Completeness Theorems, Representation Theorems: What's the Difference?

David Makinson

Dept. of Philosophy, Logic & Scientific Method, London School of Economics,

david.makinson@googlemail.com

Abstract: Most areas of logic can be approached either semantically or syntactically. Typically, the approaches are linked through a completeness or representation theorem. The two kinds of theorem serve a similar purpose, yet there also seems to be some residual distinction between them. In what respects do they differ, and how important are the differences? Can we have one without the other? We discuss these questions, with examples from a variety of different logical systems.

1. Introduction: Syntax versus Semantics

Usually, the first serious course that a student takes in logic will introduce classical propositional and predicate calculi. The class learns that there are two ways of approaching such systems: semantic (alias model theoretic) or syntactic (alias axiomatic, postulational). Typically, the two are made to work like chopsticks. The teacher takes one of the two presentations as a firm base. The other is then introduced and the two are linked by means of a completeness theorem that establishes their equivalence. This theorem is often the culminating point of the course.

The decision which of the two presentations to treat as basic is very much a matter of personal preference, influenced by philosophical perspectives and pedagogical experience. In the case of classical propositional logic it is customary to begin semantically with the definition of a tautology, and then show how this coincides with an approach in terms of axioms (or axiom schemes) and derivation rules. On the other hand, in the case of intuitionistic propositional logic, it is more common to proceed in the reverse direction, first indicating how one might question certain of the axioms of the classical system, then forming a reduced axiom set, and finally showing how the resulting set of derivable formulae may be characterized semantically, say in terms of relational model structures or a suitable family of algebras. Of course, there are also maverick authors who do the reverse in each case.

As time goes on, the student also learns that the distinction between 'semantic' and 'syntactic' is not set in stone. There are presentations such as that of semantic decomposition trees (alias semantic tableaux) which can be seen as somewhere between the two. Notwithstanding their semantic name, there is something syntactic about these trees, and indeed it is possible to map the account into Gentzen sequent systems on the one hand – a flagship of the syntactic fleet – and into truth-tables on the other. Moreover, when taught completeness proofs, one learns that even such a paradigmatically semantic object as a classical valuation can be identified with a syntactic item, namely a set of formulae that is well-behaved with respect to each of

Printed from: *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz.* Ed. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson, 2007. www.fil.lu.se/hommageawlodek

the propositional connectives, and also learns that the set of tautologies coincides with the intersection of all such sets.

Most logicians have come to accept that the distinction between semantic and syntactic approaches is one of perspective and convenience more than of metaphysics. Roughly speaking, a semantic approach is seen as considering the relationship of formulae of the system to objects 'outside' it (e.g. the two classical truth-values or, in the first-order context, elements of a domain of discourse and relations over the domain), while syntactic accounts consider formulae 'on their own terms', working with e.g. sets, sequences, rules, and trees made up of them, without appealing to external items like truth-values. But this is understood as an intuitive guide rather than a rigorous separation, and items of one kind are also used as mathematically indistinguishable substitutes for those of the other.

More important than what the objects of the two kinds of the approach *are*, is the manner in which they are *used*. Thus, a syntactic approach will usually put forward inductive definitions of its key notions such as the set of all logical truths, while a semantic approach will usually define the same notions as the intersection of a family of inductively defined sets (those formulae true under each valuation).

What about the distinction between completeness and representation theorems? Here too it would be wrong to proceed doctrinally. It is more a matter of articulating how these notions have come to be used by logicians and mathematicians and isolating typical differences. Nor need the exercise end up with a formal definitions. Contrary to a widely held view, the search for a formal definition corresponding in all points with current usage can be an endless pursuit with diminishing returns. A well known example is the attempt by epistemologists to construct a formal account of what constitutes knowledge, as contrasted with belief. What we need is clarification of essentials; rough insight can be more useful than tortuous formal definition.

We will begin by considering well-known examples of completeness theorems in classical and related logics, and compare them with associated representation theorems. This leads to an approximate articulation of differences. We will then go on to look at some less widely known examples where the contrast presents itself in a rather different light. We take these examples from the theories of 'logical friendliness' and of uncertain inference.

The discussion will not lead us to cut-and-dried definitions of either completeness or representation theorems. But it will provide us with some elementary distinctions, basic insights, and an appreciation of some of the different things that can be going on under these two names.

2. Three Perspectives: Logical Equivalence, Logical Truth, Logical Consequence

Monuments of nature or of man such as Mount Fuji or the Sydney Opera House can be looked at from many directions. The engraver Hokusai immortalized a hundred views of Mount Fuji, and photographers have seized a thousand facets of the Opera House. Logic too can be looked at from many angles; over the last century and a half there have been changes in the perspective chosen.

From 1847, the date of publication of George Boole's pioneering *The Mathematical Analysis of Logic*, we can distinguish between three epochs, focussing attention in three places. For Boole himself and many of his successors in the second half of the nineteenth century, centre court is occupied by *equations*. If we translate from an algebraic context to a more conventionally logical one, we can say that these logicians gave prime attention to a relation between propositional formulae: *the relation of classical equivalence*.

In 1879, with the publication of Frege's *Begriffsschrift*, this perspective changed radically. The relation of equivalence became secondary: cameras were directed instead at a distinguished set of formulae: *the set of all logically true formulae*. This perspective became more and more popular, flourished through much of the twentieth century, and is still standard in most texts of logic for students of mathematics.

But already in the 1920s and 1930s a third perspective was being developed by Tarski and Gentzen. They returned to a relation between formulae, but a different one: *the relation of classical consequence*. This angle was further developed by the so-called Polish school of logic after the Second World War. For many logicians today (including the present author) it provides the most convenient thread to follow.

Of course, for classical logic and some of its neighbours, all three of these notions – the relation of logical equivalence, the set of logical truths, and the relation of logical consequence – are easily inter-definable. But there are other contexts in which the account in terms of a distinguished set of formulae is unable to register the subtleties of either of the two relational ones, and indeed it can happen in some cases that the account in terms of equivalence cannot reconstitute the asymmetries of consequence. From the point of view of universal logic, the approach in terms of consequence appears to be the most fine-grained. In the present discussion of completeness versus representation theorems, we will usually formulate our remarks in terms of it.

3. The Classical Case: Representation for Boolean Algebras, Completeness for Classical Logic

In the theory of Boolean algebras, the representation theorem can take several forms. One tells us that every Boolean algebra is isomorphic to a field of sets; another tells us that every Boolean algebra is isomorphic to a subalgebra of a direct product of copies of the two-element Boolean algebra. These theorems date from the work of Stone, Birkhoff and others in the 1930s.

In classical propositional logic the most common formulation of the completeness theorem tells us that every tautology (i.e. formula true under all suitably well-behaved valuations into the set of the two truth-values) is derivable in an appropriate axiom system (i.e. may be obtained by repeated applications of chosen derivation rules to selected formulae serving as axioms). This result also dates back to the early twentieth century. When stated in terms of consequence, completeness likewise tells us that for any propositional formulae a,x, if a tautologically implies x then the pair (a,x) may be obtained from chosen pairs serving as axioms by repeated applications of chosen derivation rules.

It has for long been recognized that these are saying similar things, one in the language of the algebraist and the other in the language of the logician. Indeed, certain differences are merely incidental, without significance. Boolean algebras are usually defined as structures satisfying certain equations, while the above formulations of the completeness theorem use a distinguished set of formulae, or the non-symmetric relation of consequence over them. But it is perfectly possible to define Boolean algebras as algebraic structures with a unit element satisfying certain conditions, or in terms a partial ordering corresponding to logical consequence. Likewise, as we have already mentioned, it is possible to formulate classical logic and in particular its completeness theorem in terms of an equivalence relation, corresponding to the equations of Boolean algebras. So that is not a substantive difference.

One of the first to study systematically the interconnection between completeness and representation was Helena Rasiowa in her celebrated volume of 1963, *The Mathematics of Metamathematics*, co-authored with Roman Sikorski. Another was Paul Halmos in the papers collected in his *Algebraic Logic* of 1962 (see also the reminiscences in his autobiography of 1985). As they independently pointed out, the formulae of classical logic can be seen as forming a Boolean algebra under a suitably defined equivalence relation, and this turns out to be the free Boolean algebra on a countable set of generators. The logician's valuations are in effect homomorphisms from the free Boolean algebra into the two-element one, and the completeness theorem thus comes out as an immediate consequence of the second of the two representation theorems mentioned above. In this way, central parts of classical propositional logic may be seen as fragments of a more comprehensive theory of Boolean algebras.

From this example, one may hazard the following rough portraits of completeness and representation theorems.

- A completeness theorem for a formal language states that a semantically defined set of expressions (or relation between them) of a formal language is included in one that is presented syntactically. Typically, the latter set is defined inductively, as the closure of an explicitly given list ('axioms') under explicitly given Horn rules ('derivation rules'). In general, such a theorem is of interest only if we already have the converse inclusion (soundness), which usually is proven by a straightforward induction that rides on the back of the definition of the syntactically presented set.

- A representation theorem for a class of mathematical structures (e.g. algebras) states that every structure in that class is isomorphic to some structure in a distinguished proper subclass. Typically, the subclass of structures is in some sense more 'concrete' than the class as a whole, as are fields of sets compared to Boolean algebras in general (or even more saliently, the groups of transformations that figure in representation theorems of group theory).

Several contrasts emerge from these portraits. In the first place, a completeness theorem relates a language to a structure or family of structures, while a representation theorem relates a family of structures to one of its proper subfamilies. For this reason, a completeness theorem belongs to logic, while a representation theorem belongs to mathematics, even though it can have direct implications for logic.

Moreover, representation theorems appear to be more powerful, in general, than the corresponding completeness theorems. In the classical context, at least, the latter may be obtained as a corollary of the former, but there is no visible way of proceeding in the reverse direction.

Finally, in a representation theorem the definition of the distinguished subclass of structures is free to take a wider variety of forms than is customary for the notion of derivability. While the latter is usually required to be inductive, and the induction is often expressed using Horn conditions, there is no such constraint on the definition of the distinguished subclass of structures in a representation theorem.

This picture holds up well for a number of other examples. Essentially the same pattern emerges in modal logic, irrespective of whether, on the semantic side, we are looking at Boolean algebras with operators, topological structures, or relational frames in the style of Kripke. It also shows itself in some well-known subsystems of classical logic, such as intuitionistic logic (again irrespective of whether the semantics is algebraic or relational) and for various systems of paraconsistent (alias relevance) logics. In this way, something along the lines of this picture has become part of the folklore.

But there are also contexts where parts of the picture do not fit so well. In the next section we recall an example of a completeness theorem where the definition of derivability is quite unusual in form, though still a Horn rule, and where no underlying representation theorem appears to be available. Then we give an example of a representation theorem that links formulae with structures and so is already situated on the logical rather than the purely mathematical level. Moreover, in that example the representation theorem is significant and difficult to prove while a directly corresponding completeness theorem is trivial and uninteresting.

4. Friendliness: Completeness without Visible Representation

The concept of friendliness was introduced in Makinson (2005a) with a more detailed account following in Makinson (2005b). It is a relation between formulae of classical logic that generalizes, in a natural way, the standard notion of classical consequence.

Recall the definition of classical consequence in propositional logic. Let A be any set of formulae, and x any individual formula. Then x is said to be a classical consequence of A, written $A \models x$, iff for every valuation v on all letters of the language, if $v(A) = 1$ (shorthand for $v(a) = 1$ for all $a \in A$) then $v(x) = 1$.

Trivially, the only elementary letters (alias propositional variables) that count here are those occurring in A or in x. Write $E(A)$ for the set of elementary letters that occur in A; likewise $E(x)$ for those occurring in x, and $E(A,x)$ for those occurring in $A \cup \{x\}$. Then the definition of classical consequence may be rephrased as follows: $A \models x$ iff for every partial valuation v on $E(A)$, if $v(A) = 1$ then $v^+(x) = 1$ for every extension v^+ to $E(A,x)$.

Expressed in this last way, classical consequence is a $\forall\forall$ concept. Friendliness is just the corresponding $\forall\exists$ one. We say that A is *friendly* to x and write $A \mid\approx x$ iff every partial valuation v on $E(A)$ with $v(A) = 1$ may be extended to a partial valuation v^+ on $E(A,x)$ with $v^+(x) = 1$.

Evidently, this definition is just as semantic as is that of classical consequence itself. Can it be given a syntactic characterization?

A little reflection indicates that if it can, the syntactic conditions will have to be rather different from those to which we are accustomed. Unlike its ∀∀ counterpart, the relation of friendliness is not closed under uniform substitution of arbitrary formulae for elementary letters (briefly: substitution). For a trivial counter-example observe that $p \mid\approx p \wedge q$ where p,q are distinct elementary letters, while $p \mid\approx\!/\; p \wedge \neg\, p$ obtained by substituting p for q. Consequently, the relation cannot be characterized by taking an explicit set of expressions that is closed under substitution, and closing it under rules that are also closed under substitution, as we do in classical logic and many of its neighbours.

However, it turns out that we can do the job in the following way. On the one hand, friendliness satisfies the following three rules:

(1) Whenever $A \models x$ then $A \mid\approx x$
(2) Whenever $A \cup \{b\} \mid\approx x$ and $A \cup \{\neg b\} \mid\approx x$ then $A \mid\approx x$
(3) Whenever $A \models\!/\; \neg x$ and for each elementary letter $p \in E(A)$, either $A \models p$ or $A \models \neg p$, then $A \mid\approx x$.

Conversely, friendliness is the least relation |≈ over classical formulae that satisfies these three rules. Despite the negative antecedent of the last condition, such a least relation exists, and is the intersection of all relations satisfying the three conditions.

Of the three conditions, the second is a typical Horn rule, closed under substitution and paradigmatically syntactic. The first can be regarded as semantic or syntactic as we like, according as we give classical consequence a semantic or syntactic reading.

The third rule is the most interesting. As far as the relation |≈ is concerned, it is a Horn rule: its negative antecedent concerns only the relation of classical consequence, not friendliness (and so better called a 'side-condition' rather than an antecedent or premise). But it has an internal complexity that is not customary: its formulation uses (in the metalanguage) negation, universal quantification and disjunction). Moreover, the set of all ordered pairs (A,x) such that the rule guarantees that $A \mid\approx x$, is not closed under substitution. Finally, the rule is computationally ghastly, and if a similar rule is used in a first-order version of friendliness, it will not even be semi-decidable.

Can we call this characterization of friendliness a completeness theorem? In the author's view, that would be a natural and legitimate way of speaking. To be sure, one might hope for a simpler third rule, and perhaps one can be found. We would then have two completeness theorems, one in some sense better than the other.

Is there any interesting representation theorem for underlying this completeness theorem? The main candidates appear to be the standard representation theorems for

Boolean algebras, mentioned above, but it is difficult to see any natural way of obtaining the completeness theorem for friendliness from them.

Thus, this example departs from the standard picture in two ways: we have a completeness theorem but apparently no underlying representation theorem, and one of the rules generating the syntactic notion of derivability is more unwieldy than is customary.

5. A Celebrated Representation Theorem with a Trivial Completeness Counterpart

There are also contexts where one can prove an important and non-trivial representation theorem, but where the corresponding completeness theorem is quite trivial and degenerate.

This kind of situation arises for several kinds of logic that have been studied in the last quarter century, notably the logic of belief change in the style of AGM (Alchourrón, Gärdenfors, Makinson 1985), and logics of uncertain inference, whether qualitative (alias nonmonotonic logic) or quantitative (probabilistic consequence relations). We will illustrate the phenomenon with one of the best-known examples – the qualitative logic of uncertain inference formulated in terms of preferential models in the manner of KLM (Kraus, Lehmann and Magidor 1990).

Syntactically, we are looking at relations |~ between formulae of classical propositional logic that satisfy a certain small collection of Horn rules, i.e. rules of the form: *whenever* a_1 |~ x_1 *and…* a_n |~ x_n *then* b |~ y. Here, $n \geq 0$ and the n antecedents of the rule are possibly supplemented by side-conditions involving classical consequence but not mentioning |~.

To be precise, KLM consider the following family of Horn rules. The second and third rules use side-conditions; in the third one =||= means classical equivalence.

a |~ *a*	(Reflexivity)
whenever *a* |~ *x* and *x* |= *y*, then *a* |~ *y*	(Right Weakening)
whenever *a* |~ *x* and *a* =||= *b*, then *b* |~ *x*	(Left Classical Equivalence)
whenever *a* |~ *x* and *a* |~ *y*, then $a \wedge x$ |~ *y*	(Cautious Monotony)
whenever *a* |~ *x* and *b* |~ *x*, then $a \vee b$ |~ *x*	(Disjunction in the Premises)
whenever *a* |~ *x* and *a* |~ *y*, then *a* |~ $x \wedge y$	(Conjunction in Conclusion).

These rules do not define a unique consequence relation. Rather, they define an infinite family of such relations, all of which are supraclassical in the sense that (considered as relations, i.e. as sets of ordered pairs) they include the relation of classical consequence. The KLM representation theorem tells us: A relation |~ between formulae of classical propositional logic satisfies the KLM postulates iff it is generated from some preferential structure by means of Shoham's minimality rule.

Here, a *preferential structure* is a triple $(S,<,\lambda)$, where S is any non-empty set (heuristically, of 'states' or 'worlds'), < is any relation over S (typically, at least irreflexive and transitive, but these constraints are not necessary for the theorem), and λ is a 'labelling function', which associates with each state $s \in S$ a classical valuation

$\lambda(s)$, also written v_s, on formulae into the two-element set {0,1}. To ensure satisfaction of Cautious Monotony, it is assumed as part of the definition that the structure is stoppered (alias smooth) in the sense that for every formula *a*, if $v_s(a) = 1$ then either *s* is itself minimal among the states *t* with $v_t(a) = 1$, or there is a state $s' < s$ that is minimal among those states.

When $(S,<,\lambda)$ is a preferential structure then it generates a consequence relation |~ between formulae of classical propositional logic by Shoham's minimality rule: *a* |~ *x* iff $v_s(x) = 1$ for every $s \in S$ such that *s* is minimal among the states *t* with $v_t(a) = 1$. Such a relation is called a *preferential consequence relation*.

It is straightforward to verify that every preferential consequence relation satisfies all of the KLM rules. The representation theorem, we recall, tells us the converse: every relation between classical formulae satisfying all of the KLM rules is a preferential consequence relation. This result is justly celebrated: it is a significant fact with a far from trivial proof.

Interestingly, it already goes a little beyond the kind of representation theorem that we saw for classical logic. There is no talk of isomorphism here. The relation |~ is *itself* a preferential consequence relation; in other words, it can itself be generated from some preferential structure using Shoham’s rule.

Also of interest is that the representation theorem does not belong to some underlying algebraic or purely model-theoretic level, as in the classical and modal cases. It is about relations |~ between classical formulae, and is thus very much a result of logic itself.

To be sure, Karl Schlechta (2004) has subsequently shown that we can dig deeper, and see the KLM representation theorem as a reflection of another one that functions on a purely mathematical level, to the effect that every structure of a certain kind (defined in terms of selection functions) may be generated from a structure of another kind (defined in terms of relations). We thus have two representation theorems, one deeper than the other. But it remains standard to refer to the KLM result itself as a representation theorem, despite its explicit concern with expressions of a formal language.

Is there a completeness theorem corresponding to this representation theorem? In particular, can we show that an expression *a* |~ *x* is derivable from the KLM postulates iff it holds in every preferential model?

The answer to this last question is positive, but trivially true and of no interest. For classical consequence is itself a preferential consequence relation (take < to be the empty relation in the definition of a preferential model), and by the first of the KLM postulates it is the least one. So *a* |~ *x* holds in every preferential model iff *a* |= *x*. Likewise, *a* |~ *x* is derivable from the KLM postulates iff *a* |= *x*.

What is happening? Essentially, it appears to be the following. The representation theorem quantifies over relations |~. It tells us that for every such relation, it satisfies the KLM postulates iff it is generated by some preferential model. On the other hand, the completeness theorem distributes the quantifier over relations to each side of the equivalence after adding an initial quantification over pairs of formulae. It states that for any pair (a,x) of formulae, (a,x) is in every relation satisfying the KLM postulates iff it is in every relation that is generated by some preferential model.

Similarly in the classical case. The representation theorem for Boolean algebras says that a structure is a Boolean algebra iff it is isomorphic to a subalgebra of a direct product of copies of the two-element Boolean algebra. From this it follows that an arbitrary equation is valid in *all* Boolean algebras iff it is valid in *all* such subdirect products. And the latter holds iff it is valid in the two-element algebra itself – which is just the classical completeness theorem. Here again we are distributing the quantifier through the equivalence and adding an initial quantification, this time over equations.

Distributing a universal quantification to each side of an equivalence evidently weakens it: $\forall x(\varphi)\leftrightarrow\forall x(\psi)$ is weaker than $\forall x(\varphi\leftrightarrow\psi)$. This is why completeness theorems are generally weaker than their representation counterparts. And in some cases, such as that of preferential consequence, the distribution of the quantifier washes out everything of interest.

Such loss in the wash also occurs in the context of probabilistic consequence relations (see Hawthorne and Makinson, to appear) as well as for some other kinds of qualitative consequence relations for uncertain reasoning. It likewise takes place for belief revision under the AGM paradigm, as can be seen by applying the translation of Gärdenfors and Makinson (1991). In each of these cases, and for the same underlying reason, we have a trivial completeness theorem in which both left and right collapse into classical consequence.

6. Completeness of the KLM Postulates over the Domain of Horn Rules

However, it would be misleading to leave the question of completeness for preferential consequence with only this negative observation. For, although completeness for expressions of the form a |~ x is trivial, we can also formulate another completeness result, this time more significant. It concerns a wider class of expressions, namely Horn rules. These, we recall, are rules of the form: *whenever* a_1 |~ x_1 *and…* a_n |~ x_n *then* b |~ y, where $n \geq 0$, and the n antecedents of the Horn rule are possibly accompanied by side-conditions involving classical consequence but not |~.

First, we look at the syntactic side. Consider any such Horn rule *whenever* a_1 |~ x_1 *and …* a_n |~ x_n *then* b |~ y, and any set H of Horn rules (each possibly with side conditions involving classical consequence only). We say that the former is *derivable* from the latter iff, roughly speaking, its consequent can be obtained from its antecedents and the rules in H by iterated detachments. To be precise, iff there is a finite tree whose root is labelled with b |~ y, each of whose leaves is labelled with one of a_1 |~ x_1,…, a_n |~ x_n (or with a fact about classical consequence, if some of the rules in H make use of side-conditions) and such that each non-leaf is labelled by (an instance of) the consequent of one of the Horn rules in H and has its parents labelled by the (corresponding instance of) the antecedents (or side conditions) of the same Horn rule.

This definition is rather tedious when written out in full, but an easy result of universal logic (and of logic programming) tells us that it is equivalent to the following more succinct one. The Horn rule *whenever* a_1 |~ x_1 *and…* a_n |~ x_n *then b* |~ *y* is derivable from the set *H* of Horn rules iff it is satisfied by the least relation (i.e. the intersection of all relations) satisfying all rules in *H*.

When formulated over the domain of Horn rules, the completeness theorem for preferential consequence can be stated as follows: A Horn rule *whenever* a_1 |~ x_1 *and…* a_n |~ x_n *then b* |~ *y* is derivable from the KLM postulates whenever it holds in all preferential models. This result follows immediately from the KLM representation theorem. Since statements *a* |~ *x* are themselves Horn rules (with $n = 0$) this completeness theorem covers the previous one as a special case. But whereas the special case is trivial (left and right hand sides collapsing into classical consequence) the more general version is not.

One lesson that we can learn from this example is that whenever we formulate a completeness theorem, we should always be careful to specify the set of expressions to which it applies. Variation in the set of expressions envisaged can have major consequences.

7. Representation for Pivotal-Valuation Consequence

We take this opportunity to present a solution to a representation problem that was left open in the author's *Bridges from Classical to Nonmonotonic Logic*. It concerns a family of consequence relations that are supraclassical, but still monotonic; the family may be seen as a bridge between classical consequence and the preferential consequence relations of Kraus, Lehmann and Magidor.

These consequence relations are defined semantically in a very simple manner; the problem is to give them a syntactic characterization. Let *V* be the set of all classical valuations on propositional formulae into {0,1}, and let *W* be any subset of *V*. For any set *A* of formulae and individual formua *x*, we say that *x* is a *pivotal-valuation consequence* of *A* modulo *W* and write $A \models_W x$, iff $v(x) = 1$ for every valuation $v \in W$ with $v(A) = 1$.

Classical consequence is evidently a pivotal-valuation consequence (the case $W = V$), and the total relation is also one (case $W = \emptyset$). The problem is: What syntactic conditions characterize the family of all pivotal-valuation consequence relations $\models_W$ for $W \subseteq V$?

For the more restricted family of pivotal-valuation consequence relations $\models_W$ where *W* is a definable subset of *V* (i.e. there is a set *F* of formulae such that *W* is the set of all valuations satisfying *F*), the answer is straightforward: these relations are just the compact supraclassical closure relations satisfying the rule of Disjunction in the Premises (alias OR). This result appears to have been part of the folklore for some time, but the first formal statement and proof that the author knows of is Rott (2001) section 4.4 observation 5; there is a direct and simple proof in Makinson (2005c) chapter 2.

But when we drop the assumption that *W* is definable, compactness can fail. The remaining conditions (supraclasical closure relation satisfying OR) continue to hold, but they are not sufficient to ensure representation as pivotal-valuation consequence relations - an example establishing this was given in Schlechta (1992). We may therefore ask: Is there a condition that we can add to that of being a supraclassical closure relation satisfying Disjunction in the Premises, that does the job?

There is indeed one, though it is far from being a Horn condition, and one might even hesitate to describe it as fully syntactic. It was first formulated by Makinson (1994) as part of an analysis of Schlechta's example mentioned above. We call it the *Capping condition*. It requires that whenever $A \mid\!\not\sim x$ then there is a maxiconsistent set $A^+ \supseteq A$ with $A^+ = \{y : A^+ \mid\!\sim y\}$ and $x \notin A^+$. We can express it more elegantly in the language of consequence operations. Writing $C(A)$ for $\{y : A \mid\!\sim y\}$, it requires that whenever $x \notin C(A)$ then there is a maxiconsistent set $A^+ \supseteq A$ with $x \notin A^+ = C(A^+)$.

That this Capping condition provides a representation theorem was in effect shown by Ben-Naim (2005, also 2006 Proposition 55), in a rather roundabout way. We give a direct proof, using the language of operations.

Theorem. The pivotal-valuation consequence operations are just the supraclassical closure operations that satisfy the rule of Disjunction in the Premises and also the Capping condition.

Proof. Left-to-right: Let W be any set of valuations, with Cn_W the pivotal consequence operation that it determines. We need to show that it satisfies the listed conditions, of which the interesting one is Capping. Suppose $x \notin Cn_W(A)$. Then by the definition of pivotal-valuation consequence there is a valuation $w \in W$ with $w(A) = 1$, $w(x) = 0$. Put A^+ to be the characteristic set of w, i.e. $A^+ = \{y : w(y) = 1\}$. We claim that it has the desired properties. Clearly A^+ is a maxiconsistent set, and since $w(A) = 1$, $w(x) = 0$, we have $A^+ \supseteq A$, $x \notin A^+$. To show that $A^+ = Cn_W(A^+)$, suppose $z \notin A^+$; we need to show that $z \notin Cn_W(A^+)$. Since $A^+ = \{y : w(y) = 1\}$ we have $w(z) = 0$ while $w(A^+) = 1$, so since $w \in W$ we have $z \notin Cn_W(A^+)$ as desired.

For the converse, let C be any supraclassical closure relation satisfying OR and the Capping condition. Put W to be the family of all valuations w that are characteristic functions of maxiconsistent sets X such that $X = C(X)$. We claim that $C = Cn_W$, i.e. $C(B) = Cn_W(B)$ for every set of B of formulae.

To show $C(B) \subseteq Cn_W(B)$, suppose that $z \in C(B)$. Let w be any characteristic function of a maxiconsistent set X such that $X = C(X)$, and suppose $w(B) = 1$. We need to show that $w(z) = 1$. Since $w(B) = 1$ and X is the characteristic set of w, $B \subseteq X$ so since C is by hypothesis a closure operation and so monotonic, $C(B) \subseteq C(X) = X$, so since $z \in C(B)$ we have $z \in X$ so $w(z) = 1$ as desired.

To show $Cn_W(B) \subseteq C(B)$, suppose that $z \notin C(B)$; we need to show that $z \notin Cn_W(B)$, i.e. we need to find a $w \in W$ with $w(B) = 1$ but $w(z) = 0$. By the definition of W, it suffices to find a maxiconsistent set X with $B \subseteq X = C(X)$ and $z \notin X$. But this given by the hypothesis that C satisfies the Capping condition, and the proof is complete. □

References

Alchourrón, Carlos, Peter Gärdenfors, and David Makinson 1985, "On the logic of theory change: partial meet contraction and revision functions" *The Journal of Symbolic Logic* 50: 510-530.

Ben-Naim, Jonathan 2005, “Pivotal and pivotal-discriminative consequence relations” *J. Logic and Computation* 15: 679-700.

Ben-Naim, Jonathan 2006, “Autour de la caractérisation de raisonnements de sens commun en présence d’informations incertaines”. Thèse, Université d’Aix-Marseille, 28 avril 2006.

Boole, George 1847, *The Mathematical Analysis of Logic, Being an Essay Toward a Calculus of Deductive Reasoning* (Cambridge: Macmillan).

Frege, Gottlob 1879, *Begriffsschrift* (Halle: Nebert).

Halmos, Paul 1962, *Algebraic Logic* (New York: Chelsea).

Halmos, Paul 1985, *I Want to be a Mathematician: an Automathography* (New York: Sporinger-Verlag).

Hawthorne, James and David Makinson (*Studia Logica,* to appear), “The quantitative/qualitative watershed for rules of uncertain inference”.

Kraus, Sarit, Daniel Lehmann and Menachem Magidor 1990, “Nonmonotonic reasoning, preferential models and cumulative logics” *Artificial Intelligence* 44: 167-207.

Makinson, David 1994, “General Patterns in Nonmonotonic Reasoning”, in *Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3*, ed. Gabbay, Hogger and Robinson (Oxford University Press), pages 35-110.

Makinson, David 2005a, “Logical friendliness and sympathy” in *Logica Universalis*, ed. J.-Y. Beziau (Basel: Birkhauser Verlag), pages 191-205.

Makinson, David 2005b, “Friendliness for logicians”, in *We Will Show Them! Essays in Honour of Dov Gabbay, vol 2,* ed S. Artemov, H. Barringer, A. Garcez, L. Lamb, J. Woods (London: King’s College Publications), pages 259-292.

Makinson, David 2005c, *Bridges from Classical to Nonmonotonic Logic* (London: King’s College Publications. Series: Texts in Computing, vol 5).

Makinson, David and Peter Gärdenfors 1991, "Relations between the logic of theory change and nonmonotonic logic", in *The Logic of Theory Change* eds A. Fuhrmann and M. Morreau (Berlin: Springer) pages 185-205.

Rasiowa, Helena and Roman Sikorski 1963, *The Mathematics of Metamathematics*. (Warsaw: Polska Akademia Nauk, Monografie Matematyczne).

Rott, Hans 2001, *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning* (Oxford: Clarendon Press, Oxford Logic Guides n ° 42).

Schlechta, Karl 1992, "Some results on classical preferential models" *Journal of Logic and Computation* 2: 676-686.

Schlechta, Karl 2004, *Coherent Systems* (Amsterdam: Elsevier).

Dept. of Philosophy, Logic & Scientific Method, London School of Economics,
Houghton Street, London WC2A 2AE, United Kingdom
david.makinson@googlemail.com

12 September 2006