linked-to-each-other (X, Y):- linked-to(X, Y), linked-to (Y, X).

This book is intended as an introduction to ILP. It is aimed at a reader with a basic knowledge of logic programming who wishes to learn the basics of new field of ILP, e.g. a knowledge engineer interested in the automatic synthesis of knowledge bases for expert systems. As such, it plays a useful role in gathering together information on ILP system and explaining them in a reasonably clear and concise way. The authors both have extensive experience of theoretical and experimental ILP. They originate from the AI Laboratory of the Jozef Stefan Institute (Ljubjana, Slovenia), which under its director Ivan Bratko, has had a long record of high quality work in both applied machine learning and logic programming.

The book is divided into two main parts: an introduction to the theory behind ILP systems, and a survey of the applications of ILP systems, the introduction being the most useful part. Much of the information given in this section is difficult to obtain from original sources (like too much of AI, many of the important papers have been published in hard to obtain conference proceedings and reports). The original technical papers are also in general quite difficult to read, as they assume a wide knowledge of machine learning and logic programming theory. The book describes in detail the basic ILP techniques of: relative least general generalization (RLGG), inverse resolution, search of refinement graphs, etc. Then proceeds to describe the best known ILP programs: MIS, FOIL, GOLEM, LINUS (developed at Ljubjana), etc.

The second part of the book describes examples of the use of ILP systems in practise. This part of the book I found less interesting, and I soon grew tired of reading page after page of experimental description. I would have been more convinced by it all if there was clearer evidence that the ILP systems were succeeding where traditional propositional systems had failed. I suggest that more work needs to be done in comparing ILP and existing propositional systems. I was also unconvinced by the extensive treatment given to the m-estimate method of measuring model fit—better methods now exist.

As a small quibble. Why is there a rather blurred picture of a strand of DNA on the cover of the book? I could find no reference to DNA in the text, the nearest being the description of an application of GOLEM to the problem of predicting a protein's secondary structure. Could not a more appropriate picture be found?

In conclusion, is ILP a chimera or the future of machine learning and knowledge acquisition? It is too early to tell, but this book provides the best available general introduction to the subject, and should be read by everyone interested in this exciting new area.

Reviewed by Ross D. King, Biomolecular Modelling Laboratory, Imperial Cancer Research Fund, London, UK

The phenomenon of commonsense reasoning: nonmonotonicity, action and information by Dimitrios Thanassas, Ellis Horwood, UK, 1992, pp 254, £37.95, ISBN 0-13-663634-9.

This book is, or claims to be, about commonsense reasoning. An example of commonsense reasoning that I feel has some relevance to this review is the following: if having read half of a certain book one has found nothing to lead one to believe that the book is worth reading, then it is reasonable not to bother reading the other half. I have to confess that I find myself in rather this position with respect to the book currently under review. I have read the first six chapters, and have not found the stomach to read the remaining four. If on that account the following review is discredited on the grounds that the reviewer could not be bothered to read the book through to the end, then so be it; nonetheless, I believe that what I have to say about the first half is quite likely to be of relevance to the second half also.

What, then, do we find in the first six chapters? The first chapter, a general critique of the current state of artificial intelligence (AI), need not detain us. The second is a broad survey of theories of

commonsense reasoning in AI, ranging from default reasoning, through reasoning about action, knowledge representation and problem solving to analogical reasoning. The coverage is reasonable, but the explanations are very unclear, and I doubt whether a reader not already acquainted with this material would be able to make much of it.

Chapter 3 introduces the author's own system, the "typicality logic". Despite the many negative comments that I have to make about the book as a whole (or rather, as a half), I must in fairness say that the basic idea of the typicality logic is an attractive one. Thanassas introduces a *typicality operator* "Δ" which can be prefixed to any sentence to indicate that that sentence is typically true, as, for example "$\Delta(bird(x) \rightarrow fly(x))$". One then performs inferences in the normal way, simply ignoring occurrences of "Δ" until one encounters a contradiction. Then, and only then, does one seek to remove the contradiction by identifying where an exception to a typicality sentence has occurred. As Thanassas says, "we are not afraid of contradictions as long as we know how to handle them."

As I have said, there is something attractive about the simplicity of this idea, and it does seem to correspond to the way we go about things in at least some of our everyday thinking. Unfortunately, the detailed presentation of the idea is so sloppy that I find it impossible to determine whether he has really got it to work. This sloppiness pervades the book, and is inexcusable.

Early on in the same chapter, for instance, he represents Moore (1985) as saying "typically I do not have an older brother." Actually, Moore did not say this: what he said was that if he had an elder brother then he would surely know about it. Be that as it may, Thanassas then goes on to represent the sentence he attributes to Moore in the form "Typically if x is human then x is not my brother", which he says is "a slight paraphrase" of Moore's example but "with the same meaning". But in fact it has a completely different meaning, both from Moore's actual example and from Thanassas's garbled account of it.

In Chapter 4, Thanassas moves on to apply his typicality logic to some of the classical problems in nonmonotonic reasoning, notably the Anomalous Extension Problem. This is really a class of problems, here represented exclusively by the so-called Yale Shooting Problem. His account of how typicality logic handles this problem is of some interest, in particular in showing how the typicality logic in itself is compatible with a range of different strategies for resolving contradictions, such as "prefer the conclusion that uses the fewer typicality instances, and from the aborted conclusion, abort those typicality instances that directly yielded the contradictory conclusion". As before, I do not think that anyone not already familiar with the Yale Shooting Problem could come to an understanding even of what the problem is, let alone how to solve it, on the basis of Thanassas's exposition at the beginning of the chapter. And sometimes he seems to go out of his way to mislead: for example, in defining the meaning of the event WAIT as "to signify a period of time during which nothing changes"—if this were really how WAIT was defined then the problem would simply not arise, since the anomalous extension is precisely the one in which something (the state of the gun) *does* change during the WAIT period.

In the middle of this chapter there is a short section on "prototypical objects". I had hoped that this might throw some light on the currently fashionable area of prototype theory. Unfortunately, the formal definition of what it is for some object to be a prototypical member of class P cannot possibly be right, since it has the consequence that any object which is *not a member of* P *at all* must be a prototypical member of that class! In this case, the formula can be mended by adjusting the scope of one of the connectives, but this only goes to illustrate how again and again Thanassas bungles the presentation of formal ideas, either making them unintelligible or manifestly failing to say what he really means.

Chapter 5 presents a nonmonotonic theory of causation. The first section presents a brief "literature review", citing Hume, Locke, Kant, Georgeff and Shoham—an incongruous juxtaposition to be sure, with little to convince me that Thanassas has read these authors in any depth. I derived a certain wry amusement from his statement that "the modality K has a precise interpretation in Shoham's logic of chronological ignorance", since in my own critiques of Shoham's work (Galton, 1988, 1991) one of the major problems that I singled out was precisely the

lack of any such precise interpretation! And Thanassas himself gives the lie to his own statement by immediately going on to say that Shoham's K "is similar to the modal operator 'necessary' or 'know'"—as if anything *could* simultaneously be similar to two such dissimilar notions as necessity and knowledge. Like Shoham (whom I have elsewhere criticized for this), Thanassas seems to think that causality is inherently bound up with epistemic notions such as knowledge and belief, but it does not seem to me that the "belief revision" type of considerations he invokes have anything specifically to do with causation *per se*.

The formal theory of causation that follows get off to a disastrous start (to my mind) with the statement that the relation *cause(s,l)* should imply, amongst other things, that "in the typical case, if s is true then l is also true". This flies in the face of the way we use the notion of cause in everyday reasoning. For example, there is nothing at all unusual about a statement such as

The man's careless driving caused five policemen and a dog to be killed.

But no-one in their right mind, surely, would insist that for this to be true it has to be the case that careless driving *typically* results in that (or indeed any) particular combination of fatalities. So it seems to me that Thanassas's formal account of causation goes wrong right at the start.

Later on in the chapter, Thanassas introduces the notions of *enabling* and *preventing*, which he rightly sees as being intimately bound up with causation. However, I find it extraordinary that having gone to the trouble of propounding quite an elaborate formal characterization of "*cause*" in his typicality logic, he then defines both "*enable*" and "*prevent*" in terms of the material conditional alone, *without reference* to "*cause*" or the typicality logic. Indeed, according to his Definition 3, since it is in fact the case that "if there are unicorns then Galileo is dead" (where "if" is interpreted as the material conditional), then not only must Galileo's death *enable* unicorns to exist, but also (equating "dead" with "not alive") it is the existence of unicorns that prevents Galileo from being alive! As if this were not bad enough, an immediate consequence of the definition, Thanassas' Theorem 1, quite fails to accord with the usual meanings of "enable" and "prevent": from the fact that having a car enables me to travel to London, say, it by no means follows (as according to the theorem it should) that not having a car would prevent me from travelling to London, since I could just as well go by coach or train.

The whole discussion of causation seems to me to be deeply flawed: the author is not alone in this, though, since the same flaws recur again and again in the AI literature. The idea of comparing causation with material implication, for example, only makes sense on the assumption that causation is a relation between propositions. Surely, though, it is a relation between particular *facts*. Causal *laws* try to summarize causal relations that hold (or tend to hold) between facts answering certain descriptions—this can, admittedly, make causation look like a relation between descriptions of facts, and hence something closer to propositions. But it is the *fact* that I threw a particular brick at a particular window on a particular occasion which is the cause of the fact that the window broke on that occasion; the more general statement that under suitable circumstances bricks thrown at windows cause them to break is, if true, true because of all those other particular facts, and not the other way round.

Thanassas uncritically inherits from some of his predecessors the confusions between statements of the forms:

1. The actual occurrence e_1 of type E_1 at place p_1 and time t_1 caused the actual occurrence e_2 of type E_2 at place p_2 and time t_2.
2. If an occurrence of type E_1 were to occur at place p_1 and time t_1 then it would cause an occurrence of type E_2 at place p_2 and time t_2.
3. Occurrences of type E_1 in circumstances of type C do always/typically/sometimes cause occurrences of type E_2.

I do not see how one can claim to produce an adequate formal analysis of causation without displaying at least some sensitivity to distinctions such as these; yet Thanassas in this chapter signally fails to do so.

Chapter.6, on reasoning about actions, is not much better. Thanassas tries to distinguish between facts and events, and exemplifies the former by *loud*(*noise*, t_1) (which is presumably meant to say that there is a loud noise at time t_1, though the only way I can interpret it is as saying that the entity "noise" was loud at t_1, which is rather different), and the latter by, *inter alia*, *love* (x, y, t_1, t_2). Of the latter he says that "since if x loves y in [t_1 . . . t_2] then x loves y at any time point t between t_1 and t_2, this event conforms to the definition of a fact but does not conform to the definition of an event", which makes me wonder why he wants to classify it as an event in the first place.

The confusions pile up thick and fast. Thanassas speaks of "the fact *loud* (*noise*, t_1)" but he also refers to a fact's "changing its truth value". How can the fact that there was a loud noise at t_1 change its truth value? This is what Quine (1960) calls an *eternal sentence*: its truth value does not depend upon the time at which it is uttered. Some of the problems arise, perhaps, from thinking that there is a meaningful dichotomy between facts and events. To my mind, the significant distinction is between *states* and events; and *both* of these can participate equally in facts, for example the fact that it is now dark (a fact about a state) and the fact that the sun set an hour ago (a fact about an event).

At this point I did not see much point in reading further. I hesitate to condemn the *substance* of the book as worthless, although as I have indicated it contains many gross conceptual confusions. The problem is that the material is so badly presented that it is sometimes impossible to see what the author is trying to say. There are faults both in the way the material is organized and at the level of detail; there are distracting misprints and spelling errors on every page. This reflects badly on author and publisher alike.

Indeed, I believe that the publisher should accept the major part of the blame on this score: the author is evidently not a native English speaker, and is inexperienced in publishing. Yet there is no sign that the book has been read through by a copy editor: the impression one has is that the author has submitted camera-ready copy and the publisher has printed it without further ado. Whether I am right or wrong in this surmise makes little difference: a book with this many errors should never have been published, and it is surely in the publisher's own interests to impose some more rigorous quality control than appears to have been exercised in this case.

Speaking of errors, and returning now to the material itself rather than the presentation: who originated this nonsense about Australian penguins being able to fly (p.42)? How can a scientific researcher seriously work with a sentence such as "red birds are usually Australian penguins"? Of course, these are only light-hearted illustrative examples and not really about penguins. But one of these days, someone is going to read this sort of stuff and come away believing that Australian penguins really do fly, or that they are red. There is an intellectual irresponsibility here that should not be allowed to pass without comment. It betrays a cavalier approach to reality which indicates that its perpetrator is working in a fantasy world where real facts do not matter: it is all just an intellectual game.

References

Galton, A, 1988. "Review of Y. Shoham *Reasoning about Change*". *Artificial Intelligence Review* **2** 285–288.

Galton, A, 1991. "A critique of Yoav Shoham's theory of causal reasoning". In: *Proceedings of the Ninth National Conference on Artificial Intelligence* AAAI Press/MIT Press, pp. 355–359.

Moore, RC, 1985. "Semantical considerations on nonmonotonic logic". *Artificial Intelligence* **25** 75–94.

Quine, WVO, 1960. *Word and Object*. MIT Press.

Reviewed by Antony Galton, Department of Computer Science, University of Exeter, UK

Do the right thing—studies in limited rationality by Stuart Russell and Eric Wefald, MIT Press, Cambridge, MA, £24.75, ISBN 0-262-18144-4.

As the cover blurb of this book points out, most intelligent systems face the same dilemma as the hero of Spike Lee's film—knowing just what is the right thing to do. Previous theories of rationality