Finite-Model Theory—a Personal Perspective*

Ronald Fagin
IBM Almaden Research Center
650 Harry Road
San Jose, California 95120-6099†

Appeared in: Theoretical Computer Science 116 (1993), pp. 3–31

Abstract

Finite-model theory is a study of the logical properties of finite mathematical structures. This paper is a very personalized view of finite-model theory, where the author focuses on his own personal history, and results and problems of interest to him, especially those springing from work in his Ph.D. thesis. Among the topics discussed are:

1. *Differences between the model theory of finite structures and infinite structures.* Most of the classical theorems of logic fail for finite structures, which gives us a challenge to develop new concepts and tools, appropriate for finite structures.

2. *The relationship between finite-model theory and complexity theory.* Surprisingly enough, it turns out that in some cases, we can characterize complexity classes (such as NP) in terms of logic, where there is no notion of machine, computation, or time.

3. *0-1 laws.* There is a remarkable phenomenon, which says that certain properties (such as those expressible in first-order logic) are either almost surely true or almost surely false.

4. *Descriptive complexity theory.* Here we consider how complex a formula must be to express a given property.

In recent years, there has been a re-awakening of interest in finite-model theory. One goal of this paper is to help "fan the flames" of interest, by introducing more researchers to this fascinating area.

---

*A preliminary version of this paper appeared in the Proceedings of the 3rd International Conference on Database Theory, Dec. 1990, Springer-Verlag Lecture Notes in Computer Science 470, pp. 3-24.

†email: fagin@almaden.ibm.com

1 Introduction

Model theory is a study of the logical properties of mathematical structures. Model theory is a key area of mathematical logic, and has a rich body of results (see, for example, Chang and Keisler [CK90] or Shelah [She78]).

An important part of mathematics (as opposed to mathematical logic) is the study of finite structures, such as finite graphs or finite groups. However, the *model theory* of finite structures is, in my opinion, quite underdeveloped compared to unrestricted model theory. Why is this? There are (at least) two reasons: philosophical and technical. The philosophical reason is that logic was developed to provide a solid foundation for mathematics, which includes the study of *all* structures, both finite and infinite. There is no particular reason to restrict attention to finite structures. In fact, as observed by Gurevich [Gur88], if anything it is the *infinite* structures that need the most attention, since problems seem most likely to arise there! The second reason is technical: almost none of the key theorems and tools of model theory, such as the Completeness Theorem and the Compactness Theorem, apply to finite structures. This issue is discussed in Section 3. Among the few tools of model theory that still work on finite structures are the *games* of Ehrenfeucht [Ehr61] and Fraïssé [Fra54] (for an introduction to Ehrenfeucht-Fraïssé games and some of their applications to finite-model theory, see [AF90, pp. 122–126]). Chang and Keisler note that "in almost all of the deeper theorems in model theory the key to the proof is to construct the right kind of model" [CK90, p. 2]. These models almost invariably turn out to be *infinite* structures.

The main challenge of finite-model theory is to develop new concepts and associated tools, including concepts and tools which may be meaningful only for finite structures. One example (discussed in Section 5) is the relationship to complexity classes. Another example (discussed in Section 6) is the idea of considering the asymptotic probabilities of sentences as the cardinality of the universe grows.

My 1973 Ph.D. thesis [Fag73] (which was later published as the papers [Fag74, Fag75a, Fag75b, Fag75c, Fag76]) dealt exclusively with finite-model theory. I was disappointed that the field languished for years afterwards: very few papers were published in the area in the next ten years or so. However, in the mid-1970's, Hájek [Háj75, Háj77] discussed the importance of finite model theory. In fact, in [Háj77], Hájek made the explicit proposal to "develop logic (classical and generalized) modified by allowing only finite models." In recent years, there has been a re-awakening of interest in finite-model theory. There are several possible reasons for this renewed interest. One reason is the connection with complexity theory (see Section 5).

Another reason is the connection with databases and logic programming. In fact, it is possible to think of a database as simply a finite structure. This has not stopped many researchers, including me, from considering infinite databases when it is thereby possible to obtain interesting results. For example, in one key construction in [Fag82] I take the direct product of an infinite family of relations, each of which has at least two tuples, so the result is not only infinite but even uncountable. I shudder to think of a database practitioner's reaction to an uncountable database! Another example is in a paper of Vardi [Var82b], where Beth's Theorem (which fails for finite structures) is invoked. In the early days of database theory, I pushed database theory in the direction of model theory. In contrast, others, such as Ullman, were pushing the field towards algorithmic questions. Each of these directions has remained important for database theory. I will not discuss database theory in this paper. The interested reader is referred to textbooks by Ullman [Ull89] and Maier [Mai83], and survey papers by Kanellakis [Kan90], Chandra [Cha88], and by Vardi and myself [FV86].

Perhaps the actual reason for the current interest in finite model theory is that critical mass is starting to be achieved. I hope that this paper can help add to the critical mass.

Definitions are given in Section 2. Some differences between the model theory of finite structures and infinite structures are discussed in Section 3. Spectra and generalized spectra are defined in Section 4. The relationship between (generalized) spectra and complexity theory is explored in Section 5. In Section 6, there is a discussion of 0-1 laws. Descriptive complexity issues appear in Section 7. Section 8 contains conclusions, including a classification by Vardi of finite-model theory into three lines of research.

2 Definitions and conventions

The definitions in this section are informal and incomplete. For a careful development, see any good logic textbook, such as Enderton [End72] or Shoenfield [Sho67].

A *language* $\mathcal{L}$ (sometimes called a *similarity type*, a *signature*, or a *vocabulary*) is a finite set $\{P_1, \ldots, P_s\}$ of relation symbols, each of which has an arity. Constant or function symbols are not allowed; this is unimportant except in the case of 0-1 laws in Section 6, where the effect of allowing constant or function symbols is mentioned.

An *$\mathcal{L}$-structure* (or *structure over $\mathcal{L}$*, or simply *structure*) is a set A (called the *universe*), along with a mapping associating a relation R_i over A with each $P_i \in \mathcal{L}$, where R_i has the same arity as P_i, for $1 \leq i \leq s$. We may call R_i the *interpretation of P_i*. The structure is called *finite* if A is. For convenience, the universe of a finite structure will always be taken to be $\{0, 1, \ldots, n-1\}$ for some natural number n.

For definitions of a first-order sentence (where, intuitively, the only quantification is over members of the universe, and not over, say, sets of members of the universe) and what it means for a structure $\mathcal{A}$ to *satisfy* a sentence σ, written $\mathcal{A} \models \sigma$, see Enderton [End72] or Shoenfield [Sho67]. If $\mathcal{A}$ satisfies σ, then $\mathcal{A}$ is a *model* of σ (and a *finite model* if $\mathcal{A}$ is a finite structure). A sentence is *satisfiable* (resp. *finitely satisfiable*) if it has a model (resp. finite model). It is *valid* (resp. *valid over finite structures*) if it is satisfied by every structure (resp. finite structure) over the language of σ. We note that throughout this paper, we allow equality as a special predicate symbol, which is not considered to be a member of the language $\mathcal{L}$, and which always has the standard interpretation.

Let Σ be a set of sentences, and let σ be a single sentence. Then Σ *logically implies* σ, written $\Sigma \models \sigma$, if every structure that satisfies (every sentence of) Σ also satisfies σ.

Equivalently, Σ logically implies σ if there is no "counterexample structure" that satisfies Σ but not σ.

3 Some differences in the finite case

As noted earlier, many classical theorems of model theory fail for finite structures (see [Gur84, Gur90]). In this section, some of the differences between finite-model theory and (unrestricted) model theory are considered.

One of the great early contributions to mathematical logic is Gödel's Completeness Theorem. Gödel considers a particular proof system for first-order sentences, and writes $\Sigma \vdash \sigma$ if there is a proof of σ from Σ in this proof system. He then proves

Theorem 3.1: (Completeness Theorem) $\Sigma \models \sigma$ *iff* $\Sigma \vdash \sigma$.

One reason this result is remarkable is that it shows that a universal search over all possible structures $\mathcal{A}$ (checking to see whether every structure $\mathcal{A}$ that satisfies Σ also satisfies σ) is equivalent to the existence of a finite object (a proof of σ from Σ).

Because we can systematically list all proofs, we obtain the following important consequence of the Completeness Theorem.

Corollary 3.2: *The set of valid first-order sentences is r.e. (recursively enumerable).*

However, the next theorem says that we cannot replace "r.e." by "recursive".

Theorem 3.3: (Church's Theorem) *Assume that the language $\mathcal{L}$ contains some relation symbol that is not unary.*[1] *Then the set of valid first-order sentences over $\mathcal{L}$ is not recursive.*

Since a set is recursive precisely if it and its complement are r.e., the following is an immediate corollary of Corollary 3.2 and Theorem 3.3:

Corollary 3.4: *Assume that the language $\mathcal{L}$ contains some relation symbol that is not unary. Then the set of valid first-order sentences over $\mathcal{L}$ is r.e. but not co-r.e.*

Interestingly enough, exactly the opposite behavior is true about finite structures.

Theorem 3.5: (Trakhtenbrot [Tra50]) *Assume that the language $\mathcal{L}$ contains some relation symbol that is not unary.*[2] *Then the set of first-order sentences over $\mathcal{L}$ valid over finite structures is co-r.e. but not r.e.*

[1] In Church's original version of this theorem, the language was the language $\{+, \times\}$ of number theory. By techniques of Kalmár [Chu56, pp. 271-273], all we need is some relation symbol that is not unary.

[2] As with Church's Theorem, Trakhtenbrot used a richer language. As before, from Kalmar's results it follows that all that is needed is some relation symbol that is not unary (see e.g. Vaught [Vau60]).

The easy (and well-known) part of Theorem 3.5 is that the set of first-order sentences valid over finite structures is co-r.e. The reason this holds is that to find whether σ is not valid, it is possible to consider systematically every finite structure $\mathcal{A}$ over the language of σ to see whether $\mathcal{A} \not\models \sigma$ (it is easy to see that it is decidable to determine if $\mathcal{A} \not\models \sigma$). This makes it possible to list all the sentences that are not valid.

The hard part of Theorem 3.5, that the set of first-order sentences over finite structures is not r.e. (if the language contains some relation symbol that is not unary) tells us that there is no completeness theorem for finite structures, since as we saw, completeness implies that the set of valid sentences is r.e.

If a sentence is valid over all structures, then it is certainly valid when we consider only finite structures. (Intuitively, it is easier to be valid over only finite structures than over all structures, since there are less possible counterexamples.) However, the converse is false. An example of a first-order sentence that is valid over finite structures but not over all structures is a first-order sentence that says "If $<$ is a linear order, then it has a largest element". Of course, it also follows from Corollary 3.4 and Theorem 3.5 that the set of first-order sentences valid over all structures is not the same as the set of first-order sentences valid over finite structures.

A sentence is said to be *finitely controllable* if it is either unsatisfiable or finitely satisfiable. A set of sentences is finitely controllable if every member is. Thus, if $\mathcal{S}$ is finitely controllable, then every member of $\mathcal{S}$ that is satisfiable is finitely satisfiable. An important reason why finitely controllable classes are interesting is given by the next theorem (which is also well-known).

Theorem 3.6: *Let $\mathcal{S}$ be a recursive, finitely controllable set of first-order sentences. The decision problem for satisfiability of members of $\mathcal{S}$ is decidable.*

Proof: Let T be the set of satisfiable members of $\mathcal{S}$. Assume that $\varphi \in \mathcal{S}$. If φ is satisfiable, then by finite controllability, φ is finitely satisfiable. As in the easy part of Theorem 3.5, we can "prove" that φ is finitely satisfiable by giving a finite model for φ. This shows that T is r.e. If φ is not satisfiable, then we can give a proof of $\neg\varphi$, by the Completeness Theorem. This shows that T is co-r.e. Since T is both r.e. and co-r.e., it follows that T is recursive. ∎

Church's Theorem tells us that the satisfiability problem is undecidable in general (since, as mentioned above, a sentence is satisfiable iff its negation is not valid). However, Theorem 3.6 tells us that the satisfiability problem is decidable if we restrict our attention to recursive, finitely controllable classes. In fact, finite controllability is one of the main techniques used to prove decidability of the satisfiability problem for special classes of sentences. For example, by restricting attention to prenex normal form sentences with certain quantifier prefixes, we can obtain a finitely controllable class (this point is discussed more in Section 6). See Dreben and Goldfarb [DG79] for a survey of results in this area (see also Gurevich [Gur76] and the papers cited in [Gur76]).

Probably the most important corollary of the Completeness Theorem is the Compactness Theorem.

Theorem 3.7: (Compactness Theorem) *Let Σ be a set of first-order sentences. If every finite subset of Σ is satisfiable, then Σ is satisfiable.*

Proof: If Σ is not satisfiable, then $\Sigma \models \mathit{false}$, where *false* is some logically false sentence such as $p \wedge \neg p$. By the Completeness Theorem, $\Sigma \vdash \mathit{false}$. Thus, there is a proof of *false* from Σ in Gödel's proof system. Since every proof is of finite length in Gödel's proof system, only a finite subset $\Sigma' \subseteq \Sigma$ is used in the proof. So $\Sigma' \vdash \mathit{false}$. By the Completeness Theorem again, $\Sigma' \models \mathit{false}$. So Σ' is not satisfiable. ∎

The Compactness Theorem, which is one of the most powerful tools in the arsenal of model theory, is not true for finite structures. It is instructive to see an example. For each positive integer k, define σ_k to be a first-order sentence that says "There are at least k points." For example, we can take σ_3 to be

$$\exists x_1 \exists x_2 \exists x_3((x_1 \neq x_2) \wedge (x_1 \neq x_3) \wedge (x_2 \neq x_3)).$$

Let $\Sigma = \{\sigma_1, \sigma_2, \sigma_3, \ldots\}$. It is easy to see that Σ is not finitely satisfiable, although every finite subset of Σ is finitely satisfiable.

Two structures $\mathcal{A}$ and $\mathcal{B}$ over the same language are said to be *elementarily equivalent* if for every first-order sentence σ in this language, $\mathcal{A} \models \sigma$ iff $\mathcal{B} \models \sigma$. Elementary equivalence is an important model-theoretic notion, since if two structures are elementarily equivalent, then they cannot be distinguished by a first-order sentence. We close this section with another well-known theorem that is unique to finite structures (and that is easy to show). This theorem shows that elementary equivalence is uninteresting for finite structures.

Theorem 3.8: *For every finite $\mathcal{L}$-structure $\mathcal{A}$, there is a first-order sentence $\sigma_{\mathcal{A}}$ such that an arbitrary $\mathcal{L}$-structure $\mathcal{B}$ is isomorphic to $\mathcal{A}$ iff $\mathcal{B} \models \sigma_{\mathcal{A}}$.*

Thus, each finite structure is characterized up to isomorphism by a first-order sentence. Of course, no such theorem is true in general about infinite structures; consider, for example, nonstandard models of arithmetic. Theorem 3.8 depends on our assumption that the language is finite. We note that if the language were allowed to be infinite, then it is not hard to show that a variation of Theorem 3.8 holds, where instead of a single first-order sentence $\sigma_{\mathcal{A}}$ characterizing $\mathcal{A}$, a set $\Sigma_{\mathcal{A}}$ of first-order sentences characterizes $\mathcal{A}$. So even if the language were infinite, elementary equivalence is uninteresting for finite structures.

4 Spectra and generalized spectra

Since this paper is a personal perspective on finite-model theory, let me describe how I got interested in finite-model theory. In January of 1968, in a beautifully-taught introductory logic class that I took as a first-year graduate student in mathematics at Berkeley, the professor, Herbert Enderton, mentioned the notion of the *spectrum* of a first-order sentence, which is the set of cardinalities of its finite models. That is, if σ is a first-order sentence, and if n is a natural number, then n is in the spectrum of σ precisely if there is a structure $\mathcal{A}$ that satisfies σ where the cardinality of the universe of $\mathcal{A}$ is n. The notion of a spectrum was introduced by Scholz [Sch52]. As an example, if φ is a first-order sentence that gives the conjunction of the field axioms (φ says that $+$ and $\times$ are associative and commutative, that $\times$ distributes over $+$, etc.), then it is well-known that the spectrum of φ is the set of powers of primes.[3] It is easy to see that a spectrum is decidable: to decide if n is in the spectrum of the first-order sentence σ, systematically consider every finite structure with universe $\{0, 1, \ldots, n-1\}$ over the language of σ (there are only a finite number of such structures), and check to see whether any of them satisfies σ. If so, then n is in the spectrum of σ; if not, then n is not in the spectrum of σ. Although every spectrum is a decidable set of positive integers, the converse is false. There are several ways to see this. Just as we define the spectrum of a first-order sentence, we can similarly define the spectrum of a second-order sentence, and see, by an identical argument, that it is decidable. However, Bennett [Ben62] showed that there is a set S of positive integers that is the spectrum of a second-order sentence but not of a first-order sentence. Such a set S is a decidable set that is not a (first-order) spectrum. Another decidable set of positive integers that is not a spectrum can be obtained by forming the diagonal set D such that $n \in D$ iff n is not in the nth spectrum. Scholz [Sch52] posed the problem of characterizing spectra, and Asser [Ass55] posed the following key problem:

Open problem (Asser problem): Is the class of spectra closed under complement? That is, is the complement of every spectrum also a spectrum?

Because of Enderton's wonderful class (along with another magical logic class I had taken as an undergraduate at Dartmouth under Donald Kreider), I decided in 1969 to write my thesis in mathematical logic. My thesis advisor was Robert Vaught. My plan was to study spectra, and in particular Asser's problem.

If $\sigma(Q_1, \ldots, Q_k)$ is a first-order sentence,[4] then one way to view the spectrum of $\sigma(Q_1, \ldots, Q_k)$ is as the set of finite models of $\exists Q_1 \ldots \exists Q_k \sigma(Q_1, \ldots, Q_k)$ (since all of the relation symbols in the language have been quantified out, a model is simply a structure with universe of size n over the empty language, and we can identify such a structure with the natural number n). Around 1970, I expanded my investigations to *generalized spectra*, or equivalently Σ_1^1 *classes over finite structures*, where some but (unlike spectra)

[3] Since by assumption there are only relation symbols in the language, take $+$ and $\times$ to be ternary relations rather than binary functions.

[4] When I write $\sigma(Q_1, \ldots, Q_k)$ instead of just σ, this is to convey the fact that σ is over the language $\{Q_1, \ldots, Q_k\}$.

not necessarily all of the relation symbols in the language are existentially quantified out. Thus, a Σ_1^1 *sentence* is a sentence of the form $\exists Q_1 ... \exists Q_k \sigma(P_1, \ldots, P_s, Q_1, \ldots, Q_k)$, where $\sigma(P_1, \ldots, P_s, Q_1, \ldots, Q_k)$ is a first-order sentence and where the Q_i's are relation symbols (these are referred to as the *extra* relation symbols). This Σ_1^1 sentence refers to structures over the language $\{P_1, \ldots, P_s\}$. As an example, I now describe a Σ_1^1 sentence that says that a graph is 3-colorable. In this sentence, the three colors are represented by Q_1, Q_2, and Q_3. Let P represent the graph relation, and let $\sigma(P, Q_1, Q_2, Q_3)$ say "Each point has exactly one color, and no two points with the same color are connected by an edge". Then the Σ_1^1 sentence $\exists Q_1 \exists Q_2 \exists Q_3 \sigma(P, Q_1, Q_2, Q_3)$ says "The graph is 3-colorable". A class $\mathcal{C}$ of (finite) structures over some fixed language is said to be Σ_1^1 if it is the class of all (finite) structures that obey some fixed Σ_1^1 sentence. The notion of Σ_1^1 class is the same as Tarski's [Tar54] notion of *projective class (PC)*. In the finite case (where attention is restricted to finite structures), a Σ_1^1 class is called a *generalized spectrum*. Whereas a spectrum is a set of *numbers*, a generalized spectrum is a set of *structures*. For example, we just showed that the finite 3-colorable graphs form a generalized spectrum.

5 Relation to complexity classes

Just as a spectrum is decidable, it is easy to see by a very similar argument that the set of (encodings of) structures in a generalized spectrum is decidable. As mentioned earlier, spectra are *special* decidable sets, and a similar comment applies to generalized spectra. I began wracking my brains to try to understand what made these sets special. All of a sudden it occurred to me that they were not just decidable, but they were *very decidable*. I had heard rumblings about the exciting new field of complexity theory from my roommate John Gill (I, like most mathematics graduate students at the time, had never taken a course in computer science), and I decided that complexity theory must be able to make precise this notion of *very decidable*. I immediately began studying complexity theory, and I soon realized that in a precise sense, the class of generalized spectra are exactly the class NP (nondeterministic polynomial time), while spectra are the class NE (nondeterministic exponential time).[5] Thus, not only are these classes "very decidable", but in fact they precisely characterize certain complexity classes. This is quite interesting, since spectra and generalized spectra are defined in terms of logic, without using any notion of machine, computation, or time (or for that matter, any notion of polynomial or exponential). We shall see later why generalized spectra are exponentially simpler (in terms of complexity) than spectra. Intuitively, it corresponds to the fact that the size of the input is exponentially bigger for generalized spectra than for spectra. For example, it takes around n^2 bits to encode a graph on n points, whereas it takes only around $\log_2 n$ bits to encode the number n.

The result that spectra coincide with NE was also obtained by Jones and Selman

[5]There are various notions of "exponential time" in the literature. In this paper, it means time $2^{k\ell}$ for some constant k, where ℓ is the length of the input.

[JS74]. (I should note that in my thesis, I erroneously attributed this result, based on hearsay reports, to Bennett, although he proved no such result.) The result on spectra gives evidence that Asser's problem (as to whether the complement of every spectrum is a spectrum) is hard: it is equivalent to the problem of whether nondeterministic exponential time is closed under complement (that is, whether the complement of every set recognizable in nondeterministic exponential time is also recognizable in nondeterministic exponential time). Questions as to whether nondeterministic time classes are closed under complement seem to be very difficult to resolve. (Immerman [Imm88] and Szelepcsényi [Sze88] recently discovered the extremely surprising result that nondeterministic *space* classes are closed under complement.)

The equivalence between NP and generalized spectra is of interest because of the fundamental importance of the class NP in theoretical computer science (see Cook [Coo71], Levin [Lev73], Karp [Kar75], and, for a comprehensive discussion, Garey and Johnson [GJ79]). As we shall see, this equivalence may allow us to convert a problem in complexity theory into an equivalent problem in logic, and vice versa. In particular, results in one field can sometimes be exploited to help the other field.

To obtain a precise statement of the equivalence between NP and generalized spectra, we must encode structures by strings over some finite alphabet. For example, for simplicity assume that $\mathcal{L} = \{P_1, P_2\}$, where P_1 is a binary relation symbol and P_2 is a ternary relation symbol. Let $\mathcal{A}$ be an $\mathcal{L}$-structure with universe $\{0, \ldots, n-1\}$, where the interpretation of P_1 is R_1 and the interpretation of P_2 is R_2. Define b_1 to be a string of 0's and 1's of length n^2, where the $(in+j)$th member of the sequence b_1 is 1 precisely if $(i, j) \in R_1$ (for $0 \leq i \leq n-1$ and $1 \leq j \leq n$). Thus, b_1 encodes R_1. Similarly, let b_2 encode R_2. The encoding $enc(\mathcal{A})$ of $\mathcal{A}$ is the string $a\#b_1\#b_2$, where a is the binary encoding of n, and $\#$ is a new symbol. If $\mathcal{Q}$ is a set of $\mathcal{L}$-structures, then let $Enc(\mathcal{Q}) = \{enc(\mathcal{A}) \mid \mathcal{A} \in \mathcal{Q}\}$.

Theorem 5.1: [Fag74] *Let $\mathcal{L}$ be a nonempty language, and let $\mathcal{Q}$ be a set of finite $\mathcal{L}$-structures that is closed under isomorphism. Then $\mathcal{Q}$ is a generalized spectrum iff $Enc(\mathcal{Q})$ is in NP.*

Note that the assumption of Theorem 5.1 that $\mathcal{L}$ is nonempty is needed, since otherwise we are in the spectrum case, where the corresponding complexity class is NE rather than NP.

Just as there is a simple way to encode each structure by a string of symbols, there is also a simple way to encode each string of symbols (which we assume for simplicity is a string of 0's and 1's) by a structure. Let $\mathcal{L} = \{U, <\}$, where U is a unary relation symbol and $<$ is a binary relation symbol. We encode the string $s = s_1 \cdots s_n$, where each s_i is either 0 or 1, by the $\mathcal{L}$-structure $\mathcal{A}$ with universe $\{0, \ldots, n-1\}$, where the interpretation of $<$ in $\mathcal{A}$ is a linear order of the universe, and where for each i, the ith element in the linear order is in the interpretation of U precisely if $s_i = 1$. We can then obtain a variation of Theorem 5.1, which says that a set of strings is in NP iff the set of structures that encode these strings is a generalized spectrum.

The reason that a generalized spectrum, defined by $\exists Q_1 ... \exists Q_k \sigma(P_1, \ldots, P_s, Q_1, \ldots, Q_k)$, over a nonempty language $\mathcal{L} = \{P_1, \ldots, P_s\}$, corresponds to an NP set is that to decide if a structure $\mathcal{A}$ over $\mathcal{L}$ satisfies this sentence, a nondeterministic polynomial-time Turing machine can simply "guess" the relations corresponding to the extra relation symbols $Q_1, \ldots, Q_k$, and then verify that the structure that is the result of expanding $\mathcal{A}$ to include the new relations satisfies σ. This verification runs in deterministic polynomial time: all of the nondeterminism comes at the beginning, in guessing the extra relations. As for the other direction, the generalized spectrum that corresponds to an NP set is defined, roughly speaking, by a Σ_1^1 sentence $\exists {<} \exists T \varphi$, where φ says that $<$ is a linear order and that T is a nondeterministic Turing machine that accepts the encoding of the structure. The purpose of the linear order is to define an ordering on k-tuples of members of the universe $\{0, 1, \ldots, n-1\}$ for some fixed k. These k-tuples are used to index both time and tape squares (there are at most n^k steps, and hence at most n^k tape squares touched). Note that if the language $\mathcal{L}$ is nonempty, the time is polynomial, because, say, in the case of a graph, the length of the input is roughly n^2 (the number of bits needed to represent a binary relation on n points). A similar proof works for the spectrum case, but here $n^k = 2^{k \log_2 n}$ steps is exponential time, since the size of the input is roughly $\log_2 n$ in the case of spectra. For detailed proofs, see [Fag74].

Once we consider generalized spectra rather than just spectra, the natural question is the generalization of Asser's problem:

Open problem (Generalized Asser problem): Is the class of generalized spectra closed under complement? That is, is the complement of every generalized spectrum also a generalized spectrum?

For example, we saw that the class of 3-colorable finite graphs is a generalized spectrum. Is the class of finite graphs that are *not* 3-colorable also a generalized spectrum? It follows from the equivalence of generalized spectra with NP that the generalized Asser problem is equivalent to a fundamental problem in complexity theory.

Theorem 5.2: [Fag74] *The class of generalized spectra is closed under complement iff NP is closed under complement.*

Probably the most important problem in theoretical computer science is the question of whether P=NP. It is a trivial fact that P is closed under complement (by reversing accepting and rejecting states). It follows immediately that P=NP implies that NP is closed under complement. Therefore, a negative resolution of the generalized Asser problem (which I believe is the actual situation) implies a negative resolution of P=NP. It can be shown by techniques of Savitch [Sav70] that (1) if P=NP, then E=NE (where E represents deterministic exponential time), and (2) if NP is closed under complement then so is NE.[6] Now E is closed under complement for the same trivial reason that P

[6]These results show that we can "go up". It is an open problem as to whether we can "go down", that is, whether the converse of (1) or (2) holds. In fact, relativized versions of the converses of (1)

is closed under complement. So from (1) we see that P=NP implies that NE is closed under complement. Therefore, a negative resolution of the Asser question (which I also believe is the actual situation) implies a negative resolution of P=NP. Indeed, some people believe that resolving the Asser question is much harder than resolving P=NP!

Theorems 5.1 and 5.2 are "interdisciplinary" theorems, since they relate concepts in two distinct fields (logic and complexity theory). Such interdisciplinary theorems have (at least) two uses. One is to show the equivalence of important problems in two distinct fields. For example, Theorem 5.2 tells us that the generalized Asser problem is equivalent to a fundamental problem in complexity theory. Perhaps an even more important application of an interdisciplinary theorem is to prove a "pure" theorem (in only one field) by taking advantage of results from the other field. Let us consider an example, of a theorem of pure logic that would be of interest to a logician who does not care about complexity theory, that follows easily from Theorem 5.1.

Let $\mathcal{N}$ be the class of finite graphs that are not 3-colorable. If the complement of every generalized spectrum is a generalized spectrum, then of course $\mathcal{N}$ is a generalized spectrum. The interesting fact is that the converse holds. That is, if there are any "counterexamples" at all to the generalized Asser problem, then $\mathcal{N}$ is such a counterexample.

Theorem 5.3: [Fag74] *The complement of every generalized spectrum is a generalized spectrum iff the class of graphs that are not 3-colorable is a generalized spectrum.*

This result, which is a statement in pure logic that does not mention complexity theory at all, follows easily from Theorem 5.1 and from the fact [GJS76] that graph 3-colorability is an NP-complete problem (as with NP, for information on NP-completeness see Cook [Coo71], Levin [Lev73], Karp [Kar75], and, for a comprehensive discussion, Garey and Johnson [GJ79]). Any NP-complete class of finite graphs (such as the class of finite graphs with a Hamilton cycle) can be used instead of the class of 3-colorable graphs in Theorem 5.3.

As mentioned earlier, we can show that each set of positive integers in NE is the spectrum of a first-order sentence $\varphi(<, T)$, where φ says that $<$ is a linear order and that T is a nondeterministic Turing machine that accepts n in time exponential in the length of n. If the Turing machine is deterministic, then the sentence $\varphi(<, T)$ in the construction is *categorical*, which means that up to isomorphism, there is at most one model of each finite cardinality. Let us define a *categorical spectrum* to be the spectrum of a categorical sentence. From what we have said, we obtain:

Theorem 5.4: [Fag74] *Every set of positive integers in E is a categorical spectrum.*

---

and (2) are false, with a suitable choice of oracle. In the case of (1), this oracle result was claimed by Dekhtyar [Dek76] without proof, and proven independently by Wilson [Wil80] (see also [BWMR82]). In the case of (2), this oracle result was shown by Hartmanis, Immerman, and Sewelson [HIS85].

It is not clear as to whether the same is true about all spectra.

Open problem: Is every spectrum a categorical spectrum?

If P=NP, so that (as noted above) E=NE, we know from Theorem 5.4 that the open problem is true, that is, every spectrum is a categorical spectrum. This makes the question all the more intriguing, as to whether such a hypothesis is really needed.

Following Valiant [Val76], we say that an *unambiguous* Turing machine is a nondeterministic Turing machine that has at most one accepting computation for each input. Valiant defined the complexity class UP to consist of those languages accepted by an unambiguous Turing machine in polynomial time [Val76]. Similarly, let us define the complexity class UE to consist of those languages accepted by an unambiguous Turing machine in exponential time. Thus, the complexity class UE lies "between" E and NE. Theorem 5.4 can be strengthened to say that every set of positive integers in UE is a categorical spectrum. It is not clear as to whether the converse is true.

Open problem: Are categorical spectra precisely the sets of positive integers in UE?

The type of characterization of Theorem 5.1 has been obtained for a number of complexity classes (see Immerman [Imm87] and a survey by Immerman [Imm89]). The most interesting complexity class to consider is P (deterministic polynomial time). Ever since Cobham [Cob64] and Edmonds [Edm65], the class P has been often identified with the class of "feasible" problems. An *ordered structure* is one with a "built-in linear order", that is, a structure over a language that contains a binary relation symbol $<$, where the interpretation of $<$ in the structure is a linear order of the universe. Immerman [Imm86] and Vardi [Var82a] have independently obtained a theorem (Theorem 5.5 below) analogous to Theorem 5.1, where the complexity class of interest is P rather than NP.[7] However, for Immerman and Vardi's characterization to succeed, it is necessary to restrict attention to ordered structures. Instead of existential second-order sentences, they use sentences in *fixpoint logic* (this is first-order logic, augmented by the least fixpoint operator). An example of a fixpoint sentence is

$$\forall x \forall y \mu Rxy[(x = y)) \vee \exists z (Pxz \wedge Rzy)]. \tag{1}$$

We can think of R in (1) as being the transitive closure of P, since intuitively R is the least relation such that Rxy holds iff either $x = y$ or there is z such that Pxz and Rzy. Therefore, (1) says that the graph represented by P is strongly connected, since it says that every pair (x, y) of points is in the transitive closure. The idea of Immerman and Vardi's construction is to define a Turing machine computation by using fixpoints. The linear order that is defined existentially in the proof of the "only if" direction of Theorem 5.1 cannot be defined by a fixpoint, but needs to be built in.

Theorem 5.5: (Immerman [Imm86], Vardi [Var82a]) *Let $\mathcal{Q}$ be a set of finite ordered $\mathcal{L}$-structures that is closed under isomorphism. Then $\mathcal{Q}$ is definable by a fixpoint sentence iff $Enc(\mathcal{Q})$ is in P.*

[7] Livchak [Liv82] obtained a related result; see Gurevich [Gur84] for a discussion.

Chandra and Harel [CH82] have shown that if we were to drop the word "ordered" in Theorem 5.5, then the theorem would be false. In particular, they show that with a fixpoint sentence it is not possible to express evenness ("The cardinality of the universe is even"). That is, there is no fixpoint sentence σ about, say, graphs, such that a graph G satisfies σ precisely if G has an even number of nodes.

Open problem: Find a logic that gives a characterization analogous to Theorem 5.5 for P, where it is not necessary to restrict to ordered structures.

Chandra and Harel [CH82] pose as an open problem a technical result closely related to this open problem. Namely, they ask if the class of polynomial-time computable queries is r.e. (requirements are that there must be a recursive procedure for going from a "query" to a polynomial-time algorithm for evaluating it, and that a query must behave the same on isomorphic inputs). Gurevich [Gur88] gives a formal statement of what it means for a logic to capture P, and conjectures that there is no such logic. Cai, Fürer, and Immerman [CFI89] show that a candidate logic ("first-order + least fixpoint + counting") does not capture all polynomial-time properties of graphs.

We have seen that complexity theory can be helpful to the logician. What about the other direction? Can logic be used as a tool to obtain pure complexity-theoretic results? By using logic, Ajtai [Ajt83] proved the important result (which was obtained independently by Furst, Saxe, and Sipser [FSS81] without using logic) that the parity function cannot be realized by a family of bounded-depth, polynomial-size circuits with unbounded fan-in. As another example, by thinking in terms of logic, Immerman [Imm88] discovered the extremely surprising result that we mentioned earlier that nondeterministic space classes are closed under complement (this result, too, was obtained independently without using logic, by Szelepcsényi [Sze88]).

6 0-1 Laws

In this section, I discuss a topic that is unique to *finite*-model theory, namely 0-1 laws. As before, I begin with my personal history.

Let S be the spectrum of the first-order sentence σ. Asser's problem asks whether the complement $\overline{S}$ of S is also a spectrum. Since complementation corresponds to taking a negation, the first and most naive thing to try in resolving Asser's problem is to consider the negation $\neg\sigma$, in the hope that the spectrum of $\neg\sigma$ might be $\overline{S}$. Of course, this doesn't work. For example, let φ be the conjunction of the field axioms, so that the spectrum of φ is the set of powers of primes. The spectrum of $\neg\varphi$ is *not* the set of numbers that are not powers of primes. Instead, it is easy to see that the spectrum of $\neg\varphi$ is the set of all positive integers. In fact, the negation $\neg\varphi$ is a very uninteresting sentence, since it is very easy to violate some field axiom. While playing with such examples, I began to wonder whether we have the following general phenomenon:

(*) "If σ is a first-order sentence, then either σ or $\neg\sigma$ is very uninteresting."

The way I made precise the notion of "very uninteresting" was by considering probabilities. If φ is the conjunction of the field axioms, then $\neg\varphi$ is "almost surely true", since intuitively, almost any way one defines the plus and times relations, the result will be a non-field. Define $\mu_n(\sigma)$ to be the fraction of structures with universe $\{0, 1, \ldots, n-1\}$ that satisfy a sentence σ. Equivalently, if we probabilistically generate a structure with n nodes, where each possible tuple appears with probability $\frac{1}{2}$, independently of the other tuples, then $\mu_n(\sigma)$ is the probability that the structure satisfies σ. The sentence σ is *almost surely true* if its asymptotic probability is 1, that is, if $\mu_n(\sigma)$ converges to 1 as n goes to infinity.

If we identify "very uninteresting" with "almost surely true", then (*) above says that every first-order sentence or its negation is almost surely true. Remarkably enough, this is indeed the case!

Theorem 6.1: (0-1 law) *Let σ be a first-order sentence. Either σ or $\neg\sigma$ is almost surely true.*

This is called a 0-1 law since it says that $\mu_n(\sigma)$ always converges as n goes to infinity, and to either 0 or 1. Moshe Vardi has made an interesting observation about the 0-1 law. There are three possibilities for a sentence: it can be *surely true* (valid), it can be *surely false* (unsatisfiable), or it can be neither. The third possibility (where a sentence is neither valid nor unsatisfiable) is the "common" case. When we consider asymptotic probabilities, there are *a priori* three possibilities: it can be *almost surely true* (asymptotic probability 1), it can be *almost surely false* (asymptotic probability 0), or it can be neither (either because there is no limit, or because the limit exists and is not 0 or 1). Again, we might expect the third possibility to be the common case. The 0-1 law says that the third possibility is not only not the common case, but it is in fact impossible!

If a logic (such as first-order logic) has a 0-1 law, then an immediate corollary is that there is no sentence in the logic that can express evenness. That is, there is no sentence σ in the logic about, say, graphs, such that a graph G satisfies σ precisely if G has an even number of nodes. This is because if σ expressed evenness, then $\mu_n(\sigma)$ would be 0 if n were odd, and 1 if n were even, so $\mu_n(\sigma)$ would not converge. The fact that no first-order sentence can express evenness can also be proven by other means, such as an elimination of quantifiers argument or an Ehrenfeucht-Fraïssé game argument. However, there is another logic that I shall discuss shortly where the only reason we know that evenness is not expressible is because there is a 0-1 law.

We know that existential second-order logic can express evenness (since the equivalence between generalized spectra and NP tells us that existential second-order logic can express any NP property, and in particular evenness). Therefore, existential second-order logic (or any logic that is rich enough to be able to express evenness) does not have a 0-1 law.

The 0-1 law was first proved by Glebskiĭ, Kogan, Liogon'kiĭ, and Talanov and published in Russian in 1969 [GKLT69]; an English translation appeared in 1972. Their proof uses an elimination of quantifiers argument. Without knowing about their results, I proved the 0-1 law in 1971, and wrote a summary for the American Mathematical Society Notices in 1972 [Fag72]. Although it appeared in my thesis in 1973, the journal version [Fag76] did not appear until 1976, due to refereeing delays. I now describe my proof, which is quite different from that of Glebskiĭ et al.

For ease in description, let us assume for now that the language consists of a single binary relation symbol P. Thus, we can view the language as talking about directed graphs, where P denotes the edge relation. Everything we say generalizes in a natural way to arbitrary languages. Let us define a set T of *extension axioms*, which say that for each finite set X of points, and each possible way that a point $y \notin X$ could relate to X in terms of atomic formulas, there indeed is such a point y. For example, if X contains exactly two points x_1 and x_2, such an extension axiom would be

$$\forall x_1 \forall x_2 (x_1 \neq x_2 \Rightarrow \exists y (y \neq x_1 \wedge y \neq x_2 \wedge \psi(x_1, x_2, y))),$$

where $\psi(x_1, x_2, y)$ is a conjunction of four atomic formulas: one of Px_1y and $\neg Px_1y$, one of Pyx_1 and $\neg Pyx_1$, one of Px_2y and $\neg Px_2y$, and one of Pyx_2 and $\neg Pyx_2$. For example, one of the 16 possibilities for $\psi(x_1, x_2, y)$ is

$$Px_1y \wedge \neg Pyx_1 \wedge Px_2y \wedge Pyx_2.$$

According to Lynch [Lyn80], this theory T was discovered in 1958 by Jaśkowski (unpublished), and it was proven $\aleph_0$-categorical (that is, every two countable models are isomorphic) by Ehrenfeucht and Ryll-Nardzewski (unpublished) using Cantor's "back-and-forth" argument. Given any two countable models $\mathcal{A}$ and $\mathcal{B}$ of T, the idea of the back-and-forth argument is to build up an isomorphism, step-by-step, between $\mathcal{A}$ and $\mathcal{B}$ by using the extension axioms. The first published proof that T is $\aleph_0$-categorical is due to Gaifman [Gai64], who also uses the back-and-forth argument.

Let $\mathcal{R}$ be the unique (up to isomorphism) countable graph that satisfies T. This graph was studied by Rado [Rad64], and is sometimes called the *Rado graph*. It is also sometimes called the *random countable graph*, since with probability 1, this graph is generated by a random process where we start with a countable set of nodes, and where each possible edge appears with probability $\frac{1}{2}$, independently of the other edges [ER63, Gai64] (see also [ES74, pp. 98-99]).

A useful feature of the random countable graph (that will be exploited later) is that every countable graph is embeddable in it. Thus, let us say that a graph $\mathcal{H}$ with universe H is a *subgraph* of the graph $\mathcal{G}$ if the edges of $\mathcal{H}$ are precisely those edges of $\mathcal{G}$ where both endpoints are in H. We say that graph $\mathcal{H}$ is *embeddable* in graph $\mathcal{G}$ if $\mathcal{H}$ is isomorphic to a subgraph of $\mathcal{G}$. A graph is *universal* [Jón56] if every countable graph is embeddable in it. The following well-known proposition follows easily by making use of the extension axioms.

Proposition 6.2: *The random countable graph is universal.*

My proof of the 0-1 law now proceeds by proving the following theorem.

Theorem 6.3: [Fag76] *Let $\mathcal{R}$ be the random countable graph, let T be the set of extension axioms, and let σ be a first-order sentence. The following are equivalent.*

1. *$\mathcal{R} \models \sigma$, that is, σ is true in $\mathcal{R}$.*

2. *$T \models \sigma$, that is, σ is a logical consequence of T.*

3. *σ is almost surely true.*

Proof: (1) $\Rightarrow$ (2): Assume that $\mathcal{R} \models \sigma$, but $T \not\models \sigma$. So $T \cup \{\neg\sigma\}$ is consistent. Now T, and hence $T \cup \{\neg\sigma\}$, clearly has no finite models, so $T \cup \{\neg\sigma\}$ has an infinite model. By the Löwenheim-Skolem Theorem [End72], $T \cup \{\neg\sigma\}$ has a countable model $\mathcal{R}'$. By $\aleph_0$-categoricity of T, it follows that $\mathcal{R}'$ is isomorphic to $\mathcal{R}$. But this is impossible, since $\mathcal{R} \models \sigma$, but $\mathcal{R}' \models \neg\sigma$.

(2) $\Rightarrow$ (3): By a simple combinatorial argument [Fag76, p. 52], it follows that every member of T is almost surely true. Assume that $T \models \sigma$. By the Compactness Theorem, for some finite subset $T' \subseteq T$ we have $T' \models \sigma$. Since every member of T, and hence T', is almost surely true, so is the (finite) conjunction of the members of T'. Since $T' \models \sigma$, it follows that σ is almost surely true.

(3) $\Rightarrow$ (1): Assume that $\mathcal{R} \not\models \sigma$. Since $\mathcal{R}$ is a structure and $\mathcal{R} \not\models \sigma$, we know that $\mathcal{R} \models \neg\sigma$. By the proof we have already given that (1) $\Rightarrow$ (2) $\Rightarrow$ (3), where we let $\neg\sigma$ play the role of σ, it follows that $\neg\sigma$ is almost surely true. Therefore, it is certainly not the case that σ is almost surely true. ∎

The equivalence of (1) and (3) is called the *transfer property*. The transfer property says that a first-order sentence about graphs is almost surely true (where we are referring to finite graphs) precisely if it is true about the (infinite) graph $\mathcal{R}$. The 0-1 law follows immediately from the transfer property. This is because if σ is a first-order sentence, then either $\mathcal{R} \models \sigma$ or $\mathcal{R} \models \neg\sigma$, so by the transfer property, either σ or $\neg\sigma$ is almost surely true. An interesting feature of my proof of the 0-1 law is that, even though the 0-1 law is a theorem of finite-model theory, where the Compactness Theorem does not apply, the proof uses the Compactness Theorem!

As we shall discuss, the transfer property has proven useful in proving 0-1 laws for other logics (as has the use of the extension axioms).

The equivalence of (1) and (2) appears in Gaifman [Gai64]. It follows easily from this equivalence that T is *complete*, that is, for every sentence σ in the language of T, either $T \models \sigma$ or $T \models \neg\sigma$. Since T is also r.e., it follows that the set of logical consequences of T, that is, the set of sentences that are almost surely true, is decidable (to decide if σ or $\neg\sigma$ is a consequence of T, just list all proofs involving only hypotheses in T, and look for either σ or $\neg\sigma$ as a conclusion). This contrasts in an interesting way with Trakhtenbrot's Theorem. Although we just showed that the set of first-order sentences that are *almost surely true* is decidable, Trakhtenbrot's Theorem tells us that the set of first-order sentences that are *surely true* (valid over finite structures) is undecidable.

Grandjean has characterized the complexity of deciding if a first-order sentence is almost surely true.

Theorem 6.4: (Grandjean [Gra83]) *The problem of deciding if a first-order sentence is almost surely true is PSPACE-complete.*

The 0-1 law has been extended to logics richer than first order. Blass, Gurevich and Kozen [BGK85] and Talanov and Knyazev [TK84] have shown that the 0-1 law holds for fixpoint logic. (Note that this is consistent with Chandra and Harel's result [CH82] that we mentioned earlier that in fixpoint logic, it is not possible to express evenness.) Kolaitis and Vardi have shown that the 0-1 law holds for iterative logic [KV87] (which subsumes fixpoint logic), for the infinitary logic $L^{\omega}_{\infty\omega}$ (which subsumes iterative logic), where disjunctions and conjunctions may be infinitely long, but where there are only a finite number of distinct variables in a sentence [KV90b], and for certain interesting fragments of existential second-order logic [KV87, KV88]. Let us now focus on this last result, about fragments of existential second-order logic.

A *prefix class* is a class of first-order sentences (in prenex normal form) defined by a quantifier prefix. Several prefix classes are of special interest:

- The *Bernays-Schönfinkel class*, where the quantifier structure is $\exists^*\forall^*$ (that is, the existential quantifiers all precede the universal quantifiers).
- The *Ackerman class*, where the quantifier structure is $\exists^*\forall\exists^*$ (that is, there is exactly one universal quantifier).
- The *Gödel class*, where the quantifier structure is $\exists^*\forall\forall\exists^*$ (that is, there are exactly two universal quantifiers, and they are adjacent).

A prefix class is called *solvable* if the decision problem for satisfiability of sentences in the class is decidable. The only solvable prefix classes are the Bernays-Schönfinkel class and the Ackerman class [DG79, Lew79, Gol84].[8] These classes were proven solvable by showing that they are finitely controllable (see Theorem 3.6). Of course, they are the only finitely controllable prefix classes, since finite controllability implies solvability.

The 0-1 law does not hold for existential second-order logic (since it is possible to express evenness in this logic, that is, with a Σ^1_1 sentence). In fact, Kaufmann and

[8] Our attention is restricted to sentences with equality. Otherwise, there is an additional solvable prefix class, namely the Gödel class without equality.

Shelah [KS85] showed that the 0-1 law fails for monadic second-order logic, and Kaufmann [Kau87] extended this to the existential monadic second-order case. Kolaitis and Vardi systematically considered whether or not the 0-1 law holds for existential second-order sentences $\exists Q_1 ... \exists Q_k \sigma$, where we restrict the prefix class for the first-order part σ. They proved that the 0-1 law holds when the first-order part is in either the Bernays-Schönfinkel class [KV87] or the Ackerman class [KV88],[9] and that it fails in certain other cases. They made the bold conjecture that the 0-1 law holds for the fragment of existential second-order logic where the first-order part must be in a given prefix class precisely when the prefix class is solvable. The final step in proving that their conjecture is indeed correct was given by Pacholski and Szwast [PS89], who showed that the 0-1 law fails for the fragment where the first-order part is in the Gödel class.

Recently, Kolaitis and Vardi [KV90a] discovered at least a partial explanation for this surprising correspondence. The connection they found is between the *tools* used for proving the 0-1 law on the one hand, and the tools used for proving the solvability of the prefix class on the other hand. In both of the cases where there is a 0-1 law (where the prefix class is the Bernays-Schönfinkel class or the Ackerman class), Kolaitis and Vardi proved the 0-1 law by proving (the analogue of) the transfer property. Further, as noted above, the Bernays-Schönfinkel class and the Ackerman class are the only finitely controllable prefix classes, and finite controllability is the tool used classically to prove solvability (via Theorem 3.6). Kolaitis and Vardi have found a simple argument, which I shall give shortly, that the transfer property implies finite controllability of the prefix class. One way to view this is that each prefix class $\mathcal{C}$ is either bad or very good. Being bad means that (a) $\mathcal{C}$ is not solvable, and (b) the 0-1 law fails for the fragment where the first-order part is in $\mathcal{C}$. Being very good means that the fragment where the first-order part is in $\mathcal{C}$ satisfies the transfer property. The transfer property simultaneously implies (a) the 0-1 law for the fragment (this is immediate) and (b) solvability of $\mathcal{C}$ (since the transfer property implies finite controllability of $\mathcal{C}$, which implies solvability of $\mathcal{C}$).

I now give Kolaitis and Vardi's argument that the transfer property implies finite controllability of the prefix class. Assume that we are dealing with sentences $\exists Q_1 ... \exists Q_k \sigma$ about graphs, where σ is in a given prefix class, and assume that σ is satisfiable. We wish to show that σ is finitely satisfiable. Since σ is satisfiable, it follows by the Löwenheim-Skolem Theorem that σ has a countable model, and so $\exists Q_1 ... \exists Q_k \sigma$ has a countable model $\mathcal{A}$. By Proposition 6.2, it follows that $\mathcal{A}$ is embeddable in the random countable graph. Therefore, the relativized sentence $\exists A \exists Q_1 ... \exists Q_k \sigma^A$ holds for the random countable graph, where A is a new unary relation symbol (intuitively, all first-order quantifiers are restricted to ranging through A). Since σ^A is in the same prefix class as σ, it follows by the transfer property that $\exists A \exists Q_1 ... \exists Q_k \sigma^A$ is almost surely true, and in particular has a finite model. It follows easily that σ has a finite model.

Considering logics other than first-order is just one direction that the 0-1 law has been

[9] The Ackerman case is the one alluded to earlier where we know that evenness is not expressible because of the 0-1 law, but where it is not clear how to show that evenness is not expressible without using the 0-1 law.

extended. I will now mention some other directions, all of which deal with first-order logic. In [Fag76] I showed that the 0-1 law holds also when we consider "unlabeled" rather than "labeled" structures (that is, where two structures are considered different precisely when they are nonisomorphic). Compton [Com87] has considered restricted classes of structures, and relates the existence of a 0-1 law with convergence properties of the exponential generating series for the class. Shelah and Spencer [SS88] have considered random graphs where the edge probability, instead of being a constant (such as $\frac{1}{2}$), is instead a function $p(n)$ of the number n of nodes. They show the fascinating result that if $p(n) = n^{-\alpha}$, where $0 < \alpha < 1$, then there is a 0-1 law iff α is irrational. This is related to the fact that $p(n) = n^{-\alpha}$ is a *threshold function* for the presence of some graph precisely if α is rational. For a discussion of threshold functions, see Erdös and Spencer [ES74]. For a comprehensive survey on 0-1 laws, see Compton [Com88].

If constant or function symbols were allowed into the language, then there would not be a 0-1 law. For example, if c is a constant symbol and U a unary relation symbol, then Uc has probability $\frac{1}{2}$. Similarly, if f is a unary function symbol, then $\forall x(f(x) \neq x)$ has asymptotic probability $\frac{1}{e}$ [Fag76]. Lynch [Lyn85] has shown that if the language contains only unary function symbols, then there is an asymptotic probability, even though, as we just saw, it need not be 0 or 1. Compton, Henson, and Shelah [CHS87] have shown that if the language consists of a single binary function symbol, then there is not necessarily an asymptotic probability.

Although there are now a number of cases where 0-1 laws are known to hold, and a number of cases where 0-1 laws are known not to hold, we don't really understand the deep reasons underlying this phenomenon. This leads to a very interesting (although necessarily murky) question, posed by Irv Traiger.

Question: What *really* causes there to be a 0-1 law?

7 Descriptive complexity

The *descriptive complexity* of a class is, informally, the complexity of describing the class in some logical formalism. For example, we might consider whether a class is first-order definable, and if so, we might want to ask, say, how many quantifiers are required. Hartmanis suggested the name "descriptive complexity" to Immerman, who used it in [Imm89]. This section will focus on descriptive complexity, where we are particularly interested in second-order quantifiers.

By definition, every generalized spectrum is defined by an existential second-order sentence $\exists Q_1 ... \exists Q_k \sigma(P_1, \ldots, P_s, Q_1, \ldots, Q_k)$, where $\sigma(P_1, \ldots, P_s, Q_1, \ldots, Q_k)$ is first-order. In my struggles with the Asser problem and generalized Asser problem, I began to wonder whether the problems would be easier if I were to restrict the extra relation symbols $Q_1, \ldots, Q_k$. I found that one case, indeed, is much more tractable than the general case. This tractable case occurs when the extra relation symbols are all restricted to being unary. Such generalized spectra are called *monadic*. It is not surprising that the monadic case should be simpler, since results like Church's Theorem and Trakhtenbrot's Theorem depend on the language containing some relation symbol that is not unary. Monadic generalized spectra are referred to as *monadic NP* in [FSV93].

We saw that the class of finite 3-colorable graphs is a monadic generalized spectrum. From this and Theorem 5.3, we know that the generalized Asser question is equivalent to the problem of whether there is a monadic generalized spectrum whose complement is not a generalized spectrum. I was able to show that there is a monadic generalized spectrum whose complement is not a *monadic* generalized spectrum. To show this, I first proved the next theorem, which gives us a natural example of a generalized spectrum that is not monadic.

Theorem 7.1: [Fag75a] *The class of finite connected graphs is not a monadic generalized spectrum.*

This proof of Theorem 7.1 appears in [Fag75a]. It involves an Ehrenfeucht-Fraïssé game argument and a characterization of (monadic) generalized spectra by Ehrenfeucht-Fraïssé games. Recently, Stockmeyer, Vardi and I obtained a much simpler proof [FSV93].

Let $\mathcal{Q}$ be the class of finite connected graphs. It is not hard to see [Fag75a, p. 96] that $\mathcal{Q}$ is a generalized spectrum with a single extra binary relation symbol (we already knew that $\mathcal{Q}$ is a generalized spectrum by Theorem 5.1, since $Enc(\mathcal{Q})$ is in P, and hence in NP). Although $\mathcal{Q}$ is a generalized spectrum, Theorem 7.1 tells us that $\mathcal{Q}$ is not a *monadic* generalized spectrum. The complement of $\mathcal{Q}$ (the class of finite graphs that are not connected) is a monadic generalized spectrum, via

$$\exists A(\exists x Ax \wedge \exists x \neg Ax \wedge \forall x \forall y(Pxy \Rightarrow (Ax \Leftrightarrow Ay))). \tag{2}$$

Thus, we can indeed answer a weak version of the generalized Asser problem.

Theorem 7.2: [Fag75a] *The class of monadic generalized spectra is not closed under complement.*

Some comments are in order about Theorem 7.1. In the nonfinite case, the analogue of Theorem 7.1 can be proven by a standard compactness argument. In fact, this same argument shows that not only is the class of connected graphs not definable by a *monadic* Σ_1^1 sentence, it is not definable by any Σ_1^1 sentence whatsoever. The argument is as follows. Assume that connectedness were definable by the Σ_1^1 sentence $\exists Q_1 ... \exists Q_k \sigma$, where σ is first-order. Let s and t be new constant symbols,[10] and let γ_k be a first-order sentence that says "There is no path between s and t of length at most k". Let Σ be the set $\{\sigma, \gamma_1, \gamma_2, \gamma_3, \ldots\}$. Clearly, every finite subset of Σ is satisfiable. By the Compactness Theorem, Σ is satisfiable. But this is impossible, since this gives us a graph that is connected, but where there is no path from s to t.

[10]Here we are allowing constant symbols into the language.

If all we care about is showing that connectedness is not first-order definable (as opposed to Σ_1^1 definable) in the finite case, then the result is much easier (although not as easy as the nonfinite case, where, as we just saw, there is an easy compactness argument). This result can be shown using an argument similar to one Fraïssé [Fra67] used to show that if T is the set of true first-order sentences in the language of successor about the successor relation on the natural numbers, then T is not equivalent to a single first-order sentence. This result can also be shown with an Ehrenfeucht-Fraïssé game argument similar to (but easier than) the argument I used to prove Theorem 7.1. It is interesting to note that this result (i.e., the finite first-order case) keeps getting rediscovered, for example by Aho and Ullman [AU79] (see survey in [GV85]).

Theorem 7.1 holds whether we consider directed or undirected graphs. There are times, however, when the directed and undirected cases are different. If s and t denote distinguished points in a directed (resp. undirected) graph, then a graph is (s,t)-*connected* if there is a directed (resp. undirected) path from s to t. In 1986, Paris Kanellakis pointed out to me that the class of (s,t)-connected *undirected* graphs is a monadic generalized spectrum.[11] I was surprised, since (s,t)-connectivity is very close in spirit to connectivity, which by Theorem 7.1 is not a monadic generalized spectrum (in fact, it is easy to fool oneself into believing that the proof of Theorem 7.1 can be modified so that we can replace "connected" by "(s,t)-connected"). To see that the class of (s,t)-connected undirected graphs is a monadic generalized spectrum, let ψ be a first-order sentence that says "The set A contains both s and t; every member of A except for s and t has an edge to precisely two members of A; and s (resp. t) has an edge to exactly one member of A."[12] I will now show that the Σ_1^1 sentence $\exists A\psi$ says that the graph is (s,t)-connected. For, if the graph is (s,t)-connected, and if A is taken to consist of those vertices on a shortest path from s to t, then ψ holds. Conversely, if ψ holds, then there is a path starting at s that passes through only vertices in A. The path must end somewhere, since the graph is finite; however, the only place it can end is at t. So the graph is (s,t)-connected.

Kanellakis asked whether the same is true about *directed* graphs. As Ajtai and I showed recently, the answer is "No".

Theorem 7.3: [AF90] *The class of (s,t)-connected finite directed graphs is not a monadic generalized spectrum.*

Let $\mathcal{Q}$ be a generalized spectrum defined by $\exists Q_1 \ldots \exists Q_k \sigma(P_1,\ldots,P_s,Q_1,\ldots,Q_k)$. Thus, the extra relation symbols are $Q_1,\ldots,Q_k$, and $\mathcal{Q}$ is a class of structures over $\mathcal{L} = \{P_1,\ldots,P_s\}$. Define $\mathcal{F}_m(\mathcal{L})$ to consist of all such generalized spectra, where the arity of each of the extra relation symbols is at most m. So there is a hierarchy

$$\mathcal{F}_1(\mathcal{L}) \subseteq \mathcal{F}_2(\mathcal{L}) \subseteq \mathcal{F}_3(\mathcal{L}) \subseteq \ldots$$

[11] Again we are allowing the constant symbols s and t into the language.

[12] I am here assuming that the language consists of a binary relation P where it is "built-in" that P represents a set of unordered pairs. Without this assumption, we can simply add a clause to ψ that says that the graph is undirected, namely $\forall x \forall y(Pxy \Leftrightarrow Pyx)$.

(The inclusion $\mathcal{F}_m(\mathcal{L}) \subseteq F_{m+1}(\mathcal{L})$ holds because of well-known techniques of simulating m-ary relations by $(m+1)$-ary relations.) If $\mathcal{L} = \emptyset$ (so that we are considering spectra), then it is well-known that by an elimination of quantifiers argument, it follows that $\mathcal{F}_1(\emptyset)$ contains only finite and cofinite sets. However, it is easy to see that the set of even positive integers is a member of $\mathcal{F}_2(\emptyset)$. Hence, we have the strict inclusion $\mathcal{F}_1(\emptyset) \subset \mathcal{F}_2(\emptyset)$. Similarly, for arbitrary languages $\mathcal{L}$, the class of $\mathcal{L}$-structures where the cardinality of the universe is even is in $\mathcal{F}_2(\mathcal{L})$ but not $\mathcal{F}_1(\mathcal{L})$. (Theorem 7.1 gives us another generalized spectrum in $\mathcal{F}_2(\mathcal{L})$ but not $\mathcal{F}_1(\mathcal{L})$ when $\mathcal{L}$ consists of a single binary relation symbol.)

Open problem: Is the hierarchy $\mathcal{F}_1(\mathcal{L}) \subseteq \mathcal{F}_2(\mathcal{L}) \subseteq \mathcal{F}_3(\mathcal{L}) \subseteq \ldots$ strict for every language $\mathcal{L}$? For some language $\mathcal{L}$?

One way to try to prove that the hierarchy is strict is to show that the hierarchy interleaves with a complexity hierarchy that is known to be strict. Thus, one might hope to exploit Cook's result [Coo72] that nondeterministic time complexity n^α is strictly more powerful than nondeterministic time complexity n^β, for $\alpha > \beta > 1$. The problem is that there is no k for which it is clear that every member of $\mathcal{F}_2(\mathcal{L})$ can be recognized nondeterministically in time n^k. For example, consider a formula $\exists Q\varphi$, where Q is a binary relation symbol, and where φ is first-order and has a quantifier prefix of 17 first-order quantifiers. Let $\mathcal{A}$ be a structure with universe of size n. What is the time to decide nondeterministically whether $\mathcal{A} \models \exists Q\varphi$? Under the naive approach, we would first guess the binary relation corresponding to Q, and then cycle through all n^{17} possibilities for the 17 variables quantified by the first-order quantifiers. Although the first step (guessing the binary relation) takes time only n^2, the next step takes time n^{17}.

Although I was not able to show that the hierarchy is strict, I was able to show that if the hierarchy collapses at some point, then it collapses from that point on:

Theorem 7.4: [Fag75b] *Assume that $\mathcal{F}_p(\mathcal{L}) = \mathcal{F}_{p+1}(\mathcal{L})$. Then $\mathcal{F}_q(\mathcal{L}) = \mathcal{F}_p(\mathcal{L})$ for each $q \geq p$.*

That is, either the hierarchy is strict, so that

$$\mathcal{F}_1(\mathcal{L}) \subset \mathcal{F}_2(\mathcal{L}) \subset \mathcal{F}_3(\mathcal{L}) \subset \ldots,$$

or else there is p such that

$$\mathcal{F}_1(\mathcal{L}) \subset \ldots \subset \mathcal{F}_p(\mathcal{L}) = \mathcal{F}_{p+1}(\mathcal{L}) = \mathcal{F}_{p+2}(\mathcal{L}) = \ldots.$$

Ajtai has proven the following result.

Theorem 7.5: (Ajtai [Ajt83]) *If $\mathcal{L}$ contains an m-ary relation symbol P, then the class of $\mathcal{L}$-structures where the P relation has an even number of tuples is not in $\mathcal{F}_{m-1}(\mathcal{L})$.*

However, this class (where the P relation has an even number of tuples) is clearly a generalized spectrum (in fact, Ajtai [Ajt83] shows that if $m \geq 2$, then this class is in $\mathcal{F}_m(\mathcal{L})$, although we do not need this fact). This gives the following corollary.

Corollary 7.6: *If $\mathcal{L}$ contains an m-ary relation symbol, then $\mathcal{F}_1(\mathcal{L}) \subset \mathcal{F}_2(\mathcal{L}) \subset \ldots \subset \mathcal{F}_{m-1}(\mathcal{L}) \subset \mathcal{F}_m(\mathcal{L})$.*

Proof: Find $q \geq m$ so large that the class of $\mathcal{L}$-structures where the P relation has an even number of tuples is in $\mathcal{F}_q(\mathcal{L})$ (as noted earlier, we could actually take $q = m$ if $m \geq 2$). By Theorem 7.5, we know that $\mathcal{F}_{m-1}(\mathcal{L}) \neq \mathcal{F}_q(\mathcal{L})$. If $\mathcal{F}_{i-1}(\mathcal{L}) = \mathcal{F}_i(\mathcal{L})$ for some $i \leq m$, then by Theorem 7.4, we would have $\mathcal{F}_{m-1}(\mathcal{L}) = \mathcal{F}_q(\mathcal{L})$, a contradiction. ∎

Two choices of the language $\mathcal{L}$ are of special interest: when $\mathcal{L}$ is empty (which corresponds to spectra), and when $\mathcal{L}$ consists of a single binary relation symbol P (which corresponds to generalized spectra about graphs). In these cases, Corollary 7.6 gives us no information beyond the simple facts, which we noted earlier, that $\mathcal{F}_1(\emptyset) \subset \mathcal{F}_2(\emptyset)$ and $\mathcal{F}_1(\{P\}) \subset \mathcal{F}_2(\{P\})$. For all we know, every spectrum is in $\mathcal{F}_2(\emptyset)$, and every generalized spectrum about graphs is in $\mathcal{F}_2(\{P\})$. That is, it seems possible that the only extra relation symbols that are required are binary. In fact, it is even conceivable that only *one* extra binary relation symbol suffices. The following open problems are from [Fag75b].

Open problem: Is there any spectrum that is not the spectrum of a sentence over the language of a single binary relation symbol? That is, does a single binary relation symbol suffice?

Open problem: Is there any generalized spectrum about graphs that cannot be defined by $\exists Q\varphi$, where Q is a binary relation symbol and φ is first-order? That is, does a single extra binary relation symbol suffice?

Some further results in descriptive complexity on finite models have been proven. For example, Immerman has obtained results on the number of quantifiers [Imm81] and the number of variables [Imm82] needed to define certain classes. Turán [Tur84] has obtained various descriptive complexity results, including the result that the class of finite graphs with a Hamilton cycle cannot be defined by a monadic second-order sentence (where we allow arbitrary quantification over sets). However, descriptive complexity has not "caught on" to the same extent as the work on relationships with complexity classes (Section 5) and on 0-1 laws (Section 6). I hope that there will be more work in this area.

8 Conclusions

Finite-model theory is that branch of model theory that focuses on *finite* structures. Moshe Vardi has suggested to me a classification of finite-model theory into three lines of research.

The first line of research he calls *negative*; here we consider theorems of model theory that fail for finite-model theory. Some such results are easy to find (such as the failure of the Compactness Theorem). Other such results are a little more difficult to prove, but still have simple, elegant proofs. For example, consider the Substructure Preservation Theorem [CK90], which says that if a first-order sentence σ has the property that every substructure of a model of σ is a model of σ, then σ is equivalent to a universal sentence.

This theorem fails for finite structures [Tai59] (see also [Gur84]). Still other such results are much harder to prove. A nice example is Ajtai and Gurevich's result [AG87] that Lyndon's Theorem (which says that monotone and positive classes coincide) fails for finite structures.

The second line of research could be called *preservative*; here we consider theorems of model theory that continue to hold for finite-model theory. Again, some such results are easy (such as theorems that tell us that Ehrenfeucht-Fraïssé games are necessary and sufficient for resolving expressibility in a given logic). Kolaitis recently showed me another easy (but slightly surprising) example. Although Craig's Theorem and Beth's Theorem fail for finite structures, Kolaitis observed that the closely related Robinson Consistency Theorem holds (the proof follows easily from Theorem 3.8).[13] Some results in the preservative line of research are harder to prove. One example is Theorem 7.1, which says that the class of finite connected graphs is not a monadic generalized spectrum. (As we noted, the same result holds in the nonfinite case, by a simple compactness argument.) Ajtai and Gurevich [AG89] showed that a Datalog query is first-order iff it is bounded (for definitions of these terms, see [GMSV87]). In the nonfinite case, this result follows by a straightforward compactness argument; however, the proof is much more difficult in the finite case.

Other possible results in the preservative line of research are extremely difficult, and their resolution would be a major breakthrough. Probably the best example is given by the generalized Asser problem, which is equivalent by Theorem 5.2 to the question of whether NP is closed under complement. However, the corresponding question in the nonfinite case, as to whether every Σ_1^1 class is closed under complement, is well-known to have a negative answer. As a counterexample, the class of (not necessarily finite) graphs that are not connected is Σ_1^1, via (2) of Section 7. However, I already noted that the complement (the class $\mathcal{Q}$ of connected graphs) is not Σ_1^1. Actually, since $\mathcal{Q}$ is not first-order, the fact that $\mathcal{Q}$ is not Σ_1^1 is an instance of a general phenomenon. Specifically, in the nonfinite case it follows from the Craig Interpolation Theorem that if a class and its complement are both Σ_1^1, then the class is first-order definable. A logic, such as first-order logic, that obeys this property is called Δ-closed; cf. [Bar85]. Of course, first-order logic is not Δ-closed if we restrict our attention to finite structures: for example, evenness is Σ_1^1, as is the complement ("oddness"), but, as we have noted, evenness is not first-order definable.

The final line of research Vardi calls *positive*; here we consider results that are unique to finite-model theory. Good examples are the results of Section 5 (the relationship to complexity classes) and of Section 6 (on 0-1 laws). Another example is Immerman's result [Imm86] that for finite structures, fixpoint logic is closed under complement, and the related result by Gurevich and Shelah [GS86] that for finite structures, different natural fixed-point logics all have the same expressive power. A recent nice example was

[13]This contrasts in an interesting way with results in "abstract model theory" (cf. [Bar85, Mak85]), where Robinson consistency is a very strong property, and in fact is equivalent to the combination of compactness and Craig interpolation.

obtained by Kolaitis [Kol90]. As noted earlier, Beth's Theorem (which says that implicit definability is equivalent to explicit definability) fails for finite structures. Rather than stopping there, Kolaitis explores the expressive power of implicit definability on finite structures, and obtains very interesting results.

Finite-model theory is a fascinating area, on the borderline of logic, computer science, and combinatorics. I hope and believe that it will flourish even more in the future.

Acknowledgments: I am grateful to Yuri Gurevich, Phokion Kolaitis, Steven Lindell, and Moshe Vardi, for interesting and useful discussions. I would like to thank them, along with Miki Ajtai, Kevin Compton, Joe Halpern, Neil Immerman, Russell Impagliazzo, Jim Lynch, and Bob Vaught for reading a draft of the paper and giving me comments. Eric Allender and Ron Book provided some useful bibliographic information. Finally, I would like to thank Paris Kanellakis and Serge Abiteboul for suggesting that I write this paper.

References

[AF90] M. Ajtai and R. Fagin. Reachability is harder for directed than for undirected finite graphs. *Journal of Symbolic Logic*, 55(1):113–150, March 1990.

[AG87] M. Ajtai and Y. Gurevich. Monotone versus positive. *Journal of the ACM*, 34:1004–1015, 1987.

[AG89] M. Ajtai and Y. Gurevich. DATALOG vs. first-order logic. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 142–146, 1989.

[Ajt83] M. Ajtai. Σ^1_1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[Ass55] G. Asser. Das Repräsentantenproblem im Prädikatenkalkül der ersten Stufe mit identität. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 1:252–263, 1955.

[AU79] A. V. Aho and J. D. Ullman. Universality of data retrieval languages. In *Proc. 6th ACM Symp. on Principles of Programming Languages*, pages 110–117, 1979.

[Bar85] J. Barwise. Model-theoretic logics: background and aims. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 3–23. Springer-Verlag, Berlin/New York, 1985.

[Ben62] J. H. Bennett. *On spectra.* PhD thesis, Princeton University, 1962.

[BGK85] A. Blass, Y. Gurevich, and D. Kozen. A zero–one law for logic with a fixed point operator. *Information and Control*, 67:70–90, 1985.

[BWMR82] R. Book, C. Wilson, and X. Mei-Rui. Relativizing time, space, and time-space. *SIAM Journal on Computing*, 11:571–581, 1982.

[CFI89] J. Cai, M. Fürer, and N. Immerman. An optimal lower bound on the number of variables for graph identification. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 612–617, 1989.

[CH82] A. Chandra and D. Harel. Structure and complexity of relational queries. *Journal of Computer and System Sciences*, 25:99–128, 1982.

[Cha88] A. Chandra. Theory of database queries. In *Proc. 7th ACM Symp. on Principles of Database Systems*, pages 1–9, 1988.

[CHS87] K. J. Compton, C. W. Henson, and S. Shelah. Nonconvergence, undecidability, and intractability in asymptotic problems. *Annals of Pure and Applied Logic*, 36:207–224, 1987.

[Chu56] A. Church. *Introduction to Mathematical Logic*, volume I. Princeton University Press, Princeton, N.J., 1956.

[CK90] C. C. Chang and H. J. Keisler. *Model Theory.* North-Holland, Amsterdam, 3rd edition, 1990.

[Cob64] A. Cobham. The intrinsic computational difficulty of functions. In Y. Bar-Hillel, editor, *Proc. 1964 International Congress for Logic, Methodology, and Philosophy of Science*, pages 24–30. North-Holland, Amsterdam, 1964.

[Com87] K. J. Compton. A logical approach to asymptotic combinatorics I: First-order properties. *Advances in Mathematics*, 65:65–96, 1987.

[Com88] K. Compton. 0-1 laws in logic and combinatorics. In I. Rival, editor, *Proc. 1987 NATO Adv. Study Inst. on algorithms and order*, pages 353–383. Reidel, Dordrecht, Netherlands, 1988.

[Coo71] S. A. Cook. The complexity of theorem proving procedures. In *Proc. 3rd ACM Symp. on Theory of Computing*, pages 151–158, 1971.

[Coo72] S. A. Cook. A hierarchy for nondeterministic time complexity. In *Proc. 4th ACM Symp. on Theory of Computing*, pages 187–192, 1972.

[Dek76] M. I. Dekhtyar. On the relativization of deterministic and nondeterministic complexity classes. In *Proc. 5th Conference on Mathematical Foundations of Computer Science*, Lecture Notes in Computer Science, Vol. 45, pages 255–259. Springer-Verlag, Berlin/New York, 1976.

[DG79] D. Dreben and W. D. Goldfarb. *The Decision Problem: Solvable Classes of Quantificational Formulas.* Addison-Wesley, 1979.

[Edm65] J. Edmonds. Paths, trees, and flowers. *Canadian J. Math.*, 17:449–467, 1965.

[Ehr61] A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Math.*, 49:129–141, 1961.

[End72] H. B. Enderton. *A Mathematical Introduction to Logic.* Academic Press, New York, 1972.

[ER63] P. Erdös and A. Rényi. Asymmetric graphs. *Acta Math. Acad. Sci. Hung. Acad. Sci.*, 14:295–315, 1963.

[ES74] P. Erdös and J. Spencer. *Probabilistic Methods in Combinatorics.* Academic Press, New York, 1974.

[Fag72] R. Fagin. Probabilities on finite models. *Notices of the American Mathematical Society*, page A714, October 1972. Abstract number 72T-E90.

[Fag73] R. Fagin. *Contributions to the model theory of finite structures.* PhD thesis, University of California at Berkeley, 1973.

[Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings, Vol. 7*, pages 43–73, 1974.

[Fag75a] R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.

[Fag75b] R. Fagin. A spectrum hierarchy. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:123–134, 1975.

[Fag75c] R. Fagin. A two-cardinal characterization of double spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:121–122, 1975.

[Fag76] R. Fagin. Probabilities on finite models. *Journal of Symbolic Logic*, 41(1):50–58, 1976.

[Fag82] R. Fagin. Horn clauses and database dependencies. *Journal of the ACM*, 29(4):952–985, October 1982.

[Fra54] R. Fraïssé. Sur quelques classifications des systèmes de relations. *Publ. Sci. Univ. Alger. Sér. A*, 1:35–182, 1954.

[Fra67] R. Fraïssé. *Cours de Logique Mathematique.* Gauthier-Villars and E. Nauwelaerts, Paris, 1967. English translation is *Course in Mathematical Logic*, Reidel, Dordrecht, Netherlands, 1973, translated by D. Louvish.

[FSS81] M. Furst, J. B. Saxe, and M. Sipser. Parity, circuits, and the polynomial time hierarchy. In *Proc. 22nd IEEE Symp. on Foundations of Computer Science*, pages 260–270, 1981.

[FSV93] R. Fagin, L. Stockmeyer, and M. Y. Vardi. On monadic NP vs. monadic co-NP. In *Proc. 8th IEEE Conf. on Structure in Complexity Theory*, pages 19–30, 1993. To appear in *Information and Computation.*

[FV86] R. Fagin and M. Y. Vardi. The theory of data dependencies: a survey. In *Mathematics of Information Processing, Proceedings of Symposia in Applied Mathematics*, volume 34, pages 19–72. American Mathematical Society, 1986.

[Gai64] H. Gaifman. Concerning measures in first-order calculi. *Israel Journal of Mathematics*, 2:1–18, 1964.

[GJ79] M. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness.* W. Freeman and Co., San Francisco, Calif., 1979.

[GJS76] M. Garey, D. S. Johnson, and L. J. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.

[GKLT69] Y. V. Glebskiĭ, D. I. Kogan, M. I. Liogon'kiĭ, and V. A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Kibernetika*, 2:17–28, 1969.

[GMSV87] H. Gaifman, H. Mairson, Y. Sagiv, and M. Y. Vardi. Undecidable optimization problems for database logic programs. In *Proc. 2nd IEEE Symp. on Logic in Computer Science*, pages 106–115, 1987.

[Gol84] W. D. Goldfarb. The Gödel class with equality is unsolvable. *Bull. Amer. Math. Soc. (New Series)*, 10:113–115, 1984.

[Gra83] E. Grandjean. Complexity of the first–order theory of almost all structures. *Information and Control*, 52:180–204, 1983.

[GS86] Y. Gurevich and S. Shelah. Fixed-point extensions of first-order logic. *Annals of Pure and Applied Logic*, 32:265–280, 1986.

[Gur76] Y. Gurevich. The decision problem for standard classes. *Journal of Symbolic Logic*, 41:460–464, 1976.

[Gur84] Y. Gurevich. Toward logic tailored for computational complexity. In M. Richter et al., editors, *Computation and Proof Theory*, Lecture Notes in Math. 1104, pages 175–216. Springer-Verlag, Berlin/New York, 1984.

[Gur88] Y. Gurevich. Logic and the challenge of computer science. In E. Börger, editor, *Current Trends in Theoretical Computer Science*, pages 1–57. Computer Science Press, Rockville, Md., 1988.

[Gur90] Y. Gurevich. On finite model theory. In S. R. Buss and P. J. Scott, editors, *Perspectives in Computer Science*. Birkhauser, 1990. Proceedings of Feasible Mathematics Workshop, Cornell University, June 1989.

[GV85] H. Gaifman and M. Y. Vardi. A simple proof that connectivity is not first-order. *Bulletin of the European Association for Theoretical Computer Science*, 26:43–45, June 1985.

[Háj75] P. Hájek. On logics of discovery. In *Proc. 4th Conference on Mathematical Foundations of Computer Science, Springer-Verlag Lecture Notes in Computer Science 32*, pages 30–45, 1975.

[Háj77] P. Hájek. Generalized quantifiers and finite sets. In *Prace Nauk. Inst. Mat. Politech. Wroclaw No. 14 Ser. Konfer. No. 1*, pages 91–104, 1977.

[HIS85] J. Hartmanis, N. Immerman, and J. Sewelson. Sparse sets in NP-P – EXPTIME vs. NEXPTIME. *Information and Control*, 65:159–181, 1985.

[Imm81] N. Immerman. Number of quantifiers is better than number of tape cells. *Journal of Computer and System Sciences*, 22:384–406, 1981.

[Imm82] N. Immerman. Upper and lower bounds for first-order expressibility. *Journal of Computer and System Sciences*, 25:76–98, 1982.

[Imm86] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68:76–98, 1986.

[Imm87] N. Immerman. Languages that capture complexity classes. *SIAM Journal on Computing*, 16(4):760–778, 1987.

[Imm88] N. Immerman. Nondeterministic space is closed under complement. *SIAM Journal on Computing*, 17:935–938, 1988.

[Imm89] N. Immerman. Descriptive and computational complexity. In J. Hartmanis, editor, *Computational Complexity Theory, Proc. Symp. in Applied Math. 38*, pages 75–91. American Mathematical Society, 1989.

[Jón56] B. Jónsson. Universal relational systems. *Math. Scand.*, 4:137–142, 1956.

[JS74] N. D. Jones and A. L. Selman. Turing machines and the spectra of first-order formulas with equality. *Journal of Symbolic Logic*, pages 139–150, 1974.

[Kan90] P. C. Kanellakis. Elements of relational database theory. In A. R. Meyer, M. Nivat, M. S. Paterson, D. Perrin, and J. van Leeuwen, editors, *The handbook of theoretical computer science*. North-Holland, Amsterdam, 1990. Also available as Brown University Technical Report CS-89-39, 1989.

[Kar75] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1975.

[Kau87] M. Kaufmann. Counterexample to the 0-1 law for existential monadic second-order logic. CLI Internal Note 32, Computational Logic Inc., December 1987.

[Kol90] Ph. G. Kolaitis. Implicit definability on finite structures and unambiguous computations. In *Proc. 5th IEEE Symp. on Logic in Computer Science*, pages 168–180, 1990.

[KS85] M. Kaufmann and S. Shelah. On random models of finite power and monadic logic. *Discrete Mathematics*, 54:285–293, 1985.

[KV87] Ph. G. Kolaitis and M. Y. Vardi. The decision problem for the probabilities of higher-order properties. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 425–435, 1987.

[KV88] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws and decision problems for fragments of second-order logic. In *Proc. 3rd IEEE Symp. on Logic in Computer Science*, pages 2–11, 1988.

[KV90a] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws for fragments of second-order logic. Research Report RJ 7508, IBM, 1990.

[KV90b] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws for infinitary logics. In *Proc. 5th IEEE Symp. on Logic in Computer Science*, pages 156–167, 1990.

[Lev73] L. A. Levin. Universal sorting problems. *Problemy Peredaci Informacii*, 9:115–116, 1973. In Russian. English translation in *Problems of Information Transmission* 9:265–266.

[Lew79] H. R. Lewis. *Unsolvable Classes of Quantificational Formulas*. Addison-Wesley, 1979.

[Liv82] A. B. Livchak. The relational model for systems of automatic testing. *Automatic Documentation and Mathematical Linguistics*, 4:17–19, 1982.

[Lyn80] J. Lynch. Almost sure theories. *Annals of Mathematical Logic*, 18:91–135, 1980.

[Lyn85] J. Lynch. Probabilities of first-order sentences about unary functions. *Trans. American Mathematical Society*, 287:543–568, 1985.

[Mai83] D. Maier. *The Theory of Relational Databases*. Computer Science Press, Rockville, Md., 1983.

[Mak85] J. A. Makowsky. Compactness, embeddings and definability. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 645–716. Springer-Verlag, Berlin/New York, 1985.

[PS89] L. Pacholski and W. Szwast. The 0-1 law fails for the class of existential second-order Gödel sentences with equality. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 280–285, 1989.

[Rad64] R. Rado. Universal graphs and universal functions. *Acta Arith.*, 9:331–340, 1964.

[Sav70] W. J. Savitch. Relationships between nondeterministic and deterministic tape complexities. *Journal of Computer and System Sciences*, 4:177–192, 1970.

[Sch52] H. Scholz. Problem #1: Ein ungelöstes problem in der symbolischen logik. *Journal of Symbolic Logic*, 17:160, 1952.

[She78] S. Shelah. *Classification Theory and the Number of Non-isomorphic Models*. North-Holland, Amsterdam, 1978.

[Sho67] J. R. Shoenfield. *Mathematical Logic*. Addison-Wesley, Reading, Mass., 1967.

[SS88] S. Shelah and J. Spencer. Zero-one laws for sparse random graphs. *Journal of the American Mathematical Society*, 1:97–115, 1988.

[Sze88] R. Szelepcsényi. The method of forced enumeration for nondeterministic automata. *Acta Informatica*, 26:279–284, 1988.

[Tai59] W. W. Tait. A counterexample to a conjecture of Scott and Suppes. *Journal of Symbolic Logic*, 24:15–16, 1959.

[Tar54] A. Tarski. Contributions to the theory of models I,II. *Indagationes Mathematicae*, 16:572–588, 1954.

[TK84] V. A. Talanov and V. V. Knyazev. The asymptotic truth value of infinite formulas. In *Proc. All-Union seminar on discrete mathematics and its applications*, pages 56–61, Moscow, 1984. See Math. Rev. 89g:03054.

[Tra50] B. A. Trakhtenbrot. Impossibility of an algorithm for the decision problem in finite classes. *Doklady Akademii Nauk SSSR*, 70:569–572, 1950.

[Tur84] G. Turán. On the definability of properties of finite graphs. *Discrete Mathematics*, 49:291–302, 1984.

[Ull89] J. D. Ullman. *Database and Knowledge-Base Systems, Volumes I and II.* Computer Science Press, 1989.

[Val76] L. Valiant. Relative complexity of checking and evaluating. *Information Processing Letters*, 5:20–23, 1976.

[Var82a] M. Y. Vardi. The complexity of relational query languages. In *Proc. 14th ACM Symp. on Theory of Computing*, pages 137–146, 1982.

[Var82b] M. Y. Vardi. On decomposition of relational databases. In *Proc. 23rd IEEE Symp. on Foundations of Computer Science*, pages 176–185, 1982.

[Vau60] R. L. Vaught. Sentences true in all constructive models. *Journal of Symbolic Logic*, 25(1):39–53, March 1960.

[Wil80] C. Wilson. Relativization, reducibilities, and the exponential hierarchy. Master's thesis, Toronto, 1980.