Book review *To appear in the Bulletin of Symbolic Logic*

Dov M. Gabbay, C.J. Hogger, & J.A. Robinson (eds), *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 3: *"Nonmonotonic Reasoning and Uncertain Reasoning"*, Oxford University Press, Oxford 1994, xix–529 pp.

The emergence, over the last twenty years or so, of so-called "non-monotonic" logics represents one of the most significant developments both in logic and artificial intelligence. These logics were devised in order to represent *defeasible reasoning*, i.e., that kind of inference in which reasoners draw conclusions tentatively, reserving the right to retract them in the light of further evidence.

Formally, this can be captured by considering a relation $\mid\!\sim$ of logical consequence holding between sets Γ of sentences (of a given language) and sentences φ (of that language). Statements of the form $\Gamma \mid\!\sim \varphi$ are taken to express the fact that φ is warranted on the basis premises in Γ. The relation $\mid\!\sim$ is *monotonic* (in its first argument) if $\Gamma \mid\!\sim \varphi$ always implies $\Delta \mid\!\sim \varphi$, whenever Γ is a subset of Δ.

The relation $\models$ of logical consequence of standard, first-order logic (and of many other logics as well) is monotonic in this sense. Thus, it appears that defeasible reasoning cannot quite be formalized in any logic having a monotonic consequence relation: once a conclusion φ is warranted on the basis of a given set Γ of premises, there is no "retracting" φ, no matter how many additional premises we put in Γ, i.e., no matter how much "further evidence" we can gather. Hence the need to develop alternative frameworks in which such inferences can be adequately formalized.

Of course, it is quite a tall order to supply a relation of logical consequence that is in some sense alternative to the one of classical first-order logic. One has to measure oneself with the stunning success of first-order logic and many of its extensions and variations, beginning with the work of Frege and Dedekind all the way up to the most rarefied reaches of model theory and proof theory.

Monotonicity is, in a sense, the birthmark of first-order logic. Beginning with Frege's *Begriffsschrift*, first-order logic was developed for the purpose of formalizing mathematical reasoning, and of course there is no retracting conclusions reached in the course of such an enterprise. But since the very beginning there was also another streak in the history of modern symbolic logic, which has to do with the many instances in which symbolic logic is used to formalize everyday reasoning, and indeed to a large extent, successfully so. Therefore, the emergence of non-monotonic logics over the course of the last two decades can also be viewed as the surfacing of a trend that has been there all along in the development of the field.

Pioneering work in the field of non-monotonic logics was carried out (among others) by J. McCarthy, D. McDermott & J. Doyle, and R. Reiter. The realization (which was hardly new) that ordinary first-order logic was inadequate to represent defeasible reasoning was for the first time accompanied by several proposals of formal frameworks within which one could at least begin to talk about defeasible inferences in a precise way, with the long-term goal of providing for defeasible reasoning an account that could at least approximate the degree of success of modern symbolic logic in the formalization of mathematical reasoning.

The publication of a monographic issue of the *Artificial Intelligence Journal* in 1980 can be regarded as the "coming of age" of defeasible formalisms.

The excellent volume edited by Gabbay, Hogger, and Robinson can be viewed as an updated version of that seminal 1980 issue of *Artificial Intelligence*, and will be a standard reference for years to come. The book also serves to fill a lacuna in the widely popular and influential *Handbook of Philosophical Logic*, also edited by D. Gabbay together with F. Guenthner for the Kluwer publishing house. In the planning stages of the *Handbook of Philosophical Logic*, a decision was made to leave non-monotonic reasoning out of the third volume (to be titled "Alternatives to Classical Logic" and eventually published in 1986), probably because the field did not appear to be mature enough for such an inclusion. Obviously, things are quite different now.

The volume edited by Gabbay, Hogger, & Robinson comprises the following contributions:

1. Matthew Ginsberg, *AI and Nonmonotonic Reasoning*, pp. 1–33;
2. David Makinson, *General Patterns in Nonmonotonic Reasoning*, pp. 35–110;
3. John F. Horty, *Some Direct Theories of Nonmonotonic Inheritance*, pp. 111–187;
4. David Poole, *Default Logic*, pp. 189–215;
5. Kurt Konolige, *Autoepistemic Logic*, pp. 217–295;
6. Vladimir Lifschitz, *Circumscription*, pp. 297–352;
7. Donald Nute, *Defeasible Logic*, pp. 353–395;
8. Henry E. Kyburg, Jr, *Uncertainty Logics*, pp. 397–438;
9. D. Dubois, J. Lang, and H. Prade, *Possibilistic Logic*, pp. 439–513.

The book also contains an informative preface by the main editor, Dov Gabbay, and an extensive index.

The first contribution, by Matthew Ginsberg, serves also as a general introduction to range of phenomena to be modeled, collectively referred to as defeasible reasoning. One such example is *non-monotonic inheritance*: whenever we have a body of knowledge taxonomically organized, we presuppose that subclasses inherit properties from their superclasses: dogs have lungs because they are mammals, and all mammals have lungs. However, there can be exceptions, which can interact in complex ways. Mammals, by and large, don't fly; bats are an exception, and they do fly; baby bats are exceptional bats, in that they do not fly (does that make them unexceptional mammals?). When we infer that Stellaluna, being a baby bat, does not fly, we are resolving all these potential conflicts based on a *specificity* principle: more specific information overrides more generic information.

Another example is the *closed world assumption*: you need to travel from Oshkosh to Minsk, so you consult your travel agent, who, not surprisingly, tells you that there are no direct flights. How does the travel agent know? In a sense, he doesn't: his database does not list any direct flights between Oshkosh and Minsk, and he assumes that his database is *complete*.

In both these examples, we have an attempt to *minimize* the extension of a given predicate ("exceptional-specimen" in the first case, "flight-between" in the second), and a moment's thought reveals that the such a minimization needs to take place not with respect to what the database explicitly contains but with respect to what it implies.

Ginsberg goes on to review the main approaches to the problem of formalizing such inferences, indicating what the main issues are (how to handle conflicting default inferences) and how the so-called "quantitative" or "probabilistic" approaches differ from the "qualitative" or "symbolic" ones. It is worth mentioning here that a major difference in this respect is the fact that qualitative approaches all strive, one way or another, to achieve some form of *cut*: if φ can be defeasibly inferred from $\Gamma \cup \{\psi\}$, and ψ can be defeasibly inferred from Γ, then φ can also be defeasibly inferred from Γ. The import of cut, as is well known from proof theory, is that the length of an inference does not matter for the cogency of the conclusion, and we are free to adjoin our conclusions to our original premises without any *increase* in inferential strength. Notably, this principle fails for probabilistic approaches, in which the probability associated with a conclusion decreases with the length of the inference.

The second contribution in the book, by David Makinson, is noteworthy because it focuses on the abstract properties of a defeasible consequence relation. We know that monotonicity will have to go, but what about the other structural properties? Going back to pioneering work by Gabbay, Makinson points out the importance of the *converse of cut*, a principle referred to as "cautious monotony" and that in classical logic is subsumed under monotony proper: if both φ and ψ can be defeasibly inferred from Γ, then φ can be defeasibly inferred from $\Gamma \cup \{\psi\}$. In other words, we can put our theorems back into our premise set with no *decrease* of inferential power. Makinson argues that this principle is crucial if we want a well-behaved relation of defeasible consequence, and proceeds to show that it fails in many non-monotonic inference formalisms.

The third contribution, by John Horty, is a survey of a variety of approaches to non-monotonic inheritance networks. Such networks are collections of nodes and directed ("is-a") links representing taxonomic information. When exceptions are allowed, the network is interpreted *defeasibly*. Although the language of non-monotonic networks is expressively limited by design, such networks represent an extremely useful setting in which to test one's intuitions and methods for handling defeasible information. Horty does a particularly good job at explaining the difference between "cautious" and "bold" approaches to defeasible reasoning (these approaches are also sometimes referred to as "skeptical" and "credulous," respectively) . This chapter in many ways brings order to a varied field in which it is not always easy to discern the relationships between the different approaches.

David Poole's chapter on Reiter's default logic surveys what is perhaps the most flexible of all the tools devised to represent defeasible inference. A *default* is a *defeasible inference rule* of the form

$$\frac{\alpha : \beta}{\gamma},$$

(for formulas α, β, γ): if α is known, and there is no evidence that β might be false, then the rule allows the inference of γ. The sentences α, β, and γ are respectively known as the prerequisite, justification, and conclusion of the default.

As is clear, application of the rule requires that a consistency condition be satisfied, and rules can interact in complex ways. In particular it is possible that application of a rule might cause the consistency condition to fail (as when β is $\neg\gamma$). Reiter's default logic uses the notion of an *extension* to make precise the idea that the consistency condition has to be met both before and after the rule is applied. Given a set Γ of defaults, an extension is, roughly (and in typical circular fashion), a maximal subset Δ of Γ whose conclusions both imply all the pre-requisites of defaults in Δ and are consistent with all the justifications of defaults in Δ.

Default logic is intimately connected with certain versions of modal logic, known as *autoepistemic logics*, which are the subject of Kurt Konolige's chapter. The idea is that we can sometimes reach conclusions about the state of the world using facts concerning our own knowledge or ignorance. For instance, I can conclude that I do not have a sister given that if I did I would probably know about it, and nothing to that effect is present in my "knowkedge base". Of course, such a conclusion is defeasible, because there is always the possibility of learning new facts.

In order to make these intuitions precise, consider a modal language in which the necessity operator L is interpreted as "it is known that". The following abbreviations are useful: given a set S of sentences, $\overline{S}$ denotes the complement of S, LS denotes $\{L\phi : \phi \in S\}$, $\neg S$ denotes $\{\neg\phi : \phi \in S\}$, and S_0 denotes the subset of S composed of those sentences containing no occurrences of L. The central notion in autoepistemic logic is that of an *extensions* of a theory S, i.e., a consistent and self-supporting sets of beliefs that can be reasonably be entertained on the basis of S. Formally, an extension of S is a theory T satisfying

$$T = \{\phi : S \cup LT_0 \cup \neg L\overline{T_0} \vdash_{K45} \phi\}.$$

In the definition, LT_0 is a set of "introspective axioms" expressing what is known; $\neg L\overline{T_0}$ is a set of "negative introspective axioms," expressing what is not known, and $\vdash_{K45}$ is derivability in the modal system $K45$ (interestingly, although extensions are closed under $S5$ consequence, replacing $K45$ by $S5$ in the definition collapses autoepistemic logic into monotonic $S5$). In general, autoepistemic logic provides a rich and interesting framework, whose mathematical properties and connections to other non-monotonic formalisms are well worth exploring. The interested reader is referred to Konolige's chapter for further details.

Vladimir Lifschitz's chapter surveys yet another non-monotonic formalism based on tools and ideas derived from classical logic. *Circumscription*, originally introduced by John McCarthy, is based on the intuition that, all other things being equal, extensions of predicates should be *minimal*. This can be connected to the usual examples from non-monotonic reasoning by considering principles such as "all normal birds fly". Here we are trying to minimize the extension of the abnormality predicate, and assume that x is normal unless we have positive information to the contrary.

Formally, this can be represented using second order logic, as follows. Given predicates P and Q, let $P \leq Q$ abbreviate $\forall x(Px \to Qx)$ and, similarly, let $P < Q$ abbreviate $P \leq Q \wedge Q \not\leq P$. If $A(P)$ is a formula containing occurrences of a predicate P, then the circumscription of P in A is the second-order sentence $A^*(P)$:

$$A(P) \wedge \neg\exists p[A(p) \wedge p < P].$$

This special case of circumscription would seem not to be sufficient in practice, for in many applications one needs to minimize the extension of P, while allowing the extension of certain other predicates $Q_1, \dots, Q_n$ to vary. In other words, what would seem to be needed is the apparently more general circumscriptive formula

$$A(P; Q_1, \dots, Q_n) \wedge \neg\exists p, q_1, \dots, q_n[A(p; q_1, \dots, q_n) \wedge p < P];$$

but as Lifschitz shows, this case can be reduced to the basic one. From the point of view of applications, circumscription has a major shortcoming, namely the absence of a complete inference procedure: a shortcoming it shares with many other non-monotonic representation formalisms, and which is due to the second-order nature of the circumscriptive formula. All the more crucial, then, is one of Lisfschitz's own contributions to the field: the identification of conditions on the formula $A(P)$ that are sufficient to guarantee that the circumscriptive formula $A^*(P)$ has a first-order equivalent. For instance, if $A(P)$ is the formula $\forall x(F(x) \rightarrow P(x))$, then $A^*(P)$ is equivalent to $\forall x(F(x) \leftrightarrow P(x))$.

Nute's chapter is a survey of several different systems of non-monotonic reasoning, from Nute's own defeasible logic and Delgrande's conditional logic, to the Geffner-Pearl theory of probabilistic entailment, and Pollock's theory of warrant (perhaps the most philosophically motivated of the lot). It is worth mentioning here that Delgrande's conditional logic builds upon classical work of David Lewis and Robert Stalnaker. A non-truth-functional binary connective $>$ is introduced, where $p > q$ represents the counterfactual implication of p and q. Several plausible sets of axioms of increasing strength are considered, and completeness is established with respect to Stalnaker's "selector function" semantics as well as Lewis's "system of spheres" semantics. The connection to non-monotonic reasoning is provided, among other things, by the fact that $>$ fails to be left-monotonic: the inference from $p > r$ to $p \wedge q > r$ is, in general, not warranted.

The last two chapters of the *Handbook* are dedicated to probabilistic approaches to non-monotonic reasoning, which belong to a quite different tradition from the other theoretical frameworks found elesewhere in the book. In particular, Henry Kyburg's contribution on uncertainty logics provides an excellent introduction to the topic, with an extensive treatment of conditional probabilities (the chapter is unfortunately rife with typos, the most potentially misleading of which is in the definition of Jeffrey conditionalization given in the last displayed formula on p. 423). So-called possibilistic logics constitute the topic of the chapter by Dubois, Lang & Prade. This is an approach derived from Zadeh's "fuzzy-logic", and the authors do a good job explaining how possibilistic logics are connected to Zadeh's original work.

So far the contents of the book. Perhaps it is not inappropriate to take this opportunity briefly to assess the present development of the field of non-monotonic logic. There seem to be three major issues that face any attempt to develop a logic which is adequate to represent defeasible reasoning. The first is a condition of material adequacy: we want a system of non-monotonic logic to capture a broad enough range of examples, and to do justice to at least the most entrenched of our intuitions on the subject. The second issues has to do with the extent to which the resulting system is "well-behaved" from a logical point of view: how easy it is to define a relation of logical consequence and what abstract properties it has. The third set of issues has to do with the complexity (logical, computational, or otherwise) of the resulting framework.

Where do we stand with respect to these three issues, after twenty or more years of investigations into non-monotonic logic? With respect to the first set of issues, the whole development of non-monotonic logic has been driven by the desire to account for a rich and well-chosen array of examples. Of course, no single framework can aspire to be universal in this respect, but in general non-monotonic logics appear to be better motivated and more materially adequate than a host of other logical frameworks.

Quite a bit of work has been carried out in recent years pointing out the extent to which non-monotonic logics have failed or succeeded to come up with a well-behaved relation of logical consequence. As David Makinson points out in the book under review, practitioners of the field have encountered mixed success. In particular, one abstract property, cautious monotony, appears at the same time to be crucial and elusive for many of the frameworks to be found in the literature so far. Part of the problem, perhaps, is that work aimed at overcoming this kind of limitations can sometimes be in tension with the requirement of material adequacy, running the risk of pushing non-monotonic logic closer and closer to its monotonic (and better behaved) counterpart.

Finally, the third set of issues appears to be the most difficult among the ones that have been singled out. Non-monotonic logics appear to be stubbornly intractable with respect to the corresponding problem for classical logic. For instance, in the case of default logic, one obvious source of complexity is the ubiquity of consistency checks, and circumscription owes its intractability to the second-order nature of the circumscriptive axiom. But there are other, often overlooked, sources of complexity that are purely combinatorial. For instance, even when the language is restricted to conjunctions of literals (making consistency check feasible), still the problem of determining whether a given default theory has an extension is highly intractable (NP-complete, to be precise, as shown by Kautz and Selman), seemingly because it requires checking all possible sequences of firings of defaults. Although some important work has been done trying to make various non-monotonic formalism more tractable, this is perhaps the problem on which progress has been slowest in coming.

G. Aldo Antonelli
Logic & Philosophy of Science
University of California
Irvine, CA 92697-5100
aldo@uci.edu