

Conditional entailment: bridging two approaches to default reasoning

Hector Geffner

T.J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598, USA

Judea Pearl

Computer Science Department, UCLA, Los Angeles, CA 90024, USA

Received January 1991

Revised June 1991

Abstract

Geffner, H. and J. Pearl, Conditional entailment: bridging two approaches to default reasoning, *Artificial Intelligence* 53 (1992) 209–244.

In recent years, two conceptually different interpretations of default expressions have been advanced: *extensional* interpretations, in which defaults are regarded as prescriptions for extending one's set of beliefs, and *conditional* interpretations, in which defaults are regarded as beliefs whose validity is bound to a particular context. The two interpretations possess virtues and limitations that are practically orthogonal to each other. The conditional interpretations successfully resolve arguments of different "specificity" (e.g., "penguins don't fly in spite of being birds") but fail to capture arguments of "irrelevance" (e.g., concluding "red birds fly" from "birds fly"). The opposite is true for the extensional interpretations.

This paper develops a new account of defaults, called *conditional entailment*, which combines the benefits of the two interpretations. Like prioritized circumscriptions, conditional entailment resolves arguments by enforcing priorities among defaults. However, instead of having to be specified by the user, these priorities are extracted automatically from the knowledge base. Similarly, conditional entailment possesses a sound and complete proof theory, based on interacting arguments and amenable to implementation in conventional ATMSs.

1. Introduction

An important task in any formal characterization of plausible reasoning is to account for the set of conclusions that a given set of facts and defaults legitimizes. This requires a language for expressing facts and defaults, and a

specification of how expressions in such a language are to be interpreted. Classical logic is not well suited for the task because of its failure to accommodate the nonmonotonic character of default reasoning, according to which conclusions may be retracted when additional information becomes available. This “limitation” of classical logic led to the development of *nonmonotonic* logics, in which theorems may come and go as additional axioms are considered [24, 26, 27, 37].

In the context of nonmonotonic logics, defaults are interpreted as rules which permit us to *extend* a set of beliefs in the absence of conflicting evidence. Such evidence may originate from either known facts or from other, conflicting defaults, and may give rise to multiple “extensions”. However, as pointed out by Reiter and Criscuolo [39] and Hanks and McDermott [14], these “extensions” do not always reflect equally plausible scenarios. For that reason, the encoding of defaults in nonmonotonic logics is usually augmented with suitable cancellation axioms, priorities or similar devices, which render the spurious scenarios inconsistent or less preferred.

More recently, it has been noted that some of the problems caused by interacting defaults can be solved by interpreting defaults as conditional assertions [5, 9, 16]. Whereas “extensional” interpretations regard the default “if p then normally q ” as a soft reason to believe q given the truth of p , conditional interpretations regard it as a hard constraint to believe q in a limited context defined by p and possibly some background knowledge. As a result, conditional interpretations properly “dissolve” certain spurious conflicts among defaults, such as those arising from the rules “if a then c ” and “if a and b then $\neg c$ ”, when both a and b are known to be true. In that context, the second and “more specific” default is applicable, while the first one is not.

In spite of these virtues, however, conditional interpretations fail to account for a number of desirable inferences which are captured by extensional formalisms. This limitation has to do with the way *irrelevant* information is handled. For instance, given a default “if p then q ” both extensional and conditional interpretations will conclude q given the evidence p ; conditional interpretations, however, are unable to maintain that conclusion when an additional but irrelevant piece of evidence e is taken into account. The reason is that conditional interpretations treat all evidence as *relevant* unless otherwise proven, and hence refrain from maintaining q in the presence of e . Thus, while extensional interpretations generate conflicts out of braveness, conditional interpretations eliminate conflict out of hesitancy.

The question arises whether a unifying framework can be developed which combines the virtues of both the extensional and conditional interpretations. An earlier attempt in this direction was a proposal to enhance a conditional interpretation of defaults based on probabilities with a syntactic criterion for distinguishing relevant from irrelevant evidence [9] (see also [5]). However, while the results, for the most part, were satisfactory, the theoretical underpin-

nings were not. A promising theoretical attempt has recently been advanced by Lehmann [17] and Pearl [32], who view defaults as partial specifications for ranking worlds. However, while the resulting account is better motivated, it is also less successful: some useful inferences fail to be captured, while anomalous ones are introduced.¹

In this paper we develop an alternative model-theoretic interpretation of defaults, called *conditional entailment*, which finally closes the gap between extensional and conditional interpretations and which exhibits the best features of both. Conditional entailment is closely related to prioritized circumscription except that priorities among defaults are not provided by the user but are automatically extracted from the knowledge base. Conditional entailment thus shows that the difference between the conditional (probabilistic or model-theoretic) and extensional interpretations can be reduced to a particular *ordering* on defaults; an idea previously suggested by Pearl [32].

This paper is organized as follows. First we briefly review two conditional interpretations of defaults, one based on probabilities and the other on model-preference (Section 2). We then define the semantics of conditional entailment (Section 3), consider illustrative examples (Section 4), and develop a proof theory (Section 5). The proof theory permits us to compute conditional entailment by considering competing arguments of varying strength; a sketch of an implementation is presented in Section 6. We then discuss related work in Section 7, and summarize the main ideas and some of their limitations in Section 8.

2. Conditional interpretations of defaults

This section provides a brief survey of recent work on conditional interpretations of defaults.² Conditional interpretations assume that default theories are structured into two components: a *background context* K containing generic information about the domain of interest, and an *evidence set* E containing information specific to the particular situation at hand. Intuitively, K contains the relevant *rules*, while E contains the relevant *facts* or *observations*. For instance, in the canonical “birds fly, penguins don’t” example, we would include the strict and defeasible generics, “penguins are birds”, “birds fly”, and “penguins don’t fly” in K , leaving facts such as “Tweety is a bird”, “Tim weighs three pounds”, etc. in the evidence set E .

The distinction between background and evidence, although not widely acknowledged, appears in one way or another in most systems that handle “specificity” preferences among defaults. For example, the distinction appears

¹ See Section 6.

² See [7] for details and proofs.

in [5, 35, 41] as necessary versus contingent facts, in [29] as strict rules versus facts, and in [11] as reified predicates (rules) versus terms (facts). From a probabilistic point of view, as we will see, this distinction takes a natural form: including a sentence s in the background amounts *constraining its prior probability* to one, while including it in the evidence set amounts *conditionalizing* on it.

The background context K of a default theory $T = \langle K, E \rangle$ also has two components: a set L of sentences, and a set D of defaults. Defaults are denoted by expressions of the form $p \rightarrow q$, where p and q are sentences denoting the default antecedent and consequent respectively. The expression $\text{dog}(\text{fido}) \rightarrow \text{can_bark}(\text{fido})$, for instance, represents a default stating that “normally, if Fido is a dog, Fido can bark”. We use *default schemas* of the form $p(x) \rightarrow q(x)$, where p and q are wffs with free variables among those of x , to denote the collection of defaults $p(a) \rightarrow q(a)$ that results from replacing x by all tuples a of ground terms in the language.

In the probabilistic interpretation of defaults [1, 7, 30], the background K is viewed as imposing a constraint over probability distributions which are later conditioned on the evidence E . We call these distributions ε -admissible, where ε is a real parameter, and define them as follows:

Definition 2.1. A probability distribution P_K is ε -admissible relative to a background $K = \langle L, D \rangle$ when P_K assigns unit probability to every (strict) sentence s in L , i.e. $P_K(s) = 1$, and probabilities $P_K(q|p) > 1 - \varepsilon$ and $P_K(p) > 0$ to each default $p \rightarrow q$ in D .

In other words, a probability distribution is ε -admissible when it renders the sentences in L *certain*, while leaving a range ε of uncertainty for the defaults in D . If the conditional probability of a proposition p given a body of evidence E approaches one as ε approaches zero, then the proposition is said to be ε -entailed by the theory $T = \langle K, E \rangle$:

Definition 2.2. A proposition p is ε -entailed by a default theory $T = \langle K, E \rangle$ when for any $\varepsilon' > 0$, there exists an $\varepsilon > 0$, such that $P_K(p|E) > 1 - \varepsilon'$ for any ε -admissible probability distribution P_K .

ε -entailment is nonmonotonic relative to E and captures several of the essential aspects of defaults [7, 30]. Effective procedures for testing ε -entailment can be found in [13].

An alternative conditional interpretation of defaults appeals to models instead of probabilities. The role of probability distributions is filled by *preferential model structures* [16, 19, 23, 40]:

Definition 2.3. A *preferential model structure* is a pair $\langle \mathcal{I}, < \rangle$, where \mathcal{I} denotes

a non-empty collection of interpretations, and “ $<$ ” denotes an irreflexive and transitive order relation over \mathcal{I} .

For a particular structure $\langle \mathcal{I}, < \rangle$ and two interpretations M and M' in \mathcal{I} , the notation $M < M'$ is read as saying that M is *preferred* to M' . Moreover, when M is a model of $T = \langle K, E \rangle$ (i.e., a model of both L and E), and \mathcal{I} contains no model of T preferred to M , then M is said to be a *preferred model* of T in that structure. We will also say sometimes that M is a preferred model of the evidence E , leaving the background K implicit.

In the same way that the probabilistic interpretation considers only *admissible* probability distributions, the model-theoretic interpretation considers only *admissible* preferential model structures, viewing defaults as constraints over the preference relation “ $<$ ”:

Definition 2.4. A well-founded³ preferential model structure $\langle \mathcal{I}, < \rangle$ is *admissible* relative to a background $K = \langle L, D \rangle$ iff every interpretation in \mathcal{I} satisfies L , and for every default $p \rightarrow q$ in D ,

- (a) \mathcal{I} contains a model of p , and
- (b) q is true in all preferred models of p .

The preferred models in the admissible structures determine what is *preferentially entailed* by a default theory:

Definition 2.5. A default theory $T = \langle K, E \rangle$ *preferentially entails* (*p-entails*) a proposition p iff p is true in all the preferred models of E in every preferential model structure admissible with K .

Thus, while in the definition of ε -entailment the background K determines the admissible probability distributions which are then conditioned upon the evidence E , in the definition of p-entailment, the background K determines the admissible preferential model structures from which the preferred models of E are selected. Interestingly, as suggested early by Adams [1, 2], and noted recently by Lehmann and Magidor [18], ε -entailment and p-entailment coincide for finite propositional languages, and they accept an elegant and simple proof theory:

Theorem 2.6. For finite propositional languages, if $T = \langle K, E' \rangle$ is a default theory, the following conditions are equivalent:

- (1) T ε -entails q ,
- (2) T p-entails q ,

³ A structure $\langle \mathcal{I}, < \rangle$ is well-founded relative to a background K , if for every theory $T = \langle K, E \rangle$ and every nonpreferred model M of T in \mathcal{I} , \mathcal{I} contains a preferred model M' of T , such that $M' < M$.

- (3) *the expression $E' \vdash_K q$ is derivable from the rules:*
- (3.1) *Defaults: $p \vdash_K q$ if $p \rightarrow q \in D$;*
 - (3.2) *Deduction: if $E, L \vdash p$, then $E \vdash_K p$;*
 - (3.3) *Augmentation: if $E \vdash_K p$ and $E \vdash_K q$, then $E, p \vdash_K q$;*
 - (3.4) *Reduction: if $E \vdash_K p$ and $E, p \vdash_K q$, then $E \vdash_K q$;*
 - (3.5) *Disjunction: if $E, p \vdash_K r$ and $E, q \vdash_K r$, then $E, p \vee q \vdash_K r$.*

Some of these rules appear, in different form, in several logics of conditionals (see [28]), and in a “minimal” nonmonotonic logic proposed by Gabbay [6]. Pearl [31] has referred to rules (3.1)–(3.5) as a default reasoning *core*, suggesting that they constitute a basic set of principles that any reasonable account of defaults must obey. Rules (3.3) and (3.4) express the property called *cumulativity* by Makinson [23], namely, that theories $T = \langle K, E \rangle$ and $T' = \langle K, E + \{p\} \rangle$ should give rise to the same conclusions if p is a defeasible consequence of T . For examples illustrating the inferences captured by rules (3.1)–(3.5) see [9, 30].

As discussed in the introduction, rules (3.1)–(3.5) capture patterns of inference that escape traditional nonmonotonic logics, but miss patterns involving independence assumptions which the latter do capture (e.g., concluding that a *red* bird flies, given that birds fly). The extension of the core that we will pursue in this paper provides the benefits of the two approaches and originates from the following simple observation.

Let us say that we encode defaults $p \rightarrow q$ as “abnormality” sentences of the form $p \wedge \neg \text{ab}_i \Rightarrow q$ with unique abnormality predicates ab_i . Namely, we take a default theory $T = \langle K, E \rangle$ with a background $K = \langle L, D \rangle$ and map it into a new theory $T' = \langle K', E \rangle$ with a background $K' = \langle L', D' \rangle$, where D' is empty and L' comprises the formulas in L as well as the abnormality sentences corresponding to the defaults in D . Furthermore, let us say that we define the consequences of T in terms of the models of T' which are *minimal* in the set of abnormalities they sanction (as in [25]).

By restricting attention to models that are minimal we automatically capture the “independencies” that characterize extensional nonmonotonic formalisms. On the other hand, as with other extensional formalisms, the minimal model semantics will miss desirable patterns such as specificity preferences, that are captured by both ε -entailment and p-entailment. Yet, since for finite propositional languages ε -entailment and p-entailment are completely characterized by rules (3.1)–(3.5), it is thus natural to ask which rules among (3.1)–(3.5) are violated by the minimal model semantics.

It turns out that under some reasonable assumptions, rule (3.1) is the only one that escapes the minimal model semantics. In other words, the minimization of “abnormality” renders a semantics that complies with all the rules sanctioned by ε -entailment and p-entailment, except rule (3.1). Thus, if we could modify the minimal model semantics so as to render rule (3.1) valid, we

would reap the benefits of both extensional and conditional interpretations. This is indeed what we set out to do in the next section. The mechanism for validating rule (3.1) will be to impose certain *priorities* among abnormalities, reflecting the structure of K . Then, by considering only (minimal) models that permit abnormalities of lower priority, we will get an entailment relation that is stronger than both p-entailment and minimal models, and that captures both the conditional and nonmonotonic aspects of defaults. We call this entailment relation *conditional entailment*.

3. Conditional entailment

3.1. Preliminary definitions

As a matter of convenience, conditional entailment will be defined over the class of default theories $T = \langle K, E \rangle$ in which each default in K is of the form $p \rightarrow \delta$, where δ is a distinguished assumption of “normality”. Arbitrary default theories $T' = \langle K', E \rangle$ can be encoded in this format by replacing each default schema $p(x) \rightarrow q(x)$ in K' by a sentence $p(x) \wedge \delta_i(x) \Rightarrow q(x)$ ⁴ and a default schema $p(x) \rightarrow \delta_i(x)$, where δ_i denotes a new and unique assumption predicate which summarizes the normality conditions required for concluding $q(x)$ from $p(x)$. This way of encoding defaults is similar to McCarthy’s [25] “abnormality” formulation and Poole’s [34] default naming conventions. Our choice of assumptions as primitive objects is just a matter of convenience. Unlike McCarthy and Poole, however, we retain the default schema $p(x) \rightarrow \delta_i(x)$ in addition to the sentence $p(x) \wedge \delta_i(x) \Rightarrow q(x)$, as such schemas will tell us the contexts in which it is safe to adopt assumptions of the form $\delta_i(a)$ for ground terms a . Recall that under interpretations that validate rule (3.1) a default schema $p(x) \rightarrow \delta_i(x)$ allows us to derive the assumption $\delta_i(a)$ in the context $T = \langle K, \{p(a)\} \rangle$ regardless of what other information is in K . This ability to unconditionally assert $\delta_i(a)$ in T would be lost, however, if we didn’t have the default schema $p(x) \rightarrow \delta_i(x)$.⁵

We will call the theories structured in the above format, *assumption-based default theories*. In the context of these theories, conditional entailment seeks to *maximize* assumptions, while ensuring that rule (3.1) is valid.

We will represent assumptions by literals of the form $\delta_i(a)$, where a is a tuple of ground terms, and will use the symbols δ, δ', \dots , as variables ranging over the assumptions in the underlying language \mathcal{L} . $\Delta_{\mathcal{L}}$ will refer to the collection of all such assumptions.

⁴ Free variables are assumed to be universally quantified; thus $p(x) \wedge \delta_i(x) \Rightarrow q(x)$ stands for the closed formula $\forall x. p(x) \wedge \delta_i(x) \Rightarrow q(x)$.

⁵ From a representational point of view, this choice does not impose an additional burden on the user, as we may as well assume, as we will often do, that for each sentence $p(x) \wedge \delta_i(x) \Rightarrow q(x)$ in K , K contains a default schema $p(x) \rightarrow \delta_i(x)$.

Given a default theory $T = \langle K, E \rangle$ with background $K = \langle L, D \rangle$, we identify the models of T as the interpretations that satisfy the sentences in both L and E . We will also refer to the set of assumptions violated by an interpretation M as the *gap* of M , and denote it as $\Delta[M]$.

3.2. Model theory

Conditional entailment is a specialization of preferential entailment for the case in which the *preference order on interpretations* is determined by a given *priority ordering on assumptions*. We call the resulting structures *prioritized preferential structures* and define them as follows:

Definition 3.1. A *prioritized preferential structure* is a quadruple $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, < \rangle$, where $\mathcal{I}_{\mathcal{L}}$ stands for the set of interpretations over the underlying language \mathcal{L} , $\Delta_{\mathcal{L}}$ stands for the set of assumptions in \mathcal{L} , “ $<$ ” stands for an irreflexive and transitive priority relation over $\Delta_{\mathcal{L}}$, and “ $<$ ” is a binary relation over $\mathcal{I}_{\mathcal{L}}$, such that for two interpretations M and M' , $M < M'$ holds iff $\Delta[M] \neq \Delta[M']$ and for every assumption δ in $\Delta[M] - \Delta[M']$ there exists an assumption δ' in $\Delta[M'] - \Delta[M]$ such that $\delta < \delta'$.

We will require that priority orderings do not contain infinite ascending chains $\delta_1 < \delta_2 < \delta_3 < \dots$. This will guarantee that the order on interpretations induced by a given priority ordering is both strict and partial:

Lemma 3.2. If the quadruple $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, < \rangle$ is a prioritized preferential structure, then the pair $\langle \mathcal{I}_{\mathcal{L}}, < \rangle$ is a preferential model structure.

The resulting order “ $<$ ” on interpretations regards the relation $\delta < \delta'$ as a preference to sustain the assumption δ' over the assumption δ in cases of conflict. A similar mapping from *predicate* priorities to preferences occurs in Przymusiński’s [36] characterization of the perfect model semantics of logic programs and McCarthy’s [25] prioritized circumscription (see also [20]).

Figure 1 illustrates the preference on two interpretations M and M' de-

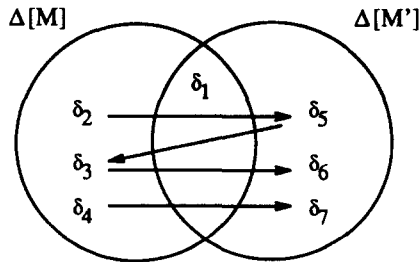


Fig. 1. Preference on interpretations in prioritized structures: $M < M'$.

terminated by an arbitrary priority ordering on assumptions. An arrow connecting an assumption δ_i to an assumption δ_j expresses that δ_i has lower priority than δ_j , i.e., $\delta_i < \delta_j$. To check then whether M is preferred to M' , it is sufficient to check that each assumption δ in $\Delta[M] - \Delta[M']$ is linked by an arrow to an assumption δ' in $\Delta[M'] - \Delta[M]$ (provided that $\Delta[M] \neq \Delta[M']$). Note that assumptions violated by both M and M' (e.g., δ_1) play no role in determining the preferences between M and M' .

An important feature of the preference on interpretations determined by any prioritization is that preferred models M are guaranteed to be *minimal* in the set $\Delta[M]$ of assumptions that they violate:

Lemma 3.3. *In any prioritized preferential structure, if M is a preferred model of a theory T then M is minimal in $\Delta[M]$, i.e., there is no model M' of T such that $\Delta[M'] \subset \Delta[M]$.*

Furthermore, models M and M' with identical gaps are not distinguished; if M is a preferred model of T , for example, so will be M' if $\Delta[M'] = \Delta[M]$. For this reason, we will often talk about *classes of models*, where a class is non-empty collection of models with a common gap, and say that a class \mathcal{C}_1 is preferred to a class \mathcal{C}_2 if any one model in \mathcal{C}_1 is preferred to any one model in \mathcal{C}_2 . Since the preference on interpretations depends on the gaps of the interpretations only, this will guarantee that any *other* model in \mathcal{C}_1 will be preferred to any *other* model in \mathcal{C}_2 as well.

While the minimality of preferred models endows conditional entailment with the features common to traditional nonmonotonic logics, the focus on a particular class of priority orderings which reflect the structure of K will account for the desirable features of conditional interpretations.

As discussed in Section 2, the aim of imposing a priority ordering on assumptions will be to validate rule (3.1), which permits us to derive the consequent δ of a default $p \rightarrow \delta$, given the evidence $E = \{p\}$. This rule does not necessarily hold in all minimal models because in the context $T = \langle K, \{p\} \rangle$, there may be sets of assumptions Δ' which are logically inconsistent with δ . Thus we can choose a maximal assumption set by adopting Δ' and rejecting δ , or by adopting δ and rejecting some assumption in Δ' . If the first option is chosen, however, rule (3.1) is violated. The effect of the priorities will be to make the second option preferred over the first one.

Let us say that a set of assumptions Δ is *in conflict with a default* $p \rightarrow \delta$ in K if Δ and δ are logically inconsistent in the context $T = \langle K, \{p\} \rangle$, i.e., $L, p, \Delta \vdash \neg \delta$. Then, we will simply require that any such Δ contain an assumption δ' weaker than δ , i.e., $\delta' < \delta$:

Definition 3.4. A priority order “<” over $\Delta_{\mathcal{L}}$ is *admissible* relative to a

background context K iff every set Δ of assumptions in conflict with a default $p \rightarrow \delta$ in K contains an assumption δ' such that $\delta' < \delta$.

Thus, if the priority ordering “ $<$ ” in the structure $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, < \rangle$ is *admissible* and $p \rightarrow \delta$ is a default in K , we are guaranteeing that for any model M' of $T = \langle K, \{p\} \rangle$ where δ does not hold there would be another model M of T preferred to M' where δ *does* hold. Roughly, M can be constructed from M' by making δ true, and the assumptions which are weaker than δ , false. Indeed, the absence of one such model M would imply the presence of a set of assumptions which includes δ but no assumption weaker than δ , which is logically inconsistent. This, however, would contradict the admissibility of “ $<$ ”.

Priorities which are admissible can also be thought as reflecting the “specificity” of defaults. For example, two defaults $p_1 \rightarrow q$ and $p_2 \rightarrow \neg q$, encoded by means of assumptions as sentences $p_1 \wedge \delta_1 \Rightarrow q$ and $p_2 \wedge \delta_2 \Rightarrow \neg q$ and defaults $p_1 \rightarrow \delta_1$ and $p_2 \rightarrow \delta_2$, will automatically constrain the priority of δ_1 to be higher than the priority of δ_2 if $p_1 \rightarrow q$ is more “specific” than $p_2 \rightarrow \neg q$, i.e., $L \vdash (p_1 \Rightarrow p_2)$. This is because, the assumption set $\Delta = \{\delta_2\}$ will be *in conflict with the default* $p_1 \rightarrow \delta_1$, and thus $\delta_2 < \delta_1$ would be required to hold if the ordering “ \leq ” is admissible.

We will define the *admissible prioritized structures* as the structures induced by admissible priority orderings:

Definition 3.5. A prioritized structure $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, < \rangle$ is admissible with a background $K = \langle L, D \rangle$ iff the priority ordering “ $<$ ” is admissible with K .

Like preferential entailment, *conditional entailment* is defined in terms of the preferred models of the admissible structures, except that now the admissible *prioritized* preferential structures are considered:

Definition 3.6. A proposition q is *conditionally entailed* by a default theory $T = \langle K, E \rangle$, iff q holds in all the preferred models of T in every *prioritized* preferential structure admissible with K .

Conditional entailment combines the two target notions: minimality and conditionality. Indeed, for *finite propositional languages* the following result can be shown:⁶

⁶ The restriction to finite propositional languages is required in order to guarantee that the preferential structure determined by a given priority ordering is well-founded, and thus, validates rules (3.3) and (3.4). That particular restriction, however, can be relaxed (e.g., we may just assume a finite number of assumptions), and it is *not* required for conditional entailment to be well-behaved (see below).

Theorem 3.7. *If an assumption-based default theory $T = \langle K, E \rangle$ preferentially entails a proposition p , then T also conditionally entails p .*

Namely, conditional entailment captures all inference patterns sanctioned by preferential entailment and ϵ -entailment (see Theorem 2.6), and as we will see, other patterns as well. For this result to be meaningful, though, we also need to show that conditional entailment remains well-behaved as long as preferential entailment is well-behaved. Namely, if preferential entailment sanctions a proposition q , we not only want conditional entailment to sanction q , but also *not* to sanction its negation. To guarantee this it is sufficient to show that if a background K is *preferentially consistent* (i.e., there is a preferential structure admissible with K) then K will also be *conditionally consistent*, i.e., there will be a prioritized structure admissible with K . This will be true as long as the use of assumptions in $K = \langle L, D \rangle$ complies with a simple syntactic restriction, namely, that assumption predicates δ_i are used *only* to encode defaults schemas $p_i(x) \rightarrow q_i(x)$ as sentences $p_i(x) \wedge \delta_i(x) \Rightarrow q_i(x)$ and defaults $p_i(x) \rightarrow \delta_i(x)$. If this restriction is enforced, an assumption $\delta_i(a)$ will be logically equivalent to the material conditional $p_i(a) \Rightarrow q_i(a)$ in all minimal models.

If we say that a background K is *pure* when it complies with this last restriction, and that a default theory T is *consistent* when it does not entail both a proposition and its negation, the following result obtains:⁷

Theorem 3.8. *For a pure background K , an assumption-based default theory $T = \langle K, E \rangle$ is consistent relative to preferential entailment only if it is consistent relative to conditional entailment.*

Another way to look at this result, is that if we encode defaults by means of assumptions, conditional entailment will not derive a contradiction as long as no contradiction is derivable from rules (3.1)–(3.5). From the example above, for instance, if *each* of the defaults $p_1 \rightarrow q$ and $p_2 \rightarrow \neg q$ is more specific than the other, i.e., $L \vdash (p_1 \equiv p_2)$, both $\delta_2 < \delta_1$ and $\delta_1 < \delta_2$ would have to hold, which is impossible. Thus, any such theory will be inconsistent, and any pair of contradictory propositions will be conditionally entailed. To see that a similar contradiction can be derived from rules (3.1)–(3.5), consider the expressions $p_1 \vdash_K q$ and $p_2 \vdash_K \neg q$. Both are derivable by means of rule (3.1) (defaults) and deductive closure. Furthermore, since by rule (3.2) (deduction) we can obtain $p_1 \vdash_K p_2$ and $p_2 \vdash_K p_1$, by rule (3.3) (augmentation) we get both $p_1, p_2 \vdash_K q$ and $p_1, p_2 \vdash_K \neg q$, which are also contradictory.

⁷ An example of a non-pure background is one containing two defaults $p_1 \rightarrow \delta_1$ and $p_2 \rightarrow \delta_2$ together with the sentence $\neg(\delta_1 \wedge \delta_2)$. This background is preferentially consistent but conditionally inconsistent. The sentence $\neg(\delta_1 \wedge \delta_2)$, however, does not originate from the encoding of any default, and thus violates the condition of “purity”. For more details and a proof of this result, see [7].

4. Examples

In this section we will illustrate how conditional entailment works on a number of examples. For each of the examples, we will first determine the admissible priorities by looking at the defaults in K , and then consider the preferred classes of models for different contexts $T = \langle K, E \rangle$ of interest. The determination of priorities can thus be understood as a “compile-time” operation whose results are good for any particular set of observations.

In the presentation below we will introduce some abbreviations. First, we will use the notation $E \vdash_K p$, for a background $K = \langle L, D \rangle$, as an abbreviation of $E, L \vdash p$. Thus, a set of assumptions Δ' will be in conflict with a default $p \rightarrow \delta$ in K , if $p, \Delta' \vdash_K \neg \delta$. When this happens, we will also say that δ *dominates* Δ' . As we will find in Section 6, if δ dominates all the sets of assumptions which support its negation, then δ will be conditionally entailed. Note that dominance patterns are determined by the information in the background context K only, but they apply to arbitrary contexts $T = \langle K, E \rangle$.

As a final abbreviation, we will write $\Delta' < \delta$ to express that “there exists a δ' in Δ' such that $\delta' < \delta$ ”. The admissibility of the priority order “ $<$ ” with respect to K thus amounts to testing whether the relation $\Delta' < \delta$ holds when δ dominates Δ' . Clearly this test needs be applied to *minimal* Δ' only.

Example 4.1. Consider a background context $K = \langle L, D \rangle$ with sentences (Fig. 2):

$$\begin{aligned} b(x) \wedge \delta_1(x) &\Rightarrow f(x) , \\ p(x) \wedge \delta_2(x) &\Rightarrow \neg f(x) , \\ p(x) &\Rightarrow b(x) , \\ r(x) &\Rightarrow b(x) , \end{aligned}$$

and defaults $b(x) \rightarrow \delta_1(x)$ and $p(x) \rightarrow \delta_2(x)$. We can read the symbols b , f , p , and r , as standing for the predicates “bird”, “fly”, “penguin”, and “red bird”, respectively.

From the definition and the notions introduced above, a priority ordering “ $<$ ” will be admissible with K if the relation $\Delta < \delta$ holds for any minimal assumption set Δ *dominated* by δ . First note that there is no assumption set Δ in conflict with instances of the default schema $b(x) \rightarrow \delta_1(x)$, since for any

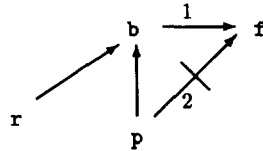


Fig. 2. Strict specificity.

ground term a there are interpretations which satisfy K and $b(a)$ and which violate no assumption. As a result, assumptions of the form $\delta_1(a)$ do not dominate any assumption set, and thus impose no constraints on the admissible priority orderings. Assumptions of the form $\delta_2(a)$, on the other hand, dominate a single minimal set $\Delta = \{\delta_1(a)\}$, as $\Delta = \{\delta_1(a)\}$ is in conflict with the default $p(a) \rightarrow \delta_2(a)$, i.e., $p(a), \Delta \vdash_K \neg \delta_2(a)$. As a result, a priority order “ $<$ ” will be admissible with K iff the relation $\delta_1(a) < \delta_2(a)$ is satisfied for every ground term a in the language. We also write in these cases $\delta_1(x) < \delta_2(x)$. As in the example above, these priorities reflect the fact that defaults about penguins are more “specific” than defaults about birds.

Provided with this characterization of the structures admissible with K , we can now turn to analyze the propositions which are conditionally entailed in the different contexts $T = \langle K, E \rangle$ of interest. For example, for an individual Tim, denoted by t , the preferred models of $b(t)$ in K are the models which violate no assumption. As a result, both $\delta_1(t)$ and $\delta_2(t)$ are conditionally entailed by $b(t)$, as are the propositions $f(t)$ and $\neg p(t)$ that they support (i.e., Tim is presumed to be a normal flying bird, and therefore, not a penguin).

A different scenario arises if we consider the evidence $p(t)$ instead of $b(t)$. In this case, every interpretation satisfying the evidence and the background context is forced to render one of assumptions $\delta_1(t)$ or $\delta_2(t)$ false. Thus, two classes of minimal models arise: a class \mathcal{C}_1 comprised of the models M_1 which violate the assumption $\delta_1(t)$, and a class \mathcal{C}_2 comprised of the models M_2 which violate the assumption $\delta_2(t)$. However, models in the former class are preferred to models in the latter class since

$$\Delta[M_2] - \Delta[M_1] = \{\delta_2(t)\},$$

$$\Delta[M_1] - \Delta[M_2] = \{\delta_1(t)\},$$

$$\delta_1(t) < \delta_2(t)$$

(see Definition 3.1). It follows then, that \mathcal{C}_1 represents the class of preferred models of $p(t)$ in K , and therefore, that the propositions $\delta_2(t)$ and $\neg f(t)$ are conditionally entailed. Similar conclusions are legitimized in this context by preferential entailment and ε -entailment. Indeed, $p(t) \vdash_K \neg f(t)$ follows immediately from the application of rule (3.1).

Consider now instead the scenario in which the target context is enhanced with the information that Tim is a red bird, i.e., $T' = \langle K, E' \rangle$, with $E' = \{p(t), r(t)\}$. In this case, neither ε -entailment nor p -entailment constrain the preferred models of T' . Conditional entailment, on the other hand, guarantees that the preferred models of T' are minimal, and therefore, that they belong to one of the two classes \mathcal{C}_1 and \mathcal{C}_2 of minimal models, where \mathcal{C}_i stands for the collection of models M of T' which violate the assumption $\delta_i(t)$, i.e., $\Delta[M] = \{\delta_i(t)\}$. As we showed above, however, models in \mathcal{C}_1 are preferred to

models in \mathcal{C}_2 .⁸ As a result, the assumption $\delta_2(\mathbf{t})$ and the proposition $\neg f(\mathbf{t})$ are conditionally entailed by T' . Note that, on the other hand, neither proposition is legitimized by either ε -entailment, p -entailment, or minimal models.

The example above illustrates different contexts built on top of a background which forces every admissible priority ordering “ $<$ ” to satisfy the relation $\delta_1(a) < \delta_2(a)$ for all ground terms a in the language. This means that every admissible priority relation “ $<$ ” must include all tuples of the form $\langle \delta_1(a), \delta_2(a) \rangle$. Such relations may include additional tuples as well, e.g. $\langle \delta_1(a), \delta_2(b) \rangle$, but those tuples are not necessary for the relations to be admissible. We will say that an admissible priority relation is *minimal* when no set of tuples can be deleted without violating the admissibility constraints. For instance, in the example above, there is a *single* minimal admissible ordering which includes *all and only* the tuples of the form $\langle \delta_1(a), \delta_2(a) \rangle$ for ground atoms a . It is natural to ask then whether conditional entailment can be computed by restricting attention to *minimal* admissible priority orderings only. The answer is yes. Indeed, if we can obtain an admissible priority ordering “ $<$ ” by deleting certain tuples from an admissible priority ordering “ $<'$ ”, the preferred models in the structure $\langle \mathcal{I}_{\mathcal{F}}, <', \Delta_{\mathcal{F}}, <' \rangle$ will be a subset of the preferred models of the structure $\langle \mathcal{I}_{\mathcal{F}}, <, \Delta_{\mathcal{F}}, < \rangle$. Thus, if we say that an admissible prioritized preferential structure $\langle \mathcal{I}_{\mathcal{F}}, <, \Delta_{\mathcal{F}}, < \rangle$ is *minimal* if the relation “ $<$ ” is a minimal admissible priority ordering, the following alternative characterization of conditional entailment results:

Lemma 4.2. *A proposition q is conditionally entailed by a default theory $T = \langle K, E \rangle$ iff q holds in all preferred models of T of every minimal prioritized preferential structure admissible with K .*

In the example above, we can thus compute conditional entailment by considering a *single* structure $\langle \mathcal{I}_{\mathcal{F}}, <, \Delta_{\mathcal{F}}, < \rangle$, where the priority ordering is such that $\delta < \delta'$ holds iff $\delta = \delta_1(a)$ and $\delta' = \delta_2(a)$ for some ground term a in the language. Often, however, multiple minimal structures need to be considered (see Example 4.4 below).

Example 4.3. We now consider a slightly different background K comprising the sentences:

$$\begin{aligned} a \wedge \delta_1 &\Rightarrow w, & u \wedge \delta_3 &\Rightarrow a, \\ u \wedge \delta_2 &\Rightarrow \neg w, & f \wedge \delta_4 &\Rightarrow a. \end{aligned}$$

⁸ These two classes do not contain the same models as in the context above, yet, since they possess the same gaps, and the preference relation depends only on such gaps, they are ranked in the same way.

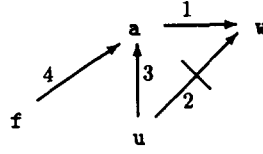


Fig. 3. Default specificity.

(see Fig. 3). As usual, we assume that for each sentence of the form $p_i \wedge \delta_i \Rightarrow q_i$, K also contains a default of the form $p_i \rightarrow \delta_i$; each such pair thus encoding the default $p_i \rightarrow q_i$.

The background K has the same structure as Example 4.1 (Fig. 2) except that all rules are now defeasible. These rules can be understood as expressing, for example, that “adults work”, “university students do not work”, “university students are adults”, and “Frank Sinatra fans are adults”.

There are two relevant dominance relations in this background context. First, the assumption δ_2 dominates the set $\Delta = \{\delta_1, \delta_3\}$ as Δ is in conflict with the default $u \rightarrow \delta_2$, i.e., $u, \Delta \vdash_K \neg \delta_2$. Likewise, the assumption δ_3 dominates the set $\{\delta_1, \delta_2\}$. Thus, any priority ordering “ $<$ ” admissible with K must be such that both relations $\{\delta_1, \delta_3\} < \delta_2$ and $\{\delta_1, \delta_2\} < \delta_3$ hold. Moreover, due to the asymmetric and transitive character of priority orderings, such constraints imply $\delta_1 < \delta_2$ and $\delta_1 < \delta_3$. To show that this is the case, let us first assume $\delta_2 < \delta_3$. Then, by the asymmetry of priority orderings we must have $\delta_3 \not< \delta_2$, and therefore, from the constraints above, $\delta_1 < \delta_2$. Now assume $\delta_2 \not< \delta_3$. If $\delta_1 < \delta_2$ does not hold, the constraints above imply $\delta_1 < \delta_3$ and $\delta_3 < \delta_2$, contradicting the transitivity of “ $<$ ”. Thus, in either case the relation $\delta_1 < \delta_2$ must hold. By symmetry, we conclude that $\delta_1 < \delta_3$ also holds.⁹

With this space of admissible priority orderings, let us first consider a context $T = \langle K, E \rangle$, with $E = \{f\}$. Since there is an interpretation that satisfies T and every assumption in the language, the single preferred class in every admissible prioritized structure is the class of models which violate no assumption. In particular, the assumptions δ_1 and δ_4 are conditionally entailed by T , as are the propositions a and w that they support. These inferences involve default chaining, a pattern which is not sanctioned by either ε -entailment or p -entailment.

A different situation arises when the proposition u is observed. The context $T' = \langle K, E' \rangle$, with $E' = \{f, u\}$, gives rise to three classes \mathcal{C}_i of minimal models M_i , $i = 1, 2, 3$, with gaps $\Delta[M_1] = \{\delta_1\}$, $\Delta[M_2] = \{\delta_2\}$, and $\Delta[M_3] = \{\delta_3, \delta_4\}$. However, since $\delta_1 < \delta_2$ and $\delta_1 < \delta_3$ hold, any model M_1 in \mathcal{C}_1 is preferred to any model M_2 in \mathcal{C}_2 and any model M_3 in \mathcal{C}_3 . Hence \mathcal{C}_1 is the class of preferred models of T' , and therefore, all assumptions other than δ_1 are conditionally entailed by T' , as are the propositions that those assumptions support.

⁹ Note that on “specificity” grounds the priority of δ_3 over δ_1 does not appear justified. However, without it, the conclusion a from u would not be warranted.

Example 4.4. Let K represent the hierarchy depicted in Fig. 4:

$$\begin{aligned} a \wedge \delta_1 &\Rightarrow b, \\ a \wedge \delta_2 &\Rightarrow d, \\ b \wedge \delta_3 &\Rightarrow c, \\ c \wedge \delta_4 &\Rightarrow \neg d. \end{aligned}$$

As usual we assume that for each sentence $p_i \wedge \delta_i \Rightarrow q_i$, K also contains a default of the form $p_i \rightarrow \delta_i$.

In order to determine the admissible priority orderings we have to identify first the relevant dominance patterns. There are two such patterns in this background: the assumption δ_1 dominates the assumption set $\Delta_1 = \{\delta_2, \delta_3, \delta_4\}$, while the assumption δ_2 dominates the assumption set $\Delta_2 = \{\delta_1, \delta_3, \delta_4\}$. This is because Δ_1 is in conflict with the default $a \rightarrow \delta_1$, and Δ_2 is in conflict with the default $a \rightarrow \delta_2$. Thus, every priority ordering admissible with K must be such that both $\{\delta_2, \delta_3, \delta_4\} < \delta_1$ and $\{\delta_1, \delta_3, \delta_4\} < \delta_2$ hold. Moreover, from these two constraints and the fact that priority orderings are asymmetric and transitive, it is possible to show as above that every admissible ordering must also comply with the simplified constraints $\{\delta_3, \delta_4\} < \delta_1$ and $\{\delta_3, \delta_4\} < \delta_2$. These constraints, however, cannot be simplified further.

Let us consider now a body of evidence $E = \{a\}$. The theory $T = \langle K, E \rangle$ gives rise to four classes \mathcal{C}_i of minimal models, each one with an associated gap $\{\delta_i\}$, $i = 1, \dots, 4$. We show first that the preferred models of T in any admissible prioritized structure $\tau = \langle \mathcal{J}_\tau, <, \Delta_\tau, < \rangle$ are contained in \mathcal{C}_3 and \mathcal{C}_4 . Let M_i be a model in \mathcal{C}_i , for $i = 1, \dots, 4$, and assume that the relation $\delta_3 < \delta_1$ holds. Then, since $\Delta[M_3] - \Delta[M_1] = \{\delta_3\}$ and $\Delta[M_1] - \Delta[M_3] = \{\delta_1\}$, M_3 must be preferred to M_1 in τ and, therefore, M_1 is not a preferred model of T in τ . Assume now otherwise, that $\delta_3 \not< \delta_1$. Then from the constraint $\{\delta_3, \delta_4\} < \delta_1$ above, the relation $\delta_4 < \delta_1$ must be true. By similar arguments, it follows then that M_4 is preferred to M_1 in τ , and therefore, that M_1 , again, is not a preferred model of T in τ . Replacing M_1 by M_2 , we obtain similarly that M_2 is not a preferred model of T either. Furthermore, since neither δ_3 has higher priority than δ_4 , nor vice versa, no class among \mathcal{C}_3 and \mathcal{C}_4 is preferred to the other, and thus, \mathcal{C}_3 and \mathcal{C}_4 are the preferred classes of T . As a result, the

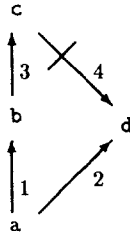


Fig. 4. Disjunctive constraints.

assumptions δ_1 and δ_2 are conditionally entailed by T , and so are the propositions b and d .

It is worth noticing in the last example that conditional entailment does not subsume inheritance reasoning. For example, most inheritance accounts would sanction the assumption δ_3 and the proposition c in the context T but neither conclusion is sanctioned by conditional entailment. The reason is that the *intended* class of models M_4 which only violate the assumption δ_4 is not preferred to the class of models M_3 which only violate the assumption δ_3 . We could in principle modify the way priorities are selected to capture these inferences; however, the value of conditional entailment as an account of default reasoning is to show what the combination of minimality and conditional considerations can and cannot account for. Examples as the one above show that default reasoning involves aspects other than minimality and conditionality. We will say more about this topic in Section 7 (see also [7]).

5. Proof theory

Conditional entailment provides a characterization of the propositions entailed by a given default theory but does not provide effective methods for computing them. In this section and in the next we will focus on such methods. We will develop a number of *syntactic* criteria for testing conditional entailment some of which are amenable to implementation in ATMS-type of systems.

As the proof theory of classical deduction is structured around the notion of *proofs*, the proof theory of conditional entailment is structured around the notion of *arguments* [22, 33]. An argument in a theory $T = \langle K, E \rangle$ is a set Δ of assumptions which is logically consistent with T , i.e., $E, \Delta \not\vdash_K \text{false}$. Moreover, Δ will be an argument *for* a proposition p if $E, \Delta \vdash_K p$, and an argument *against* p if $E, \Delta \vdash_K \neg p$. In the former case we will also say that Δ *supports* p . If Δ is not logically consistent with T , we will say that Δ is a *conflict set*. *Two arguments are in conflict* when their union is a conflict set. Moreover, an assumption δ is *contradicted* if it belongs to a *minimal* conflict set. This is equivalent to saying that δ is contradicted if there is an argument which supports the *negation* of δ . For the results below we will assume that we are given a theory $T = \langle K, E \rangle$ which gives rise to a *finite* number of contradicted assumptions. This will guarantee that the theory T will be well-founded relative to any structure $\langle \mathcal{F}, <, \Delta_{\mathcal{F}}, < \rangle$; namely, for any nonpreferred model M' of T there will be a *preferred* model M of T such that $M < M'$. Other restrictions could be used for this purpose but this one is sufficiently general for practical purposes.¹⁰

¹⁰ A finite number of contradicted assumptions means simply that all except a finite number of defaults will be applicable in the scenarios under consideration.

The first assertability condition, which permits us to derive an assumption δ that is not contradicted, is a simple consequence of the minimality of preferred models within the class of prioritized structures:

Lemma 5.1. *An assumption is conditionally entailed if there are no arguments against it.*

A similar condition is both sound and complete for circumscription in the context of “abnormality” theories [25], provided that T includes the *unique names* and *domain closure* axioms, and that “abnormality” is circumscribed at the expense of all other predicates (assumptions are then identified with “normality” literals $\neg ab_i(a)$ [10]).

In the context of conditional entailment, however, such a condition is too weak, since an assumption may face counterarguments and still be entailed. The “birds fly—penguins don’t” example above provides one such case. If Tim is a penguin, the assumption $\delta_2(t)$: “Tim does not fly, because it is a normal penguin” faces the counterargument $\delta_1(t)$: “Tim flies, because it is a normal bird”; yet $\delta_2(t)$ is conditionally entailed but $\delta_1(t)$ is not.

In order to capture these conclusions by syntactic means, we need to consider the priority orderings determined by K . Indeed, the reason the assumption $\delta_2(t)$ is entailed in spite of the conflict with $\delta_1(t)$ is because its priority is higher. The assertability conditions below take these priorities into account. Recall that we write $\Delta < \delta$ as an abbreviation of the expression “ $\exists \delta' \in \Delta$ such that $\delta' < \delta$ ”.

Lemma 5.2. *An assumption δ is conditionally entailed if for every argument Δ against δ and every admissible priority ordering “ $<$ ”, the relation $\Delta < \delta$ holds.*

Note that it is sufficient to consider only the *minimal* arguments Δ against δ ; if $\Delta < \delta$ holds, so will $\Delta' < \delta$ for any superset Δ' of Δ .

The condition introduced by Lemma 5.2 permits us now to handle the example above. Given the evidence $E = \{p(t)\}$ (“Tim is a penguin”), $\Delta = \{\delta_1(t)\}$ is the only (minimal) argument against $\delta_2(t)$, and since $\delta_2(t)$ has a higher priority than $\delta_1(t)$, Lemma 5.2 permits us to derive $\delta_2(t)$, from which “Tim does not fly” follows.

However, Lemma 5.2 is not yet complete relative to conditional entailment. This can be illustrated by converting the *strict* “links” in Example 4.1 into *default* “links”, resulting in a structure analogous to that of Example 4.3 which is depicted in Fig. 5. By arguments similar to those used in Example 4.3, we can show that the assumption $\delta_4(t)$, associated with the defeasible link $r \rightarrow b$ in the figure, is conditionally entailed by the theory $T' = \langle K', E \rangle$, with $E = \{p(t), r(t)\}$. Yet, the condition in Lemma 5.2 does not authorize this

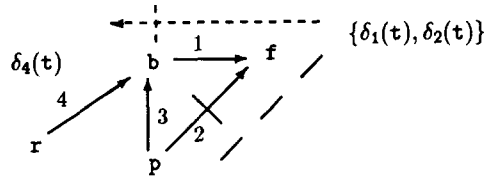


Fig. 5. $\delta_4(t)$ is entailed in spite of conflict with $\{\delta_1(t), \delta_2(t)\}$.

conclusion: $\Delta = \{\delta_1(t), \delta_2(t)\}$ is an argument against $\delta_4(t)$ for which the relation $\Delta < \delta_4(t)$ does not hold.

Intuitively, $\delta_4(t)$ is conditionally entailed by T' in spite of the counterargument Δ , because Δ contains an assumption $\delta_1(t)$ which is *defeated*. Namely, $\delta_1(t)$ is in conflict with two “stronger” assumptions $\delta_2(t)$ and $\delta_3(t)$ which knock out the argument Δ , leaving the assumption $\delta_4(t)$ unchallenged.

In order to account for these interactions, we need to consider multiple conflicts at the same time. For that, some definitions will be handy. We write below $\Delta' < \Delta$ as an abbreviation of the expression “for every δ in Δ , there is a δ' in Δ' such that $\delta' < \delta$ ”.

Definition 5.3. Given a priority ordering “ $<$ ”, an argument Δ *defeats* an argument Δ' if the two arguments are in conflict and the relation $\Delta' < \Delta$ holds. We say in this case that Δ is a *defeater* of Δ' .

Definition 5.4. An argument Δ is *protected* from a conflicting argument Δ' iff for every priority ordering admissible with K , Δ contains a defeater of Δ' .

Intuitively, when an argument Δ is protected from a conflicting argument Δ' it means that Δ is *stronger* than Δ' . For example, for the background K' depicted in Fig. 5, both assumptions $\delta_2(t)$ and $\delta_3(t)$ have a higher priority than $\delta_1(t)$. Hence, in the context $T' = \langle K', E' \rangle$ with $E = \{p(t), r(t)\}$, in which the argument $\Delta_1 = \{\delta_2(t), \delta_3(t)\}$ is in conflict with the argument $\Delta_2 = \{\delta_1(t)\}$, Δ_1 *defeats* Δ_2 . Thus, any argument which includes Δ_1 , like the argument $\Delta_0 = \Delta_1 + \{\delta_4(t)\}$ will be *protected* from Δ_2 .

When an argument is protected from every conflicting argument, we will say the argument is *stable*:

Definition 5.5. An argument is *stable* iff it is protected from every conflicting argument.

As suggested above, a stable argument is better than *any* of its competitors, and propositions supported by stable arguments are conditionally entailed:

Lemma 5.6. A proposition is conditionally entailed if it is supported by a stable argument.

For example, the arguments Δ_1 and Δ_0 above are stable, as the only arguments Δ' in conflict with either of them must contain the assumption $\delta_1(\tau)$ which is defeated by the subargument $\{\delta_2(\tau), \delta_3(\tau)\}$ of both Δ_1 and Δ_0 . Thus, according to Lemma 5.6, the assumptions $\delta_2(\tau)$, $\delta_3(\tau)$, $\delta_4(\tau)$ and all the propositions that they support will be conditionally entailed.

The notion of stable arguments is very powerful and accounts for most of the natural inferences authorized by conditional entailment. Nonetheless, Lemma 5.6 is still not complete. For example in a theory T comprised of the sentences $\delta_1 \Rightarrow \neg \delta_3$, $\delta_2 \Rightarrow \neg \delta_3$ and $\neg \delta_1 \vee \neg \delta_2$, where δ_3 has a lower priority than both δ_1 and δ_2 , the disjunction $\delta_1 \vee \delta_2$ is conditionally entailed even though it is not supported by any stable argument (neither $\Delta_1 = \{\delta_1\}$ nor $\Delta_2 = \{\delta_2\}$ are stable as they are in conflict but neither one is protected from the other).

To account for such conclusions we need to consider *disjunctive* arguments. We will accommodate disjunctive arguments by considering the assertability conditions of disjunctive collection of arguments which we call *covers*. For instance, $C = \{\Delta_1, \Delta_2\}$, with $\Delta_1 = \{\delta_1\}$ and $\Delta_2 = \{\delta_2\}$ will turn out to be a *stable cover*, thus legitimizing the disjunction $\delta_1 \vee \delta_2$ and any proposition supported by it.

We make the notion of *stable covers* precise by refining first the conditions under which an argument is protected:

Definition 5.7. An argument Δ is *strongly protected* from a conflicting argument Δ' if for every subargument Δ'_i of Δ' in conflict with Δ there exists a subargument Δ_i of Δ in conflict with Δ' such that $\Delta'_i < \Delta_i$.

Note that if an argument Δ is protected from a conflicting argument Δ'_1 but is not protected from a conflicting argument Δ'_2 , Δ will *not* be strongly protected from the union $\Delta'_1 + \Delta'_2$ even though Δ will be protected from it. The distinction between the two notions is irrelevant for stable arguments which are *both* protected and strongly protected from every conflicting argument but is needed for *disjunctive* arguments.

Let us refer to a collection of arguments as a *cover*—where a cover is to be understood as the disjunction of the arguments it contains—and let us generalize the notions of conflicts and protection as follows:

Definition 5.8. An argument Δ is *in conflict with a cover* if Δ is in conflict with every argument in the cover.

Definition 5.9. A *cover* is *protected* from a conflicting argument Δ if the cover contains an argument Δ' which is *strongly protected* from Δ .

The conditions under which a *cover* is stable can then be obtained as a generalization of the conditions under which an *argument* is stable. The only

difference is that, for the purpose of completeness, we only consider arguments Δ in conflict with C that have as many assumptions from C as possible. We call such conflicting arguments *definite* as they either include or rebut each of the assumptions which occur in C .

Definition 5.10. A *cover* is *stable* iff it is protected from every definite conflicting argument Δ .¹¹

As expected, the conditions of Lemma 5.6 can be strengthened by replacing stable arguments by stable covers. Furthermore, if we say that a proposition p is supported by a *cover* when it is supported by every argument in the cover, the following *complete* characterization of conditional entailment results:

Theorem 5.11 (Main). *A proposition p is conditionally entailed if and only if p is supported by a stable cover.*

We have thus arrived at a complete syntactic characterization of conditional entailment, which can now be computed by manipulating either models or arguments. Still, an undesirable feature of both approaches is that they rely on the identification of the set of admissible priority orderings in order to test whether relations of the form $\Delta' < \Delta$ are satisfied. This, however, is a nontrivial task. Fortunately, it is possible to replace such a test by a corresponding *syntactic* test on K .

Let us say that a *set Δ of assumptions dominates a set Δ'* if every assumption δ in Δ dominates the union $\Delta + \Delta'$. Then, due to the asymmetry and transitivity of priority orderings, the following result obtains:

Theorem 5.12 (Dominance). *For two sets of assumptions Δ and Δ' , the relation $\Delta' < \Delta$ holds in every priority ordering “<” admissible with a consistent background $K = \langle L, D \rangle$ if and only if Δ is part of a set Δ'' that dominates Δ' in K .¹²*

Theorems 5.11 and 5.12 permit us to determine whether a given proposition is conditionally entailed by purely syntactic means. To do so, we only need to look for stable covers and corresponding dominance relations.

¹¹ A consequence of this definition is that the stability of a cover C cannot be computed by considering the *minimal* arguments in conflict with C only. Rather, such arguments have to be made “definite” by extending them with as many assumptions from C as possible, which can lead to a proliferation of arguments to evaluate if C is large.

¹² A background K is consistent when there exists a priority ordering admissible with K . See Section 3.

6. Computing conditional entailment

Computing conditional entailment along the lines of the proof theory above, would still be a formidable task. A more reasonable approach would be to construct a sound but incomplete account which, by capturing most patterns of interest, will be both useful and understandable.

An obvious candidate for approximating conditional entailment is given by the propositions supported by stable arguments. From Theorem 5.11 we are guaranteed that such propositions are conditionally entailed, and that incompleteness results only from the exclusion of disjunctive arguments (covers). If we further commit ourselves to a single minimal admissible priority ordering, we find that testing whether a given set of assumptions Δ constitutes a stable argument can be easily accomplished in terms of the minimal conflict sets (nogoods) and support relations computed by ATMS-like systems (de Kleer [4]).

The first thing that we need to do is to incorporate information about priorities. However, since priorities only play a role in determining defeat relations among arguments, this information can be accommodated by partitioning each conflict sets C_i , $i = 1, \dots, n$, into two sets C_i^0 and C_i^1 , such that $C_i^0 < C_i^1$ holds and C_i^1 is maximal (such partition is unique and can easily be computed by setting C_i^0 to the set of assumptions δ in C_i which are *minimal* relative to the ordering “ $<$ ”).¹³ If C_i^1 is non-empty, we will call the pairs $\langle C_i^0, C_i^1 \rangle$, *basic defeat pairs*, as C_i^1 defeats C_i^0 and all other relevant defeat information can be inferred from them. Indeed, we can check whether an argument is stable as follows:

Theorem 6.1. *An argument Δ is stable if and only if for every conflict set C_i , $C_i \cap \Delta \neq \emptyset$, there is a basic defeat pair $\langle C_j^0, C_j^1 \rangle$ such that $C_j^0 + C_j^1 \subseteq C_i + \Delta$ and $C_j^0 \subseteq C_i - \Delta$.*

The idea is that if Δ' is a minimal argument in conflict with Δ , then there must be a conflict set C_i such that $\Delta' = C_i - \Delta$ and $C_i \cap \Delta$ is non-empty. Furthermore, if Δ is stable, Δ will contain a defeater of Δ' , namely, a subargument Δ_j in conflict with Δ' such that $\Delta' < \Delta_j$. In particular, there must be a minimal such Δ_j in Δ , and for such Δ_j , the set $\Delta_j + \Delta'$ will be a minimal conflict set (recall that we are assuming that Δ' is a *minimal* argument in conflict with Δ). Such a conflict set, call it C_j , can thus be partitioned into two sets $C_j^0 = \Delta'$ and $C_j^1 = \Delta_j$ such that $C_j^0 < C_j^1$. The pair $\langle C_j^0, C_j^1 \rangle$, however, is not necessarily a basic defeat pair, as the set C_j^1 is not necessarily maximal. Yet one such pair can easily be constructed by moving assumptions from C_j^0 to C_j^1 , and for such pair, $C_j^0 \subseteq \Delta'$ and $C_j^1 \subseteq \Delta + \Delta'$, as in Theorem 6.1.

¹³ Recall that we write $\Delta' < \Delta$ to express that for each assumption δ in Δ there is an assumption δ' in Δ' such that $\delta' < \delta$.

Thus, provided with the conflict sets computed by an ATMS, we can obtain the basic defeat pairs and test whether an argument is stable or not. This provides the basic tool for testing whether there is a stable argument that supports a given proposition p . To achieve this goal, we construct stable arguments incrementally: we adopt a set of assumptions Δ_0 that supports p and try to prove it stable. If it is not stable, we incrementally extend Δ_0 with a set of assumptions Δ_1 that defeats the counterarguments of Δ_0 , and try to show $\Delta_0 + \Delta_1$ stable, and so on.

To make this method precise, let (supports-of p) be a function that returns the minimal arguments Δ_i supporting a given atom p , and (defeaters-of Δ') be a function that returns the minimal arguments Δ_j which defeat a given argument Δ' . The function supports-of is part of the normal ATMS task, while the function defeaters-of can be computed from the basic defeat pairs $\langle C_j^0, C_j^1 \rangle$ for which $C_j^0 \subseteq \Delta'$ by setting Δ_j to $C_j^1 - \Delta'$. Furthermore, let

(choose x_i among $List$ until $p(x_i)$)

be an iterative procedure which binds x_i successively to the elements of $List$ until $p(x_i)$ evaluates to **true**, in which case the procedure returns **true**, or until the list has been exhausted, in which case the procedure returns **false**. Finally, let

(for-a-rebuttal Δ' of Δ do $p(\Delta')$)

be a function which finds a rebuttal Δ' and Δ , and evaluates to $p(\Delta')$ if one is found, and to **true** otherwise. Δ' is a rebuttal of Δ if both are in conflict but Δ does not contain a defeater of Δ' .

With these functions available, testing whether a proposition p is supported by a stable argument can be achieved by invoking the procedure (provable? p) below:

(provable? p) $\stackrel{\text{def}}{=}$ (choose Δ_i among (supports-of p)
until (proven-support Δ_i))

where

(proven-support Δ_i) $\stackrel{\text{def}}{=}$ (for-a-rebuttal Δ' of Δ_i do
(choose Δ_D among (defeaters-of Δ')
until (proven-support $\Delta_i + \Delta_D$)))

This method for testing whether a proposition p is supported by a stable argument is not aimed at being clear, rather than efficient: we build arguments supporting p , search for rebuttals of those arguments, and try to defeat those rebuttals. For finite languages, the procedure terminates and will be correct and complete. For practical purposes, however, additional refinements would be needed (e.g., caching results, evaluating certain expressions on need only, etc.)

7. Related work

The task of augmenting a conditional interpretation of defaults with assumptions about independence was first undertaken by Geffner and Pearl [9]. There we extended the core with an “irrelevance” rule which permitted us to derive the consequent q of a default $p \rightarrow q$ when both p and a *body of evidence* E *irrelevant to* q were given. The essence of that irrelevance condition was to assume E to be irrelevant to q given p when the arguments against q provided by E and p together were identical to those provided by p alone (so, E did not provide “new” arguments against q).

The proposal turned out adequate for most examples in the literature but lacked a formal justification. Thus, while five of the six inference rules had a clean probabilistic semantics, the irrelevance rule remained largely ad hoc. Conditional entailment now provides a basis for evaluating the scope and limitations of such irrelevance criteria. For example, the intuition underlying the criterion above can be understood as follows: if $p \rightarrow \delta$ is a default in K , and the arguments against δ provided by E and p together are identical to those provided by p alone, then those arguments will have to contain an assumption δ' with priority lower than δ and, thus, will not affect the status of δ which will be conditionally entailed in both contexts (Lemma 5.2).

Yet, although the intuition is the same, conditional entailment and the irrelevance account are not equivalent. In particular, the irrelevance criterion proposed in [9] is not *complete* with respect to conditional entailment (e.g., it does not sanction default contraposition),¹⁴ and examples can be constructed in which the irrelevance account is not even sound (although it remains sound for most simple examples).

An idea similar to irrelevance appears in a recent proposal by Poole [35]. Poole starts with a logical argument system, in which a proposition is accepted if it is supported by a (disjunctive) argument for which no rebuttal exists. He then introduces a dominance-like criterion that neutralizes some of these rebuttals. Poole’s proposal inherits some of the problems of the irrelevance account (e.g., its lack of formal justification), plus a few others (e.g., the resulting system, unlike the one he starts with, is not deductively closed). Other features become apparent when compared with conditional entailment. For instance, consider our example involving penguins, birds and red birds (Example 4.1), plus the default “if it looks like a penguin, it is a penguin”. Poole’s basic proposal, while able to conclude that something that looks like a penguin is a penguin and a bird, can no longer conclude that Tim is a penguin once Tim is *confirmed* to be a bird (i.e., it does not satisfy cumulativity, see Section 2).

¹⁴ In conditional entailment the only difference between a default $p \rightarrow q$ and a default **true** $\rightarrow (p \Rightarrow q)$, which clearly allows contraposition, is that the former provides a more specific context in which the material $p \rightarrow q$ can be asserted, and *thus* is stronger than the latter.

The problem here is that Poole's proposal does not capture the transitivity of default preferences, which in conditional entailment is enforced by means of priorities. The "iterated" account, which Poole introduces later compensates for this deficiency in certain cases (like in this example), but not in others (e.g., if "penguins are birds" is expressed as a default). A second problem can be pointed out in light of the proof theory of conditional entailment. In Poole's system there is no distinction between what we called *protection* and *defeat*, which sometimes yields a strange behavior. For instance, consider again the example above, plus a default "if it looks like a red bird, then it is a red bird", and a scenario in which we know that Tim is a penguin and looks like a red bird. Intuitively, we would expect to derive that Tim is a red bird. However, Poole can conclude that it is a bird, that it does not fly, but *not* that it is a red bird. In our terminology, this is because none of the arguments supporting "Tim is a red bird" *defeat* their counterarguments; yet, what really matters is that some of those arguments are *protected* from all their counterarguments, and thus are stable. Both differences point to the advantage of deriving an argument system from model-theoretic considerations, where no legitimate inference is being left out.

Other argument systems which enforce "specificity" preferences are the systems of Horty et al. [15], Nute [29], and Simari and Loui [41]. The major differences with these systems is that they regard defaults as rules of inference which do not contrapose nor do they allow reasoning by cases. Nute's system in particular implicitly embeds an interesting assumption: it permits us to derive the consequent q of a default $p \rightarrow q$ when all defaults $p' \rightarrow \neg q$ with supported antecedents p' are less "specific" than $p \rightarrow q$. Regardless of how this specificity criterion is determined, such a rule will derive a set of assumptions *only when those assumptions can be derived one at a time*. This limitation can be seen by considering the three rules

$$b(x) \rightarrow f(x), \quad p(x) \rightarrow \neg f(x), \quad p(x) \Rightarrow b(x)$$

as above, together with their "mirror" rules

$$b'(x) \rightarrow f(x), \quad p'(x) \rightarrow \neg f(x), \quad p'(x) \Rightarrow b'(x).$$

Given $p(a)$ and $p'(a)$, we would expect these rules to yield $\neg f(a)$. Yet, the rule $b(a) \rightarrow f(a)$ is no less specific than the conflicting default $p'(a) \rightarrow \neg f(a)$, and the rule $b'(a) \rightarrow f(a)$ is no less specific than the conflicting default $p(a) \rightarrow \neg f(a)$. Thus, the conclusion $\neg f(a)$ is not sanctioned by the above criterion. Conditional entailment produces the desired conclusion because, even though neither of the two "more specific" assumptions supporting the conclusion $\neg f(a)$ is stable by itself, *their union* is stable. In [8] we referred to stable arguments Δ which can be written as a list $\{\delta_1, \dots, \delta_n\}$ of assumptions, such that each prefix $\Delta_i = \{\delta_1, \dots, \delta_i\}$, $1 \leq i \leq n$,

is stable, as *linear stable arguments*. The computation of linear stable arguments is more efficient than the computation of nonlinear stable arguments, as the former can be built by proving one assumption at a time. On the other hand, as this example shows, a restriction to compute *only* linear stable arguments, implies that certain reasonable conclusions will escape derivation.

Conditional entailment is closest to two recent extensions of preferential entailment developed independently by Pearl [32] and Lehmann [17]. Like conditional entailment, Pearl and (implicitly) Lehmann rank defaults according to a dominance-like criterion, and use those rankings to infer a preference relation on models. Nonetheless Pearl and Lehmann deal with integer rankings as opposed to strict partial orders; they define the rank of a model as the rank of the highest ranked default violated by the model; and they only consider *one* (minimal) prioritized preferential structure, as opposed to many. The consequences of these choices are that preferred models are not always minimal; i.e., they do not always violate a minimal set of defaults, and conflicts among defaults that should remain unresolved, are sometimes resolved anomalously. Two examples illustrate these problems.

Given two defaults $p \rightarrow q$ and $p \rightarrow \neg r$, both accounts fail to authorize the conclusion q given $p \wedge r$. The reason is that, in the resulting world ranking, the violation of one default “costs” as much as the violation of many defaults of equal rank. This renders exceptional subclasses (e.g., penguins) unable to inherit properties from their superclasses (e.g., birds). The second class of problems stems from their commitment to a unique integer ranking on worlds. Consider for example two defaults $p \wedge s \rightarrow q$ and $r \rightarrow \neg q$ which render the status of q ambiguous in the presence of $p \wedge s \wedge r$. In Pearl’s and Lehmann’s approaches, this ambiguity is anomalously resolved *in favor* of q when a new default $p \rightarrow \neg q$ is added. The reason is that the introduction of $p \rightarrow \neg q$ automatically raises the ranking of the more specific default $p \wedge s \rightarrow q$ which thus becomes preferred to $r \rightarrow \neg q$. The extension of ε -semantics based on the principle of maximum entropy [12] remains committed to a unique integer ranking on worlds and thus inherits similar problems.

Another difference between conditional entailment and Pearl’s and Lehmann’s accounts is that the latter validate a stronger version of the augmentation rule, called *rational monotony* [18] by which conclusions of a theory $T = \langle K, E \rangle$ are also conclusions of the theory $T' = \langle K, E + \{p\} \rangle$, whenever the negation of p is not a consequence of T . Rational monotony does not hold in conditional entailment and whether it is a reasonable property of defeasible inference is an open issue. Yet, a weaker form of rational monotony does hold in conditional entailment. Let us say that a proposition q is *arguable* in T if there is an argument Δ for q such that T does not sanction $\neg \Delta$. Then, in conditional entailment, conclusions of the theory $T = \langle K, E \rangle$ will be conclusions of the theory $T' = \langle K, E + \{p\} \rangle$ when $\neg p$ is not *arguable* in T . This is

because the effect of propositions like p is simply to exclude *some* but not *all* models from each of the preferred classes of T .

Outside the conditional camp, conditional entailment is closest to prioritized circumscription. Prioritized circumscription is a refinement of parallel circumscription, originally proposed by McCarthy [25], and later developed by Lifschitz [20, 21]. Roughly, the effect of prioritized circumscription is to induce a preference for models that assign smaller extensions to predicates of higher priority. In the propositional case, the only difference between conditional entailment and prioritized circumscription is the source of the priorities: while prioritized circumscription relies on the user, conditional entailment extracts the priorities from the knowledge base itself. In the first-order case, two other technical differences arise: priorities in prioritized circumscription are on *predicates* as opposed to *literals*,¹⁵ and the minimality of models is understood in terms of “abnormal individuals” and not in terms of “abnormality literals”. The consequences of these choices are discussed in [7].

In light of the relation between the model theory of prioritized circumscription and conditional entailment, it is not surprising to find their respective proof theories related as well. An elegant proof theory for prioritized circumscription has recently been developed by Baker and Ginsberg [3]. Baker and Ginsberg address the case in which predicates are *linearly* ordered; namely, circumscribed predicates are drawn from sets P_1, P_2, \dots, P_n such that the priority of a predicate in a set P_i is higher than the priority of a predicate in a set P_j , if $j < i$. While differing in technical detail, the proof theory they present has the same *dialectical* flavor as the proof theory developed in Section 5, and both are closely related to systems of defeasible inference based on the evaluation of arguments [22, 33]. The differences are mainly in the treatment of disjunctions, which in our case are pushed completely into what we called *covers*, and our commitment to priority orderings which may be *partial* as opposed to *linear*. In this regard, the results in Sections 5 and 6 are relevant to prioritized circumscription, as they relax some of the assumptions on which the proof theory of Baker and Ginsberg is based.

8. Conclusions

Conditional entailment combines the benefits of the extensional and conditional interpretations of defaults, thus accounting for both “specificity” preferences and independence assumptions. Additionally, the inferences sanctioned by conditional entailment can be derived by an appealing proof theory, in which arguments of different strength “compete” and the stronger arguments “win”.

¹⁵ Except for the “pointwise” formulation in [21].

Conditional entailment, however, does not yet provide a complete account of default reasoning, and for certain default theories, the conclusions sanctioned by conditional entailment remain too weak. We have seen such case in Example 4.4; a better known example is the Yale Shooting Problem of Hanks and McDermott [14]. In the simplest terms, the Yale Shooting Problem can be encoded by means of a background comprised of the sentences

$$\begin{aligned} \text{loaded}_0 \wedge \delta_1 &\Rightarrow \text{loaded}_1, \\ \text{alive}_1 \wedge \delta_2 &\Rightarrow \text{alive}_2, \\ \text{shoot}_1 \wedge \text{loaded}_1 &\Rightarrow \neg \text{alive}_2, \\ \text{shoot}_1 \wedge \text{loaded}_1 &\Rightarrow \neg \delta_2, \end{aligned}$$

and the defaults $\text{loaded}_0 \rightarrow \delta_1$ and $\text{alive}_1 \rightarrow \delta_2$. In the context $T = \langle K, E \rangle$ with $E = \{\text{loaded}_0, \text{alive}_1, \text{shoot}_1\}$ we get to two classes of minimal models: one in which $\neg \text{alive}_2$ holds and the assumption δ_2 is violated, and a second in which $\neg \text{loaded}_1$ holds and the assumption δ_1 is violated. Since δ_1 is not constrained to have a priority higher than δ_2 , nor vice versa, both classes of models are equally preferred and, as in the formalisms analyzed by Hanks and McDermott, the “expected” conclusion $\neg \text{alive}_2$ is not sanctioned.

These examples show that default reasoning involves other aspects besides minimality and conditionality, which conditional entailment does not address. Of particular importance are considerations of causality and explanations. We have studied these issues in greater depth in [7].

Appendix A. Proofs

Lemma 3.2. *If the quadruple $\langle \mathcal{I}_{\mathcal{F}}, <, \Delta_{\mathcal{F}}, < \rangle$ is a prioritized preferential structure, then the pair $\langle \mathcal{I}_{\mathcal{F}}, < \rangle$ is a preferential structure.*

Proof. From Definition 3.1, the relation $M < M'$ holds iff $\Delta[M] \neq \Delta[M']$, and for every δ in $\Delta[M] - \Delta[M']$ there exists a δ' in $\Delta[M'] - \Delta[M]$, such that $\delta < \delta'$, where “ $<$ ” is an irreflexive and transitive relation which does not contain infinite ascending chains. First, note that the relation “ $<$ ” must also be irreflexive. We next show that “ $<$ ” is also transitive. Let M_1 , M_2 , and M_3 be three interpretations such that $M_1 < M_2$ and $M_2 < M_3$, and let

$$\Delta_1 = \Delta[M_1], \quad \Delta_2 = \Delta[M_2], \quad \Delta_3 = \Delta[M_3].$$

We will use the notation $\bar{\Delta}$ to denote the complement of a set Δ , i.e., $\bar{\Delta} = \Delta_{\mathcal{F}} - \Delta$, and write $\Delta_{i_1, i_2, \dots, i_n}$ to denote the intersection of the sets $\Delta_{i_1}, \dots, \Delta_{i_n}$. Furthermore, when one of the indices i is preceded by a minus sign, Δ_i is to be replaced by its complement $\bar{\Delta}_i$. Thus, for example, $\Delta_{1, -2, 3}$ stands for the intersection of the sets Δ_1 , Δ_3 and the complement $\bar{\Delta}_2$ of Δ_2 . Similarly, $\Delta_{-1, 2}$ stands for the intersection of $\bar{\Delta}_1$ and Δ_2 .

To prove transitivity we need to show that for every assumption δ in $\Delta_{1,-3}$, there is an assumption δ' in $\Delta_{-1,3}$ such that $\delta < \delta'$.¹⁶ First note that since “ $<$ ” does not contain infinite ascending chains, it is sufficient to prove this for every *maximal* element δ in $\Delta_{1,-3}$. Let then δ_1 be an arbitrary maximal element in $\Delta_{1,-3}$. We need to consider two cases:

- (1) If δ_1 belongs to $\Delta_{1,-2,-3}$, then δ_1 must also belong to $\Delta_{1,-2}$. Thus, since $M_1 < M_2$, there must be an assumption δ_2 in $\Delta_{-1,2}$ such that $\delta_1 < \delta_2$. Furthermore, let δ_2 be the maximal such element. If $\delta_2 \in \Delta_{-1,2,3}$, we are done. Otherwise, $\delta_2 \in \Delta_{-1,2,-3}$, and then since $M_2 < M_3$, there must be an assumption $\delta_3 \in \Delta_{-2,3}$ such that $\delta_2 < \delta_3$. Now, if $\delta_3 \in \Delta_{1,-2,3}$, then from $M_1 < M_2$, there must be a δ'_2 in $\Delta_{-1,2}$ such that $\delta_3 < \delta'_2$, and therefore, $\delta_2 < \delta'_2$, in contradiction with the maximality of δ_2 . Thus, $\delta_3 \in \Delta_{-1,3}$, and $\delta_1 < \delta_3$, by the transitivity of “ $<$ ”.
- (2) If δ_1 belongs to $\Delta_{1,2,-3}$, then, since $M_2 < M_3$, there must be a δ_3 in $\Delta_{-2,3}$ such that $\delta_1 < \delta_3$. Moreover, if $\delta_3 \in \Delta_{-1,-2,3}$ we are done. Otherwise, $\delta_3 \in \Delta_{1,-2,3}$, and therefore, as a result of $M_1 < M_2$, there must be a δ_2 in $\Delta_{-1,2}$ such that $\delta_3 < \delta_2$. Let δ_2 be a maximal such element. Then if δ_2 belongs to Δ_3 we are done. Otherwise, $\delta_2 \in \Delta_{-1,2,-3}$, and therefore, there must be a δ'_3 in $\Delta_{-2,3}$ such that $\delta_2 < \delta'_3$. Furthermore, δ'_3 cannot belong to Δ_1 ; otherwise, there should be another element δ'_2 in $\Delta_{-1,2}$, such that $\delta'_3 < \delta'_2$, contradicting the maximality of δ_2 . So, $\delta'_3 \in \Delta_{-1,3}$ and $\delta_1 < \delta'_3$ by transitivity of “ $<$ ”. \square

Lemma 3.3. *In any prioritized preferential structure, if M is a preferred model of a theory T then M is model of T minimal in $\Delta[M]$, i.e., there is no model M' of T such that $\Delta[M'] \subset \Delta[M]$.*

Proof. If $\Delta[M'] \subset \Delta[M]$, then $\Delta[M'] - \Delta[M] = \emptyset$, and $M' < M$ would trivially hold in every prioritized preferential structure preventing M from being preferred. \square

Theorem 3.7. *If an assumption-based default theory $T = \langle K, E \rangle$ preferentially entails a proposition p , then T also conditionally entails p .*

Proof. If T is logically inconsistent, the result is trivial. So let us assume that T is logically consistent. We will show that if $\xi = \langle \mathcal{I}_\xi, <, \Delta_\xi, < \rangle$ is a prioritized structure admissible with $K = \langle L, D \rangle$, then $\pi = \langle \mathcal{I}, < \rangle$, where $\mathcal{I} \subseteq \mathcal{I}_\xi$ stands for the collection of models of L , will be a preferential structure admissible with K . Thus if T does not conditionally entail p , T will not preferentially entail it either. Note that since we are assuming \mathcal{L} to be a finite propositional language, the preferential structure π must be well-founded. We need to show that for

¹⁶ A similar proof can be found in [36].

every default $p \rightarrow \delta$ in D , δ holds in all the preferred models of the theory $T' = \langle K, \{p\} \rangle$ in π . Again the result is trivial if T' is logically inconsistent, so we will assume otherwise. Let then M' be a model of T' in which δ does not hold, i.e., $\delta \in \Delta[M']$. We will now construct a model M preferred to M' in which δ holds. Since the preference order “ $<$ ” is well-founded, this is sufficient to prove that δ holds in all preferred models of T' . Let C stand for the collection of all minimal conflict sets in T' (i.e., minimal sets of assumptions logically inconsistent with T'), and let C' stand for the collection of all minimal conflict sets Δ in T' such that $\Delta \cap \Delta[M'] = \{\delta\}$. Since the priority ordering “ $<$ ” is admissible, any such set Δ must contain an assumption δ' such that $\delta' < \delta$. Let Δ' stand for the collection of all such assumptions δ' , and let us select M as an interpretation which satisfies T' , with a gap

$$\Delta[M] = \Delta[M'] + \Delta' - \{\delta\}.$$

There must be one such interpretation as $\Delta[M]$ is a *hitting set* for C (i.e., $\Delta[M]$ contains at least one assumption for every conflict set in C ; see [38]). Indeed, any conflict set in C not “hit” by assumptions from $\Delta[M'] - \{\delta\}$ will certainly be “hit” by assumptions from Δ' . Furthermore, $M < M'$ must hold, since

$$\Delta[M] - \Delta[M'] = \Delta', \quad \Delta[M'] - \Delta[M] = \{\delta\},$$

and for every δ' in Δ , $\delta' < \delta$ holds. \square

Lemma 4.2. *A proposition q is conditionally entailed by a default theory $T = \langle K, E \rangle$ iff q holds in all preferred models of T of every minimal prioritized preferential structure admissible with K .*

Proof. It is sufficient to show that if $<'$ is an admissible priority ordering that properly contains all the tuples in a *minimal* admissible priority ordering “ $<$ ”, then the preferred models in the structure $\xi' = \langle \mathcal{I}_{\mathcal{L}}, <', \Delta_{\mathcal{L}}, <' \rangle$ will be a subset of the preferred models of the structure $\xi = \langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, < \rangle$. Assume otherwise, that M is a preferred model in ξ' but not in ξ . Then there must be a model M' such that $M' < M$. This means that for every assumption δ in $\Delta[M'] - \Delta[M]$ there is an assumption δ' in $\Delta[M] - \Delta[M']$ such that $\delta < \delta'$. However, since $\delta < \delta'$ implies $\delta <' \delta'$, then M' would be preferred to M in ξ' as well, in contradiction with the minimality M in ξ' . \square

Lemmas 5.1 and 5.2 are special cases of Lemma 5.6, and the latter is a special case of the “if part” of Theorem 4 (Δ is a stable argument iff $C = \{\Delta\}$ is a stable cover). Recall that we are assuming that the theories under consideration give rise to a finite number of contradicted assumptions. These theories are *well-founded* in the following sense:

Lemma A.1. *If $T = \langle K, E \rangle$ gives rise to a finite number of contradicted assumptions, and $\xi = \langle \mathcal{J}_\xi, <, \Delta_\xi, < \rangle$ is a prioritized structure, then for every nonpreferred model M of T there is a preferred model M' of T such that $M' < M$.*

Proof. If T is logically inconsistent, the lemma follows trivially. So let us assume that T is logically consistent and let C stand for the collection of all minimal conflict sets that T gives rise to. It is easy to show that for every hitting set Δ for C (see above) there is a model M of T such that $\Delta[M] = \Delta$, and vice versa, that if M is a model of T then $\Delta[M]$ must include a hitting set Δ for C . Furthermore, there are only a finite number of contradicted assumptions, there must be a finite number of minimal hitting sets Δ_i , $i = 1, \dots, n$. Thus, let M_i stand for n models of T such that $\Delta[M_i] = \Delta_i$ and let \mathcal{M} stand for the collection of all such models. Furthermore, let \mathcal{M}_p denote the *minimal* collection of models in \mathcal{M} such that if $M \in \mathcal{M} - \mathcal{M}_p$ then \mathcal{M}_p contains a model M_i such that $M_i < M$. It is simple to show that such collection of models \mathcal{M}_p is unique. We show furthermore that they are all preferred models of T . Assume otherwise that there is a model M' of T preferred to some M_i in \mathcal{M}_p . This implies that the gap $\Delta[M']$ of M' contains some hitting set Δ_j , and thus, that if M' is preferred to M_i so will be M_j , in contradiction with the selection of M_i . We are thus left to show that for every nonpreferred model M of T there is a model in \mathcal{M}_p preferred to M . Two cases need to be considered.

- (1) If $\Delta[M] = \Delta_i$, $1 \leq i \leq n$, then a model will be preferred to M if and only if it is preferred to M_i above. Since M is not a preferred model of T , then M_i must belong to $\mathcal{M} - \mathcal{M}_p$, and thus, there must be a model M_j in \mathcal{M}_p preferred to M_i , and therefore, to M .
- (2) If for no i , $1 \leq i \leq n$, $\Delta[M] = \Delta_i$, then there must be one such i for which $\Delta[M] \supset \Delta_i$. In that case, $M_i < M$, and since \mathcal{M}_p must contain a model M_j preferred to M_i , by transitivity, $M_j < M$. \square

Theorem 5.11. *A proposition p is conditionally entailed if and only if p is supported by a stable cover.*

Proof. (if part) Since we are assuming that the theory $T = \langle K, E \rangle$ under consideration gives rise to a finite number of contradicted assumptions, by Lemma A.1 above, it is sufficient to show that for any model M which violates assumptions from every set Δ_i , $i = 1, \dots, n$, in the cover and any structure $\langle \mathcal{J}_\xi, <, \Delta_\xi, < \rangle$ admissible with K , there is model M' , $M' < M$, such that one of the assumption sets Δ_i is satisfied. Without loss of generality we can select M to be a minimal model, so that the set Δ' of assumptions validated by M is maximal. If there is no such minimal model M , we are done, because as stated

in Lemma A.1, there would be a minimal model M' that satisfies some Δ_i , such that $\Delta[M'] \subset \Delta[M]$, and thus $M' < M$. We assume thus that Δ' is maximal and in conflict with every set Δ_i in the cover. Since the cover is stable though, it must then contain a set Δ_i strongly protected from Δ' . That is, for every subset Δ'_j of Δ' in conflict with Δ_i , there is a subset Δ'_i of Δ_i in conflict with Δ' , such that $\Delta'_j < \Delta'_i$. That means that every set Δ'_j in Δ' in conflict with Δ_i contains an assumption δ'_j such that $\delta'_j < \delta'_i$, for some assumption δ'_i in both Δ'_i and $\Delta[M]$. Then, it is possible to build a model M' of T that satisfies Δ_i by making δ'_i true and δ'_j false. Thus for every assumption δ'_j in $\Delta[M'] - \Delta[M]$ there is an assumption δ'_i in $\Delta[M] - \Delta[M']$ such that $\delta'_j < \delta'_i$, and therefore $M' < M$.

(only-if part) If $T = \langle K, E \rangle$ gives rise to a finite number of contradicted assumptions, then T is well-founded and there are only a finite number of preferred classes. Let $\Delta_1, \Delta_2, \dots, \Delta_n$ be the *maximal* sets of assumptions validated by the preferred classes of T . Since p is conditionally entailed this means that every such set supports p . We will show now that the collection C of sets $\Delta_1, \dots, \Delta_n$ constitutes a stable cover, i.e., that for any definite conflicting argument Δ' , the cover contains an argument Δ_i strongly protected from Δ' . For that purpose, let M be a model of T satisfying Δ' and let M_i be a preferred model of T satisfying Δ_i , $1 \leq i \leq n$, such that $M_i < M$. From Lemma A.1 we know that there must be one such model. We show now that Δ_i is strongly protected from Δ' . Assume otherwise, i.e., there is a subset Δ'_j of Δ' in conflict with Δ_i such that for every set $\Delta'_i \subseteq \Delta_i$ in conflict with Δ' , $\Delta'_j \not< \Delta'_i$. This implies that the set Δ_A of assumptions δ in Δ such that $\Delta'_j < \delta$ is consistent with Δ' . Furthermore, since Δ' is a *definite* conflicting argument this means that $\Delta_A \subset \Delta'$, and therefore that for every assumption δ' in $\Delta[M] - \Delta[M_i]$, $\Delta'_j \not< \delta'$. However, this contradicts $M_i < M$; indeed, since Δ'_j is inconsistent with Δ_i , one of the assumptions δ'_j in Δ'_j must belong to $\Delta[M_i] - \Delta[M]$, and for $M_i < M$ to be true, another assumption δ_i , such that $\delta'_j < \delta_i$ must belong to $\Delta[M] - \Delta[M_i]$. \square

Theorem 5.12. *For two sets of assumptions Δ and Δ' , the relation $\Delta' < \Delta$ holds in every priority ordering “<” admissible with a consistent background $K = \langle L, D \rangle$ if and only if Δ is part of a set Δ'' that dominates Δ' in K .*

Proof. (if part) Let us recall, that we use the notation $\Delta' < \Delta$ to state that for every assumption δ in Δ there exists an assumption δ' in Δ' such that $\delta' < \delta$. Moreover, the relation “<” among *sets* of assumptions remains irreflexive and transitive, and therefore asymmetric. That is, for every priority ordering $\Delta \not< \Delta$, and if $\Delta_1 < \Delta_2$ and $\Delta_2 < \Delta_3$ hold, so does $\Delta_1 < \Delta_3$.

Let Δ stand for a collection of assumptions δ_i , $i = 1, \dots, n$. We will use the notation $\Delta_{i,j}$, for $i \leq j$, to stand for the set $\{\delta_i, \delta_{i+1}, \dots, \delta_j\}$. If $j > n$, the notation $\Delta_{i,j}$ is to be understood as $\Delta_{i,n}$, and as the empty set, if $i > n$. We show

first that if Δ dominates a set Δ' then the relation $\Delta' < \Delta$ must hold for any priority ordering “<” admissible with K . We show this by induction:

Base case: $\Delta_{i+1,n} + \Delta' < \Delta_{1,i}$, for $i = 1$. Clearly, if Δ dominates Δ' , the assumption δ_1 must dominate $\Delta_{2,n} + \Delta'$, and thus, $\Delta_{2,n} + \Delta' < \{\delta_1\}$ must hold. Thus, if $n = 1$, we are done.

Induction hypothesis: Let us assume that n is greater than one. Furthermore, let us assume as inductive hypothesis that $\Delta_{i+1,n} + \Delta' < \Delta_{1,i}$ holds for every i , $1 \leq i \leq j \leq n$. We need to show the same relation holds for $i = j$. Now, since δ_j dominates the set $\Delta + \Delta'$ we must have $\Delta + \Delta' < \delta_j$ for any admissible “<”. If we can show $\Delta_{j+1,n} + \Delta' < \delta_j$, we will be done, as otherwise there should be some assumption δ in $\Delta_{1,j-1}$ for which $\delta_j < \delta$ holds but $\Delta_{j+1,n} + \Delta' < \delta$ does not, contradicting transitivity. Similarly, if $\Delta_{j+1,n} + \Delta' < \delta_j$ did not hold, we should have $\delta < \delta_j$ for some δ in $\Delta_{1,j-1}$. Furthermore, since $\Delta_{j+1,n} + \{\delta_j\} + \Delta' < \Delta_{1,j-1}$ holds as hypothesis, the latter would imply that either $\Delta_{j+1,n} + \Delta' < \delta$ holds, in contradiction with transitivity, or that $\delta_j < \delta$ holds, in contradiction with the asymmetry of priority orderings.

(only-if part) This part of the proof is slightly more involved. We need to show that if the relation $\Delta' < \Delta$ holds for every admissible ordering with a (conditionally) consistent background context K , then Δ is part of a set that dominates Δ' . Let us first divide the assumptions in $\Delta_{\mathcal{F}}$ between those which participate in a set that dominates Δ' , which we group into a set Δ_A , from those which do not participate in a set that dominates Δ' . Furthermore, let

$$\Delta_B = \Delta' - \Delta_A \quad \text{and} \quad \Delta_C = \Delta_{\mathcal{F}} - \Delta_A - \Delta_B.$$

Note that Δ_B cannot be empty, otherwise Δ_A would dominate itself, precluding K from being consistent. Note also, that if two sets dominate Δ' , so will their union. It follows then that Δ_A dominates Δ' . Our goal will be to show that Δ is included in Δ_A . For that we will show that there is a priority ordering “<” admissible with K , such that the relation $\Delta' < \delta$ holds only if $\delta \in \Delta_A$.

Let us say that a priority ordering “<” in a background context K is admissible within a *range* Δ and a *restriction* Δ' iff every set Δ'' dominated by an assumption δ in Δ contains an assumption δ' in Δ' , such that $\delta' < \delta$ holds. The notions of *range* and *restriction* provide a finer measure of the admissibility of a priority ordering. In particular, an admissible priority ordering, must be admissible within a range $\Delta_{\mathcal{F}}$ and a restriction $\Delta_{\mathcal{F}}$. Furthermore, if a priority relation “<” is admissible within a range Δ_1 and a restriction Δ_2 , for two disjoint sets Δ_1 and Δ_2 such that $\Delta_1 + \Delta_2 = \Delta_{\mathcal{F}}$, then there must be a priority relation “<” admissible within a range Δ_1 and a restriction $\Delta_{\mathcal{F}}$, such that $\delta_2 < \delta_1$ holds only if $\delta_1 \in \Delta_1$ and $\delta_2 \in \Delta_2$. Indeed, if “<” is a priority relation admissible within a range Δ_1 and a restriction Δ_2 , the relation that results by deleting all pairs $\langle \delta_1 \notin \Delta_1, \delta_2 \notin \Delta_2 \rangle$ for which $\delta_1 < \delta_2$ holds, remains irreflexive, transitive, and admissible.

Now, let us assume that there is no priority ordering admissible within a range Δ_C and a restriction Δ_C , for Δ_C as above. It is possible to show then that there must be a non-empty subset $\overline{\Delta'_C}$ of Δ_C such that each assumption $\delta' \in \Delta'_C$ dominates the set $\Delta'_C + \overline{\Delta_C}$, where $\overline{\Delta_C}$ stands for the set of assumptions not in Δ_C ; in this case, $\Delta_A + \Delta_B$. This, however, amounts to say that Δ'_C dominates the set $\Delta_A + \Delta_B$, which by virtue of the dominance of Δ_A over Δ' and the inclusion of Δ_B in Δ' , implies that Δ'_C dominates Δ' as well, in contradiction with the maximality of Δ_A . Thus, there must be a priority ordering " $<_C$ " admissible within a range Δ_C and a restriction Δ_C , such that $\delta <_C \delta'$ holds only if both δ and δ' belong to Δ_C . Furthermore, since K is consistent, there must be a priority ordering " $<_A$ " admissible within range Δ_A and restriction Δ_\varnothing , such that $\delta <_A \delta'$ holds only if δ' belongs to Δ_A . We can thus define a relation " $<$ " such that $\delta < \delta'$ iff [$\delta <_A \delta'$] or [$\delta <_C \delta'$] or [$\delta \in \Delta_C$ and $\delta' \in \Delta_A + \Delta_B$]. It is simple to show that such a relation is a priority relation, and that it is admissible within a range $\Delta_A + \Delta_C$. Let us assume, on the other hand, that " $<$ " is not admissible within a range Δ_B . That is, there is an assumption δ in Δ_B which dominates a set Δ'_B for which the relation $\Delta'_B < \delta$ fails to hold. Note that Δ'_B cannot contain elements from Δ_C ; for otherwise the relation $\Delta'_B < \delta$ will certainly hold. Thus, $\Delta'_B \subseteq \Delta_A + \Delta_B$, so δ dominates $\Delta_A + \Delta_B$. That means, however, that the set $\Delta_A + \{\delta\}$ dominates the set Δ' , in contradiction with the assumption that Δ_A is the maximal such set. So, the ordering " $<$ " must be admissible within the range Δ_B as well, and so " $<$ " must be a priority relation admissible with K . Since $\Delta' < \Delta$ holds by hypothesis, and $\Delta' < \delta$ holds only if $\delta \in \Delta_A$, it follows then that Δ is part of a set, Δ_A , which dominates Δ' . \square

References

- [1] E. Adams, *The Logic of Conditionals* (Reiter, Dordrecht, Netherlands, 1975).
- [2] E. Adams, A note comparing probabilistic and modal logics of conditionals, *Theoria* **43** (1978) 186–194.
- [3] A. Baker and M. Ginsberg, A theorem prover for prioritized circumscription, in: *Proceedings IJCAI-89*, Detroit, MI (1989) 463–467.
- [4] J. de Kleer, An assumption-based TMS, *Artif. Intell.* **28** (1986) 127–162.
- [5] J. Delgrande, An approach to default reasoning based on a first-order conditional logic, in: *Proceedings AAAI-87*, Seattle, WA (1987) 340–345.
- [6] D. Gabbay, Theoretical foundations for non-monotonic reasoning in expert systems, in: K.R. Apt, ed., *Logics and Models of Concurrent Systems* (Springer, Berlin, 1985) 439–457.
- [7] H. Geffner, Default reasoning: causal and conditional theories, Ph.D. Thesis, Computer Science Department, UCLA, Los Angeles, CA (1989). Forthcoming, MIT Press.
- [8] H. Geffner, Beyond negation as failure, in: *Proceedings Second International Conference on Principles of Knowledge Representation and Reasoning* (1991) 218–229.
- [9] H. Geffner and J. Pearl, A framework for reasoning with defaults, Tech. Report TR-94-III, Cognitive Systems Laboratory, UCLA, Los Angeles, CA (1987); also in: H. Kyburg, R. Loui and G. Carlson, eds., *Knowledge Representation and Defeasible Inference* (Kluwer, Boston, MA, 1990).

- [10] M. Gelfond and H. Przyminska, Negation as failure: careful closure procedure, *Artif. Intell.* **30** (1986) 273–287.
- [11] M. Ginsberg, A local formalization of inheritance: preliminary report, in: *Proceedings Third International Workshop on Nonmonotonic Reasoning*, South Lake Tahoe, CA (1990) 119–130.
- [12] M. Goldszmidt, P. Morris and J. Pearl, A maximum entropy approach to nonmonotonic reasoning, in: *Proceedings AAAI-90*, Boston, MA (1990) 646–652.
- [13] M. Goldszmidt and J. Pearl, On the consistency of defeasible databases, *Artif. Intell.* **52** (1991) 121–149.
- [14] S. Hanks and D. McDermott, Nonmonotonic logic and temporal projection, *Artif. Intell.* **33** (1987) 379–412.
- [15] J.F. Horty, R.H. Thomason and D.S. Touretzky, A skeptical theory of inheritance, in: *Proceedings AAAI-87*, Seattle, WA (1987) 358–363.
- [16] S. Kraus, D. Lehmann and M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, *Artif. Intell.* **44** (1990) 167–207.
- [17] D. Lehmann, What does a conditional knowledge base entail? In: *Proceedings First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, Ont. (1989) 212–222.
- [18] D. Lehmann and M. Magidor, Rational logics and their models: a study in cumulative logic, Tech. Report, Department of Computer Science, Hebrew University, Jerusalem, Israel (1988).
- [19] D. Lewis, *Counterfactuals* (Harvard University Press, Cambridge, MA, 1973).
- [20] V. Lifschitz, Computing circumscription, in: *Proceedings IJCAI-85*, Los Angeles, CA (1985) 121–127.
- [21] V. Lifschitz, Circumscriptive theories: a logic-based framework for knowledge representation, *J. Philos. Logic* **17** (1988) 391–441.
- [22] R. Loui, Defeat among arguments: a system of defeasible inference, *Comput. Intell.* **3** (3) (1987) 100–107.
- [23] D. Makinson, General theory of cumulative inference, in: *Proceedings Second International Workshop on Nonmonotonic Reasoning*, Munich, Germany (1988) 1–18.
- [24] J. McCarthy, Circumscription—a form of non-monotonic reasoning, *Artif. Intell.* **13** (1980) 27–39.
- [25] J. McCarthy, Applications of circumscription to formalizing common-sense knowledge, *Artif. Intell.* **28** (1986) 89–116.
- [26] D. McDermott and J. Doyle, Non-monotonic logic I, *Artif. Intell.* **13** (1980) 41–72.
- [27] R.C. Moore, Semantical considerations on nonmonotonic logic, *Artif. Intell.* **25** (1985) 75–94.
- [28] D. Nute, Conditional logic, in: D. Gabbay and F. Guenther, eds., *Handbook of Philosophical Logic* (Reidel, Dordrecht, Netherlands, 1984) 387–439.
- [29] D. Nute, Defeasible reasoning: a philosophical analysis in Prolog, in: J. Fetzer, ed., *Aspects of Artificial Intelligence* (Kluwer Academic Publishers, Boston, MA, 1988) 251–287.
- [30] J. Pearl, *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann, Los Altos, CA, 1988).
- [31] J. Pearl, Probabilistic semantics for nonmonotonic reasoning: a survey, in: *Proceedings First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, Ont. (1989) 505–516. Expanded version in: R. Cummins and J. Pollock, eds., *Philosophical AI: Computational Approaches to Reasoning* (MIT Press, Cambridge, MA, 1991) 157–187.
- [32] J. Pearl, System Z: a natural ordering of defaults with tractable applications to non-monotonic reasoning, in: R. Parikh, ed., *Theoretical Aspects of Reasoning about Knowledge* (Morgan Kaufmann, San Mateo, CA, 1990) 121–135.
- [33] J. Pollock, Defeasible reasoning, *Cogn. Sci.* **11** (1987) 481–518.
- [34] D. Poole, A logical framework for default reasoning, *Artif. Intell.* **36** (1988) 27–47.
- [35] D. Poole, Dialectics and specificity: conditioning in logic-based hypothetical reasoning (preliminary report), in: *Proceedings Third International Workshop on Nonmonotonic Reasoning*, South Lake Tahoe, CA (1990) 201–209.
- [36] T. Przyminska, On the declarative semantics of stratified deductive databases and logic programs, in: J. Minker, ed., *Foundations of Deductive Databases and Logic Programming* (Morgan Kaufmann, Los Altos, CA, 1987) 193–216.

- [37] R. Reiter, A logic for default reasoning, *Artif. Intell.* **12** (1980) 81–132.
- [38] R. Reiter, A theory of diagnosis from first principles, *Artif. Intell.* **32** (1987) 57–95.
- [39] R. Reiter and G. Criscuolo, Some representational issues in default reasoning, *Int. J. Comput. Math.* **9** (1983) 1–13.
- [40] Y. Shoham, *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence* (MIT Press, Cambridge, MA, 1988).
- [41] G. Simari and R. Loui, Confluence of argument systems: Poole's rules revisited, in: *Proceedings Third International Workshop on Nonmonotonic Reasoning*, South Lake Tahoe, CA (1990) 223–232.