To appear in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, Montréal, Québec, Aug 20 - 25, 1995.

Technical Report
R-224
January, 1995

Specificity and Inheritance in Default Reasoning

Sek-Wah Tan < *tan@cs.ucla.edu* >
Judea Pearl < *judea@cs.ucla.edu* >
Computer Science Department,
University of California, Los Angeles, CA 90024

Abstract

When specificity considerations are incorporated in default reasoning systems, it is hard to ensure that exceptional subclasses inherit all legitimate features of their parent classes. To reconcile these two requirements, specificity and inheritance, this paper proposes the addition of a new rule, called coherence rule, to the desiderata for default inference. The coherence rule captures the intuition that formulae which are more compatible with the defaults in the database are more believable. We offer a formal definition of this extended desiderata and analyze the behavior of its associated closure relation which we call *coherence closure*. We provide a concrete embodiment of a system satisfying the extended desiderata, by taking the coherence closure of system Z . A procedure for computing the (unique) most compact belief ranking in the coherence closure of system Z is also described.

1 Introduction

It has been proposed [11, 9] that default reasoning systems be analyzed in terms of their (default) consequence relations. A number of *inference rules* (or axioms) have generally been accepted [13, 11] as a reasonable set of desiderata for a commonsense consequence relation. Despite the general acceptance of these desiderata, they fail to reconcile two accepted lines of reasoning, widely known as "inheritance" and "specificity". These can be illustrated by the classical Tweety example as follows: Consider the database containing four defaults, "penguins are birds", "penguins do not fly", "birds fly" and "birds have wings". "Specificity" tells us that if Tweety is a penguin, then Tweety does not fly because *penguin* is a more specific classification of Tweety than *bird*. "Inheritance", on the other hand, does equip Tweety with wings, by virtue of being a bird, albeit an exceptional bird with respect to flying ability.

The inheritance and specificity lines of reasoning depend on the interactions among the defaults in the database. An inspection of the rules proposed in past desiderata reveals that, invariably, each rule refers to the defaults database as one unit; no reference is made to specific subsets of defaults, the interaction among which produces the tension between inheritance and specificity. In this paper we propose a new rule, called *coherence*, that resolves this tension. Intuitively, the coherence rule prefers formulae that are more compatible with the defaults database. We will formalize the requirements of inheritance and specificity, and show that any consequence relation that satisfies the coherence rule (and the standard desiderata) will honor both requirements. In the next section, we review the accepted desiderata (including rational monotony) [13] before introducing the coherence rule. We will analyze the behavior of the closure of the extended desiderata which we call *coherence closure*.

In Section 3 we refine the semantics of system Z [12] to satisfy the coherence rule. First we review the semantics of system Z and the definition of belief rankings. *Coherence constraints* are then further imposed on admissible rankings to make them satisfy the coherence rule. We show that the resulting system is sound with respect to the extended desiderata. We also present a procedure for computing the *most compact* admissible belief ranking in the coherence closure. In the last section we compare related work.

2 An Extended Desiderata

2.1 The Standard Desiderata

Rules 1-5 below represent the standard desiderata which is supported by two different interpretations of defaults, probabilistic semantics [1, 13] and model preference semantics [14]. The desiderata (excluding rational monotony) offers a complete characterization [1, 13] of ϵ-entailment, a consequence relation induced by interpreting each default sentence as a statement of conditional probability assertion, infinitesimally removed from certainty.[1] By contrast, in model preference semantics a

[1] In ϵ-semantics, a default sentence is interpreted as a constraint on the infinitesimal conditional probabilities. The default conclusions are then the formulae that are forced to have extremely high probabilities by the constraints. Rational monotony is satisfied by restricting our attention to distributions that are parameterized by ϵ and are analytic in ϵ. Alternatively, the interpretation of defaults as statements in nonstandard probability theory [12, 7] also gives us rational

default $\varphi \rightarrow \psi$ is interpreted as "ψ holds in all the most preferred worlds compatible with φ". It has been shown [10] that the ability to represent preference among worlds by some numerical rank is a necessary and sufficient condition for the satisfaction of the desiderata. This confluence of two diverse interpretations offers a strong argument for the acceptance of the rules as a desiderata for default reasoning.

Given a fixed set of defaults Δ, the desiderata for its default consequence relation $\mathrel{|\!\!\sim}$ are:

Rule 1 (Logic) *If $\varphi \supset \psi$, then $\varphi \mathrel{|\!\!\sim} \psi$.*

Rule 2 (Cumulative) *If $\varphi \mathrel{|\!\!\sim} \gamma$, then $\varphi \mathrel{|\!\!\sim} \psi$ if and only if $\varphi \wedge \gamma \mathrel{|\!\!\sim} \psi$.*

Rule 3 (Cases) *If $\varphi \mathrel{|\!\!\sim} \psi$ and $\gamma \mathrel{|\!\!\sim} \psi$, then $\varphi \vee \gamma \mathrel{|\!\!\sim} \psi$.*

Rule 4 (Direct Inference) *If $\varphi \rightarrow \psi \in \Delta$, then $\varphi \mathrel{|\!\!\sim} \psi$.*

Rule 5 (Rational Monotony) *If $\varphi \mathrel{|\!\!\sim} \psi$ and it is not the case that $\varphi \mathrel{|\!\!\sim} \neg\gamma$, then $\varphi \wedge \gamma \mathrel{|\!\!\sim} \psi$.*

The logic rule says that logical conclusions are also default conclusions. The cumulative rule tells us that default conclusions are preserved when default conclusions are added to or removed from the set of facts. The cases rule says that the default conclusions of two facts also follows from their disjunction. The direct inference rule allows us to conclude the consequent of a default regardless of the contents of the database when its antecedent is all that has been learned. Finally, rational monotony captures the intuition that new observations (γ) can be assumed to be "irrelevant" (does not affect the default conclusions) unless they are implausible to begin with.

2.2 The Coherence Rule

The rationale for the coherence rule is that wffs (well formed formulae) that are more compatible with the defaults in Δ should be more believable and, in this proposal, incompatibility is measured by the set of defaults that is *falsified* by the wff. A wff ϕ *falsifies* a default $\varphi \rightarrow \psi$ if $\phi \supset \varphi \wedge \neg\psi$.

Definition 1 *Let $\Delta' \subset \Delta$ be a subset of defaults. We define $\Phi(\Delta')$ to be the wff*

$$\Phi(\Delta') = \bigwedge_{\varphi\rightarrow\psi\in\Delta'} \varphi \wedge \neg\psi \bigwedge_{\varphi\rightarrow\psi\in\Delta\setminus\Delta'} \varphi \supset \psi.$$

Note that $\Phi(\Delta')$ falsifies the defaults in Δ' and no others. Therefore it is reasonable to require that $\Phi(\Delta')$ be judged more coherent than $\Phi(\Delta'')$ whenever $\Delta' \subset \Delta''$. Thus, if we know that either $\Phi(\Delta')$ or $\Phi(\Delta'')$ is true, we should believe the former, written

$$\Phi(\Delta') \vee \Phi(\Delta'') \mathrel{|\!\!\sim}_{\Delta} \Phi(\Delta').$$

Rule 6 further extends this requirement to hold in the context of every observation ϕ:

Rule 6 (Coherence) *Let $\Delta' \subset \Delta'' \subseteq \Delta$ and ϕ be a wff. If $\phi \wedge \Phi(\Delta')$ is satisfiable then*

$$\phi \wedge (\Phi(\Delta') \vee \Phi(\Delta'')) \mathrel{|\!\!\sim}_{\Delta} \phi \wedge \Phi(\Delta').$$

monotony.

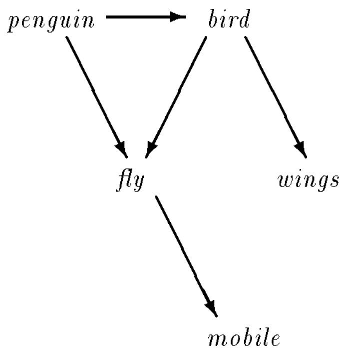

Figure 1: Penguin database, Δ.

We will call the set of rules, 1 to 6, the **extended desiderata**.

The coherence rule turns out to be equivalent to a simple constraint on preferences among worlds. For notational convenience, we will use the same symbol ω to represent the world (the truth assignment on atomic propositions) and the conjunctive clause (a conjunction of literals, a wff) whose only model is ω. We say a world ω *falsifies* a default $\varphi \rightarrow \psi$ if $\omega \models \varphi \wedge \neg\psi$. $\Delta[\omega]$ represents the set of defaults in Δ falsified by the world ω.

Theorem 1 (Coherence Constraints) *Let $\mathrel{|\!\!\sim}_{\Delta}$ be a default consequence relation that satisfies the desiderata. $\mathrel{|\!\!\sim}_{\Delta}$ satisfies the coherence rule if and only if for all worlds ω and ν, $\Delta[\omega] \subset \Delta[\nu]$ implies $\omega \vee \nu \mathrel{|\!\!\sim}_{\Delta} \omega$.*

This theorem follows from the fact that for every pair $(\omega, \Delta' \subseteq \Delta)$

$$\omega \models \Phi(\Delta') \text{ if and only if } \Delta' = \Delta[\omega].$$

It is easy to show that the coherence rule ensures the inheritance of *wings* by penguins in the example of Figure 1. This example also shows that the coherence rule is not derivable from the standard desiderata. However, instead of focusing on a single example, we will propose a general formulation of the inheritance requirement and show that it is satisfied by any default system that satisfies the extended desiderata.

Theorem 2 (Inheritance) *Let Δ be a database of defaults of the form $l \rightarrow l'$ where l and l' are literals and let $\mathrel{|\!\!\sim}_{\Delta}$ satisfy the extended desiderata. If $l_0 \rightarrow l_1 \rightarrow \ldots l_n$ is a chain of defaults in Δ and $\neg l_i$ is not the final consequent of another chain of defaults in Δ beginning with l_0 for all $i = 1, \ldots, n$, then*

1. *$\bigwedge_0^{n-1} l_i \mathrel{|\!\!\sim}_{\Delta} l_n$ and*
2. *$l_0 \mathrel{|\!\!\sim}_{\Delta} l_n$.*

The theorem says that if there is a chain of increasingly general superclasses (or properties) $l_0, \ldots, l_n$, then a member of l_0 will inherit the properties of the superclasses until the chain of inheritance is "severed" by another contradictory chain. In the database shown in Figure 1, the chain of rules

$$penguin \rightarrow bird \rightarrow wings$$

allows us to conclude that the *penguin* Tweety has *wings*. However, we cannot conclude that Tweety is *mobile* as the chain

$$\textit{penguin} \rightarrow \textit{bird} \rightarrow \textit{fly} \rightarrow \textit{mobile}$$

is "severed" by the contradictory chain

$$\textit{penguin} \rightarrow \neg \textit{fly}.$$

This behavior is in general agreement with our intuitions.

2.3 The Specificity Rule

Specificity arguments are usually defended by examples such as the one in Figure 1. We now cast these arguments formally, as a general rule and show that any default consequence relation that satisfies the extended desiderata will satisfy the specificity rule as well.

Definition 2 (Specificity) *A wff φ is* **more specific** *than φ' in Δ, written $\varphi >_\Delta \varphi'$, if $\varphi \mathrel{|\!\!\!\sim_\Delta} \varphi'$ and $\varphi' \not\mathrel{|\!\!\!\sim_\Delta} \varphi$.*

Note that specificity is not limited to strict subclass relationships; it is defined in terms of the consequence relation itself which, being based on defaults, allows for exceptions and abnormalities. We say that a default $\varphi \rightarrow \psi$ is **more specific** than another $\varphi' \rightarrow \psi'$, written $\varphi \rightarrow \psi >_\Delta \varphi' \rightarrow \psi'$ if its antecedent φ is more specific than φ'.

Rule 7 (Specificity) *If $\varphi \rightarrow \psi >_\Delta \varphi' \rightarrow \psi'$, then $\varphi \wedge \varphi' \mathrel{|\!\!\!\sim_\Delta} \psi$.*

The specificity rule captures the intuition that, when two (possibly conflicting) defaults are applicable, the more specific default prevails. It is a generalization of the standard triangle *penguin* $\rightarrow$ *bird* $\rightarrow$ *fly* where the specificity *penguin* $>_\Delta$ *bird* is dictated by a single default and yields *penguin* $\wedge$ *bird* $\mathrel{|\!\!\!\sim_\Delta} \neg$*fly*. It turns out that, in this formulation, the specificity rule follows from the cumulative and direct inference rules.

Theorem 3 (Specificity) *If $\mathrel{|\!\!\!\sim}$ is a consequence relation that satisfies the cumulative and direct inference rules, then $\mathrel{|\!\!\!\sim}$ satisfies the specificity rule.*

2.4 Other Examples

In addition to the standard penguin examples, a number of "particularly nasty" [sic] examples were considered in [4]. It was suggested that these examples illustrate some lines of reasoning that are intuitively desirable for any default consequence relation. We will consider these lines of reasoning and show that every default consequence relation that satisfy the extended desiderata will handle them correctly. The first line of reasoning captures the intuition that defaults should be assumed to be independent and that they should be activated whenever possible. Let Δ be a set of n default $\varphi_i \rightarrow \psi_i$, $i = 1, \ldots, n$. We write φ^m and ψ^m to represent $\bigwedge_1^m \varphi_i$ and $\bigwedge_1^m \psi_i$ respectively, for $m = 1, \ldots, n$.

Rule 8 $\neg\psi^n \wedge \varphi^m \mathrel{|\!\!\!\sim_\Delta} \psi^m$ *for all $m < n$.*

This means that, although not all consequents can be true together, any proper subset of the consequents should still follow from their observed antecedents.

The next line of reasoning concerns the case where all the antecedents have been observed and it is known that at least one of the consequents is false; we should be adventurous and conclude that one and only one consequent is false. This is the essence of the following rule.

Rule 9 $\neg\psi^n \wedge \varphi^n \mathrel{|\!\!\!\sim_\Delta} \bigvee_{j=1}^n (\bigwedge_{i \neq j} \psi_i \wedge \neg\psi_j)$.

The next rule says that if all the antecedents are true, a proper subset of false consequents should not interfere with the (default) conclusion of the other consequents.

Rule 10 $\varphi^n \bigwedge_{i \in I} \neg\psi_i \mathrel{|\!\!\!\sim_\Delta} \bigwedge_{i \notin I} \psi_i$ *for all $I \subset \{1, \ldots, n\}$.*

We agree with [4] that the above rules represent intuitive default inferences. The following theorem tells us that these inferences are guaranteed whenever we satisfy the extended desiderata.

Theorem 4 (Independence) *If $\mathrel{|\!\!\!\sim_\Delta}$ satisfies the extended desiderata, then $\mathrel{|\!\!\!\sim_\Delta}$ satisfies rules 8 to 10.*

3 Realizing the Extended Desiderata

Having axiomatized the desired behavior of a default consequence relation, we will now present an interpretation of defaults that satisfies the extended desiderata. The approach is to extend system Z by adding constraints on admissible belief rankings. Since system Z has been shown to be characterized by Rule 1-5, the added constraints should enforce the coherence rule. We call this new system the *coherence closure*[2] of system Z.

3.1 Rankings and System Z: A Review

Definition 3 (Belief Rankings) *A* **belief ranking** *κ is a non-negative integer function on the set of worlds Ω such that $\kappa(\omega) = 0$ for some $\omega \in \Omega$. The belief rank of a wff φ is defined as:*

$$\kappa(\varphi) = \begin{cases} \min_{\omega \models \varphi} \kappa(\omega) & \text{if } \varphi \text{ is satisfiable} \\ \infty & \text{otherwise.} \end{cases}$$

Believability is associated with a lower rank and surprise or abnormality with a higher rank. Therefore, if $\kappa(\varphi) < \kappa(\psi)$, then φ is more believable than ψ. A default $\varphi \rightarrow \psi$ is interpreted as the constraint $\kappa(\varphi \wedge \psi) < \kappa(\varphi \wedge \neg\psi)$ that is, $\varphi \wedge \psi$ is more believable than $\varphi \wedge \neg\psi$. This leads to the notion of admissibility.

Definition 4 (Admissibility) *A belief ranking κ is* **admissible** *with respect to a defaults database Δ, if*

$$\kappa(\varphi \wedge \psi) < \kappa(\varphi \wedge \neg\psi)$$

for every default $\varphi \rightarrow \psi$ in Δ. We say Δ is **consistent** *if it has a Δ-admissible belief ranking.*

Each belief ranking κ induces a default consequence relation $\mathrel{|\!\!\!\sim_\kappa}$; where $\alpha \mathrel{|\!\!\!\sim_\kappa} \beta$ if and only if $\kappa(\alpha \wedge \beta) < \kappa(\alpha \wedge \neg\beta)$. Given a defaults database, system Z bases its inference on the most compact belief ranking; the unique ranking that assigns the lowest possible rank to each world while respecting the admissibility constraints. This reflects the assumption that the worlds are considered to be as believable (normal) as possible. It can be shown that the most compact belief ranking (as well as any other specific ranking) induces a consequence relation that is characterized by the desiderata (without the coherence rule).

[2]This new system may also be called a coherent rational closure to emphasize that the rational monotony rule is also satisfied by the system.

Table 1: Falsified Defaults.

ω	$\Delta[\omega]$
$\bar{b}\bar{p}\bar{f}, \bar{b}\bar{p}f, b\bar{p}f$	
$b\bar{p}\bar{f}, bp\bar{f}$	$b \rightarrow f$
bpf	$p \rightarrow \neg f$
$\bar{b}p\bar{f}$	$p \rightarrow b$
$\bar{b}pf$	$p \rightarrow b, p \rightarrow \neg f$

3.2 The Coherence Closure

To support the inheritance line of reasoning, we now impose *coherence* constraints on the admissible belief rankings.

Definition 5 (Coherence) *Given a defaults database Δ, we say that world ω is* **more coherent** *than world ν, written $\omega <_\Delta \nu$ if $\Delta[\omega] \subset \Delta[\nu]$. A ranking κ is* **coherent** *if $\kappa(\omega) < \kappa(\nu)$ whenever $\omega <_\Delta \nu$.*

We will write $\omega < \nu$ when the defaults database Δ is not pertinent to the discussion. The transitivity of the coherence constraints follow directly from the transitivity of the subset relation.

Theorem 5 (Transitivity of Coherence) *If $\omega > \omega'$ and $\omega' > \omega''$, then $\omega > \omega''$.*

As an illustration of coherence constraints, consider the standard example, $\Delta = \{b \rightarrow f, p \rightarrow b, p \rightarrow \neg f\}$. Table 1 shows the defaults that are falsified by each of the world. The coherence constraints induced by Δ are $\bar{b}pf > \bar{b}p\bar{f}, bpf$ and $\omega > \bar{b}\bar{p}\bar{f}, \bar{b}\bar{p}f, b\bar{p}f$ for all ω with a non-empty $\Delta[\omega]$. Clearly not every ranking satisfying these constraints is admissible; we therefore require both the admissibility and coherence conditions of c-admissibility.

Definition 6 (C-Admissibility) *Let Δ be a defaults database and let κ be a belief ranking. κ is* **c-admissible** *with respect to Δ if κ is both admissible and coherent. We say that Δ is* **c-consistent** *if Δ has a c-admissible belief ranking.*

Despite the addition of coherence constraints, c-consistency turns out to be no different from consistency.

Theorem 6 (Consistency Equivalence) *A set of defaults is c-consistent if and only if it is consistent.*

In [8, p. 25], a procedure for testing the consistency of a database of defaults was presented. The procedure requires $O(|\Delta|^2)$ satisfiability tests on the material counterparts[3] of the defaults in Δ [7]. Therefore the procedure is tractable for defaults databases that have Horn material counterparts [5].

As in system Z, we select the most compact belief ranking from the set of c-admissible rankings.

Definition 7 (The κ^c Ranking) *Let Δ be a consistent set of defaults. A belief ranking κ^c is the most compact c-admissible ranking if for all worlds ω*

$$\kappa^c(\omega) \leq \kappa(\omega)$$

for every κ that is c-admissible with respect to Δ.

[3] The material counterpart of $\varphi \rightarrow \psi$ is the wff $\varphi \supset \psi$.

The following theorem shows that the most compact ranking is unique.

Theorem 7 (Minimization) *Let κ_1 and κ_2 be belief rankings. If κ_1 and κ_2 are c-admissible, then $\kappa = \min(\kappa_1, \kappa_2)$ is c-admissible.*

Definition 8 (Coherence Closure) *Given a defaults database Δ, φ* **coherently entails** *ψ, written $\varphi \mid\!\sim_c \psi$ if and only if*

$$\kappa^c(\varphi \wedge \psi) < \kappa^c(\varphi \wedge \neg\psi).$$

We call $\mid\!\sim_c$ the coherence closure of system Z.

Since $\mid\!\sim_c$ satisfies Rule 1-5 [12] and the coherence constraints, it constitutes an interpretation of the extended desiderata (by Theorem 1).

3.3 Computing the Coherence Closure

The algorithm in Table 2 computes the most compact ranking relative to Δ and runs in time $O(2^n n^3)$ where n is the size of Δ. The key to computing the κ^c ranking lies in the construction of two non-negative integer-valued functions; a ranking $Z^c(d)$ on the set of defaults Δ, and a ranking $r(\phi)$ on a set of formulae $\Phi = \{\phi = \Phi(\Delta') \mid \Delta' \subseteq \Delta\}$. The Z^c rank records the minimal rank of worlds that satisfies[4] a default while the r rank records the minimal rank required for c-admissibility. The most compact ranking function $\kappa^c(\omega)$ is then obtained from $r(\phi)$ through the assignment

$$r(\omega) = r(\phi)$$

for every world $\omega \models \phi$.

The next theorem leads to a natural scheme of computing the Z^c and r ranks.

Theorem 8 (Goldszmidt [8, Theorem 2.4]) *Δ is c-consistent if and only if there is a default $d \in \Delta'$ such that d is tolerated[5] by Δ' for every non-empty subset $\Delta' \subseteq \Delta$.*

Given a c-consistent set of defaults D, we first compute the Z^c rank of every default that is tolerated by D and update the r ranking to ensure that it satisfies the requirements for c-admissibility. Next we remove from D the default that has the smallest Z^c rank. We repeat this procedure until all the defaults have been removed.

Some terminology used in the procedure: We define $\Delta[\phi_i] = \Delta'$ whenever $\phi_i = \Phi(\Delta')$. If $\Delta[\phi_i] \subset \Delta[\phi_j]$, then ϕ_i is an *ancestor* of ϕ_j. The minimal ancestors of ϕ_i are its *parents*, and ϕ_i is a *successor* of its ancestors and a *child* of its parents. Procedure `c-rank` assigns a rank $r(\phi_i)$ to satisfiable formulae of the form $\phi_i = \Phi(\Delta')$ where $\Delta' \subseteq \Delta$. The procedure assumes that the formulae ϕ_i are ordered such that every ϕ_i has a smaller index i than all its successors and a larger index than all its ancestors. Thus $\phi_0 = \Phi(\emptyset)$ and $\phi_n = \Phi(\Delta)$.

The following theorem tells us that the ranking r is precisely the κ^c ranking.

[4] A world ω satisfies a rule $\varphi \rightarrow \psi$ if $\omega \models \varphi \wedge \psi$.

[5] A default $\varphi \rightarrow \psi$ is *tolerated* by a set S if the wff $\varphi \wedge \psi \bigwedge_i \varphi_i \supset \psi_i$ is satisfiable (where i ranges over all the rules in S).

Table 2: Procedure `c-rank`

$r(\phi_0) = 0$
`For` i `from` 1 `to` n
  $r(\phi_i) = \max\{r(\phi_j) + 1 \mid \phi_j \in parent(\phi_i)\}$
`Let` $D = \Delta$
`While` $D \neq \emptyset$
  `Let` D_t `be the set of defaults tolerated by` D
  `For` $d = \varphi \rightarrow \psi \in D_t$
    $\gamma_d = \varphi \wedge \psi \bigwedge\{\varphi_i \supset \psi_i \mid \varphi_i \rightarrow \psi_i \in D\}$
    $Z^c(d) = \min\{r(\phi_i) \mid \phi_i \wedge \gamma_d$ `is satisfiable`$\}$
    `For` i `from` 1 `to` n
      `if` $d \in \Delta[\phi_i]$ `then`
        $r(\phi_i) = \max\{r(\phi_i), Z^c(d) + 1), r(\phi_j) + 1 \mid \phi_j \in parent(\phi_i)\}$
  `Remove from` D `a default in` D_t `with the lowest` Z^c `rank`

Theorem 9 (Rank) *The belief ranking as computed by procedure* `c-rank` *is precisely the most compact c-admissible ranking* κ^c.

In an extended technical report [16], we prove the correctness of the algorithm and analyze its complexity. We also establish a more economical representation of the ranking κ^c, from which queries of the type "does φ entail ψ?" can be answered.

4 Discussion and Conclusion

Coherence closure may be viewed as a refinement of ceteris paribum (cp) admissibility condition proposed in [15]. In addition to being admissible, a cp-admissible ranking must satisfy additional constraints, requiring a world ω to be ranked lower than a world ν whenever there is a subset $\Delta' \subset \Delta$ such that ω verifies all the defaults in Δ', ν falsifies all the defaults in Δ' and ω and ν agree on all the other defaults in $\Delta \setminus \Delta'$. Clearly every cp-condition is also a coherence constraint, but not the converse.

Motivated by Brewka's preferred subtheories [3], Boutilier [2] proposed an alternative way of resolving the tension between specificity and inheritance. Each default is assigned a priority and the set of defaults Δ is partitioned into $\{\Delta^i\}$ according to their priorities. Given two worlds ω and ν, attention is focused on the maximum priority ($i = \max(\omega, \nu)$) subset of defaults Δ^i where either $\Delta^i[\omega] \subset \Delta^i[\nu]$ or $\Delta^i[\nu] \subset \Delta^i[\omega]$. The world that falsifies a smaller set of defaults, say ω, is then considered strictly more believable than the other world, ν. Since $\Delta[\omega] \subset \Delta[\nu]$ implies $\Delta^i[\omega] \subseteq \Delta^i[\nu]$ for all i and $\Delta^i[\omega] \subset \Delta^i[\nu]$ for some i, it should be clear that if $\omega < \nu$ is a coherence constraint, then ω will also be considered strictly more believable than ν by [2]. The converse however is not true and is not desirable in general. The reason is that by comparing only the maximum priority subset of defaults $\Delta^{\max(\omega,\nu)}$ an implicit interpretation is imposed on the priorities. The implicit interpretation is that the violation of a default with priority i is somehow more "damaging" than the violation of any number of defaults of lower priority. While [3] gives little indication of the source of the priorities, [2] equates the priority of a default to the Z rank [12] of the negation of its material counterpart. In the following analysis we will attempt to show that the implicit interpretation of priorities when compounded with equating priorities to the Z rank of the negated material counterpart lead to counterintuitive inferences.

Consider the familiar example containing the defaults $s \rightarrow a$ ("students are typically adults"), $a \rightarrow e$ ("adults are typically employed") and $s \rightarrow \neg e$ ("students are typically not employed") to which we add a fourth default $x \rightarrow e$ where x is arbitrary. Since $s \rightarrow \neg e$ has priority 2 and $x \rightarrow e$ has priority 1, violation of the default "students are typically not employed" is considered to be more "damaging" than the violation of the default "x's are typically employed" in [2]. This is counterintuitive as this conclusion is obtained independently of the meaning of x. Thus even if x represents the statement "sells hamburgers in MacDonald's", violation of this statement is still considered less "damaging" than the violation of the default "students are typically not employed". Such counterintuitive behavior is not present in our semantics.

In [4], an approach based on preference orderings between worlds was presented. Beginning with an extant theory of default conditionals, any default that is true in this original theory *prefers* a world in which the material counterpart is true over a world in which it is false. In the new theory, a preference between two worlds is true when it can be attributed to some default and there is no *more specific* default that has a contradictory preference. The notion of *specificity* is also defined in terms of the truth of defaults in the original theory. This reliance on an extant theory of defaults make the system unnatural and complex. The intuitiveness of such a two step definition of a default theory is questionable and lacks a good philosophical justification. It is also not known if the desiderata is satisfied by the system.

In [6], inference rules 1 to 4 were supplemented by an *irrelevance* rule. Unfortunately, the rule involves the evaluation of a meta-logical irrelevance predicate $I(\cdot)$.

Much work in default reasoning has been guided by examples rather than on generally accepted principles of reasoning. We hope our general formalization of notions such as specificity, inheritance and coherence will help change this trend. It remains to be seen, though, whether the extended desiderata is sufficient for capturing other patterns of plausible reasoning.

References

[1] E. W. Adams. *The Logic of Conditionals*. D. Reidel, Dordrecht, Holland, 1975.

[2] C. Boutilier. What is a default priority? In *Proceedings of Canadian Society for Computational Studies of Intelligence Conference*, pages 140–147, Vancouver, 1992.

[3] G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. In *Proceedings of the International Joint Conference in Artificial Intelligence (IJCAI-89)*, Detroit, 1989.

[4] J. P. Delgrande. A preference-based approach to default reasoning: Preliminary report. In *Proceedings of the Twelth National Conference on Artificial Intelligence (AAAI94)*, pages 902–908, Seattle, WA, July 31 - August 4 1994. Morgan Kaufmann.

[5] W. Dowling and J. Gallier. Linear-time algorithms for testing the satisfiability of propositional Horn formulae. *Journal of Logic Programming*, 3:267–284, 1984.

[6] H. Geffner. *Default Reasoning: Causal and Conditional Theories*. PhD thesis, University of California Los Angeles, Cognitive Systems Lab., Los Angeles, November 1989.

[7] M. Goldszmidt and J. Pearl. System Z^+: A formalism for reasoning with variable strength defaults. In *Proceedings of American Association for Artificial Intelligence Conference*, Anaheim, CA, 1991.

[8] Moisés Goldszmidt. *Qualitative Probabilities: A Normative Framework for Commonsense Reasoning*. PhD thesis, University of California Los Angeles, Cognitive Systems Lab., Los Angeles, October 1992. Available as UCLA technical report (R-190).

[9] S. Kraus, D. J. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[10] D. J. Lehmann. Rational logics and their models: A study in cumulative logics. Technical Report TR-88-16, Department of Computer Science, Hebrew University, Jerusalem, Israel, 1988.

[11] D. Makinson. General theory of cumulative inference. In M. Reinfrank, J. de Kleer, M. Ginsberg, and E. Sandewall, editors, *Nonmonotonic Reasoning: Second International Workshop*, pages 1–18. Springer-Verlag, 1989. (Lecture Notes in Artificial Intelligence 346).

[12] J. Pearl. System Z: A natural ordering of default with tractable applications to nonmonotonic reasoning. In M. Vardi, editor, *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Third Conference*, pages 121–135, San Fancisco, 1990. Morgan Kaufmann.

[13] J. Pearl. Probabilistic semantics for nonmonotonic reasoning. In R. Cummins and J. Pollock, editors, *Philosophy and AI - Essays at the Interface*, pages 157–187. Bradford Books/MIT Press, 1991.

[14] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, Cambridge, MA, 1988.

[15] S.-W. Tan. Exceptional subclasses in qualitative probablity. In R. Lopez de Mantaras and D. Poole, editors, *Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference (UAI94)*, pages 553–559, Seattle, WA, July 29 - 31 1994. Morgan Kaufmann. Available as UCLA technical report R-215.

[16] S.-W. Tan and J. Pearl. Specificity and inheritance in default reasoning (extended report). Technical Report 224, University of California, Los Angeles, Jan 1995. In preparation.