The Bulletin of Symbolic Logic
Volume 16, Number 4, Dec. 2010

VON NEUMANN, GÖDEL AND COMPLEXITY THEORY

ALASDAIR URQUHART

Abstract. Around 1989, a striking letter written in March 1956 from Kurt Gödel to John von Neumann came to light. It poses some problems about the complexity of algorithms; in particular, it asks a question that can be seen as the first formulation of the $P = ?\, NP$ question. This paper discusses some of the background to this letter, including von Neumann's own ideas on complexity theory. Von Neumann had already raised explicit questions about the complexity of Tarski's decision procedure for elementary algebra and geometry in a letter of 1949 to J. C. C. McKinsey. The paper concludes with a discussion of why theoretical computer science did not emerge as a separate discipline until the 1960s.

§1. Gödel and complexity theory.

1.1. Gödel's letter of 1956. Around 1989, a remarkable letter from Kurt Gödel to John von Neumann came to light, causing a stir in the theoretical computer science community. In August 1955 von Neumann had been diagnosed with bone cancer and in April 1956 he was admitted to Walter Reed Hospital, where he died on February 8, 1957. Gödel's last letter to him, dated 20 March 1956, may have been intended to take von Neumann's mind off his illness; a letter of 20 March 1974 to Abraham Robinson [17, pp. 204–205], written a few days before Robinson's death, strongly resembles the earlier letter to von Neumann.

Gödel asks a mathematical question in his letter that bears on the possibility that mathematics could be mechanized, in spite of the undecidability results of Church, Turing and others.

> Since, as I hear, you are feeling stronger now, I would like to take the liberty to write to you about a mathematical problem; your view on it would be of great interest to me: Obviously, it is easy to construct a Turing machine that allows us to decide, for each formula F of the restricted functional calculus and every natural number n, whether F has a proof of length n [length = number

Received March 20, 2010.
The author gratefully acknowledges the support of the Natural Sciences and Engineering Research Council of Canada.

1079-8986/10/1604-0002/$2.50

> of symbols]. Let $\psi(F, n)$ be the number of steps required for the machine to do that, and let $\varphi(n) = \max_F \psi(F, n)$. The question is, how rapidly does $\varphi(n)$ grow for an optimal machine? It is possible to show that $\varphi(n) \geq Kn$. If there really were a machine with $\varphi(n) \sim Kn$ (or even just $\sim Kn^2$) then that would have consequences of the greatest significance. Namely, this would clearly mean that the thinking of a mathematician in the case of yes-or-no questions could be completely[1] replaced by machines, in spite of the unsolvability of the Entscheidungsproblem. n would merely have to be chosen so large that, when the machine does not provide a result, it also does not make any sense to think about the problem [17, p. 375].

The problem addressed by Gödel is NP-complete, even if we restrict attention to the purely propositional fragment of first-order predicate logic (the 'restricted functional calculus' in Gödel's terminology). Furthermore, if we extrapolate from his remarks about linear-time and quadratic-time algorithms, the question that he posed to von Neumann can be seen as the earliest formulation of the $P =? NP$ problem.[2]

The fact that Gödel's problem is NP-complete is not immediately obvious, but an elegant proof due to Stephen A. Cook is given in outline in the preface to the collection edited by Peter Clote and Jan Krajíček [6], and also in the article by Buss [4]. Buss's article also contains a detailed proof of the fact that Gödel's problem is NP-complete even in the case of propositional logic—the reduction here is more intricate than in the case of predicate logic.

In keeping with his well known rationalistic optimism, Gödel conjectures that there could be a positive solution.

> Now it seems to me to be quite within the realm of possibility that $\varphi(n)$ grows that slowly. For 1.) $\varphi(n) \geq Kn$ seems to be the only estimate obtainable by generalizing the proof of the unsolvability of the Entscheidungsproblem; 2.) $\varphi(n) \sim Kn$ (or $\sim Kn^2$) just means that the number of steps when compared to pure trial and error can be reduced from N to $\log N$ (or $(\log N)^2$). Such significant reductions are definitely involved in the case of other finitist problems, e.g., when computing the quadratic remainder symbol by repeated application of the law of reciprocity. It would be interesting to know what the case would be, e.g., in determining whether a number is prime, and how significantly *in general* for finitist combinatorial problems the number of steps can be reduced when compared to pure trial and error [17, p. 375].

[1] Except for the formulation of axioms. [Gödel's own footnote]

[2] This may be an unjustified extrapolation. Buss [4, p. 60] argues that Gödel was considering quadratic time as the limit of feasible computation.

This letter is certainly most remarkable, especially given its date. Gödel pinpoints the crucial ideas that form the heart of the $P =? NP$ problem, even hinting at the key distinction between polynomial and exponential running time of algorithms. Furthermore, he raises the question of whether there is a feasible decision method for primality, an important topic of research until its solution by Agrawal, Kayal and Saxena [1]. In an article [23] reporting the rediscovery of the letter, Juris Hartmanis speculated that "these are very likely the first questions ever asked about the time required to solve problems on a deterministic Turing machine and particularly about the computational complexity of an NP complete problem." The second remark (about NP completeness) may well be correct, but in fact von Neumann had asked questions about computational complexity several years earlier (though without explicit reference to Turing machines).

One of the themes of the present article is that Gödel's remarks appear somewhat less surprising if viewed as part of von Neumann's scientific biography. In particular, there is a remarkable letter of von Neumann to J. C. C. McKinsey from 1949 that raises questions about the complexity of algorithms quite explicitly. This discussion is not intended to detract in any way from Gödel's striking insights—rather, the idea is to show that the theory of computational complexity has a rich prehistory extending back to the 1940s. As a preliminary to a discussion of von Neumann's letter, in its way just as striking as the 1956 letter of Gödel, let us fill in some of the background.

1.2. Gödel and the complexity of proofs. It is sometimes forgotten that Gödel published what is perhaps the very first result in modern complexity theory, though it is a result in the complexity of proofs, rather than the complexity of algorithms. His result, published only as an abstract [20], reports a speedup phenomenon in proof theory. The main result of the abstract is a speedup of arithmetic of order $i + 1$ (S_{i+1}) over arithmetic of order i (S_i). More precisely, for any computable function φ, there are infinitely many formulas F provable in arithmetic of order i so that if k is the length of the shortest proof of F in S_{i+1}, then the length of the shortest proof of F in S_i is greater than $\varphi(k)$.

Here, by "length," Gödel means the number of lines in the proof, rather than the number of symbols. This definition of length complicates the proof of his result, since, as Parikh observes in his comments [31] to the reprinting of the paper in Gödel's *Collected Works, Volume 1*, if we employ the well known trick of Craig [11] to axiomatize a recursively enumerable theory, any theorem can be proved in at most three lines. Craig's construction is based on the definition of a new decidable set of axioms for any such theory; all of the axioms are of the form $A \wedge (A \wedge (\cdots \wedge A)\cdots)$, where A is a theorem of the original theory, and there is an axiom of this form for any theorem. Consequently, if we include the axiom scheme $(A \wedge B) \to A$ in our axiomatization, any theorem can be proved in three lines, using *modus ponens*. This means that Gödel's result only holds for certain formulations of arithmetic, showing that care must be taken in the proof of the theorem. The first published proof of the theorem appears to be that of Buss [3]. Parikh in 1973 published a speedup theorem for second-order arithmetic over first-order arithmetic, where addition and multiplication are formalized as three-place relations; this result was extended to higher order arithmetics by Jan Krajíček [26].

It is more natural to define the length of a proof as the number of symbols in it—it is not clear why Gödel chose the less natural measure. Speedup theorems similar to Gödel's using the number of symbols as the measure of length have been proved by Mostowski [29], and Ehrenfeucht and Mycielski [15]. Further references and background are to be found in Parikh's introduction to the reprinting of Gödel's abstract in Volume 1 of the *Collected Works* [31].

§2. Von Neumann and complexity theory.

2.1. The logical theory of automata. Von Neumann's writings on computing machines show a fundamental concern for complexity considerations from the very beginning. His address to the Hixon symposium in 1948 [42] lays out the problems very clearly:

> Throughout all modern logic, the only thing that is important is whether a result can be achieved in a finite number of elementary steps or not. The size of the number of steps which are required, on the other hand, is hardly ever a concern of formal logic. Any finite sequence of correct steps is, as a matter of principle, as good as any other. It is a matter of no consequence whether the number is small or large, or even so large that it couldn't possibly be carried out in a lifetime, or in the presumptive lifetime of the stellar universe as we know it. In dealing with automata, this statement must be significantly modified. In the case of an automaton the thing which matters is not only whether it can reach a certain result in a finite number of steps at all but also how many such steps are needed [40, p. 303].

This mismatch between conventional attitudes to logic in the 1940s and the requirements of computers led von Neumann to propose a new logic of automata:

> The logic of automata will differ from the present system of formal logic in two relevant respects.
>
> 1. The actual length of "chains of reasoning," that is, of the chains of operations, will have to be considered.

2. The operations of logic (syllogisms, conjunctions, disjunctions, negations, etc., ...) will all have to be treated by procedures which allow exceptions (malfunctions) with low but non-zero probabilities [40, p. 304].

Von Neumann's comments on the complexity of computations, and his ideas for a logic of practical computability are remarkably prescient. However, the last quotation shows that his concerns were not exactly the same as those of contemporary theoretical computer science. In the current theory of computational complexity, the possibility of errors is ignored, and the computational model is that of an idealized error-free machine (it is common to consider probabilistic algorithms, but that is a different idea).

By contrast, in von Neumann's writings on computers from the 1940s and 1950s, the control of errors occupies a central place. This shows up clearly in a passage immediately following the first quotation above. He states that there are two reasons for the importance of the number of steps needed in a computation:

> First, automata are constructed in order to reach certain results in certain pre-assigned durations, or at least in pre-assigned orders of magnitude of duration. Second, the componentry employed has on every individual operation a small but nevertheless non-zero probability of failing. In a sufficiently long chain of operations the cumulative effect of these individual probabilities of failure may (if unchecked) reach the order of magnitude of unity—at which point it produces, in effect, complete unreliability [40, pp. 303–304].

It is the second reason that dominates in Von Neumann's discussions of complexity. Consequently, he is led to propose the development of a logical theory of automata that would belong to continuous rather than discrete mathematics. He begins by emphasizing the difficulties of the discrete, combinatorial approach:

> We are very far from possessing a theory of automata which deserves that name, that is, a properly mathematical-logical theory. There exists today a very elaborate system of formal logic, and, specifically, of logic as applied to mathematics.... Everybody who has worked in formal logic will confirm that it is one of the technically most refractory parts of mathematics. The reason for this is that it deals with rigid, all-or-none concepts, and has very little contact with the continuous concept of the real or of the complex number, that is, with mathematical analysis. Yet analysis is the technically most successful and best-elaborated part of mathematics. Thus formal logic is, by the nature of its approach, cut off from the best cultivated portions of mathematics, and forced onto the

> most difficult part of the mathematical terrain, into combinatorics [40, pp. 302–303].

By contrast, von Neumann envisages a theory of automata that is more like thermodynamics than discrete mathematics;

> All of this will lead to theories which are much less rigidly of an all-or-none nature than past and present formal logic. They will be of a much less combinatorial, and much more analytical, character. In fact, there are numerous indications to make us believe that this new system of formal logic will move closer to another discipline which has been little linked in the past with logic. This is thermodynamics, primarily in the form it was received from Boltzmann, and is that part of theoretical physics which comes nearest in some of its aspects to manipulating and measuring information. Its techniques are indeed much more analytical than combinatorial, which again illustrates the point that I have been trying to make above [40, p. 304].

Theoretical computer science has not developed in the direction suggested by von Neumann. Although the theory of error-correcting codes is an important part of this discipline, its techniques are those of discrete mathematics and number theory, far removed from the ideas of thermodynamics.

In December 1949, von Neumann gave a series of five lectures at the University of Illinois that continue some of the themes of his Hixon Symposium lecture of the preceding year. The lectures have survived only as an incomplete transcript of a tape recording (the tape recording itself has disappeared), but were edited by Arthur W. Burks as part of his collection of von Neumann's late work on the theory of automata [43]. In his second lecture, von Neumann again addresses questions of complexity. He discusses the relative complexity of describing an automaton, and describing its behaviour.

> I am twisting a logical theorem a little, but it's a perfectly good logical theorem. It's a theorem of Gödel that the next logical step, the description of an object, is one class type higher than the object and is therefore asymptotically [?] infinitely longer to describe [43, p. 47].

Later in the lecture, he expands a little on his earlier remarks.

> It is connected with the theory of types and with the results of Gödel. The feature is just this, that you can perform within the logical type that's involved everything that's feasible, but the question of whether something is feasible in a type belongs to a higher logical type [43, p. 51].

These passages of von Neumann's lecture are certainly far from clear. In an attempt to clarify the matter, Burks wrote to Gödel, asking whether he had discussed these points with von Neumann. He speculated that von Neumann might have been referring to Gödel's 1936 results on the length of proofs. However, he pointed out that:

> This fits everything that von Neumann says except that the result seems to go in the opposite direction: namely, the higher the type the shorter the proof [43, p. 55].

In his reply, Gödel disagreed with this reading of von Neumann's remarks:

> I have some conjecture as to what von Neumann may have had in mind in the passages you quote, but since I never discussed these matters with him it is only a guess.
>
> I think the theorem of mine which von Neumann refers to is not that on the existence of undecidable propositions or that on the lengths of proofs but rather the fact that a complete epistemological description of a language A cannot be given in the same language A, because the concept of truth of sentences of A cannot be defined in A. It is this theorem which is the true reason for the existence of undecidable propositions in the formal systems containing arithmetic [43, p. 55].

Burks admits that the recorded transcript is "mangled at this point" [43, p. 54], and it is possible that the transcription is garbled in these puzzling passages. As we shall see below, von Neumann's letter of 1949 to McKinsey tends to disconfirm Gödel's own interpretation of the allusions.

2.2. The letter to McKinsey. Although von Neumann's Hixon lecture of 1948 does prefigure some of the themes of contemporary complexity theory, a much closer anticipation appears in a remarkable letter [34, pp. 178–180] that von Neumann wrote to J. C. C. McKinsey on February 18, 1949.

McKinsey had written to von Neumann about his idea for solving n-person games by using a computer; he was working at RAND Corporation in Santa Monica, California at the time. Partly as a result of von Neumann's enthusiastic advocacy, RAND had built up a powerful research group in game theory in the late 1940s [30, pp. 116–117]—McKinsey himself published one of the first texts on the theory of games [28] in 1952.

McKinsey collaborated with Tarski in the 1940s. Tarski had completed a monograph entitled *The Completeness of Elementary Algebra and Geometry* in 1939; it had reached the stage of page proofs when the outbreak of the Second World War put an end to its publication by the Parisian publishing house Hermann et Cie. However, McKinsey undertook a revision of the work, and with the new title *A Decision Method for Elementary Algebra and Geometry*, it was published by RAND in 1948; a second, slightly revised edition appeared in 1951 [39]. In his preface to the revised edition, Tarski remarks about the new version:

> As was to be expected, it reflected the specific interests which the RAND corporation found in the results. The decision method for elementary algebra and geometry—which is one of the main results of the work—was presented in a systematic and detailed way, thus bringing to the fore the possibility of constructing an actual decision machine [39].

McKinsey's proposal was to use Tarski's decision method for elementary algebra to search for solutions of games; specifically, he seems to have advocated a special purpose machine built to implement Tarski's method. In reply, von Neumann observed that such problems could also be solved by a general purpose digital computer, but that "the mere observation that a particular type of finite problem can be solved with either type of machine is not very relevant. . . . The crucial question is that of the time required for the solution."

> Let N be the number of steps required by Tarski's decision method in the case of either one of the problems. N is, of course, a function of n and r. What kind of a function is it? How rapidly does it increase? How large is it for moderately large values of n and r? Or conversely: In what range can n and r be chosen without giving prohibitively large N's? [34, p. 179]

He goes on to discuss the speedup in computation that could be attained through the use of computers, and estimates that a speedup of the order of 10^4 to 10^5 over human computers could be attained. This figure of 10^5 then represents an upper bound on the complexity of problems suitable for solution by machine.

> My skepticism regarding the suitability of any variety of machine which is at this moment in sight for combinatorial-logical problems is just due to this: I suspect that while the five powers of 10 referred above include a lot of territory in the sense of problems of a mathematical-analytical character, they cover very little material in the case of combinatorial-logical problems. I am inclined to believe this—until I see a proof of the opposite—because the number of steps that are needed to solve a problem increases with the characteristic parameters of a problem much more quickly for problems of a high logical type than for others of a lower type. This is a rather natural conclusion from the old work of K. Gödel. Now ordinary mathematical-analytical problems are never of a high logical type: They are usually of the next type after that one of the arithmetical fundamental variable. Logical-combinatorial problems like those which you mentioned in your letter, on the other hand, are almost always of a higher type. I have not determined the type of either one of your problems, but I suppose that it will be the second or third one above the arithmetical fundamental variable [34, p. 179].

This letter to McKinsey is certainly a most striking foreshadowing of the basic ideas of computational complexity, including the key emphasis on the rate of growth of the number of steps needed, as a function of the size of the input. In contrast to the somewhat optimistic attitude in Gödel's letter of 1956, von Neumann evinces skepticism about the straightforward computer implementation of algorithms for decidable logical theories, basing his doubts on Gödel's work of 1936 on the complexity of proofs. The reference to this work of Gödel is quite explicit in the letter, making it more likely that it was in fact the theorem on lengths of proofs to which von Neumann was alluding in his Illinois lectures, in accord with Burks's conjecture.

Von Neumann's skepticism was vindicated, but only much later, with the 1972 proof by Fischer and Rabin [19] of an exponential lower bound on the complexity of the problem considered by McKinsey. Let $\mathrm{Th}(\mathbb{R})$ be the set of all sentences true in the set of real numbers, expressed in terms of addition and multiplication. Fischer and Rabin prove that there is a constant $c > 0$, so that for every decision procedure $\boldsymbol{DP}$ for $\mathrm{Th}(\mathbb{R})$, there exists an integer n_0 so that for all $n > n_0$, there is a sentence F of length n in $\mathrm{Th}(\mathbb{R})$ that requires 2^{cn} steps to decide whether or not F belongs to $\mathrm{Th}(\mathbb{R})$. The integer n_0 may be considered the point beyond which the algorithm is unfeasible in the worst case. Their proof shows that $n_0 = O(|\boldsymbol{DP}|)$, where $|\boldsymbol{DP}|$ represents the size of a program encoding the decision procedure. Fischer and Rabin remark that "computations and proofs become very long quite early in the game." Bearing in mind von Neumann's questions about the feasibility of problems of moderate size, does this result rule out applications of Tarski's algorithm?

In 1951, McKinsey, classified as a security risk because of his open homosexuality, was forced to leave RAND, joining the Philosophy Department at Stanford University [30, p. 185] [16, pp. 161,185]. It seems unlikely that any of his plans for implementation of Tarski's algorithm were carried out, particularly in view of the fact that the computer technology required was not available at RAND Corporation while McKinsey was working there. In 1950, RAND's computing equipment consisted of six IBM 604 calculators; these machines were "electronic calculating punches" that performed basic arithmetical operations (addition, subtraction, multiplication, division) when given input in the form of punched cards. Von Neumann, a consultant to RAND, convinced them to build a computer with stored program capability. The result was the JOHNNIAC, completed in 1953; it was based on the design of the machine built at the Institute for Advanced Study[3] under von Neumann's leadership, but with substantial improvements [2, pp. 92–93].

One of the earliest attempts at implementing a decision procedure for a first-order theory was that of Martin Davis, who wrote a program for the Institute for Advanced Study electronic digital computer [12]. About this

[3]The IAS machine was popularly known as the JONIAC [43, p. 12].

program, Davis later remarked: "Its great triumph was to prove that the sum of two even numbers is even" [13, p. 15]. Davis concluded his 1957 report by remarking:

> The writer's experience would indicate that with equipment presently available, it will prove impracticable to code any decision procedures considerably more complicated than Presburger's. Tarski's procedure for elementary algebra falls under this head [12, p. 223].

In any case, it soon became apparent that Tarski's method required too much computation to be practical except for quite trivial problems. Nevertheless, George Collins, a student of J. Barkley Rosser, reported on his work on programming part of the decision procedure on an IBM 704 at the famous Cornell Summer Institute in Symbolic Logic in 1957 [8]. Collins subsequently made a very striking improvement in Tarski's algorithm with his discovery of cylindrical algebraic decomposition [9]. Subsequently, Collins's ideas have been applied in many areas of pure and applied mathematics—for a partial list of these applications, the reader is referred to the excellent anthology by Caviness and Johnson [5].

§3. The origins of modern complexity theory. It is tempting to speculate that Gödel's letter of 1956 represents only a small fragment of an ongoing conversation between Gödel and von Neumann. However, there is no evidence of this, and on the whole, the interpretation seems unlikely. Von Neumann, although a colleague of Gödel at the Institute for Advanced Study, was in continual demand as a consultant to government and industry; an astonishingly long list of his consultancies from 1940 to 1957 can be found in William Aspray's history [2, p. 246]. In October 1954, he was appointed by President Eisenhower as a member of the United States Atomic Energy commission, and left Princeton on a leave of absence, discontinuing all commitments except for his chairmanship of the Strategic Missiles Evaluation Committee [2, p. 250]. It is difficult to imagine much contact between Gödel and von Neumann under these circumstances. In addition, Gödel himself remarked in his letter to Burks (see §2.1 above) that he had never discussed these matters with von Neumann.

In his 1989 article, Juris Hartmanis speculated that if von Neumann had started thinking about these issues earlier, then computer science would have had an earlier start:

> It is a pity that von Neumann was not challenged by Gödel to think about computational complexity when he was in good health. The development of theoretical computer science could have been substantially accelerated and the computer science intellectual landscape could be quite different today. A healthy, superbright Johnny could have possibly solved the notorious $P =? NP$

> problem before it became notorious or made it even more notorious than it is today, as the problem, raised by Gödel, which defied von Neumann [23].

As we have seen above, von Neumann had already been considering questions of computational complexity in the late 1940s. In spite of this, computational complexity as a defined area of research did not emerge until the 1960s. Why was this?

We can find a clue to the answer if we look at another article by Hartmanis, his fascinating reminiscences of the beginnings of theoretical computer science.

> I was attracted to pure mathematics in the United States and completed my Ph.D. work in lattice theory at the California Institute of Technology in 1955. During this period I met for the first time John von Neumann when he lectured in 1952 at Cal. Tech. ... I must admit that at that time I failed completely to understand the importance of von Neumann's work on computing machines and automata theory. His enthusiasm and brilliance were obvious, but my indoctrination in pure mathematics was too strong to be penetrated by such ideas. As a matter of fact, many of the Cal. Tech. mathematicians lamented the fact that "Johnny had given up doing real mathematics" [22].

Without the institutional setting provided by a community of researchers in theoretical computer science, and (even more importantly) independent departments of computer science, the seeds sown by von Neumann's lectures advocating the new science of automata were bound to fall on stony soil.

Modern theoretical computer science is a remarkable confluence of several different intellectual traditions, and in its early development it attracted researchers from mathematics, physics, engineering, logic and linguistics. The nascent science of algorithms, however, would have been stillborn without two external stimuli, first the rapid development and diffusion of computer technology in the 1950s, and second, the establishment in the 1960s of separate and independent departments of computer science, in response to the demand created by the first stimulus.

The years 1956 and 1957 are especially significant in the diffusion of computers and their availability to students and researchers. In an article recalling his experience in 1957 with the IBM 650, Donald Knuth notes that

> The IBM 650 was the first computer to be manufactured in really large quantities. Therefore the number of people in the world who knew about programming increased by an order of magnitude. ... It was a great machine on which to learn about machines. We had a computer architecture that was rudimentary but pleasant to use;

> and we had program masterpieces like the Bell interpreter and Poley's assembler, as examples of excellent style [24] [25, pp. 228,236].

The year 1957 was also the year of the famous Cornell Summer Institute already mentioned in connection with the computer implementations of Martin Davis and George Collins; some of the history of this conference can be found in Halmos's reminiscences [21, pp. 215–216] and the Fefermans' biography of Tarski [16, pp. 220–230]. No fewer than eighteen of the eighty-two talks summarized in the proceedings are more or less directly connected with computer-related questions—this is without counting the work of Rabin and Scott on finite automata that was to prove highly influential in the early development of theoretical computer science. The conference also included four talks on programming systems, including one by Peter Sheridan of the IBM Corporation on the "Fortran Automatic Coding System" [36].

In spite of the anticipations of some of the main concepts of complexity theory by von Neumann and Gödel, these ideas, it seems, could not form part of a generally accepted mathematical research program until the wider spread of computer technology had provided the foundation. In fact, important work on the complexity of circuits had already been carried out by Claude Shannon in the late 1940s [35], and circuit complexity was a constant concern of the early computer designers. Yet this area was considered by mathematicians and logicians to be the province of engineers, in spite of the early involvement of logicians such as Quine, who was stimulated to work on Boolean logic and circuit simplification by a consultancy at the RAND corporation in the summer of 1949 [32, p. 27]. Circuit complexity is now one of the most vigorous areas in theoretical computer science, but it only attained mathematical respectability several decades after the basic results in the area were proved by Shannon.

In spite of the separation of research communities caused by the Cold War, the development of complexity theory in the Soviet Union followed remarkably similar lines to those in the U.S.A. In the Soviet Union, research on the complexity of switching networks continued through the 1950s, the work of Shannon providing some of the earlier inspiration. Even earlier than in the West, Soviet scientists recognized the key importance of the question of whether brute force search ("perebor") was unavoidable in some cases. In fact, there was even a claim by Yablonski in 1959 to have *proved* its unavoidability; a detailed account of this intriguing episode is given by Boris Trakhtenbrot [41]. Yet in Russia, just as in the U.S.A., we find isolated attempts that do not grow into a full research programme. Trakhtenbrot reports [41, p. 391] that as early as 1956 G. S. Tseitin, at the time a 19-year-old student of Markov at Leningrad University, proved nontrivial lower and upper bounds on the complexity of some concrete problems in the context of Markov's normal algorithms. However, Tseitin's results remained unpublished.

The proposal to identify feasible with polynomial time algorithms appeared in the 1960s in the work of Cobham [7], Edmonds [14] and Rabin [33]. The framework in which to pose the most basic problem about "perebor" algorithms, that of NP-completeness, does not emerge until the early 1970s, with the independent work of Cook [10] and Levin [27]. A more detailed history of the background to the $P = ?\, NP$ question can be found in Sipser's excellent historical survey [37].

I have argued in this final section that questions about computational complexity could not form the basis of a serious mathematical research program until certain external conditions were met. Yet von Neumann contributed in a less obvious way to the development of computer science, leading ultimately to the current flourishing state of computational complexity theory. As William Aspray explains in his fine history [2], von Neumann earlier than anybody else, with the possible exception of Alan Turing, saw the potential of the digital computer as an instrument of scientific research. He was able to push through the construction of a digital computer at the Institute for Advanced Study, in spite of its "ivory tower" orientation, and the IAS computer was crucial in the spread of digital technology in the United States.

Von Neumann, before his tragically early death, made the IAS computer available to researchers in areas such as meteorology, shock waves, condensed matter physics, quantum theory, astrophysics, logic, and many other areas (for a fuller list, see [2, pp. 156–157]). He insisted that the new technology be made freely available outside the realm of commercial patents, and as a result, the IAS computer was widely copied, its direct descendants including the AVIDAC, the ILLIAC, the JOHNNIAC and the IBM 701 [2, p. 91]. For this reason alone, von Neumann must be considered one of the most important founders of modern computer science.

REFERENCES

[1] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena, *PRIMES is in P*, ***The Annals of Mathematics***, vol. 160 (2004), pp. 781–793.

[2] William Aspray, ***John von Neumann and the origins of modern computing***, The MIT Press, 1990.

[3] Samuel R. Buss, *On Gödel's theorems on lengths of proofs I: Number of lines and speedup for arithmetics*, ***The Journal of Symbolic Logic***, vol. 59 (1994), pp. 737–756.

[4] ———, *On Gödel's theorems on lengths of proofs II: Lower bounds for recognizing k symbol provability*, ***Feasible mathematics II*** (Peter Clote and Jeffrey Remmel, editors), Birkhäuser, 1995, pp. 57–90.

[5] B. F. Caviness and J. R. Johnson (editors), ***Quantifier elimination and cylindrical algebraic decomposition***, Texts and Monographs in Symbolic Comptutation, Springer-Verlag, Vienna and New York, 1998.

[6] Peter Clote and Jan Krajíček (editors), ***Arithmetic, proof theory, and computational complexity***, Oxford University Press, 1993.

[7] ALAN COBHAM, *The intrinsic computational difficulty of functions*, ***Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science*** (Yehoshua Bar-Hillel, editor), North-Holland, 1964, pp. 24–30.

[8] GEORGE E. COLLINS, *Tarski's decision method for elementary algebra*, ***Summer Institute for Symbolic Logic, Cornell University 1957***, Institute for Defense Analyses, Communications Research Division, second ed., 1960, summaries of talks presented, pp. 64–70.

[9] ———, *Quantifier elimination for real closed fields by cylindrical algebraic decomposition*, ***Automata theory and formal languages***, Lecture Notes in Computer Science, vol. 33, Springer Verlag, 1975, reprinted in [5], pp. 85-121, pp. 134–183.

[10] STEPHEN A. COOK, *The complexity of theorem-proving procedures*, ***Proceedings of the third annual ACM symposium on Theory of Computing***, 1971, pp. 151–158.

[11] WILLIAM CRAIG, *On axiomatizability within a system*, ***The Journal of Symbolic Logic***, vol. 18 (1953), pp. 30–32.

[12] MARTIN DAVIS, *A program for Presburger arithmetic*, ***Summer Institute for Symbolic Logic, Cornell University 1957***, Institute for Defense Analyses, Communications Research Division, second edition ed., 1960, summaries of talks presented, pp. 215–223.

[13] ———, *The prehistory and early history of automated deduction*, ***Automation of reasoning, volume 1: Classical papers on computational logic 1957–1966*** (Jörg Siekmann and Graham Wrightson, editors), Springer-Verlag, 1983, pp. 1–28.

[14] JACK EDMONDS, *Paths, trees, and flowers*, ***Canadian Journal of Mathematics***, vol. 17 (1965), pp. 449–467.

[15] ANDRZEJ EHRENFEUCHT and JAN MYCIELSKI, *Abbreviating proofs by adding new axioms*, ***Bulletin of the American Mathematical Society***, vol. 77 (1971), pp. 366–367.

[16] ANITA BURDMAN FEFERMAN and SOLOMON FEFERMAN, ***Alfred Tarski: Life and logic***, Cambridge University Press, 2004.

[17] Solomon Feferman, John W. Dawson Jr., Warren Goldfarb, Charles Parsons, and Wilfried Sieg (editors), ***Kurt Gödel. Collected works, volume V: Correspondence H–Z***, Oxford University Press, 2003.

[18] Solomon Feferman, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort (editors), ***Kurt Gödel. Collected works, volume I: Publications 1929–1936***, Oxford University Press, 1986.

[19] M. J. FISCHER and M. O. RABIN, *Super-exponential complexity of Presburger arithmetic*, ***Complexity of computations, SIAM–AMS Proceedings***, vol. 7, 1974, reprinted in [5], pp. 122-135, pp. 27–41.

[20] KURT GÖDEL, *Über die Länge von Beweisen*, ***Ergebnisse eines mathematischen Kolloquiums***, vol. 7 (1936), pp. 23–24, reprinted in [18] with facing English translation, entitled "On the Length of Proofs".

[21] P. R. HALMOS, ***I want to be a mathematician: An automathography***, Springer-Verlag, 1985.

[22] JURIS HARTMANIS, *Observations about the development of theoretical computer science*, ***Annals of the History of Computing***, vol. 3 (1981), pp. 42–51.

[23] ———, *Gödel, von Neumann and the $P =? NP$ problem*, ***Bulletin of the European Association for Theoretical Computer Science***, vol. 38 (1989), pp. 101–107.

[24] DONALD E. KNUTH, *The IBM* 650: *An appreciation from the field*, ***Annals of the History of Computing***, vol. 8 (1986), pp. 50–55, reprinted in [25], pp. 227-239.

[25] ———, ***Selected papers on computer science***, CSLI Lecture Notes 59, CSLI Publications and Cambridge University Press, 1996.

[26] JAN KRAJÍČEK, *On the number of steps in proofs*, ***Annals of Pure and Applied Logic***, vol. 41 (1989), pp. 153–178.

[27] LEONID LEVIN, *Universal sorting problems*, ***Problems of Information Transmission***, vol. 9 (1973), pp. 265–266.

[28] J. C. C. McKinsey, ***Introduction to the theory of games***, The RAND Series, McGraw-Hill, New York, Toronto, London, 1952.

[29] Andrzej Mostowski, ***Sentences undecidable in formalized arithmetic: An exposition of the theory of Kurt Gödel***, North-Holland, Amsterdam, 1952.

[30] Sylvia Nasar, ***A beautiful mind: A biography of John Forbes Nash, Jr., winner of the Nobel Prize in Economics, 1994***, Simon and Schuster, 1998.

[31] Rohit J. Parikh, *Introductory note to* 1936*a*, ***Kurt Gödel, Collected works, volume I: Publications 1929–1936***, Oxford University Press, 1986, pp. 394–396.

[32] W. V. Quine, *Autobiography of W. V. Quine*, ***The philosophy of W. V. Quine*** (Lewis Edwin Hahn and Paul Arthur Schilpp, editors), The Library of Living Philosophers, vol. XVIII, Open Court Publishing Company, 1986, pp. 1–48.

[33] Michael O. Rabin, *Mathematical theory of automata*, ***Proceedings of the 19th ACM Symposium in Applied Mathematics***, 1966, pp. 153–175.

[34] Miklós Rédei (editor), ***John von Neumann: Selected letters***, History of Mathematics, vol. 27, American Mathematical Society, 2005.

[35] Claude E. Shannon, *The synthesis of two-terminal switching circuits*, ***Bell System Technical Journal***, vol. 28 (1949), pp. 59–98, reprinted in [38].

[36] Peter Sheridan, *The Fortran automatic coding system*, ***Summer Institute for Symbolic Logic, Cornell University 1957***, Institute for Defense Analyses, Communications Research Division, second ed., 1960, summaries of talks presented, pp. 425–426.

[37] Michael Sipser, *The history and status of the P versus NP question*, ***Proceedings of the 24th annual ACM Symposium on the Theory of Computing***, 1992, pp. 603–618.

[38] N. J. A. Sloane and Aaron D. Wyner (editors), ***Claude E. Shannon. Collected papers***, IEEE Press, 1993.

[39] Alfred Tarski, ***A decision method for elementary algebra and geometry***, second revised ed., University of California Press, 1951, prepared for publication with the assistance of J. C. C. McKinsey. First edition published in 1948 as RAND report R-109, RAND Corporation, Santa Monica, CA.

[40] A. H. Taub (editor), ***John von Neumann. Collected works volume V: Design of computers, theory of automata and numerical analysis***, Pergamon Press, 1961.

[41] B. A. Trakhtenbrot, *A survey of Russian approaches to* Perebor (*Brute-Force Search*) *algorithms*, ***Annals of the History of Computing***, vol. 6 (1984), pp. 384–440.

[42] John von Neumann, *The general and logical theory of automata*, ***Cerebral mechanisms in behaviour—the Hixon symposium***, John Wiley, New York, 1948, reprinted in [40], pp. 288-328.

[43] ———, ***Theory of self-reproducing automata***, University of Illinois Press, Urbana and London, 1966, edited and completed by Arthur W. Burks.

DEPARTMENT OF COMPUTER SCIENCE
10 KING'S COLLEGE ROAD
UNIVERSITY OF TORONTO
TORONTO, ONTARIO M5S 3G4, CANADA
E-mail: urquhart@cs.toronto.edu