formulae *characterizing* normal situations, i.e. formulae that allow us to distinguish a normal situation from an exceptional one. Undertaking such an interpretation, the failure of (RW) and (MP) does not sound strange anymore.
Let us give an example. Assume that a normal tiger is characterized by a tawny coat with black stripes ($\Delta = \alpha \wedge \beta$), and consequently we recognize that a tiger is a normal one if we are informed that its coat is tawny and has black stripes.
The fact that normally its coat has black stripes ($\top \mid\!\sim_\Delta \beta$), with the known fact that a striped coat is not uniform ($\vdash \beta \rightarrow \gamma$), implies by MP that normally a tiger does not have an uniform coat ($\top \mid\!\sim_\Delta \gamma$). However, the non-uniformity of the tiger's coat does not give us any information about the normality of the situation: if we only know that a particular tiger has a non-uniform coat, this does not allow us to say that it is a normal tiger, since its coat could be, for example, a spotted coat. If a spotted tiger is more normal than a tiger with a uniform coat would be an open question, and hence we are not justified in adding the non-uniformity of the coat (γ) to the default set Δ.
Hence, default-assumptions acquire a prominent status between the formulae holding in normal situations, i.e. as *distinguishing features* of normality. Such a status makes them a useful formalization, for example for stereotyp-

ical reasoning, as pointed out by the tiger example just made.
This characterization of default formulae justifies also the other properties of $\triangleright$.
Let us give some examples.
The introduction of conjunction is very intuitive, since, given that a couple of formulae α and β characterizes the normality status of a situation, all the more the satisfaction of both of them, i.e. of $\alpha \wedge \beta$, will indicate that the situation we are in is a normal one. Conversely, the failure of the elimination of conjunction is desirable: if we have that a normality is characterized by the satisfaction of $\alpha \wedge \beta$, then it is not obvious at all to deem also α by itself as a default formula. The example above about the tiger is clarifying: the normal coat of a tiger must be both tawny and striped; if we simply know that the coat of a tiger is tawny, and we do not know anything about its pattern, we do not have sufficient clues to judge the normality of the tiger's coat.
The same can be said about the properties of disjunction. Cautious introduction of disjunction is intuitive: if we maintain α and β as default formulae, it is obvious that the satisfaction of at least one of them, that is, the satisfaction of $\alpha \vee \beta$, will indicate that we are moving toward normality. The same cannot be said about the traditional introduction of conjunction: if α characterizes typicality, every normal situation will also satisfy $\alpha \vee \gamma$ for every γ. However, such a formula cannot be used to characterize normality: a situation satisfying γ would also satisfy $\alpha \vee \gamma$, but, since γ does not characterize normality, the satisfaction of $\alpha \vee \gamma$ does not tell us anything about the typicality of the situation.

7.2.1 Reasoning with stereotypes

As the above example about the coat of the tiger shows, the operator $\triangleright$ seems appropriate for the management of stereotypes.
Stereotypes have played an essential role, both in philosophy and in cog-

nitive science, in the development of a theory of concepts, both for what regards their structures and their role in reasoning (for an introduction, see [28]). Stereotypes play an essential role in Putnam's social characterization of meaning ([47]), and play a role in most of the actual notions of conceptualizations: for some commentators, they are at the core of the structure of many kinds of concepts (see [43]), but such a position has often been criticized. Despite this, the role of stereotypes in many dimensions of commonsense reasoning, from the classification of an object to uncertain reasoning, is universally recognized.

> "Most actual cases of prototype phenomena simply are not used in 'identification'. They are used instead in thought - making inferences, doing calculations, making approximations, planning, comparing, making judgments, and so on - as well as in defining categories, extending them, and characterizing relations among subcategories. Prototypes do a great deal of the real work of mind, and have a wide use in rational processes." [27], p.418

Given a concept X, a stereotype can be seen as the most distinctive example of X. In a logical approach, it can be characterized by means of a set of formulae, defining the most typical properties that an element of X should satisfy.

The use of stereotypes differs from the classical definitional theory of concepts, which most authoritative reference is Locke: in the latter approach, the properties connected to a concept X are individually necessary and jointly sufficient for the application of the concept X to an item; in a stereotypical perspective toward classification, the properties connected to the stereotype of X are not strictly necessary conditions for an item to be an X, but represent simply the most prominent properties connected with the elements of the category X. The stereotypical view accounts for the vagueness of concepts: the satisfaction of every property expressed by a stereotype identifies especially good examples of a concept, but other items are recognized as elements of the category, despite not satisfying every stereotypical property.

The behaviour of default formulae makes them good candidates for modeling stereotypical properties. Such an intuition has already been suggested by Lehmann ([30]). He proposes a logical model for stereotypical reasoning: the agent starts with a set of n 'stereotypes', represented by means of a finite set of sets of default formulae $S = \{\Delta_1, \ldots, \Delta_n\}$, and the information about an individual, represented, as usual, by means of a finite set of premises A. He proposes a notion of semantic distance $d(A, \Delta)$ between actual information and stereotypical information, in order to model the classification of our individual under a stereotype represented in S. The agent associates the set of premises A to the 'nearest' default set in S (that we call S^A), allowing the agent to complete its own premises by means of the information contained in the chosen stereotype.

Lehmann proposes a set of minimal intuitive constraints that such notion of distance should satisfy; the chosen stereotype should be the one that fits better the information in the premises, and he valuate such 'fitness' referring to the overlap between the set of semantic valuations of both premises and defaults ($[A]$ and $[\Delta]$ respectively).

In particular, given a set of stereotypes S and a premise set A, he imposes that:

- For every A and S, the choosing procedure picks up just a single element in S (that is, $|S^A| = 1$). This is just a simplifying assumption.

- For every $\Delta \in S$, $d(A, \Delta)$ should be anti-monotonic with respect to $[\Delta] \cap [A]$ (the larger the overlap, the smaller the distance).

- For every $\Delta \in S$, $d(A, \Delta)$ should be monotonic with respect to $[\Delta] - [A]$ (the larger the set of situations in which the defaults can be satisfied, but not the premises, the larger the distance).

Hence, Lehmann imposes the following basic constraint over our notion of distance:

$$\left.\begin{array}{l}[\Delta'] \cap [A'] \subseteq [\Delta] \cap [A] \\ [\Delta] - [A] \subseteq [\Delta'] - [A']\end{array}\right\} \Rightarrow d(A, \Delta) \leq d(A', \Delta')$$

The chosen stereotype is the one with a minimal distance from A:

$$S^A = \{\Delta_i \in S|\ d(A, \Delta_i) \leq d(A, \Delta_j) \text{ for every } \Delta_j \in S)\}$$

The consequence operator is defined by adding the formulae in the default set $\Delta^A \in S^A$ to the premise set A.

$$A \mid\!\sim_S \beta \text{ iff } A \cup \Delta^A \vDash \beta$$

Lehmann proves that, if the distance function d satisfies the constraint above, then the choice function defines a cumulative inference relation ([30], Corollary 5.6). Lehmann maintains cumulativity as an intuitive property of this kind of reasoning, but admits that it should be tested experimentally.
However, Lehmann's model manifests a series of problems. As you can see, Lehmann's proposal associates a prototype Δ as background knowledge to a set of premises A only in case they are mutually consistent, otherwise there would be no overlap between the two sets of valuations. This is made explicit in the following extract, where we have to interpret the notation in the following way: capital Roman letters represent sets of valuations, and, in particular, given, in our notation, a premise set A and the chosen default set Δ^A, F and S^A correspond, respectively, to our sets of valuations $[A]$ and $[\Delta^A]$.

> "The reasoner will then conclude that the actual state of affairs is one of the members of the intersection $F' :=_{def} F \cap S^F$. The nonmonotonicity of the reasoning stems from the jump from F to the subset F'. Clearly, we do expect F' to be non-empty, assuming F is non-empty, since we want to avoid jumping to contradictory conclusions. It will be the task of the function that defines the best stereotype to pick a stereotype that has a non-empty intersection with the information F at hand." [30], p.51

Lehmann does not even take under consideration the possibility that every prototype in S is inconsistent with the set of premises A. Hence, from the definition of the inference relation $\mid\!\sim_S$ and the quotation above, we feel free to assume that, in case of mutual inconsistency, the distance between a set of premises and a stereotype can be set to ∞, referring by ∞ to the maximum value of the distance d.

$$d(A, \Delta) = \infty \text{ iff } A \cup \Delta \vdash \bot$$

So, in the case that every stereotype is inconsistent with the premises, the choice of the stereotype will be vacuous. That is, if A is inconsistent with every stereotype in S, $S^A = S$.

Such a model is really limited, since one of the main characteristics of the use of stereotypes is the fact that an item can be related to a stereotype also if its characteristics are not consistent with every property described in the stereotype, and we can derive our suppositions on the basis of the pieces of information codified in the stereotype which are compatible with the premises.

Hence, we should define a notion of distance able to manage also relative distances between a set of premises A and A-inconsistent default sets; in such a way, in the absence of A-consistent stereotypes, we allow the choice of the 'nearest' A-inconsistent default sets. Consequently, we should modify also the definition of the inference relation to one appropriate to the management of possible inconsistences between the premises and the chosen default set. That can be done, obviously, by referring to default-assumption inference relations:

$$A \mid\!\sim_S \beta \text{ iff } A \mid\!\sim_{\Delta^A} \beta$$

For example, knowing that Tweety is a penguin, we can reason about him using the information contained in the stereotype of bird, apart from the information that are known to be inconsistent with being a penguin (flying, nesting in trees…).

Hence, we feel the necessity to move to a notion of semantic distance able to manage also distances between sets of formulae inconsistent between them. In order to obtain this, we move to a notion of semantic distance proposed by Lehmann, Magidor and Schlechta in [32], proposed in order to model belief revision and appropriate for the measure of a sort of 'consistency distance' between formulae.
We define a total function of semantical pseudo-distance d defined over the valuations of our domain U ($d : U \times U \mapsto Z$, where Z is a generic set).
d is a pseudo-distance on U iff

(d1) The set Z is totally ordered by a strict order $<$.

This is the minimal requirement in order to define a notion of pseudo-distance. We also require that d *respects identity*, i.e. Z has a $<$-smallest element 0, and d satisfies the following property

(d2) $d(w, v) = 0$ iff $w = v$.

Pseudo-distance is a generalization of the notion of distance: it does not require that the set Z corresponds to reals, there is no need for a definition of a notion of addition of values, and it does not even require symmetry to hold (i.e. $d(w, v) = d(v, w)$ for every $w, v \in U$). The possibility of failure of such properties is interesting with respect to a 'cognitive' interpretation of the notion of distance, where we want to read such notion of distance in terms of 'informational cost' of the acceptance of different clusters of information with respect to a set of premises.
Given a finite language ℓ, we model the distance between different sets of formulae A and B on the basis of a semantical definition of distance with respect to the set of valuations $[A]$ and $[B]$ satisfying such formulae.
We define the distance between two sets of valuations T and T' as the minimal distance between the valuations in T and T'.

$$d(T,T') = min\{d(w,v)|w \in T, v \in T'\}$$

Again, given a finite set S of default sets $\{\Delta_1, \ldots, \Delta_n\}$, and a set of premises A, we define as S^A the subset of S composed by the default sets resulting as the 'nearest' ones to A.

$$S^A = \{\Delta_i \in S|\ d(A, \Delta_i) \leq d(A, \Delta_j) \text{ for every } \Delta_j \in S)\}$$

Since d is a total function, $S \neq \emptyset$ implies $S^A \neq \emptyset$.
We abandon Lehmann's simplifying assumption that S^A must be composed by a single default set; we take under consideration the possibility that a set of premises could be equally distant from distinct default sets. In such a case, our agent shall reason in a skeptical way with such conflicting default sets.
In the following, we will use as arguments of the function d both sets of formulae A or sets of valuations $[A]$. That is, we state $d(A, \Delta) = d([A], [\Delta])$ for every set of formulae.
If our function d respects identity (satisfies d2), then the distance between a set of premises and a default set mutually consistent is 0, since they share at least a valuation. Hence, the default sets which result consistent with our set of premises have, intuitively, the priority over inconsistent ones.
Having defined a pseudo-distance function d, satisfying (d1), we define $\mid\!\sim_S$ as the skeptical inference relation defined by the choice of the stereotypes with respect to the set of premises A:

$$A \mid\!\sim_S \beta \text{ iff } A \mid\!\sim_\Delta \beta \text{ for every } \Delta \in S^A$$

That is, $C_S(A) = \bigcap\{C_\Delta(A)|\Delta \in S^A\}$, where C_S is the inference operation corresponding to the relation $\mid\!\sim_S$.
It is easy to check that our inference relations satisfies some basic important properties.

Lemma 7.2.2. *Assume a pseudo-distance d and a set of stereotypes S. The inference relation $\mid\!\sim_S$ satisfies REF, LLE, RW, AND.*

Proof.

Assume a set of premises A and a set of stereotypes S. By means of our distance function d, we can identify the set S^A.

From the properties of default-assumption inference relations, we obtain the following properties holding for every $\Delta_i \in S^A$:

$A \mid\!\sim_{\Delta_i} \alpha$ for every $\alpha \in A$ (REF).
$A \mid\!\sim_{\Delta_i} \alpha$ iff $B \mid\!\sim_{\Delta_i} \alpha$, for every set B s.t. $Cl(B) = Cl(A)$ (LLE).
If $A \mid\!\sim_{\Delta_i} \alpha$ and $\vDash \alpha \to \beta$, then $A \mid\!\sim_{\Delta_i} \beta$ (RW).
If $A \mid\!\sim_{\Delta_i} \alpha$ and $A \mid\!\sim_{\Delta_i} \beta$, then $A \mid\!\sim_{\Delta_i} \alpha \wedge \beta$ (AND).

Consequently, it is easy to see that, since $C_S(A)$ is defined by the intersection of every $C_\Delta(A)$ s.t. $\Delta \in S^A$, and the above properties are all of Horn form, then $C_S(A)$ satisfies the above properties as well.

■

Given the validity of REF and RW, our inference relation will satisfy also supraclassicality, and by LLE we have also Left Conjunction ($A \mid\!\sim_S \beta$ iff $\bigwedge\{A\} \mid\!\sim_S \beta$), which allows us to treat only single-premises sequents, without loss of generality.

However, this notion of distance is not sufficient to guarantee cumulativity, since we do not have any coherence constraint that can relate the choice function with respect to possible variations in the premise set. Hence, the default sets chosen with respect to a premise set α will not possibly have any connection with the default sets chosen with respect to a premise set $\alpha \wedge \beta$. For a trivial example, assume we have a set of stereotypes $S = \{\Delta, \Delta'\}$, where $\Delta = \{\neg p, p \to r, p \to t\}$ and $\Delta' = \{\neg p, p \wedge r \to \neg t\}$. Since both Δ and Δ' are p-inconsistent, we have that $d(p, \Delta) \neq 0$ and $d(p, \Delta') \neq 0$. We can state that $d(p, \Delta) < d(p, \Delta')$ and $d(p \wedge r, \Delta') < d(p \wedge r, \Delta)$; this does not

violate neither (d1) nor (d2).
However, in such a case, we obtain $p \mid\!\sim_S r$, $p \mid\!\sim_S t$, since $C_S(p) = C_\Delta(p)$, but, given that $C_S(p \wedge r) = C_{\Delta'}(p \wedge r)$, we have that $p \wedge r \mid\!\sim_S \neg t$, against cautious monotony.

To satisfy cumulativity, it is sufficient to add a new property to the distance d. In order to introduce such a property, recall the definition of a remainder set $\Delta \bot \alpha$ (Definition 2.3.2). Moreover, define $[\Delta \bot \alpha]$ as the set of valuations satisfying an element of the remainder set $\Delta \bot \alpha$:

$$v \in [\Delta \bot \alpha] \text{ iff } v \in [B], \text{ for some } B \in \Delta \bot \alpha$$

In the following property (d3), we impose to the distance function d an intuitive constraint, stating that, given a premise α and a default set Δ, a valuation satisfying both α and a α-maxiconsistent subset of Δ is at least as near to Δ as any other valuation in $[\alpha]$.

(d3) If $w \in [\alpha] \cap [\Delta \bot \neg\alpha]$, and $v \in [\alpha]$, then $d(w, [\Delta]) \leq d(v, [\Delta])$.

It is sufficient also to use a weakened form:

(d3') There is a $w \in [\alpha] \cap [\Delta \bot \neg\alpha]$ s.t. $d(w, [\Delta]) \leq d(v, [\Delta])$ for every $v \in [\alpha]$.

We can prove that if d satisfies $(d1) - (d3')$, then we obtain a cumulative inference relation.

Lemma 7.2.3. *If $[\alpha] \subseteq [\alpha']$, then $d(\alpha', \Delta) \leq d(\alpha, \Delta)$ for every Δ.*

Proof.
$d(\alpha, \Delta) = min\{d(w, v) | w \in [\alpha], v \in [\Delta]\}$. Since $w \in [\alpha]$ implies $w \in [\alpha']$, we have that $min\{d(w, v) | w \in [\alpha'], v \in [\Delta]\} \leq min\{d(w, v) | w \in [\alpha], v \in [\Delta]\}$,

i.e. $d(\alpha', \Delta) \leq d(\alpha, \Delta)$.

■

We need to prove that, if $\alpha \mid\!\sim_S \beta$, then $S^\alpha = S^{\alpha\wedge\beta}$, that is, $d(\alpha, \Delta) \leq d(\alpha, \Delta')$ for every $\Delta, \Delta' \in S$ iff $d(\alpha \wedge \beta, \Delta) \leq d(\alpha \wedge \beta, \Delta')$.

Lemma 7.2.4. *If $\alpha \mid\!\sim_S \beta$ and $\Delta \in S^\alpha$, then $d(\alpha \wedge \beta, \Delta) = d(\alpha, \Delta)$.*

Proof.
Recall that $S^\alpha = \{\Delta \in S \mid d(\alpha, \Delta) \leq d(\alpha, \Delta') \text{ for every } \Delta' \in S\}$.
By (d3'), we have that $d([\alpha], [\Delta]) = d(w, [\Delta])$ for some $w \in [\alpha] \cap [\Delta \bot \neg\alpha]$.
$\alpha \mid\!\sim_S \beta$ implies that $\alpha \mid\!\sim_\Delta \beta$ for every $\Delta \in S^\alpha$, which implies that if $w \in [\alpha] \cap [\Delta \bot \neg\alpha]$, then $w \in [\alpha \wedge \beta]$.
Given that $d([\alpha], [\Delta]) = d(w, [\Delta])$, we have that $d(w, [\Delta]) \leq d(v, [\Delta])$ for every $v \in [\alpha]$. Since $[\alpha \wedge \beta] \subseteq [\alpha]$, we have that $d(w, [\Delta]) \leq d(v, [\Delta])$ for every $v \in [\alpha \wedge \beta]$, that is, $d(\alpha \wedge \beta, \Delta) = d(w, [\Delta]) = d(\alpha, \Delta)$.

■

Lemma 7.2.5. *If $\alpha \mid\!\sim_S \beta$, then $S^\alpha = S^{\alpha\wedge\beta}$.*

Proof.
Assume $\alpha \mid\!\sim_S \beta$.

($\Rightarrow$): If $\Delta \in S^\alpha$, then $\Delta \in S^{\alpha\wedge\beta}$.

If $\Delta \in S^\alpha$, then $d(\alpha, \Delta) \leq d(\alpha, \Delta')$ for every $\Delta' \in S$. By Lemma7.2.3, since $[\alpha \wedge \beta] \subseteq [\alpha]$, we have that $d(\alpha, \Delta) \leq d(\alpha \wedge \beta, \Delta')$ for every $\Delta' \in S$. Since $\Delta \in S^\alpha$, by Lemma7.2.4, we obtain $d(\alpha \wedge \beta, \Delta) \leq d(\alpha \wedge \beta, \Delta')$ for every $\Delta' \in S$, i.e. $\Delta \in S^{\alpha\wedge\beta}$.

($\Leftarrow$): If $\Delta \notin S^\alpha$, then $\Delta \notin S^{\alpha\wedge\beta}$.

If $\Delta \notin S^{\alpha}$, then $d(\alpha, \Delta') < d(\alpha, \Delta)$ for every $\Delta' \in S^{\alpha}$. By Lemma7.2.3, we have that $d(\alpha, \Delta') < d(\alpha \wedge \beta, \Delta)$. Since $\Delta' \in S^{\alpha}$, by Lemma7.2.4, we obtain $d(\alpha \wedge \beta, \Delta') < d(\alpha \wedge \beta, \Delta)$, i.e. $\Delta \notin S^{\alpha \wedge \beta}$.
■

Theorem 7.2.6. *Given a set of stereotypes S and a notion of distance satisfying (d1)-(d3'), the consequence relation $\mid\sim_S$ is cumulative.*

Proof.
We have to show that $\mid\sim_S$ satisfies CM and CT.
CM: assume $\alpha \mid\sim_S \beta$ and $\alpha \mid\sim_S \gamma$, which correspond to saying that $\alpha \mid\sim_\Delta \beta$ and $\alpha \mid\sim_\Delta \gamma$ for every $\Delta \in S^{\alpha}$. Since every default-assumption consequence relation $\mid\sim_\Delta$ satisfies CM, we have $\alpha \wedge \beta \mid\sim_\Delta \gamma$ for every $\Delta \in S^{\alpha}$. Given $\alpha \mid\sim_S \beta$, we have, by Lemma7.2.5, that $S^{\alpha} = S^{\alpha \wedge \beta}$, which implies that $\alpha \wedge \beta \mid\sim_\Delta \gamma$ for every $\Delta \in S^{\alpha \wedge \beta}$, i.e. $\alpha \wedge \beta \mid\sim_S \gamma$.
CT: assume $\alpha \wedge \beta \mid\sim_S \gamma$ and $\alpha \mid\sim_S \beta$. $\alpha \wedge \beta \mid\sim_S \gamma$ means that $\alpha \wedge \beta \mid\sim_\Delta \gamma$ for every $\Delta \in S^{\alpha \wedge \beta}$. $\alpha \mid\sim_S \beta$ implies, again by Lemma7.2.5, that $S^{\alpha} = S^{\alpha \wedge \beta}$. Hence, we have that $\alpha \wedge \beta \mid\sim_\Delta \gamma$ and $\alpha \mid\sim_\Delta \beta$ for every $\Delta \in S^{\alpha}$. Since every such $\mid\sim_\Delta$ satisfies CT, we have $\alpha \mid\sim_\Delta \gamma$ for every $\Delta \in S^{\alpha}$, i.e. $\alpha \mid\sim_S \gamma$.
■

This notion of distance seems appropriate for the classification of an item under a stereotype also in case of lack of mutual consistency. However, with this measure there is a problem if there are more than a stereotype coherent with the premise, since in such a case, if d satisfies identity, we have that the distance between the stereotype and the premise is always 0.
Hence, on one hand we have Lehmann's distance, that can manage the choice of a stereotype in case of consistency with the premise. On the other hand we have our use of semantical pseudo-distance, appropriate for the choice of stereotypes if they are inconsistent with the premise. One possible solution could be the combination of both notions, using Lehmann's distance in order to 'refine' our notion of distance, in case of more stereotype consistent with

the premise. This could result in a lexicographic ordering of a distance d_{lex}, built by giving precedence to the pseudo-distance proposed here, that from now on we indicate with d_1, and, in case of equality, refining with Lehmann's distance, from now on called d_2.

$$d_{lex}(\alpha, \Delta) < d_{lex}(\alpha', \Delta') \Leftrightarrow \begin{cases} d_1(\alpha, \Delta) < d_1(\alpha', \Delta') \\ \textit{or} \\ d_1(\alpha, \Delta) = d_1(\alpha', \Delta') \text{ and } d_2(\alpha, \Delta) < d_2(\alpha', \Delta') \end{cases}$$

Given a set of stereotypes S and a premise α, the previous definition of the set of the chosen stereotypes holds:

$$S^\alpha = \{\Delta_i \in S|\ d_{lex}(\alpha, \Delta_i) \leq d_{lex}(\alpha, \Delta_j) \text{ for every } \Delta_j \in S)\}$$

$\mid\!\sim_S$ is defined as for d_1, referring to default-assumption inference relations:

$$\alpha \mid\!\sim_S \beta \text{ iff } \alpha \mid\!\sim_\Delta \beta \text{ for every } \Delta \in S^\alpha$$

At this point, we can show that also this new definition of distance generates a cumulative inference relation.

From now on, given a premise α, we indicate by S^α the stereotypes chosen on the basis of d_{lex}, $S^\alpha_{d_1}$ the stereotypes chosen on the basis of d_1, and $S^\alpha_{d_2}$ the stereotypes chosen on the basis of d_2. They define, respectively, three inference relations (operations): $\mid\!\sim_S$, $\mid\!\sim_{S,d1}$, and $\mid\!\sim_{S,d2}$ (C_S, $C_{S,d1}$, and $C_{S,d2}$). Consequently, from the lexicographic definition of the d_{lex}-ordering, we have:

$$S^\alpha = \begin{cases} S^\alpha_{d_1} \text{ if } |S^\alpha_{d_1}| = 1 \\ \\ (S^\alpha_{d_1})^\alpha_{d_2} \text{ if } |S^\alpha_{d_1}| > 1 \end{cases}$$

Proposition 7.2.7. *if* $\alpha \mid\!\sim_S \beta$, *then* $S^\alpha = S^{\alpha\wedge\beta}$

Proof.
We have three possibilities.

(1) $|S^{\alpha}_{d_1}| = 1$.
(2) $|S^{\alpha}_{d_1}| > 1$ and $d_1(\alpha, \Delta) > 0$ for every $\Delta \in S^{\alpha}$.
(3) $|S^{\alpha}_{d_1}| > 1$ and $d_1(\alpha, \Delta) = 0$ for every $\Delta \in S^{\alpha}$.

(1): $|S^{\alpha}_{d_1}| = 1$ implies that $S^{\alpha}_{d_1} = S^{\alpha}$ and $C_S(\alpha) = C_{S,d1}(\alpha)$. By Lemma 7.2.5, from $\alpha \mid\!\sim_S \beta$ we obtain $S^{\alpha}_{d_1} = S^{\alpha\wedge\beta}_{d_1}$, that implies $|S^{\alpha\wedge\beta}_{d_1}| = 1$ and $S^{\alpha\wedge\beta} = S^{\alpha\wedge\beta}_{d_1}$, that is, $S^{\alpha\wedge\beta} = S^{\alpha}$.

(2): $d_1(\alpha, \Delta) > 0$ for every $\Delta \in S^{\alpha}$ implies that the default sets in S^{α} are not consistent with the premise α. Therefore, d_2 is not able to choose between them, and we have $(S^{\alpha}_{d_1})^{\alpha}_{d_2} = S^{\alpha}_{d_1}$. Again we have that $S^{\alpha}_{d_1} = S^{\alpha}$ and $C_S(\alpha) = C_{S,d1}(\alpha)$, and we can apply the argument at point (1).

(3): $d_1(\alpha, \Delta) = 0$ for every $\Delta \in S^{\alpha}$ implies that the default sets chosen by means of d_1 are consistent with α, and we can refine the choice by means of d_2.
Since $\alpha \mid\!\sim_S \beta$, we have that some default sets in $S^{\alpha}_{d_1}$ are consistent with $\alpha \wedge \beta$ (surely those in $(S^{\alpha}_{d_1})^{\alpha}_{d_2}$), and, therefore, these are the default sets in S minimally d_1-distant from $\alpha \wedge \beta$. Hence, we have $(S^{\alpha}_{d_1})^{\alpha}_{d_2} \subseteq S^{\alpha\wedge\beta}_{d_1} \subseteq S^{\alpha}_{d_1}$, that is,

$$S^{\alpha} \subseteq S^{\alpha\wedge\beta}_{d_1} \subseteq S^{\alpha}_{d_1}$$

Since every element in $(S^{\alpha}_{d_1})^{\alpha}_{d_2}$ is in $S^{\alpha\wedge\beta}_{d_1}$, we have that $(S^{\alpha\wedge\beta}_{d_1})^{\alpha}_{d_2} = (S^{\alpha}_{d_1})^{\alpha}_{d_2} = S^{\alpha}$. Now we have to use a Theorem from Lehmann. Remember that the choice function determined by d_2 selects a single default set, i.e. $|S^{\alpha}_{d_2}| = 1$ for every S and every α. In particular, we will indicate by $\Delta^{\alpha}_{d_2}$ the only default set in $S^{\alpha}_{d_2}$.

*[30], **Theorem 5.5**: If* $([\alpha] \cap [\Delta^{\alpha}_{d_2}]) \subseteq [\alpha'] \subseteq [\alpha]$, *then* $S^{\alpha'}_{d_2} = S^{\alpha}_{d_2}$

In particular, let α' be $\alpha \wedge \beta$, S be $S^{\alpha\wedge\beta}_{d_1}$, and consequently $(S^{\alpha\wedge\beta}_{d_1})^{\alpha}_{d_2} = \{\Delta^{\alpha}_{d_2}\}$; given that $\alpha \mid\!\sim_S \beta$ and $S^{\alpha} = (S^{\alpha\wedge\beta}_{d_1})^{\alpha}_{d_2}$, we have that $([\alpha] \cap [\Delta^{\alpha}_{d_2}]) \subseteq$

$[\alpha \wedge \beta] \subseteq [\alpha]$, and this, by the theorem above, implies $(S_{d_1}^{\alpha\wedge\beta})_{d_2}^{\alpha\wedge\beta} = (S_{d_1}^{\alpha\wedge\beta})_{d_2}^{\alpha}$. Combining the equations, we have $(S_{d_1}^{\alpha})_{d_2}^{\alpha} = (S_{d_1}^{\alpha\wedge\beta})_{d_2}^{\alpha\wedge\beta}$, that is

$$S^{\alpha} = S^{\alpha\wedge\beta}$$

as desired.

■

Theorem 7.2.8. *Given a set of stereotypes S and a notion of distance d_{lex} (defined lexicographically from d_1 and d_2 as above), the consequence relation $\mid\sim_S$ is cumulative.*

Proof.
Since $C_S(\alpha)$ is obtained by the intersection of default-assumption inference operations, it satisfies REF, LLE, RW, AND (see Lemma 7.2.2).
Cumulativity can be proven using Proposition 7.2.7, restating appropriately the proof of Theorem 7.2.6.

■

So, combining our notion of pseudo-distance, which seems appropriate in cases of inconsistencies between the item and the stereotypes, and Lehmann's distance, which is more appropriate in case of consistent item-stereotype pairs, we have defined a notion of distance that seems appropriate for modeling stereotypical reasoning: it gives back a cumulative inference relation in every case.

Chapter 8

Default updating

Abstract. We present a possible approach to the revision of default information.

In this chapter we shall propose a model for the revision of default-assumption sets. We will start by briefly depicting the main results in the classical theory of belief change, and then we will move to a corresponding characterization for defaults. After a brief presentation of the AGM approach to belief revision, we shall consider the logical constraints appropriate for the characterization of the revision of default sets, i.e. sets of formulae closed under $\triangleright$, and then we shall consider also the possibilities of revision models with respect to default bases, i.e. sets of formulae not necessarily closed under $\triangleright$, representing the defaults explicitly considered by a real agent.

8.1 Theory revision

Assume an agent characterized by a consistent belief base H, i.e. a set of formulae representing its explicit beliefs, and by a tarskian supraclassical closure operation Cn (relation $\vdash$); if the agent acquires a new piece of information α, it is possible for α to be Cn-inconsistent with its previous belief base H ($Cn(H, \alpha) = \bot$). To manage such a situation we can undertake one of two possible attitudes.

On one hand, we can admit that an inconsistent belief set is a realistic and not undesirable possibility, and hence we point to the definition of a reasoning process (i.e. a consequence relation) appropriate for the management of inconsistent information, avoiding the triviality of the *ex falso quod libet* property; such a position refers to the so-called *paraconsistent logics* (see e.g. [46]).

On the other hand, we can maintain logical consistency as an unrejectable requisite for satisfactory databases. If we point to such an approach, we need to define a mechanism in order to keep our database consistent after every informational update. Belief revision is a theory developed to define the logical constraints appropriate for the definition of such a mechanism.

If we add to a belief base H a formula α s.t. α and H are mutually inconsistent, we have to eliminate some of the content of the agent's database in order to restore consistency. In general, in belief revision we assume that the incoming information α has to be accepted, and so it is the previous information H that has to be revised.

The following is a typical simple example:

> "Suppose you have a database that contains, among other things, the following pieces of information (in some form of code):
>
> | | |
> |---|---|
> | The bird caught in the trap is a swan. | A |
> | The bird caught in the trap comes from Sweden | B' |
> | Sweden is part of Europe | $B' \to B$ |
> | All European swans are white | $A \wedge B \to C$ |
>
> If your database is coupled with a program that can compute logical inferences in the given code, the following fact is derivable:
>
> The bird caught in the trap is white $\quad C$

Now suppose that, *as a matter of fact*, the bird caught in the trap turns out to be black. This means that you want to add the fact $\neg C$, i.e. the negation of C, to your database. But then the database becomes inconsistent. If you want to keep the database consistent - which is normally a sound methodology - you need to *revise* it. This means that some of the beliefs in the original database must be *retracted*." ([20], p.36)

The problem of belief revision corresponds to the characterization of methodologies for modifying the database in order to keep it logically consistent. The main preoccupation for a logician is the definition of rationality postulates working as ideal constraints for the behaviour of a real agent. Such constraints are both of logical and economical kind.
On the logical side, our main interest is the *logical consistency* of the beliefs of the agent. Since we are working on the level of ideal agents, we assume that we operate on *deductively closed* sets of formulae, that is, if a database entails a sentence α, then we should consider α as included in the database. Consistency and closure have to be considered as necessary properties for the characterization of rationality and coherence in the epistemic states of an ideal agent.
We refer to deductively closed sets of formulae as *belief sets* K ($K = Cn(K)$), opposed to *belief bases* H, which are finite sets of formulae representing the explicit content of a real database.
Beyond logical principles such as consistency and closure, we have desiderata of economical kind: in particular, since we maintain information as a precious resource, we adopt the *conservativity principle*, or *maxim of minimum mutilation*: in every belief change, the loss of information should be kept minimal. We should abandon as little information as possible in order to keep the belief set consistent.
As suggested by Rott (see [53],[54]), on the basis of such desiderata we can distinguish three senses of coherence that a theory of belief revision should

satisfy. Fist of all, there is a static sense of coherence, the *synchronic* or *inferential coherence*, tied to the notion of consistency and closure and imposing constraints on the isolated belief sets: the agent's beliefs should be 'well balanced', i.e. we should avoid not closed and not consistent epistemic states.
Then, we can identify a relational, *diachronic*, kind of coherence, with respect to the move from an epistemic state to a new one, which has to account for the economical constraints imposed by the conservativity principle.
In the end, we can also identify a third sense of coherence, pointing toward the inclusion of every potential update of an epistemic state into a comprehensive rational framework: while diachronic coherence is aimed to the comparison of the prior and the posterior belief sets in the case of a revision by a particular input, *dispositional coherence* compares the potential revisions of a belief set with respect to different but logically related formulas: for example, given a belief set K, it is natural to relate the behaviour if the revision of K with respect to a formula $\alpha \wedge \beta$ to its revision with respect to the single formulas α and β.
Moreover, we have a basic principle connected to the formalism we chose in defining the epistemic state of an agent.

Definition 8.1.1. *Principle of categorial matching ([20], p.37):*
The representation of a belief state after a belief change has taken place should be the same format as the representation of the belief state before the change.

The principle of categorial matching simply states that, starting with an epistemic state characterized by a particular formal structure, the output of every belief change operation should be an epistemic state with the same kind of formal structure: for example if we start from a closed belief set, we have to obtain a new closed belief set, and, analogously, if we start with a finite belief base, every change has to result in a finite belief base. Such a principle seems quite intuitive and obvious, and is considered necessary in particular for the iteration of belief-change operations. However, it is often not very simple to be satisfied, especially if we work with epistemic structures involving special

expectation relations between formulas, as the entrenchment relation we are going to see below, which has to be redefined after every change.
The connection between nonmonotonic logics and belief revision theory is very strong (see [37]). In particular, the meet contraction presented below have a clear connection with default-assumption models, since they resemble the same construction by means of maxiconsistent sets.

On the theoretical side, we can define two main approaches to the problem of belief revision: the *coherentist* approach and the *foundationalist* one. Notwithstanding that they both generally concord with the coherence constraints above, they are distinguished with respect to the characterization of the epistemic state of the agent.
On one hand, coherentists identify the epistemic state of the agent with its belief set, i.e. with a closed theory; such a theory is confronted with new information in an holistic way, and every formula in the theory has the same epistemological status.
On the other hand, foundationalists emphasise the role of the belief base of the agent, i.e. the finite set of formulas the agent explicitly 'has in mind', the real database of the agent. Notwithstanding the fact that the rationality constraints of revision have to be formulated with respect to ideal conditions of coherence (closure and consistency), the operations of belief revision have to be modeled with respect to the changes of the belief base; given an agent with a belief base H and a new piece of information α, the formulae in H, the *basic formulae*, front the new information with a different status with respect to the formulae in $Cn(H)$, called *derivative formulae*, which presence in the belief set of the agent depends only on their derivability from some formulas in H.
In order to define the ideal rational desiderata of belief revision, we have to work at a high level of idealization; this means that we have to refer to logical closed sets of formulae as the representation of the epistemic states of an agent. Referring to the belief sets instead of the belief bases in order to define

the rationality postulates of revision operations, we aim to a representation of the behaviour of an ideal reasoner when it is forced to reorganize its beliefs in front of new incoming information. The coherentist approach points toward the definition of an ideal prescriptive model defining what rational agents should do, abstracting from the restrictions of their real reasoning capabilities.

Hence, we could see the coherentist approach as positioned on a higher level of idealization, abstracting from a distinction between explicit and implicit information in an agent's epistemic state. It's from such a perspective that Alchourròn, Gärdenfors and Makinson have proposed in [1] some rationality postulates for belief revision, known as the AGM approach, that have been recognized as the standard view in the field.

In such a characterization of belief change operations, we assume that the possible change operations on a belief set consist in the addition or the elimination of pieces of information. With respect to the way we undertake such additions and eliminations, we can identify three kinds of operations:

1. *Expansion*: given a belief set K, we want to add a formula α to K. The addition of α to K is done without any consistency constraint.

2. *Contraction*: given a belief set K, we want to eliminate a formula α from K.

3. *Revision*: given a belief set K, we want to add a formula α to K, but we want the belief set to remain consistent.

The AGM theory presents a series of rationality postulates for these three operations, assuming we are working with closed belief sets, without considering any relation between formulae besides a closure operator; we assume as our closure operation a supraclassical monotonic consequence operator Cn (and the corresponding relation $\vdash$).

We are going to briefly present the main results of the classical theory of belief revision in order to be able to introduce default revision. Notwithstanding, assuming the characterization of nonmonotonic systems given in the previous chapters, as generated by a knowledge set and a default set, we can deem the following results of theory revision as appropriate tools for the revision of the knowledge set $K_{\mid\sim}$ associated to a nonmonotonic inference relation $\mid\sim$.

8.1.1 Expansion

Given that in the operation of expansion we do not assume any consistency constraint, such an operation is usually the most simple one to be defined, and generally it results in quite a trivial operation. Obviously, it is possible for the result of an operation of expansion to be an inconsistent belief set, written $K_{\perp}$, i.e. a deductively closed set corresponding to the entire language ℓ ($K_{\perp} = \ell$).

We indicate the operation of expansion by means of a function $+$, that is, $+ : \mathcal{T} \times \ell \mapsto \mathcal{T}$, where $\mathcal{T} \subseteq \wp(\ell)$ is the set of the Cn-theories in ℓ. Given a belief set K, K^{+}_{α} is the belief set obtained by the addition of α to K.

There are six rationality postulates that such an operation should satisfy.

$(K+1)$ K^{+}_{α} is a belief set.

$(K+1)$ imposes the respect of the principle of categorial matching, forcing the result of expansion to be again a belief set, i.e. a set of formulae closed under Cn.

$(K+2)$ $\alpha \in K^{+}_{\alpha}$.

$(K+2)$ is the *success postulate*: since expansion is finalized to the addition of α in our epistemic state, α must be a member of the new belief set.

$(K+3)$ $K \subseteq K^{+}_{\alpha}$.

$(K+3)$ is a postulate referring to informational economy: since we do not impose any consistency constraint, we do not have any reason to abandon information: it guarantees that no previous information is affected by the expansion of K with α.

$$(K+4) \quad \text{If } \alpha \in K, \text{ then } K^+_\alpha = K.$$

$(K+4)$ captures a degenerate case: if the information to be added is already in the agent's belief set, nothing changes.

$$(K+5) \quad \text{If } K \subseteq K', \text{ then } K^+_\alpha \subseteq K'^+_\alpha.$$

$(K+5)$ is known as the *monotonicity postulate.* If $K \subseteq K'$, then the addition of α to K does not add anything that is not included also in K'^+_α.

$(K+6)$ For all belief sets K and all sentences α, K^+_α is the smallest belief set that satisfies $(K+1)-(K+5)$.

$(K+6)$ is known as the *minimality postulate.* It imposes that the new belief set does not contain any extra information with respect to the addition of α to K.
It is easy to show that the operation satisfying these postulates simply corresponds to the closure of $K \cup \{\alpha\}$ under Cn.

Theorem 8.1.1. *([17], Theorem 3.1)*
The expansion function + satisfies (K+1)-(K+6) iff $K^+_\alpha = Cn(K \cup \{\alpha\})$.

Modeling expansion by means of this definition, the iteration of the operation is very simple to analyze given the properties of a monotonic supraclassical operator Cn. For example, the operation of expansion is commutative.

$$(K^+_\alpha)^+_\beta = (K^+_\beta)^+_\alpha$$

Moreover, the successive expansion of K by means of two formulae α and

β corresponds to a single expansion of K by means of $\alpha \wedge \beta$.

$$(K^{+}_{\alpha})^{+}_{\beta} = K^{+}_{\alpha\wedge\beta}$$

8.1.2 Contraction

The contraction of a belief set K with respect to a formula α consists in the elimination of α from K without adding any new piece of information. Given the closure assumption, i.e. that the epistemic state of the agent is represented by means of a closed belief set, such an operation is not as trivial as expansion: we want α to be no more Cn-derivable in our epistemic state, and so we have to eliminate from K every set of formulae H s.t. $H \subseteq K$ and $H \vdash \alpha$. In this operation, the economic principle of conservativity plays a fundamental role, since we want to avoid the derivation of α but at the same time eliminating from K as few pieces of information as possible. However, as argued by Rott in [52], there has often been a misunderstanding in the role of the AGM postulates with respect to conservativity: notwithstanding that most of the commentators have indicated the AGM theory as an approach centered on the principle of minimum mutilation, AGM logical constraints do not fully account for the respect of such economical constraints, or better, they fully account for such constraints just for trivial cases; at most, we could say that they determine the compatibility of the contraction operation with the principle of minimum mutilation. As we are going to see below, there are constructions, such as full meet contraction, that satisfy the AGM postulates without respecting conservativity. For now, economical constraints remain mainly an extra-logical problem, not to be valuated by means of logical constraints but with respect to the particular constructions of the belief change operators.

As we are going to see in the next section, the main theoretical justification of the operation of contraction is with respect to its role in the definition of the revision operation, and it has been debated if we can recognize 'pure' cases of

contraction, i.e. if there are cases in which an agent simply eliminates a belief in α from its belief set, without adding the belief in $\neg\alpha$. In general, however, we recognize the possibility of the move from an epistemic state in which α is believed to an epistemic state in which we simply suspend our judgment about α, since we could have received information forcing us to raise doubts about the truth value of α, without driving us towards a commitment with respect to the truth of $\neg\alpha$. Moreover, we can imagine the use of contraction in counterfactual reasoning, imaging situations in which we do not assume the truth of α, everything else remaining the same.
We indicate by $\dot{-}$ a contraction function over belief sets ($\dot{-} : \mathcal{T} \times \ell \mapsto \mathcal{T}$): given a belief set K, $K^{\dot{-}}_{\alpha}$ is the belief set obtained by contracting K with respect to α.
Again, we have six *basic AGM postulates* for rational contraction:

$(K\dot{-}1)$ $K^{\dot{-}}_{\alpha}$ is a belief set *Closure*

This first principle imposes the respect of the principle of categorial matching, stating that the resulting epistemic state is represented by a Cn-theory.

$(K\dot{-}2)$ $K^{\dot{-}}_{\alpha} \subseteq K$ *Inclusion*

Contraction simply eliminates information from K, so every sentence in $K^{\dot{-}}_{\alpha}$ must already be in K.

$(K\dot{-}3)$ If $\alpha \notin K$, then $K^{\dot{-}}_{\alpha} = K$ *Vacuity*

This postulate treats the trivial case: if the sentence to be eliminated is not in K, then nothing needs to be changed, and the result of the operation is the starting belief set.

($K\dot{-}4$) If $\not\vdash \alpha$, then $\alpha \notin K^{\dot{-}}_{\alpha}$ *Success*

This is the *success postulate* for contraction: the set resulting from a contraction must not contain the contracted sentence. The claim 'If $\not\vdash \alpha$' indicates that we assume the underlying logic Cn as fixed, unchangeable by means of the belief-contraction operation, and hence we cannot eliminate a Cn-tautology from a Cn-closed belief set.

($K\dot{-}5$) If $\alpha \in K$, then $K \subseteq (K^{\dot{-}}_{\alpha})^{+}_{\alpha}$ *Recovery*

($K\dot{-}5$) relates contraction with expansion, and it states that putting back a previously contracted sentence should not result in any loss of information. This property is the most controversial principle of the AGM model, and its satisfaction is not considered as necessary as the fulfillment of the other basic postulates. Furthermore, as we shall see in the next section, its failure does not affect the desired properties of the revision operator. A contraction operation satisfying every basic postulates apart from ($K\dot{-}5$) is called a *withdrawal* (see [36]).

($K\dot{-}6$) If $\vdash \alpha \leftrightarrow \beta$, then $K^{\dot{-}}_{\alpha} = K^{\dot{-}}_{\beta}$ *Extensionality*

This last basic postulate states that contraction has to be a *syntax-independent* operation, i.e. that the contraction of logically equivalent formulas has to give the same result.

Besides the six basic postulates, there are two more postulates proposed for contraction, finalized to the management of the contraction of the conjunctions of formulae.
The management of the giving up of conjunctions represents the discriminating task for the contraction operation with respect to economical considerations: if we want to eliminate $\alpha \wedge \beta$ from our belief set, it is sufficient to

eliminate just one between α and β. If we aim at satisfying the principle of minimum mutilation, we should eliminate just one of the two, but in that case we would need a way to single out the 'weakest' of the two beliefs. Even if we are provided with such a way, α and β could be equally 'strong' beliefs: in such a case, we should abandon both of them, and it would be debatable if we should keep or not the belief in $\alpha \vee \beta$, in order to minimize the loss of information.
The postulates $(K\dot{-}7)$ and $(K\dot{-}8)$ are both of economical kind, and they are aimed to state the minimal rationality desiderata in order to relate the contraction of a conjunction with the contractions of its component formulas.

$$(K\dot{-}7) \quad K^{\dot{-}}_{\alpha} \cap K^{\dot{-}}_{\beta} \subseteq K^{\dot{-}}_{\alpha\wedge\beta}$$

$(K\dot{-}7)$ states that, since both the contractions of K with respect to α and with respect to β eliminate $\alpha \wedge \beta$ from the epistemic state of the agent, the beliefs that are both in $K^{\dot{-}}_{\alpha}$ and in $K^{\dot{-}}_{\beta}$ are also in $K^{\dot{-}}_{\alpha\wedge\beta}$, otherwise we would have the certainty that we have eliminated too much information from K in order to contract with respect to $\alpha \wedge \beta$.

$$(K\dot{-}8) \quad \text{If } \alpha \notin K^{\dot{-}}_{\alpha\wedge\beta}, \text{ then } K^{\dot{-}}_{\alpha\wedge\beta} \subseteq K^{\dot{-}}_{\alpha}$$

If $\alpha \notin K^{\dot{-}}_{\alpha\wedge\beta}$, then this means that every piece of information in K apt to derive α has been eliminated from $K^{\dot{-}}_{\alpha\wedge\beta}$. Hence we assume that the formulas in $K^{\dot{-}}_{\alpha\wedge\beta}$ are all also in $K^{\dot{-}}_{\alpha}$, otherwise we would have the certainty that our contraction operation eliminates too much information from K in order to contract with respect to α.
From $(K\dot{-}7)$, $(K\dot{-}8)$ and the basic postulates, we obtain the following principle (see [1], Observation 6.5), stating that the contraction with respect to a conjunction $\alpha \wedge \beta$ corresponds to the contraction with respect to one of its two conjuncts, or to their intersection:

Either $K^{\dot{-}}_{\alpha\wedge\beta} = K^{\dot{-}}_{\alpha}$ or $K^{\dot{-}}_{\alpha\wedge\beta} = K^{\dot{-}}_{\beta}$ or $K^{\dot{-}}_{\alpha\wedge\beta} = K^{\dot{-}}_{\alpha} \cap K^{\dot{-}}_{\beta}$

Such a rule is compatible with the economical considerations depicted above: to contract K with respect to a conjunction, it is sufficient to contract it with respect to one of the conjuncts, if we are able to identify the less 'important', otherwise we have to eliminate both of them.

8.1.3 Revision

The revision operation is the core of the belief revision theory. By means of such an operation, we point to the addition of a piece of information to the epistemic state of an agent, avoiding the formation of inconsistent belief sets. Hence, contrary to simple expansion, such an operation could involve also the elimination of previously held beliefs for consistency sake.
We indicate by $*$ a revision function over belief sets ($* : \mathcal{T} \times \ell \mapsto \mathcal{T}$): given a belief set K, K^*_α is the belief set obtained from the revision of K in order to introduce α.
Again, we have six *basic AGM postulates* for rational revision:

$(K * 1)$ K^*_α is a belief set *Closure*

As for the previous operations, revision has to give back a Cn-closed theory in order to respect the principle of categorical matching.

$(K * 2)$ $\alpha \in K^*_\alpha$ *Success*

The *success postulate* for revision, analogous to that for expansion.

$(K * 3)$ $K^*_\alpha \subseteq K^+_\alpha$ *Expansion*

$(K * 4)$ If $K \nvdash \neg\alpha$, then $K^+_\alpha \subseteq K^*_\alpha$ *Preservation*

$(K*3)$ and $(K*4)$ define the relation between revision and expansion. If the sentence α is consistent with K, revision amounts to expansion, since there is not any necessity to eliminate any piece of information contained in K. The only case in which the revision operation has to differ from the expansion operation is if K and α are inconsistent, i.e. $K^{+}_{\alpha} = K_{\perp}$.

$(K*5)$ $\quad K^{*}_{\alpha} = K_{\perp}$ if and only if $\vdash \neg\alpha$ $\quad$ *Consistency*

The preservation of consistency is always satisfied, except in the case where α is itself a Cn-contradictory formula.

$(K*6)$ $\quad$ If $\vdash \alpha \leftrightarrow \beta$, then $K^{*}_{\alpha} = K^{*}_{\beta}$ $\quad$ *Extensionality*

Also revision has to be syntax-independent.
Beyond basic postulates, there are less elementary conditions apt to the management of iterated operations of revision. In particular, $(K*7)$ and $(K*8)$ are obtained generalizing $(K*3)$ and $(K*4)$ in order to constraint the behaviour of iterated belief revision. Again, the regulating idea is that, whenever possible, the revision function should behave as the expansion function.

$(K*7)$ $\quad K^{*}_{\alpha\wedge\beta} \subseteq (K^{*}_{\alpha})^{+}_{\beta}$.

$(K*8)$ $\quad$ If $\neg\beta \notin K^{*}_{\alpha}$, then $(K^{*}_{\alpha})^{+}_{\beta} \subseteq K^{*}_{\alpha\wedge\beta}$.

Following the distinctions made by Rott about the coherence of a revision operator, we can say that $(K*1)$ and $(K*5)$ embody the notion of inferential coherence, stating that the epistemic states of an agent are characterized by means of coherent, deductively closed sets of formulas; $(K*3)$ and $(K*4)$ are principles of diachronic coherence, constraining the move from a belief set to a revised one in a way compatible with the satisfaction of the principle of minimal change; finally, $(K*7)$ and $(K*8)$ formalize dispositional coherence, constraining the behaviour of the revision operator with respect to different potential inputs, in particular relating the addition of a conjunction to the addition of its conjuncts.

The revision function can be modeled referring to the expansion and contraction functions by means of the *Levi identity*:

$$K^*_\alpha = (K^{\dot{-}}_{\neg\alpha})^+_\alpha$$

The Levi identity proposes a very intuitive reduction of revision to contraction and expansion. Such a formulation states that to revise K with respect to α, we have to first contract K with respect to $\neg\alpha$ in order to guarantee the consistency with respect to α, and then to add α. Note that contraction and expansion have to be performed in the order indicated above, otherwise there is no guarantee for the preservation of as much information of K as possible: if we state $K^*_\alpha = (K^+_\alpha)^{\dot{-}}_{\neg\alpha}$, and α is inconsistent with respect to K, we obtain $\ell^{\dot{-}}_{\neg\alpha}$, which is a contraction operation that has no longer a link with the original set K.
The behaviour of the Levi identity is very intuitive, given a well-behaved contraction function: if α and K are consistent, i.e. $K \nvdash \neg\alpha$, then $(K^{\dot{-}}_{\neg\alpha}) = K$, and consequently $K^*_\alpha = K^+_\alpha$; otherwise, if α and K are inconsistent, the contraction by $\neg\alpha$ guarantees that we modify K in the right way in order to introduce α consistently.
Moreover, assuming the Levi identity construction, Gärdenfors has proved two theorems relating directly the behaviour of a revision function $*$ to that of the underlying contraction function $\dot{-}$.

Theorem 8.1.2. *([17], Theorem 3.2)*
If the contraction function $\dot{-}$ *satisfies* $(K\dot{-}1)$ - $(K\dot{-}4)$ *and* $(K\dot{-}6)$ *and the expansions satisfy* $(K+1)$-$(K+6)$*, then the revision function* $*$ *obtained by means of the Levi Identity satisfies* $(K*1)$ - $(K*6)$.

Theorem 8.1.3. *([17], Theorem 3.3)*
Suppose that the assumptions of Theorem 8.1.2 are fulfilled. Then (a) if $(K\dot{-}7)$ *is satisfied,* $(K*7)$ *is satisfied for the defined revision function, and (b) if* $(K\dot{-}8)$ *is satisfied,* $(K*8)$ *is satisfied for the defined revision function.*

Such results stress the appropriateness of the Levi Identity construction in order to define a revision function. Note that the satisfaction of the postulate $(K\dot{-}5)$ is not necessary for the revision function to be well-behaving, so we can also use withdrawal operations to define good revision functions. Given the general triviality of the expansion function, the focus of the AGM approach moves toward the characterization of satisfying contraction operations.

8.1.4 Meet contractions and entrenchment relations

There are different ways proposed in order to define a contraction operation appropriate for a specific belief set. Two (strongly connected) methods are the most intuitive and satisfying: meet contractions and entrenchment relations.

Meet contractions

By the principle of conservativity, a contraction operation should be minimal, that is, in order to contract a belief set K with respect to a formula α, we should look for the largest subset of K not implying α. Such subsets of K are contained in the *remainder sets* of our belief set (see Definition 2.3.2): given a set of formulae A and a formula α, the *remainder set of A by α* is the set containing all the 'biggest' subsets of A not implying α.
As was said in Chapter 2, the link between the elements of the remainder sets and the maxiconsistent sets of Definition 2.3.1, which are at the basis of the default-assumption approach, is immediate: a set B is in the remainder set of

A with respect to α if and only if it is a $\neg\alpha$-maxiconsistent subset of A. The most popular approach to the formalization of contraction works with such remainder sets: given a belief set K and a formula α to be eliminated, we can define a contracted belief set K' on the basis of the information contained in the elements of the remainder set $K\bot\alpha$. This can be done by means of different procedures, following a distinction analogous to the one between skeptical and choice approaches to nonmonotonic inference presented in Chapter 2.
It is easy to show that the elements of the remainder set of a deductive closed set are themselves deductively closed. That means that the remainder set of a belief set A contains only other belief sets, i.e. every element of $A\bot\alpha$ can satisfy the principle of categorial matching, corresponding to the postulate $(K\dot{-}1)$.

Proposition 8.1.4. *([21], Observation 1.48) If A is logically closed, and $B \in A\bot\alpha$, then B is logically closed.*

It is also immediate from Definition 2.3.2 that every element of $A\bot\alpha$, if considered as the set resulting from the contraction of A with respect to α, satisfies $(K\dot{-}2)$ - $(K\dot{-}4)$ and $(K\dot{-}6)$, and consequently it is a plausible candidate for the contracted belief set.
Hence, if $A\bot\alpha$ contains only a single element, such a set will be the natural candidate for the belief set resulting from the contraction of A with respect to α. The possibility of different options in the definition of the contraction function arise when $A\bot\alpha$ contains more than one element. In such a case we have two limiting cases: *full meet contraction* and *maxichoice contraction*.
The former is a radical form of contraction: we consider every element of the remainder set and we define the contracted belief set by intersecting *all* of them.

$$K^{\dot{-}}_{\alpha} = \begin{cases} \bigcap(K\bot\alpha) & \textit{when } \nvdash \alpha \\ K & \textit{otherwise} \end{cases}$$

At the extreme opposite case, we have the maxichoice approach, in which, given a theory K and a formula to be eliminated α, we identify the contracted belief set $K \dot{-} \alpha$ with a single element of the remainder set $K \bot \alpha$.

$$K_{\alpha}^{\dot{-}} = \begin{cases} K' \in (K \bot \alpha) & when \nvdash \alpha \\ K & otherwise \end{cases}$$

Obviously maxichoice contraction is the most faithful operation with respect to the conservativity principle, since, choosing a single element of the remainder set, we eliminate only a minimal set of formulae from our belief set.

Both of these constructions satisfy the basic postulates $(K\dot{-}1) - (K\dot{-}6)$, but their behaviour often sounds counterintuitive (see [21], Section 2.4). In particular, if we work with closed belief sets, full meet contraction manifests a trivial behaviour.

Lemma 8.1.5. *([17], Lemma 4.9)*
Assume a belief set K. If $K_{\alpha}^{\dot{-}}$ is defined by a full meet contraction and $\alpha \in K$, then $\beta \in K_{\alpha}^{\dot{-}}$ iff $\beta \in K$ and $\neg\alpha \vdash \beta$.

$$If\ \alpha \in K,\ then\ K \dot{-} \alpha = K \cap Cn(\{\neg\alpha\})$$

If we contract α from K in this way, we are left with the propositions of K that are also logical consequences of $\neg\alpha$. Every piece of information that is consistent with $\neg\alpha$, but is not implied by such formula, is automatically eliminated from the contracted belief set. Such triviality in the contraction function is reflected by the revision function generated by means of the Levi Identity.

Corollary 8.1.6. *([17], Corollary 4.10)*
Assume a belief set K. If a revision function $*$ is defined from a full meet contraction $\dot{-}$ by means of the Levi Identity, then, for any α s.t. $\neg\alpha \in K$, K_{α}^{*} will contain only α together with its logical consequences.

$$If\ \neg\alpha \in K,\ then\ K * \alpha = Cn(\{\alpha\})$$

Hence, revising closed theories, the addition of a piece of information inconsistent with the original belief set results in the complete loss of the

information contained in the original belief state K of the agent.
Note that the full-meet approach is connected with such anomalies only in case we are working with belief sets. If, otherwise, we are working with finite belief bases, both full-meet and maxichoice approaches behave in a intuitive way, simply representing respectively the strongest and the weakest form of contraction.
Instead, working with belief sets, the most satisfying contraction operations can be defined taking under consideration only *some* of the elements of the remainder set. *Partial meet contraction* can be modeled by means of a *selection function* S, in order to identify which elements of the remainder set have to be used in the definition of the contracted belief set.

$$K_\alpha^{\dot{-}} = \begin{cases} \bigcap S(K \bot \alpha) & \textit{when } \nvdash \alpha \\ K & \textit{otherwise} \end{cases}$$

Obviously, full meet and maxichoice approaches are special cases of partial meet contraction, defining respectively $S(K\bot\alpha) = (K\bot\alpha)$ and $S(K\bot\alpha) = \{K'\}$ for some $K' \in K\bot\alpha$. The satisfaction of the basic postulates for contraction is characterized by the generation of the contraction function by means of the partial meet approach.

Theorem 8.1.7. *([17], Theorem 4.13)*
For every belief set K, $\dot{-}$ corresponds to a partial meet contraction function iff $\dot{-}$ satisfies $K\dot{-}1$ - $K\dot{-}6$ for contraction over K.

The characterization of the choice function S determines the behaviour of the generated contraction function.

Entrenchment relations

An alternative, but strictly connected, way for generating satisfying contraction operations relies upon the use of *entrenchment relations*.
When we have to give up a belief α, we have to guarantee on one hand that α is no more derivable from our set of beliefs, and on the other hand that we

have eliminated just the beliefs necessary in order to avoid the derivability of α. Such requirement implies that often we have to undertake a choice between different formulas. For example, if we have to contract a belief set K with respect to a formula $\alpha \wedge \beta$, it is sufficient to eliminate just one between α and β in order to avoid the derivability of $\alpha \wedge \beta$ from our belief set.

If we are in a situation such that we can choose which formula to give up in a set, we will prefer to give up the formula that has as little explanatory power and overall information as possible. In order to identify such a formula, we can define a binary relation $<_e$, an entrenchment relation (see [17], Chapter 4) determining an ordering between formulas; 'β is more entrenched than α' ($\alpha <_e \beta$) should be interpreted as stating that β is more useful in enquiry or deliberation, or has more "epistemic value" than α; practically, we are more inclined to give up α than β. Hence the notion of entrenchment can be defined by means of the following correspondence with a contraction function:

$$\alpha <_e \beta \text{ iff } \alpha \notin K^{\dot{-}}_{\alpha\wedge\beta} \; and \; \beta \in K^{\dot{-}}_{\alpha\wedge\beta}$$

That is, if we have to withdrawal the formula $\alpha \wedge \beta$, we prefer to give up α from our belief set. As can be seen from this statement, the entrenchment relation is relative to a belief set K, and every modification occasioned to our belief set causes also a modification in the entrenchment relation between the formulae of our language.

We are interested in the definition of an entrenchment relation between formulas starting from an epistemic state, and, furthermore, the definition of a contraction function from such an entrenchment relation.

Exploiting the correspondence between nonmonotonic consequence and belief revision, Bochmann has proposed an intuitive association (*ie-mapping*) between entrenchment relations $<_e$ and preferential inference relations $\mid\!\sim$ ([5], p.143):

(IE) $\alpha \mid\!\sim \beta$ iff $\alpha \rightarrow \neg\beta <_e \alpha \rightarrow \beta$

(EI) $\alpha <_e \beta$ iff $\neg(\alpha \wedge \beta) \mid\!\sim \beta$

If our agent is defined by means of an inference relation $\mid\!\sim$, we can use the EI-transformation in order to define the correlated entrenchment relation. Given an entrenchment relation as $<_e$, there are different ways for generating a correlated contraction function. Gärdenfors has proposed a satisfying definition of a contraction operation $\dot{-}_G$ based on entrenchment relations:

$$\beta \in H_{\alpha}^{\dot{-}_G} \text{ iff } \beta \in H \text{ and either } \alpha < (\alpha \vee \beta) \text{ or } \alpha \in Cn(\emptyset)$$

Another good proposal is from Rott:

$$\beta \in H_{\alpha}^{\dot{-}_R} \text{ iff } \beta \in H \text{ and either } \alpha < \beta \text{ or } \alpha \in Cn(\emptyset)$$

In general, both of these proposals are considered as interesting limiting cases of contraction, since Gärdenfors' proposal usually generates a too weak contraction operation, while Rott's contraction is too radical. Lindström and Rabinowicz ([35]) have proposed that an intuitive theory-contraction function $\dot{-}$ should be positioned somewhere between $\dot{-}_R$ and $\dot{-}_G$. That is, given a belief set K, we should have for every α:

$$K_{\alpha}^{\dot{-}_R} \subseteq K_{\alpha}^{\dot{-}} \subseteq K_{\alpha}^{\dot{-}_G}$$

8.2 Base revision

If we refer to belief revision from a foundationalist perspective, we point to the definition of the revision of a finite set of beliefs H instead of a closed theory K, notwithstanding that consistency constraints have to be stated with respect of the closure of the base $Cn(H)$.
A base-revision operation can still be defined by means of the Levi identity, referring to base-expansion and base-contraction operations. The definition of the base expansion operation is generally trivial, since set-theoretic union operation is intuitively satisfying. Given a belief base H, and referring by $+_b$ to a base-expansion operation, we can simply state

$$H_{\alpha}^{+_b} = H \cup \{\alpha\}$$

Given that $Cn(Cn(H) \cup \{\alpha\}) = Cn(H \cup \{\alpha\})$, this operation corresponds to the expansion operation $+$ defined on the theory level, that is:

$$Cn(H_{\alpha}^{+_b}) = (Cn(H))_{\alpha}^{+}$$

More effort is needed in order to model contraction on the base level, since the methods identified at the theory level result not appropriate.
In particular, the main efforts have been made in order to define at the base level an appropriate ordering between formulae corresponding to the entrenchment one. Both Rott ([51]) and Bochmann ([6]) have proposed an analogous kind of relations, called by Bochmann *dependence relations*.
Following Rott, in contracting a belief base H by a formula α (obtaining a new base $H' \subseteq H$), we want to eliminate from H only the basic formulae that are essential for the presence of α in $Cn(H)$. Moreover, if in the contraction we eliminate a formula β from H, we would like to eliminate also every formula finding its own justification only in our previous belief in β. This demand is expressed by the following filtering condition.

Definition 8.2.1 (Simple Filtering [51])**.** *The contraction of one's beliefs by ϕ should not contain any sentences that were believed "just because" ϕ was believed.*

In this statement the sensible part becomes the definition we give to "just because", which, in a foundationalist perspective, can be characterized in the following way.

Definition 8.2.2 (Derivative Just Because [51])**.** *A sentence ψ is in $K = Cn(H)$* just because *ϕ is in H if and only if ϕ is not in $Cn(\emptyset)$, and ψ is in K but not in $Cn(G)$, for all $G \subseteq H$ such that ϕ is not in $Cn(G)$.*

This definition means that, starting from a belief set K, it is not possible for the agent to believe ψ without believing also ϕ.
On the basis of such a notion, Rott defines a notion of entrenchment appropriate for a foundationalist approach.

Definition 8.2.3 (Basic entrenchment $\leq_b$ [51]). *Given a belief base H and a (monotonic supraclassical) closure operation Cn, ϕ is more entrenched than ψ ($\psi \leq_b \phi$) iff ψ is not a logical truth ($\psi \notin Cn(\emptyset)$) and for every set $G \subseteq H$ that does not imply ϕ ($\phi \notin Cn(G)$) there is a set $G' \subseteq H$ that does not imply ψ ($\psi \notin Cn(G')$) and $G \subseteq G'$.*

Given that the closure operation Cn satisfies monotonicity, it is easy to see that the definition above corresponds exactly to the notion of *dependence relation* suggested by Bochmann in [5] and [6].

Definition 8.2.4 (Dependence relation $\leq_d$). *Given a belief base H and a (monotonic supraclassical) closure operation Cn, the belief in ψ depends upon the belief in ϕ ($\psi \leq_d \phi$) iff ψ is not a logical truth ($\psi \notin Cn(\emptyset)$) and, for every set $G \subseteq H$, if $\psi \in Cn(G)$, then $\phi \in Cn(G)$.*

Proposition 8.2.1. *Given a belief base H and a (monotonic supraclassical) closure operation Cn, $\psi \leq_b \phi$ if and only if $\psi \leq_d \phi$ for every $\psi, \phi \in \ell$.*

Proof.

($\Rightarrow$): if $\psi \leq_b \phi$, we have that $\psi \notin Cn(\emptyset)$ and for every set $G \subseteq H$ that does not imply ϕ ($\phi \notin Cn(G)$) there is a set $G' \subseteq H$ that does not imply ψ ($\psi \notin Cn(G')$) and $G \subseteq G'$. By the monotonicity of Cn, we have that for every set $G \subseteq H$ s.t. $\phi \notin Cn(G)$, $\psi \notin Cn(G)$. That is, for every set $G \subseteq H$, if $\psi \in Cn(G)$, then $\phi \in Cn(G)$. $\psi \leq_d \phi$.

($\Leftarrow$): if $\psi \leq_d \phi$, we have that $\psi \notin Cn(\emptyset)$ and, for every set $G \subseteq H$, if $\psi \in Cn(G)$, then $\phi \in Cn(G)$. Consequently, for every set $G \subseteq H$, if $\phi \notin Cn(G)$, then $\psi \notin Cn(G)$. Hence, if $\phi \notin Cn(G)$, there is a set G' s.t. $G \subseteq G' \subseteq H$ and $\psi \notin Cn(G')$. $\psi \leq_b \phi$.
■

This dependence relation states that it is not possible to contract a belief base H by ϕ without withdrawing also ψ, and it is an appropriate relation

in order to account for the filtering condition.
Using such an ordering between formulae, Rott has proven ([51], Observation 5, Observation 6) that, working with belief bases instead that with belief sets, the contraction operation defined by means of Rott's definition is preferable to the one defined by Gärdenfors' definition; that is, given a dependence relation $\leq_d$, we can define a basic contraction operation $\dot{-}_b$ by means of the following rule.

$$\beta \in H_\alpha^{\dot{-}_b} \text{ iff } \beta \in H \text{ and either } \alpha <_d \beta \text{ or } \alpha \in Cn(\emptyset).$$

However, it remains a very 'rude' form of contraction.

8.3 Default revision

Now that we have presented some classical notions and tools from belief-revision theory, we can move toward a characterization of the revision of the default information associated to an agent reasoning nonmonotonically.
Recall from the previous chapter that the main intuition underlying our characterization of normality is based on the kind of "epistemic variation" that can be tolerated by a given preferential ordering (and therefore a given default-assumption consequence relation). It is therefore natural to compare our normality operator with the epistemic change operations constituting the AGM approach to theory change, namely expansion, contraction and revision. As we have seen, the AGM model, which aims at characterizing the epistemic behaviour of ideally rational agents, is centered around two key constraints: Logical closure and consistency. The former imposes that, given a set K, an agent should behave as if it accepted not only the information contained in K but also all of its logical consequence. The latter amounts to the requirement that no logical inconsistency should arise after the correct instantiation of any of the three epistemic operations. However, as seen in the previous chapter (see Theorem 7.2.1)[1], if the set of default-assumptions

[1]It is immediate to note the strict correspondence between this theorem and Corollary 8.1.6, since they resemble the same semantical construction.

Δ is closed under classical consequence ($\Delta = Cl(\Delta)$), the behaviour of the generated default-assumption consequence relation results trivial. In order to avoid this, we shall weaken the requirement of logical closure to closure under the normality operator $C_{\triangleright}$ which takes a set of sentences as argument and returns its closure under $\triangleright$ as value, that is:

$$C_{\triangleright}(A) = \{\alpha | A \triangleright \alpha\}.$$

In what follows, it will be useful to make the following terminological distinction. Analogously to beliefs, we shall refer to the finite set of default-assumptions Δ which determines a default-assumption consequence relation as the *default base*, while we shall call *default sets* those default bases D which are closed under the normality operator, that is to say that $D = C_{\triangleright}(D)$. Since we are going to use $C_{\triangleright}$ as our closure operation, we will confront our change operations with postulates obtained reinterpreting the AGM postulates with respect to $C_{\triangleright}$.

Dealing with default formulae, there is an important difference from the classical case: Theorem 7.1.14 states that for every default set we have a minimal corresponding default base. This result, that associates every $\triangleright$-theory with a distinct minimal base and that has no correspondent for classical theories, surely softens the polarity between the coherentist and the foundationalist approaches. Hence our policy will be slightly different from the traditional approach to belief revision: we shall define the rational desiderata with respect to $\triangleright$-theories, but then we shall propose some update operations referring to minimal default bases.

8.3.1 Expansion

Expansion formalizes the epistemic operation of simply adding a sentence to a default set D. We will use the symbol $\pm$ to represent default expansion: a function $\pm$ s.t. $\pm : \mathcal{D} \times \ell \mapsto \mathcal{D}$, where $\mathcal{D} \subseteq \wp(\ell)$ is the set of the $C_{\triangleright}$-theories in ℓ. Given a default set D, $D^{\pm}_{\alpha}$ is the belief set obtained by the addition of

α to D.
That is, if an agent acquires the information that α characterizes the normal situations, then it will simply add α to D and close this set under $C_{\triangleright}$:

$$D^{\pm}_{\alpha} := C_{\triangleright}(D \cup \{\alpha\})$$

The normality expansion operator thus defined satisfies the relevant AGM postulates, reformulated in order to account for the new closure operation.

$(D \pm 1)$ $\quad D^{\pm}_{\alpha}$ is a default set $\quad$ *Closure*

$(D \pm 1)$ forces the result of expansion to be a default set, i.e. closed under $C_{\triangleright}$.

$(D \pm 2)$ $\quad \alpha \in D^{\pm}_{\alpha}$ $\quad$ *Success*

$(D \pm 2)$ corresponds to the so-called *success postulate* and it follows by the reflexivity of $\triangleright$.

$(D \pm 3)$ $\quad D \subseteq D^{\pm}_{\alpha}$ $\quad$ *Inclusion*

$(D \pm 3)$ guarantees that no previous information is affected by the expansion of D with α. Again, this follows by reflexivity.

$(D \pm 4)$ $\quad$ If $\alpha \in D$, then $D^{\pm}_{\alpha} = D$ $\quad$ *Vacuity*

$(D \pm 4)$ captures an aspect of informational economy for the trivial cases: if the information to be added is already in the agent's default set, nothing changes. It clearly follows from the definition of $\triangleright$.

$(D \pm 5)$ $\quad$ If $D \subseteq D'$, then $D^{\pm}_{\alpha} \subseteq D'^{\pm}_{\alpha}$ $\quad$ *Monotonicity*

$(D \pm 5)$ is the *monotonicity postulate*. If $D \subseteq D'$, then the adding of α to D does not add anything that is not included also in $D'^{\pm}_{\alpha}$. It follows from the monotonicity of $\rhd$: $D \subseteq D'$ implies $D \cup \alpha \subseteq D' \cup \alpha$, so $C_{\rhd}(D \cup \{\alpha\}) \subseteq C_{\rhd}(D' \cup \{\alpha\})$.

$(D \pm 6)$ For all belief sets D and all sentences α, $D^{\pm}_{\alpha}$ is the smallest belief set that satisfies $(D \pm 1) - (D \pm 5)$ *Minimality*

This *minimality postulate* imposes that the new belief set does not contain any extra information with respect to the addition of α to D. To see that $(D \pm 6)$ holds, let H be such that $H \subseteq D^{\pm}_{\alpha}$. Assume H satisfies $(D \pm 2)$ and $(D \pm 3)$, that is $D \cup \{\alpha\} \subseteq H$. If H satisfies also $(D \pm 1)$, then $H = C_{\rhd}(H)$; given $D^{\pm}_{\alpha} = C_{\rhd}(D \cup \{\alpha\})$, we have $D^{\pm}_{\alpha} \subseteq H$, by the monotonicity of $\rhd$. Contradiction.

We conclude that the default expansion operator qualifies as an AGM-expansion operator.

8.3.2 Contraction

Let's now turn to the problem of removing a sentence from a given default set, that is to say to the operation of default contraction. As in the AGM case, the problem of contraction is two-fold: on the one hand a specific sentence needs removing from a default set; on the other hand, we need to to make sure that its deduction is blocked in the new set. Of course there can be many ways to achieve this latter result and the key heuristic principle to do so is again the principle of informational economy.
For the moment we will limit ourselves to a very general definition of a successful default contraction in order to check if such an operation is apt to the satisfaction of the basic AGM postulates.
Given a default set D and a sentence α, we define a function *contr* as follows:

$$contr_D(\alpha) = \begin{cases} \emptyset & if\ \alpha \notin D\ or\ \alpha = \top\ or\ \alpha = \bot, \\ \{B \subseteq D \mid \alpha \notin C_{\triangleright}(D - B)\} & otherwise \end{cases}$$

Such a definition informally recalls the *kernel* approach to belief contraction (see [21], pp.88-92), since the function $contr_D$ is analogous to the *incision function* proposed by Hansson; kernel contraction is, in fact, a generalization of the more popular AGM approach to contraction.
Note that the clause relative to $\alpha = \top$ and $\alpha = \bot$ accounts for the fact that both classical tautologies and contradictions are $\triangleright$-valid sentences, and thus cannot be removed from a default set. We can now define normality contraction by letting

$$D^-_\alpha := C_{\triangleright}(D - contr_D(\alpha)).$$

Such a definition imposes economical constraints just for the trivial cases of the contraction of a $\triangleright$-valid formula or of a formula which already is not in D. As for default expansion, default contraction is an AGM contraction operation, as we can ascertain by checking the relevant postulates in turn.

$(D-1)$ D^-_α is a default set *Closure*

This immediately follows from the closure under $C_{\triangleright}$.

$(D-2)$ $D^-_\alpha \subseteq D$ *Inclusion*

Contraction simply eliminates information from D so that every sentence in D^-_α must already be in D. It follows from $D - contr_D(\alpha) \subseteq D$ and the monotonicity of $\triangleright$.

$(D-3)$ If $\alpha \notin D$, then $D^-_\alpha = D$ *Vacuity*

If the sentence to be eliminated is not in D, then nothing needs changing. It is satisfied by the clause that if $\alpha \notin D$, then $contr_D(\alpha) = \emptyset$.

$(D-4)$ If $\not\triangleright\alpha$, then $\alpha \notin D^-_\alpha$ *Success*

The set resulting from a contraction must not contain the contracted sentence, unless it is a $\triangleright$-valid sentence (a classical tautology or contradiction), which cannot be eliminated. This is guaranteed by the definition of $contr(\alpha)$.

$(D-5)$ If $\alpha \in D$, then $D \subseteq (D^-_\alpha)^+_\alpha$ *Recovery*

Putting back a previously contracted sentence should not result in any loss of information. This property is not generally satisfied by the normality contraction operator. Assume we have to withdraw a formula $\alpha \wedge \beta$ from a default set D, and also α and β are in D. Obviously we have to eliminate as well one between α and β. However, given the failure of the $E\wedge$-rule, reintroducing $\alpha \wedge \beta$ we are not able to recover also α and β.
However, as in the classical case, this need not worry us too much, as the intuitive appeal of the recovery postulate is a controversial issue.

$(D-6)$ If $\vdash \alpha \leftrightarrow \beta$, then $D^-_\alpha = D^-_\beta$ *Extensionality*

Contraction should behave well with respect to classical equivalence. It follows from Right Logical Equivalence that $\beta \notin D^-_\alpha$ and $\alpha \notin D^-_\beta$, but this alone does not guarantee that $D^-_\alpha = D^-_\beta$. However, this can be ensured in various ways depending on the particular construction at hand.
The properties $(D-1)$-$(D-4)$ and $(D-6)$ are all desirable for a contraction of default. A word apart is needed for the translation of properties $K\dot{-}7$ and $K\dot{-}8$, which, given the particular behaviour of $\triangleright$ with respect to conjunction and disjunction, have to be reformulated.
The direct translations would be:

$(D-7) \quad D^-_\alpha \cap D^-_\beta \subseteq D^-_{\alpha\wedge\beta}$

$(D-8) \quad$ If $\alpha \notin D^-_{\alpha\wedge\beta}$, then $D^-_{\alpha\wedge\beta} \subseteq D^-_\alpha$

Since we do not have that $\alpha \wedge \beta \rhd \alpha$, the presence of $\alpha \wedge \beta$ in D does not imply the presence of α and β in D. This implies that the formulation of $(D-7)$ is not satisfying: if we have that $\alpha \wedge \beta \in D$, $\alpha, \beta \notin D$, we want $D^-_\alpha = D^-_\beta = D$, while $D^-_{\alpha\wedge\beta} \subset D$, violating $(D-7)$. On the other hand, if $\alpha, \beta \in D$, the presence of $\alpha \wedge \beta$ in D is justified by its derivability from α and β by $I\wedge$. Hence, in order to eliminate $\alpha \wedge \beta$, we will have to eliminate at least one between α and β, making $(D-7)$ a desirable property.
It is notable that, given the particular behaviour of $\rhd$, the assumption of a foundational perspective is obligated since the basic or derivative status of a formula can be discerned also in the closed default set and influences the mechanism of contraction.
Hence, we need a new formulation of $(D-7)$ in order to account for the fact that $\alpha \wedge \beta$ is not an independent formula, but it is supported by the presence in D of α and β. Hence, we have to modify our postulate as:

$(D-7a) \quad$ If $\alpha, \beta \in D$, then $D^-_\alpha \cap D^-_\beta \subseteq D^-_{\alpha\wedge\beta}$

On the other hand, the direct translation of $(K\dot{-}8)$ in $(D-8)$ seems intuitively satisfactory.
However, since in $\rhd$ the behaviour of the disjunction is analogous to the behaviour of the conjunction, we need analogous postulates for the disjunction, given the validity of $CI\vee$, but not of the classical $I\vee$.

$(D-7a_\vee) \quad$ If $\alpha, \beta \in D$, then $D^-_\alpha \cap D^-_\beta \subseteq D^-_{\alpha\vee\beta}$

$(D-8_\vee) \quad$ If $\alpha \notin D^-_{\alpha\vee\beta}$, then $D^-_{\alpha\vee\beta} \subseteq D^-_\alpha$

These properties impose binds with respect to the contraction of different sentences of a default set. They are not satisfied automatically by our general definition of contraction, and we have to check if they are satisfied by the particular contraction construction at hand.

8.3.3 Revision

The problem of revising a default set D consists in adding to D a sentence which is potentially inconsistent with it, without affecting the consistency of D. We formalize normality revision by means of a function *, which takes a default set D and a sentence α as inputs and returns a new consistent default set containing α. Here, as in the AGM approach, consistency means classical consistency, since the classical consistency of our default set is an intuitive desideratum.
First of all, then, we need to make sure that our closure operator is *consistency-preserving*, that is, for any given classically consistent default base K, the closure operator should return a consistent default set. This is not immediate in our setting since classical contradictions are $\rhd$-valid sentences and so $C_{\rhd}$ is not consistency-preserving. However, we can easily constrain the closure operator $\rhd$ to ensure consistency preservation. This is done by letting

$$K \rhd' \phi \Leftrightarrow \forall \mathfrak{M} \in \mathcal{C}, \text{ if } \mathfrak{M} \Vdash \psi \forall \psi \in K \,\text{and}\, \exists w \in W \,\text{s.}t.\, K_w = K, \text{ then } \mathfrak{M} \Vdash \phi \text{ and } w \vDash \phi.$$

We can prove the following proposition.

Proposition 8.3.1.

$$C_{\rhd'}(\Delta) = \begin{cases} C_{\rhd}(\Delta) - \bot & if\ K\, is\ consistent \\ \ell & otherwise \end{cases}$$

Proof. If K is inconsistent, then there is no valuation satisfying it and vacuously allows everything to follow. Thus $\rhd'$ is an explosive operator.

Otherwise, if Δ is consistent, there is a valuation w such that $w \vDash \psi$ for any $\psi \in K$. Since $\Delta_w = K$, then $w \leq_{\varepsilon_\Delta} v$ holds for every $v \in U$. We have to show that $C_{\triangleright'}(\Delta) = C_{\triangleright}(\Delta) - \bot$. $C_{\triangleright'}(\Delta) \subseteq C_{\triangleright}(\Delta) - \bot$ is immediate, since $C_{\triangleright'}(\Delta) \subseteq C_{\triangleright}(\Delta)$ and $\bot \notin C_{\triangleright'}(\Delta)$ hold. To show that $C_{\triangleright}(\Delta) - \bot \subseteq C_{\triangleright'}(\Delta)$, assume that there is a $\psi \neq \bot$ s.t. $\psi \in C_{\triangleright}(\Delta)$ and $\psi \in C_{\triangleright'}(\Delta)$. Since $\psi \neq \bot$, there must be a valuation $v \in U$ such that $v \vDash \psi$. Moreover, the fact that $\psi \in C_{\triangleright}(K)$ forces $t \vDash \psi$ for every $t \leq_{\varepsilon_\Delta} v$. Now it follows from $w \leq_{\varepsilon_\Delta} v$ that $w \vDash \psi$. Hence, we get $\psi \in C_{\triangleright'}(\Delta)$. Contradiction. □

The above Proposition shows that, in the case of a consistent default base, the constrained closure operator does preserve consistency, but it does so at the price of eliminating only contradictions, which are irrelevant to the construction of maximally consistent sets.

Consistency preservation and explosion make $\triangleright'$ an intuitively appealing operator in the characterization of normality: it would surely be counterintuitive if an ideally rational agent could hold a sentence as a *normal contradiction*.

Returning to the formal properties, it is immediate to note that $\triangleright'$ satisfies exactly the same structural properties as $\triangleright$, apart, obviously, from Contradiction. It is likewise easy to see that $\triangleright'$ behaves exactly as $\triangleright$ insofar as the properties of expansion and contraction are concerned.
We could use $C_{\triangleright'}$ to characterize *normality revision* via the so-called Levi Identity (see e.g. [17] p.69), which defines revision by means of a combination of expansion and contraction:

$$D^*_\alpha := (D^-_{\neg\alpha})^+_\alpha.$$

The Levi Identity formalizes the two-step operation of revision, where the initial contraction guarantees the consistency of the result while the final expansion guarantees the success of the revision.

It often happens that revision is studied as a primitive operation $*$ which is required to satisfy the following postulates:

(* 1) D^*_α is a default set.
Closure for revision.

(* 2) $\alpha \in D^*_\alpha$.
The *success postulate* of revision.

(* 3) $D^*_\alpha \subseteq D^+_\alpha$. (* 4) If $D \nvdash \neg\alpha$, then $D^+_\alpha \subseteq D^*_\alpha$.
(* 3) and (* 4) define the relation between revision and expansion. If the sentence α is consistent with D, revision amounts to expansion.

(* 5) $D^*_\alpha = D_\perp$ if and only if $\vdash \neg\alpha$.
The preservation of consistency is always satisfied, except in the case where α is itself a contradiction.

(* 6) If $\vdash \alpha \leftrightarrow \beta$, then $D^*_\alpha = D^*_\beta$.
Revision is syntax-independent.

Theorems 8.1.2 and 8.1.3 point out that the revision operator defined via the Levi Identity satisfies the AGM postulates. Of course, in the AGM model this result is obtained taking classical consequence as closure operator.

However, since we point to a revision operation apt to the consistency preservation of the default set D, i.e. $(*5)$, a revision operation obtained by the combination of $\pm$ and $-$ by means of the Levi Identity is not satisfying, given the failure of (RW) for $\rhd$. For example, assume a singleton default set $\Delta = \{\beta\}$, and that $\vdash \beta \rightarrow \neg\alpha$ and $\nvdash \neg\alpha \rightarrow \beta$ (i.e. $\neg\alpha$ is a consequence of β, but they are not Cn_U-equivalent). Given the failure of (RW), we have that $\neg\alpha \notin C_{\rhd'}(\beta)$. Assume we want to add α to Δ. Since $\neg\alpha \notin C_{\rhd'}(\beta)$, by (* 3) and (* 4), we have that $C_{\rhd'}(\Delta)^*_\alpha = C_{\rhd'}(\Delta)^\pm_\alpha$, that is, $C_{\rhd'}(\Delta)^*_\alpha = C_{\rhd'}(\beta, \alpha)$. Here we have a default base $\Delta' = \{\beta, \alpha\}$ s.t. $\alpha \wedge \neg\alpha \notin C_{\rhd'}(\Delta')$, but, by $(I\wedge)$, $\alpha \wedge \beta \in C_{\rhd'}(\Delta')$ that, given $\vdash \beta \rightarrow \neg\alpha$, is an inconsistent formula.

So, to obtain a revision operation apt to keep a default set consistent, we have to contract it with respect to the monotonic operation Cn in order to guarantee Cn-consistency.
Notwithstanding, the expansion operation has to be done with respect to $\rhd'$, since the success of an expansion of a default base Δ has to be valuated with respect to its default closure.

The default-contraction operation maintains a value by itself. Notwithstanding the uncertain status of contraction operation per se, we think that the simple withdrawal of a default formula, not induced by the introduction of a new conflicting default, is a plausible epistemic event, that needs to be analyzed.
Hence, we have to model a default-contraction function apt to obtain the Cn-consistency of the revised default-set, i.e. we need a function, which we will call $\div$, that, given a default set D and a formula α to be contracted, gives back a default set $D^{\div}_{\alpha}$ s.t. $D^{\div}_{\alpha} \nvdash \alpha$, instead of $D^{\div}_{\alpha} \not\triangleright' \alpha$ as above.
The desired properties of our contraction function have to be restated with respect to $\triangleright'$-closure, but with respect to Cn-consistency. In particular, we need to reformulate the third, the fourth and the fifth desiderata. However, since we move from a contraction operation with respect to $\triangleright$ to a contraction based on $\vdash$-derivability, the postulate 7 has to be treated in its original form:

($\div$ 1) $D^{\div}_{\alpha}$ is a default set.

($\div$ 2) $D^{\div}_{\alpha} \subseteq D$

($\div$ 3) If $D \nvdash \alpha$, then $D^{\div}_{\alpha} = D$

($\div$ 4) If $\nvdash \alpha$, then $\alpha \notin D^{\div}_{\alpha}$

($\div$ 5) If $\alpha \in D$, then $D \subseteq (D^{\div}_{\alpha})^{+}_{\alpha}$

($\div$ 6) If $\vdash \alpha \leftrightarrow \beta$, then $D^{\div}_{\alpha} = D^{\div}_{\beta}$

($\div$ 7) $D^{\div}_{\alpha} \cap D^{\div}_{\beta} \subseteq D^{\div}_{\alpha\wedge\beta}$

($\div$ 8) If $\alpha \notin D^{\div}_{\alpha\wedge\beta}$, then $D^{\div}_{\alpha\wedge\beta} \subseteq D^{\div}_{\alpha}$

We can construct such a function $\div$ on the basis of a classical contraction function $\dot{-}$ respect to Cn-theories, where Cn represent the monotonic core of the default assumption system:

$$D^{\div}_{\alpha} := Cn(D)^{\dot{-}}_{\alpha} \cap D$$

Proposition 8.3.2. *If the classical contraction function* $\dot{-}$ *satisfies* $(K\dot{-}1)$*-*$(K\dot{-}4)$ *and* $(K\dot{-}6)$*, then the default contraction function* $\div$ *satisfies* $(\div\ 1)$*-*$(\div\ 4)$ *and* $(\div\ 6)$

Proof.

$(\div\ 1)$: $D^{\div}_{\alpha}$ is a default set.

We have to show that $Cn(D)^{\dot{-}}_{\alpha} \cap D = C_{\rhd'}(Cn(D)^{\dot{-}}_{\alpha} \cap D)$. Given the reflexivity of $C_{\rhd'}$, $Cn(D)^{\dot{-}}_{\alpha} \cap D \subseteq C_{\rhd'}(Cn(D)^{\dot{-}}_{\alpha} \cap D)$ is obvious. We have to prove that $C_{\rhd'}(Cn(D)^{\dot{-}}_{\alpha} \cap D) \subseteq Cn(D)^{\dot{-}}_{\alpha} \cap D$.

If Δ is Cn-consistent, from Theorem 7.1.11 and Proposition 8.3.1 , we have that $C_{\rhd'}(\Delta) = (\Delta)^{\wedge\vee} - \bot$, where $(\Delta)^{\wedge\vee}$ is the closure of Δ under disjunction and conjunction. Since $Cn(D)^{\dot{-}}_{\alpha}$ is always a Cn-consistent set, so is $Cn(D)^{\dot{-}}_{\alpha} \cap D$, and hence $C_{\rhd'}(Cn(D)^{\dot{-}}_{\alpha} \cap D) = (Cn(D)^{\dot{-}}_{\alpha} \cap D)^{\wedge\vee} - \bot$.

It is sufficient to prove that $Cn(D)^{\dot{-}}_{\alpha} \cap D$ is closed under disjunction and conjunction:

Assume $\alpha, \beta \in Cn(D)^{\dot{-}}_{\alpha} \cap D$, then $\alpha, \beta \in Cn(D)^{\dot{-}}_{\alpha}$ and $\alpha, \beta \in D$.

If $\dot{-}$ satisfies $(K\dot{-}1)$, then $Cn(D)^{\dot{-}}_{\alpha}$ is a Cn-theory. By Cn-closure, $\alpha, \beta \in Cn(D)^{\dot{-}}_{\alpha}$ implies $\alpha \vee \beta \in Cn(D)^{\dot{-}}_{\alpha}$ and $\alpha \wedge \beta \in Cn(D)^{\dot{-}}_{\alpha}$.

By $C_{\rhd'}$-closure, $\alpha, \beta \in D$ implies $\alpha \vee \beta \in D$ and $\alpha \wedge \beta \in D$.

Then, if $\alpha, \beta \in Cn(D)^{\dot{-}}_{\alpha} \cap D$, we have $\alpha \vee \beta \in Cn(D)^{\dot{-}}_{\alpha} \cap D$ and $\alpha \wedge \beta \in Cn(D)^{\dot{-}}_{\alpha} \cap D$.

Hence $D^{\div}_{\alpha}$ is a default set and $(\div\ 1)$ is satisfied.

$(\div\ 2)$: $D^{\div}_{\alpha} \subseteq D$.

It is obviously satisfied by the definition.

($\div$ 3): if $D \nvdash \alpha$, then $D^{\div}_{\alpha} = D$.
Since $\dot{-}$ satisfies ($K\dot{-}3$), if $D \nvdash \alpha$, then $Cn(D)^{\dot{-}}_{\alpha} = Cn(D)$. So, $D^{\div}_{\alpha} = Cn(D) \cap D$, which is D.

($\div$ 4): if $\nvdash \alpha$, then $\alpha \notin D^{\div}_{\alpha}$.
Since $\dot{-}$ satisfies ($K\dot{-}4$), if $\nvdash \alpha$, then $\alpha \notin Cn(D)^{\dot{-}}_{\alpha}$, i.e. $\alpha \notin D^{\div}_{\alpha}$.

($\div$ 6): if $\vdash \alpha \leftrightarrow \beta$, then $D^{\div}_{\alpha} = D^{\div}_{\beta}$
Given that $\dot{-}$ satisfies ($K\dot{-}6$), if $\vdash \alpha \leftrightarrow \beta$, then $Cn(D)^{\dot{-}}_{\alpha} = Cn(D)^{\dot{-}}_{\beta}$, that is, $D^{\div}_{\alpha} = D^{\div}_{\beta}$.

■

The same cannot be proved about ($\div$ 5): if $\alpha \in D$, then $D \subseteq (D^{\div}_{\alpha})^{\pm}_{\alpha}$
It is easy to show that the fact that the classical contraction function $\dot{-}$ satisfies the postulate ($K\dot{-}5$) does not guarantee that the corresponding default contraction satisfies the correspondent postulate ($\div$ 5).
Assume we have a default assumption system defined by an empty knowledge set (i.e. its monotonic core is Cl) and its default set is

$$D = \{\top, \alpha, \beta, \alpha \wedge \beta, \alpha \vee \beta\}$$

It is easy to see that, modulo logical equivalence, D is closed under $C_{\triangleright'}$.
Assume a classical AGM contraction function $\dot{-}$ satisfying ($K\dot{-}5$): $Cl(D) \subseteq Cl(Cl(D)^{\dot{-}}_{\alpha} \cup \{\alpha\})$.
We define $\div$ with respect to $\dot{-}$:

$$D^{\div}_{\phi} = Cl(D)^{\dot{-}}_{\phi} \cap D$$

Suppose we want to contract D with respect to $\alpha \wedge \beta$. The operator $\dot{-}$, and consequently also the operator $\div$, will eliminate $\alpha \wedge \beta$ and at least one between α and β, or both of them. Assume the $\dot{-}$ eliminates only β (the other two cases are analogous). It is easy to check the validity of ($K\dot{-}5$) in

this case, i.e. $Cl(D) \subseteq Cl(Cl(D)^{\dot{-}}_{\alpha\wedge\beta} \cup \{\alpha \wedge \beta\})$.
However, we have that

$$D^{\div}_{\alpha\wedge\beta} = Cl(D)^{\dot{-}}_{\alpha\wedge\beta} \cap D = \{\top, \alpha, \alpha \vee \beta\}$$

Hence we have that

$$(D^{\div}_{\alpha\wedge\beta})^{\pm}_{\alpha\wedge\beta} = C_{\rhd'}(\{\top, \alpha, \alpha \vee \beta\} \cup \{\alpha \wedge \beta\})$$

which implies that $\beta \notin (D^{\div}_{\alpha\wedge\beta})^{\pm}_{\alpha\wedge\beta}$, i.e. ($\div$ 5) fails.
Again, we recall that the status of the recovery postulate as a desiderate rational constraint is debatable, and that the failure of such a property does not damage the properties of the resulting revision operation.
On the other hand, the postulates ($\div$ 7) and ($\div$ 8) are implied by the satisfaction of $K\dot{-}7$ and $K\dot{-}8$ by the classical operator $\dot{-}$.

Proposition 8.3.3. *Assume the preconditions in the above proposition are satisfied. Then,*
- if the classical contraction function $\dot{-}$ satisfies ($K\dot{-}7$), then the default contraction function $\div$ satisfies ($\div$ 7).
- if the classical contraction function $\dot{-}$ satisfies ($K\dot{-}8$), then the default contraction function $\div$ satisfies ($\div$ 8).

Proof.
($\div$ 7): $D^{\div}_{\alpha} \cap D^{\div}_{\beta} \subseteq D^{\div}_{\alpha\wedge\beta}$.

Assume that we have a classical contraction operator satisfying ($K\dot{-}7$), that is

$$Cn(D)^{\dot{-}}_{\alpha} \cap Cn(D)^{\dot{-}}_{\beta} \subseteq Cn(D)^{\dot{-}}_{\alpha\wedge\beta}$$

This implies that

$$(Cn(D)^{\dot{-}}_{\alpha} \cap Cn(D)^{\dot{-}}_{\beta}) \cap D \subseteq Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \cap D$$

that is equivalent to

$$(Cn(D)^{\dot{-}}_{\alpha} \cap D) \cap (Cn(D)^{\dot{-}}_{\beta} \cap D) \subseteq Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \cap D$$

which is

$$D^{\div}_{\alpha} \cap D^{\div}_{\beta} \subseteq D^{\div}_{\alpha\wedge\beta}$$

($\div$ 8): if $\alpha \notin D^{\div}_{\alpha\wedge\beta}$, then $D^{\div}_{\alpha\wedge\beta} \subseteq D^{\div}_{\alpha}$.

$\div$ is defined over an operation $\dot{-}$ satisfying ($K\dot{-}8$) and $\div$ satisfies ($\div$ 3), ($\div$ 2) by the previous proposition.
Hence, by ($K\dot{-}8$) we have that

$$\alpha \notin Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \text{ implies } Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \subseteq Cn(D)^{\dot{-}}_{\alpha}$$

Assume $\alpha \notin D^{\div}_{\alpha\wedge\beta}$, which is $\alpha \notin Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \cap D$.
That means that $\alpha \notin Cn(D)^{\dot{-}}_{\alpha\wedge\beta}$ or $\alpha \notin D$.
Assume that $\alpha \notin D$, which by ($\div$ 3) implies that $D^{\div}_{\alpha} = D$. Since by ($\div$ 2) we have that $D^{\div}_{\alpha\wedge\beta} \subseteq D$, we also have that $D^{\div}_{\alpha\wedge\beta} \subseteq D^{\div}_{\alpha}$.
Assume that $\alpha \notin Cn(D)^{\dot{-}}_{\alpha\wedge\beta}$. Then, by ($K\dot{-}8$), we have that $Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \subseteq Cn(D)^{\dot{-}}_{\alpha}$, which implies $Cn(D)^{\dot{-}}_{\alpha\wedge\beta} \cap D \subseteq Cn(D)^{\dot{-}}_{\alpha} \cap D$, i.e. $D^{\div}_{\alpha\wedge\beta} \subseteq D^{\div}_{\alpha}$.
■

Hence, we assume a default contraction function $\div$, built over a *Cn*-contraction operation $\dot{-}$, and again we define the default-revision by means of the Levy-identity:

$$D^{*}_{\alpha} := (D^{\div}_{\neg\alpha})^{\pm}_{\alpha}$$

We can check that, given a well-behaving contraction operation $\div$, the new revision operation satisfies the AGM logical postulates.

Theorem 8.3.1. *If the default contraction function* $\div$ *satisfies* ($\div$ 1)-($\div$ 4) *and* ($\div$ 6)*, and the normality expansion function* $\pm$ *satisfies* ($\pm$ 1)-($\pm$ 6)*, then the normality revision function* $*$*, defined via the Levi identity, satisfies (* 1)-(* 6).*

Proof.

It is sufficient to restate the proof of Theorem 3.2 in [17].

(* 1) and (* 2) are obviously satisfied by postulates ($\pm$ 1) and ($\pm$ 2) of default expansion.

Postulate (* 3) follows from the fact that $\div$ satisfies ($\div$ 2), i.e. $D^{\div}_{\neg\alpha} \subseteq D$.

Assume $D \nvdash \neg\alpha$; hence, by ($\div$ 3), $D^{\div}_{\alpha} = D$, and $(D)^{\pm}_{\alpha} = (D^{\div}_{\neg\alpha})^{\pm}_{\alpha}$, i.e. (* 4).

(* 5) follows from ($\div$ 4).

(* 6) follows directly from ($\div$ 6).

■

8.4 Default-base revision

Now we shall move to the definition of plausible default-revision operators at the level of default bases, and we will see if such operators satisfy the rationality constraints defined above.

8.4.1 Default-base expansion

Assume we have a finite set Δ of default formulae and we want to add to it a formula ϕ. Base expansion can be modeled simply by adding ϕ to the set Δ.

We address to base-expansion by means of the symbol $\pm_b$:

$$\Delta^{\pm_b}_{\alpha} = \Delta \cup \{\alpha\}$$

We can immediately see that the operation of expansion $\pm$ defined for default theories corresponds exactly to the closure of the operation of expansion $\pm_b$ defined for bases.

Proposition 8.4.1. *The theory-expansion operation $\pm$ corresponds exactly to the closure under $C_{\triangleright}$ of the base-expansion operation $\pm_b$, i.e., for every Δ*

and every α:

$$C_{\triangleright}(\Delta)^{\pm}_{\alpha} = C_{\triangleright}(\Delta^{\pm_b}_{\alpha})$$

Proof.
By definition, $C_{\triangleright}(\Delta)^{\pm}_{\alpha}$ corresponds to $C_{\triangleright}(C_{\triangleright}(\Delta) \cup \{\alpha\})$. Since $\triangleright$ is a tarskian relation, we have that $C_{\triangleright}(C_{\triangleright}(\Delta) \cup \{\alpha\}) = C_{\triangleright}(\Delta \cup \{\alpha\})$, i.e. $C_{\triangleright}(\Delta)^{\pm}_{\alpha} = C_{\triangleright}(\Delta^{\pm_b}_{\alpha})$.
■

On the semantical side, expansion corresponds to a simple operation of intersection between orderings: the preorder generated by the expanded base corresponds to the intersection of the preorder generated by the original base with the preorder generated by the added proposition.

Proposition 8.4.2. *For every Δ and every α:*

$$\varepsilon_{\Delta^{\pm_b}_{\alpha}} = \varepsilon_{\Delta} \cap \varepsilon_{\alpha}$$

Proof.
Since $\Delta^{\pm_b}_{\alpha} = \Delta \cup \{\alpha\}$, $\varepsilon_{\Delta^{\pm_b}_{\alpha}} = \varepsilon_{\Delta \cup \{\alpha\}}$:

$$\varepsilon_{\Delta^{\pm_b}_{\alpha}} = \{(w, v) \in U \times U |\ v \vDash \psi \ \Rightarrow \ w \vDash \psi \text{ for every } \psi \in \Delta^{\pm_b}_{\alpha}\}\}$$

that is

$$\varepsilon_{\Delta^{\pm_b}_{\alpha}} = \{(w, v) \in U \times U |\ v \vDash \psi \ \Rightarrow \ w \vDash \psi \text{ for every } \psi \in \Delta \text{ and } v \vDash \alpha \ \Rightarrow \ w \vDash \alpha\}.$$

Since

$$\varepsilon_{\Delta} = \{(w, v) \in U \times U |\ v \vDash \psi \ \Rightarrow \ w \vDash \psi \text{ for every } \psi \in \Delta\}$$

and

$$\varepsilon_{\alpha} = \{(w, v) \in U \times U |\ v \vDash \alpha \ \Rightarrow \ w \vDash \alpha\}$$

it is immediate that the pairs of worlds in $\varepsilon_{\Delta^{+}_{\alpha}}$ are exactly the pairs of worlds that are both in ε_{Δ} and in ε_{α}:

$$\varepsilon_{\Delta^{\pm_b}_{\alpha}} = \varepsilon_{\Delta} \cap \varepsilon_{\alpha}$$

■

In general, we can see that the preorder generated by a default base Δ corresponds to the intersection of the preorders generated by every formula in Δ.

Proposition 8.4.3. *For every default base Δ,*

$$\varepsilon_\Delta = \bigcap\{\varepsilon_\phi | \phi \in \Delta\}$$

Since the operation of intersection between sets is commutative, we have that the expansion operation $\pm_b$ is commutative too.
Note that the relation between the default base and the generated default-assumption inference relations is not monotonic, that is,

$$\Delta \subseteq \Delta' \not\Rightarrow \mid\!\sim_\Delta \subseteq \mid\!\sim_{\Delta'}$$

For example, assume a default base $\Delta = \{p\}$ (and an empty knowledge set K) and take as premise the formula $\neg(p \wedge q)$. We have that $\neg(p \wedge q)$ is consistent with p and so $C_\Delta(\neg(p \wedge q)) = Cl(\neg(p \wedge q), p)$, which, for example, validates $\neg(p \wedge q) \mid\!\sim_\Delta p \wedge \neg q$. If we expand Δ in $\Delta' = \{p, q\}$, now we have two $\neg(p \wedge q)$-maxiconsistent subsets of Δ', i.e. $\{p\}$ and $\{q\}$, and we can easily see that now $\neg(p \wedge q) \not\mid\!\sim_{\Delta'} p \wedge \neg q$. Hence, augmenting our default information we do not necessarily augment our inferential power.

8.4.2 Default-base contraction

As we have seen , we can conceive two kinds of default contraction.
First we have a *pure* contraction operation $-$, that, aiming at the simple elimination of a formula ϕ from the default set, is modeled referring to $C_\triangleright$ as closure operation.
Alternatively, we can define a contraction operation $\div$, proposed in order to generate a proficient revision function by means of the Levi Identity. Since

revision is finalized to allow a consistent update of the default set, we need a stronger contraction operation, s.t. the contracted formula ϕ not only is withdrawn from the agent's default set $C_{\triangleright}(H)$, but it is also no longer present in its classical (or supraclassical) closure $Cn(H)$.
Hence we are going to describe two different basic contraction operation, that we will distinguish as *pure* default-base contraction $-_b$ and *derived* default-base contraction $\div_b$.

Pure default-base contraction

To define an operator for pure default-base contraction we have to refer to $\triangleright$-derivability.
Hence, we shall define a notion of remainder set $(\Delta \bowtie \alpha)$ for default formulae, that is,

$$\Delta \bowtie \alpha = \{\Delta' \subseteq \Delta | \Delta' \not\triangleright \alpha \text{ and } \Delta'' \triangleright \alpha \text{ for every } \Delta'' \text{ s.t. } \Delta' \subset \Delta'' \subseteq \Delta\}$$

We can define default-base contraction by means of a full-meet approach (recall that working with finite bases the full-meet approach does not return trivial results):

$$\Delta_{\alpha}^{-_b} = \begin{cases} \bigcap(\Delta \bowtie \alpha) & when \ \not\triangleright \alpha \\ \Delta & otherwise \end{cases}$$

We can check if the closure of this operation satisfies the rationality postulates defined above. That is, if setting $D = C_{\triangleright}(\Delta)$ and $D_{\alpha}^{-} = C_{\triangleright}(\Delta_{\alpha}^{-_b})$ we satisfy the rationality postulates defined for default contraction $-$.

Proposition 8.4.4. *The default contraction operation $-$, defined from the $\triangleright$-closure of a full-meet default-base contraction operation $-_b$, satisfies the postulates* $(D-1)-(D-4)$, $(D-6)$, $(D-7a)$, $(D-7a_{\vee})$, $(D-8)$, *and* $(D-8_{\vee})$.

Proof.
$(D-1)$ is obviously satisfied because D_{α}^{-} is defined as closed under $\triangleright$.

$(D-2)$ is satisfied because $\Delta_{\alpha}^{-b} \subseteq \Delta$ and $\rhd$ is monotone.
$(D-3)$ is satisfied as well, since $\alpha \notin C_{\rhd}(\Delta)$ implies $\Delta \bowtie \alpha = \Delta$.
$(D-4)$ is implied by the construction of the set $\Delta \bowtie \alpha$.
The satisfaction of the extensionality property $(D-6)$ is guaranteed by the fact that $C_{\rhd}$ is closed under logical equivalence, and consequently the contractions of logically equivalent formulas generate the same remainder set.
For $(D-7a)$, note that, if $\alpha, \beta \in D$, then every set in $\Delta \bowtie (\alpha \wedge \beta)$ is also in $\Delta \bowtie \alpha$ or in $\Delta \bowtie \beta$. Consequently, we have that $\Delta \bowtie (\alpha \wedge \beta) \subseteq \Delta \bowtie \alpha \cup \Delta \bowtie \beta$. This implies that $\bigcap(\Delta \bowtie \alpha) \cap \bigcap(\Delta \bowtie \beta) \subseteq \bigcap(\Delta \bowtie (\alpha \wedge \beta))$. By the monotonicity of $\rhd$, $D_{\alpha}^{-} \cap D_{\beta}^{-} \subseteq D_{\alpha\wedge\beta}^{-}$.
The same argument holds for $(D-7a_{\vee})$.
For $(D-8)$, assume $\alpha \notin D_{\alpha\wedge\beta}^{-}$, that is, $\alpha \notin C_{\rhd}(\Delta_{\alpha\wedge\beta}^{-b})$. If $\alpha \notin D_{\alpha\wedge\beta}^{-}$, we have two possibilities: that is, $\alpha \notin D$ or $\alpha \in D$. In the first case, we have $\Delta \bowtie (\alpha \wedge \beta) = \Delta \bowtie \alpha = \Delta$, and, consequently, $D_{\alpha\wedge\beta}^{-} = D_{\alpha}^{-} = D$. In the other case, the elimination of α from Δ has been necessary in order to eliminate $\alpha \wedge \beta$, that is, we have both α and β in D. This implies that every set in $\Delta \bowtie \alpha$ is also in $\Delta \bowtie \alpha \wedge \beta$. Hence, $\bigcap(\Delta \bowtie (\alpha \wedge \beta)) \subseteq \bigcap(\Delta \bowtie \alpha)$, and $D_{\alpha\wedge\beta}^{-} \subseteq D_{\alpha}^{-}$.
Analogously for $D-8_{\vee}$.
■

In order to refine full-meet base contraction into partial-meet contraction, we would need to define a choice function or an entrenchment relation appropriate for this situation, but we do not see any way to build them: the formulae in a default base H are all independent with respect to the operator $\rhd$, given the $\rhd$-minimality of the base, and they all play the same role in the definition of the epistemic state of the agent.

Derived default-base contraction

In Section 8.3.3, we have defined the derived contraction operation $\div$ assuming a classical theory-contraction operator $\dot{-}$ and restricting its action to a $\triangleright$-closed set D (i.e., $D^{\div}_{\alpha} := Cn(D)^{\dot{-}}_{\alpha} \cap D$). We can use the same approach for the definition of a derived contraction operation $\div_b$ over default bases.
One possibility is to use classical full-meet contraction, referring to classical remainder sets $\Delta \bot \alpha$. This possibility works well, but it is a strong form of contraction. We could try to refine the operation.
Recall the dependence relation $\leq_d$ presented in Section 8.2. Bochmann ([5], p.145) has defined an easy mapping (*id-mapping*) that makes explicit the relation between preferential inference relations and the corresponding dependence relations:

ID $\quad \alpha \mid\!\sim \beta \quad$ iff $\quad \neg(\alpha \wedge \beta) \leq_d \neg\alpha$
DI $\quad \alpha \leq_d \beta \quad$ iff $\quad \neg(\alpha \wedge \beta) \mid\!\sim \neg\alpha$

However, we can show that, if we are working with default formulae, such dependence relation manifests quite a trivial behaviour, since it is completely determined by the property of *Dominance*, that is, it corresponds with the monotonic core $\vdash_{\mid\sim}$ of the inference relation $\mid\!\sim$: the presence of a formula α between our defaults depends upon the presence of a formula β just if β is a consequence of α.
To show this, we need also to recall a known property of remainder sets, the Upper Bound property, stating that, given a set A and a formula α, if a subset of A is consistent with $\neg\alpha$, then it is the subset of an element of the remainder set $A \bot \alpha$.

Lemma 8.4.5. *(Upper Bound Property)([21], Postulate 1.37)*
If $B \subseteq A$ and $B \nvdash \alpha$, then there is some B' s.t. $B \subseteq B' \in A \bot \alpha$.

Lemma 8.4.6. *Assume a preferential inference relation $\mid\!\sim$ determined by a set of defaults Δ and a knowledge set K. Let $\rho, \sigma \in \Delta$.*
Then $\sigma \vdash_K \rho$ if and only if $\sigma \leq_d \rho$.

Proof.
($\Leftarrow$): If $\sigma \leq_d \rho$, then $\sigma \vdash_K \rho$.

If $\sigma \leq_d \rho$, then we have $\neg(\sigma \wedge \rho) \mid\!\sim \neg\sigma$ by id-mapping. This implies that for every $\Delta' \subseteq \Delta$ s.t. Δ' is $\neg(\sigma \wedge \rho)$-maxiconsistent, then $\sigma \notin \Delta'$.
Assume $\sigma \nvdash_K \rho$. Then σ is consistent with $\neg\rho$, i.e. is consistent with $\neg(\sigma \wedge \rho)$, and, by the upper bound property, there is a $\Delta' \subseteq \Delta$ s.t. Δ' is maximally $\neg(\sigma \wedge \rho)$-consistent and $\sigma \in \Delta'$. Contradiction. Hence, $\sigma \vdash_{\mid\sim} \rho$.

($\Rightarrow$): If $\sigma \vdash_K \rho$, then $\sigma \leq_d \rho$.

Assume $\sigma \vdash_K \rho$. By reflexivity and AND, we have $\sigma \vdash_K \sigma \wedge \rho$. By contraposition, $\neg(\sigma \wedge \rho) \vdash_K \neg\sigma$. Since $\vdash_K$ is the monotonic core of $\mid\!\sim$, we obtain $\neg(\sigma \wedge \rho) \mid\!\sim \neg\sigma$, that, by id-mapping, implies $\sigma \leq_d \rho$.
■

Also if we move to the entrenchment relation $<_e$ defined by means of ie-mapping in Section 8.1.4, we obtain the same result, respect to the strict part of the dependence relation.

Lemma 8.4.7. *Assume a preferential inference relation $\mid\!\sim$ determined by a set of defaults Δ and a knowledge set K. Let $\rho, \sigma \in \Delta$.*
Then $\sigma <_e \rho$ if and only if $\sigma \vdash_K \rho$ and $\rho \nvdash_K \sigma$.

Proof.
($\Rightarrow$): If $\sigma <_e \rho$, then $\sigma \vdash_K \rho$ and $\rho \nvdash_K \sigma$.

If $\sigma <_e \rho$, then $\neg(\sigma \wedge \rho) \mid\!\sim \rho$ by ie-mapping.
Assume $\sigma \nvdash_K \rho$. This implies that σ is consistent with $\neg\rho$, and consequently with $\neg(\sigma \wedge \rho)$. Hence, by the upper bound property, there is a $\Delta' \subseteq \Delta$ s.t. Δ' is maximally $\neg(\sigma \wedge \rho)$-consistent, $\sigma \in \Delta'$ and, obviously, $\rho \notin \Delta'$.
Hence we have that $\neg(\rho \wedge \sigma) \not\mid\!\sim \rho$. Contradiction.

Assume $\rho \vdash_K \sigma$. Then, by REF and AND, $\rho \vdash_K \rho \wedge \sigma$, and , by the monotonicity of $\vdash_K$, $\neg(\rho \wedge \sigma) \wedge \rho \vdash_K \rho \wedge \sigma$. This implies $\neg(\rho \wedge \sigma) \wedge \rho \mid\!\sim \rho \wedge \sigma$, and, by CT, $\neg(\rho \wedge \sigma) \mid\!\sim \rho \wedge \sigma$. Contradiction.

($\Leftarrow$): If $\sigma \vdash_K \rho$ and $\rho \nvdash_K \sigma$, then $\sigma <_{e'} \rho$.
Since $\sigma \vdash_K \rho$ and $\rho \nvdash_K \sigma$ we have that ρ is consistent with $\neg(\sigma \wedge \rho)$, while σ is not. Hence, we have that ρ is in every $\neg(\sigma \wedge \rho)$-maxconsistent subset of Δ, while σ is not. Consequently, we have that $\neg(\sigma \wedge \rho) \mid\!\sim \rho$ and, by ie-mapping, $\sigma <_e \rho$.
■

Then we have that, working with a preferential inference relation $\mid\!\sim$, as long as we are dealing with default formulae the strict part of the dependence relation $\leq_d$ and the entrenchment relation $<_e$ correspond each other, and, moreover, they are completely determined by the monotonic core $\vdash_{\mid\sim}$.

Theorem 8.4.8. *Assume we have a default-assumption system $\mathfrak{S} = \langle \Delta, K \rangle$, where $\vdash_K$ is its monotonic core, $\leq_d$ is its dependence relation (determined by means of id-mapping), and $<_e$ is its entrenchment relation (determined by means of ie-mapping). Then, if $\sigma, \rho \in \Delta$, we have that*

$$\sigma <_d \rho \text{ iff } \sigma <_e \rho \text{ iff } [\sigma \vdash_K \rho \text{ and } \rho \nvdash_K \sigma].$$

Hence, we can use the relation $\leq_d$ to model derived default-base contraction for defaults, but it is not a very elaborated relation. Moreover, using the definition of contraction presented by Rott (see Section 8.2),

$$\beta \in \Delta_\alpha^{\dot{-}_b} \text{ iff } \beta \in \Delta \text{ and either } \alpha <_d \beta \text{ or } \alpha \in Cn(\emptyset),$$

we obtain a contraction operation resulting really too strong, since we would mantain in $\Delta_\alpha^{\dot{-}_b}$ only those formulae implied by α but not implying α.
We feel the need of a more efficient default-base contraction operation.

8.5 Conclusion

Since we have taken note of the incompatibility between default formulae and classical closure, in this chapter we have proposed a new interpretation of the classical theory of belief revision, in order to deal with the update of default assumptions. Such a change of perspective has given good results in the reinterpretation of the classical rationality postulates, but it must be further elaborated in order to deal efficiently with the revision of finite default-bases.

Chapter 9

Conclusions

The subject of this thesis, the logical characterization of defeasible reasoning, falls within the relatively recent (but, anyhow, unstable) reconciliation between epistemology and logics, which have met each other again in the field of Artificial Intelligence.

In particular, we have focused on how effective is the use of background formulae for modeling the unexpressed information that an agent uses in its everyday reasoning.

The use of maxiconsistent sets of default formulae can be seen as a simple but expressive and powerful tool. Such an approach has revealed itself as efficient and pervasive, since it has been used as a basis for incisive logical models of different epistemic phenomena.

Poole, in [45], has utilized default-assumption sets to model abductive processes: starting from a set A of known facts, an observation β, and a set Δ of default information representing possible hypothesis, we identify the possible explanations of β with the A-maxiconsistent subsets of Δ implying β. That is, given A and Δ, β is explicable iff it is derivable from A and Δ in a credulous way (see Section 2.2).

Datteri, Hosni and Tamburrini, in [9], have emphasized a possible role of default-assumption inference relations in modeling machine learning processes. In particular, it is possible to use the default-assumption approach

as a good logical model of a classical proposal in machine learning, i.e. the decision tree algorithm ID3 (see [48]).
Finally, we have seen in the previous chapter that a most classical approach to belief revision is based on meet contraction, that means that it is based on the identification of maxiconsistent sets of formulae and has an immediate connection with the structure of default-assumption inference relations (see [39], pp.44-45).

In this thesis we have tried to analyze more deeply the role of background information in modeling defeasible reasoning. We have identified a strong connection between the default-assumption approach and the most rigorous characterization of defeasible reasoning, that is, the consequentialist perspective delineated by Gabbay and Makinson ([15, 38]).
Referring to the strict connection with the class of preferential models, that have turned out to be an extremely proficient tool for the characterization of nonmonotonic inference relations, we have seen that the class of default-assumption inference relations comprehends the greatest part of the interesting nonmonotonic inference relations, since we exclude only those relations representable exclusively by means of non-injective preferential models.
Hence, the use of default-formulas turns out to be a very comprehensive logical tool, that, moreover, has a very simple formal structure.
The strict connection between default-assumption and preferential approaches results fruitful for both the perspective: on one hand, in Chapter 6, we have defined how to use the tools supplied by the default-assumption formalization in order to generate interesting preferential models; on the other hand, in Chapter 7, we have used the connection with preferential models in order to delineate in a precise way the behaviour of default-assumption sets.

Many of the results presented here would deserve further deepening.
The construction of interesting preferential models, given a conditional base of sequents (Chapter 6), is still an open field, since, as we have seen, $\mathbb{R}$-

closure and $\mathbb{W}$-closure have a clear problem in property-heritage that yet has not been solved, and the use of default assumptions could be a promising way to face it.

The problem of default-revision would need further work as well, not only in the definition of alternative characterizations of revision, but, in particular, we should look for some regularities connecting the variation of a default set Δ with the variation of the corresponding inference relation $\mid\!\sim_\Delta$, and further on, how the combined variation of the knowledge set K and of the default set Δ of an agent influence the generated inference relation.

The default revision presented in the last chapter enlightens what is probably a new problem in the definition of appropriate rational constraints in logical epistemic models. Up to now, logicians have modeled the management of epistemic information referring to its most natural form, beliefs, i.e. formulae that the agent maintains as true. Consequently, the rational constraints for the management of epistemic information, in our particular case for revision, have been formulated referring only to the idealizations traditionally linked to beliefs, that is, classical consistency and closure. Instead, actual models often distinguish between the different *epistemic attitudes* the agent can entertain toward information, that is, the agent treats each piece of information according to the role it plays in reasoning and deliberation processes; for example, the same formula can be treated as a belief, a desire, a goal, a commitment, and so on, and such different attitudes determine radically different roles the formula plays in the reasoning and deliberation 'machinery' of the agent. In the last two chapters we have shown that, since such different roles determine a different treatment of the information, it is possible that the idealizations traditionally connected to beliefs (as classical closure) result inappropriate for other epistemic attitudes, as in the case of default information. Such departures from classical constraints can determine the need for a rearrangement of the logical tools developed on the basis of such idealizations. This is exactly what we have done in the case of default formulae: analyzing the default-assumption approach, which is judged as a satisfying formalization, we can see that classical closure is not appropriate for the treatment of default formulae, and that such a behaviour is not counter-intuitive; consequently, there is the necessity of modifying appropriately every rational constraint defined on the basis of classical closure, as those of belief-revision theory, in order to apply them in the management of default information.

Lastly, the model of stereotypical reasoning presented at the end of Chapter 7 sounds promising, and further elaboration of the proposed notion of semantical distance, or of alternative notions, seems worth more study. Moreover, we conjecture that the model presented here for the characterization of stereotypical reasoning could also result proficient in other fields, for example in the logical analysis of deliberation processes: it seems plausible that a typical problem of deliberative and deontic logics, that is, the distinction between *goals*, *desires* and *duties* on one hand, and *collateral effects* on the other, could be formalized analogously to the distinction between *stereotypical properties* and simple *expectations*, modeled by means of the normality operator $\triangleright$; moreover, a notion of semantical distance, not necessarily strictly similar to the one presented here, could be useful in a valuation of the 'nearness' of alternative goals or desires to the actual situation, that is, which of them should presumptively be easier to be satisfied.

Bibliography

[1] Alchourròn C. E., Gärdenfors P., Makinson D. (1985), *On the logic of theory change: Partial meet contraction and revision functions*, Journal of Symbolic Logic, 50, pp.510-530.

[2] Alchourròn C. E., Makinson D. (1985), *On the logic of theory change: safe contraction*, Studia Logica, 44, pp.405-422.

[3] Benthem J. van (1996), *Exploring Logical Dynamics*, CSLI Publications.

[4] Bezzazi H., Makinson D., Pino Pérez R. (1997), *Beyond Rational Monotony: Some Strong Non-Horn Rules for Nonmonotonic Inference Relations*, Journal of Logic and Computation 7(5), pp.605-631.

[5] Bochman A. (2001), *A Logical Theory of Nonmonotonic Inference and Belief Change*, Springer.

[6] Bochman A. (2002), *Entrenchment vs. Dependence: Coherence and Foundations in Belief Change*, Journal of Logic, Language and Information, 11, pp.3-27.

[7] Bratman M. E. (1990), *What is Intention?*, in Cohen P. R., Morgan J., Pollack M. E., *Intentions in Communication*, MIT Press, pp.15-31.

[8] Brewka G. (1994), *Adding Properties and Specificity to Default Logic*, Artificial Intelligence 98(1-2), pp.209-235.

[9] Datteri E., Hosni H., Tamburrini G. (2006), *An inductionless, default based account of machine learning*, in Magnani L. (ed.), *Model-Based Reasoning in Science and Engineering*, College Publications, pp. 379-399.

[10] Davidson D. (1963), *Action, Reasons and Causes*, reprinted in Davidson D. (1980), *Essays on Actions and Events*, Oxford University Press, pp.3-19.

[11] Freund M. (1993), *Injective Models and Disjunctive Relations*, Journal of Logic and Computation 3(3), pp.231-247.

[12] Freund M. (1998), *Preferential Reasoning in the Perspective of Poole Default Logic*, Artificial Intelligence 98(1-2), pp.209-235.

[13] Freund M., Lehamnn D. (1996), *On Negation Rationality*, Journal of Logic and Computation 6(2), pp.263-269.

[14] Freund M., Lehamnn D., Morris P. (1991), *Rationality, Transitivity, and Contraposition*, Artificial Intelligence 52, pp.191-203.

[15] Gabbay D. (1985), *Theoretical Foundations for nonmonotonic reasoning in expert systems*, in Apt K.(ed.), *Logics and Models of Concurrent Systems*, Springer.

[16] Gabbay D., Woods J. (2005), *The Practical Turn in Logic*, in Gabbay D., Guenthner F. (eds.) (2005), *Handbook of Philosophical Logic*, vol.13, pp.15-122, Springer.

[17] Gärdenfors P. (1988), *Knowledge in Flux*, MIT Press.

[18] Gärdenfors P., Makinson D. (1988), *Revisions of knowledge systems using epistemic entrenchment*, in Vardi M. (ed.), *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, Morgan Kaufmann, pp.83-95.

[19] Gärdenfors P., Makinson D. (1994), *Nonmonotonic Inference Based on Expectations*, Artificial Intelligence, 65, pp.197-245.

[20] Gärdenfors P., Rott H. (1995), *Belief Revision*, in Gabbay D., Hogger C., Robinson J. (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming, vol.4*, pp.35-132, Clarendon Press.

[21] Hansson S. O. (1999), *A Textbook of Belief Dynamics. Theory Change and Database Updating*, Kluwer Academic Publishers.

[22] Hintikka J. (1962), *Knowledge and Belief. An Introduction to the Logic of the Two Notions*, Cornell University Press.

[23] Konolidge K. (1988), *On the Relation between Default and Autoepistemic Logic*, Artificial Intelligence, 35, pp.343-382.

[24] Konolige K. (1994), *Autoepistemic Logic*, in Gabbay D. M., Hogger C. J., Robinson J. A. (eds.) (1994), *Handbook of Logic in Artificial Intelligence and Logic Programming - vol. 3*, Clarendon Press, pp.217-295.

[25] Kraus S., Lehmann D., Magidor M. (1990), *Nonmonotonic reasoning, preferential models and cumulative logic*, Artificial Intelligence 44, pp.167-207.

[26] Kripke S. (1963), *Semantical Analysis of Modal Logic I. Normal Modal Propositional Calculi*, Zeitschr. f. math. Logik und Grundlagen d. Math., 9, pp.67-96.

[27] Lakoff G. (1987), *Cognitive Models and Prototype Theory*, in Laurence S, Margolis E. (eds.) (1999), *Concepts*, pp. 391-431, MIT Press.

[28] Laurence S, Margolis E. (1999), *Concepts and Cognitive Science*, in Laurence S, Margolis E. (eds.) (1999), *Concepts*, pp. 3-81, MIT Press.

[29] Lehmann D. (1995), *Another Perspective on Default Reasoning*, Annals of Mathematics and Artificial Intelligence, 15, pp.61-82.

[30] Lehmann D. (1998), *Stereotypical Reasoning: Logical Properties*, L. J. of the IGPL, 6 (1), pp.49-58.

[31] Lehmann D., Magidor M. (1992), *What does a Conditional Knowledge Base Entail?*, Artificial Intelligence, 55, pp.1-60.

[32] Lehmann D., Magidor M., Schlechta K. (2001), *Distance Semantics for Belief Revision*, The Journal of Symbolic Logic, 66 (1), pp.295-317.

[33] Levi I. (1997), *Contraction and Informational Value*, Columbia University Press.

[34] Lin F., Shoham Y. (1990), *Epistemic Semantics for Fixed-Points Non-Monotonic Logics*, in Parikh R. (ed.) (1990), *Theoretical Aspects of Reasoning about Knowledge - TARK1990*, Morgan Kaufmann.

[35] Lindström S., Rabinowicz W. (1991), *Epistemic Entrenchment with Incomparabilities and Relational Belief Revision*, in Fuhrmann, Morreau (eds.) (1991), *The Logic of Theory Change*, Springer-Verlag, pp. 208-228.

[36] Makinson D. (1987), *On the Status of Postulate of Recovery in the Logic of Theory Change*, Journal of Philosophical Logic, 16, pp.383-394.

[37] Makinson D. (1993), *Five Faces of Minimality*, Studia Logica, 52, pp.339-380.

[38] Makinson D. (1994), *General Patterns in Nonmonotonic Reasoning*, in Gabbay D., Hogger C., Robinson J. (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming, vol.3*, pp.35-110, Clarendon Press.

[39] Makinson D. (2005) *Bridges from Classical to Nonmonotonic Logic*, King's College Text in Computing, volume 5, King's College Publications.

[40] McCarthy J. (1968), *Programs with Common Sense*, in Brachman R.J., Levesque H.J. (eds.) (1985), *Readings in Knowledge Representation*, pp.299-308, Morgan Kaufmann.

[41] McCarthy J. (1980), *Circumscription - A form of non-monotonic reasoning*, Artificial Intelligence, 13, pp.27-39.

[42] McDermott D., Doyle J. (1980), *Non-monotonic Logic I*, Artificial Intelligence, 13, pp.41-72.

[43] Osherson D., Smith E. (1981), *On the Adequacy of the Prototype Theory as a Theory of Concepts*, in Laurence S, Margolis E. (eds.) (1999), *Concepts*, pp. 261-278, MIT Press.

[44] Paris J. B. (2003), *Nonmonotonic Logic*, Tech. Rep. MT4181/5181, University of Manchester, School of Mathematics.

[45] Poole D. (1988), *A Logical Framework for Default Reasoning*, Artificial Intelligence 36, pp.27-47.

[46] Priest G., Routley R., Norman J. (1989), *Paraconsistent Logic. Essays on the Inconsistent*, Philosophia Verlag.

[47] Putnam H. (1988), *Representation and Reality*, MIT Press.

[48] Quinlan J. R. (1983), *Induction of Decision Trees*, Machine Learning, 1(1), pp.81106.

[49] Reiter R. (1978), *On Closed World Data Bases*, in Gallaire H., Minker J. (eds.), *Logic and Data Bases*, pp. 119-140, Plenum.

[50] Reiter R. (1980), *A Logic for Default Reasoning*, Artificial Intelligence, 13, pp. 81-132.

[51] Rott H. (1998), *"Just Because": Taking Belief Bases Seriously*, Publications of the Universiteit van Amsterdam.

[52] Rott H. (2000), *Two Dogmas of Belief Revision*, The Journal of Philosophy, 97(9), pp. 503-522.

[53] Rott H. (2001), *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*, Oxford University Press.

[54] Rott H. (2003), *Basic Entrenchment*, Studia Logica, 73, pp. 257-280.

[55] Shoham Y. (1988), *Reasoning about Change*, MIT Press.

[56] Simon H. A. (1996), *The Sciences of Artificial*, 3^{rd} edition, MIT Press.

[57] Smith E., Osherson D., Rips L., Keane M. (1988), *Combyning Prototypes: A Selective Modification Model*, reprinted in Laurence S, Margolis E. (eds.) (1999), *Concepts*, pp. 355-390, MIT Press.

[58] Tarski A. (2002), *On the Concept of following logically*, History and Philosophy of Logic, 23, pp. 155-196.