large differences in $Bel(\theta)$. For a metric he proposes the "well known Blaschke metric." This metric was apparently developed for differential geometry—my mathematical encyclopedia does not go into more detail. This is one instance among several in which the depth of the mathematics seems out of proportion to the problem being considered.

One principle mentioned in this chapter, that of "calibration," could be more thoroughly discussed. I. J. Good has written on the calibration problem extensively; and various approaches to it have been analyzed and criticized by Seidenfeld; neither is referred to.

Chapter 8 concerns belief revision, in the sense of updating: Given some new constraints, how does *Bel* change? Conditionalization, Jeffrey updating, and minimum cross entropy are considered.

Chapter 9 is a short chapter on independence. Since the constraints embodied in independence assertions need not be all linear, this chapter does not fit neatly into the general thrust of the book; it appears to be included mainly because belief nets represent a popular formalism, and belief nets crucially depend on judgments of independence.

Chapter 10 presents results, some new, concerning the computational complexity of the "inference processes" being discussed. It includes discussion of random polynomial time processes as well as P and NP. This is the richest and most rewarding chapter of the book. But remember that even here we are concerned primarily with a finite, propositional, language and a belief function subject to linear constraints. That most of the results are decidedly negative could be considered discouraging for the deductive approach to uncertainty management.

The last two chapters concern the predicate calculus. Chapter 11 presents a formalization of the problem, and proves a useful representation theorem; Chapter 12 discusses some "principles of predicate uncertain reasoning." These are symmetry principles suggested by W. E. Johnson and discussed extensively by R. Carnap. Laplace's law of succession ($P(a_{n+1}) = (r + 1)/(n + 2)$, corresponding to Carnap's predictive inference in the λ continuum with $\lambda = 2$) is given. The discussion of quantified formulas might have been improved by reference in this chapter to the work of Scott and Krauss (*Assigning probabilities to logical formulas*, ***Aspects of inductive logic***, North-Holland, 1966, pp. 219–264), in addition to the reference to Gaifman. Nevertheless it is gratifying to see Carnap receiving the credit he deserves; his work is often neglected in modern discussions of "uncertain reasoning."

The most serious deficiency of content is the lack of a discussion of statistical inference. The author is aware of this, but offers the excuse that to do so would be to "move on to ever thinner ice. Just assuming belief to be a probability is ... already a grand idealization" (p. 195). Yet one may speculate that just as the intelligent agent can have some intuitions that conform to probabilities without being an accurate probability machine, so he may have some sound statistical intuitions that may provide foundations for probabilities, without being an accurate statistics machine. We are hardly in a position to neglect any generally agreed-upon and intuitively appealing sources of credence.

Many pages of the book are devoted to proofs of theorems. These proofs are not easy to follow. In part this may be do to differences between British and American styles of mathematical proof, but in part it is due to the heavyweight mathematics employed (for example, the Blaschke metric already mentioned). It is true that greater generality, and greater generalizability, are obtained in this way; but since the cases treated (finite propositional languages with linear belief constraints) are so constrained anyway, it is not clear that the mathematical game is worth the candle.

Nevertheless, for the mathematically sophisticated reader, many interesting results and useful techniques are represented. In particular, Chapter 10 on complexity results is a valuable addition to the literature of this area. As mentioned above, belief nets get short shrift (because the constraints they embody are not linear), and statistics is neglected. While it may seem that too little attention is paid to the predicate calculus, it could be argued that in real applications propositional representations will do just as well.

HENRY E. KYBURG

ALLAN RAMSAY. ***Formal methods in artificial intelligence.*** Cambridge tracts in theoretical computer science, no. 6. Cambridge University Press, Cambridge etc. 1988, ix + 279 pp.

Many universities teach artificial intelligence (AI) by having one undergraduate course that introduces students to a very wide variety of topics, usually including search and search heuristics, representational systems (including formal logic), problem solving, vision, expert systems, language, learning, connectionism, abduction, robotics, and planning. They then have further AI courses, often at the graduate level, that follow up one or another of these topics. The book under review is intended for the follow-up course in "Logical Methods in AI."

Although the title of the book under review suggests that it is concerned with the whole panoply of formal and theoretical foundations of AI (as opposed to being oriented towards implementation or towards cognitive science, perhaps), this is not true. In point of fact the book is exclusively concerned with logic. In this, it omits other "formal" areas of AI such as the theory of search algorithms, the mathematical background to neural nets and to vision, the physics behind robotics, computational learning theory, probability theory, and qualitative physics—all of which, plus many other topics, might correctly be called "formal methods in AI." The book is also not interested in applying the explained formal logic to topics in AI. Although occasionally such applications are mentioned (e.g. STRIPS-planning systems, temporal logic as a semantics for programs, and truth maintenance systems), there is only rarely much more than a reference pointing to the literature.

As a logic book then, as a logic book to be used to orient graduate computer science students into AI, how good is it? In brief: it is the best one around. This is not to say it is without shortcomings; rather, it is a matter of all other books being aimed at a different market. For example, Turner's ***Logics for artificial intelligence*** (LVI 339) introduces non-classical logics to "logically mature" students; Chang and Lee's ***Symbolic logic and mechanical theorem proving*** (Academic Press, 1973), Loveland's ***Automated theorem proving*** (XLV 629), Wos, Overbeek, Lusk, and Boyle's ***Automated reasoning*** (LI 464, LIX 1437) are directed towards automated theorem proving and introduce logic in a format aimed at that end; Paulson's ***Logic and computation*** (Cambridge University Press, 1987) is interested in the use of logic for testing program correctness; and Genesereth and Nilsson's ***Logical foundations of artificial intelligence*** (LV 1304), the main competition for this book, is more interested in applying logic to some knowledge representation areas of AI. Ramsay is the only one simply teaching logic in detail to AI students. One might argue that students would be just as well served by a logic textbook written for philosophy or mathematics students; but when compared only to those books purporting to teach logically-oriented AI, Ramsay's is notably better.

Ramsay starts with a nice introduction that distinguishes semantics and syntax in a way accessible to the beginning logic student. He then develops propositional logic in three formats: as a (Hilbert) axiom system, as a (Gentzen) sequent system, and as a (Beth) tableau system. He sketches proofs of the equivalence of these formulations, and closes the chapter with some interesting remarks on the "point" of the different formulations. He views the Hilbert formulation as "a purely formal set of rules for manipulating sets of symbols which are given a meaning from outside"; the Gentzen formulation as "an attempt to give precise form to the patterns of reasoning which are recognised as being valid"; and the Beth formulation as working "directly with the intended semantics of the language." The present reviewer would have liked to see the Gentzen system further developed into a modern-style Fitch system—which is more in accord with the way elementary logic is presented in philosophy departments. One rather disconcerting habit Ramsay has in his presentation of proof sketches is to characterize issues he is not going to present in detail as being "rather long-winded." For example, he says this of the proof of the cut-elimination theorem for Gentzen systems (which is also claimed "not [to] introduce any major new ideas"), and the students are referred to Gentzen's original paper(!).

Ramsay presents the syntax of the first-order logic (FOL) in a vaguely computer-influenced way, for example with variables prefixed by underlines and lower-case predicate symbols with pairs of subscripts (rather than, say, super- and subscripts). The semantics of FOL is briefly presented (three pages), where truth in a model is defined only for sentences (quantified formulas are given a substitutional interpretation, pp. 38–39). A Hilbert formulation of the syntax with a rule of generalization (GEN) is presented, and a deduction theorem for FOL in this formulation is proved. This formulation of the syntax allows free variables to occur in premisses (and conclusions), although GEN is blocked if the variable of generalization occurs in premisses.

So one might ask what type of soundness theorem Ramsay is going to prove, since there are syntactically correct derivations with free variables while his semantics accommodates only sentences. For one thing, *weak* soundness is not true: $\vdash (Fx \rightarrow (Gx \rightarrow Fx))$, yet no model satisfies $(Fx \rightarrow (Gx \rightarrow Fx))$ and so not $\models (Fx \rightarrow (Gx \rightarrow Fx))$.

Ramsay's strategy for dealing with soundness (his Theorem 3.2) is to modify the semantic definition of 'true in a model' in such a way as to have "formulae with free variables ... be thought of as implicitly universally quantified" (p. 43). But this cannot be taken literally, lest it convert the semantics from being "too weak for the syntax" into one that is "too strong for the syntax"; for, according to this proposal $Fx \models \forall xFx$ would be semantically valid (interpreting the free variable in the premiss as universally quantified), yet not $Fx \vdash \forall xFx$, due to the restriction on GEN. Instead of directly appealing to the

simple notion of being universally quantified, in the proof of Theorem 3.2, Ramsay makes two rather subtle moves. Let A^* be the universal closure of A. In the first of these moves, he sets out to prove three things: (1) that if A is an axiom then A^* holds in any model; (2) that the rules of inference (including GEN) preserve truth-in-a-model for universal closures, that is, if $P_1, \ldots, P_n \therefore C$ is a rule of inference and if $P_1^*, \ldots, P_n^*$ are each true in every model then C^* is true in every model; and (3) that "any proof preserves truth in a model." Part (3) of this move is not quite right. Suppose we are given that $\Gamma \vdash A$. Consider the case where A is an (instance of an) axiom. Here Ramsay says that A is true in every model ("that's what we proved in part 1"); but in fact this was *not* proved in part 1. Rather it was proved that A^* held in each model. So in fact, if A is an axiom then what we know is that $\Gamma \vDash A^*$. Similarly, for the case where A was a member of Γ: Ramsay says "if $\langle W, V\rangle \vDash \Gamma$ then $\langle W, V\rangle \vDash A$, since A is one of Γ." But for the interesting case where A has free variables and is a member of Γ, then Γ is not satisfied by any model. At this point Ramsay's "second move" comes to play: Theorem 3.2 is restricted to the case where Γ is a set of *sentences*. So we do not have an unrestricted soundness theorem for the system; instead the theorem is that if Γ contains only sentences, then if $\Gamma \vdash A$, then $\Gamma \vDash A^*$. Note that the condition that A^* be universally closed cannot be dropped: $\forall xFx \vdash Fx$ is syntactically valid, and the premisses are all sentences, yet the conclusion is not true in any model. Just as our earlier intuitive assessment said: Ramsay's syntax is "too strong" for his semantics and he has not succeeded in redefining the crucial semantic notions to make up for this.

But Ramsay seems to think that he has proved more. For example, he claims that a corollary of this theorem is weak soundness ("if A is provable from an empty set of hypotheses, then any model at all must be a model of A"). But what Ramsay has proved is that if $\vdash A$ then $\vDash A^*$. Note also that Ramsay relies on his Theorem 3.2 in his proof of the consistency of FOL (Theorem 3.4, p. 45).

Ramsay includes a brief account of functions, identity, and arithmetic. Rather than introduce function symbols as a primitive vocabulary, Ramsay chooses to suppose that the relations which are functional in nature can be identified somehow, and that when they are existentially quantified the result can be abbreviated as (what we normally would call) a function. Here is what he says. "We can supplement the language [FOL] with a special notation for talking about functional relationships without changing any of its basic characteristics. Suppose P is an ordinary relation and R is a functional one. Then the sentence $\exists \_x(R(\_x, t_1, \ldots, t_n)$ & $P(t_{n+1}, \ldots, \_x, \ldots t_{n+m})$ can only be true for at most one instantiation of $\_x$. We can abbreviate this sentence to $P(t_{n+1}, \ldots, F_R(t_1, \ldots, t_n), \ldots, t_{n+m})$, where F_R is simply a piece of notation indicating that this sentence is an abbreviation of the first form." Identity is then introduced with this comment. "There was very little point in having equality until now, since in the basic notation only identical constants could denote identical individuals. Once we have functional notation it does become worth adding axioms for equality. These axioms are all very obvious, but they do need to be spelled out." And we are then given axioms for reflexivity, symmetry, transitivity, and the collection of all "functional identities" (sometimes called Euclid's law when presented as a rule of inference: that if the arguments to a function are identical then so is the result of applying the function). Ramsay says that the functional identities "follow from our introduction of functional notation as an abbreviation for existential quantification over function relations, and hence can be derived from the basic axioms" of FOL plus reflexivity, symmetry, and transitivity. This is incorrect. First, Ramsay does *not* have existential quantification over function relations; instead he allows quantification over arguments of a relation, and the relation may happen to be a functional one. There is furthermore no syntactic recognition of functional relations: the notion of being functional is given semantically with no way to use it syntactically. Without explicit quantification over functions, Ramsay needs to add some version of Leibniz's law as a rule. As the system is now presented, $a = b \rightarrow (P(a) \leftrightarrow P(b))$ cannot be proved.

The chapter on FOL continues with clear proofs of various important theorems (the extension lemma, the Herbrand universe lemma) culminating in the completeness theorem. A clear presentation of Turing machines is given; the undecidability of FOL is claimed to follow from the unsolvability of the halting problem ("though the details are very long-winded"). At the end of the chapter a brief nod is made in the direction of fuzzy logics and relevance logics. Although this chapter is good for its clear presentation of the metatheory of FOL, it puzzles the reviewer as to why there is no alternative development of FOL as (say) a natural deduction system—as was done for the propositional logic.

Chapter 4 is devoted to (automated) theorem proving, and is prefaced with an essay on why "the Undecidability Theorem is not as damning a result as it first seems." This is followed by a lengthy discussion of various unification algorithms, ending with one in which an *occurs check* is "almost free." Backward chaining, negation as failure, and resolution are introduced; and various resolution-restricting

strategies and ordering strategies are given. The connection method is described (and advocated) as "an intricate encoding of the sort of reasoning embodied in tableau proofs." Although the material in this chapter is crucial to an AI student, and although it is very clearly expressed, it remains puzzling to the reviewer why Ramsay did not choose to introduce these various resolution and connection graph methods as distinct developments of FOL (in the same way that he introduced the three different developments of propositional logic). Indeed, it is difficult to discern Ramsay's attitude toward "theorem proving" as he uses the term. Certainly the results of a resolution proof or a connection proof would not constitute a proof as defined in the formal account of Chapter 3. Perhaps Ramsay thinks of all such methods as some sort of semantic discovery technique, essentially relying on the soundness and completeness of FOL?

Chapter 5 concerns modal logic. Ramsay considers the axioms K, T, S4 and B, plus the rule NEC. He develops a standard possible-worlds semantic theory of these (Kripke-normal) logics, and gives quick indications of what computer-oriented areas these logics have been used in: temporal reasoning, the blocks world, planning domains. After an extremely brief (two-page) introduction to modal FOL, including the concepts of rigid designators and the Barcan formula (he uses the biconditional version without considering using only one or the other of the conditionals), Ramsay moves directly to theorem-proving methods for modal logics. Here he considers a modal semantic tableaux method that keeps explicit track of "which world we are in," and also a "translation method" in which modal formulas are represented as first-order formulas about the relations amongst possible worlds presupposed by any particular formula to be proved. The version chosen for the tableaux method is not optimal, in the reviewer's opinion, since it involves the deletion of branching steps that have already been constructed. As Ramsay remarks "this [method] is slightly long-winded." Ramsay does not remark on the fact that his tableaux method will generate an infinite tree in an attempt to prove (for example) $MLp \rightarrow LMp$ in S4, nor does he comment on the issue of "loops" or "infinite branches" generally in modal logics. His discussion leaves one with the impression that these problems are somehow due to quantifiers, as in FOL, and not to the modal operators. This chapter concludes with short but interesting "applications" of modal logic to reasoning about knowledge, to planning (STRIPS-style), and to belief structures.

The book continues with a chapter on temporal logic in which, first, various modal axioms for time are discussed, and then the McCarthy–Hayes situation calculus is introduced (together with the crucial notion of "frame axioms"). The chapter closes with a comparison of points vs. intervals as foundations for a theory of time.

Following the chapter on time is a rather longer and more complex chapter on non-monotonic reasoning. A nice introduction to some sorts of default reasoning is followed by an introduction to Reiter's default logic, including the semantics (default extensions) and the proof theory. (Ramsay says the whole enterprise "is depressing in the general case" but that "it often happens that we can find default proofs in particular cases.") McCarthy's notion of circumscription is introduced and reasonably carefully explained. A brief explanation of the relationship between default logic and circumscription is provided by discussing Shoam's concept of preferential satisfaction in a non-monotonic frame. The chapter closes with an interesting and deep discussion of truth maintenance systems. In the reviewer's opinion this chapter is the best student-oriented introduction to the formal aspects of default reasoning available, lacking only some better examples that are clearly worked out. It is certainly more "available" to students than the discussion in Genesereth and Nilsson, its only serious competitor in this arena.

The last serious chapter concerns properties, and is mostly concerned with the use of the λ-calculus to represent natural language semantics, especially by way of Montague semantics. There is an interesting (if elementary) account of some issues in natural language semantics, such as compositionality and type raising, and it could be used as a pointer for the formally-minded student to interesting linguistic developments.

The final chapter, *Alternative views*, briefly discusses "constructive logic," mentioning intuitionism and relevance logic. An account of game-theoretic semantics is presented, and claims are made for the possible relevance to these contructivist views. The last few pages of the chapter introduce and make fun of fuzzy logic.

The book has a number of irritating typographical errors, but I discovered that most of them were caught in an errata sheet. (But change both occurrences of '$W \models$' in line 2 of "Part 2" on p. 44 to '$\langle W, V\rangle \models$', change 'A' in the line immediately before Theorem 3.4 on p. 45 to '$\neg$A', and change 'is' to 'it' in (1.1) of p. 66.) Various infelicities in the bibliography are not, however, mentioned on this sheet. (See, e.g., the entries for Lewis, Robbin, Shortliffe, and Pelletier).

Overall this book deserves a place in graduate or senior undergraduate courses in formally-oriented AI, either as a main or supplementary textbook or else as a place from which instructors can crib lectures.

I would like to express appreciation to Bernard Linsky, Herbert Enderton, and David Sharp for discussions concerning issues raised in this review. FRANCIS JEFFRY PELLETIER

RAYMOND M. SMULLYAN. ***First-order logic.*** Corrected republication of XL 237. Dover Publications, New York 1995, xii + 158 pp.

In reconciling the contrary viewpoints of syntax and semantics, Gödel's completeness theorem lies at the heart of mathematical logic. Over the years numerous approaches to proving completeness have been explored, the efforts justified by the theorem's primary importance. Smullyan's ***First-order logic***, a corrected reprinting of the 1968 original (XL 237), explains many of these approaches in detail.

For example, the method of *analytic tableaux* runs roughly as follows. Let φ be a sentence of predicate logic, and suppose we want to prove that either φ is satisfiable or its negation is provable. Try to satisfy φ by building a tree: to satisfy $\theta \wedge \psi$ one needs to satisfy both θ and ψ; to satisfy $\theta \vee \psi$ one can branch and try to satisfy either one of the two; to satisfy $\exists x\, \theta(x)$ one tries to satisfy $\theta(c)$ for a suitable constant c; and so on. If every branch of the tree yields a contradiction at some finite stage, we have the desired proof of $\neg\varphi$. On the other hand, if the construction has been done carefully, an infinite branch yields what is called a *Hintikka set*. These sets, though not maximal, are saturated downwards (e.g. if $\theta \wedge \psi \in S$ then $\theta \in S$ and $\psi \in S$, though not necessarily conversely), and from such a set it is easy to build a model of φ.

Such constructions are the essence of Smullyan's book. Part I covers propositional logic and various proofs of compactness, including the tableaux method described above. Part II introduces predicate logic and its semantics, extends the tableaux method to that domain, and abstracts the general notion of an *analytic consistency property* from the resulting completeness proof. There is also a discussion of various Hilbert-style axiomatizations of predicate logic, and of approaches to proving completeness using witnessing constants in the style of Henkin and Hasenjaeger. Part III treats Gentzen-style proof systems and the cut-elimination theorem (Hauptsatz), Craig's interpolation lemma, and some interesting variants of Gentzen-style deduction.

Two of the book's most salient features are its thoroughness and its rich use of notation. In most cases varying approaches to proving a theorem are compared and contrasted, and their underlying ideas are abstracted and generalized. So, for example, variations of analytic tableaux using signed and unsigned formulas are considered; analytic (and synthetic) consistency properties are introduced to generalize the notion of consistency in a proof system without cut (and with cut, respectively); and the abstract notions of "regular" and "magic" sets are extracted from the Henkin–Hasenjaeger completeness proofs described above. This style has its downside: at times the abstraction and terminology are overwhelming and make casual reading difficult. But all in all the exposition is clear and organized, and the subject matter is pithy and well chosen.

I can see the book being useful in the following capacities:

1. *As a supplementary logic text.* The prose is friendly and readable, with helpful side-remarks and examples. With selective cutting and pruning, the chapters on propositional logic and tableaux would make interesting supplementary reading for an introductory logic course, and consistency properties provide a uniform method of proving completeness, compactness, the interpolation lemma, Robinson's joint consistency theorem, and the like.

2. *As a general logic reference.* The book surveys a number of important syntactic constructions and methods, providing some historical perspective as well.

3. *As a proof theory reference.* Since analytic tableaux are essentially just Gentzen-style cut-free proofs turned upside-down (Chapter XI discusses the exact relationship), the book offers a transparent proof of the completeness of cut-free Gentzen systems. Proof theorists will further appreciate, for example, a complete Hilbert-style system that avoids the use of modus ponens, a discussion of the extended Hauptsatz (sometimes referred to as the midsequent lemma), and a constructive proof of the interpolation lemma.

4. *As a theoretical computer science reference.* Tableaux methods have made their way to the fields of automated theorem proving and proof complexity. Though the book does not cover resolution methods, for example, it provides good background to the study of such topics.

This Dover paperback edition is nicely bound and features the same attractive typesetting of the original edition. At a list price of about $8, it is well worth the cost. JEREMY AVIGAD