the syntactic representation. The syntactic representation is the original string of words or a parse tree with these words labelling the leaves. The semantic representation must capture not this grammatical structure but its content. Confusingly, this is usually done by a logical formula; so the linguist's semantics is the logician's syntax!

Computer scientists use the word "semantics" to describe the mapping from a programming language to a mathematical theory. Ironically, this turns the logician's usage on its head. Logical semantics translates a mathematical formula into a program for calculating a truth value; computer science semantics translates a program into a mathematical formula.

Because of their remark on p. 36 of *Deduction* (see para. 1 above), I will assume that J-L & B intend the word "semantics" in the logician's sense. I assume that their mental models are based on Tarski's models of logical theories; that their deductive mechanism is an attempt to reason in the model theory in contrast to rule-based mechanisms that reason in the proof theory. I claim that it is not possible to do this.

3. Is semantic reasoning possible? If we regard Tarskian models as part of the real world, then reasoning with them would entail physically manipulating the real world. This has limited utility. It is not possible to conduct forward planning, hypothetical reasoning, counterfactual reasoning or abstract reasoning by manipulating the current world state. We must reason by manipulating an internal representation of the world.

At this point the problems of intentionality emerge, that is, we need a semantics to map this internal representation onto its meaning. This remains true even if the internal representation is based on a Tarskian model. Calling the manipulation procedure "semantic" does not affect the situation.

Basing a computational reasoning mechanism on Tarskian models presents problems for a finite computer. For example, some models have an infinite domain of objects. Some reasoning involves proving that an infinite collection of objects has a property. Some reasoning involves the representation and use of incomplete or vague information. These problems are solved in rule-based mechanisms by the use of quantifiers, variables, disjunction, and so forth. Some equivalent device is needed in model-based reasoners if they are to have the same reasoning power. J-L & B use such devices in their mental-model mechanism. For example, infinite numbers of objects are represented by a finite number of tokens; incomplete information is represented by having alternative models to cover the range of possibilities.

4. Are rule- and model-based reasoners different in kind? One paradigmatic example of a rule-based deductive system is a resolution-based theorem prover. The rules are formulae of predicate calculus in clausal form representing the axioms of the theory and the negation of the conjecture. The conjecture is proved by *reductio ad absurdum;* the clauses are "resolved" together, usually exhaustively, until the empty clause is derived.

However, resolution can also be viewed as a systematic attempt to check that none of the models of the theory provide a counterexample to the conjecture. The fact that resolution can be viewed in this way goes back to a metalogical theorem of Herbrand's. If the attempt to prove the conjecture fails after a finite search then a counterexample to the conjecture can be read off automatically from the failed attempt. Thus resolution can be viewed both as a rule-based and as a model-based mechanism!

This potential duality was brought home to me forcibly as a result of my first foray into automatic theorem proving. I built a model-based theorem prover for arithmetic called SUMS (Bundy 1973). Its model consisted of a representation of the "real line" as used by mathematicians in informal blackboard arguments. The hypotheses of the theorem were represented by placing points in appropriate positions on this "real line" and the conclusion was then read off from the model.

As I tried to get SUMS to prove harder and harder theorems, this simple idea became more and more elaborate. For example, consequences of the original hypotheses had to be propagated around the model before the conclusion could be read off. The natural propagation mechanism was forward-chaining with rules. After a while I realised that I had just built yet another rule-based mechanism. SUMS was now similar to a standard semantic tableau prover with a bottom-up search strategy. SUMS' progression from model-based to rule-based was incremental. There was no point at which the nature of its reasoning dramatically changed in kind.

5. Conclusion. I have argued that there is no difference in kind between the mental-model deduction mechanism of J-L & B and rule-based mechanisms. Indeed, it is possible to view many deduction mechanisms as simultaneously of both types. The issue of intentionality arises with both types of mechanism, and is not finessed by the use of a model-based approach. To the best of my knowledge J-L & B make no claim to the contrary. However, others may erroneously draw that conclusion from the free use of words like "semantics," "model," and so forth. For this reason I recommend that the word "semantics" be used with extreme caution. It is a highly ambiguous term and has great potential to mislead.

None of this detracts from Johnson-Laird & Byrne's significant contribution in defining a new deduction mechanism and providing evidence for its psychological validity.

ACKNOWLEDGMENT
The work reported in this commentary was supported by a SERC Senior Fellowship to the author.

Mental models and nonmonotonic reasoning

Nick Chater
Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, Scotland
Electronic mail: *nicholas@cogsci.ed.ac.uk*

Johnson-Laird & Byrne (J-L & B) are equivocal concerning the scope of mental-model theory. On the one hand, they are careful to note that mental models are aimed primarily at explaining deduction, although commonsense inference is not deductive in character. On the other hand, they contend that mental-model theory solves the problem of nonmonotonic reasoning, which is *not* deductive and is characteristic of commonsense inference. This equivocation requires clarification: An account of deductive reasoning casts light on a fascinating if rather arcane human ability; an account of nonmonotonic inference in general would be little short of a theory of thinking. It is not clear, therefore, exactly what J-L & B see as the domain of the mental-model account. I shall argue that mental-model theory does not in fact address the problem of understanding commonsense nonmonotonic reasoning, still less provide a solution to it.

Everyday, commonsense reasoning may be conceived of as a species of inference to the best explanation: It involves inferring from given information to what best explains and is best explained by that information (Fodor 1983). Such inference is nonmonotonic, because the addition of new information can invalidate what were previously plausible conclusions. So, for example, the plausible inference from hearing the sound of purring behind the door to the conclusion that the cat is trapped in the cellar is immediately overridden if I catch sight of the cat in the garden. The premise on which the inference is based, the purring, need not be withdrawn, although another explanation for this fact may be sought. By contrast, in monotonic reasoning, the conclusion of a valid argument can only be challenged if one of its premises is false.

Providing an account of inference to the best explanation is very difficult. Inference to the best explanation encompasses theory confirmation in science, the problem of inferring the scientific theory which best fits the available data. Understanding inference to the best explanation also presupposes a solution to the notorious frame problem (McCarthy & Hayes 1969). Indeed the problem of nonmonotonic reasoning and the frame problem really constitute the same problem looked at from different points of view. Where nonmonotonic reasoning concerns which conclusions *should* be revised when new information is added, the frame problem involves deciding which conclusions *need not* be revised.

Nonmonotonic reasoning has proved resistant to a vast number of extremely ingenious proposals within artificial intelligence. The principal line of attack has been to attempt to provide a *nonmonotonic* logic which captures commonsense, rather than deductive, inferences (e.g., Hanks & McDermott 1987; McCarthy 1980; Reiter 1980; Shoham 1987). These logics have been dogged by two serious problems (McDermott 1987; Oaksford & Chater 1991). First, it has not been possible to capture the nonmonotonic inferences that people routinely draw (specifically, most computational methods tend only to draw extremely weak conclusions). Second, these methods have been plagued by computational intractability due to explicit or implicit reliance on consistency checking. J-L & B's treatment of nonmonotonic reasoning addresses neither of these difficulties, but I shall concentrate on the first and most fundamental.

Throughout their book J-L & B give examples of how mental models can be constructed and checked to generate inferences licensed by a standard logic (typically, propositional or predicate calculus). This is not surprising, because they are primarily concerned with deductive reasoning. The rules governing the building of models and searching for alternative models ensure that the notion of the validity of the logic concerned is respected. A mental-model account of nonmonotonic reasoning would require similar rules for model construction and search; but how could these rules be chosen? By analogy with the deductive case, we might expect these constraints to respect the notion of the validity of some nonmonotonic logic. J-L & B rightly do not attempt to follow this line, since there is no appropriate nonmonotonic logic to which appeal can be made. Yet, without an underlying logic of some kind, the processes of building mental models will be without any constraint or justification.

There is also a more specific worry. Inference on the basis of a failure to find countermodels, which is at the heart of the mental-model account, appears inapplicable to nonmonotonic reasoning, because, by definition, such countermodels invariably exist: If there were no model in which the premises are true and the conclusion false, the reasoning would be valid according to a standard, deductive, monotonic logic, and would not be an instance of nonmonotonic reasoning at all. Furthermore, not only do countermodels exist, but people find it easy to generate these countermodels on demand. For example, returning to the inference that the cat is in the cellar on the basis of hearing a purring sound, it is easy to generate a host of countermodels in which this inference is not correct. For example, the sound may be produced by a different cat, a tape recorder, a person doing a cat imitation, the wind, and so on. Quite generally, it is easy to find countermodels for everyday nonmonotonic inferences. This means that if nonmonotonic inference proceeded from a failed search for countermodels, no nonmonotonic inferences would be drawn at all! Notice that the mental-model theorist cannot argue that only plausible models are constructed, or that the single most plausible model is chosen, because the problem of finding such models is simply a restatement of the problem of inference to the best (i.e., most plausible) explanation (i.e., model). If mechanisms for assessing plausibility could be assumed, then the problem of nonmonotonic reasoning would already have been solved, and invoking mental models would be unnecessary (for a discussion of more detailed proposals concerning how mental models might be applied to nonmonotonic reasoning, see Chater & Oaksford 1993; Garnham 1993).

What leads J-L & B to claim that they have solved the problem of nonmonotonic reasoning? Simply the observation that inferences made on the basis of a particular model may not follow deductively from the premises given; the addition of later information which demands that the set of models be searched more fully can lead to such a conclusions' being withdrawn. This certainly shows how mental models can make logically invalid inferences and later recover from them, but it says nothing about the real problem in hand. The problem of nonmonotonic reasoning is not to explain how it is possible to jump to conclusions and later revise them; this could be achieved trivially by modifying any standard method of proof. The problem is to explain how people are able to jump to *sensible* conclusions and revise these *appropriately* as new information is encountered, and indeed to elucidate what it means for a conclusion to be sensible or appropriate. Regarding these questions, artificial intelligence and cognitive science have had depressingly little to say; unfortunately, mental-model theory appears to have nothing further to add.

Some difficulties about deduction

L. Jonathan Cohen
The Queen's College and Sub-faculty of Philosophy, Oxford University, Oxford OX1 4AW, England

Johnson-Laird & Byrne's (J-L & B's) arguments are beset with a number of serious difficulties. I have space to mention two.

The first can be stated quite briefly. According to the authors (p. 22), "to deduce is to maintain semantic information, to simplify, and to reach a new conclusion." But by this criterion even *modus ponens* should seem an "odd or improper" (p. 21) deduction because it does not maintain semantic information. For example, the conclusion "There is a triangle" obviously does not maintain all the semantic information contained in the premises "If there is a circle, there is a triangle" and "There is a circle." So either J-L & B are wrong to assume (e.g., p. 41) that such a *modus ponens* inference is a respectable example of deduction, or they are wrong to hold that deduction ought to maintain semantic information. If they abandon the assumption that such a *modus ponens* inference is a respectable example of deduction, they exclude what would normally be acknowledged as a veritable prototype of deduction and leave obscure what kinds of mental processes they are in fact investigating under the rubric of "deduction." But if they abandon the view that deduction ought to maintain semantic information, they must give some other explanation for most people's reluctance to treat detachment of an asserted conjunct and similar trivialities as full-blooded deductive processes.

The second difficulty is a more complex matter. Consider a proposed inference from

> If John's automobile is a Mini, John is poor, and, if John's automobile is a Rolls, John is rich

to

> Either, if John's automobile is a Mini, John is rich, or, if John's automobile is a Rolls, John is poor.

It is easy to imagine circumstances (automobile prices, customer preferences, etc.) in which the premise of this proposed inference is true and the conclusion is false. So the inference appears to be invalid, and we seem somehow to be capable of having reached this evaluation by searching in our minds for imaginable models that can function as counterexamples. But, if we were to implement J-L & B's own model-theoretic, counterexample-seeking algorithm, we ought apparently to come up with a different evaluation of the inference's validity: The inference