Supraclassical Consequence Relations
Tolerating Rare Counterexamples

Willem Labuschagne[1], Johannes Heidema[2], and Katarina Britz[3]

[1] University of Otago, Dunedin, New Zealand,
willem@cs.otago.ac.nz,
[2] University of South Africa, Pretoria, South Africa,
[3] Centre for AI Research, CSIR Meraka Institute, Pretoria
University of KwaZulu-Natal, Durban, South Africa

Abstract. We explore a family of supraclassical consequence relations obtained by varying the criteria according to which counterexamples to classical entailment may be deemed tolerable. This provides a different perspective on the rational consequence relations of nonmonotonic logic, as well as introducing new kinds of entailment with a diversity of potential contextual applications.

Keywords: Nonmonotonic logic, Preference order, Induction, Abduction, Supraclassical consequence, Rational consequence, Dual preferential consequence, Correlative preferential consequence

1 Introduction

Classical logic focuses on the relatively cautious consequence relation $\models$, which is used to represent inferences from premisses φ to consequences ψ that preserve truth: $\varphi \models \psi$ if and only if every model of φ is a model of ψ. Mathematical reasoning mainly employs classical consequence.

In contrast, nonmonotonic logic belongs to a tradition in logic that considers $\models$ merely a useful cognitive reference point among several consequence relations that may be of interest. As Johan van Benthem [12] puts it:

> "The idea that logic is about just one notion of 'logical consequence' is actually one very particular historical stance. It was absent in the work of the great pioneer Bernard Bolzano, who thought that logic should chart the many different consequence relations that we have, depending on the reasoning task at hand."

When defining alternative consequence relations, one may choose to go subclassical or to go supraclassical or both. To go subclassical is to disallow some inferences from premiss to consequence that would be legitimate according to $\models$. One may wish, for example, to introduce a constraint of pertinence between premiss and conclusion as in [2], which results in a stricter criterion than $\models$. The result is that some of the classically permitted inferences become illegitimate.

Nonmonotonic logic chooses instead to go supraclassical, adopting a criterion less strict than $\models$ so as to accommodate various forms of common-sense reasoning in which agents compensate for limited information by using heuristic rules of thumb. Arguably, the most important thing about nonmonotonic logic is not its nonmonotonicity but its supraclassicality. Supraclassical consequence relations are ampliative extensions of $\models$, i.e. if $\mid\!\sim$ is supraclassical, then $\varphi \models \psi$ is a sufficient condition for $\varphi \mid\!\sim \psi$. Thus supraclassical consequence relations build on the pairs in $\models$ by adding new inference-pairs legitimised by the agent's heuristic information. It is these additional inference-pairs that may need to be retracted in the light of new evidence, giving rise to nonmonotonicity.

The endeavour to develop useful supraclassical logics was set forth in the landmark paper [8], in which the extra heuristic information needed for venturing beyond $\models$ was encoded in a preference order on states. Consequence relations satisfying different groups of postulates were described. One particular group of postulates (in some sense the strongest) characterised the rational consequence relations [9], which have since earned a privileged position in virtue of their close connection with AGM belief change [13]. However, focusing only on rational consequence relations would give nonmonotonic logic a rather monolithic face.

When we reflect on everyday human reasoning, it becomes apparent that this complex scene comprises diversified, even disparate, subfields, potentially exceeding Charles Sanders Peirce's division of reasoning into deduction, induction, and abduction [5, 6]. Along one dimension the intentions and goals of the reasoning agents may differ: to induce a plausible prediction; to abduce a plausible cause; to test the plausibility of some purported entailment; to speculate (but not too wildly); etc. Along another dimension the heuristic information embodied in a specific preference order on states may be expressing a comparison relative to diverse attributes of those states: normality; typicality; likelihood; frequency of occurrence; resemblance to some real or ideal state; closeness in time according to the plan for some important event; place in some causal order in accord with relevant laws of nature; degree of compliance with some norm; etc.

We shall describe a family of supraclassical entailment relations among which are the rational consequence relations that constitute the industry standard for nonmonotonic logic. The family also includes hitherto unexamined relations that deserve scrutiny because of the very natural way in which they arise within our semantic framework. The product is an articulated range of consequence relations all sharing the preferential paradigm but varying in aptness for specific contexts.

The novelty of the contribution resides in the method by which different members of the family are generated. Whereas [8] arrives at different consequence relations by varying the type of order relation on states or the association between states and valuations, we shall take these to be fixed as if for rational consequence relations, and instead use the preference relation on states to control, in a nuanced way, the addition of new inference-pairs to $\models$.

2 Preliminaries

Henceforth L_A denotes a propositional language generated from the atomic sentences in A by the usual set of connectives $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$, with $\top$ designating an arbitrary tautology. The language is equipped with a semantics in the manner familiar from nonmonotonic logic [8, 9], so that the semantic structure comprises a set S of states, the set W_A of all valuations $w : A \longrightarrow \{0, 1\}$, a labelling function $\ell : S \longrightarrow W_A$, and a suitable order relation $\preccurlyeq$ on S. As is customary, $\preccurlyeq$ encodes heuristic information about preference, typicality, or likelihood that cannot in general be expressed propositionally in L_A (but see Section 8).

We shall deviate from [8] and [9] in three ways.

Firstly, we shall assume that A and S are finite. We consider this finiteness assumption to be justified both by an interest in everyday reasoning (as opposed to metamathematical applications) and as a technical simplification appropriate for an initial scrutiny of a new landscape. (A consequence of the finiteness assumption is that smoothness for $\preccurlyeq$ is automatic.)

Secondly, we shall assume that $\preccurlyeq$ is a total preorder on S (i.e. is reflexive on S, transitive, and connected in the sense that for all $s, t \in S$ at least one of $s \preccurlyeq t$ and $t \preccurlyeq s$ is the case). Total preorders provide a unified framework for both nonmonotonic logic and belief revision, and the strict modular partial orders used in [9] are just strict versions of total preorders.

Finally, we take states higher up in the order to be more preferred, more typical, or more likely to occur. For historical reasons the order is often inverted, as in [8] and [9], but we follow Shoham ([11], p.74) in respecting the common intuition that 'up' is 'more' when 'more' qualifies positive attributes such as normality, typicality, or likelihood rather than negative qualities such as abnormality, atypicality, or unlikelihood.

Satisfaction is defined as usual. For every sentence $\varphi \in L_A$, $Mod(\varphi)$ denotes the set of models of φ, i.e. the set of all $s \in S$ for which $\ell(s)$ renders φ true, and $Max(\varphi)$ the maximal models of φ with respect to $\preccurlyeq$.

For purposes of illustration we shall take $A = \{p, q\}$ with $S = W_A = \{11, 10, 01, 00\}$ where 10 denotes the state (i.e. valuation) in which p is true and q is false, and so forth. As an interpreted language, we may consider the system of interest to be either the Traffic System in which p abbreviates *The light for oncoming traffic is red* and q stands for *The oncoming car stops*, or the Light-Fan System in which p abbreviates *The light is on* and q stands for *The fan is on*.

3 From counterexamples to supraclassicality

Since $\varphi \models \psi$ if and only if $Mod(\varphi) \subseteq Mod(\psi)$, a strictly supraclassical consequence relation must have at least one pair (φ, ψ) for which $\varphi \not\models \psi$ and thus for which there is some model u of φ that fails to be a model of ψ. Think of state u as a *counterexample* to $\varphi \models \psi$, and thus in some sense a 'bad guy'. More generally, $Mod(\varphi \wedge \neg\psi)$ is the set of bad guys (i.e. counterexamples).

Not all bad guys are equally bad. The total preorder $\preccurlyeq$ on S stratifies the states into levels stacked from the least normal (most rare, most atypical) to the most normal (most likely, most typical). In this context some counterexample states may be relatively exceptional (less normal, less typical, or less likely than others). The guiding idea behind supraclassicality is to tolerate such not-so-very-bad guys, just as in life we commonly view those guilty of an uncharacteristic lapse of judgment less harshly than habitual criminals.

We wish to allow an inference-pair (φ, ψ) into the consequence relation as long as the counterexamples in $Mod(\varphi \wedge \neg\psi)$ are not maximally bad. The new insight of which we take advantage is that the worst counterexample states, those that are maximal, may be identified in four different ways. Maximality is relative not only to the ordering but also to the subset within which an element is considered to be maximal. Different supraclassical relations may be obtained by varying the subsets of S against which counterexamples are evaluated. These subsets must be supersets of $Mod(\varphi \wedge \neg\psi)$.

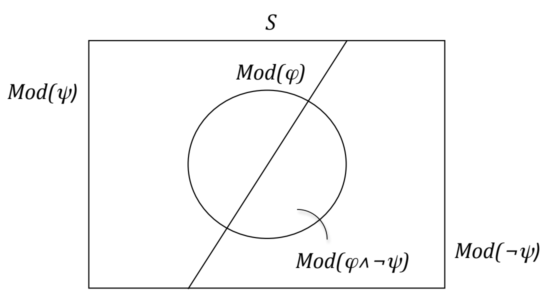

Fig. 1. Important subsets of S

On the face of it, there are four obvious supersets of $Mod(\varphi \wedge \neg\psi)$ within which counterexamples could be maximal:

1. $X_1 = Mod(\varphi)$
2. $X_2 = Mod(\neg\psi)$
3. $X_3 = Mod(\varphi) \cup Mod(\neg\psi)$
4. $X_4 = S$.

As a generic symbol for a supraclassical consequence relation we use $\mid\!\sim$ decorated by a subscript designating one of the four supersets of $Mod(\varphi \wedge \neg\psi)$. Thus we explore four supraclassical consequence relations, $\mid\!\sim_1$, $\mid\!\sim_2$, $\mid\!\sim_3$ and $\mid\!\sim_4$, which are all defined similarly:

$$\varphi \mid\!\sim_i \psi \text{ if and only if } Mod(\varphi \wedge \neg\psi) \cap Max(X_i) = \varnothing.$$

Relative to a fixed semantics, we may speak of 'the' relation $\mid\!\sim_i$ rather than 'a' relation $\mid\!\sim_i$.

4 Rational Consequence

Consider $\mid\!\sim_1$. In this case the set of counterexamples $Mod(\varphi\wedge\neg\psi)$ is viewed as a subset of $X_1 = Mod(\varphi)$, and thus $\varphi \mid\!\sim_1 \psi$ if and only if $Mod(\varphi\wedge\neg\psi)\cap Max(\varphi) = \varnothing$.

It follows that $\varphi \mid\!\sim_1 \psi$ if and only if every maximal model of φ is a model of ψ. In other words, $\mid\!\sim_1$ is the familiar rational consequence relation of nonmonotonic logic, defined as in [9].

A great deal is known about the properties of rational consequence relations and how they differ from those of $\models$. For purposes of comparison, we give a swift and selective recapitulation, beginning with some familiar properties of $\models$ that fail to hold for $\mid\!\sim_1$.

The relation $\mid\!\sim_1$ is famously *nonmonotonic*, i.e. there exist instances of $\mid\!\sim_1$ such that for some sentences $\varphi, \psi, \alpha \in L_A$ we have $\varphi \mid\!\sim_1 \psi$ but not $\alpha \wedge \varphi \mid\!\sim_1 \psi$. As hinted earlier, we should not ascribe overwhelming importance to nonmonotonicity. For some forms of reasoning, nonmonotonicity may be appropriate; for others, it may not. We shall see that the supraclassical consequence relation $\mid\!\sim_2$ is in fact monotonic, and that this appears appropriate given that its utility is in a sense complementary to that of $\mid\!\sim_1$.

Contraposition fails for $\mid\!\sim_1$, since it is possible to have $\varphi \mid\!\sim_1 \psi$ while failing to have $\neg\psi \mid\!\sim_1 \neg\varphi$. This has significance for $\mid\!\sim_2$, as we shall see.

Also, $\mid\!\sim_1$ fails the Ramsey test relative to the conditional $\rightarrow$ since it is not in general the case that $\varphi \mid\!\sim_1 \psi$ if and only if $\top \mid\!\sim_1 (\varphi \rightarrow \psi)$. This failure gains an interesting twist when we discuss $\mid\!\sim_4$ later.

Various other properties of $\models$ fail for $\mid\!\sim_1$, but will not be addressed here.

Turning to properties that do hold in general for $\mid\!\sim_1$, we may note the observance of a weaker form of monotonicity, rational monotonicity, which accounts for the *rational* in the name given to this type of consequence relation:

$$\text{if } \varphi \mid\!\sim_1 \psi \text{ and it is not the case that } \varphi \mid\!\sim_1 \neg\alpha, \text{ then } \alpha \wedge \varphi \mid\!\sim_1 \psi.$$

Although unqualified monotonicity fails for $\mid\!\sim_1$, the property called *right weakening*, which is a sort of dual of monotonicity obtained by replacing $\wedge$ by $\vee$ and left with right, does hold for $\mid\!\sim_1$:

$$\text{if } \varphi \mid\!\sim_1 \psi \text{ then } \varphi \mid\!\sim_1 \psi \vee \alpha.$$

We shall see the relevance of the duality between monotonicity and right weakening when we examine $\mid\!\sim_2$.

We may also make an algebraic observation. The reader will recall that in classical propositional logic, the equivalence classes of sentences form a Boolean algebra (hereafter referred to as the Lindenbaum-Tarski algebra) having $\models$ as the order relation, conjunction as meet, disjunction as join, the class $\top$ of tautologies as maximum, and the class $\bot$ of contradictions as minimum. The set of consequences of a premiss φ under $\models$ is a filter of the algebra, and in the other direction the set of premisses entailing a fixed consequence under $\models$ forms an ideal. This classical picture is neatly bisected by $\mid\!\sim_1$ and $\mid\!\sim_2$.

In respect of $\mid\!\sim_1$, we note that for a fixed premiss φ the set $\{\psi \mid \varphi \mid\!\sim_1 \psi\}$ is a filter in the Lindenbaum-Tarski algebra of propositions, just as in the classical case. In contrast, for a fixed conclusion ψ the set $\{\varphi \mid \varphi \mid\!\sim_1 \psi\}$ is not necessarily an ideal, because of nonmonotonicity.

Concluding our brief overview, the significance of the rational consequence relation $\mid\!\sim_1$ is that arguably it formalises inductive reasoning of the kind we might call *singular predictive inference* in order to distinguish it from other meanings of "induction", such as learning a default rule from a finite set of instances (generalisation). An example will serve to illustrate, but a deeper analysis may be found in [3] and [4], as well as in Section 7.

Example 1. Consider the language $L_{\{p,q\}}$ with a semantics consisting of $S = \{11, 10, 01, 00\}$ and the total preorder $\preccurlyeq$ that stratifies S into two levels, with 10 and 01 on the bottom and 11 and 00 at the top. Think of this as a knowledge representation language for the Traffic System. The preorder $\preccurlyeq$ depicts the heuristic that it is normal for the oncoming traffic to stop if their traffic light is red (the state 11) and it is normal for the oncoming traffic to continue driving without stopping if their traffic light is not red (the state 00). Suppose we are waiting at the intersection and observe that p is the case, i.e. that the light for oncoming cross traffic is red. We need to decide whether q is the case, i.e. whether the oncoming car will stop. Although $p \not\models q$, the reason we get to work every morning is that we are able to predict that the oncoming car will stop and so proceed fearlessly to cross the intersection ourselves. Our prediction is sanctioned by $\mid\!\sim_1$, since $p \mid\!\sim_1 q$. The prediction may of course turn out to be wrong, because $\mid\!\sim_1$ relies on uncertain heuristic information. In the case of a disastrous falsification of q, the disaster embodies a tragic, though presumably rare, counterexample to $p \models q$. Had the prediction instead been sanctioned by $\models$, we could have proceeded across the intersection secure in the knowledge that no drunk driver would run the red light. Sadly, everyday decision-making seldom enjoys the luxury of sufficient information to dispense with $\mid\!\sim_1$ and rely on $\models$.

5 Dual Preferential Consequence

Rational consequence in [9] (Section 3) corresponds to *modularity* of the preferential order, which is there taken to be a strict partial order on S, and in our exposition to *totality* of the preferential preorder. In [9] "preferential consequence" (and related terminology) does not connote modularity of the corresponding order. In this article, however, we use "preferential" throughout while always staying with the stipulation in Section 2 that $\preccurlyeq$ is a *total* preorder on S.

Consider $\mid\!\sim_2$. In this case the set of counterexamples $Mod(\varphi \wedge \neg\psi)$ is viewed as a subset of $X_2 = Mod(\neg\psi)$, and thus $\varphi \mid\!\sim_2 \psi$ if and only if $Mod(\varphi \wedge \neg\psi) \cap Max(\neg\psi) = \varnothing$.

By the definition, although it is possible that some model of φ may fail to satisfy ψ, that model is not to be a typical (i.e. maximal) model of $\neg\psi$ but instead is required to be somewhat atypical among the states that falsify ψ.

It is not hard to see that $\varphi \mathrel{|\!\sim}_2 \psi$ if and only if $Mod(\varphi) \subseteq S \setminus Max(\neg\psi)$. Hence $\mathrel{|\!\sim}_2$ is precisely the dual preferential consequence relation studied in [1].

The relation $\mathrel{|\!\sim}_2$ is related to the rational consequence relation $\mathrel{|\!\sim}_1$ in a manner that contraposition would have rendered trivial had this property held for $\mathrel{|\!\sim}_1$:

$$\varphi \mathrel{|\!\sim}_2 \psi \text{ if and only if } \neg\psi \mathrel{|\!\sim}_1 \neg\varphi.$$

This intimate connection between $\mathrel{|\!\sim}_1$ and $\mathrel{|\!\sim}_2$ invites a quick comparison of features.

We first note that whereas $\mathrel{|\!\sim}_1$ is nonmonotonic, monotonicity holds for $\mathrel{|\!\sim}_2$:

$$\text{if } \varphi \mathrel{|\!\sim}_2 \psi \text{ then } \alpha \wedge \varphi \mathrel{|\!\sim}_2 \psi.$$

As previously observed, monotonicity is in some sense a dual of right weakening, and there is a general pattern of properties holding for $\mathrel{|\!\sim}_2$ if they are the duals of properties holding for $\mathrel{|\!\sim}_1$, as explained in [1]. For example, although right weakening fails to hold for $\mathrel{|\!\sim}_2$, the weaker form of right weakening that is the dual of rational monotonicity does hold:

$$\text{if } \varphi \mathrel{|\!\sim}_2 \psi \text{ and it is not the case that } \neg\alpha \mathrel{|\!\sim}_2 \psi \text{ then } \varphi \mathrel{|\!\sim}_2 \psi \vee \alpha.$$

Algebraically, for a fixed conclusion ψ the set $\{\varphi \mid \varphi \mathrel{|\!\sim}_2 \psi\}$ is an ideal of the Lindenbaum-Tarski algebra, just as in the classical case, but for a fixed premiss φ the set of consequences $\{\psi \mid \varphi \mathrel{|\!\sim}_2 \psi\}$ may fail to be a filter.

Finally, the significance of $\mathrel{|\!\sim}_2$ is that it arguably formalises a kind of abductive reasoning, i.e. $\varphi \mathrel{|\!\sim}_2 \psi$ can be interpreted as "φ partially explains ψ". An example will serve to illustrate, until Section 7.

Example 2. Consider again the language $L_{\{p,q\}}$ with a semantics consisting of $S = \{11, 10, 01, 00\}$ and the total preorder $\preccurlyeq$ that stratifies S into two levels, with 10 and 01 on the bottom and 11 and 00 at the top. Think of this as a knowledge representation language for the Light-Fan System, where p stands for "The light is on" and q for "The fan is on". Now suppose we observe that $\neg p \wedge q$ is the case. The semantic constraint $\mathrel{|\!\sim}_2$ admits several different explanations amongst which is the sentence $(p \vee q) \wedge \neg(p \wedge q)$, which we may abbreviate by $p + q$. To see that $p + q \mathrel{|\!\sim}_2 \neg p \wedge q$, note that although $10 \in Mod(p + q)$ and 10 is a counterexample to $p + q \models \neg p \wedge q$, 10 is not maximal in $Mod(\neg(\neg p \wedge q))$. Intuitively, the observation that the light is off while the fan is on has been explained by the conjecture that only one component can be on at a time.

6 Correlative Preferential Consequence

Consider $\mathrel{|\!\sim}_3$. In this case the set of counterexamples $Mod(\varphi \wedge \neg\psi)$ is viewed as a subset of $X_3 = Mod(\varphi) \cup Mod(\neg\psi)$, and thus $\varphi \mathrel{|\!\sim}_3 \psi$ if and only if $Mod(\varphi \wedge \neg\psi) \cap Max(Mod(\varphi) \cup Mod(\neg\psi)) = \varnothing$.

Recalling that $Mod(\varphi) \cup Mod(\neg\psi) = Mod(\psi \rightarrow \varphi)$, we get that

$$\varphi \mathrel{|\!\sim}_3 \psi \text{ if and only if } Mod(\varphi \wedge \neg\psi) \subseteq S \setminus Max(\psi \rightarrow \varphi).$$

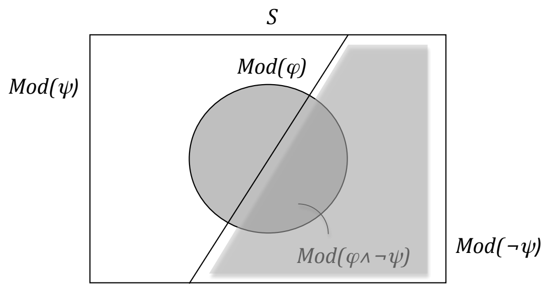

Fig. 2. The shaded region depicts $X_3 = Mod(\varphi) \cup Mod(\neg\psi)$

As far as the authors are aware, the relation $\mid\!\sim_3$ has not previously been studied, and its intuitive meaning, properties, and potential uses remain to be fully elucidated. Nevertheless, we are able to report some preliminary insights.

One striking aspect is that the counterexample states in $Mod(\varphi \wedge \neg\psi)$ are interdicted from the set of maximal models of $\psi \to \varphi$, giving the following surprising connection between $\mid\!\sim_3$ and the rational consequence relation $\mid\!\sim_1$:

$$
\begin{aligned}
\varphi \mid\!\sim_3 \psi \quad & \text{if and only if} \quad Mod(\neg(\varphi \to \psi)) \subseteq S \setminus Max(\psi \to \varphi) \\
& \text{if and only if} \quad Max(\psi \to \varphi) \subseteq S \setminus Mod(\neg(\varphi \to \psi)) \\
& \text{if and only if} \quad Max(\psi \to \varphi) \subseteq Mod(\varphi \to \psi) \\
& \text{if and only if} \quad (\psi \to \varphi) \mid\!\sim_1 (\varphi \to \psi) \\
& \text{if and only if} \quad (\psi \to \varphi) \mid\!\sim_1 (\varphi \leftrightarrow \psi).
\end{aligned}
$$

If $\varphi \mid\!\sim_3 \psi$ says that the truth of the conditional $\psi \to \varphi$ renders plausible the converse conditional $\varphi \to \psi$, one is reminded of psychological experiments showing that under some circumstances humans have a tendency to infer the converse from a conditional premiss in exactly this way [7, pages 51–54]. Perhaps $\mid\!\sim_3$ represents thought patterns deriving from mental models in the sense of Johnson-Laird — a topic for future research.

Further reflection reveals that, since $\preccurlyeq$ is a total preorder on S, the set $Max(\psi \to \varphi)$ from which counterexamples are interdicted must be one of the following three sets: $Max(\varphi)$, $Max(\neg\psi)$, or $Max(\varphi) \cup Max(\neg\psi)$. This allows us to prove that $\mid\!\sim_3$ is related to both $\mid\!\sim_1$ and $\mid\!\sim_2$ in an elegantly balanced way:

- If $Max(\psi \to \varphi) = Max(\varphi)$ then $Mod(\varphi \wedge \neg\psi)$ has no member in $Max(\varphi)$ and so $\varphi \mid\!\sim_1 \psi$
- If $Max(\psi \to \varphi) = Max(\neg\psi)$ then $Mod(\varphi \wedge \neg\psi)$ has no member in $Max(\neg\psi)$ and so $\varphi \mid\!\sim_2 \psi$
- If $Max(\psi \to \varphi) = Max(\varphi) \cup Max(\neg\psi)$ then both $\varphi \mid\!\sim_1 \psi$ and $\varphi \mid\!\sim_2 \psi$.

Summarising, if $\varphi \mid\!\sim_3 \psi$ then $\varphi \mid\!\sim_1 \psi$ or $\varphi \mid\!\sim_2 \psi$ or both.
Conversely, if both $\varphi \mid\!\sim_1 \psi$ and $\varphi \mid\!\sim_2 \psi$, then $\varphi \mid\!\sim_3 \psi$.
Overall: $(\mid\!\sim_1 \cap \mid\!\sim_2) \subseteq \mid\!\sim_3 \subseteq (\mid\!\sim_1 \cup \mid\!\sim_2)$.

7 More on Rational, Dual, and Correlative Preferential Consequence

In our introduction we referred to Peirce's articulation of reasoning into deduction, induction, and abduction (concisely delineated in [10]). Classical $\models$ is the canonical way to secure "deduction". In Example 1 we illustrated how $\varphi \mathrel{|\!\sim_1} \psi$ may be one way to assign a formal meaning to the following notion: "from evidence φ we induce — defeasibly but plausibly — the prediction that also ψ". And in Example 2 we propounded $\varphi \mathrel{|\!\sim_2} \psi$ for "from evidence ψ we abduce the hypothesis φ as a plausible partial explanation of ψ". In $|\!\sim_3$ we have a type of balanced, correlated combination of $|\!\sim_1$ and $|\!\sim_2$, roaming the interval between $|\!\sim_1 \cap |\!\sim_2$ and $|\!\sim_1 \cup |\!\sim_2$. But a sharper focus on φ as the given evidence versus ψ as the given evidence may enhance intuitive appreciation of the dual roles of $|\!\sim_1$ and $|\!\sim_2$.

When glossing $\varphi \mathrel{|\!\sim_1} \psi$ as "the information (evidence, observation) expressed by φ is given and affords, quite plausibly, that we have ψ too", φ is given, fixed, presumably reliable, and the actual state (if pertinent) is one in $Mod(\varphi)$. In contrast, ψ is now much more fluttery. Many different predictions may be underwritten by φ. When accepting defeasibility of prediction ψ by tolerating some of the counterexamples in $Mod(\varphi \wedge \neg\psi)$, the only dependable information we have against which to evaluate them sits embedded in $Mod(\varphi)$, which contains all of these counterexamples. So it is reasonable to tolerate only those counterexamples not maximally likely in $Mod(\varphi)$, i.e. those in $Mod(\varphi) \setminus Max(\varphi)$. The actual state (if pertinent) then would likely sit in $Max(\varphi)$ and not be a counterexample — a comforting thought.

When glossing $\varphi \mathrel{|\!\sim_2} \psi$ as "the information (evidence, observation) expressed by ψ is given and is afforded, quite plausibly, by the explanatory hypothesis φ", ψ is given, fixed, presumably reliable, and the actual state (if pertinent) is one in $Mod(\psi)$. Many hypotheses may constitute plausible if partial explanations of ψ, among them φ. (The monotonicity of $|\!\sim_2$ in this explanatory context seems agreeable: if φ plausibly and partially explains ψ, then so does $\varphi \wedge \alpha$. A next stage in one's abductive endeavour may then be the somewhat controversial search for the *best* explanation of ψ.) All counterexamples to $\varphi \models \psi$ falsify ψ, sit in $Mod(\neg\psi)$, but among them we tolerate only those states that do not add insult to injury by being very likely to occur. So to be tolerable a counterexample must not be "maximally bad", i.e. must not belong to $Max(\neg\psi)$.

In a correlative preferential entailment $\varphi \mathrel{|\!\sim_3} \psi$, we may be in a context where the agent's aim is neither to predict ψ from φ, nor to explain ψ by φ. No extra evidence supports either φ or ψ in an unbalanced way. When tolerating some counterexamples, the information in $\preccurlyeq$ now plays a balanced role with regard to proscribing states that support φ but violate ψ and is used to interdict those that are seriously embarrassing the entailment of ψ by φ by their maximally prominent presence in the crowd of kindred states which support φ or violate ψ.

The information embodied in $(S, \preccurlyeq, \ell)$ determines in a unique way all three consequence relations $|\!\sim_1$, $|\!\sim_2$, and $|\!\sim_3$ on the sentences of L_A. We now assume that $\ell : S \longrightarrow W_A$ is injective and then show that $\preccurlyeq$ can be recovered from each of $\mid\!\sim_1$ and $\mid\!\sim_2$.

A *literal* or *diagrammatic sentence* is any atomic sentence in A or the negation of such an atom. The *state description* or *diagram* of state $s \in S$ is the sentence $\delta(s)$ that is the conjunction of all the literals made true by $\ell(s)$. For instance, if $\ell(s) = 10$, then $\delta(s) = p \wedge \neg q$. Let s and t be any two (possibly equal) states. We demonstrate that any information about s and t in $\preccurlyeq$ can be retrieved from the corresponding relation $\mid\!\sim_1$ on sentences that are state descriptions or disjunctions of such. Remember that $Mod(\delta(s)) = \{s\}$. No undue cognitive torment is incurred when verifying the following:

$$
\begin{aligned}
&s \prec t \quad \text{iff} \quad \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(t) \text{ and not } \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(s) \\
&s = t \quad \text{iff} \quad \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(s) \text{ and } \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(t) \\
&s \preccurlyeq t,\ t \preccurlyeq s, \text{ and } s \neq t \text{ iff not } \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(s) \text{ and not } \delta(s) \vee \delta(t) \mid\!\sim_1 \delta(t).
\end{aligned}
$$

So, when ℓ is one-to-one, then $\mid\!\sim_1$ harbours exactly the same information as $\preccurlyeq$. Since $\mid\!\sim_2$ is the dual or contrapositive relation of $\mid\!\sim_1$ ($\varphi \mid\!\sim_2 \psi$ iff $\neg\psi \mid\!\sim_1 \neg\varphi$), $\mid\!\sim_2$ on Boolean combinations of state descriptions also contains the same heuristic information as $\preccurlyeq$.

8 Contextual Rules

Consider $\mid\!\sim_4$. In this case the set of counterexamples $Mod(\varphi \wedge \neg\psi)$ is viewed as a subset of $X_4 = S$, and thus $\varphi \mid\!\sim_4 \psi$ if and only if $Mod(\varphi \wedge \neg\psi) \cap Max(S) = \varnothing$.

While this entailment relation is quite new, it has a simple intuitive basis. Recall that in the case of classical entailment we have, if $\top$ denotes your favourite tautology,

$$\varphi \models \psi \text{ if and only if } \top \models (\varphi \rightarrow \psi).$$

Now we observe that although, as previously noted, it is not in general the case that $\varphi \mid\!\sim_1 \psi$ if and only if $\top \mid\!\sim_1 (\varphi \rightarrow \psi)$, the latter condition is of independent interest, because:

$$
\begin{aligned}
\varphi \mid\!\sim_4 \psi \quad &\text{if and only if} \quad Mod(\varphi \wedge \neg\psi) \subseteq S \setminus Max(S) \\
&\text{if and only if} \quad Max(S) \subseteq Mod(\varphi \rightarrow \psi) \\
&\text{if and only if} \quad \top \mid\!\sim_1 (\varphi \rightarrow \psi).
\end{aligned}
$$

This relationship between $\mid\!\sim_4$ and $\mid\!\sim_1$ affords a new and surprising way to understand the former, namely as deduction from a knowledge base. Suppose an agent has background information expressed by a sentence κ. This background information can be encoded into a preference order on states — simply take the dichotomous total preorder on S that places all models of κ on the upper level and all nonmodels of κ on the lower level. Since $Max(S) = Max(\top) = Mod(\kappa)$ it follows that

$$\kappa \models \psi \quad \text{if and only if} \quad \top \mid\!\sim_1 \psi$$

where we have used the dichotomous preorder induced by κ for defining the rational consequence relation $\mid\!\sim_1$.

Hence in the context of background information κ we see that

$$\begin{aligned} \varphi \mid\!\sim_4 \psi \quad & \text{if and only if} \quad \top \mid\!\sim_1 (\varphi \rightarrow \psi) \\ & \text{if and only if} \quad \kappa \models (\varphi \rightarrow \psi). \end{aligned}$$

The real surprise lurks in the converse. Suppose we start (not with sentence κ, but) with some $\preccurlyeq$ and consider the corresponding $\mid\!\sim_4$. The simplification $\preccurlyeq^*$ of $\preccurlyeq$ into the dichotomous preference order on S with $Max(S)$ in the top level and $S \setminus Max(S)$ in the bottom level yields again the same $\mid\!\sim_4$. And $\preccurlyeq^*$ (or $\preccurlyeq$) yields the corresponding explicit object-language background informational sentence $\kappa = \bigvee\{\delta(s) \mid s \in Max(S)\}$. Now $\preccurlyeq^*$, $\mid\!\sim_4$, and κ harbour exactly the same information. And, against the background of this informational context, we may construe $\varphi \mid\!\sim_4 \psi$ as the conditional rule $\varphi \rightarrow \psi$.

9 Future Research

Impelled by the conviction that it would be a mistake to seek a single 'correct' consequence relation for the formalisation of common-sense reasoning and the belief that appropriate candidates would be supraclassical, we have described a coherent family of supraclassical consequence relations $\mid\!\sim_i$ within a single unifying framework.

In so doing, we tolerated all but the most habitual criminals in the set $Mod(\varphi \wedge \neg\psi)$ of counterexamples. Accordingly we may think of the family of $\mid\!\sim_i$ as the 'liberal' supraclassical relations, each of which includes inference-pairs (φ, ψ) as long as the counterexample states in $Mod(\varphi \wedge \neg\psi)$ are not maximal in the relevant superset of $Mod(\varphi \wedge \neg\psi)$.

There is an intriguing extension of this family. Tolerating only those criminals whose transgressions are very rare would deliver 'conservative' supraclassical relations, from each of which inference-pairs (φ, ψ) are excluded unless the counterexample states, if any, are all minimal in the relevant superset.

Recall the four supersets of $Mod(\varphi \wedge \neg\psi)$ within which counterexamples could be either 'not maximal' or 'minimal': $X_1 = Mod(\varphi)$, $X_2 = Mod(\neg\psi)$, $X_3 = Mod(\varphi) \cup Mod(\neg\psi)$, and $X_4 = S$.

As a generic symbol for a conservative supraclassical entailment relation we may use $\mid\!\approx$, with appropriate subscript. The conservative $\mid\!\approx_1$, $\mid\!\approx_2$, $\mid\!\approx_3$ and $\mid\!\approx_4$ would most naturally be defined by

$$\varphi \mid\!\approx_i \psi \text{ if and only if } Mod(\varphi \wedge \neg\psi) \subseteq Min(X_i).$$

However, there is a subtle problem with this. We would wish to allow the total preorder $\preccurlyeq$ on S to be $S \times S = \{(s, s') \mid s, s' \in S\}$, the relation of complete preferential equity between all states, in order to include the case of an agent with no discriminatory heuristic information at all. Unfortunately, $Min(X_i) = Max(X_i)$ in this case, which causes the constraint $Mod(\varphi \wedge \neg\psi) \subseteq Min(X_i)$ to violate the guiding intuition that counterexample states should certainly not be maximally typical or maximally likely in X_i.

We therefore introduce the notation $Min\,\cancel{Max}\,(X_i) = Min(X_i) \setminus Max(X_i)$ and define the four conservative entailment relations by

$$\varphi \mathrel{|\!\!\approx}_i \psi \text{ if and only if } Mod(\varphi \wedge \neg\psi) \subseteq Min\,\cancel{Max}\,(X_i).$$

In the limiting case of an agent with no heuristic information (i.e. when $\preccurlyeq$ is $S \times S$), we now have that both the liberal consequence relations $\mathrel{|\!\sim}_i$ and the conservative consequence relations $\mathrel{|\!\!\approx}_i$ collapse to $\models$.

Precisely how the conservative consequence relations are related to the liberal consequence relations is as yet *terra* very much *incognita*.

References

1. Britz, K., Heidema, J., Labuschagne, W.: Semantics for dual preferential entailment. *Journal of Philosophical Logic* **38** (2009) 433–446
2. Britz, K., Heidema, J., Varzinczak, I.: Constrained consequence. *Logica Universalis* **5** (2011) 327–350
3. Freund, M.: On the notion of concept I *Artificial Intelligence* **172** (2008) 570–590
4. Freund, M.: On the notion of concept II *Artificial Intelligence* **173** (2009) 167–179
5. Hauser, N., Kloesel, C.: *The Essential Peirce*, Volume 1 (1867–1893), Bloomington: Indiana University Press, 1992
6. Hauser, N., Kloesel, C.: *The Essential Peirce*, Volume 2 (1893–1913), Bloomington: Indiana University Press, 1998
7. Johnson-Laird, P.: *Mental Models*, Cambridge, MA: Harvard University Press, 1983
8. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* **44** (1990) 167–207
9. Lehmann, D., Magidor, M.: What does a conditional knowledge base entail? *Artificial Intelligence* **55** (1992) 1–60
10. Rodrigues, C.T.: The method of scientific discovery in Peirce's philosophy: Deduction, induction, and abduction. *Logica Universalis* **5** (2011) 127–164
11. Shoham, Y.: *Reasoning about change: Time and causation from the standpoint of artificial intelligence*, Cambridge, MA: The MIT Press, 1988
12. van Benthem, J.: Logic and reasoning: Do the facts matter? *Studia Logica* **88** (2008) 67–84
13. Zhang, D. et al.: Nonmonotonic reasoning and multiple belief revision. *Proceedings IJCAI 97*, Volume 1, Los Altos, CA: Morgan Kaufmann, 1997, 95–100