Electronic Colloquium on Computational Complexity, Report No. 111 (2009)

Model-Theoretic Characterization of Complexity Classes

Walid Gomaa

Faculty of Engineering
Alexandria University, Alexandria, Egypt
wgomaa@alex.edu.eg

Abstract. Model theory is a branch of mathematical logic that investigates the logical properties of mathematical structures. It has been quite successfully applied to computational complexity resulting in an area of research called descriptive complexity theory. Descriptive complexity is essentially a syntactical characterization of complexity classes using logical formalisms. However, there are still much more of model theory technologies that have not yet been explored by complexity theorists, especially the subarea of classification/stability theory. This paper is divided into two parts. The first part quickly surveys the main results of descriptive complexity theory. In the second part we introduce the field of classification/stability theory, then give the outlines of a research project whose aim is to apply this theory to give a semantical characterization of complexity classes. This would initiate a brand new research area.

keywords: descriptive complexity, complexity theory, classification theory, model theory, logic

1 Introduction

Model theory is the branch of mathematical logic that investigates the relationship between formal languages and their semantics. From a mathematical perspective model theory studies the logical properties of mathematical structures. It was developed in the early years of the twentieth century by pioneering logicians such as Löwenheim (1915), Skolem (1920), Gödel (1930), Tarski (1931), and Malcev (1936). It gained its distinguishing identity as a separate branch of mathematical logic with the work of Henkin, Robinson, and Tarski in the late 1940's and early 1950's [9].

Since the beginning till the early 1950's, model theory exclusively investigated infinite mathematical structures. This was due to two main reasons: (1) model theory was developed as part of the earnest effort towards developing a solid foundation of mathematics and (2) many of the nice abstract logical properties of classes of infinite structures axiomatized by first-order logic fail in the finite case.

Finite structures started getting attention in 1950; this year witnessed the birth of *finite model theory* identified with Trakhtenbrot's result stating that

ISSN 1433-8092

logical validity over finite models is not recursively enumerable. A direct consequence of this result is the failure of completeness over classes of finite structures. Other negative results (as opposed to the infinite case) concerning the model-theoretic behavior of classes of finite models include the failure of compactness, some preservation theorems such Łos-Tarski and Lyndon's theorems, Craig's theorem, and Beth's theorem.

Despite these negative results, finite model theory survived and established itself as an independent field of study. Two things might have helped that: the need for developing a foundation of finite mathematics and the birth of *descriptive complexity theory*.

2 Descriptive Complexity Theory

In the following subsections we give a quick survey of descriptive complexity.

2.1 Algorithmic complexity

In 1952 H. Scholz, G. Kreisel, and L. Henkin [31] defined the *spectrum* of a first-order sentence σ to be the set of natural numbers n such that σ has a finite model of cardinality n. They asked for a characterization of such spectrum. Subsequently, several people had worked on the *computational aspects* of this problem such as G. Asser in 1955 [1], A. Mostowski in 1956 [29], and J. Bennett in 1962 [5]. They had shown that the spectra lies between the third and fourth levels of the Grzegorczyk hierarchy (the third level coincides with the class of functions computable in linear space and the fourth level coincides with the class of elementary computable functions). Then N. Jones and A. Selman in 1972 [24, 25] showed that spectra and context-sensitive languages are closely related; in particular they proved that the spectra are just those sets acceptable by non-deterministic Turing machines in exponential time $2^{O(n)}$.

In the early seventies Ronald Fagin studied *generalized spectra*: a generalized spectrum is the class of finite models of an existential second-order sentence. Consequently in 1974 he proved his celebrated theorem that generalized spectra coincides with the complexity class NP; in other words NP can be exactly captured by existential second-order logic [13]. This is considered by many to mark the beginning of the new research area of *descriptive complexity* which investigates how hard it is to express the computational problem in some logical formalism. Hence, the resources considered are basically logical such as the number of object variables, quantifier depth, type, and alternation, sentence length (finite vs. infinite), inductive vs. non-inductive capabilities, etc.

The work done by Fagin in 1974 and most of the work done afterwards have been leaning towards computation in the computation/logic duality. This can be described as *the computerization of logic*. The converse direction, *the logicization of computation*, has recently started to get attention with the work of A. Blass, Y. Gurevich, and S. Shelah on the *Choiceless Polynomial Time* model [6]. As a direct consequence of Fagin's theorem it follows that $coNP$ is exactly captured by universal second-order logic. Fagin's result has also been generalized by Stockmeyer [42] to show that the whole of the polynomial hierarchy is exactly captured by second-order logic.

What about the complexity-theoretic behavior of first-order logic FOL? Although this logic is very powerful over infinite structures, its expressibility over finite structures is surprisingly too weak. It has to be augmented with additional logical resources in order to be useful enough to capture important complexity classes. Such resources include among others generalized quantifiers, transitive closure operators, inductive capabilities, infinitude of formulas.

The simplest operator to add to FOL is the *deterministic transitive closure* DTC. Let $G = (V, E)$ be a graph, then

$$\begin{aligned} DTC(G) = \{&(a,b) \in E\colon \textit{there exists } n > 0 \textit{ and } a_0, \ldots, a_n \in V \\ &\textit{such that } a_0 = a \textit{ and } a_n = b \textit{ and for each } i < n, \\ &a_{i+1} \textit{ is the unique } a \textit{ for which } (a_i, a) \in E\} \end{aligned}$$

In other words, $DTC(G)$ is the set of ordered pairs (a, b) such that there is a *deterministic* path in G from vertex a to vertex b.

An ordered structure is a structure with a built-in binary relation symbol (not necessarily in the vocabulary) that is always interpreted as a linear ordering over the universe. Let $FO(DTC)$ denote first-order logic augmented with the DTC operator and let L denote the logspace complexity class. Then L can be logically characterized as follows.

Theorem 1 ([21]). *L over ordered finite structures is captured by $FO(DTC)$.*

The next step is to augment FOL with the more powerful (*transitive closure*) TC operator. TC behaves similarly to DTC except that the path between any two vertices in the closure need not be deterministic (vertices along the path may have multiple neighbors). As expected this adds a nondeterministic capability to the logic which will be seen below when related to a complexity class. First, we give an example that illustrates the working of the TC operator.

Example 1. *Consider the language of groups. A simple group is a nontrivial group whose only normal subgroups are the trivial group and the group itself. The following $FO(TC)$ sentence defines the class of simple groups.*

$$\exists x(x \neq 1) \wedge \forall x \left(x \neq 1 \to \forall y[TC_{u,v} \exists w(u(w^{-1}xw)) = v](1, y)\right)$$

The first conjunct expresses the fact that the group is non-trivial and the second one implies that the normal closure of any element, excluding the identity, equals the whole group.

Let NL denote nondeterministic logspace. As expected, the following holds.

Theorem 2 ([21]). *NL over ordered finite structures is captured by $FO(TC)$.*

Using $FO(TC)$, N. Immerman [22] and R. Szelepcsényi [43] have discovered the surprising fact that nondeterministic space classes are closed under complementation.

Another important operator added to FOL in order to improve its expressibility power is the *least fixpoint operator*. The resulting logic has inductive capabilities; it is called *fixpoint logic* and denoted by LFP. Consider a first-order language $\mathcal{L}$ and let S be a relation symbol of arity n such that $S \notin \mathcal{L}$. Consider an $(\mathcal{L} \cup \{S\})$-formula $\varphi(\bar{x}, S)$ where $|\bar{x}| = n$ and φ is positive in S. Let $\mathcal{A}$ be a finite $\mathcal{L}$-structure. Define the following inductive sequence of subsets of A^n.

$$S^0 = \emptyset$$
$$S^{i+1}\{\bar{a} \in A^n \colon \mathcal{A} \models \varphi(\bar{a}, S^i)\}$$

By the positivity of S, the sequence S^i is monotonically increasing, that is $S^0 \subseteq S^1 \subseteq \cdots$. By the finiteness of $\mathcal{A}$ this sequence will stop after a finite number of steps, that is there exists some j such that $S^j = S^k$ for any $k \geq j$. Let j be minimal, then S^j is called the least fixpoint of $\varphi(\bar{x}, S)$, and is usually denoted by S^∞. N. Immerman [20] and M. Vardi [44] have independently characterized the polynomial time class P as follows.

Theorem 3 ([20, 44]). *P over ordered finite structures is captured by LFP.*

This theorem fails if the ordering is removed, that is LFP is too weak to capture the whole of P. It is still an open problem whether there exists a logic that characterizes P, the problem can be stated more formally as follows [8].

Question 1 *Is there a recursively enumerable listing of a set of Turing machines that accept exactly all the polynomial-time graph properties?*

Relaxing LFP by removing the monotonicity condition would result in a new logic called *partial fixpoint logic* PFP. Note that in such logic the fixpoint operator is not guaranteed to converge. Divergence can be detected on finite structures after a finite number of iterations, in this case the fixpoint is taken to be the empty set. Let *PSpace* denote the class of problems decidable in polynomial space, then

Theorem 4 ([18, 19, 44]). *PSpace over ordered finite structures is captured by PFP.*

2.2 Circuit complexity

Circuit complexity is the branch of computational complexity that uses circuits of boolean logic gates as its model of computation. Circuits are directed acyclic graphs in which the input bits are placed at the leaves and signals proceed up the circuit toward the root, thus a gate is never reused during a computation [23].

Assume a polynomially bounded function $t(n)$. Let $AC[t(n)]$ denote the uniform family of boolean circuits $\{C_n\}_{n\in\mathbb{N}}$ that have polynomial size, whose depths are bounded by $O(t(n))$, and consist of AND and OR gates that have unbounded fanin. Let $FO[t(n)]$ denote the set of properties definable by a uniform sequence of sentences $\{\sigma_n\}$ such that each sentence σ_n has length $O(t(n))$ and has a bounded number of variables independent of n. Let $FO[t(n)](<,+,\times)$ denote $FO[t(n)]$ extended with linear ordering and arithmetic. Then AC can be logically characterized as follows.

Theorem 5 ([23]). *$AC[t(n)]$ is captured by $FO[t(n)](<,+,\times)$.*

The logical resource mainly used in this theorem to characterize AC is the length of the defining first-order sentence. Instead of using one finite sentence to define a given problem, we now use an infinite increasing chain of sentences, each for a given input length, to describe it. This actually parallels the nature of the circuit model of computation where a given problem is solved by an infinite increasing chain of circuits.

In the remaining part of this subsection we shall return to the use of single finite sentences. First we extend FOL with the *majority quantifier*. Let $\varphi(x,\bar{y})$ be a first-order formula with free variables $x,\bar{y}$. Then $(Maj\ x)\varphi(x,\bar{y})$ is a formula in $FO(Maj)$ interpreted as follows: $\mathcal{A} \models (Maj\ x)\varphi(x,\bar{b})$ iff $|\{a \in A\colon \mathcal{A} \models \varphi(a,\bar{b})\}| \geq \lceil\frac{|A|}{2}\rceil$.

The circuit class TC^0 is the class of problems decidable by a family of boolean circuits that have polynomial size, constant depth, and unbounded fanin; the gates allowed in such circuits are: AND, OR, NOT, and *threshold gates* (a threshold gate returns 1 if at least half of its inputs are 1, and 0 otherwise). The following theorem logically characterizes TC^0.

Theorem 6 ([4]). *TC^0 is captured by $FO(Maj;<,+,\times)$.*

Another descriptive characterization of TC^0 has been obtained through the use of counting quantifiers. Let FOC denote FOL augmented with counting quantifiers which are defined as follows. Let $\varphi(x,\bar{y})$ be a first-order formula with free variables $x,\bar{y}$. Then $\exists^{\geq z}x\varphi(x,\bar{y})$ is a formula in FOC where x is bounded and $z,\bar{y}$ are free. Given a structure $\mathcal{A}$ and parameters $\bar{b} \in A$, then $\mathcal{A} \models \exists^{\geq z}x\varphi(x,\bar{b})$ iff $|\{a \in A\colon \mathcal{A} \models \varphi(a,\bar{b})\}| \geq z$. Then TC^0 can be characterized as follows.

Theorem 7 ([4]). *TC^0 is captured by $FOC(<,+,\times)$.*

An important open problem in complexity theory is whether $TC^0 \subsetneq L$. One step towards the solution of this problem has been done by M. Ruhl in [30]. As mentioned above TC^0 can be characterized by $FOC(<,+,\times)$. Ruhl proved that the logic resulting from the absence of multiplication is too weak to capture L, in other words $FOC(<,+) \subsetneq L$. The proof is based on giving an explicit winning strategy for the duplicator in an Ehrenfeucht-Fraïssé game over two graphs, one of which has a deterministic path between two designated vertices whereas the other graph does not have such path. However, the proof breaks down when adding the multiplication operation to the language because of the combinatorial nature of the resulting structure which is much more involved than the case with only the addition operation.

3 Semantical Characterization of Complexity Classes

So far the application of model theory to computational complexity has been within the realm of descriptive complexity theory. As seen above descriptive complexity is essentially a *syntactical characterization* of complexity classes. However, there is a rich deep arena of research in model theory which, to the best of our knowledge, has not yet been explicitly and satisfactorily applied to computational complexity. This is the research area of *classification/stability* theory.

3.1 Historical background

Classification theory is a sophisticated very well-developed subarea of pure model theory. The work in classification theory was initiated in 1965 by M. Morley's celebrated categoricity theorem [28]: A first-order theory in a countable language is categorical in an uncountable cardinal κ (has only one model of cardinality κ up-to isomorphism) if and only if it is categorical in all uncountable cardinals. Saharon Shelah's taxonomy of first-order theories by the *stability classification* laid the foundation for most model-theoretic research in the last 40 years [2]. He first introduced the dichotomy of stable vs. unstable theories in 1969 [32] as an aid to count the number of non-isomorphic models of a theory in a given cardinality. In the following few years he focused on the classification problem [33, 36, 35, 34, 37] which , roughly speaking, aims at categorizing first-order complete theories into two kinds: those whose models have a good structure theory and those whose models are ill-behaved [16]. The main results for the countable first-order case had been reached by the mid 1980's [38, 40]. Every countable theory T falls into one of two classes: either (1) T is intractable, that is, for all sufficiently large κ, T has 2^κ models (which is the maximum possible) or (2) the number of models of T of cardinality κ is bounded well below 2^κ; furthermore every model of T can be decomposed as a tree of countable models [2]. For several accounts of the subject the interested reader should consult S. Shelah [39], W. Hodges [15], D. Lascar [26], and J. Baldwin [3].

This line of research has resulted in a huge rich body of model-theoretic concepts, technologies, tools, and theorems. In the following subsections we will introduce the classification/stability theory, its restriction to finite models, and its potential application to computational complexity which would lead to a novel research area that gives a *semantical characterization of complexity classes* (as opposed to descriptive complexity).

3.2 Classification/Stability theory

The mainstream research in model theory since the seventies have been focusing on the development of *classification/stability* theory for classes of infinite structures. Let $\mathscr{K} = (\mathbb{K}, \preceq_{\mathscr{K}})$ be a class $\mathbb{K}$ of structures with the same vocabulary along with a partial ordering $\preceq_{\mathscr{K}}$ defined over $\mathbb{K}$. For example, $\mathscr{K}$ could be the class of models of a first-order theory with the regular elementary substructure relation. Let λ be a cardinal and let $\mathscr{K}_\lambda = \{\mathcal{M} \in \mathscr{K} : |M| = \lambda\}$. Then classification theory aims at answering questions about $\mathscr{K}_\lambda$ of the following nature [14].

1. Is $\mathscr{K}_\lambda \neq \emptyset$?
2. Does $\mathscr{K}_\lambda \neq \emptyset$ imply that $\mathscr{K}_{\lambda^+} \neq \emptyset$?
3. If $\mathscr{K}$ is λ^+-categorical ($|\mathscr{K}_{\lambda^+}| = 1$ up-to isomorphism), does that imply it is λ-categorical? (downward transfer of categoricity)
4. If $\mathscr{K}$ is λ-categorical, does that imply it is λ^+-categorical? (upward transfer of categoricity)
5. What are the possible functions $\lambda \longmapsto |\mathscr{K}_\lambda|$?
6. Under what conditions on $\mathscr{K}$ it is possible to find a nice *independence* relation on subsets of every $\mathcal{M} \in \mathscr{K}$? (this is a generalization of linear independence in vector spaces or algebraic independence in fields)

First-order *stability theory* is the main technology used to develop a classification theory for first-order logic: the study of the internal structure of models was developed to provide classifications of those models. First-order stable classes behave very nicely and have a well-defined *dimension theory* [7] based on an independence relation called *forking* [7, 27].

Let $\mathcal{L}(\tau)$ denote first-order logic with vocabulary τ.

Definition 1 (Types). Let $\mathcal{M}$ be a τ-structure and let $a \in M$.

1. Define *the type (sometimes called the pure type) of a inside $\mathcal{M}$ as follows.*

$$tp(a, \mathcal{M}) = \{\varphi(x) \in \mathcal{L}(\tau) \colon \mathcal{M} \models \varphi(a)\}$$

2. Let $B \subseteq M$. Define *the type of a over B inside $\mathcal{M}$ as follows.*

$$tp(a/B, \mathcal{M}) = \{\varphi(x, \bar{b}) \in \mathcal{L}(\tau \cup \{\bar{b}\}) \colon \bar{b} \in B \text{ and } \mathcal{M} \models \varphi(a, \bar{b})\}$$

Definition 2 (Stability).

1. Let T be a first-order theory and let λ be an infinite cardinal. We say that T is *stable in λ if for every $\mathcal{M} \models T$ and for every $A \subseteq M$ such that $|A| \leq \lambda$, the total number of types over A that are realized in $\mathcal{M}$ is at most λ. In other words the number of types over sets of cardinality at most λ grows polynomially with λ. (as opposed to the exponential growth in the case of unstability)*
2. Let $\mathscr{K}$ be a class of structures. We say that $\mathscr{K}$ is stable in λ if for every $\mathcal{M} \in \mathscr{K}$ and for every $A \subseteq M$ with $|A| \leq \lambda$, the total number of 'types' over A realized in $\mathcal{M}$ is at most λ.

The second part of this definition is more general since $\mathscr{K}$ need not be first-order axiomatizable and consequently the notion of $\mathscr{K}$-'type' is more general (need not be first-order). For example, it can be the class of models of a theory defined in any logic such as $\mathcal{L}_{\omega_1\omega}$, or even need not be associated with any logic such as the case of *abstract elementary classes* [2, 14, 41] (where a class is defined by a set of purely semantical axioms).

Roughly speaking, a set of elements have the same type means that they are very much structurally alike as individuals. So stability implies polynomial, rather than exponential, growth of the total number of really distinct elements in terms of the number of parameters used to induce their types. Stability in general implies a well-behaved class. For example in the first-order case stability implies the existence of infinite arbitrarily large sets of indiscernible elements; *indiscernibility* is much stronger measure of similarity among a set of elements than a type is.

3.3 Stability theory for finite models

One of the fundamental difficulties to developing model theory, and in particular classification theory, for finite structures is the choice of an appropriate 'submodel' relation. In classification theory for elementary classes (the first-order case), the notion of *elementary submodel* ($\mathcal{M} \preceq \mathcal{N}$) has been used quite successfully; it is a strengthening of the notion of submodel ($\mathcal{M} \subseteq \mathcal{N}$). Unfortunately, for finite structures $\mathcal{M} \preceq \mathcal{N}$ always implies that $\mathcal{M} = \mathcal{N}$. Moreover, in many cases even $\mathcal{M} \subseteq \mathcal{N}$ implies $\mathcal{M} = \mathcal{N}$ (e.g., when $\mathcal{N}$ is a group of prime order) [12]. So a substitute is needed.

To the best of our knowledge there have been only three attempts to developing a classification/stability theory for finite models. The first of these is given in an unpublished article by D. Ensley and R. Grossberg [12]. In this article the authors started a project whose ultimate goal is to have a decomposition theorem for finite structures like the theorem for finite abelian groups; a finite structure be decomposed into a class of structures that have certain nice properties (in some sense this class converges to the original structure). The approach taken is basically *syntactical*. They directly adapt the stability-theoretic notions developed for elementary classes to the finite case through cardinality restrictions. For example, $(\mathcal{M}, \varphi(x, y))$ has the *n-order property* if there exists $\{a_i : i < n\} \subseteq M$ such that $\mathcal{M} \models \varphi(a_i, a_j)$ if and only if $i < j$. Given some apriori properties they define a class $\mathscr{K} = (\mathbb{K}, \preceq_{\mathscr{K}})$ that satisfies three basic axioms: (i) a finitary restricted version of elementary substructure (and consequently a restricted form of the Tarski-Vaught test), (ii) a finitary restricted version of saturation (realizing its own induced types), and (iii) a property similar to the one in first-order model theory which guarantees that types over models are stationary (have unique nonforking extensions). Some properties of $\preceq_{\mathscr{K}}$ are proved such as transitivity and coherence. A sort of independence relation called 'stable amalgamation' is defined and its symmetry is proved.

The second attempt for developing a classification theory for classes of finite structures is given by T. Hyttinen in [17]. The approach taken here is more semantical. In his redefinition of stability-theoretic notions, finiteness is treated as potential (non-uniform) infinity as opposed to both the actual infinity assumed in first-order model theory and the actual finiteness assumed in the first approach mentioned above. For example, in his version of the order property, a formula $\varphi(x, y)$ has this property if for every $n < \omega$, there are a model $\mathcal{M} \in \mathscr{K}$ and $\{a_i : i < n\} \subseteq M$ such that $\mathcal{M} \models \varphi(a_i, a_j)$ if and only if $i < j$ (also a *coherent* sequence is defined that unifies all these partial finite sequences into one sequence of length ω). $\mathscr{K}$ is defined to satisfy three basic axioms: (i) a sort of bounded elementary equivalence among all models in $\mathscr{K}$ (for example with respect to the bounded-variable logic $\mathcal{L}^n$), (ii) a sort of amalgamation (to enable an effective definition of types), and (iii) finiteness of the number of types. Some stability-theoretic notions and results were obtained in this framework, for example: (i) the equivalence between stability and the absence of the order property in $\mathscr{K}$ and (ii) the equivalence between stability and the well-definededness of Shelah's *ω-rank* [17] for all types in $\mathscr{K}$.

The third attempt is due to M. Djordjević in [10]. His approach is syntactical, however, finite models is not the main focus of the paper. The main goal is to study complete $\mathcal{L}^n$-theories (theories in bounded-variable logic) and to transfer concepts and methods from first-order stability theory to this context. Finite models are investigated in Section 4 after the machinery of the $\mathcal{L}^n$-stability theory for infinite models have been developed. Two main results concerning finite models are obtained: (i) a finitary version of the downward Löwenheim-Skolem theorem assuming the $\mathcal{L}^n$-theory is ω-categorical, ω-stable, and does not have function symbols and (ii) the existence of arbitrarily large finite models and finite amalgamation.

None of the aforementioned articles has tried to make any connection with the computational complexity-theoretic properties of finite models.

3.4 Stability-theoretic characterization of complexity classes

Understanding the internal structure of finite models is necessary to understand the computational complexity-theoretic nature of these models and to provide complexity-theoretic classifications for them. Therefore, the connection between the two fields, stability and complexity, is intuitive and natural. In the following we will try to give the outlines of a research project aiming at connecting the two fields.

Any such project may be divided into two interdependent parts: (1) developing a stability theory for classes of finite structures and (2) developing the relationship between the stability-theoretic properties of these classes and their complexity-theoretic ones, these latter properties can be either computational or descriptive.

Two main goals should be considered when defining $\mathscr{K} = (\mathbb{K}, \preceq_{\mathscr{K}})$, and in particular defining $\preceq_{\mathscr{K}}$,: (1) the definition be as purely semantical as possible (independent of any particular logical formalism) taking the axiom system developed for abstract elementary classes [2, 14, 41] as a starting point and (2) the definition be as closely related as possible to complexity-theoretic notions. Through this philosophy there will be no fixed syntactical characterization of $\preceq_{\mathscr{K}}$, it is just required that $\preceq_{\mathscr{K}}$ satisfies a set of semantical conditions such as closure under isomorphism, coherence, some version of the downward Löwenheim-Skolem axiom, etc. The restricted form of elementary substructure or the bounded variable substructure as both defined in the aforementioned articles may satisfy these conditions making the intended axiom system comprehensive and inclusive of many frameworks especially from the complexity-theoretic perspective. If the semantical axiom system comes up too general to be useful, then syntactical ingredients should be considered.

Developing the right axiom system for $\mathscr{K}$ is the most crucial task as it lays the foundation for all later work. So this system must be tested thoroughly against many frameworks of complexity-theoretic interest such as different classes of graphs (planar, excluding a minor, with hamiltonian cycles, etc), groups, linear orderings, etc.

Given the axiom system for $\mathscr{K}$, the second major task is to define the basic model-theoretic concepts such as *types* and *saturation*. A set of elements have the same type means that they are very much structurally alike as individuals, hence the concept of 'types' has natural complexity-theoretic consequences; for example, having the same type can be a crucial hint for any algorithmic procedure operating over the structure containing these elements. Based on the axiom system and those basic definitions, primary model-theoretic theorems can be derived such as uniqueness of saturated models (models that contain realizations of their own induced types).

The third major task is to develop the stability-theoretic notions, tools, technologies, and theorems for $\mathscr{K}$ and correlate them with the complexity-theoretic properties of $\mathscr{K}$. The work done in the three aforementioned articles would provide ideas and hints for developing the stability-theoretic notions. As mentioned above, in model theory stability vs. non-stability corresponds to polynomial vs. exponential growth of the number of types. Stability on classes of finite structures should be defined in a similar, though finer, way.

An important concept in stability theory is the notion of an *indiscernible sequence/set*, it is a much stronger measure of similarity among a set of elements than a type is; a group of elements form an indiscernible set if they are very much structurally alike, not just as individuals, but as tuples formed arbitrarily. The existence of a large set of indiscernible elements implies that the model containing them is very simple, hence nicely behaved from the complexity-theoretic perspective. Two possible approaches can be taken when correlating the complexity-theoretic and the stability-theoretic properties of $\mathscr{K}$: (1) a purely semantical approach: roughly speaking a stable class has a small number of types and many indiscernible elements, hence less structural variations inside the models of the class, therefore the computational properties of the class are less complex and (2) semantical-syntactical approach: the idea is to find a syntactical characterization of stable classes within the context of some logical formalism (similar to preservation theorems), then use the relatively rich body of results in descriptive complexity theory to find the corresponding computational complexity-theoretic properties of $\mathscr{K}$.

The fourth major task is to develop the deeper more advanced concepts of stability and correlate them with complexity theory. Among these notions are the order property, the independence property, and the cover property. These are combinatorial concepts the existence of any of which in $\mathscr{K}$ typically implies non-stability. Another concept that can be borrowed from first-order stability theory is *Morley rank* [7, 11]. It is an ordinal assigned to definable subsets of structures in $\mathscr{K}$ which indicates the complexity of their logical definability. Ranking from the complexity-theoretic perspective should indicate how complex (from a logical and/or computational perspective) a set is.

Another major advancement in first-order stability theory is the development of an *independence* relation among the subsets of the structures in $\mathscr{K}$. This relation is a generalization of linear independence in vector spaces or algebraic independence in fields. Several such relations have been developed such as splitting, dividing, and forking. Correlating independence among sets with their complexity-theoretic behavior is natural and intuitive. For example, a vertex u is independent from v over w could mean that there is no path from u to v through w.

The developments done in [17] for these sophisticated notions could provide starting points, ideas, and hints.

References

1. G. Asser. Das repräsentenproblem im Prädikatenkalkül der ersten Stüfe Identitat. *Zeitschrift Für Math. Logik and Grundlagen Der Math*, 1:252–263, 1955.
2. J. Baldwin. *Categoricity in Abstract Elementary Classes*. Monograph, available at http://www.math.uic.edu/ jbaldwin/model.html.
3. J. Baldwin. *Fundamentals of Stability Theory*. Springer, 1988.
4. D. Barrington, N. Immerman, and H. Straubing. On Uniformity within NC^1. *Journal of Computer and System Sciences*, 41(3):274–306, 1990.
5. J. Bennett. *On Spectra*. PhD thesis, Princeton University, 1962.
6. A. Blass, Y. Gurevich, and S. Shelah. Choiceless Polynomial Time. *Annals of Pure and Applied Logic*, 100:141–187, 1999.
7. S. Buechler. *Essential Stability Theory*. Springer, 2002.
8. J. Cai, M. Fürer, and N. Immerman. An Optimal Lower Bound on the Number of Variables for Graph Identification. *Combinatorica*, 12(4):389–410, 1992.
9. C. Chang and H. Keisler. *Model Theory*. North Holland, 3 edition, 1990.
10. M. Djordjević. Finite Variable Logic, Stability, and Finite Models. *Journal of Symbolic Logic*, 66(2), 2001.
11. V. D. Dries. Introduction to Model-Theoretic Stability. Lecture notes available at http://www.math.uiuc.edu/ vddries/.
12. D. Ensley and R. Grossberg. Finite Models, Stability, and Ramsey's Theorem. http://arxiv.org/abs/math.LO/9608205.
13. R. Fagin. Generalized First-Order Spectra and Polynomial-Time Recognizable Sets. In R. Karp, editor, *Complexity of Computation*, volume 7, pages 43–73. SIAM-AMS, 1974.

14. R. Grossberg. Classification Theory for Abstract Elementary Classes. *Contemporary Mathematics*, 302, 2002.
15. W. Hodges. What is a Structure Theory? *Bulletin of London Mathematical Society*, 19:209–237, 1987.
16. W. Hodges. *Model Theory*. Cambridge University Press, 1993.
17. T. Hyttinen. On Stability in Finite Models. *Archive for Mathematical Logic*, 39:89–102, 2000.
18. N. Immerman. Number of Quantifiers is Better than Number of Tape Cells. *Journal of Computer and System Sciences*, 22(3):384–406, 1981.
19. N. Immerman. Upper and Lower Bounds for First-Order Expressibility. *Journal of Computer and System Sciences*, 25:76–98, 1982.
20. N. Immerman. Relational Queries Computable in Polynomial Time. *Information and Control*, 68:86–104, 1986.
21. N. Immerman. Languages That Capture Complexity Classes. *SIAM Journal of Computing*, 16(4):760–778, 1987.
22. N. Immerman. Nondeterministic Space is Closed under Complement. *SIAM Journal on Computing*, 17(5):935–938, 1988.
23. N. Immerman. *Descriptive Complexity*. Springer, 1998.
24. N. Jones and A. Selman. Turing Machines and the Spectra of First-Order Formulas with Equality. In *Proceedings of the 4th Annual ACM Symposium on Theory of Computing*, pages 157–167, 1972.
25. N. Jones and A. Selman. Turing Machines and the Spectra of First-Order Formulas. *Journal of Symbolic Logic*, 39(1):139–150, 1974.
26. D. Lascar. Stability in Model Theory. *Pitman Monographs and Surveys in Pure and Applied Mathematics*, 36, 1987.
27. D. Lascar and B. Poizat. An Introduction to Forking. *Journal of Symbolic Logic*, 44(3):330–350, 1979.
28. M. Morley. Categoricity in Power. *Transactions of the American Mathematical Society*, 114(2):514–538, 1965.
29. A. Mostowski. Concerning a Problem of H. Scholz. *Zeitschrift Für Math. Logik and Grundlagen Der Math*, 2:210–214, 1956.
30. M. Ruhl. Counting and Addition Cannot Express Deterministic Transitive Closure. In *LICS'99: 14th Annual IEEE Symposium on Logic in Computer Science, IEEE Computer Society*, pages 326–334, 1999.
31. H. Scholz, G. Kreisel, and L. Henkin. Problems. *Journal of Symbolic Logic*, 17(2):160, 1952.
32. S. Shelah. Stable Theories. *Israel Journal of Mathematics*, 7:187–202, 1969.
33. S. Shelah. Finite Diagrams Stable in Power. *Annals of Mathematical Logic*, 2:69–118, 1970.
34. S. Shelah. Stability, the f.c.p., and Superstability; Model Theoretic Properties of Formulas in First Order Theory. *Annals of Mathematical Logic*, 3:271–362, 1971.
35. S. Shelah. The Number of Non-Isomorphic Models of an Unstable First-Order Theory. *Israel Journal of Mathematics*, 9:473–487, 1971.
36. S. Shelah. Categoricity of Uncountable Theories. In *Proceedings of Tarski Symposium*, pages 187–203, 1974.
37. S. Shelah. Why There are Many Non-isomorphic Models for Unsuperstable Theories. In *Proceedings of International Congress of Mathematicians*, pages 259–263, 1974.
38. S. Shelah. *Classification Theory and the Number of Non-Isomorphic Models*. North-Holland, 1978.

39. S. Shelah. Classification of First Order Theories which have a Structure Theorem. *Bulletin of the American Mathematical Society*, 12(2):227–232, 1985.
40. S. Shelah. Around Classification Theory of Models. *Lecture Notes in Mathematics*, 1182, 1986.
41. S. Shelah. *Classification Theory for Abstract Elementary Classes*. College Publications, 2009.
42. L. Stockmeyer. The Polynomial-Time Hierarchy. *Theoretical Computer Science*, 3:1–22, 1977.
43. R. Szelepcsényi. The Method of Forced Enumeration for Nondeterministic Automata. *Acta Informatica*, 26:279–284, 1988.
44. M. Vardi. The Complexity of Relational Query Languages. In *Proc. of the 14th ACM Symp. on Theory of Computing*, pages 137–146, 1982.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de