On the Classical Decision Problem *

Yuri Gurevich †

- *Author:* Hello, my friend. What is on your mind today?

- *Quisani:* Decidable and undecidable fragments of first-order logic. People refer to this field as Entscheidungsproblem or the classical decision problem. I wanted to see a global picture, without going into too many details, and failed. You worked in the field, didn't you? Can you shed some light?

- A: I am surprised. Decidability, you joked the other day, is a red herring.

- Q: I did? Well, there is no doubt in my mind that feasibility is the real issue, but I keep bumping into the classical decision problem. Most recently, this happened when I looked up a paper of Kolaitis and Vardi [KV] on the 0–1 law. (This law is, by the way, another issue I would like to discuss with you sometime.) Besides, the two issues – decidability and feasibility – are obviously related. Undecidability implies nonfeasibility, and nonfeasibility proofs (e.g. proofs of completeness for NP or exponential time) are often fashioned after undecidability proofs.

  Speaking about surprises, I was surprised too. Apparently, the classical decision problem was tremendously popular among logicians. Even Gödel worked on it. This puzzles me. Logicians are so philosophically minded. Why all that interest in what seems to be a rather technical question?

- A: Let me start from the beginning. The original Entscheidungsproblem was posed, I guess, by Hilbert. It may be stated as a satisfiability or validity problem: Given a first-order formula ϕ, decide whether ϕ is satisfiable (respectively, valid). Proof theorists usually prefer the validity version whereas model theorists prefer the satisfiability version. I am more used to the satisfiability version; let me choose it to be the default.

- Q: What first-order formulas are you talking about? Do you allow equality, function symbols? They make a big deal out of such details in that field.

*Logic in Computer Science Column, The Bulletin of EATCS, October 1990

†Partially supported by NSF grant CCR 89-04728. Address: EECS Department, University of Michigan, Ann Arbor, MI 48109-2122, USA

- A: These details do not matter for the original Entscheidungsproblem, but they will matter later, so let us fix a version of first-order logic without equality or individual constants or function symbols. (First-order logic comes with a deduction mechanism, but details of the deduction mechanism will be irrelevant for our purposes.) Without loss of generality, we may restrict attention to sentences, i.e., formulas without free individual variables. Let me also clarify the terminology. Recall that the collection of predicates (i.e. relation symbols) of a sentence ϕ is its *signature*. A sentence ϕ of some signature σ is *satisfiable* if there exists a structure of signature σ that satisfies ϕ, and ϕ is *valid* (or logically true) if every σ-structure satisfies ϕ. It is easy to see that the two versions of Entscheidungsproblem are easily reducible each to the other. Hilbert called Entscheidungsproblem "the fundamental problem of mathematical logic" [DG].

- Q: Sounds very important indeed.

- A: At the time the notion of algorithm was not formalized. Algorithms usually meant feasible algorithms, I think. Just imagine you have a feasible decision algorithm for Entscheidungsproblem. You would be able to solve numerous mathematical problems including those most famous.

- Q: Name one.

- A: The great "theorem" of Fermat. Its negation is expressed by an existential sentence $\phi = (\exists x, y, z, u)\Phi(x, y, z, u)$ in the language of Peano Arithmetic where $\Phi(x, y, z, u)$ states that x, y, z are positive and $u > 2$ and $x^u + y^u = z^u$. Choose a finitely axiomatizable fragment PA_0 of Peano Arithmetic sufficiently rich to prove $\Phi(a, b, c, n)$ or $\neg\Phi(a, b, c, n)$, whichever is true, for any specific quadruple a, b, c, n of natural numbers, and let α_0 be the conjunction of axioms of PA_0. It is easy to see that the great "theorem" fails if and only if the implication $\alpha_0 \rightarrow \phi$ is valid. Now use your algorithm.

- Q: Sorry, I do not remember exactly what Peano Arithmetic is. Is it well known that there exists a finitely axiomatizable fragment of Peano Arithmetic sufficiently rich for our purpose?

- A: Peano Arithmetic is a standard first-order formalization of the arithmetic of natural (i.e. nonnegative integer) numbers. It is described in many logic textbooks; see Kleene's book [Kl] for example. It has a small number of specific axioms and one axiom schema that formalizes the induction principle. One well-known finitely axiomatizable fragment of Peano Arithmetic sufficiently rich for our purpose is Robinson's system [Kl] formulated by Raphael Robinson.

- Q: All right. Allow me to check that the great "theorem" fails if and only if the implication $\alpha_0 \rightarrow \phi$ is valid. First I suppose that the great "theorem" fails and a quadruple a, b, c, n is a counter-example. Then $\Phi(a, b, c, n)$ is true, PA_0

proves $\Phi(a, b, c, n)$ and PA_0 proves ϕ. Hence the implication $\alpha_0 \to \phi$ is valid. Next I suppose that the implication $\alpha_0 \to \phi$ is valid. Then PA_0 proves ϕ, and – if PA_0 is consistent – ϕ is true, and the great "theorem" fails. Fine. You need the consistency of PA_0, but natural numbers with usual arithmetical operations (which you probably want to represent by relations) form, I understand, a model for Peano Arithmetic and therefore for PA_0, so there is no problem there.

I like your argument independently of the decidability issue. It shows that the great "theorem" fails if and only if its negation is provable in a fragment of Peano Arithmetic. Thus, proving that the great "theorem" is independent from, say, Peano Arithmetic would mean that it is true. This is interesting.

- A: This is not my argument of course. It is folklore. Notice that the argument uses only that the great "theorem" is expressible by universal sentences of Peano Arithmetic. The Riemann hypothesis is expressible by a universal sentence of Peano Arithmetic too though this is not obvious at all [DMR]. It follows that the Riemann hypothesis fails if and only if its negation is provable in Peano Arithmetic. The decision algorithm for first-order logic would decide the Rieman hypothesis as well. Of course, the applicability of the decision algorithm would not be restricted to problems expressible by a universal sentences of Peano Arithmetic.

- Q: You have made your point. What happened after Hilbert posed the problem?

- A: The classical decision problem was indeed very popular with logicians. There were plenty of positive and negative results [Ch2]. Some good mathematics was done along the way too. For example, Ramsey's Theorem, so popular in combinatorics, was proved in a paper related to a case of the classical decision problem.

- Q: Wait, I thought you were still talking about the period *before* the formalization of algorithm. How could one prove negative results at that time?

- A: The same way that many negative complexity results are proved today. The key word is "reduction". A class K of sentences is called a *reduction class* (for satisfiability) if there exists an algorithm that, given an arbitrary sentence ϕ, produces a sentence ϕ' in K such that ϕ' is satisfiable if and only if ϕ is. K is called decidable (for satisfiability) if the satisfiability problem SAT(K) – given a sentence in K, decide whether it is satisfiable – is decidable.

- Q: Let me see. We deal with total recursive reductions. Is it true that the decision problem for any recursively enumerable set reduces to the Entscheidungsproblem?

- A: Absolutely.

- Q: Then the satisfiability problem for any reduction class is complete for recursive enumerability with respect to total recursive reductions. I realize that these notions were not known at the time, and you don't need these notions to understand that the satisfiability problem for a reduction class is as difficult as the whole Entscheidungsproblem.

- A: Right. When Church and Turing formalized the notion of algorithm and proved the undecidability of the original Entscheidungsproblem [Ch1, Tu], reduction classes were proven to be undecidable. (More exactly, the satisfiability problem for any reduction class is undecidable.) But the field did not die, though the focus shifted. The classical decision problem became sort of a metaproblem: Which fragments of first-order logic (more exactly, which classes of sentences) are decidable?

- Q: Why didn't the field die? The original Entscheidungsproblem was a specific question. Church and Turing answered the question. I guess, it took some time for the Church-Turing thesis to sink in and become accepted. But why didn't the field die after that? Why did the metaproblem attract attention after that?

- A: First of all, I doubt that Hilbert saw the original Entscheidungsproblem as a yes-no question. He might think about an open-ended problem of mechanizing mathematics. The ambitious attempt to mechanize mathematics via a decision algorithm for first-order logic failed. Does this mean that the field should be abandoned? Of course not. One should try to see what can be done. It is natural to try to isolate special cases of interest where mechanization is possible. There are many ways to define the syntactic complexity of logic formulas. It turns out that, with respect to some natural definitions, sentences of low syntactic complexity suffice to express many mathematical problems. For example, many mathematical problems can be formulated with very few quantifier alternations. The decision problem for such classes of sentences is of interest.

- Q: Why do you speak about mathematics only. These days logic is widely used in computer science as well.

- A: You are absolutely right, thank you. This is a very good point.

- Q: I can see that decidable cases may be of interest, but then the issue of the complexity of decision algorithms arises.

- A: Of course. Also, instead of decidability, one can speak about tractability in one sense or another. But this is a separate issue. Let me resume the attempt to justify the metaproblem.

  Undecidable classes may be of interest too. (Recall that undecidability means nonfeasibility as well.) Of most interest are undecidable classes that are minimal in some appropriate sense. (I will return to minimal undecidable classes.)

They delineate the realm of decidability. Minimal undecidable classes may be seen also in the context of a broader study of minimal combinatorial bases for computation.

Tradition is probably another justification for the metaproblem. When Church and Turing solved the original Entscheidungsproblem, the field had a life of its own. Too much was invested and achieved, too much tradition was involved. Many open problems remained.

There are also direct and indirect applications of results and methods of the field and, as you mentioned yourself, the classical decision problem pops up from time to time in some seemingly unrelated areas like the 0-1 law. I do not want to overdo my point. The days when the classical decision problem was in the center of logicians' attention are long gone. Still, it is an important problem, I think.

- Q: What was your own motivation for working on the metaproblem?

- A: My motivation, I am afraid, was not philosophical. The Ural University, my *alma mater*, had no logic tradition. You may say that I came to logic by accident: A friend of mine gave me Kleene's "Introduction to Metamathematics" [Kl] as a birthday present. I was fascinated and awed and wanted badly to work in logic. The classical decision problem was considered very important by many Soviet logicians and, I guess, I accepted the importance of the problem without questioning it.

- Q: I see a problem with your metaproblem: There are continuum many different sets of sentences. Obviously, you don't want to consider all of them. This would occupy you for a while. Let me analyze the situation a little; my attention wanes if I keep listening quietly for too long. If we want to cut a nice decision problem out of the metaproblem, we should restrict attention to classes presentable, in some fixed way, by constructive objects. One possibility is to consider finite classes. But no, this case is degenerate: Of course, there exists a decision algorithm, given by a finite table, for any particular finite class.

  How about exploiting the entailment relation? For each sentence α, let K_α be the class of implications $\alpha \to \phi$. The validity problem for K_α can be reformulated as follows: Given a sentence ϕ, decide whether ϕ is a consequence of α. (Sorry, here validity is more appropriate than satisfiability.) This gives rise to the following fragment of the metaproblem which is a perfect decision problem all by itself: Given a sentence α decide whether the validity problem for K_α is decidable. I love the sound of it, but suspect that this problem is undecidable. The desired decision algorithm would probably do too much mathematics: α may incorporate (have as conjuncts) axioms of groups or fields, etc. Some people may lose their jobs.

- A: You are right, this problem is undecidable.
- Q: Is the proof difficult?
- A: It depends where do you start. Do you know Tarski's notion of essentially undecidable theories?
- Q: No.
- A: A theory is essentially undecidable if every consistent extension of it obtained by adding finitely many axioms is undecidable; Robinson's system, mentioned above, is essentially undecidable [Kl]. Thus, there exists a satisfiable sentence α (e.g. the conjunction of Robinson'a axiom) such that, for every sentence ϕ, the conjunction $\alpha \wedge \phi$ is satisfiable if and only if the validity problem for $K_{\alpha\wedge\phi}$ is undecidable. In particular, the validity problem for K_α is undecidable, but this problem is reducible to your problem. Take any ϕ. It is a consequence of α if and only if $\alpha \wedge \neg\phi$ is inconsistent if and only if the validity problem for $K_{\alpha\wedge\neg\phi}$ is decidable.
- Q: I see. So which cases of the classical decision problem were attacked by logicians?
- A: Logicians were interested in classes given by simple syntactic restrictions. In 1915, Löwenheim gave a decision procedure for the satisfiability of sentences with only unary predicates and proved that sentences with only binary predicates form a reduction class; the negative result was sharpened by Herbrand in 1931: 3 binary predicates suffice, and by Kalmar in 1936: 1 binary predicate suffices. You can find references in [Ch2].
- Q: Interesting. This may explain why graph theory is difficult.
- A: The first-order theory of one binary relation easily reduces to the first-order theory of one binary relation that is symmetric and reflexive. Thus the theory of graphs is undecidable.
- Q: You mentioned quantifier alternations above.
- A: Yes, this brings us to prefix classes. Recall that every sentence can be written in the prenex form, i.e., with all quantifiers up front. For example,

  $$\forall x(\exists y R(x,y) \wedge \exists z R(z,x))$$

  is equivalent to

  $$\forall x \exists y \exists z (R(x,y) \wedge R(z,x)).$$

  In 1920, Skolem showed that $\forall^*\exists^*$ sentences, i.e., prenex sentences with quantifier prefixes $\forall \dots \forall \exists \dots \exists$ form a reduction class. In 1928, Bernays and Schönfinkel gave a decision procedure for the satisfiability of $\exists^*\forall^*$ sentences. In the same year, Ackerman gave a decision procedure for $\exists^*\forall\exists^*$ sentences. Independently, Gödel, Kalmar and Schütte (in 1932, 1933 and 1934 respectively), published decision procedures for the $\exists^*\forall^2\exists^*$; Gödel proved also that $\forall^3\exists^*$ sentences form a reduction class. Again, references can be found in [Ch2].

- Q: This seems to cover all prefix classes. No, you may have something like $\exists^*\forall\exists\forall^{17}$. Also, can a class given by one specific prefix be a reduction class?

- A: Yes. Suranyi proved that $\forall^3\exists$ sentences form a reduction class. In his 1959 book [Su], he summarized a huge work on reduction classes given by restrictions on the prefix and/or the signature. In 1962, Büchi found an amazingly simple proof of the Church-Turing Theorem which established that $\forall\exists\forall\exists$ sentences form a reduction class. In the same year, Kahr, Moore and Wang sharpened his result: $\forall\exists\forall$ suffices. See references in [Le]. This takes care of all prefix classes.

  For, let Π be an arbitrary set of prefixes and K be the class of all sentences with prefixes in Π. If one of those prefixes contains $\forall\exists\forall$ as a subprefix (not necessarily contiguous subprefix) then, by Kahr-Moore-Wang's Theorem, K is a reduction class. We can suppose therefore that universal quantifiers form a contiguous block in any prefix π in Π. If that block has at least 3 universal quantifiers and is followed by an existential quantifier then, by Kahr's Theorem, K is a reduction class. Thus, we can further suppose that every prefix in Π is of the form $\exists^i\forall^2\exists^j$ or $\exists^i\forall^j$. Since the $\exists^*\forall^2\exists^*$ class and the $\exists^*\forall^*$ are decidable, K is decidable.

- Q: This is beautiful: All one has to remember is 2 undecidable or 2 decidable prefix classes.

- A: There is, by the way, sort of an *a priori* reason for the possibility of a complete solution of the decision problem for prefix classes. Instead of decidability, you may speak about easy decidability in one sense or another, you may speak about the 0–1 law as in the paper of Kolaitis and Vardi, etc. All I need that the collection of good classes (decidable, easy decidable, etc) is closed downward: A subclass of a good class is good. Then problem of characterizing good classes has a complete solution. You asked me to shed some light on the jungle of results on the classical decision problem. I think this may shed some light.

- Q: You sound suspiciously enthusiastic about this *a priori* possibility of a complete solution. Is this your own result?

- A: Yes, I developed a whole theory around it [Gu1], but the central notion of that theory turned out to be discovered and rediscovered many times before me.

- Q: What is it?

- A: It is actually a very useful notion of well partially ordered sets; I called them tightly ordered. A partially ordered set is *well partially ordered* (shortly, *wpo*) if every nonempty subset has at least one, but only a finite number, of minimal elements. A wpo set has no infinite descending chains, and every collection of incomparable elements of a wpo set is finite.

- Q: Well ordered sets are wpo. Can you give me substantially different examples?

- A: Quantifier prefixes form a wpo set under the following order: $\pi_1 \leq \pi_2$ if π_1 is a not necessarily contiguous substring of π_2. This is a simple example of a much more general phenomenon [Kr]. Further, call a prefix set Π a *prefix type* if it is closed under (not necessarily contiguous) subprefixes. In other words, Π is a prefix type if and only if it contains all prefixes π_1 such that $\pi_1 \leq \pi_2$ for some $\pi_2 \in \Pi$. The collection of prefix types, ordered by inclusion, is wpo.

- Q: I think I see your point. Suppose we split prefix types into good and bad. For example, good types are those that give a class of sentences decidable in some suitable sense. Since types are well partially ordered, there exists a finite collection of minimal bad types. Any bad type Π includes a minimal one, doesn't it?

- A: Of course. Just consider the collection of bad subtypes of Π. It is nonempty and therefore contains a minimal member.

- Q: Thus, the collection of minimal bad types gives a complete solution to the problem of characterizing bad types. This is your point, isn't it?

- A: Yes.

- Q: I have a nasty thought. What if some of the minimal bad types do not lend themselves to a nice description?

- A: Fortunately, this is not the case. Call a prefix type *special* if it contains all prefixes or can be described by a string in the alphabet $\forall, \exists, \forall^*, \exists^*$ where $\forall^*$ (resp. $\exists^*$) stands for a block of $\forall$'s (resp. $\exists$'s) of arbitrary length. It is not difficult to check that the union of every ascending chain of a special types is special. Thus, every good special type is included in some maximal good special type. The maximal good special types are incompatible and therefore – since prefix types are well partially ordered – the number of maximal good special types is finite. Thus, good types can be nicely characterized directly.

  Further, every type Π is the union of finitely many special types, namely the maximal special subtypes of Π. In particular, every minimal bad type is the union of a finitely many special types. Further, if we suppose that, as it often happens, the union of any two good types is good then every minimal bad type is special.

- Q: Do signature classes form a wpo set?

- A: Yes. Also, prefix-signature classes, ordered in a natural way (allow me to skip the exact definition of that partially ordered set) form a wpo set. It turns out that there are 9 minimal undecidable prefix-signature classes. Two of the 9 minimal classes are obtained from the two minimal undecidable prefix classes by restricting the signature to at most one binary predicate and arbitrarily many unary predicates. In the case of the remaining 7 classes the signature comprises one binary predicate. Two of the 7 prefix types are $\forall^*\exists\forall$ and $\forall\exists\forall^*$. Two other prefix types are $\exists^*\forall^3\exists$ and $\forall^3\exists^*$. The remaining three prefix types are $\exists^*\forall\exists\forall$, $\forall\exists^*\forall$ and $\forall\exists\forall\exists^*$. The history of this classification is described in [Le]; I happen to be the one to complete it.

- Q: OK. I see the pattern. What about maximal decidable classes?

- A: Let me exclude classes given by a finite prefix type and a finite signature; in the absence of function symbols, the satisfiability problem for such a class is decidable in a trivial way. Any remaining prefix-signature class is decidable if and only if it is included in the union of the two maximal decidable special prefix classes (namely, the $\exists^*\forall^*$ class and the $\exists^*\forall^2\exists^*$ class) and the one maximal decidable signature class (namely, the class of sentences with unary predicates).

- Q: Great. Did you develop your version of the theory of wpo sets first and then tried to complete the classification of prefix-signature classes?

- A: No, I wasn't that smart. Only when the classification of prefix-signature classes was completed, I started to wonder why the completion was at all possible.

- Q: But did you apply the theory of wpo sets to decidability questions?

- A: In a sense. It encouraged me to look into the variants of classical decision problem when equality or function symbols are allowed. That same paper [Gu1] on wpo sets gives also a complete characterization of decidable and undecidable fragments of first-order logic with function symbols but without equality (though the decidability of the class of $\exists^*\forall\exists^*$ sentences was proved in a separate paper).

- Q: What about the case with equality and without function symbols? I can see that function symbols matter, but does the equality matter? I would be surprised if it does.

- A: The question is whether the three decidable classes remain decidable. For two of them, the known decision procedures worked well in the case of equality as well. The one remaining class is the prefix class $\exists^*\forall^2\exists^*$. Gödel wrote that his decision procedure should work also in the case of equality. This turned out to be not so obvious. It was settled only in 1984, when Goldfarb proved that, in the presence of equality, the prefix class $\forall^2\exists$ is undecidable; moreover, the class of $\forall^2\exists$ sentences with one binary predicate and arbitrarily many unary predicates and the class of $\forall^2\exists^*$ sentences with one binary predicate are undecidable [Go].

- Q: Finally, what about the case when you have equality and function symbols?

- A: I was able only to settle this case modulo a conjecture that the class of sentences with one unary function symbol and arbitrary predicates given by the prefix type $\exists^*\forall\exists^*$, is decidable. The conjecture was proven in 1977 by Saharon Shelah [Sh].

- Q: I am not sure I can take any more of this stuff today. Allow me just a couple of quick questions. First, you did not say anything about finite models. You are a great fan of finite models, aren't you?

- A: Trakhtenbrot proved that the set of sentences satisfiable on finite models is undecidable [Tr]. All results mentioned above remain true if satisfiability is replaced by satisfiability on finite models.

- Q: I understand there are numerous cases not covered by results above.

- A: You bet. Among most important cases not covered by the results above, I would mention the decision problem for Horn formulas and Krom formulas. In this connection, see Börger's book [Bo] and relevant references there. I mentioned already the book [Le] of Lewis. Another important book is [DG]. I have reservations about it [Gu3]; it gives a view of the field which is too – what is a right word? – idiosyncratic. But it is an important and useful book.

Acknowledgement. It is a pleasure to thank Andreas Blass, Kevin Compton, Warren Goldfarb, Harry Lewis and Moshe Vardi for commenting on a draft of this paper. In particular, their comments inspired and gave arguments for the discussion on the value of the classical decision problem after Church's and Turing's results.

References

[Bo] "Computability, Complexity, Logic", North-Holland, Amsterdam, 1989.

[Ch1] Alonzo Church, "A note on Entscheidungsproblem", J. Symbolic Logic 1 (1936), 40–41; correction 1 (1936), 101–102.

[Ch2] Alonzo Church, "Introduction to Mathematical Logic", Princeton University Press, Princeton, N.J., 1956.

[DG] Burton Dreben and Warren D. Goldfarb, "The decision problem: Solvable classes of quantificational formulas", Addison-Wesley, Reading, MA, 1979.

[DMR] Martin Davis, Yuri Matijasevich and Julia Robinson, "Hilbert's tenth problem. Diophantine equations: Positive aspects of negative solutions", Proceedings of Symposia in Pure Mathematics 28, American Mathematical Society, Providence, RI, 1976, 323–378.

[Go] Warren D. Goldfarb, "The unsolvability of the Gödel class", JSL 49 (1984), 1237–1252.

[Gu1] Y. Gurevich, "The decision problem for logic of predicates and operations", Algebra and Logic 8 (1969), pages 284–308 of Russian original, pages 160–174 of English translation.

[Gu2] Y. Gurevich, "The decision problem for standard classes", Journal of Symbolic Logic 41 (1976), 460–464.

[Gu3] Y. Gurevich, A review of two books on the decision problem, Bulletin of American Mathematical Society 7 (1982), 273–277

[Kl] Stephen Cole Kleene, "Introduction to Metamathematics," North-Holland, Amsterdam, 1952.

[Kr] Joseph B. Kruskal, "The theory of well-quasi-ordering: A frequently discovered concept", J. of Combinatorial Theory 13, no. 3 (1972), 297–305.

[KV] Phokion G. Kolaitis and Moshe Y. Vardi, "0–1 laws and decision problems for fragments of second-order logic", Information and Computation 87 (1990), 302–338.

[Le] Harry R. Lewis, "Unsolvable classes of quantificational formulas," Addison-Wesley, Reading, MA, 1979.

[Sh] Saharon Shelah, "Decidability of a portion of the predicate calculus", Israel J. Math. 28 (1977), 32–44.

[Su] János Surányi, "Reduktionstheorie des Entscheidungsproblems im Prädikatenkalkul der ersten Stufe." Verlag der Ungarischen Akademie der Wissenschaften, Budapest, 1959.

[Tr] Boris A. Trakhtenbrot, "Impossibility of an algorithm for the decision problem in finite classes", Dokl. Akad. Nauk SSSR 70 (1950), 569–572.

[Tu] Alan M. Turing, "On computable numbers with an application to the Entscheidungsproblem," Proc. London Math. Soc. 42 (1937) 230–265; correction, 43 (1937), 544–546.