The Review of Modern Logic
Volume 9 Numbers 3 & 4 (December 2003–August 2004) [Issue 30], pp. 181–190.

Egon Börger, Erich Grädel, Yuri Gurevich
The Classical Decision Problem
Berlin/Heidelberg/New York: Springer-Verlag (Universitext), 2001
x + 482 pp. ISBN 3540423249

REVIEW

R. GREGORY TAYLOR

The Classical Decision Problem first appeared in a 1997 hard-cover edition within the Springer series *Perspectives in Mathematical Logic*. The book under review is a soft-cover reissue within Springer's *Universitext* series.

The decision problem for first-order predicate logic—Hilbert's *Entscheidungsproblem*—is this: Does there exist an effective procedure for deciding whether an arbitrary first-order sentence S is logically valid or, alternatively, whether S is satisfiable. These alternative formulations are equivalent given that S is valid if and only if $\neg S$ is unsatisfiable. Assuming that effectiveness is captured by the technical notion of partial recursive function (the Church–Turing Thesis), it was shown in the mid-1930s by Church and also by Turing that there is no effective decision procedure of the desired sort (Church–Turing Theorem). In light of this negative result, one proceeds to ask about subclasses C of the collection of all first-order sentences, although, in general, a procedure for deciding validity for sentences in C might exist in the absence of a procedure for satisfiability for C. Similarly, demonstrating that there is no decision procedure for validity for C is compatible with the existence of a decision procedure for satisfiability for C. In any case, the *Entscheidungsproblem* is now recast as a classification problem: Which subclasses are decidable for satisfiability (validity) and which are undecidable?

This classification problem is now completed—at least if one considers standard subclasses only, *i.e.*, those determined by quantifier prefix and by which predicate and function symbols may occur. After an introductory first chapter, Part I of the book, comprising Chapters 2–5, is concerned with (minimal) undecidable subclasses, where "undecidable" tends to mean "no decision procedure for satisfiability." Part II

(Chapters 6–8) takes up the search for (maximal) decidable subclasses and has the complexity of decision procedures as a major focus. We summarize each of Chapters 2–8 below.

Chapter 2, entitled simply "Reductions," presents what is essentially Turing's proof of the unsolvability of the *Entscheidungsproblem*, which then serves as a model for the various undecidability proofs appearing later in Part I. The general pattern of such proofs is as follows.

Step 1. Choose a known undecidable problem $\mathcal{P}$.
Step 2. Present a technique such that, given an arbitrary instance π of $\mathcal{P}$, one can effectively construct a *reduction sentence* S_π.
Step 3. Prove that π has a positive solution just in case S_π is satisfiable (or valid or contradictory).
Step 4. Conclude that the satisfiability (or validity or unsatisfiability) problem for the class of reduction sentences S_π is undecidable.

In the case of Turing's proof, Steps (1)–(4) above become

Step 1′. Consider the Halting Problem for Turing Machines, already known to be unsolvable assuming the Church–Turing Thesis.
Step 2′. Present an effective technique such that, given arbitrary Turing machine M, one constructs a first-order reduction sentence ρ_M.
Step 3′. Demonstrate that M, starting in standard initial configuration, halts just in case ρ_M is contradictory.
Step 4′. Conclude that the satisfiability problem for the class $\{\rho_M | M \text{ is a Turing machine}\}$ is undecidable, from which the undecidability of the *Entscheidungsproblem* follows immediately.

An exercise on page 23 points the reader in the direction of an alternative proof from the undecidability of the general word problem for Thue processes. Chapter 2 also introduces one of the overall themes of the book, namely, that the combinatorial structure of the reduced decision problem $\mathcal{P}$ is mirrored in the syntactic features of the class of reduction sentences.

Showing a class to be undecidable frequently involves showing it to be a so-called *reduction class*. Suppose that, for class C of sentences, we possess an algorithm (or *reduction procedure*) $\mathcal{A}$ that transforms an arbitrary first-order sentence S into sentence $\mathcal{A}(S)$ in C such that S is satisfiable (valid) if and only if $\mathcal{A}(S)$ is. Then C is said to be a *reduction class for satisfiability (validity)*. (Note that a reduction procedure for satisfiability yields a reduction procedure for validity, thus restoring duality.) Intuitively, deciding satisfiability (validity) for such a class C is as hard as deciding it in general. So reduction classes are undecidable classes of maximal computational complexity in this sense.

The first result considered concerning reduction classes, due to Aanderaa and Börger, concerns the class $\exists\forall\exists\forall \cap$ KROM $\cap$ HORN with binary predicate symbols only, where the utilized reduction sentence S_ψ for first-order ψ is a logical description of the behavior of a certain deterministic two-register machine M for input some numeric encoding of ψ. "$\exists\forall\exists\forall$" speaks to the prefix of any member, assumed to be in prenex normal form, of the reduction class and "KROM $\cap$ HORN" to its quantifier-free matrix. This result in turn yields a proof of a proposition, attributed to Kreisel, stating that there exists no recursive "interpolation function." (Such a function, given first-order formulae ψ and ϕ with equality where $\psi \models \phi$, returns an interpolant for ψ and ϕ.) The formalization of instructions of two-register machines involved in the Aanderaa-Börger Theorem (Theorem 2.1.15) plays a rôle at later points in the text, *e.g.*, in the Chapter 5 proof of the Lewis–Goldfarb result mentioned below.

Trakhtenbrot's Theorem, which says that Fin-sat(FO), Non-sat(FO), Inf-axioms(FO) are pairwise recursively inseparable sets of first-order sentences, is shown to follow from the recursive inseparability of two sets of Turing machines related to the Halting Problem. (An infinity axiom is a satisfiable first-order sentence having no finite models.) This transfer of recursive inseparability is effected by a certain reduction property involving a reduction sentence as in $(2')$. So the proof method here is a variant of Turing's method of logical description. On page 17, Trakhtenbrot's Theorem is described as the analogue of Church–Turing for finite satisfiability, which makes sense given that X, Y recursively inseparable implies X, Y not recursive. The more usual presentation of Trakhtenbrot's Theorem—no decision procedure for Fin-sat(FO)—is thus implicit in the page 17 remark. The logical descriptions chosen in the proof of Trakhtenbrot's Theorem do not fix their models up to isomorphism. For example, no part of the description of Turing machines states that, at any time t, the machine described is in one and only one state at t or that one and only one tape square is being scanned. It is the authors' contention that these incomplete descriptions facilitate simpler proofs and smoother complexity correlations between machine computations and simulating logical deductions.

The concept of a *conservative reduction* is introduced, whereby one intends a mapping that (two-way) preserves both satisfiability as well as finite satisfiability. $\exists\forall\exists\forall \cap$ KROM $\cap$ HORN and $\forall\exists\exists\forall \cap$ KROM $\cap$ HORN with binary predicate symbols only are mentioned as examples of *conservative reduction classes.* The relationship between logical expressability and computational complexity is investigated beginning in

§2.2.1. Using sentences of propositional logic to describe Turing machine computations, the authors give a quick proof of the Cook–Levin Theorem.

Trakhtenbrot's Theorem quite naturally leads one to ask after the cardinalities of the models of the various sentences ψ of Fin-sat(FO). The class of these natural numbers is the *spectrum* of ψ. Since any model of a second-order (SO) sentence having no nonlogical vocabulary is just a domain, there is a natural correspondence between the spectra of first-order sentences and the class of finite models of existential second-order sentences in which no nonlogical predicate constants occur. So such a class of models, assumed closed under isomorphism, may itself be regarded as a spectrum. By permitting nonlogical predicate constants to occur in existential second-order sentences, one obtains the notion of *generalized spectrum*. The celebrated theorem of Fagin is proved, stating that generalized spectra are precisely the (isomorphism-closed) classes K of finite structures of nonempty signature that are accepted by some nondeterministic Turing machine in time polynomially bounded in the size of their respective domains. Alternatively, Σ_1^1-SO is said to *capture* NP. Complexity class P has been shown to be captured in a weaker sense by various classes of sentences (logics). Namely, if one restricts attention to classes K of successor structures, then SO-HORN and Σ_1^1-HORN both capture P. This is also true of least-fixed-point logic FO + LFP for ordered structures.

The last part of Chapter 2 is given over to a justification of the standard classification scheme, in terms of quantifier prefixes and nonlogical vocabulary, used to organize the nondenumerable realm of subclasses of first-order sentences. *Generalized prefix sets* are collections of words over the four-letter alphabet $\{\forall, \exists, \forall^*, \exists^*\}$. Such a set is *closed (downward)* if it contains every (not necessarily contiguous) substring of any member. It then turns out that any nontrivial ("does not contain every prefix") closed prefix set—set of strings over two-letter alphabet $\{\forall, \exists\}$—is describable as a finite union of maximal generalized prefix sets. For example, the downward closure $\overline{\Pi}$ of the prefix set $\Pi = \{\forall^n\exists^m\forall^k | 0 \leq n \leq m \leq k\} \cup \{\exists^n\forall^m\exists^k | 0 \leq n \leq m \leq k\}$ is just $(\forall^*\exists^*\forall^*) \cup (\exists^*\forall^*\exists^*)$. The entire realm of closed prefix sets is well partial ordered by inclusion. In turn, the entire realm of (not necessarily closed) prefix sets is well quasi-ordered by the (nonantisymmetric) domination relation whereby Π_2 *dominates* Π_1 provided that $\overline{\Pi}_1 \subseteq \overline{\Pi}_2$. The authors go on to describe a well-quasi-ordering of the sequences used to characterize the range of nonlogical predicate and function symbols occurring within classes of first-order sentences. We content ourselves with two examples of the notation used to denote standard classes. First, take the (undecidable) Goldfarb class $[\exists^*\forall^2\exists, (0,1), (0)]_=$, discussed below. Here $\exists^*\forall^2\exists$ means that the prefix of any sentence in the class, assumed to be in prenex normal form, consists of arbitrarily many existential quantifiers followed by at most two universal quantifiers followed possibly by a single existential quantifier. Then $(0,1)$ indicates that the matrix can contain no occurrences of monadic predicate symbols but may have occurrences of (up to) one dyadic predicate symbol (but no n-adic predicate symbols for $n > 2$). Finally, (0) means no function symbols at all. The subscript indicates that occurrences of the equality predicate are allowed. Sentences in the (decidable) Ramsey class without identity $[\exists^*\forall^*, all, (0)]$, or just $[\exists^*\forall^*, all]$, may involve occurrences of arbitrarily many n-adic predicate symbols for any $n \geq 1$, and occurrences of $=$ are not permitted.

The two well-quasi-orderings of the last paragraph together yield a well-quasi-ordering of the realm of prefix–vocabulary classes used throughout the book. In summary, Gurevich's Classifiability Theorem asserts that if downward closed collection $\mathfrak{D}$ of prefix–vocabulary classes is closed under finite union, then $\mathfrak{D}$ and $\mathfrak{D}^c$ have other finiteness properties as well. For instance, suppose $\mathfrak{D}$ is the collection of prefix–vocabulary classes with satisfiability problem in PSPACE. Then $\mathfrak{D}^c$ is the upward closure under the well-quasi-ordering of a finite collection of minimal (downward) closed classes.

Chapter 3 takes up the satisfiability problem for prefix–vocabulary classes of sentences of pure predicate logic—without $=$ or function symbols. Certain *classical solvable cases* were identified very early on, namely, Löwenheim's class $[all, (\omega)]$ of *relational monadic sentences*—(ω) for "arbitrarily many", Ackermann's $[\exists^*\forall\exists^*, all]$, $[\exists^*\forall^*, all]$ of Bernays and Schönfinkel, and $[\exists^*\forall^2\exists^*, all]$, shown decidable by each of Gödel, Kalmár, and Schütte. (These decidable classes, the last of which alone is maximal, are covered in Chapter 6.) A generation later other classes of sentences of pure predicate logic were shown to be undecidable. Regarding the latter, the Classifiability Theorem predicts that there will be just finitely many minimal undecidable (downward) closed prefix–vocabulary classes—let $\mathfrak{D}$ in the preceding paragraph be the collection of all decidable prefix–vocabulary classes. There turn out to be nine such minimal undecidable classes. Two of them—the Kahr class $[\forall\exists\forall, (\omega, 1)]$ and the Gurevich class $[\forall\exists\forall\exists^*, (0,1)]$—are central to the story. First, a certain domino (tiling) problem, known to be undecidable, is described and shown to be reducible to (finite) satisfiability for Kahr's class. Subsequently, Kahr's class is shown reducible to each of five other prefix–vocabulary classes. Finally, Kahr is reducible to Gurevich, in turn reducible to the remaining two classes of the nine. Each reduction is conservative, which means that finite satisfiability for each of the nine prefix–vocabulary classes is undecidable as well. (In an appendix by C. Allauzen and B. Durand, new proofs of the undecidability of several domino problems are presented.)

In Chapter 4 the authors survey the situation with regard to standard classes of sentences of full predicate logic—either with functions but without =, with = but without functions, or with both. Again, the hypothesis of the Classifiability Theorem is met so that a finite classification must be expected—this time there are seven minimal undecidable prefix–vocabulary classes. First, the Halting Problem for Two-Register Machines is reduced to (the Herbrand subclass of) $[\forall, (0), (2)]_=$, in its turn reduced to (the Herbrand subclass of) $[\forall, (0), (0, 1)]_=$. This takes care of classes with both functions and =. Next, $[\forall\exists\forall, (0, \omega)]$, shown undecidable in Chapter 3, is reduced first to two minimal classes with functions but without =. These results are all due to Gurevich. Completing the classification was left to Goldfarb, who in the early 1980s showed that $[\forall\exists\forall, (0, \omega)]$ is reducible, this time, to three minimal classes with = but without functions. The most challenging of these reductions involves the class $[\forall^2\exists, (\omega, 1)]_=$. The interesting history of this reduction class, long assumed decidable due to an erroneous remark in a 1933 paper of Gödel, is reviewed briefly in the historical section at the end of the chapter. All the reductions mentioned here are conservative so that the classification with regard to satisfiability in the finite coincides with that of satisfiability generally.

Chapter 5, entitled "Other Undecidable Cases," surveys results concerning the (finite) satisfiability problem for classes of prenex first-order sentences characterized, nonstandardly, by the structure of their quantifier-free matrices. A first section covers Krom and Horn structure. (Any conjunct within the matrix of a Krom formula, assumed to be in prenex normal form, is a disjunction of at most two literals. Any conjunct within the matrix of a Horn formula is a disjunction involving at most a single positive literal.) In the proof of Theorem 5.1.2, five prefix classes with Krom matrix and without functions or equality are shown to be undecidable—in fact to be reduction classes. In particular, using two distinct ways of modeling machine configurations as atomic formula, (a version of) the Halting Problem for Two-Register Machines is reduced to the Lewis classes $[\forall^2\exists\forall, (0, 0, \omega)] \cap \text{KROM} \cap \text{HORN}$ and $[\forall\exists\forall^2, (0, 0, \omega)] \cap \text{KROM} \cap \text{HORN}$, respectively. By the fact that the Halting Problem is complete for r.e. sets and given that, by Gödel

Completeness, the valid formulae of first-order logic themselves constitute an r.e. set, we have that the Lewis classes are reduction classes for satisfiability.

Section 5.2 is a brief survey of proof techniques and results concerning classes determined by restrictions upon the number of occurring atomic subformulæ. For example, the authors present a proof of the result of Lewis and Goldfarb stating that the class of sentences with prefix in $\forall\exists\forall^*$ and matrix of the form $\pi_1 \wedge \pi_2 \rightarrow \pi_3 \wedge \pi_2 \rightarrow \pi_4 \wedge \neg\pi_5$ is a reduction class, where each of $\pi_1, \ldots, \pi_5$ is an atomic formula involving one and the same k-adic predicate symbol for constant k fixed by the size of the instruction set of some assumed universal two-register machine. A corollary (Exercise 5.2.7) asserts the undecidability of the class of equality- and function-free $\forall^*\exists$-sentences whose matrices are of the form $(\pi_1 \vee \pi_2 \vee \pi_3) \wedge (\neg\pi_4 \vee \neg\pi_5) \wedge (\neg\pi_4 \vee \pi_6)$, where each of $\pi_1, \ldots, \pi_6$ is atomic. The next subsection covers Wirsing's proof that the subclass of $[\forall^6, (0), (0,1)]_=$ comprising all and only sentences with matrices of the form $s = t \wedge u \neq v$ is a conservative reduction class. A related result establishes the undecidability of the subclass of $[\forall^6, (0,0,0,1), (2,1)] \cap$ KROM $\cap$ HORN with matrices conjoining just one atomic formula, one negated atomic formula, and one disjunction of the form $\neg\rho \vee \sigma$, where ρ and σ are both atomic.

Chapter 6 brings us back to standard classes. It tells the complete story with respect to those that are decidable and have the so-called *Finite Model Property*—any satisfiable sentence in the class has a finite model. (In older literature such classes are also said to be *finitely controllable.*) The authors begin by setting to one side the *essentially finite* classes, *viz.*, those involving a finite set of possible prefixes (no $\exists^*$ or $\forall^*$), a finite collection of predicate symbols, and no function symbols. If a linear ordering of CNFs of possible matrices (up to relettering) is assumed, then each sentence in the class can be identified with a sentence with identical prefix and with that one among equivalent CNFs that is minimal in the ordering. But then there are only finitely many cases to consider, and we can assume that a Turing machine "decider" M possesses all the answers (table lookup). Note that, given the peculiar manner in which decision problems are framed, we need not know what the table actually is in the case of a particular essentially finite class in order to assert decidability. Also, worst case, M consumes space logarithmic in the length of its input sentence.

Most of Chapter 6 is taken up with showing that each of (1)–(5) below is decidable for (finite) satisfiability.

(1) Ramsey class $[\exists^*\forall^*, all, (0)]_=$

(2) Gödel–Kalmár–Schütte class $[\exists^*\forall^2\exists^*, all, (0)]$
(3) Löb–Gurevich class $[all, (\omega), (\omega)]$
(4) Gurevich–Maslov–Orevkov class $[\exists^*\forall\exists^*, all, all]$
(5) Gurevich class $[\exists^*, all, all]_=$

Since any recursive class C with the Finite Model Property is such that $\mathrm{Sat}(C)$ is then recursive and since each of (1)–(5) is obviously recursive, we need only see that each of (1)–(5) has the Finite Model Property, which is what Chapter 6 is about.

As for complexity, one can usually extract, from the proof that C has the Finite Model Property, a bound m on the size of a satisfying model of $S \in C$, which in turn yields a bound on nondeterministic computation time that is polynomial in (number of occurring universal quantifiers, model-size bound m and) sentence length n. For example, the proof of Theorem 6.2.43 states that membership in $\mathrm{Sat}([\exists^*\forall^2\exists^*, all, (0)])$ (cf. (2)) is decidable nondeterministically in $2^{O(n/\log\, n)}$ steps, and a matching lower bound is given earlier in Theorem 6.2.13. This puts it outside NP. But, as shown in Theorem 6.2.45, due to Grädel, there is a large family of nontrivial subclasses of $[\exists^*\forall^2\exists^*, all, (0)]$ whose satisfiability problems are in NP.

The final section §6.5 of Chapter 6 gives a classification of prefix–vocabulary classes with respect to the Finite Model Property itself, which is a departure from the "decision problem" theme of the book—but a very interesting one. There are nine maximal classes with the property and ten minimal classes with infinity axioms. (Most, but not all, of these classes figured earlier in the classification with respect to the decision problem for satisfiability.) Since an essentially finite class may contain an infinity axiom, these classes cannot be ignored this time around. Another important difference is that, to date, this classification is incomplete: except in a few cases, it is presently unknown which of the essentially finite subclasses of the Goldfarb class $[\exists^*\forall^2\exists^*, all, (0)]_=$ contain infinity axioms.

Back to decidability, just two maximal decidable standard classes contain infinity axioms, and Chapter 7 is devoted to demonstrating decidability for these two classes. Since the Finite Model Property is not involved, the techniques of Chapter 7 differ greatly from those of Chapter 6. The first class considered is $[all, (\omega), (1)]_=$, and $\exists x \forall y \forall z (fy \neq x \wedge (fy = fz \rightarrow y = z))$, which says that f is injective but not surjective, is an infinity axiom here. The proof of decidability appeals to Rabin's Theorem stating that the <u>s</u>econd-order monadic theory <u>S</u>2S of the infinite binary tree—"2S" for "two successors (children)"—is decidable. That result in turn uses Büchi's proof that the second-order monadic theory S1S of one successor is decidable. (Incidentally, decidability here means that, given any formula $\psi(X_1, \ldots, X_n)$ of the language of S1S, we can effectively determine whether $S1S \models \exists X_1 \ldots \exists X_n \psi(X_1, \ldots, X_n)$.) The treatment is largely self-contained, which means that Büchi automata and Rabin's tree automata are covered. The last half of Chapter 7 is given over to a proof that the Shelah class $[\exists^*\forall\exists^*, all, (1)]_=$ is decidable for satisfiability. This is apparently the only published proof, and Shelah himself advised the authors regarding its presentation.

One uses L_k for the restriction of first-order logic to formulae with predicate symbols but no function symbols and such that at most variables $x_1, \ldots, x_k$ can occur in them. Since L_3, with or without equality, extends the minimal undecidable Kahr class $[\forall\exists\forall, (\omega, 1)]$, it follows that L_k is undecidable for $k \geq 3$. On the other hand, in a paper published in 1975, Mortimer showed that L_2 is decidable—indeed has the Finite Model Property. The authors give what we are told is a new, simpler, proof of this result, and the upper bound $2^{O(n)}$ on (nondeterministic) complexity extractable from their proof is also better. This material is presented at the beginning of the final Chapter 8, entitled "Other Decidable Cases," which otherwise covers nonstandard classes of first-order sentences of interest to those working in computer science and linguistics. Proofs are given for the decidability of the Aanderaa–Lewis class $[\forall\exists\forall] \cap \text{KROM}$ and the Maslov class $[\exists^*\forall^*\exists^*] \cap \text{KROM}$. A host of related complexity results are provided, *e.g.*, the satisfiability problem for the Aanderaa–Lewis class is NLOGSPACE-complete, as shown by Denenberg and Lewis in 1984. Open questions remain here: it is not known whether $[\forall\exists\forall\exists^k] \cap \text{KROM}$ with $k > 0$ and $[\forall\exists\forall\exists^*] \cap \text{KROM}$ are decidable for satisfiability.

It is hoped that the reader now has some idea of what is in *The Classical Decision Problem.* As for the manner in which the material is presented, the authors, in the book's preface, describe their effort as that of "combining the features of a research monograph and a textbook." They suggest that the book—or selected chapters of it—might be used for an introductory course on decision problems, undecidability, and the complexity of decision procedures. Indeed, numerous exercises are provided in every chapter. On the other hand, considerable sophistication regarding automata theory and finite model theory is presupposed. For example, the authors' descriptions of specific machines and models usually amount to mere sketches. So there is usually a lot to think about in making sense of the authors' arguments. This is part of what makes this book so enjoyable. But reading it could be a frustrating experience for a student reader with inadequate preparation.

The book's encyclopedic, exhaustively annotated, bibliography is a particularly useful feature. Each entry includes a brief summary ranging from a single sentence to half a page. The bibliography alone is worth the price of this remarkable book.

Department of Mathematics and Computer Science, Manhattan College, Riverdale, NY 10471
E-mail address: gregory.taylor@manhattan.edu