

Non-monotonic inference*

Keith Frankish

Abstract

Non-monotonic inference is inference that is defeasible: in contrast with deductive inference, the conclusions drawn may be withdrawn in the light of further information, even though all the original premises are retained. Much of our everyday reasoning is like this, and a non-monotonic approach has applications to a number of technical problems in artificial intelligence. Work on formalizing non-monotonic inference has progressed rapidly since its beginnings in the 1970s, and a number of mature theories now exist – the most important being default logic, autoepistemic logic, and circumscription.

Main text

In most logical systems inferences cannot be invalidated simply by the addition of new premises. If an inference can be drawn from a set of premises *S*, then it can also be drawn from any larger set incorporating *S*. The truth of the original premises guarantees the truth of the inferred conclusion and the addition of extra premises cannot undermine it. This property is known as *monotonicity*. (The term is a mathematical one; a monotonic sequence is one whose terms increase but never decrease, or vice versa.) *Non-monotonic* inference lacks this property. The conclusions drawn are provisional, and new information may lead to the withdrawal of a previous conclusion, even though none of the original premises is retracted.

Much of our everyday reasoning is non-monotonic. We frequently jump to conclusions on the basis of partial information, relying on rough generalizations – that people usually mean what they say, that machines usually work as they are designed to, that objects usually stay where they are put, and so on. We treat these conclusions as provisional, however, and are prepared to retract them if we learn that the cases we are dealing with are atypical. To take an example that is ubiquitous in the literature, if we know that Tweety is a bird then we may infer that Tweety flies, since we know that birds typically fly, but we shall withdraw this conclusion if we learn that Tweety is an atypical bird – a penguin, say.

An important feature of inferences like this is that they are sensitive to the absence of information as well as to its presence. Because we lack information that Tweety is atypical, we assume that he is not and proceed on that basis. (That, of course, is why the acquisition of new information can undermine the inference.) In standard logics, by contrast, inference is sensitive only to information that is explicitly represented, and we would need to add the premise that Tweety is not atypical in order to reach the conclusion that he flies. This feature of non-monotonic inference makes it highly

* This is the author's version of an article published in K. Brown (Editor-in-Chief), *Encyclopedia of Language & Linguistics, Second Edition*, volume 8, (pp. 672-75), Elsevier, 2006. This version may not reflect changes resulting from the publishing process, such as copy-editing, corrections, and formatting.

useful. We do not have the time or mental capacity to collect, evaluate, and process all the potentially relevant information before deciding what to do or think. (Think of everything we would need to know in order to be sure that Tweety is not atypical – that he is not a penguin, not an ostrich, not a hatchling, not injured, not tethered to the ground, and so on.)

Because of its central role in commonsense reasoning, non-monotonic inference has attracted much interest from researchers in artificial intelligence – in particular, from those seeking to model human intelligence in computational terms. The challenge has been to *formalize* non-monotonic inference – to describe it in terms of a precisely-defined logical system which could then be used to develop computer programs that replicate everyday reasoning.

Non-monotonic logic also has applications to more specific problems in artificial intelligence, among them the so-called *frame problem* (McCarthy and Hayes 1969; see also FRAME PROBLEM entry). In order to plan how to reach its goals, an artificial agent will need to know what will and what will not change as a result of each action it might perform. But the things that will *not* change as a result of an action will be very numerous and it would be impracticable to list them all in the system's database. A more efficient solution would be for the system to reason non-monotonically, using the rule of thumb that actions leave the world unchanged except in those respects in which they are known to alter it. Another application is in the area of database theory. Here it is often convenient to operate with the tacit assumption that a database contains all the relevant information and to treat as false any proposition that cannot be proved from it. This is known as the *closed world assumption* (Reiter 1978). Again, this involves a non-monotonic inference relation, since the addition of new data may permit the derivation of a proposition which was previously underivable and had thus been classified as false. Further areas of application include reasoning about natural kinds, diagnostic reasoning, and natural language processing (Reiter 1987, McCarthy 1986).

Work on formalizing non-monotonic inference has progressed rapidly since its beginnings in the 1970s, and there is now a large body of mature work in the area, much of it highly technical in character. (For a collection of seminal papers, see Ginsberg 1987; for surveys of the field, see Brewka et al. 1997 and the articles in Gabbay et al. 1994. Antoniou 1997 offers a relatively accessible introduction to the area.)

One of the most important non-monotonic formalisms is *default logic*, developed by Raymond Reiter (Reiter 1980). This involves supplementing first-order logic with new rules of inference called *default rules*, which have the form

$$\frac{p : q}{r}$$

where p is known as the *prerequisite*, q as the *justification*, and r as the *consequent*. Such rules are to be read, 'If p , and if it is consistent with the rest of what is known to assume that q , then conclude that r '. In simpler cases ('normal defaults'), q and r are

the same, so the rule says that, given the prerequisite, the consequent can be inferred, provided it is consistent with the rest of one's data. Thus the rule that birds typically fly would be represented as

$$\frac{\text{Bird}(x) : \text{Flies}(x)}{\text{Flies}(x)}$$

which says that if x is a bird and the claim that x flies is consistent with what we know, then we can infer that x flies. Given that all we know about Tweety is that he is a bird, we can therefore infer that he flies. The inference is non-monotonic, since if we subsequently acquire information that is inconsistent with the claim that Tweety flies, then the rule will cease to apply to him.

The application of default rules is tricky, since it is necessary to check their justifications for consistency, not only with one's initial data, but also with the consequents of any other default rules that may be applied. The application of one rule may thus block that of another. To solve this problem, Reiter introduces the notion of an *extension* for a *default theory*. A default theory consists of a set of premises W and a set of default rules D . An extension for a default theory is a set of sentences E which can be derived from W by applying as many of the rules in D as possible (together with the rules of deductive inference) without generating inconsistency. An extension of a default theory can be thought of as a reasonable development of it.

Another approach, closely related to default logic, is *autoepistemic logic* (Moore 1985). This turns on the idea that we can infer things about the world from our introspective knowledge of our own minds (hence the term 'autoepistemic'). From the fact that I do not believe that I owe you a million pounds, I can infer that I do not owe you a million pounds, since I would surely know if I did. Building on this idea, autoepistemic logic represents rules of thumb as implications of claims about one's own ignorance. For example, the rule that birds typically fly can be represented as the conditional claim that if something is a bird and one does not believe that it *cannot* fly, then it does fly. Given introspective abilities, one can use this claim to draw the defeasible conclusion that Tweety flies, based on one's ignorance of reasons to think he cannot. This approach can be formalized using the apparatus of modal logic, with the modal operator L interpreted as 'It is believed that'.

A third approach is *circumscription* (McCarthy 1980, 1986; see also Lifschitz 1994). This involves formulating rules of thumb with abnormality predicates and then restricting the extension of these predicates – *circumscribing* them – so that they apply to only those things to which they must apply, given the information currently available. Take the Tweety case again. We render the rule of thumb that birds typically fly as the following conditional, where 'Abnormal' signifies abnormality with respect to flying ability:

$$\forall x(\text{Bird}(x) \ \& \ \neg\text{Abnormal}(x) \rightarrow \text{Flies}(x)).$$

This does not, of course, allow us to infer that Tweety flies, since we do not know that he is not abnormal with respect to flying ability. But if we add axioms which circumscribe the abnormality predicate so that it applies to only those things which

are currently known to be abnormal in this way, then the inference can be drawn. This inference is non-monotonic, since if we were to add the premise that Tweety is abnormal with respect to flying ability, then the extension of the circumscribed predicate would expand to include Tweety and the inference would no longer go through. Unlike the other strategies mentioned, circumscription can be formulated using only the resources of first-order predicate calculus, though its full development requires the use of second-order logic, allowing quantification over predicates.

Each of these approaches has its own strengths and weaknesses, but there are some general issues that affect them all. One problem is that they all allow for the derivation of multiple incompatible sets of conclusions from the same premises. In default logic, for example, a theory may have different extensions depending on the order in which the rules are applied. The standard example is the theory that consists of the premises *Nixon is a Quaker* and *Nixon is a Republican*, together with the default rules *Quakers are typically pacifists* and *Republicans are typically not pacifists*. Each of these rules blocks the application of the other, since the consequent of one is incompatible with the justification of the other. The theory thus has two incompatible extensions – one in which the first rule is applied and containing the conclusion that Nixon is a pacifist, the other in which the second rule is applied and containing the conclusion that Nixon is not a pacifist. Similar results occur with the other approaches mentioned.

The existence of multiple extensions is not in itself a weakness – indeed it can be seen as a strength. We might regard each extension as a reasonable extrapolation from the premises and simply plump for one of them. (This is known as the *credulous* strategy. The alternative *sceptical* strategy is to endorse only those claims that appear in every extension.) In some cases, however – particularly in reasoning involving time and causality – a plausible theory generates an unacceptable extension. Much work has been devoted to this problem – one strategy being to set priorities among default rules, which determine the order in which they are applied and so restrict the conclusions that can be drawn. (A much-discussed case is the *Yale shooting problem*, introduced in Hanks and McDermott 1987, where the rule of thumb that living things typically stay alive generates an unexpected conclusion in reasoning about the effects of using a firearm. For a description of the scenario, see FRAME PROBLEM entry and for discussion, see Shanahan 1997.)

A second problem concerns implementation. The goal of much work in this area is to build artificial non-monotonic reasoning systems, but the formal approaches that have been devised are not easy to implement. Default logic, for example, requires checking sets of sentences for consistency, and there is no general procedure for computing such checks. Restricted applications have been devised, however, and aspects of non-monotonic reasoning have been effectively implemented using the techniques of *logic programming* (programming in languages based on formal logic).

A third issue concerns the piecemeal character of much work in this area. Theories have been developed and elaborated in response to particular problem cases and with relatively little attention to the features of the inference relations they generate. Recent work has begun to address this issue and to identify properties that are desirable in

any non-monotonic inference relation (see, for example, Makinson 1994). One of these is that adding a conclusion back into the premise set should not undermine any other conclusions – a property known as *cautious monotonicity*.

Finally, a word about probabilistic logics. Most non-monotonic formalisms embody a *qualitative* approach: premises and conclusions are treated as either true or false, as in deductive logic. But probabilistic logics, in which propositions are assigned continuous probability values, can also be used to model certain types of non-monotonic inference. In probabilistic reasoning it is common to treat a conclusion as warranted if its probability, given the premises, exceeds a certain threshold. This inference relation is non-monotonic, since the addition of new premises may lead to a readjustment of one's probability assignments, with the result that a conclusion which previously passed the threshold now falls short of it. However, although probabilistic logics are well-suited for modelling reasoning under uncertainty (see Shafer and Pearl 1990), it is unlikely that they can do all the work required of a theory of non-monotonic inference (McCarthy 1986).

Bibliography

- Antoniou, G. (1997). *Nonmonotonic Reasoning*. Cambridge, Mass.: MIT Press.
- Brewka, G., Dix, J., Konolige, K. (1997). *Nonmonotonic Reasoning: An Overview*. Stanford: CSLI Publications.
- Gabbay, D. M., Hogger, C. J., & Robinson, J. A. (Eds.). (1994). *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford: Oxford University Press.
- Ginsberg, M. L. (Ed.). (1987). *Readings in Nonmonotonic Reasoning*. Los Altos, CA: Morgan Kaufmann.
- Hanks, S., & McDermott, D. (1987). Nonmonotonic logic and temporal projection. *Artificial Intelligence*, 33, 379-412.
- Lifschitz, V. (1994). Circumscription. In Gabbay, Hogger & Robinson (Eds.), pp. 298-352.
- Makinson, D. (1994). General patterns in nonmonotonic reasoning. In Gabbay, Hogger & Robinson (Eds.), pp. 35-110.
- McCarthy, J. (1980). Circumscription – a form of non-monotonic reasoning. *Artificial Intelligence*, 13, 27-39. Reprinted in Ginsberg 1987.
- McCarthy, J. (1986). Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28, 89-116. Reprinted in Ginsberg 1987.
- McCarthy, J., & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer & D. Mitchie (Eds.), *Machine Intelligence 4* (pp. 463-502). Edinburgh: Edinburgh University Press. Reprinted in Ginsberg 1987.
- Moore, R. C. (1985). Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25, 75-94. Reprinted in Ginsberg 1987.

- Reiter, R. (1978). On closed world data bases. In H. Gallaire & J. Minker (Eds.), *Logic and Data Bases* (pp. 55-76). New York: Plenum. Reprinted in Ginsberg 1987.
- Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13, 81-132. Reprinted in Ginsberg 1987.
- Reiter, R. (1987). Nonmonotonic reasoning. *Annual Review of Computer Science*, 2, 147-186.
- Shafer, G., & Pearl, J. (1990). *Readings in Uncertain Reasoning*. San Mateo, CA: Morgan Kaufmann.
- Shanahan, M. (1997). *Solving the Frame Problem*. Cambridge, Mass.: MIT Press.