Logical Dilemmas: The Life and Work of Kurt Gödel. *By John W. Dawson, Jr.* A K Peters, Ltd., Wellesley, MA, 1997. $49.95. xiv+361 pp., hardcover. ISBN 1-56881-025-3.

Kurt Gödel was one of the most important mathematicians of the twentieth century and certainly the most important logician of the century. His work was of the highest importance to modern logic and one of the cornerstone foundations of computer science. Thus it is particularly welcome that Dawson has written a book-length biography of Gödel which contains much information about the life and work of Gödel that has not been widely known.

The external circumstances of Gödel's life are easily related. He was born in 1907 in Brünn (now Brno, in the Czech Republic) in a German-speaking community. In 1924, he moved to Vienna to attend the University of Vienna and, in 1929, he completed his doctoral dissertation containing the completeness theorem. In 1930-1931, he proved and published the incompleteness theorems; this work became his Habilitationsschrift. In Vienna, Gödel came in contact with many logicians and mathematicians, including P. Furtwängler, H. Hahn, M. Schlick, R. Carnap, O. Neurath, K. Menger, O. Tausky-Todd and many others; Gödel was also a member of the "Vienna Circle" although he did not subscribe to its logical positivist philosophy. He traveled between Princeton and Vienna from 1933 until 1940, holding visiting positions at the Institute for Advanced Studies (IAS) in Princeton and a docent position in Vienna, and also spent a half-year in Notre Dame. In 1938 he married Adele Nimbursky (née Porkert), a former dancer. In 1940 he moved permanently to Princeton, eventually becoming a full professor at the IAS, where he continued research in logic, physics and philosophy. Gödel tended to be private, reserved and very cautious; he suffered from hypochondria, paranoia and a distrust of medical doctors. He died in 1978 of "malnutrition and inanition" due to self-starvation.

Of course, the real reason for our interest in Gödel is the significance of his work in logic. His most important accomplishments are probably the completeness and compactness theorems for first-order logic, the incompleteness theorems, and the consistency of the axiom of choice and the continuum hypothesis.

The completeness theorem for first-order logic states the equivalence of the syntactic notion of "consistency" and the semantic notion of "satisfiability." A set Γ of first-order sentences is said to be *consistent* if there is no proof of a contradiction from Γ. The set is said to be *satisfiable* if there is some model (i.e., a structure or an interpretation) in which every sentence in Γ is true. The completeness theorem is the fact that the set Γ is satisfiable if and only if it is consistent. The completeness theorem can also be stated in terms of implication: we say that "Γ logically implies ϕ", written $\Gamma \models \phi$, if the sentence ϕ is true in every model in which Γ is true, and we say that "ϕ is provable from Γ", written $\Gamma \vdash \phi$, provided there is a proof of ϕ from the assumptions in Γ. The completeness theorem then states that $\Gamma \models \phi$ holds if and only if $\Gamma \vdash \phi$ holds. Thus the completeness theorem implies that, for first-order logic at least, all valid reasoning can be captured by the usual notions of first-order provability.

Since proofs are finite, the compactness theorem is an immediate consequence of the completeness theorem; namely, Γ is satisfiable if and only if every finite subset of Γ is satisfiable, or equivalently, $\Gamma \models \phi$ if and only if there is a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models \phi$.

Loosely speaking, the incompleteness theorems of Gödel imply that there is no way to give an axiomatization for all of mathematics or even for the first-order theory of the integers which is sufficiently strong to prove all true theorems. A theory T is said to be *complete* if, for every sentence ϕ, either $T \vdash \phi$ or $T \vdash \neg\phi$. If neither condition holds, then ϕ is independent of the theory T. (This terminology is standard but a little unfortunate, since the notion of a theory being complete is quite distinct from the notion of the completeness of first-order logic discussed in the previous paragraph.) Gödel originally stated the incompleteness theorems for ω-consistent theories containing as a subtheory the extension of Peano arithmetic to finite type theory, and subsequently noted that it also applies to ω-consistent theories containing first-order Peano arithmetic. Subsequent developments of Rosser, of Tarski, Mostowski and Robinson, and of others has shown that the incompleteness theorems hold for any consistent extension of R. Robinson's theory Q; the theory Q has the constant symbol 0 (zero) and function symbols for successor, addition and multiplication and a small set of axioms defining the basic properties of zero and the three functions. The incompleteness theorem states that there is no deductively closed theory containing Q which is both consistent and decidable. A special case of this is that there is no decidable set of true axioms about the integers which implies all true first-order statements about the integers; likewise, there is no decidable set of consistent axioms for set theory which gives a complete theory. Even more remarkably, for any consistent, recursive axiomatization $T \supseteq Q$ for the integers, Gödel's second incompleteness theorem gives a construction of a true formula Con_T which is unprovable in T. The formula Con_T expresses the fact that T is consistent — in order to construct Con_T, Gödel had to show that metamathematical concepts such as provability and consistency can be formalized appropriately as first-order statements about integers. This formalization was based on what is now called "Gödel numbering", namely, the coding of logical concepts such as formulas and proofs by integers.

As is well-known, Gödel's proof of the incompleteness theorem involved a formula G_T which asserts of itself that it has no T-proof. G_T can be shown to be T-provably equivalent to Con_T for sufficiently strong theories T.

Gödel's consistency proof for the Axiom of Choice (AC) and the Generalized Continuum Hypothesis (GCH) showed that $\mathit{ZF} + \mathit{AC} + \mathit{GCH}$ is consistent provided that the base set theory ZF is consistent. Gödel's proof method involved showing that a subclass of the universe of all sets, the class of "constructible sets," forms a model of $\mathit{ZF} + \mathit{AC} + \mathit{GCH}$. The fact that AC and GCH are independent of ZFC was proved later by P. Cohen.

Gödel also worked on the foundations of intuitionistic logic; most notably, his Dialectica interpretation gave a quantifier reduction procedure which allowed a reduction of intuitionistic first-order arithmetic to a quantifier-free theory of higher-type functionals. G. Kreisel later extended this to a "no-counter example" interpretation for classical logic.

Gödel had wide-ranging interests in philosophy, but here we will discuss only his views on mathematics and human intuition, particularly his opinion of the significance of the (in)completeness theorems to the foundations of mathematics and to the possibility of artificial intelligence. Gödel held a strong Platonistic viewpoint of mathematics — that is, he believed that mathematical objects such as integers, reals and sets have an actual existence independent of human thought. In the posthumously published article based on his 1951 Gibbs lecture, he wrote [2, p.320]

> [Mathematical] concepts form an objective reality of their own, which we cannot create or change, but only perceive and describe.

In other words, Gödel believed that mathematical research is a process of discovery, not a process of creation. Related to his Platonistic philosophy, Gödel believed that human minds and thought processes are not merely algorithmic and that even human *mathematical* intuition surpasses any Turing machine. Of course, Gödel was very concerned to reconcile this view with Turing's definition of algorithms and with the second incompleteness theorem. The problem is that Turing's definition of algorithm is so general that it is difficult to quantify how human mathematical reasoning could surpass an arbitrary Turing machine. On the other hand, if human mathematical cognition is Turing computable, then there are true mathematical statements (e.g., of the form Con_T or G_T, for some appropriate T which simulates human cognition) which are beyond the capabilities of human reasoning. This latter possibility was one which Gödel believed ought to be false.

There have have been several recent, popular accounts of relationships between the incompleteness theorems and artificial intelligence. In particular, D. Hofstadter's popular and influential book *Gödel, Escher, Bach* draws analogies between the self-referential sentences used in the proof of the second incompleteness theorems and the consciousness or self-awareness of humans. More recently, a more dubious argument has been proposed by R. Penrose (similar to an earlier, less-ambitious argument of Lucas) in his books *The Emperor's New Mind* and *Shadows of the Mind*. Penrose argues that Gödel's second incompleteness theorem provides a proof that human minds are not algorithmic and that therefore artificial intelligence is impossible, even *in principle,* for Turing machines. Indeed, Penrose claims that his argument against artificial intelligence has rigor close to the level of mathematical proof. Penrose's proof has been widely criticized and we'll discuss one problem in his proof below; however, first, it is interesting that Gödel examined exactly the same issues in the Gibbs lecture article. Gödel came to the conclusion, which he calls "mathematically established," that [2, p.310]

> Either ... the human mind (even within the realm of pure mathematics) infinitely surpasses any finite machine, or else there exist absolutely unsolvable diophantine problems ...

Penrose argues that the second case cannot hold (or rather, argues that the second case cannot hold for the unprovable sentences obtained from the second incompleteness theorem). Gödel, however, would presumably not have accepted this argument.

In this reviewer's opinion, both Penrose and Gödel are making a subtle, but serious error; namely, they fail to adequately account for the fact that humans, unlike Turing machines, do not have unequivocally distinguishable "accepting states" for deciding mathematical truth. In other words, humans may make errors in mathematical reasoning, which may later be corrected.[1] For this reason, the Turing-Gödel undecidability/incompleteness phenomena do not apply to humans in the same way that they apply to Turing machines. From these considerations, we conclude that the Gödel incompleteness theorems cannot be used to resolve the question of whether human reasoning is algorithmic or of whether true artificial intelligence is possible.

The book under review provides a comprehensive and detailed account of Gödel's life and accomplishments. The author, John Dawson, was responsible for the multiyear project of cataloguing Gödel's papers, the so-called *Nachlass.* Gödel saved nearly all his paperwork and his Nachlass contains "library request slips, luggage tags, crank correspondence, ..." as well as copies of his mathematical correspondence and draft copies of Gödel's letters and even unsent letters. Thus Dawson was able to draw on a wealth of knowledge and documentation about Gödel's life. The book also draws heavily from other sources, most notably Oskar Morgenstern's diaries covering the last thirty years of Gödel's life, which contain a new and illuminating viewpoint of Gödel.

Dawson's book is accessible to a reader with little mathematical knowledge. It contains an interesting account of the logical research leading up to Gödel's completeness and incompleteness theorems, and a brief history of the development of axiomatic set theory. It unfortunately contains only a cursory treatment of the technical aspects of Gödel's theorems and their proofs, and for this the reader will have to seek elsewhere. There are a number of excellent textbooks in mathematical logic which fully cover Gödel's completeness and incompleteness theorems; in addition, popular, readable yet technically correct accounts are given by Nagel and Newman [3] of the incompleteness theorems and by Crossley et al.[1] of the completeness and incompleteness theorems and the consistency of the axiom of choice and the continuum hypothesis. Finally, readers who are interested in reading original papers by Gödel can find his published and unpublished works in the *Collected Works*, of which three have appeared so far. (Dawson is also one of the editors for the *Collected Works* volumes.)

Dawson handles the delicate matter of Gödel's mental health straightforwardly and without undue sensationalism. We learn that Gödel's promotion to full professor at the IAS was delayed because of members' concerns about his mental stability and fears that his legalistic outlook would impede the delibera-

[1] Penrose addresses this point at length in *Shadows of the Mind*, but his arguments are unconvincing.

tions of the IAS faculty. Gödel's paranoia, his fears that certain mathematicians wanted to assassinate him, his hypochondria, his chronic fear of being poisoned, and his distrust of doctors are all documented. These negative character traits are balanced not only by Gödel's mathematical brilliance and life-long work in logic, philosophy and physics, but also by Gödel's relationships with his wife, with his mother, with colleagues such as Morgenstern and Einstein, and with younger logicians. The net result is a balanced and sympathetic portrait of Gödel's life and personality.

Dawson's biography of Gödel is provocative and interesting on several fronts, and is highly recommended to anyone with an interest in logic, the foundations of mathematics or the history of mathematics.

Samuel R. Buss
Department of Mathematics
University of California, San Diego

[1] J.N. Crossley, C.J. Ash, C.J. Brickhill, J.C. Stillwell and N.H. Williams, *What is Mathematical Logic?*, Oxford University Press, 1972, 82 pp.

[2] K. Gödel, Some basic theorems on the foundations of mathematics and their implications, in K. Gödel, *Collected Works, Volume III: Unpublished Essays and Lectures*, S. Feferman, J.W. Dawson Jr., W. Goldfarb, C. Parsons, and R.N. Solovay, eds., Oxford University Press, 1995, pp. 290–323. Includes an introductory note by G. Boolos.

[3] E. Nagel and J.R. Newman, *Gödel's Proof*, New York University Press, 1958, 118 pp.