A Series of Revisions of DAVID POOLE's Specificity

CLAUS-PETER WIRTH
FRIEDER STOLZENBURG
FB AI, Hochschule Harz, 38855 Wernigerode, Germany
wirth@logic.at
fstolzenburg@hs-harz.de

October 17, 2014

Abstract

In the middle of the 1980s, DAVID POOLE introduced a semantical, model-theoretic notion of specificity to the artificial-intelligence community. Since then it has found further applications in non-monotonic reasoning, in particular in defeasible reasoning. POOLE tried to approximate the intuitive human concept of specificity, which seems to be essential for reasoning in everyday life with its partial and inconsistent information. His notion, however, turns out to be intricate and problematic, which — as we show — can be overcome to some extent by a closer approximation of the intuitive human concept of specificity. Besides the intuitive advantages of our novel specificity orderings over POOLE's specificity relation in the classical examples of the literature, we also report some hard mathematical facts: Contrary to what was claimed before, we show that POOLE's relation is not transitive in general. The first of our specificity orderings (CP1) captures POOLE's original intuition as close as we could get after the correction of its technical flaws. The second one (CP2) is a variation of CP1 and presents a step toward similar notions that may eventually solve the intractability problem of POOLE-style specificity relations. The present means toward deciding our novel specificity relations, however, show only slight improvements over the known ones for POOLE's relation, and further work is needed for testing the viability of a workaround we suggest for applications in practice.

Keywords: Artificial Intelligence, Positive-Conditional Specification, Non-Monotonic Reasoning, Specificity, Defeasible Reasoning

[small letters]

Contents

1 Introduction

A possible explanation of how humans manage to interact with reality — in spite of the fact that their information on the world is partial and inconsistent — mainly consists of the following two points:

1. Humans use a certain amount of rules for default reasoning and are aware that some arguments relying on these rules may be defeasible.
2. In case of the frequent conflicting or even contradictory results of their reasoning, they prefer more specific arguments to less specific ones.

An intuitive concept of specificity plays an essential rôle in this explanation, which is interesting because it seems to be highly successful in practice, even if it were just an epiphenomenon providing an *ex eventu* explanation of human behavior.

On the long way approaching this proven intuitive human concept of specificity, the first milestone marks the development of a semantical, model-theoretic notion of specificity having passed first tests of its usefulness and empirical validity. Indeed, at least as the first step, a semantical, model-theoretic notion will probably offer a broader and better basis for applications in systems for common sense reasoning than notions of specificity that depend on peculiarities of special calculi or even on extra-logical procedures. ~~This holds in particular if the results of these systems are to be accepted by human users or even by the human society.~~ [not necessary]

David Poole has sketched such a notion as a binary relation on arguments and evaluated its intuitive validity with some examples in [Poole, 1985]. Poole's notion of specificity was given a more appropriate formalization in [Simari & Loui, 1992]. The properties of this formalization were examined in detail in [Stolzenburg &al., 2003].

In this paper, before we give a specification of the formal requirements on any reasonably conceivable relation of specificity in § 5, we present a detailed analysis of the reasons behind our *intuition that* Poole*'s specificity is a first step on the right way* (§ 4). We expect that the results of this analysis will carry us even beyond this paper to future improved concepts of specificity, especially w.r.t. efficiency, but also w.r.t. intuitive adequacy. We hope that the closer we get to human intuition, the more efficiently our concepts can be implemented, simply because they seem to run so well on the human hardware, which — by all that we know today — is pretty slow.

Moreover, in § 6, we clearly disambiguate Poole's specificity from slightly improved versions such as the one in [Simari & Loui, 1992], and introduce a novel specificity relation (CP1), which presents a major correction of Poole's specificity because it removes a crucial shortcoming of Poole's original relation (P1) and its slight improvements (P2, P3), namely their lack of transitivity.

Furthermore, in § 7, we present several examples that are to convince every carefully contemplating reader of the superiority of our novel specificity relation CP1 w.r.t. human intuition.

Finally, in § 8, we discuss efficiency issues and introduce the specificity ordering CP2, a variation of CP1, which presents a first step toward similar notions that may finally solve the intractability problem of Poole-style specificity relations, for which we also present a workaround which remains to be evaluated in practice; and then we conclude with § 9.

2.2 Secondary Aspects of our Logic

Remark 2.3 (Negation Symbol "¬")
The negation symbol "¬", which occurs in Definition 2.1 and which seemingly gets us beyond the definite rules of positive-conditional specifications by admitting literals instead of just atoms, does not have any effect, however, on the *derivations* and *theories* considered in this paper (cf. Definition 2.2).

For instance, the literal ¬flies(edna) may actually be considered as the atom resulting from application of the predicate ¬flies to the constant symbol edna.

On the other hand, if we write an atom A as $A = \mathsf{true}$, and a negated atom $\neg A$ as the equational atom $A = \mathsf{false}$, for the data type BOOLEan given by the constructors true and false, then the rules of our specification can be seen as *positive*-conditional equational specifications in the framework for positive/negative-conditional specification found in [WIRTH & GRAMLICH, 1994], [WIRTH, 1997; 2009].[1] [AHA! does no use [S, [in one pair of brack

In the application context, of course, the literals ¬flies(edna) and flies(edna) will be considered to be *contradictory* (cf. Definition 2.4), but this is a secondary and non-essential notion built on top of our derivations and theories, which do not rely on this notion.

As a consequence, none of the results in this paper relies on this special negation symbol. To the contrary, in the weakness of our logical theories we see an indication for the generality of our results (cf. Remark 2.5).

To distinguish the inactive negation here from negation as failure and from any other form of negation playing an active rôle in derivation, the symbol "∼" is sometimes used in the literature of defeasible logic in place of our more standard symbol "¬".

Definition 2.4 (Contradictory Sets of Rules)
A set of rules Π is called *contradictory* if there is an atom A such that $\Pi \vdash \{A, \neg A\}$; otherwise Π is *non-contradictory*.

Remark 2.5 (Weakness of Our Logical Theories)
On the one hand, $\{\mathsf{A},\ \neg\mathsf{A} \Leftarrow \mathsf{A}\}$ is contradictory according to Definitions 2.2 and 2.4.

On the other hand, $\{\mathsf{A} \Leftarrow \neg\mathsf{A},\ \neg\mathsf{A} \Leftarrow \mathsf{A}\}$ is non-contradictory according to these definitions, although we can infer both A and $\neg\mathsf{A}$ from $\{\mathsf{A} \Leftarrow \neg\mathsf{A},\ \neg\mathsf{A} \Leftarrow \mathsf{A}\}$ in classical (i.e. two-valued) logic.

For the case of our very limited formal language, our notions of consequence and contradiction are equivalent both to intuitionistic logic and to the three-valued logic where $\neg$ and $\wedge$ are given as usual,[2] but (following neither KLEENE nor ŁUKASIEWICZ) implication has to be defined via

$$(A \Leftarrow \mathsf{TRUE}) = A, \qquad (A \Leftarrow \mathsf{FALSE}) = \mathsf{TRUE}, \qquad (A \Leftarrow \mathsf{UNDEF}) = \mathsf{TRUE}.$$

[1] Note, however, that derivability is invariant under this equivalence transformation on atoms only if our specification is non-contradictory in the sense of Definition 2.4, in which case also the equational specification is consistent in the sense of $\mathsf{true} \neq \mathsf{false}$.

[2] The standard interpretation is that TRUE is 1, UNDEF is $\frac{1}{2}$, FALSE is 0, $\neg A$ is $1-A$, and $A \wedge B$ is $\min\{A, B\}$. In other words: $\neg\mathsf{TRUE} = \mathsf{FALSE}$, $\neg\mathsf{UNDEF} = \mathsf{UNDEF}$, $\neg\mathsf{FALSE} = \mathsf{TRUE}$; $\mathsf{TRUE} \wedge A = A$, $\mathsf{UNDEF} \wedge \mathsf{TRUE} = \mathsf{UNDEF}$, $\mathsf{UNDEF} \wedge \mathsf{UNDEF} = \mathsf{UNDEF}$, $\mathsf{UNDEF} \wedge \mathsf{FALSE} = \mathsf{FALSE}$, $\mathsf{FALSE} \wedge A = \mathsf{FALSE}$.

[may be omitted]

Remark 2.8 (Minimality and Non-Contradiction of Arguments)
Some authors (cf. e.g. [STOLZENBURG &AL., 2003], [CHESÑEVAR &AL., 2003]) require all arguments

1. to be minimal arguments, and
2. to be non-contradictory.

Because non-minimal as well as contradictory arguments often occur in practical situations, there is no use-oriented justification for any of these requirements.

For requirement 1 there is no conceptual justification, either, because the non-minimal arguments become inessential by our preference on specific arguments, in the sense that for every argument there must be a minimal sub-argument that is at least as specific, cf. Corollaries 6.7, 6.13, and 8.11.

Because being contradictory is only a secondary aspect of our logic (cf. § 2.2), there is no conceptual justification for requirement 2, either.

To obtain a more general setting in the comparison of arguments, we omit these restrictions in the context of this paper, where they turned out to be completely superfluous. Thus, the omission of these requirements has no effect on the results of this paper.

2.5 Quasi-Orderings

We will use several binary relations comparing arguments according to their specificity. For any relation written as $\precsim_N$ ("being more or equivalently specific w.r.t. N"), we set

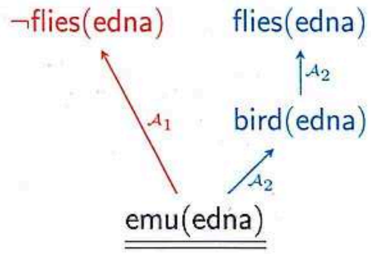
$$\begin{array}{lcll}
\succsim_N & := & \{\, (X,Y) \mid Y \precsim_N X \,\} & \text{("less or equivalently specific"),} \\
\approx_N & := & \precsim_N \cap \succsim_N & \text{("equivalently specific"),} \\
<_N & := & \precsim_N \setminus \succsim_N & \text{("properly more specific"),} \\
\leq_N & := & <_N \cup \{\, (X,X) \mid X \text{ is an argument} \,\} & \text{("more specific or equal"),} \\
\triangle_N & := & \left\{ (X,Y) \;\middle|\; \begin{array}{l} X,Y \text{ are arguments with} \\ X \not\precsim_N Y \text{ and } X \not\succsim_N Y \end{array} \right\} & \text{("incomparable w.r.t. specificity").}
\end{array}$$

A *quasi-ordering* is a reflexive transitive relation. An *(irreflexive) ordering* is an irreflexive transitive relation. A *reflexive ordering* (also called: "partial ordering") is an anti-symmetric quasi-ordering. An *equivalence* is a symmetric quasi-ordering.

Corollary 2.9 *If $\precsim_N$ is a quasi-ordering, then $\approx_N$ is an equivalence, $<_N$ is an ordering, and $\leq_N$ is a reflexive ordering.*

Let us see what happens to Example 3.2 if we ~~get the idea that emus may actually be some sort of dinosaurs and~~ doubt that they are birds.

Example 3.3 (Renamed Subsystem of Example 3 of [POOLE, 1985])



$$\begin{aligned}
\Pi^{\mathrm{F}}_{3.3} &:= \{\ \mathsf{emu(edna)}\ \}, \qquad \Pi^{\mathrm{G}}_{3.3} := \emptyset,\\
\Delta_{3.3} &:= \left\{\begin{array}{l} \neg\mathsf{flies}(x) \leftarrow \mathsf{emu}(x),\\ \mathsf{flies}(x) \leftarrow \mathsf{bird}(x),\\ \mathsf{bird}(x) \leftarrow \mathsf{emu}(x)\end{array}\right\},\\
\mathcal{A}_1 &:= \{\ \neg\mathsf{flies(edna)} \leftarrow \mathsf{emu(edna)}\ \}.\\
\mathcal{A}_2 &:= \left\{\begin{array}{l} \mathsf{flies(edna)} \leftarrow \mathsf{bird(edna)},\\ \mathsf{bird(edna)} \leftarrow \mathsf{emu(edna)}\end{array}\right\}.
\end{aligned}$$

We have

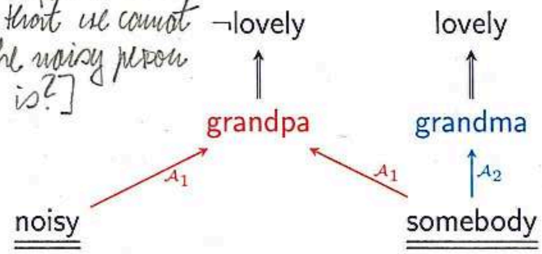
$$\mathfrak{T}_{\Pi_{3.3}} = \{\mathsf{emu(edna)}\}, \qquad \mathfrak{T}_{\Pi_{3.3}\cup\Delta_{3.3}} = \{\mathsf{bird(edna)}, \mathsf{flies(edna)}, \neg\mathsf{flies(edna)}\} \cup \mathfrak{T}_{\Pi_{3.3}}.$$

Now it is not clear anymore whether we should prefer $(\mathcal{A}_1, \neg\mathsf{flies(edna)})$ to $(\mathcal{A}_2, \mathsf{flies(edna)})$. Both arguments are now based on $\mathsf{emu(edna)}$, but it is not clear whether the less specific $\mathsf{bird(edna)}$ — that has dropped out of $\mathfrak{T}_{\Pi_{3.3}}$ — can still be considered as a basis for $(\mathcal{A}_2, \mathsf{flies(edna)})$. We will further discuss this in Example 6.20.

Let us now suppose that we have a lovely grandma and a grouchy grandpa, stay at their house, and somebody is coming into the house noisily, but we cannot see who it is.

[how does it fit to the statement, that we cannot see who the noisy person is?]

Example 3.4



$$\begin{aligned}
\Pi^{\mathrm{F}}_{3.4} &:= \{\ \mathsf{somebody},\ \mathsf{noisy}\ \},\\
\Pi^{\mathrm{G}}_{3.4} &:= \left\{\begin{array}{l} \mathsf{lovely} \Leftarrow \mathsf{grandma},\\ \neg\mathsf{lovely} \Leftarrow \mathsf{grandpa}\end{array}\right\},\\
\Delta_{3.4} &:= \mathcal{A}_1 \cup \mathcal{A}_2.\\
\mathcal{A}_1 &:= \{\ \mathsf{grandpa} \leftarrow \mathsf{somebody}\wedge\mathsf{noisy}\ \},\\
\mathcal{A}_2 &:= \{\ \mathsf{grandma} \leftarrow \mathsf{somebody}\ \}.
\end{aligned}$$

Let us compare the specificity of the arguments $(\mathcal{A}_1, \neg\mathsf{lovely})$ and $(\mathcal{A}_2, \mathsf{lovely})$. We have

$$\mathfrak{T}_{\Pi_{3.4}} = \{\mathsf{somebody}, \mathsf{noisy}\}, \qquad \mathfrak{T}_{\Pi_{3.4}\cup\Delta_{3.4}} = \{\mathsf{grandma}, \mathsf{grandpa}, \mathsf{lovely}, \neg\mathsf{lovely}\} \cup \mathfrak{T}_{\Pi_{3.4}}.$$

Now, because there is somebody who is noisy according to the current situation given by $\Pi^{\mathrm{F}}_{3.4}$, it is probably grandpa because his characterization is more specific. Thus, it is intuitively clear that we would prefer $(\mathcal{A}_1, \neg\mathsf{lovely})$ as the more specific argument to $(\mathcal{A}_2, \mathsf{lovely})$. We will further discuss this in Example 6.21.

4.3 The Intuitive Rôle of Activation Sets in the Definition of Specificity

If we want to classify a derivation with defeasible rules according to its specificity, then we have to isolate the defeasible part of the derivation and look at its input formulas, so that we can see how specific these input formulas are. The input formulas are the set of those literals on which the defeasible part of the derivation is based, called the *activation set* for the defeasible part of the derivation. In our framework of defeasible positive-conditional specification, the only relevant property of an activation set can be the conjunction of its literals which we can represent by the set itself.[5]

For instance, in Example 3.2 of § 3, the argument $(\mathcal{A}_1, \neg\mathsf{flies}(\mathsf{edna}))$ is based only on the activation set $\{\mathsf{emu}(\mathsf{edna})\}$, whereas the argument $(\mathcal{A}_2, \mathsf{flies}(\mathsf{edna}))$ can also be based on the activation set $\{\mathsf{bird}(\mathsf{edna})\}$, or on the union of these sets.

Moreover, in Example 3.4 of § 3, the argument $(\mathcal{A}_1, \neg\mathsf{lovely})$ is based only on the activation set $\{\mathsf{somebody}, \mathsf{noisy}\}$, whereas the argument $(\mathcal{A}_2, \mathsf{lovely})$ can also be based on the (smaller) activation set $\{\mathsf{somebody}\}$.

4.3.1 Modulo Which Theory are Activation Sets to be Compared? ?

Because all literals of an activation set have been derived from the given specification, it does not make sense to compare activation sets w.r.t. the models of the entire specification. Indeed, only a comparison w.r.t. the models of a sub-specification can show any differences between them.

Therefore, we have to find out which parts of a specification $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$ are to be excluded from the comparison of activation sets.

We want to have the *entire* set Π^{G} available for our comparison of activation sets, for the following reasons: The general and strict part Π^{G} of our specification represents the necessary and stable kernel of our rules, independent of the concrete situation under consideration given by Π^{F}, and independent of the uncertainty of our default rules Δ. Moreover, it is hardly meaningful to exclude any proper rule from Π^{G} (i.e. any rule from Π^{G} that is not just a literal); the technical reason for this will be given right at the beginning of § 4.4.3.

[5] A formal definition of an activation set is not needed here and would be harmful to intuition, but several different formal notions of activation sets will be found in Definition 6.1 of § 6.1 and also in Definition 8.7 of § 8.3.1.

may be omitted

4.4 Isolation of the Defeasible Parts of a Derivation

If $(\mathcal{A}, L)$ is an argument (cf. § 4.2), then there is a derivation of L which is based only on those instances of defeasible rules which are contained in $\mathcal{A}$. Such an argument ignores the concrete derivation, and therefore suits our model-theoretic intentions (cf. § 1). With such an argument as an abstraction of a derivation, however, we lose the possibility to isolate the actual defeasible parts of the derivation. Such a loss is typical for abstractions in general; in our case, however, the discussion of this loss in § 4.4.1 will turn out to be conceptually crucial and result in several different formal notions of activation sets (found in Definition 6.1 of § 6.1 and also in Definition 8.7 of § 8.3.1). [too many forward references]

4.4.1 Isolation of Actual Defeasible Parts in And-Trees

Let us compare this set $\mathcal{A}$ with an *and-tree of the derivation*. Every node in such a tree is labeled with the conclusion of an instance of a rule, such that its children are labeled exactly with the elements of the conjunction in the condition of this instance.

Definition 4.1 (And-Tree)
Let $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$ be a defeasible specification (cf. § 2.3), and let L be a literal.
An *and-tree* T for L [and for the derivation of $\Phi \vdash \{L\}$] w.r.t. $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$ is a finite, rooted tree, where every node is labeled with a literal, satisfying the following conditions:

1. The root node of T is labeled with L.
2. For each node N in T that is labeled with a literal L', there is a strict or defeasible rule $(L''_0 \Leftarrow L''_1 \wedge \ldots \wedge L''_k) \in \Pi \cup \Delta$, such that $L' = L''_0\sigma$ for some substitution σ [with $(L''_0\sigma \Leftarrow L''_1\sigma \wedge \ldots \wedge L''_k\sigma) \in \Phi$]. Moreover, the node N has exactly k child nodes, which are labeled with $L''_1\sigma, \ldots, L''_k\sigma$, respectively.

[notion not explained, but okay]

This standard and very simple formal notion of an and-tree is meant to capture a single derivation for a single argument. It must not be confused with the compact multi-graphs that come as a synopsis with our examples (such as the ones in § 3).[6]

An isolation of the defeasible parts of an and-tree of the derivation may now proceed as follows:

- Starting from the root of the tree, we iteratively erase all applications of strict rules. This results in a set of trees, each of which has the application of a defeasible rule at the root.
- Starting now from the leaves of these trees, we again erase all applications of strict rules. This results in a set of trees where all nodes *all* of whose children are leaves result from an application of a defeasible rule.

[I do not understand. They may be strict rules in between, right?]

[6] These sophisticated multi-graphs illustrate several derivations for several arguments in parallel, share sub-graphs, and may have =-edges between occurrences of the same literal L to represent alternative derivations of L (cf. Example 6.9 in § 6.2 as well as Examples 7.4 and 7.5 in § 7.2). Because these synopses are redundant in all examples, we do not provide a formalization for these multi-graphs; ~~either they become clear by immediate intuition or the reader had better ignore them.~~

[There is no reason to insult the reader.]

4.4.4 First Effect: Simplified Second Sketch of a Notion of Specificity

The first effect is that we immediately realize that every model of Π^{G} in the model class that is represented by the activation set $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$ is also in the model class represented by the activation set $\{\mathsf{Q}(\mathsf{a})\}$.

Indeed, this growth toward the leaves will immediately add $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$ as a further activation set for every argument with the activation set $\{\mathsf{Q}(\mathsf{a})\}$. By this effect it is just made explicit that an argument that can be based on the activation set $\{\mathsf{Q}(\mathsf{a})\}$ can also be based on the activation set $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$. Thus — provided that there are no other activation sets — an argument that can be based on the activation set $\{\mathsf{Q}(\mathsf{a})\}$ is less or equivalently specific compared to any argument that can be based on $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$.

Therefore — if we admit the effect of a growth toward the leaves on our activation sets — we may simplify[10] the comparison of activation sets in our first sketch of a notion of specificity of § 4.3.2 as follows:

> An argument $(\mathcal{A}_1, L_1)$ is [properly] *more specific than* an argument $(\mathcal{A}_2, L_2)$ if, for each activation set H_1 for $(\mathcal{A}_1, L_1)$, this set H_1 is also an activation set for $(\mathcal{A}_2, L_2)$ [but not vice versa].

4.4.5 Second Effect: Preference of the "More Concise"

The second effect, however, is that an argument $(\mathcal{A}_2, L_2)$ that gets along with $\{\mathsf{Q}(\mathsf{a})\}$ becomes even *properly* less specific than an argument $(\mathcal{A}_1, L_1)$ that actually requires $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$ and does not get along with $\{\mathsf{Q}(\mathsf{a})\}$,[11] simply because $(\mathcal{A}_2, L_2)$ has the additional activation set $\{\mathsf{Q}(\mathsf{a})\}$.

The resulting preference of $(\mathcal{A}_1, L_1)$ to $(\mathcal{A}_2, L_2)$ as being properly more specific is usually called *preference of the "more concise"*, cf. e.g. [STOLZENBURG &AL., 2003, p. 94], [GARCÍA & SIMARI, 2004, p.108]. Although — to the best of our knowledge — this notion has never been formally defined, roughly speaking it is — for an instantiated rule $\mathsf{Q}(\mathsf{a}) \Leftarrow \mathsf{P}_0(\mathsf{a}) \wedge \cdots \wedge \mathsf{P}_{n-1}(\mathsf{a})$ of the specification — the preference of an argument that gets along with the conclusion $\{\mathsf{Q}(\mathsf{a})\}$ of the instantiated rule as an activation set, instead of actually requiring the condition $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \}$.

For instance, in Example 3.2 of § 3, an argument that gets along with $\{\mathsf{bird}(\mathsf{edna})\}$ is properly less specific than one that actually requires $\{\mathsf{emu}(\mathsf{edna})\}$, in the sense that $\mathsf{emu}(\mathsf{edna})$ is more concise than $\mathsf{bird}(\mathsf{edna})$.

---

[10] Note that we have replaced here the option to choose some activation set $H_2 \subseteq \mathfrak{T}_{H_1 \cup \Pi^{\mathrm{G}}}$ of the first sketch with the restrictive determination $H_2 := H_1$. This simplifying restriction applies here for the following reason: If $H_2 \subseteq \mathfrak{T}_{H_1 \cup \Pi^{\mathrm{G}}}$ is an activation set for $(\mathcal{A}_2, L_2)$, then H_1 is an activation set for $(\mathcal{A}_2, L_2)$ as well, provided that we admit the first effect of a growth toward the leaves via Π^{G} on our activation sets.

[11] This can happen only if we have $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n{-}1\}\ \} \not\subseteq \{\mathsf{Q}(\mathsf{a})\}$, i.e. only if $n \neq 0$.

4.4.6 Preference of the "More Precise"

If we consider an argument requiring an activation set $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in \{0,\ldots,n\}\ \}$ to be *properly* more specific than an argument that gets along with a proper subset $\{\ \mathsf{P}_i(\mathsf{a}) \mid i \in I\ \}$ for some index set $I \subsetneq \{0,\ldots,n\}$, then the resulting preference is usually called *preference of the "more precise"*, cf. e.g. [STOLZENBURG &AL., 2003, p. 94], [GARCÍA & SIMARI, 2004, p.108]. An example for the preference of the "more precise" is Example 3.4 of § 3.

There is, however, an exception from this preference to be observed, namely the case that we actually can derive the set from its subset with the help of Π^{G}. In this case, the above-mentioned growth toward the leaves with rules from Π^{G} again implements the approximation of the subclass relation among model classes via the one among activation sets.[13]

Apart from this exception, there is again a problem, namely that it is not the case that

$$\bigwedge_{i\in I} \mathsf{P}_i(\mathsf{a}) \not\models \bigwedge_{i\in\{0,\ldots,n\}} \mathsf{P}_i(\mathsf{a})$$

would be explicitly given by the specification via $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$.

Nevertheless — if we do not just want to see it as a matter-of-fact property of notions of specificity in the style of POOLE — we could justify also the preference of the "more concise" by imposing the following best practice on positive-conditional specification:

If we want to exclude the above non-consequence, then we ought to specify, for each $j \in \{0,\ldots,n\}\backslash I$, a rule like $\mathsf{P}_j(x) \Leftarrow \bigwedge_{i\in I} \mathsf{P}_i(x)$.

$$\longleftarrow \bigwedge_{i\in I} \mathsf{P}_i(\mathsf{a}) \longrightarrow$$
$$\longleftarrow \bigwedge_{i\in\{1,\ldots,n\}} \mathsf{P}_i(\mathsf{a}) \longrightarrow$$
$$\longleftarrow \mathsf{P}_k(\mathsf{a}) \longrightarrow$$

4.4.7 Conclusion on the Preferences

[style

After all, even if you do not buy our justification of the preference of the "more concise" and the "more precis[e]" you can still follow our investigations into the properties of these preferences w.r.t. POOLE's model-theoretic notion of specificity and our correction of this notion in the following sections.

[12] There is one exception to this justification, however, namely in the practice of *logic programming*: If $\mathsf{Q}(x) \Leftarrow \mathsf{P}_0(x) \wedge \cdots \wedge \mathsf{P}_{n-1}(x)$ is the only rule of the specification with Q as the predicate symbol of the conclusion, then it is standard in PROLOG to consider this implication as an implementation of a full equivalence defining the predicate Q. This corresponds to the \emph{closed world assumption}.

This is different in our context of *positive-conditional specification* here, however, where we can add and ought to add the rules $\mathsf{P}_i(x) \Leftarrow \mathsf{Q}(x)$ $(i \in \{0,\ldots,n-1\})$ to our specification, simply because we are not concerned with the non-termination problem of logic programming resulting from such a specification of the full equivalence (cf. § 2.1).

An alternative which is given also in logic programming is to omit the above indicated rule and to replace each occurrence of each $\mathsf{Q}(t)$ with $\mathsf{P}_0(t) \wedge \cdots \wedge \mathsf{P}_{n-1}(t)$, respectively.

Moreover, in the frequent case that several cases of the definition of a predicate are spread over several rules, the implications definitely tend to be proper also in logic programming, because, roughly speaking, the defined predicate is given as the proper disjunction of the conditions of the several rules.

[very lengthy footnote ⇒ insert in text?]

[13] This approximation was discussed in § 4.4.4 and will be demonstrated in Example 7.7 of § 7.

6 Formalizations of Specificity

6.1 Activation Sets

A generative, bottom-up (i.e. from the leaves to the root) derivation with defeasible rules can now be split into three phases of derivation of literals from literals. This splitting follows the discussion in § 4.4.1 on how to isolate the defeasible parts of a derivation (phase 2) from strict parts that may occur toward the root (phase 3) and toward the leaves (phase 1):

(phase 1) First we derive the literals that provide the basis for specificity considerations.

In our approach we derive the set $\mathfrak{T}_{\Pi}$ here. POOLE takes the set $\mathfrak{T}_{\Pi\cup\Delta}$ instead.

(phase 2) On the basis of

- a subset H of the literals derived in phase 1,
- the first item $\mathcal{A}$ of a given argument $(\mathcal{A}, L)$, and
- the general rules Π^{G},

we derive a further set of literals $\mathfrak{L}$: with the property $H \cup \mathcal{A} \cup \Pi^{\mathrm{G}} \vdash \mathfrak{L}$.

(phase 3) Finally, on the basis of $\mathfrak{L}$, the literal of the given argument $(\mathcal{A}, L)$ is derived:

$$\mathfrak{L} \cup \Pi \vdash \{L\}.$$

[but literals in H can be used to derive L in Poole's approach] 2

In POOLE's approach, phase 3 is empty and we simply have $\mathfrak{L} = \{L\}$. In our approach, however, it is admitted to use the facts from Π^{F} in phase 3, in addition to the general rules from Π^{G}, which were already admitted in phase 2.

With implicit reference to our sets $\Pi = \Pi^{\mathrm{F}} \cup \Pi^{\mathrm{G}}$ and Δ, the phases 2 and 3 can be more easily expressed with the help of the following notions.

Definition 6.1 ([Minimal] [Simplified] Activation Set)
Let $\mathcal{A}$ be a set of ground instances of rules from Δ, and let L be a literal.
H is a *simplified activation set for* $(\mathcal{A}, L)$ if $L \in \mathfrak{T}_{H\cup\mathcal{A}\cup\Pi^{\mathrm{G}}}$.
H is an *activation set for* $(\mathcal{A}, L)$ if $L \in \mathfrak{T}_{\mathfrak{L}\cup\Pi}$ for some $\mathfrak{L} \subseteq \mathfrak{T}_{H\cup\mathcal{A}\cup\Pi^{\mathrm{G}}}$.
H is a *minimal [simplified] activation set for* $(\mathcal{A}, L)$ if H is an [simplified] activation set for $(\mathcal{A}, L)$, but no proper subset of H is an [simplified] activation set for $(\mathcal{A}, L)$.

Corollary 6.2 *Let $\mathcal{A}$ be a set of ground instances of rules from Δ, and let L be a literal. Every simplified activation set for $(\mathcal{A}, L)$ is an activation set for $(\mathcal{A}, L)$.*

Roughly speaking, an argument is now more (or equivalently) specific than another one if each of its activation sets is also an activation set for the other argument. Note that this follows the simplified second sketch of a notion of specificity displayed in § 4.4.4, not the first one displayed in § 4.3.2.

6.2 POOLE's Specificity Relation P1; its Minor Corrections P2, P3

In this section we will define the binary relations $\lesssim_{\mathrm{P1}}$, $\lesssim_{\mathrm{P2}}$, $\lesssim_{\mathrm{P3}}$ of "being more or equivalently specific according to DAVID POOLE" with implicit reference to our sets of facts and of general and defeasible rules (i.e. to Π^{F}, Π^{G}, and Δ, respectively).

The relation $\lesssim_{\mathrm{P1}}$ of the following definition is precisely POOLE's original relation $\geq$ as defined at the bottom of the left column on Page 145 of [POOLE, 1985]. See § 5 for our reasons to write "$\gtrsim$" instead of "$\geq$" as a first change. Moreover, as a second change required by mathematical standards, we have replaced the symbol "$\gtrsim$" with the symbol "$\lesssim$" (such that the smaller argument becomes the more specific one), so that the relevant well-foundedness becomes the one of its ordering $<$ instead of the reverse $>$.

Definition 6.3 ($\lesssim_{\mathrm{P1}}$: DAVID POOLE's Original Specificity)
$(\mathcal{A}_1, L_1) \lesssim_{\mathrm{P1}} (\mathcal{A}_2, L_2)$ if $(\mathcal{A}_1, L_1)$ and $(\mathcal{A}_2, L_2)$ are arguments, and if, for every $H \subseteq \mathfrak{L}_{\Pi\cup\Delta}$ that is a simplified activation set for $(\mathcal{A}_1, L_1)$ but not a simplified activation set for $(\mathcal{A}_2, L_1)$, H is also a simplified activation set for $(\mathcal{A}_2, L_2)$.

The relation $\lesssim_{\mathrm{P2}}$ of the following definition is the relation $\succeq$ of Definition 10 on Page 94 of [STOLZENBURG &AL., 2003] (attributed to [POOLE, 1985]). Moreover, the relation $>_{\mathsf{spec}}$ of Definition 2.12 on Page 132 of [SIMARI & LOUI, 1992] (attributed to [POOLE, 1985] as well) is the relation $<_{\mathrm{P2}} \ := \ \lesssim_{\mathrm{P2}} \setminus \gtrsim_{\mathrm{P2}}$.

Definition 6.4 ($\lesssim_{\mathrm{P2}}$: Standard Version of DAVID POOLE's Specificity)
$(\mathcal{A}_1, L_1) \lesssim_{\mathrm{P2}} (\mathcal{A}_2, L_2)$ if $(\mathcal{A}_1, L_1)$ and $(\mathcal{A}_2, L_2)$ are arguments, and if, for every $H \subseteq \mathfrak{L}_{\Pi\cup\Delta}$ that is a simplified activation set for $(\mathcal{A}_1, L_1)$ but not a simplified activation set for $(\emptyset, L_1)$, H is also a simplified activation set for $(\mathcal{A}_2, L_2)$.

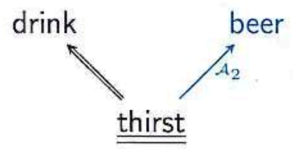
Why is Def. 6.3 counter-intuitive? Explain!

The only change in Definition 6.4 as compared to Definition 6.3 is that "$(\mathcal{A}_2, L_1)$" is replaced with "$(\emptyset, L_1)$". We did not encounter any example yet where this intuitively most appropriate correction of the counter-intuitive variant "$(\mathcal{A}_2, L_1)$" of Definition 6.3 makes any difference to "$(\emptyset, L_1)$" in Definition 6.4 (which is standard in the publications of the last two decades), and leave it as an exercise to construct one.

The relations $\lesssim_{\mathrm{P1}}$ and $\lesssim_{\mathrm{P2}}$ were not meant to compare arguments for literals that do not need any defeasible rules — or at least they do not show an intuitive behavior on such arguments, as shown in Example 6.5.

Example 6.5 (Minor Flaw of $\lesssim_{\mathrm{P1}}$ and $\lesssim_{\mathrm{P2}}$)

$\Pi^{\mathrm{F}}_{6.5} \ := \ \{ \ \mathsf{thirst} \ \}$,
$\Pi^{\mathrm{G}}_{6.5} \ := \ \{ \ \mathsf{drink} \Leftarrow \mathsf{thirst} \ \}$,
$\Delta_{6.5} \ := \ \mathcal{A}_2$.
$\mathcal{A}_2 \ := \ \{ \ \mathsf{beer} \leftarrow \mathsf{thirst} \ \}$.

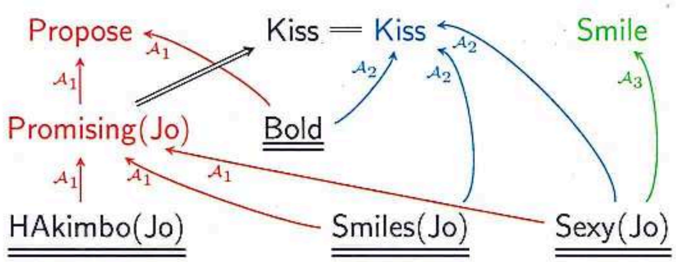


Let us compare the specificity of the arguments $(\mathcal{A}_2, \mathsf{beer})$ and $(\emptyset, \mathsf{drink})$.

Example 6.9 (Counterexample to the Transitivities: "Choose one action!")
Suppose you meet the sexy girl Jo in a lift for a very short time, you smile at her, and she smiles back with a head akimbo. Since smiling, kissing, and proposing are mutually exclusive actions of your mouth, you have to make up your mind quickly what to do next, depending on your current level of boldness.[26]

$$\Pi^{\mathrm{F}}_{6.9} := \{\ \mathsf{Bold},\ \mathsf{HAkimbo}(\mathsf{Jo}),\ \mathsf{Smiles}(\mathsf{Jo}),\ \mathsf{Sexy}(\mathsf{Jo})\ \},$$
$$\Pi^{\mathrm{G}}_{6.9} := \{\ \mathsf{Kiss} \Leftarrow \mathsf{Promising}(G)\ \},$$
$$\Delta_{6.9} := \left\{ \begin{array}{l} \mathsf{Smile} \leftarrow \mathsf{Sexy}(G), \\ \mathsf{Kiss} \leftarrow \mathsf{Bold} \wedge \mathsf{Smiles}(G) \wedge \mathsf{Sexy}(G), \\ \mathsf{Promising}(G) \leftarrow \mathsf{HAkimbo}(G) \wedge \mathsf{Smiles}(G) \wedge \mathsf{Sexy}(G), \\ \mathsf{Propose} \leftarrow \mathsf{Promising}(G) \wedge \mathsf{Bold} \end{array} \right\}.$$
$$\mathcal{A}_1 := \left\{ \begin{array}{l} \mathsf{Promising}(\mathsf{Jo}) \leftarrow \mathsf{HAkimbo}(\mathsf{Jo}) \wedge \mathsf{Smiles}(\mathsf{Jo}) \wedge \mathsf{Sexy}(\mathsf{Jo}) \\ \mathsf{Propose} \leftarrow \mathsf{Promising}(\mathsf{Jo}) \wedge \mathsf{Bold} \end{array} \right\},$$
$$\mathcal{A}_2 := \{\ \mathsf{Kiss} \leftarrow \mathsf{Bold} \wedge \mathsf{Smiles}(\mathsf{Jo}) \wedge \mathsf{Sexy}(\mathsf{Jo})\ \},$$
$$\mathcal{A}_3 := \{\ \mathsf{Smile} \leftarrow \mathsf{Sexy}(\mathsf{Jo})\ \}.$$

Compare the specificity of the arguments $(\mathcal{A}_1, \mathsf{Propose})$, $(\mathcal{A}_2, \mathsf{Kiss})$, $(\mathcal{A}_3, \mathsf{Smile})$!



Lemma 6.10 *There are*

- *a specification* $(\Pi^{\mathrm{F}}_{6.9}, \Pi^{\mathrm{G}}_{6.9}, \Delta_{6.9})$ *without any negative literals (i.e., a fortiori,* $\Pi^{\mathrm{F}}_{6.9} \cup \Pi^{\mathrm{G}}_{6.9} \cup \Delta_{6.9}$ *is non-contradictory), and*
- *arguments* $(\mathcal{A}_1, L_1)$, $(\mathcal{A}_2, L_2)$, $(\mathcal{A}_3, L_3)$ *with respective minimal sets* $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ *(i.e.,* $(\mathcal{A}'_i, L_i)$ *is not an argument for any proper subset* $\mathcal{A}'_i \subsetneq \mathcal{A}_i$*),*

such that $(\mathcal{A}_1, L_1) \lesssim_{\mathrm{P3}} (\mathcal{A}_2, L_2) \lesssim_{\mathrm{P3}} (\mathcal{A}_3, L_3) \not\gtrsim_{\mathrm{P1}} (\mathcal{A}_1, L_1)$
and $(\mathcal{A}_1, L_1) \not\gtrsim_{\mathrm{P1}} (\mathcal{A}_2, L_2) \not\gtrsim_{\mathrm{P1}} (\mathcal{A}_3, L_3)$.

[26] The nullary predicate Bold could actually be removed from all rules and facts of this example, which would still remain a counterexample to the transitivities; to the contrary, it would even improve its status by becoming a *minimal* counterexample. A renaming of the resulting minimal counterexample was presented as Example 5.8 in [Wirth & Stolzenburg, 2013; 2014]. The reasons we prefer the non-minimal counterexample are the following. The minimal counterexample looks artificial because the single general strict rule lacks any effect on activation sets for arguments for $(\mathcal{A}_1 \cup \mathcal{A}_2, \mathsf{Kiss})$. Moreover, the minimal counterexample used to confuse the audience during presentations because the names of its predicates mixed up different cognitive categories.

[This is not interesting for most of the readers.]

Moreover, this consequence is also immediate for the relation $\succeq$ [STOLZENBURG &AL., 2003, Definition 10, p. 94] and for the relation $>_{\rm spec}$ [SIMARI & LOUI, 1992, Definition 2.12, p.132], simply because we can replace $\succeq$ and $>_{\rm spec}$ with $\precsim_{\rm P2}$ and $<_{\rm P2}$ in the context of Example 6.9, respectively.

Although transitivity of these relations is strongly suggested by the special choice of their symbols and seems to be taken for granted in general, we found an actual statement of such a transitivity only for the relation $\sqsupseteq$ of Definition 2.22 on Page 134 of [SIMARI & LOUI, 1992], namely in "Lemma 2.23" [SIMARI & LOUI, 1992, p.134].[27]

Finally, note that those readers who do not see a proper conflict in our counter-example just should add to Example 6.9 some general rules such as $\mathsf{Execute} \Leftarrow \mathsf{Kiss}$, $\neg\mathsf{Execute} \Leftarrow \mathsf{Smile}$, $\neg\mathsf{Execute} \Leftarrow \mathsf{Propose}$, say to model the situation in one of the areas of today's planet Earth where an unmarried woman who raises the wish to kiss has to be executed.

[can be omitted]

6.4 Our Novel Specificity Ordering CP1

In the previous section, we have seen that *minor corrections* of POOLE's original relation P1 (such as P2, P3) do not cure the (up to our finding of Example 6.9) hidden or even denied deficiency of these relations, namely their lack of transitivity. Our true motivation for a *major correction* of P3 was not this formal deficiency, but actually an informal one, namely that it failed to get sufficiently close to human intuition, which will become even more clear in § 7 than in §§ 3 and 4.

For these reasons, we now define our major correction of POOLE's specificity — the binary relation $\precsim_{\rm CP1}$ — with implicit reference to our sets of facts and of general and defeasible rules (i.e. to $\Pi^{\rm F}$, $\Pi^{\rm G}$, and Δ, respectively) as follows.

Definition 6.12 ($\precsim_{\rm CP1}$: 1st Version of our Specificity Relation)
$(\mathcal{A}_1, L_1) \precsim_{\rm CP1} (\mathcal{A}_2, L_2)$ if $(\mathcal{A}_1, L_1)$ and $(\mathcal{A}_2, L_2)$ are arguments, and we have

1. $L_1 \in \mathfrak{T}_\Pi$ or
2. $L_2 \notin \mathfrak{T}_\Pi$ and every $H \subseteq \mathfrak{T}_\Pi$ that is an [minimal][28] activation set for $(\mathcal{A}_1, L_1)$ is also an activation set for $(\mathcal{A}_2, L_2)$.

[27] According to the rules of good scientific and historiographic practice, we pinpoint the violation of this "lemma" now as follows. Non-transitivity of $\sqsupseteq$ follows here immediately from the non-transitivity of the relation $\geq_{\rm spec}$ of Definition 2.15, which, however, is not identical to the above-mentioned relation $\succeq$, but actually a subset of $\succeq$, because it is defined via a peculiar additional equivalence $\approx_{\rm spec}$ introduced in Definition 2.14 [SIMARI & LOUI, 1992, p.132], namely via $\geq_{\rm spec} := >_{\rm spec} \cup \approx_{\rm spec}$ [SIMARI & LOUI, 1992, Definition 2.15, p.132f.]. Directly from Definition 2.14 of [SIMARI & LOUI, 1992], we get $\approx_{\rm spec} \subseteq \approx_{\rm P2}$. Thus, by Corollary 6.8, we get $\geq_{\rm spec} \subseteq \precsim_{\rm P2} \subseteq \precsim_{\rm P1}$; and so (recollecting $<_{\rm P2} \subseteq >_{\rm spec} \subseteq \geq_{\rm spec}$) the result

$$(\mathcal{A}_1, L_1) <_{\rm P2} (\mathcal{A}_2, L_2) <_{\rm P2} (\mathcal{A}_3, L_3) \not\precsim_{\rm P1} (\mathcal{A}_1, L_1)$$

of Lemma 6.10 gives us the following counterexample to transitivity:

$$(\mathcal{A}_1, L_1) \geq_{\rm spec} (\mathcal{A}_2, L_2) \geq_{\rm spec} (\mathcal{A}_3, L_3) \not\leq_{\rm spec} (\mathcal{A}_1, L_1).$$

[28] Note that the omission of the optional restriction to *minimal* activation sets for $(\mathcal{A}_1, L_1)$ in Definition 6.12 has no effect on the extension of the defined notion, simply because the additional non-minimal activation sets for $(\mathcal{A}_1, L_1)$ will then be activation sets for $(\mathcal{A}_2, L_2)$ *a fortiori*.

6.5 Relation between the Specificity Relations P3 and CP1

Theorem 6.16 *Let $\Pi^{<2}$ be the set of rules from Π that are unconditional or have exactly one literal in the conjunction of their condition.*

Let $\Pi^{\geq 2}$ be the set of rules from Π with more than one literal in their condition.

$\precsim_{\mathrm{P3}} \subseteq \precsim_{\mathrm{CP1}}$ holds if one (or more) of the following conditions hold:

1. *For every $H \subseteq \mathfrak{T}_\Pi$ and for every set $\mathcal{A}$ of ground instances of rules from Δ, and for $\mathfrak{L} := \mathfrak{T}_{H \cup \mathcal{A} \cup \Pi^{\mathrm{G}}}$, we have $\mathfrak{T}_{\mathfrak{L} \cup \Pi} \subseteq \mathfrak{L} \cup \mathfrak{T}_\Pi$.* ~~[Does this always hold??]~~
2. *For each instance $L \Leftarrow L'_0 \wedge \ldots \wedge L'_{n+1}$ of each rule in $\Pi^{\geq 2}$ with $L \notin \mathfrak{T}_{\Pi^{<2}}$, we have $L'_j \notin \mathfrak{T}_{\Pi^{<2}}$ for all $j \in \{0, \ldots, n+1\}$.* [strange numbering]
3. *For each instance $L \Leftarrow L'_0 \wedge \ldots \wedge L'_{n+1}$ of each rule in $\Pi^{\geq 2}$, we have $L'_j \notin \mathfrak{T}_\Pi$ for all $j \in \{0, \ldots, n+1\}$.*
4. *We have $\Pi^{\geq 2} = \emptyset$.* [strange numbering]

Note that if we had improved $\precsim_{\mathrm{P3}}$ only w.r.t. phase 1 of § 6.1, but not w.r.t. phase 3 in addition, then Theorem 6.16 would not require any condition at all (See the proof). This means that a condition becomes necessary by our correction of simplified activation sets to non-simplified ones, but not because of the major changes (A) and (B) of § 6.4.

Proof of Theorem 6.16

First let us show that condition 2 implies condition 1. To this end, let $H \subseteq \mathfrak{T}_\Pi$, let $\mathcal{A}$ be a set of ground instances of rules from Δ, and set $\mathfrak{L} := \mathfrak{T}_{H \cup \mathcal{A} \cup \Pi^{\mathrm{G}}}$. For an *argumentum ad absurdum*, let us assume $\mathfrak{T}_{\mathfrak{L} \cup \Pi} \not\subseteq \mathfrak{L} \cup \mathfrak{T}_\Pi$. Because of $\Pi^{\mathrm{F}} \subseteq \mathfrak{T}_{\Pi^{<2}}$, we have $\mathfrak{L} \cup \Pi = \mathfrak{L} \cup \Pi^{\mathrm{F}} \cup \Pi^{\mathrm{G}} \subseteq \mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \cup \Pi^{\mathrm{G}}$, and thus $\mathfrak{T}_{\mathfrak{L} \cup \Pi} \subseteq \mathfrak{T}_{\mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \cup \Pi^{\mathrm{G}}}$, and thus $\mathfrak{T}_{\mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \cup \Pi^{\mathrm{G}}} \not\subseteq \mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}}$ (because otherwise $\mathfrak{T}_{\mathfrak{L} \cup \Pi} \subseteq \mathfrak{T}_{\mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \cup \Pi^{\mathrm{G}}} \subseteq \mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \subseteq \mathfrak{L} \cup \mathfrak{T}_\Pi$). Now $\mathfrak{L}$ is closed under Π^{G} by definition. Moreover, $\mathfrak{T}_{\Pi^{<2}}$ is closed under $\Pi^{<2}$ by definition and under $\Pi^{\geq 2}$ by condition 2. Because both of the sets of literals $\mathfrak{L}$ and $\mathfrak{T}_{\Pi^{<2}}$ are closed under Π^{G} — but nevertheless their union is not closed under Π^{G} according to $\mathfrak{T}_{\mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}} \cup \Pi^{\mathrm{G}}} \not\subseteq \mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}}$ — there must be an inference step *essentially based on both sets in parallel.* More precisely, this means that there must be an instance $L \Leftarrow L'_1 \wedge \ldots \wedge L'_n$ of a rule from Π^{G} with $L \notin \mathfrak{L} \cup \mathfrak{T}_{\Pi^{<2}}$, and some $i, j \in \{1, \ldots, n\}$ with $L'_i \in \mathfrak{L} \setminus \mathfrak{T}_{\Pi^{<2}}$ and $L'_j \in \mathfrak{T}_{\Pi^{<2}} \setminus \mathfrak{L}$. Then $L \Leftarrow L'_1 \wedge \ldots \wedge L'_n$ must actually be an instance of a rule from $\Pi^{\geq 2}$, and $L \notin \mathfrak{T}_{\Pi^{<2}}$, but $L'_j \in \mathfrak{T}_{\Pi^{<2}}$ in contradiction to condition 2.

[Why not simply mention only cond. 1?]

As condition 2 implies condition 1, condition 3 trivially implies condition 2, and condition 4 trivially implies condition 3, it now suffices to show the claim that $(\mathcal{A}_1, L_1) \precsim_{\mathrm{CP1}} (\mathcal{A}_2, L_2)$ holds under condition 1 and the assumption of $(\mathcal{A}_1, L_1) \precsim_{\mathrm{P3}} (\mathcal{A}_2, L_2)$. By this assumption, $(\mathcal{A}_1, L_1)$ and $(\mathcal{A}_2, L_2)$ are arguments and $L_2 \in \mathfrak{T}_\Pi$ implies $L_1 \in \mathfrak{T}_\Pi$. If $L_1 \in \mathfrak{T}_\Pi$ holds, then our claim holds as well. Otherwise, we have $L_1, L_2 \notin \mathfrak{T}_\Pi$, and it suffices to show the sub-claim that H is an activation set for $(\mathcal{A}_2, L_2)$ under the additional sub-assumption that $H \subseteq \mathfrak{T}_\Pi$ is an activation set for $(\mathcal{A}_1, L_1)$. Under the sub-assumption we also have $H \subseteq \mathfrak{T}_{\Pi \cup \Delta}$ because of $\mathfrak{T}_\Pi \subseteq \mathfrak{T}_{\Pi \cup \Delta}$, and, for $\mathfrak{L} := \mathfrak{T}_{H \cup \mathcal{A}_1 \cup \Pi^{\mathrm{G}}}$, we have $L_1 \in \mathfrak{T}_{\mathfrak{L} \cup \Pi}$, and then, by condition 1, $L_1 \in \mathfrak{L} \cup \mathfrak{T}_\Pi$. Then, by our current case of $L_1, L_2 \notin \mathfrak{T}_\Pi$, we have $L_1 \in \mathfrak{L}$. Thus, H is a *simplified* activation set for $(\mathcal{A}_1, L_1)$.

$(\emptyset, \neg\mathsf{flies(edna)})$, then we have $\mathsf{emu(edna)} \in H$, and thus H is a simplified activation set also for $(\mathcal{A}_2, \mathsf{flies(edna)})$.

All in all, by Theorem 6.16, we get $(\mathcal{A}_1, \neg\mathsf{flies(edna)}) <_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{flies(edna)})$
and $(\mathcal{A}_1, \neg\mathsf{flies(edna)}) \lesssim_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{flies(edna)})$.

Example 6.20 *(continuing Example 3.3 of § 3)*
We have $(\mathcal{A}_2, \mathsf{flies(edna)}) \lesssim_{\mathrm{CP1}} (\mathcal{A}_1, \neg\mathsf{flies(edna)})$ because $\neg\mathsf{flies(edna)} \notin \mathfrak{T}_{\Pi_{3.3}}$ and, for every activation set $H \subseteq \mathfrak{T}_{\Pi_{3.3}}$ for $(\mathcal{A}_2, \mathsf{flies(edna)})$, we get $\mathsf{emu(edna)} \in H$, and so H is an activation set also for $(\mathcal{A}_1, \neg\mathsf{flies(edna)})$.

Nevertheless, we have $(\mathcal{A}_2, \mathsf{flies(edna)}) \not\lesssim_{\mathrm{P3}} (\mathcal{A}_1, \neg\mathsf{flies(edna)})$, because $\{\mathsf{bird(edna)}\} \subseteq \mathfrak{T}_{\Pi_{3.3} \cup \Delta_{3.3}}$ is a simplified activation set for $(\mathcal{A}_2, \mathsf{flies(edna)})$, but neither for $(\emptyset, \mathsf{flies(edna)})$, nor for $(\mathcal{A}_1, \neg\mathsf{flies(edna)})$.

We have $(\mathcal{A}_1, \neg\mathsf{flies(edna)}) \lesssim_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{flies(edna)})$, because of $\mathsf{flies(edna)} \notin \mathfrak{T}_{\Pi_{3.3}}$ and because, if $H \subseteq \mathfrak{T}_{\Pi_{3.3} \cup \Delta_{3.3}}$ is a simplified activation set for $(\mathcal{A}_1, \neg\mathsf{flies(edna)})$, but not for $(\emptyset, \neg\mathsf{flies(edna)})$, then we have $\mathsf{emu(edna)} \in H$ and thus H is a simplified activation set also for $(\mathcal{A}_2, \mathsf{flies(edna)})$.

All in all, by Theorem 6.16, we get $(\mathcal{A}_1, \neg\mathsf{flies(edna)}) \approx_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{flies(edna)})$
and $(\mathcal{A}_1, \neg\mathsf{flies(edna)}) <_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{flies(edna)})$.

From a conceptual point of view, we have to ask ourselves, whether we would like the two *defeasible* rule instances in $\mathcal{A}_2 = \{\ \mathsf{flies(edna)} \leftarrow \mathsf{bird(edna)},\ \mathsf{bird(edna)} \leftarrow \mathsf{emu(edna)}\ \}$ to reduce the specificity of $(\mathcal{A}_2, \mathsf{flies(edna)})$ as compared to a system that seems equivalent for the given argument for $\mathsf{flies(edna)}$, namely the argument $(\{\mathsf{flies(edna)} \leftarrow \mathsf{emu(edna)}\}, \mathsf{flies(edna)})$.

Does the specificity of a defeasible reasoning step really reduce if we introduce intermediate literals (such as $\mathsf{bird(edna)}$ between $\mathsf{flies(edna)}$ and $\mathsf{emu(edna)}$)?

According to human intuition, this question has a negative answer, as we have already explained in Remark 4.2 at the end of § 4.4.5.[32]

Example 6.21 *(continuing Example 3.4 of § 3)*
We have $(\mathcal{A}_2, \mathsf{lovely}) \not\lesssim_{\mathrm{CP1}} (\mathcal{A}_1, \neg\mathsf{lovely})$ because $\mathsf{lovely} \notin \mathfrak{T}_{\Pi_{3.4}}$ and because $\{\mathsf{somebody}\} \subseteq \mathfrak{T}_{\Pi_{3.4}}$ is an activation set for $(\mathcal{A}_2, \mathsf{lovely})$, but not for $(\mathcal{A}_1, \neg\mathsf{lovely})$.

We have $(\mathcal{A}_1, \neg\mathsf{lovely}) \lesssim_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{lovely})$ because of $\mathsf{lovely} \notin \mathfrak{T}_{\Pi_{3.4}}$ and because, if $H \subseteq \mathfrak{T}_{\Pi_{3.4} \cup \Delta_{3.4}}$ is a simplified activation set for $(\mathcal{A}_1, \neg\mathsf{lovely})$, but not for $(\emptyset, \neg\mathsf{lovely})$, then we have $\{\mathsf{somebody}, \mathsf{noisy}\} \subseteq H$, and so H is also a simplified activation set for $(\mathcal{A}_2, \mathsf{lovely})$.

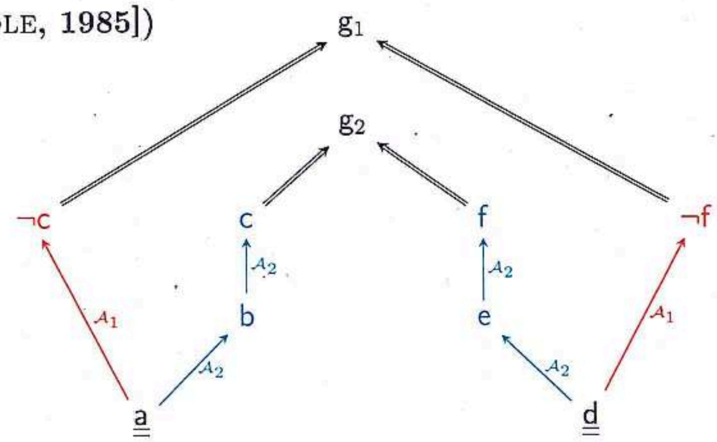
All in all, by Theorem 6.16, we get $(\mathcal{A}_1, \neg\mathsf{lovely}) <_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{lovely})$
and $(\mathcal{A}_1, \neg\mathsf{lovely}) <_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{lovely})$.

Note that we can nicely see here that the condition that H is not a simplified activation set for $(\emptyset, \neg\mathsf{lovely})$ is relevant in Definition 6.6. Without this condition we would have to consider the simplified activation set $\{\mathsf{grandpa}\}$ for $(\mathcal{A}_1, \neg\mathsf{lovely})$, which is not an activation set for $(\mathcal{A}_2, \mathsf{lovely})$; and so, contrary to our intuition, $(\mathcal{A}_1, \neg\mathsf{lovely})$ would not be more specific than $(\mathcal{A}_2, \mathsf{lovely})$ w.r.t. $\lesssim_{\mathrm{P3}}$ anymore.

[32] Moreover, Examples 7.1 and 7.2 will exhibit a strong reason to deny this question: the requirement of monotonicity w.r.t. conjunction. Furthermore, see Example 7.3 for another example that makes even clearer why defeasible rules should be considered for their global semantical effect instead of their syntactical fine structure.

Example 7.1 (Example 6 of [POOLE, 1985])

$$\Pi^{\mathrm{F}}_{7.1} := \{\ \mathsf{a},\ \mathsf{d}\ \},$$
$$\Pi^{\mathrm{G}}_{7.1} := \left\{ \begin{array}{l} \mathsf{g}_1 \Leftarrow \neg\mathsf{c}\wedge\neg\mathsf{f}, \\ \mathsf{g}_2 \Leftarrow \mathsf{c}\wedge\mathsf{f} \end{array} \right\},$$
$$\Delta_{7.1} := \mathcal{A}_1 \cup \mathcal{A}_2.$$
$$\mathcal{A}_1 := \left\{ \begin{array}{l} \neg\mathsf{c} \leftarrow \mathsf{a}, \\ \neg\mathsf{f} \leftarrow \mathsf{d} \end{array} \right\}.$$
$$\mathcal{A}_2 := \left\{ \begin{array}{l} \mathsf{b} \leftarrow \mathsf{a}, \\ \mathsf{c} \leftarrow \mathsf{b}, \\ \mathsf{e} \leftarrow \mathsf{d}, \\ \mathsf{f} \leftarrow \mathsf{e} \end{array} \right\}.$$



Let us compare the specificity of the arguments $(\mathcal{A}_1, \mathsf{g}_1)$ and $(\mathcal{A}_2, \mathsf{g}_2)$.

We have $(\mathcal{A}_1, \mathsf{g}_1) \approx_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{g}_2)$ because $H \subseteq \mathfrak{T}_{\Pi_{7.1}} = \{\mathsf{a}, \mathsf{d}\}$ is an activation set for $(\mathcal{A}_i, \mathsf{g}_i)$ if and only if $H = \{\mathsf{a}, \mathsf{d}\}$.

We have $(\mathcal{A}_1, \mathsf{g}_1)\ \Delta_{\mathrm{P3}}\ (\mathcal{A}_2, \mathsf{g}_2)$ for the following reasons: $\{\mathsf{a}, \neg\mathsf{f}\} \subseteq \mathfrak{T}_{\Pi_{7.1}\cup\Delta_{7.1}}$ is a simplified activation set for $(\mathcal{A}_1, \mathsf{g}_1)$, but neither for $(\emptyset, \mathsf{g}_1)$, nor for $(\mathcal{A}_2, \mathsf{g}_2)$. $\{\mathsf{a}, \mathsf{f}\} \subseteq \mathfrak{T}_{\Pi_{7.1}\cup\Delta_{7.1}}$ is a simplified activation set for $(\mathcal{A}_2, \mathsf{g}_2)$, but neither for $(\emptyset, \mathsf{g}_2)$, nor for $(\mathcal{A}_1, \mathsf{g}_1)$.

POOLE [1985] considers the same result for $\lesssim_{\mathrm{P1}}$ as for $\lesssim_{\mathrm{P3}}$ to be "seemingly unintuitive", because, as we have seen for the isomorphic sub-specification in Example 3.3 of § 3, we have both $(\mathcal{A}_1, \neg\mathsf{c}) <_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{c})$ and $(\mathcal{A}_1, \neg\mathsf{f}) <_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{f})$.

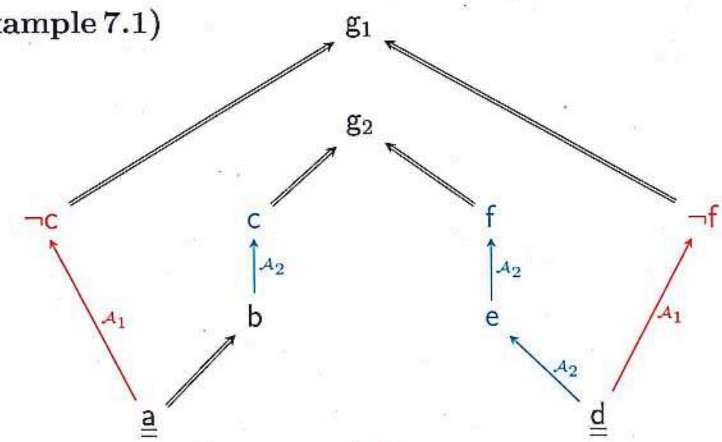
Indeed, as already listed as an essential requirement in § 5, the conjunction of two respectively more specific derivations should be more specific.

On the other hand, considering $\lesssim_{\mathrm{CP1}}$ instead of $\lesssim_{\mathrm{P3}}$, the conjunctions of two equivalently specific argument are equivalently specific — exactly as one intuitively expects.

By turning the defeasible rule $\mathsf{b}\leftarrow\mathsf{a}$ of Example 7.1 into a strict general rule, we obtain the following example.

Example 7.2 (1st Variation of Example 7.1)

$$\Pi^{\mathrm{F}}_{7.2} := \{\ \mathsf{a},\ \mathsf{d}\ \},$$
$$\Pi^{\mathrm{G}}_{7.2} := \left\{ \begin{array}{l} \mathsf{g}_1 \Leftarrow \neg\mathsf{c}\wedge\neg\mathsf{f}, \\ \mathsf{g}_2 \Leftarrow \mathsf{c}\wedge\mathsf{f}, \\ \mathsf{b} \Leftarrow \mathsf{a} \end{array} \right\},$$
$$\Delta_{7.2} := \mathcal{A}_1 \cup \mathcal{A}_2.$$
$$\mathcal{A}_1 := \left\{ \begin{array}{l} \neg\mathsf{c} \leftarrow \mathsf{a}, \\ \neg\mathsf{f} \leftarrow \mathsf{d} \end{array} \right\}.$$
$$\mathcal{A}_2 := \left\{ \begin{array}{l} \mathsf{c} \leftarrow \mathsf{b}, \\ \mathsf{e} \leftarrow \mathsf{d}, \\ \mathsf{f} \leftarrow \mathsf{e} \end{array} \right\}.$$



Let us compare the specificity of the arguments $(\mathcal{A}_1, \mathsf{g}_1)$ and $(\mathcal{A}_2, \mathsf{g}_2)$.

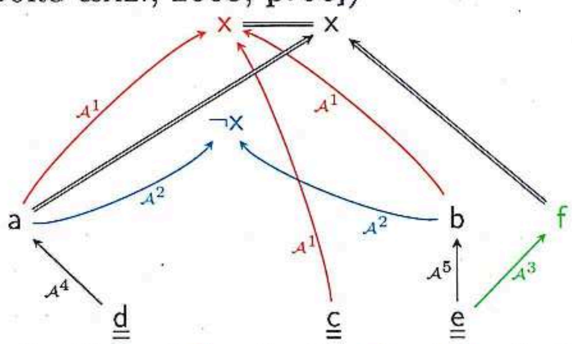
We have $(\mathcal{A}_2, \mathsf{g}_2) \not\lesssim_{\mathrm{CP1}} (\mathcal{A}_1, \mathsf{g}_1)$ because $\{\mathsf{b}, \mathsf{d}\} \subseteq \mathfrak{T}_{\Pi_{7.2}} = \{\mathsf{a}, \mathsf{b}, \mathsf{d}\}$ is an activation set for $(\mathcal{A}_2, \mathsf{g}_2)$, but not for $(\mathcal{A}_1, \mathsf{g}_1)$.

We have $(\mathcal{A}_1, \mathsf{g}_1) \lesssim_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{g}_2)$ because, for every activation set $H \subseteq \mathfrak{T}_{\Pi_{7.2}}$ for $(\mathcal{A}_1, \mathsf{g}_1)$, we have $\{\mathsf{a}, \mathsf{d}\} \subseteq H$; and so H is also an activation set for $(\mathcal{A}_2, \mathsf{g}_2)$.

We again have $(\mathcal{A}_1, \mathsf{g}_1)\ \Delta_{\mathrm{P3}}\ (\mathcal{A}_2, \mathsf{g}_2)$, for the same reason as in Example 7.1. Thus, the situation for $\lesssim_{\mathrm{P3}}$ is just as in Example 7.1, and just as "seemingly unintuitive" for exactly implement the intuition that defeasible rules should be considered for their global semantical effect instead of their syntactical fine structure.

Example 7.4 (Example 11 from [STOLZENBURG &AL., 2003, p. 96])

$\Pi^{\mathrm{F}}_{7.4} \;:=\; \{\, \mathsf{c},\ \mathsf{d},\ \mathsf{e} \,\},$
$\Pi^{\mathrm{G}}_{7.4} \;:=\; \{\, \mathsf{x} \Leftarrow \mathsf{a}\wedge\mathsf{f} \,\},$
$\Delta_{7.4} \;:=\; \mathcal{A}^1 \cup \mathcal{A}^2 \cup \mathcal{A}^3 \cup \mathcal{A}^4 \cup \mathcal{A}^5.$
$\mathcal{A}^1 \;:=\; \{\, \mathsf{x} \leftarrow \mathsf{a}\wedge\mathsf{b}\wedge\mathsf{c} \,\}.$
$\mathcal{A}^2 \;:=\; \{\, \neg\mathsf{x} \leftarrow \mathsf{a}\wedge\mathsf{b} \,\}.$
$\mathcal{A}^3 \;:=\; \{\, \mathsf{f} \leftarrow \mathsf{e} \,\}.$
$\mathcal{A}^4 \;:=\; \{\, \mathsf{a} \leftarrow \mathsf{d} \,\}.$
$\mathcal{A}^5 \;:=\; \{\, \mathsf{b} \leftarrow \mathsf{e} \,\}.$



Compare the specificity of the arguments $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x})$, $(\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$, $(\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x})$!

We have $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) <_{\mathrm{CP1}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x}) \approx_{\mathrm{CP1}} (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x})$, because of $\mathsf{x}, \neg\mathsf{x} \not\in \mathfrak{T}_{\Pi_{7.4}}$, and because any activation set $H \subseteq \mathfrak{T}_{\Pi_{7.4}} = \{\mathsf{c}, \mathsf{d}, \mathsf{e}\}$ for any of $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x})$, $(\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$, $(\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x})$ contains $\{\mathsf{d}, \mathsf{e}\}$, which is an activation set only for the latter two.

This matches our intuition well, because the first of these arguments essentially requires the "more precise" $\mathsf{c}\wedge\mathsf{d}\wedge\mathsf{e}$ instead of the less specific $\mathsf{d}\wedge\mathsf{e}$.

We have $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) \vartriangle_{\mathrm{P3}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x}) \vartriangle_{\mathrm{P3}} (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x}) \vartriangle_{\mathrm{P3}} (\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x})$, however.[34] This means that $\lesssim_{\mathrm{P3}}$ cannot compare these counterarguments and cannot help us to pick the more specific argument.

What is most interesting under the computational aspect is that, for realizing

$$(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) \not\lesssim_{\mathrm{P3}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x}),$$

we have to consider the simplified activation set $\{\mathsf{d}, \mathsf{f}\} \subseteq \mathfrak{T}_{\Pi_{7.4}\cup\Delta_{7.4}}$ for $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x})$. This means that here — to realize that $\mathsf{f} \in \mathfrak{T}_{\Pi_{7.4}\cup\Delta_{7.4}}$ — we have to take into account the defeasible rule of $\mathcal{A}^3$, which is not part of any of the two arguments under comparison.[35]

Note that such considerations are not required, however, for realizing the properties of $\lesssim_{\mathrm{CP1}}$, because defeasible rules not in the given argument can be completely ignored when calculating the minimal activation sets as subsets of $\mathfrak{T}_{\Pi}$ instead of $\mathfrak{T}_{\Pi\cup\Delta}$. In particular, the complication of *pruning* — as discussed in detail in [STOLZENBURG &AL., 2003, § 3.3] — does not have to be considered for the operationalization of $\lesssim_{\mathrm{CP1}}$.

derivation and-trees

inline or omit

[34] Because $\{\mathsf{d}, \mathsf{f}\} \subseteq \mathfrak{T}_{\Pi_{7.4}\cup\Delta_{7.4}}$ is a simplified activation set for $(\mathcal{A}^4, \mathsf{x})$, but neither for $(\emptyset, \mathsf{x})$, nor for $(\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$, we have $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) \not\lesssim_{\mathrm{P3}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x}) \not\gtrsim_{\mathrm{P3}} (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x})$.
Because of $\quad (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x}) \not\lesssim_{\mathrm{CP1}} (\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) \not\gtrsim_{\mathrm{CP1}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$,
we have $\quad (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x}) \not\lesssim_{\mathrm{P3}} (\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}) \not\gtrsim_{\mathrm{P3}} (\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$
by Theorem 6.16. Because $\{\mathsf{b}, \mathsf{c}, \mathsf{d}\} \subseteq \mathfrak{T}_{\Pi_{7.4}\cup\Delta_{7.4}}$ is a simplified activation set for $(\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x})$ and $(\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x})$, but for none of $(\emptyset, \neg\mathsf{x})$, $(\emptyset, \mathsf{x})$, $(\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x})$, we have

$$(\mathcal{A}^2{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \neg\mathsf{x}) \not\lesssim_{\mathrm{P3}} (\mathcal{A}^3{\cup}\mathcal{A}^4, \mathsf{x}) \not\gtrsim_{\mathrm{P3}} (\mathcal{A}^1{\cup}\mathcal{A}^4{\cup}\mathcal{A}^5, \mathsf{x}).$$

[35] Have a look at Figure 1 in § 6.1 to see that the effect of f proceeds here only via the set F, but not via the usage of the set H at the bottom of Figure 1.

Example 7.6 (Variation of Example 7.2)

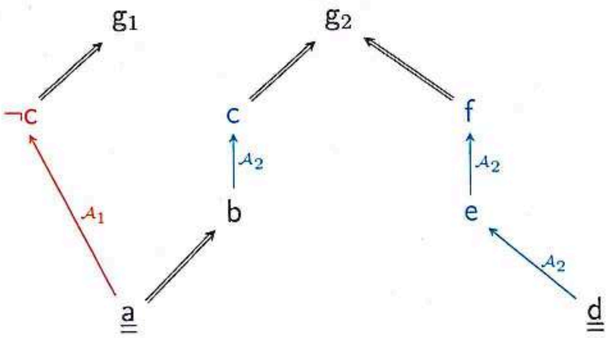
$$\Pi^{\mathrm{F}}_{7.6} := \{\ \mathsf{a},\ \mathsf{d}\ \},$$

$$\Pi^{\mathrm{G}}_{7.6} := \left\{\begin{array}{l} \mathsf{g}_1 \Leftarrow \neg\mathsf{c}, \\ \mathsf{g}_2 \Leftarrow \mathsf{c}\wedge\mathsf{f}, \\ \mathsf{b} \Leftarrow \mathsf{a} \end{array}\right\},$$

$$\Delta_{7.6} := \mathcal{A}_1 \cup \mathcal{A}_2.$$

$$\mathcal{A}_1 := \{\ \neg\mathsf{c} \leftarrow \mathsf{a}\ \}.$$

$$\mathcal{A}_2 := \left\{\begin{array}{l} \mathsf{c} \leftarrow \mathsf{b}, \\ \mathsf{e} \leftarrow \mathsf{d}, \\ \mathsf{f} \leftarrow \mathsf{e} \end{array}\right\}.$$



[may also be omitted in journal version]

Let us compare the specificity of the arguments $(\mathcal{A}_1, \mathsf{g}_1)$ and $(\mathcal{A}_2, \mathsf{g}_2)$.

We have $(\mathcal{A}_1, \mathsf{g}_1) \vartriangle_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{g}_2)$ for the following reasons: $\{\mathsf{a}\} \subseteq \mathfrak{T}_{\Pi_{7.6}} = \{\mathsf{a}, \mathsf{b}, \mathsf{d}\}$ is an activation set for for $(\mathcal{A}_1, \mathsf{g}_1)$, but not for $(\mathcal{A}_1, \mathsf{g}_1)$; $\{\mathsf{b}, \mathsf{d}\} \subseteq \mathfrak{T}_{\Pi_{7.6}}$ is an activation set for $(\mathcal{A}_2, \mathsf{g}_2)$, but not for $(\mathcal{A}_2, \mathsf{g}_2)$.

By Theorem 6.16 we also get $(\mathcal{A}_1, \mathsf{g}_1) \vartriangle_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{g}_2)$.

In this example the two intuitive reasons for specificity — super-conjunction (preference of the "more precise") and implication via a strict rule (preference of the "more concise") — are in an irresolvable conflict, which goes well together with the fact that neither $\lesssim_{\mathrm{CP1}}$ nor $\lesssim_{\mathrm{P3}}$ can compare the two arguments.

7.4 Global Effect matters more than Fine Structure

The following example nicely shows that any notion of specificity based only on single defeasible rules (without considering the context of the general strict rules as a whole) cannot be intuitively adequate. [only side-topic of this paper, thus may be omitted in article]

Example 7.7 (Example from Page 95 of [Stolzenburg &al., 2003])

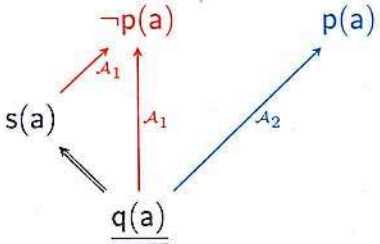
$$\Pi^{\mathrm{F}}_{7.7} := \{\ \mathsf{q}(\mathsf{a})\ \},$$

$$\Pi^{\mathrm{G}}_{7.7} := \{\ \mathsf{s}(x) \Leftarrow \mathsf{q}(x)\ \},$$

$$\Delta_{7.7} := \left\{\begin{array}{l} \mathsf{p}(x) \leftarrow \mathsf{q}(x), \\ \neg\mathsf{p}(x) \leftarrow \mathsf{q}(x)\wedge\mathsf{s}(x) \end{array}\right\},$$

$$\mathcal{A}_1 := \{\ \neg\mathsf{p}(\mathsf{a}) \leftarrow \mathsf{q}(\mathsf{a})\wedge\mathsf{s}(\mathsf{a})\ \},$$

$$\mathcal{A}_2 := \{\ \mathsf{p}(\mathsf{a}) \leftarrow \mathsf{q}(\mathsf{a})\ \}$$



Let us compare the specificity of the arguments $(\mathcal{A}_1, \neg\mathsf{p}(\mathsf{a}))$ and $(\mathcal{A}_2, \mathsf{p}(\mathsf{a}))$.

We have $(\mathcal{A}_1, \neg\mathsf{p}(\mathsf{a})) \approx_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{p}(\mathsf{a}))$, because of $\mathsf{p}(\mathsf{a}), \neg\mathsf{p}(\mathsf{a}) \notin \mathfrak{T}_{\Pi_{7.7}} = \{\mathsf{q}(\mathsf{a}), \mathsf{s}(\mathsf{a})\}$, and because, for $H \subseteq \mathfrak{T}_{\Pi_{7.7}\cup\Delta_{7.7}}$, $i \in \{1, 2\}$, $L_1 := \neg\mathsf{p}(\mathsf{a})$, and $L_2 := \mathsf{p}(\mathsf{a})$, we have the logical equivalence of $H = \{\mathsf{q}(\mathsf{a})\}$ on the one hand, and of H being a minimal simplified activation set for $(\mathcal{A}_i, L_i)$ but not for $(\emptyset, L_i)$, on the other hand.

By Theorem 6.16, we also get $(\mathcal{A}_1, \neg\mathsf{p}(\mathsf{a})) \approx_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{p}(\mathsf{a}))$.

This makes perfect sense because $\mathsf{q}(\mathsf{a}) \wedge \mathsf{s}(\mathsf{a})$ is not at all strictly "more precise" than $\mathsf{q}(\mathsf{a})$ in the context of $\Pi^{\mathrm{G}}_{7.7}$.

Note that nothing is changed here if $\mathsf{s}(x) \Leftarrow \mathsf{q}(x)$ is replaced by setting $\Pi^{\mathrm{G}}_{7.7} := \{\mathsf{s}(\mathsf{a})\}$. If $\mathsf{s}(x) \Leftarrow \mathsf{q}(x)$ is replaced by setting $\Pi^{\mathrm{G}}_{7.7} := \emptyset$ and $\Pi^{\mathrm{F}}_{7.7} := \{\mathsf{q}(\mathsf{a}), \mathsf{s}(\mathsf{a})\}$, however, then we get both $(\mathcal{A}_1, \neg\mathsf{p}(\mathsf{a})) <_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{p}(\mathsf{a}))$ and $(\mathcal{A}_1, \neg\mathsf{p}(\mathsf{a})) <_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{p}(\mathsf{a}))$.

This also speaks for our admission of literals (i.e. unconditional rules) to Π^{G}.[37]

Example 8.1 (Minimal argument with two minimal and-trees/activation sets)

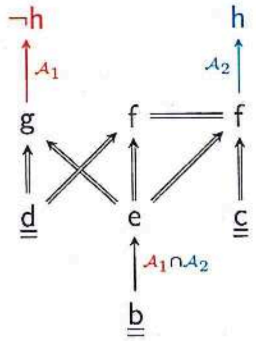
$\Pi^{\mathrm{F}}_{8.1} := \{\ \mathsf{b},\ \mathsf{c},\ \mathsf{d}\ \}$,

$$\Pi^{\mathrm{G}}_{8.1} := \left\{ \begin{array}{l} \mathsf{f} \Leftarrow \mathsf{c}\wedge\mathsf{e}, \\ \mathsf{f} \Leftarrow \mathsf{d}\wedge\mathsf{e}, \\ \mathsf{g} \Leftarrow \mathsf{d}\wedge\mathsf{e}, \end{array} \right\},$$

$\Delta_{8.1} := \mathcal{A}_1 \cup \mathcal{A}_2$.

$$\mathcal{A}_1 := \left\{ \begin{array}{l} \neg\mathsf{h} \leftarrow \mathsf{g}, \\ \mathsf{e} \leftarrow \mathsf{b} \end{array} \right\}.$$

$$\mathcal{A}_2 := \left\{ \begin{array}{l} \mathsf{h} \leftarrow \mathsf{f}, \\ \mathsf{e} \leftarrow \mathsf{b} \end{array} \right\}.$$



The argument $(\mathcal{A}_1, \neg\mathsf{h})$ has $\{\mathsf{b},\mathsf{d}\}$ as the only minimal activation set that is a subset of $\mathfrak{T}_{\Pi_{8.1}} = \Pi^{\mathrm{F}}_{8.1}$. $\{\mathsf{b},\mathsf{d}\}$ is also a minimal activation set for $(\mathcal{A}_2, \mathsf{h})$. On the other hand, $\{\mathsf{b},\mathsf{c}\}$ is an activation set for $(\mathcal{A}_2, \mathsf{h})$, but not for $(\mathcal{A}_1, \neg\mathsf{h})$. Thus, we get $(\mathcal{A}_1, \neg\mathsf{h}) <_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{h})$.

Because either d or c is in an and-tree of the argument $(\mathcal{A}_2, \mathsf{h})$ (but never both!), a comparison of two fixed and-trees does not suffice.

Moreover note that we have $(\mathcal{A}_1, \neg\mathsf{h}) \vartriangle_{\mathrm{P3}} (\mathcal{A}_2, \mathsf{h})$, because of the simplified activation sets $\{\mathsf{g}\}$ and $\{\mathsf{f}\}$, respectively.

Furthermore note that the only minimal activation set for the minimal argument $(\{\mathsf{e}\leftarrow\mathsf{b}\}, \mathsf{f})$ is $\{\mathsf{b}\}$, which, however, is not a simplified activation set for that argument.

The reason for the complication of an element-by-element comparison of and-trees is that we consider a very general setting of defeasible reasoning in this paper. Indeed, we admit

1. more than one condition literal in rules, i.e. conditions containing more than one literal, and
2. non-empty sets of *background knowledge*, i.e. general rules, not only facts. [no \emph]

Typically, only restricted cases are considered: Conditions have always to be singletons in [GELFOND & PRZYMUSINSKA, 1990], no background knowledge is allowed in [DUNG & SON, 1996], and both restrictions are present in [BENFERHAT & GARCIA, 1997].

8.2.3 Path Criteria?

Before we come to the computation of activations sets via goal-directed derivations in § 8.3, let us have a closer look here at the path criterion of [STOLZENBURG &AL., 2003, § 3.4].

Definition 8.2 (Path)
For a leaf node N in an and-tree T, we define the *path* in T through N as the empty set if N is the root, and otherwise as the set consisting of the literal labeling N, together with all literals labeling its ancestors except the root node. Let $\mathsf{Paths}(T)$ be the set of all paths in T through all leaf nodes N.

With this notion of paths, the quasi-ordering $\trianglelefteq$ on and-trees can be given as follows:

Definition 8.3 ([STOLZENBURG &AL., 2003, Definition 23])
$T_1 \trianglelefteq T_2$ if T_1 and T_2 are two and-trees, and for each $t_2 \in \mathsf{Paths}(T_2)$ there is a path $t_1 \in \mathsf{Paths}(T_1)$ such that $t_1 \subseteq t_2$.

8.3 Toward a More Efficiently Realizable Notion of POOLE-Style Specificity

Contrary to our ~~toy~~ small examples in the previous sections, examples of a practically relevant size require notions of specificity that can be decided efficiently.

As we are mainly interested in the more specific arguments, i.e. in the minimal elements of our specificity ordering, we may admit variations of our specificity ordering CP1 that offer better chances for an efficient implementation, but do not relevantly differ w.r.t. these minimal elements.

Therefore, in this section, we will introduce another correction (CP2) of POOLE's specificity relation, which offers some advantages for the computation of the respective activation sets; whereas our specificity ordering CP1 offers only the minor advantages over P1, P2, P3 we have already described in §§ 8.1 and 8.2.1.

More precisely, our plan for this section is to obtain another quasi-ordering $\lesssim_{\mathrm{CP2}}$ by slight modification of our quasi-ordering $\lesssim_{\mathrm{CP1}}$, such that the two do not differ in any of our previous examples, and such that $\lesssim_{\mathrm{CP2}}$ may mirror our intuition on specificity according to the analysis in § 4 even more closely in some aspects. Finally, we will try to develop a more efficient procedure for deciding the specificity quasi-ordering $\lesssim_{\mathrm{CP2}}$ than those known for any of $\lesssim_{\mathrm{P1}}$, $\lesssim_{\mathrm{P2}}$, $\lesssim_{\mathrm{P3}}$, $\lesssim_{\mathrm{CP1}}$.

The crucial step in such a procedure is the computation of activation sets. For a goal-directed, SLD-resolution-like computation of activation sets we cannot keep our restriction to arguments that are ground. For this reason, we now have to modify our notion of a derivation by disallowing the instantiation of variables in our definition of $\mathfrak{T}_{\Pi}$ and $\vdash$ (cf. Definition 2.2) as already hinted at in Remark 2.7 at the end of § 2.4. As a compensation, we then may add a hat over a set of rules Π, such that $\hat{\Pi}$ denotes the set of all instances of Π.

8.3.1 Immediate Activation Sets

As a first step — since the workaround via path criteria failed in § 8.2.3 — we now have to find a new notion of an *immediate* activation set such that there are less[39] and more easily computable immediate activation sets for a given argument than (non-immediate) activation sets according to Definition 6.1 of § 6.1. Our idea here is to avoid SLD-resolution steps that expand a goal clause by *inessential* applications of rules in the sense of the following definition, where we again apply the simple concept of an and-tree according to Definition 4.1 of § 4.4.1.

Definition 8.6 (Inessential Application of an Instance of a Rule)
The application of the instance $L{\Leftarrow}C$ of a rule in an and-tree is *inessential (in the and-tree)* if there is a node between the root (inclusively) and the application (including the node labeled with L) that is labeled with an element of $\mathfrak{T}_{\hat{\Pi}}$.

[39] There are indeed never more (cf. Corollary 8.8(4)), and typically much less immediate activation sets than activation sets.

Definition 8.7 ([Minimal/Weakly] *Immediate* Activation Set)
Let $\mathcal{A}$ be a set of instances of rules from Δ, and let L be a literal.
H is an *immediate activation set for* $(\mathcal{A}, L)$ if $H \subseteq \mathfrak{L}_{\hat{\Pi}}$ and there is a (possibly empty) set of literals $\mathfrak{L}$, such that both of the following two items hold:

1. For each $L' \in \mathfrak{L}$ there is an and-tree for the derivation of $H \cup \mathcal{A} \cup \hat{\Pi}^{\mathrm{G}} \vdash \{L'\}$ in which
   (a) the root is labeled with L' and generated by an element of $\mathcal{A}$, and
   (b) every literal L'' that labels a non-leaf node or the root satisfies $L'' \notin \mathfrak{L}_{\hat{\Pi}}$, and
   (c) every literal $L'' \notin \mathcal{A}$ that labels a leaf node satisfies $L'' \in \mathfrak{L}_{\hat{\Pi}}$,[41]

   such that the set of literals labeling the leaves of these trees is a subset of $H \cup \mathfrak{L}_{\hat{\Pi}^{\mathrm{G}}} \cup \mathcal{A}$.
2. There is an and-tree for the derivation of $\mathfrak{L} \cup \hat{\Pi} \vdash \{L\}$, such that each literal L'' labeling a node in a path from the root to a leaf labeled with an element from $\mathfrak{L}$ satisfies $L'' \notin \mathfrak{L}_{\hat{\Pi}}$.

H is a *minimal immediate activation set for* $(\mathcal{A}, L)$ if H is an immediate activation set for $(\mathcal{A}, L)$, but no proper subset of H is an immediate activation set for $(\mathcal{A}, L)$.

H is a *weakly immediate activation set for* $(\mathcal{A}, L)$ if $H \subseteq \mathfrak{L}_{\hat{\Pi}}$ and there is an immediate activation set H' with $H' \subseteq \mathfrak{L}_{H \cup \hat{\Pi}^{\mathrm{G}}}$ for $(\mathcal{A}, L)$.

Corollary 8.8 *Let $\mathcal{A}$ be a set of instances of rules from Δ, and let L be a literal.*

1. *If H is an [weakly] immediate activation set, then we have $H \subseteq \mathfrak{L}_{\hat{\Pi}}$.*
2. *If H is a minimal immediate activation set, then we have $H \subseteq \mathfrak{L}_{\hat{\Pi}} \backslash \mathfrak{L}_{\hat{\Pi}^{\mathrm{G}}}$.*
3. *Every immediate activation set for $(\mathcal{A}, L)$ is a weakly immediate activation set for $(\mathcal{A}, L)$.*
4. *Every [weakly] immediate activation set for $(\mathcal{A}, L)$ is an activation set*[42] *for $(\mathcal{A}, L)$.*
5. *Every minimal activation set for $(\mathcal{A}, L)$ that is an immediate activation set for $(\mathcal{A}, L)$ is a minimal immediate activation set for $(\mathcal{A}, L)$.*

[41] Here "literal $L'' \notin \mathcal{A}$" means that L'' is a literal that is not a literal in $\mathcal{A}$, i.e. no conclusion of an unconditional rule from $\mathcal{A}$. Note that, by (a), this excludes an overlap of (b) and (c): If the root is a leaf, then, by (a), it is labeled with a literal from $\mathcal{A}$.

[42] Instead of the otherwise required condition that $\mathcal{A}$ *is ground,* we assume here — and will do so in what follows without further mentioning — that the definition of an activation set in Definition 6.1 of § 6.1 refers (just as Definition 8.7 of immediate ones and just as we have changed arguments and derivations in this section) to sets also of *non-ground* instances of defeasible rules in the first element of arguments, but with non-instantiating derivations and theories.

Remark 8.13 (Relaxation to a *Weakly* Immediate Activation Set is Crucial)
Note that we cannot straightforwardly require H to be a (non-weakly) immediate activation set for $(\mathcal{A}_2, L_2)$ in Definition 8.10, because otherwise our new relation CP2 would already fail to pass Example 3.2 of § 3, in the sense that both arguments there would be incomparable.[44 45]

Theorem 8.14 $\lesssim_{\mathrm{CP2}}$ *is a quasi-ordering on arguments.*

and $(\mathcal{A}_2, L_2) \lesssim$

Proof of Theorem 8.14
$\lesssim_{\mathrm{CP2}}$ is a reflexive relation on arguments because of Corollary 8.11.

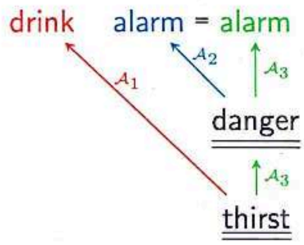
To show transitivity, let us assume $(\mathcal{A}_1, L_1) \lesssim_{\mathrm{CP2}} (\mathcal{A}_2, L_2) \lesssim_{\mathrm{CP2}} (\mathcal{A}_3, L_3)$.

According to Definition 8.10, because of $(\mathcal{A}_1, L_1) \lesssim_{\mathrm{CP2}} (\mathcal{A}_2, L_2)$, we have $L_1 \in \mathfrak{T}_{\hat\Pi}$ — and then immediately the desired $(\mathcal{A}_1, L_1) \lesssim_{\mathrm{CP2}} (\mathcal{A}_3, L_3)$ — or we have $L_2 \notin \mathfrak{T}_{\hat\Pi}$. The latter case excludes the first option in Definition 8.10 as a justification for $(\mathcal{A}_2, L_2) \lesssim_{\mathrm{CP2}} (\mathcal{A}_3, L_3)$. Thus, it now suffices to consider the case that $L_i \notin \mathfrak{T}_{\hat\Pi}$ for all $i \in \{1, 2, 3\}$.

Suppose that H is an immediate activation set for $(\mathcal{A}_1, L_1)$. It suffices to show that H is a weakly immediate activation set for $(\mathcal{A}_3, L_3)$, i.e. to find an immediate activation set $H'' \subseteq \mathfrak{T}_{H \cup \hat\Pi^{\mathrm{G}}}$ for $(\mathcal{A}_3, L_3)$. Because of our supposition, the first step of our original assumption, and the case considered, H is a weakly immediate activation set for $(\mathcal{A}_2, L_2)$, i.e. there is an immediate activation set $H' \subseteq \mathfrak{T}_{H \cup \hat\Pi^{\mathrm{G}}}$ for $(\mathcal{A}_2, L_2)$. Then, because of the second step of our original assumption and the case considered, there is an immediate activation set $H'' \subseteq \mathfrak{T}_{H' \cup \hat\Pi^{\mathrm{G}}}$ for $(\mathcal{A}_3, L_3)$. Because of the monotonicity of our logic and the closedness of our theories, we now have $H'' \subseteq \mathfrak{T}_{H' \cup \hat\Pi^{\mathrm{G}}} \subseteq \mathfrak{T}_{\mathfrak{T}_{H \cup \hat\Pi^{\mathrm{G}}} \cup \hat\Pi^{\mathrm{G}}} = \mathfrak{T}_{H \cup \hat\Pi^{\mathrm{G}}}$, i.e. $H'' \subseteq \mathfrak{T}_{H \cup \hat\Pi^{\mathrm{G}}}$, as was to be shown.

Q.e.d. (Theorem 8.14)

Example 8.15 ($\lesssim_{\mathrm{CP1}}$ vs. $\lesssim_{\mathrm{CP2}}$)



$\Pi^{\mathrm{F}}_{8.15} := \{\ \mathsf{thirst},\ \mathsf{danger}\ \}$, $\Pi^{\mathrm{G}}_{8.15} := \emptyset$, $\Delta_{8.15} := \mathcal{A}_1 \cup \mathcal{A}_3$.

$\mathcal{A}_1 := \{\ \mathsf{drink} \leftarrow \mathsf{thirst}\ \}$.

$\mathcal{A}_2 := \{\ \mathsf{alarm} \leftarrow \mathsf{danger}\ \}$.

$\mathcal{A}_3 := \mathcal{A}_2 \cup \{\ \mathsf{danger} \leftarrow \mathsf{thirst}\ \}$.

First note that — because of $\Pi^{\mathrm{G}}_{8.15} = \emptyset$ — the two notions of an immediate and a weakly immediate activation set coincide here.

We have $\mathfrak{T}_{\hat\Pi_{8.15}} = \Pi^{\mathrm{F}}_{8.15}$. Moreover, we have

$$(\mathcal{A}_2, \mathsf{alarm}) <_{\mathrm{CP1}} (\mathcal{A}_3, \mathsf{alarm}) \approx_{\mathrm{CP2}} (\mathcal{A}_2, \mathsf{alarm}):$$

There is only one minimal activation set for $(\mathcal{A}_2, \mathsf{alarm})$ that is a subset of $\mathfrak{T}_{\hat\Pi_{8.15}}$, namely $\{\mathsf{danger}\}$. It is also a minimal *immediate* activation set for $(\mathcal{A}_2, \mathsf{alarm})$; to see this, take $\mathfrak{L} := \{\mathsf{alarm}\}$ in Definition 8.7. There are only two minimal activation sets for $(\mathcal{A}_3, \mathsf{alarm})$ that are subsets of $\mathfrak{T}_{\hat\Pi_{8.15}}$, namely $\{\mathsf{danger}\}$ and $\{\mathsf{thirst}\}$, but only the first one is an immediate activation set for $(\mathcal{A}_3, \mathsf{alarm})$. Note that $(\mathcal{A}_2, \mathsf{alarm})$ is *strictly* more specific than $(\mathcal{A}_3, \mathsf{alarm})$ in the sense of $(\mathcal{A}_2, \mathsf{alarm}) \not\gtrsim_{\mathrm{CP1}} (\mathcal{A}_3, \mathsf{alarm})$ by the inessential[46] application of the rule $\mathsf{danger} \leftarrow \mathsf{thirst}$ of $\mathcal{A}_3$, which is not admitted for *immediate* activation sets.

[44] See the discussion at the end of Example 8.15.

[45] It might also be interesting to see that the slight modification (via "weakly"), which we need here, occurred already in our first intuitive sketch of a notion of specificity in § 4.3 — long before the development of the CP2 notion (cf. [WIRTH & STOLZENBURG, 2013, § 3.2]).

[46] This means inessential in the sense of Definition 8.6.

Example 8.17 *(continuing Example 7.7)*
Indeed, the only noticeable change at all occurs in Example 7.7, where $\{\mathsf{q(a)}\}$ is a minimal activation set for $(\mathcal{A}_1, \neg\mathsf{p(a)})$, but not an *immediate* activation set. Nevertheless, because $\{\mathsf{q(a)}\}$ is a *weakly* immediate activation set for $(\mathcal{A}_1, \neg\mathsf{p(a)})$, and because the only immediate activation set for $(\mathcal{A}_1, \neg\mathsf{p(a)})$ is $\{\mathsf{q(a)}, \mathsf{s(a)}\}$ (which is a weakly immediate activation set for $(\mathcal{A}_2, \mathsf{p(a)})$, though not $\{\mathsf{q(a)}\}$, the only (non-weakly) immediate one), we have $(\mathcal{A}_1, \neg\mathsf{p(a)}) \approx_{\mathrm{CP2}} (\mathcal{A}_2, \mathsf{p(a)})$, just as we have $(\mathcal{A}_1, \neg\mathsf{p(a)}) \approx_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{p(a)})$.

Example 8.18 *(continuing Example 8.1 of § 8.2.2)*
(Minimal argument with two minimal *immediate* activation sets)
It is obvious that a minimal argument can easily have two minimal activation sets that are incomparable w.r.t. $\subseteq$. For instance, already in Example 3.2 of § 3, the minimal argument $(\mathcal{A}_2, \mathsf{flies(edna)})$ has two minimal [simplified] activation sets, namely $\{\mathsf{bird(edna)}\}$ and $\{\mathsf{emu(edna)}\}$, from which, however, only $\{\mathsf{bird(edna)}\}$ is an *immediate* activation set. Indeed, minimal arguments can have more than one minimal *immediate* activation set only if conditions of *general* rules directly contribute to the leaves of the isolated defeasible part as described in § 4.4.1.[47] This is happens in Example 8.1 of § 8.2.2 for the minimal argument $(\mathcal{A}_2, \mathsf{h})$: The general rule $\mathsf{f}{\Leftarrow}\mathsf{c}{\wedge}\mathsf{e}$ contributes the leaf c to the isolated defeasible part with root h, the inner nodes f and e, and the set of leaves $\{\mathsf{b}, \mathsf{c}\}$, which is one minimal immediate activation set of $(\mathcal{A}_2, \mathsf{h})$. Moreover, the general rule $\mathsf{f}{\Leftarrow}\mathsf{d}{\wedge}\mathsf{e}$ contributes the leaf d to the isolated defeasible part with root h, the inner nodes f and e, and the set of leaves $\{\mathsf{b}, \mathsf{d}\}$, which is the other minimal immediate activation set of $(\mathcal{A}_2, \mathsf{h})$, and also the only one for $(\mathcal{A}_1, \neg\mathsf{h})$. Thus, we get both $(\mathcal{A}_1, \neg\mathsf{h}) <_{\mathrm{CP1}} (\mathcal{A}_2, \mathsf{h})$
and $(\mathcal{A}_1, \neg\mathsf{h}) <_{\mathrm{CP2}} (\mathcal{A}_2, \mathsf{h})$.

8.3.2 Special Cases with Simple Activation-Set Computation

A typical problem in practical application is to classify rules automatically as being facts, general rules, or defeasible rules. We briefly discuss the trivial forms of such a classification now.

The first trivial form of classification is to take all proper rules as defeasible rules. Note that the following lemma (motivated by Example 8.18 of § 8.3.1) reduces the task of computing activation sets to the simpler task of computing minimal arguments.

Lemma 8.19 *Assume that all rules in Π^{G} are just literals (i.e. have empty conditions). Let $(\mathcal{A}, L)$ be a minimal argument. Let $\mathfrak{L}$ be the set of all condition literals of all rules in $\mathcal{A}$. Then $(\mathcal{A}, L)$ has a unique minimal activation set H; and this H is actually a minimal immediate activation set for $(\mathcal{A}, L)$ and equal to $\mathfrak{L} \cap \hat{\Pi}^{\mathrm{F}} \setminus \hat{\Pi}^{\mathrm{G}}$.*

[47] Technically, it is possible to enforce a unique immediate activation set for each minimal argument by including the instances also of the *general* rules of the isolated defeasible part into the first element of the arguments. Intuitively, however, this is not reasonable because it leads to unintendedly incomparable arguments.

8.3.3 A Step Toward Operationalization of Immediate Activation Sets

Let us assume that the sets of our predicate and function symbols are enumerable and contain only symbols with finite arities. This assumption does not seem to restrict practical application.

It is straightforward to enumerate for a given input literal — say in a top-down SLD-resolution style — the and-trees of all possible derivations of instances of this input literal, and to interleave this enumeration of and-trees with the enumeration of all ground instances of each and-tree, and finally to enumerate for each ground instance of an and-tree all activation sets for all contained arguments and the ground instance of the input literal labeling the root. Indeed, this is possible because $\mathfrak{T}_{\hat{\Pi}}$ is enumerable (i.e. *semi-decidable*) by our above assumption.

To do the same for all *immediate* activation sets, we have to require the *co-semi-decidability* of $\mathfrak{T}_{\hat{\Pi}}$, because, in general, we cannot output an activation set supposed to be an immediate one before we have established that the literals labeling the ancestors of the nodes of its literals really do *not* occur in $\mathfrak{T}_{\hat{\Pi}}$.

So let us assume the decidability of $\mathfrak{T}_{\hat{\Pi}}$ for the remainder of this section.[48]

It is much harder, however, to enumerate all activation sets in an SLD-like derivation style *directly*, i.e. without storing the intermediate and-trees and their instances. Although *immediate* activation sets offer a crucial advantage for a direct enumeration in principle (because they admit to cut off inessential[49] derivations of literals), the imperative, tail-recursive procedure we will sketch in this section (cf. Figure 2) still needs further refinement. This procedure enumerates the immediate activation sets *directly*, unless it sometimes outputs the character string `"breach"`, which indicates that some immediate activation sets may be missing.

We present the procedure of Figure 2 here mainly because we want to concretize the tasks that still remain to be solved for obtaining a POOLE-style notion of specificity that admits a sufficiently efficient operationalization, and because our solution of these tasks in § 8.3.4 may not be the only way to solve them.

Let us assume that *picking* elements from sets satisfies some fairness restriction in the sense that every element will be picked eventually. Moreover, let us assume that we have a procedure to decide $\mathfrak{T}_{\hat{\Pi}}$. Furthermore, let us assume that L is a literal with $L \notin \mathfrak{T}_{\hat{\Pi}}$.

Under these assumptions, the SLD-like procedure immediate-activation-sets(L) of Figure 2 has the following two properties:

1. If it outputs $(H, (A, I))$ then $I \notin \mathfrak{T}_{\hat{\Pi}}$ is an instance of L, we have $A \neq \emptyset$, and $H \subseteq \mathfrak{T}_{\hat{\Pi}}$ is an immediate activation set for the argument (A, I).
2. If it never outputs `"breach"`, then, for each instance $L\varrho \notin \mathfrak{T}_{\hat{\Pi}}$ with an immediate activation set $H' \subseteq \mathfrak{T}_{\hat{\Pi}}$ for an argument $(\mathcal{A}, L\varrho)$, it outputs some $(H, (A, I))$ such that there is a substitution μ with $(H, (A, I))\mu = (H', (\mathcal{A}, L\varrho))$. As this is similar to what is called a "most general unifier", we may speak of all *maximally general*, immediate activation sets with arguments here.

[48] We will relax this restriction in § 8.3.4.

[49] This means inessential in the sense of Definition 8.6.

Remark 8.21 (Restriction to Ground Conclusions Prevents "breach")
In the special case that the conclusions of all rules of $\Pi^{\mathrm{G}} \cup \Delta$ with non-empty condition are ground, however, the call of the procedure immediate-activation-sets(L) is guaranteed not to output "breach", simply because then only ground literals can enter the set of the program variable O', which are immediately removed again by the line before the tail-recursive call.

Remark 8.22 (Restriction to Ground Input Literals Does *Not* Prevent "breach")
Note that a restriction to input literals that are ground does not really solve the crucial problem (of which the program variables O, O' have to take care in Figure 2) that a literal with free variables may be not in $\mathfrak{T}_{\hat{\Pi}}$, whereas some of its instances actually are in $\mathfrak{T}_{\hat{\Pi}}$. The main source of the free variables here are the *extra-variables*, i.e. the free variables that occur in the condition but not in the conclusion of a rule. Such rules with extra-variables and non-ground conclusions, however, are standard in positive-conditional specification, just as in logic programming. A single extra-variable in an arbitrary rule of $\Pi^{\mathrm{G}} \cup \Delta$ can force SLD-resolution to work on non-ground goals even for a ground input literal.

Some examples may be more appropriate here than a proof of the soundness of the procedure of Figure 2 (that enumerates a maximally general, immediate activation set for each immediate activation set unless it sometimes indicates "breach"), because we see the procedure only as a step in a further development toward a tractability that is sufficient in practice. Therefore, we will give some examples here on how the procedure

$$\mathsf{immediate\text{-}activation\text{-}sets}(L)$$

works for certain literals $L \notin \mathfrak{T}_{\hat{\Pi}}$, namely by

listing all calls of the auxiliary procedure immediate-activation-sets-helper.

[may be restrict to one example]

Example 8.23 *(continuing Example 7.7 of § 7.4)*
Let us start with Example 7.7 of § 7.4, which we recently reconsidered in Example 8.17. A call of immediate-activation-sets($\neg$p(a)) results in a call of immediate-activation-sets-helper with the argument quintuple ($\{(\neg\mathsf{p(a)}, 2)\}$, $\emptyset$, $\emptyset$, $\emptyset$, $\neg\mathsf{p(a)}$), where the only rule whose conclusion is unifiable with the only goal literal is a defeasible one, namely $\neg\mathsf{p}(x)\leftarrow\mathsf{q}(x)\wedge\mathsf{s}(x)$ from $\Delta_{7.7}$. We can take ξ and σ as the identity and $\{x\mapsto\mathsf{a}\}$, respectively. The program variable B' will be set to 1, and the tail-recursive call will have the argument tuple

$$(\ \emptyset, \ \emptyset, \ \{\mathsf{q(a)}, \mathsf{s(a)}\}, \ \{\neg\mathsf{p(a)}\leftarrow\mathsf{q(a)}\wedge\mathsf{s(a)}\}, \ \neg\mathsf{p(a)} \ ).$$

This call immediately terminates by outputting the immediate activation set $\{\mathsf{q(a)}, \mathsf{s(a)}\}$ for the argument ($\{\neg\mathsf{p(a)}\leftarrow\mathsf{q(a)}\wedge\mathsf{s(a)}\}$, $\neg\mathsf{p(a)}$). As all calls are terminated now and there was no output of "breach", this means that we have enumerated all immediate activation sets for the input literal.

Example 8.24 *(continuing Example 3.3 of § 3)*
Let us now come to Example 3.3 of § 3. A call of immediate-activation-sets(flies(y)) results in a call of immediate-activation-sets-helper with the argument quintuple

$$(\ \{(\mathsf{flies}(y), 2)\}, \ \emptyset, \ \emptyset, \ \emptyset, \ \mathsf{flies}(y) \ ),$$

where the only rule whose conclusion is unifiable with the only goal literal is a defeasible

[This sounds too pessimistic to me.

8.3.4 A Specificity Relation Based on Given And-Trees

We see no straightforward procedure to decide $\lesssim_{\mathrm{CP2}}$. Even worse, we see neither a procedure to semi-decide it, nor a procedure to co-semi-decide it. A positive answer can be given if the procedure of Figure 2 terminates for the first argument of $\lesssim_{\mathrm{CP2}}$ without outputting `"breach"`. A negative answer can be given if, for an immediate activation set enumerated for the first argument, the derivation for testing the property of being a weakly immediate activation set for the second argument terminates with failure. In general, even if we assume $\mathfrak{T}_{\hat{\Pi}}$ to be decidable, none of these terminations is guaranteed.[50]

In such a situation it is clearly appropriate to relax our requirement of a *model-theoretical* specificity relation a bit. So we replace the fancied decision procedure for $\mathfrak{T}_{\hat{\Pi}}$ with the test whether the literal has a derivation from those instances of Π which can be found in some and-tree occurring in *a finite set of and-trees fixed in advance.* For the solution we are aiming at, it is crucial that this given finite set of and-trees cannot be further extended during related specificity considerations. A good candidate may be the set of those and-trees that our derivation procedure has been able to construct within a certain time limit.

Then we can replace each of the three elements of our specification $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$ with the sets of those instances of their elements that are actually applied in our finite set of and-trees, resulting in the new specification $(\Pi^{\mathrm{F}}_{\mathrm{G}}, \Pi^{\mathrm{G}}_{\mathrm{G}}, \Delta_{\mathrm{G}})$. The further considerations must use these three finite sets without any further instantiation. This means that their rules are to be considered to be ground and this is what the lower index "G" stands for. [difficult to distinguish from upper index]

We again abbreviate $\Pi_{\mathrm{G}} := \Pi^{\mathrm{F}}_{\mathrm{G}} \cup \Pi^{\mathrm{G}}_{\mathrm{G}}$, and also replace the typically undecidable set $\mathfrak{T}_{\hat{\Pi}}$ with finite set $\mathfrak{T}_{\Pi_{\mathrm{G}}}$.

Note that hardly anything has changed for our set of defeasible rules, because arguments work anyway with instances that are ground, or are at least treated as if they were ground (cf. Remark 2.7 in § 2.4), and we can hardly consider an argument that is not contained in some and-tree we have constructed in advance.

There is a major change, however, for the set Π of strict rules. The situation here is similar to an expansion w.r.t. a *champ fini* in HERBRAND's Fundamental Theorem,[51] and we have reason to hope that the effect of this change can be neglected in practice, provided that a sufficient number of the proper instances is considered. Note that, for first-order logic, the depth limit n for terms required for HERBRAND's Property C to establish a sentential tautology (i.e. the natural number n for the *champ fini* of order n) is not computable in the sense of a *total* recursive function. Even if we knew the smallest such n, however, the number of terms of depth smaller than n would still be too high for practical feasibility in general. This means that it is crucial to chose the instances of our rules in a clever way, say from the successful proofs delivered by a theorem-proving system within a sufficient time limit.

[50] Both of these terminations can be guaranteed, however, under most restrictive conditions, such as the one that the conclusions of every defeasible rule from $\Pi^{\mathrm{G}} \cup \Delta$ with a non-empty condition are ground (cf. Remark 8.21). ?

[51] Cf. [HERBRAND, 1930], [WIRTH &AL., 2009; 2014], [WIRTH, 2012; 2014]. inline

inline

With the modifications described above, let us now come back to our procedure of Figure 2. As noted before, there cannot be any output of `"breach"` anymore, because our new sets of general strict and defeasible rules, i.e. the sets $\Pi^{\mathrm{G}}_{\mathrm{G}}$ and Δ_{G}, are now ground by definition. After the resulting simplifications, the procedure immediate-activation-sets-helper now may be replaced with the procedure ground-immediate-activation-sets-helper sketched in Figure 3.

To ensure termination of ground-immediate-activation-sets-helper we additionally have to store the path of the and-tree and exit without further output if we encounter a literal for a second time on the same path.

Remark 8.28 (Considerations on Complexity)
Regarding time complexity of the procedure of Figure 3 extended with the storage of the and-tree for ensuring termination mentioned above, only the following preliminary remarks apply in this early state of development.

From practical experience, complexity is not relevant yet: Our straightforward PROLOG (cf. e.g. [CLOCKSIN & MELLISH, 2003])) implementation of the algorithm in Fig. 3, available by the author (which prefers simplicity of coding over efficiency) computes, compares, and sorts — without any noticeable delay in the answer — all minimal immediate activation sets for all minimal arguments for all literals of $\mathfrak{T}_{\Pi_{\mathrm{G}}\cup\Delta_{\mathrm{G}}}\backslash\mathfrak{T}_{\Pi_{\mathrm{G}}}$, for a specification $(\Pi^{\mathrm{F}}_{\mathrm{G}}, \Pi^{\mathrm{G}}_{\mathrm{G}}, \Delta_{\mathrm{G}})$ of all instances required for a superset of all examples in this paper.

Regarding the theoretical worst case, which will hardly ever occur in practice, the following first estimate may be not completely irrelevant. Let n be the number of different literals in all conclusions of all rules of $\Pi_{\mathrm{G}} \cup \Delta_{\mathrm{G}}$. With our mentioned mechanism for ensuring termination, it is obvious that n limits the maximal depth of the SLD-like search tree. Let m be the maximal number of all condition literals of all rules with an identical conclusion. It is obvious that m limits the maximal number of children of any node in the SLD-like search tree, cumulated over the whole run. This means that the maximal size of the cumulated search tree is $m^{n-1}-1$, i.e. $\mathrm{O}(m^n)$. Luckily, this LANDAU-O limits also the size of the theory $\mathfrak{T}_{\Pi_{\mathrm{G}}}$ (which we pre-compute in our PROLOG implementation) and all other efforts at each node. Therefore, the whole algorithm is $\mathrm{O}(m^n)$.

Now we can compute the finite set of all minimal[53] immediate activation sets of all minimal arguments for a given input literal w.r.t. our ground specification $(\Pi^{\mathrm{F}}_{\mathrm{G}}, \Pi^{\mathrm{G}}_{\mathrm{G}}, \Delta_{\mathrm{G}})$. All what is left for deciding $\precsim_{\mathrm{CP2}}$ is to check whether each of the computed immediate activation sets whose defeasible rules are part of the first argument is a weakly immediate activation set for the second argument. This is straightforward, although it is not clear yet which implementation will be optimal.

We should not forget, however, that the specification $(\Pi^{\mathrm{F}}_{\mathrm{G}}, \Pi^{\mathrm{G}}_{\mathrm{G}}, \Delta_{\mathrm{G}})$ is only a reasonably constructed sub-specification of our original specification $(\Pi^{\mathrm{F}}, \Pi^{\mathrm{G}}, \Delta)$, which actually stands for $(\hat{\Pi}^{\mathrm{F}}, \hat{\Pi}^{\mathrm{G}}, \hat{\Delta})$. Practical tests have to show whether such an omission of infinitely many instances can be viable without deteriorating our specificity ordering. Theoretically, such a viability can only be guaranteed for the special case that the number of instances of the rules of the specification is finite (up to renaming of variables).

[53]First we filter the immediate activations sets from Figure 3 by removing all elements from $\mathfrak{T}_{\hat{\Pi}^{\mathrm{G}}}$, and then, for each minimal argument, we compare them w.r.t. $\subseteq$ and remove all supersets.

Moreover, we have to distinguish between orderings for comparing conflicting arguments w.r.t. specificity and orderings for comparing arguments w.r.t. a form of subsumption, such as the quasi-ordering of being "more conservative" found in [BESNARD & HUNTER, 2001, Definition 3.3, p. 206], [BESNARD &AL., 2013, Definition 6, p. 50]). There, roughly speaking, an argument $(\mathcal{A}_1, L_1)$ is *more conservative* than an argument $(\mathcal{A}_2, L_2)$ if $\mathcal{A}_1 \subseteq \mathcal{A}_2$ and $\{L_2\} \vdash \{L_1\}$. So if our opponent accepts the argument $(\mathcal{A}_2, L_2)$, then he also has to accept our more conservative argument $(\mathcal{A}_1, L_1)$, because we need less presuppositions and our result follows from our opponent's result. In many practical situations, however, the less conservative argument will be preferred. For instance, when we ask a question-answering system (such as LogAnswer [FURBACH &AL., 2010]) for the mother of PIERRE FERMAT, then — as an answer — we prefer[56] the less conservative argument $(\mathcal{A}, \mathsf{Mother}(\text{CLAIRE DE LONG}, \text{PIERRE FERMAT}))$ to

$$(\mathcal{A}, \exists x.\ \mathsf{Mother}(x, \text{PIERRE FERMAT})).$$

Moreover, the arguments $(\mathcal{A}, \mathsf{Mother}(\text{FRANÇOISE CAZENEUVE}, \text{PIERRE FERMAT}))$ and

$$(\mathcal{A}, \mathsf{Mother}(\text{CLAIRE DE LONG}, \text{PIERRE FERMAT})),$$

are incomparable in the "more conservative"-quasi-ordering. Even worse, for a non-trivial derivability relation, i.e. in a non-contradictory theory, this quasi-ordering cannot compare arguments with conflicting results L_1, L_2 by definition, and none of the arguments of our examples can be compared by this quasi-ordering.

~~It has become clear in several discussions that the main~~ One obstacle for an acceptance of one of our relations $\precsim_{\mathrm{CP1}}$ or $\precsim_{\mathrm{CP2}}$ as a replacement for $\precsim_{\mathrm{P3}}$ ~~in the scientific community~~ ~~is~~ may be the change this brings to Example 3.3 of § 3: Some scientists working in the field for a longer time have become used to the preference given by $\precsim_{\mathrm{P3}}$ in this most popular ~~toy~~ example — so much that they now consider that preference a must. Note that the situation in Example 3.3 is actually most unstable under the two following aspects: [no fluff needed

1. The preference chosen by $\precsim_{\mathrm{P3}}$ in Example 3.3 has justifications that are intuitive and valid, but are in general uncorrelated to specificity, such as the preference of conservativeness or even the non-model-theoretic preference of defeasible derivations of shorter length. In particular in this example, such intuitive justifications easily contaminate the readers' intuition w.r.t. specificity. Moreover, as the arguments in Example 3.3 are not incomparable, but just equivalent according to $\precsim_{\mathrm{CP1}}$, we can easily combine $\precsim_{\mathrm{CP1}}$ lexicographically with another ordering, say "minimum in the ordering of the natural numbers, for all and-trees, of the maximal length of defeasible paths", and so recover the traditional preference of Example 3.3.

2. The situation of the example is chaotic in the sense that different preferences result from minor changes that may escape the readers' disambiguation. For instance, if we add the general rule of the example that precedes Example 3.3 (i.e. of Example 3.2), then the preference chosen by $\precsim_{\mathrm{P3}}$ is chosen by $\precsim_{\mathrm{CP1}}$ and $\precsim_{\mathrm{CP2}}$ as well. Moreover, if we alternatively add $\mathsf{bird}(\mathsf{edna})$ as a fact, then we can embed the example injectively into Example 8.15 of § 8.3.1, and then the preference chosen by $\precsim_{\mathrm{P3}}$ is again chosen by $\precsim_{\mathrm{CP1}}$, whereas the arguments become incomparable w.r.t. $\precsim_{\mathrm{CP2}}$.

Already the examples in § 7 show, however, that $\precsim_{\mathrm{P3}}$ almost always fails to prefer any argument in slightly bigger examples, not to speak of big ones. Indeed, $\precsim_{\mathrm{P3}}$ can be considered a reasonable choice only if we restrict our considerations to ~~tiny~~ small examples. Moreover,

[BESNARD &AL., 2013] Philippe Besnard, Éric Grégoire, and Badran Raddaoui. A conditional logic-based argumentation framework. 2013. In [LIU &AL., 2013, pp. 44–56]. http://dx.doi.org/10.1007/978-3-642-40381-1_4.

[CHESÑEVAR &AL., 2003] Carlos I. Chesñevar, Jürgen Dix, Frieder Stolzenburg, and Guillermo R. Simari. Relating defeasible and normal logic programming through transformation properties. *Theoretical Computer Sci.*, 290:499–529, 2003. Received Jan. 8, 2001; rev. Nov. 9, 2001. http://dx.doi.org/10.1016/S0304-3975(02)00033-6.

[CLOCKSIN & MELLISH, 2003] William F. Clocksin and Christopher S. Mellish. *Programming in* PROLOG. Springer, 2003. 5th edn. (1st edn. 1981).

[DUNG & SON, 1996] Phan Minh Dung and Tran Cao Son. An argumentation-theoretic approach to reasoning with specificity. 1996. In [AIELLO &AL., 1996, pp. 506–517]. ?

[DUNG, 1995] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and *n*-person games. *Artificial Intelligence*, 77:321–358, 1995. http://dx.doi.org/10.1016/0004-3702(94)00041-X.

[FURBACH &AL., 2010] Ulrich Furbach, Ingo Glöckner, and Björn Pelzer. An application of automated reasoning in natural-language question answering. *AI Comm.*, 23:241–265, 2010.

[GABBAY & WOODS, 2004ff.] Dov Gabbay and John Woods, editors. *Handbook of the History of Logic.* North-Holland (Elsevier), 2004ff..

[GABBAY &AL., 1997] Dov Gabbay, Rudolf Kruse, Andreas Nonnengart, and Hans-Jürgen Ohlbach, editors. *Proc. 1st Int. Joint Conf. on Qualitative and Quantitative Practical Reasoning, 1997, June 9–12, Bad Honnef (Germany)*, number 1244 in Lecture Notes in Computer Science. Springer, 1997.

[GABBAY, 2002] Dov Gabbay, editor. *Handbook of Philosophical Logic.* Kluwer (Springer Science+Business Media), 2002. 2nd edn.

[GARCÍA & SIMARI, 2004] Alejandro J. García and Guillermo R. Simari. Defeasible logic programming: An argumentative approach. *Theory and Practice of Logic Programming, Cambridge Univ. Press*, 4:95–138, 2004.

[GELFOND & PRZYMUSINSKA, 1990] Michael Gelfond and Halina Przymusinska. Formalization of inheritance reasoning in autoepistemic logic. *Fundamenta Informaticae*, XIII:403–443, 1990.

[GILLMAN, 1987] Leonard Gillman. *Writing Mathematics Well.* The Mathematical Association of America, 1987.

[GOLDFARB, 1970] Warren Goldfarb. Review of [HERBRAND, 1968]. *The Philosophical Review*, 79:576–578, 1970.

[HEIJENOORT, 1971] Jean van Heijenoort. *From* FREGE *to* GŒDEL*: A Source Book in Mathematical Logic, 1879–1931.* Harvard Univ. Press, 1971. 2nd rev. edn. (1st edn. 1967).

Irrthum und Schein. Johann Wendler, Leipzig, 1764. Vol. I (Dianoiologie oder die Lehre von den Gesetzen des Denkens, Alethiologie oder Lehre von der Wahrheit) (`http://books.google.de/books/about/Neues_Organon_oder_Gedanken_Uber_die_Erf.html?id=ViS3XCuJEw8C`) & Vol. II (Semiotik oder Lehre von der Bezeichnung der Gedanken und Dinge, Phänomenologie oder Lehre von dem Schein) (`http://books.google.de/books/about/Neues_Organon_oder_Gedanken_%C3%BCber_die_Er.html?id=X8UAAAAAcAAj`). Facsimile reprint by Georg Olms Verlag, Hildesheim (Germany), 1965, with a German introduction by HANS WERNER ARNDT.

[LIU &AL., 2013] Weiru Liu, V. S. Subrahmanian, and Jef Wijsen, editors. *Proc. 7th Int. Conf. on Scalable Uncertainty Management (SUM 2013), Washington (DC), Sept. 16–18, 2013*, number 8078 in Lecture Notes in Computer Science. Springer, 2013.

[MODGIL & PRAKKEN, 2014] Sanjay Modgil and Henry Prakken. The $ASPIC^+$ framework for structured argumentation: a tutorial. *Argument & Computation*, 5:31–62, 2014. `http://dx.doi.org/10.1080/19462166.2013.869766`.

[POOLE, 1985] David L. Poole. On the comparison of theories: Preferring the most specific explanation. 1985. In [JOSHI, 1985, pp. 144–147].

[PRAKKEN & VREESWIJK, 2002] Henry Prakken and Gerhard Vreeswijk. Logics for defeasible argumentation. 2002. In [GABBAY, 2002, pp. 218–319].

[RAEDT &AL., 2012] Luc De Raedt, Christian Bessière, Didier Dubois, Patrick Doherty, Paolo Frasconi, Fredrik Heintz, and Peter J. F. Lucas, editors. *Proc. 20th European Conf. on Artificial Intelligence (ECAI), Aug. 27–31, 2012, Montepellier, France*, number 242 in Frontiers in Artificial Intelligence and Applications. IOS Press, 2012. `http://ebooks.iospress.nl/volume/ecai-2012`.

[ROSENFELD, 1974] Jack L. Rosenfeld, editor. *Proc. of the Congress of the Int. Federation for Information Processing (IFIP), Stockholm (Sweden), Aug. 5–10, 1974*. North-Holland (Elsevier), 1974.

[SIMARI & LOUI, 1992] Guillermo R. Simari and Ronald P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence*, 53:125–157, 1992. Received Feb. 1990, rev. April 1991.

[STOLZENBURG &AL., 2003] Frieder Stolzenburg, Alejandro J. García, Carlos I. Chesñevar, and Guillermo R. Simari. Computing generalized specificity. *J. Applied Non-Classical Logics*, 13:87–113, 2003. `http://www.tandfonline.com/doi/abs/10.3166/jancl.13.87-113`.

[WIRTH & GRAMLICH, 1994] Claus-Peter Wirth and Bernhard Gramlich. A constructor-based approach to positive/negative-conditional equational specifications. *J. Symbolic Computation*, 17:51–90, 1994. `http://dx.doi.org/10.1006/jsco.1994.1004`, `http://wirth.bplaced.net/p/jsc94`.

[WIRTH & STOLZENBURG, 2013] Claus-Peter Wirth and Frieder Stolzenburg. DAVID POOLE's *Specificity Revised*. SEKI-Report SR–2013–01 (ISSN 1437–4447). SEKI Publications, DFKI Bremen GmbH, ~~Safe and Secure Cognitive Systems, Cartesium, Enrique~~