Finite and Infinite Model Theory
A Historical Perspective

Abstract

We describe the progress of model theory in the last half century from the standpoint of how finite model theory might develop.

Keywords: stability theory, finite model theory

1 Introduction

History is more or less bunk.
Henry Ford
Progress, far from consisting in change, depends on retentiveness. Those who cannot remember the past are condemned to repeat it.
George Santayana [1]

In the first section of this paper we sketch a history of some developments in 'infinite' model theory in the last fifty years. This sketch does not purport to be a real history of the period but rather a selective account that lays out a certain analogy with 'finite' model theory and which focuses on those results in infinite model theory which we think will be important for the study of finite models. In particular, many important results are omitted simply because they don't fit well with the theme of this sketch. Most important, the dichotomy – finite versus infinite model theory – around which we organize the entire essay might not seem relevant to many model theorists.

In the second section we give a briefer sketch of the development of finite model theory. In the third we sketch how these stability theoretic and finite model theoretic ideas have interacted in four areas: stability theory for L^n, embedded finite model theory, L^n axiomatizability of first order theories, and 0-1 laws.

Earlier versions of several sections of this paper appeared as the written version of my talk at the Workshop on Logic and Cognitive Science held April 16 - 18, 1999 and distributed as a report of the University of Pennsylvania, Institute for Research in Cognitive Science. My thanks to Eric Rosen, Rami Grossberg, and especially the referees for a number of perceptive comments.

2 Infinite Model Theory

2.1 Infinite Model Theory: Before 1955

We briefly fix the vocabulary of this discussion; when in doubt consult [45] or [19].

- A signature L is a collection of relation and function symbols (yielding atomic formulas). By default, L will be countable and recursive.

[1] Excerpted from Compton's Reference Collection 1996 Copyright (c) 1995 Compton's NewMedia, Inc.

- A structure for that signature (L-structure) is a set with an interpretation for each of those symbols.
- The first order language associated with L is the least set of formulas containing the atomic L-formulas and closed under the Boolean operations and quantification over individuals.
- Each class $\mathbf{K}$ of models is taken to be closed under isomorphism.
- Formulas in which each variable is bound by a quantifier are called L-sentences.
- A theory T is a collection of L-sentences which is closed under logical consequence. If all models of T satisfy exactly the same L-sentences (are elementarily equivalent), T is said to be *complete*. We write $|T|$ for the number of symbols in T plus $\aleph_0$.

Until 1960, much of the work in model theory studied the properties of first order logic not the properties of first order theories. Sample results include the fundamental properties of first order logic:

THEOREM 2.1 (Löwenheim-Skolem)
If a first order theory has an infinite model, it has a model in each infinite cardinality.

THEOREM 2.2 (Compactness)
If every finite subset of a collection Σ of sentences has a model then Σ has a model.

THEOREM 2.3 (Completeness)
The collection of valid first order sentences is recursively enumerable.

There were a number of applications of these basic results:

THEOREM 2.4 (Łoś-Tarski)
A class $\mathbf{K}$ of models for a relational language can be axiomatized by universal sentences if and only if a) $\mathbf{K}$ is closed under substructure and b) if every finite substructure of a model A is in $\mathbf{K}$ then $A \in \mathbf{K}$.

THEOREM 2.5 (Lyndon)
The class of models of a first order theory is closed under homomorphism iff it can be axiomatized by positive sentences.

There were also seminal results about specific first order theories such as Tarski's proof that the real field admits quantifier elimination and Szmielev's classification of theories of Abelian groups.

2.2 Infinite Model Theory: From 1955 to 1970

Around 1960, infinite model theory took a decisive turn by concentrating on the study of complete theories. This was a natural tack in that case since first order logic is well-equipped to study theories of intrinsic mathematical interest; e.g. real closed fields, algebraically closed fields, Peano arithmetic. I will discuss in later sections the possibilities of studying theories in finite model theory. This is not such an obviously well-motivated topic; however, it is natural to consider not only all finite structures for a language but various subclasses. In this survey we focus on the importance of the switch from the study of logics to the study of families of complete first order theories and the role of the notion of type in this study. In Section 4 we discuss the developing analogous tools for finite models.

There is a useful, but confusing to the uniniated, convention concerning properties of complete theories. Each such theory is determined by one (any) of its models. Thus, an adjective, like stable, can be (and is) applied interchangably to a structure M or to the theory T of M- the set of all sentences true in M.

The key notion of elementary submodel was introduced by Tarski and Vaught in 1957 [71]. Abraham Robinson [65] had already formulated the notion of model completeness. However, the following characterization (in these words) only appeared later: A theory is model complete if it satisfies the following equivalent conditions.

THEOREM 2.6 (Robinson)
For every first order theory, T, the following are equivalent:

1. Every submodel of a model of T is an elementary submodel.
2. Every formula is equivalent in T to an existential formula.

The amalgamation property (for various categories of models) has played a central role in many areas of model theory. The symbol $A \longrightarrow B$ means there is an embedding, in the appropriate category, from A to B. Thus, in ordinary first order order model theory, $A \longrightarrow B$ denotes elementary embedding; in universal algebra, $A \longrightarrow B$ denotes embeddings; in the finite variable context, $A \longrightarrow B$ denotes L^n-elementary embedding.

DEFINITION 2.7
The class $(\mathbf{K}, \longrightarrow)$ satisfies the *amalgamation property* $(\mathbf{AP})$ if for any situation:

$$\begin{tikzcd} & N \\ A \arrow[ur] \arrow[dr] & \\ & M \end{tikzcd}$$

such that A is $\mathbf{K}$-embedded in M and N there exists an N_1

such that

$$\begin{tikzcd} & N \arrow[dr] & \\ A \arrow[ur] \arrow[dr] & & N_1 \\ & M \arrow[ur] & \end{tikzcd}$$

and the result is a commuting diagram of $\mathbf{K}$-embeddings.

The notion of a homogeneous-universal structure for a class $\mathbf{K}$ was isolated by Fraïssé [32] and Jónsson [53, 54] in the 50's. They pointed out that certain properties of the class (most importantly closure under substructure and amalgamation) allowed the construction of structures which are homogeneous (local isomorphisms extend to automorphisms) and universal (embed all structures in $\mathbf{K}$). This idea with the properties of the class expressed in suitably general form will reoccur often in our story. We expand in Subsection 4 on the role of the amalgamation property in studying theories in logic with finitely many variables. Morley and Vaught [60] discovered the proper category for considering this notion for first order logic (models of a complete first order theory with elementary embeddings as morphisms) and introduced the notion of a saturated model. They proved that in this category the 'universal-algebraic' notion of homogeneous-universal was equivalent to the 'model theoretic' notion of saturation. To define saturation we need the notion of a type.

DEFINITION 2.8
A collection of formulas p is a complete type over $A \subseteq M \models T$ if it satisfies one of the following equivalent conditions.

1. p is a maximal consistent set of formulas $\phi(x, \mathbf{a})$ with parameters $\mathbf{a}$ from A.
2. p is a member of the Stone Space of the Lindenbaum algebra of A.

For every set A contained in a model M and every natural number n, $S^n(A)$ denotes the collection of types (in n-variables) over A (the *nth Stone space* of A). $S(A)$ denotes the set of such p (where n is taken either as 1 or ambiguously as any arbitrary n; technical results show that in most cases this distinction is not important).

A model M is *κ-saturated* if for every A contained in M with $|A| < \kappa$, if $p \in S(A)$ then p is *realized* in M, i.e. some element of M satisfies p, i.e. satisfies each formula in p.

The 'monster model' is a κ-saturated model for some κ greater that the size of any set that we discuss; the use of monster models is explained in any standard text on stability theory. One important use is that the solutions of a type $p \in S(A)$ can be seen as an orbit of the group of automorphisms of the monster model which fix A.

The Lowenheim-Skolem theorem shows that a first order theory cannot hope to characterize models uniquely since any theory with infinite models must have one in each infinite cardinality. The closest approximation is to have a single model in each cardinality.

DEFINITION 2.9
A theory T is κ-categorical if all models of T with cardinality κ are isomorphic.

EXAMPLE 2.10
1. The theory of an infinite set is categorical in every cardinal.
2. The theory of the integers under successor is categorical in all uncountable cardinals.
3. Let T be the theory of a vector space V over a field F. Then T is categorical in every uncountable cardinal. If F is finite then T is $\aleph_0$-categorical.
4. The theory of an algebraically closed field is categorical in every uncountable cardinal.

The study of categoricity is one aspect of studying the number of models a theory has in each cardinality. Ehrenfeucht showed in the early 60's that the theory of the following structure has 3 countable models.

EXAMPLE 2.11
Let $\mathcal{A} = (A, <, a_n)$ be the rational numbers in the open interval $(-1, 1)$ and let a_n denote $-1/n$.

Examining this example with care will illuminate many of the concepts discussed in the next few pages. No really different example of a theory with a finite (> 1) number of countable models has been discovered. Vaught's epochal paper [74], which is still required reading, began showing the connections between types and the number of models. Among his results, Vaught showed that there is no complete theory with exactly two countable models and the following result linking the number of types with the existence of saturated models.

THEOREM 2.12
For any complete countable T, $S(\emptyset)$ is countable if and only if T has a countable saturated model.

This built on two earlier results characterizing semantic properties of theories in terms of the number of types.

THEOREM 2.13 (Łos)
For any complete countable T, if $S(\emptyset)$ is uncountable then T has $2^{\aleph_0}$ countable models.

THEOREM 2.14 (Ryll-Nardjewski et al)
For any complete countable T, $S^n(\emptyset)$ is finite for each $n < \omega$ if and only if T is $\aleph_0$-categorical.

Morley took this notion one step further in defining the concept of an ω-stable theory; we describe some important properties of such theories later. A key point is that the definition involves types over arbitrary sets not just the empty set as in the previous three theorems.

DEFINITION 2.15
A complete countable theory T is ω-stable if for every countable $A \subset M \models T$, $S(A)$ is also countable.

Morley showed that for an ω-stable theory and any A contained in a model of T,

$$|S(A)| \leq |A| + \aleph_0.$$

A priori, the collection of cardinals in which a first order theory is categorical could be arbitrary. But Morley's seminal proof of the Łos categoricity conjecture (for countable languages), Theorem 2.16, showed the situation was much more structured than that.

THEOREM 2.16 (Morley)
A complete countable theory T is categorical in one uncountable cardinal if and only if it is categorical in all uncountable cardinalities.

As part of his categoricity proof, Morley showed that all countable theories, categorical in some uncountable cardinality, were ω-stable. This demonstrates again the role of counting types in determining the properties of a theory. In view of his result we may identify '$\aleph_1$-categorical' with 'uncountably categorical'.

2.3 Infinite Model Theory: The 1970's

Two results of the very early 70's characterize the knowledge of categoricity at that time. A new approach to the study of $\aleph_1$-categorical theories isolated certain canonical examples, strongly minimal theories, and showed that every $\aleph_1$-categorical theory was built up from these examples.

Definition 2.17
The formula ϕ is strongly minimal in a theory T if in every model of T, ϕ has infinitely many solutions but for each formula ψ either $\phi \wedge \psi$ or $\phi \wedge \neg\psi$ has only finitely many solutions. A theory T is strongly minimal if $x = x$ is strongly minimal in T.

It is easy to see that

Theorem 2.18
A strongly minimal theory is categorical in any uncountable cardinality.

One of the key properties of ω-stability is that ω-stable theories admit prime models over sets:

Definition 2.19
The model M of a theory T is *prime* over A if every elementary embedding of A into a model N of T extends to an elementary embedding of M into N.

Baldwin and Lachlan [6] showed that strongly minimal sets controlled each model of an $\aleph_1$-categorical theory; this characterization played an important role in answering a question of Vaught (Theorem 2.20 2).

Theorem 2.20 (Baldwin-Lachlan)
1. Every model M of an $\aleph_1$-categorical theory T is prime over $\phi(M)$ for a fixed strongly minimal formula ϕ.
2. An uncountably categorical countable theory has either one or $\aleph_0$ countable models.

Around 1970, Glassmire [34], Ehrenfeucht [31] and Henson [43] showed that there were $2^{\aleph_0}$ theories which are $\aleph_0$-categorical. The nicest proof constructs a family of amalgamation classes such that each example is the theory of a homogeneous universal model for one of the classes. These constructions generalize the construction of rationals or the random graph [64]. All of these theories can be easily seen to be unstable (Definition 2.25).

Three problems guided much research in model theory or more precisely abstract stability theory in the 1970's. The *spectrum function* of T, $n(T, \kappa)$, is the number of models of T with cardinality κ.

Conjecture 2.21 (Morley)
With the obvious exception of $\aleph_1$- but not $\aleph_0$-categorical theories, $n(T, \kappa)$ is a non-decreasing function.

Conjecture 2.22 (Vaught)
A first order theory with uncountably many countable models has $2^{\aleph_0}$ countable models.

Question 2.23 (Morley)
Are there any finitely axiomatizable uncountably categorical complete theories?

Shelah proved Morley's conjecture by reformulating it as as classification problem. He proposed to first characterize those theories which admitted a good structure theory and then to classify the models of the 'good' theories and compute the spectrum function of such theories (which would have the same form for each theory in a given class). This program evolved into the subject of stability theory. There are many expositions of this subject; the most complete is [69], which contains the results in first order model theory attributed to Shelah in this article. The key tool for the first classification is counting types - the stability classification. The key tool for the second is the development of a general kind of independence: nonforking.

DEFINITION 2.24
For any infinite κ, a theory T is κ-stable if for every A with $|A| \leq \kappa$, $|S(A)| \leq \kappa$.

In particular, the notion of ω-stability which Morley had defined by a notion of rank was placed by Shelah into a hierarchy of all first order theories. We restrict ourselves in this account to countable languages; most of Shelah's work applied to languages of arbitrary cardinality. Sometimes (e.g. the Łos conjecture for uncountable languages) these generalizations require much more difficult arguments which lead to insight for the countable case.

DEFINITION 2.25

1. T is *superstable* if it is stable in all cardinals that are at least $2^{\aleph_0}$.
2. T is *stable* if it is stable in some cardinal.
3. T is *unstable* if it is not stable in any cardinal.

THEOREM 2.26

1. If T is ω-stable, it is stable in every infinite cardinal. Thus, every ω-stable theory is superstable.
2. Every theory is ω-stable, superstable, stable or unstable.

Instability can be defined not only by counting types but by a local property: a combinatorial property of a single formula and a countable set of points. In more detail we have the following definitions. Since every complete first order theory T has the amalgamation property (for elementary embeddings) T has a monster model and in the following we can just require the existence of sequences satisfying these properties in that monster model. In the adaptation below to finite variable logic, where amalgamation is not automatic, these definitions become problematic.

DEFINITION 2.27

1. The formula $\phi(\mathbf{x}, \mathbf{y})$ has the order property in T if there exist $\langle \mathbf{a}_i, \mathbf{b}_i : i < \omega \rangle$ in the monster model of T such that
$$\phi(\mathbf{a}_i, \mathbf{b}_j) \text{ iff } i \leq j.$$
2. The formula $\phi(\mathbf{x}, \mathbf{y})$ has the strict order property in T if there exist $\langle \mathbf{b}_i : i < \omega \rangle$ in the monster model of T such that
$$(\forall \mathbf{x})\phi(\mathbf{x}, \mathbf{b}_i) \rightarrow \phi(\mathbf{x}, \mathbf{b}_j) \text{ iff } i \leq j.$$
3. The formula $\phi(\mathbf{x}, \mathbf{y})$ has the independence property in T if there exist sequences $\langle \mathbf{b}_i : i < \omega \rangle$, $\langle \mathbf{a}_\sigma : \sigma \in 2^\omega \rangle$ such that
$$\phi(\mathbf{a}_\sigma, \mathbf{b}_j) \text{ iff } \sigma(j) = 1.$$

We say T has one of the properties in Definition 2.27 when some formula has the property in T.

Thus unstable theories involve some kind of ordering on n-tuples (the domain of the order need not be defined). Two prototypical unstable theories are the theory of the random graph (independence property but not strict order) and dense linear order (strict order but not independence property); atomless Boolean algebras have both properties.

THEOREM 2.28
The theory T has the order property if and only if it has the independence property or the strict order property.

We write $S_\phi(A)$ for the collection of types over A which contain only instances of ϕ and $\neg\phi$ and which are complete for formulas of this form.

THEOREM 2.29
T is unstable if and only if one of the following equivalent statements hold.

1. For every κ, there is an A with $|A| = \kappa$ and $|S(A)| > \kappa$.
2. There is a formula ϕ and an A with $|A| = \aleph_0$ and $|S_\phi(A)| > \aleph_0$.
3. Some formula $\phi(\mathbf{x}, \mathbf{y})$ has the order property.

We mentioned (Conjecture 2.21) Morley's conjecture that the spectrum function was essentially monotone and Shelah's strategy of turning this into a problem of classifying theories. More specifically, Shelah's astonishing proof consisted of enumerating all the possible spectrum functions. He began with a sufficient condition for a theory having the maximal number of models.

THEOREM 2.30 (Shelah)
If T is not superstable, then for all κ,

$$n(T, \kappa) = 2^\kappa.$$

The main goal of the research was reached when Shelah proved 'the main gap'.

THEOREM 2.31 (Shelah)
For each theory T, either, for sufficiently large α, $n(T, \aleph_\alpha) = 2^{\aleph_\alpha}$ or there is a $\delta(T)$ such that for sufficiently large α, $n(T, \aleph_\alpha) \leq \beth(\delta(T), \alpha)$.

Moreover, when there are few models, each model is decomposed into a tree of small models. The careful computation of lower parts of the spectrum was only finished with recent work of [39]. Shelah also made major progress on Vaught's conjecture by proving it for ω-stable theories. Buechler, building on Newelski, proved the conjecture for superstable theories of low U-rank but it remains open in general.

THEOREM 2.32 (Shelah)
An ω-stable first order theory with uncountably many countable models has $2^{\aleph_0}$ countable models.

These two results depended on many ideas. One of the most important is the classification of theories by counting types: stability. (Further classification by more technical conditions were essential to the argument.) The second was the development of forking. The notion of forking generalizes the Van Der Waerden notion of algebraic dependence in several ways. Perhaps the most important change is to drop the global requirement of transitivity of dependence (which allows one to assign a global dimension to a structure) and to replace it with a somewhat weaker notion which assigns a dimension to certain subsets (the relalizations of regular types); further arguments show the dimensions of these sets control the model.

Question 2.23 is more obviously connected to finite model theory. Here are three examples of *axioms of infinity*: first order sentences with only infinite models.

1. infinite linear order
2. $f(x)$ is an injective function; exactly one element does not have a predecessor.
3. $t(x, y)$ is a pairing function

Definition 2.33
A *complete* axiom of infinity is a first order sentence ϕ such that the consequences of ϕ are a complete first order theory and ϕ has an infinite model.

It is easy to extend 1) linear order to a complete sentence; the second sentence already axiomatizes a complete theory; Lachlan [8] extended the third to a complete stable but not superstable theory. But asking that a sentence be not merely complete but categorical is more subtle.

Theorem 2.34 (Zilber, Cherlin-Harrington-Lachlan)
No complete first order sentence is categorical in all infinite cardinalities.

Each strongly minimal set gives rise to a combinatorial geometry (or matroid) by the closure notion of algebraic closure. Geometries are classified as: trivial, locally modular, non-locally modular depending on whether the lattice of algebraically closed sets is distributive, modular (after naming a constant), or otherwise. The proof of Theorem 2.34 depended on an analysis of the geometries associated with various strongly minimal sets in a model and led to 'geometric stability theory'. One early proof relied on the classification of the finite simple groups. Zilber's argument eventually translated into a new proof of some results classifying two-transitive groups.

Theorem 2.35 (Peretyatkin)
There is an $\aleph_1$-categorical first order sentence.

Peretyatkin's seems to capture 'pairing'. This can be seen by examining the account in [45]. Refining the problem a bit leaves an open question.

Question 2.36
Is there a finitely axiomatizable strongly minimal set?

The key ideas of stability theory can be summarised as:

- Classify theories by counting types.
- Count models of stable theories by understanding dimension relations.
- Instability is caused by combinatorial properties of formulas.
- Study finer problems by analyzing combinatorial geometries.

2.4 Infinite Model Theory: The 1980's and 90's

2.4.1 'Algebraic Model Theory'

Before returning to the relation between stability theory and finite model theory, we will briefly discuss some of the main strains of model theory in the 80's and 90's. This work is much more algebraic - both in application and in methods. It continues the tradition of the school of Abraham Robinson (which we have shamefully neglected in this discussion) but also depends either in spirit, o-minimality, or directly, applications to algebraic geometry, on the fundamental insights of Shelah.

DEFINITION 2.37
A theory, T, is o-minimal if the universe of each model of T is linearly ordered and in every model every definable subset is a union of intervals.

EXAMPLE 2.38
1. $(R, +, \cdot)$ (Tarski)
2. $(R, +, \cdot, e^x)$ (Wilkie)
3. Many other expansions of the reals: (see the survey [70] and book [28])

There is an extremely highly developed theory of o-minimal structures with important applications to real algebraic geometry and to analysis; we describe some connections with finite model theory in Subsection 4.2.

A more direct connection with finite model theory is the work of Cherlin, Lachlan and Hrushovski [57, 21, 23] which analyzes in great detail the finite approximations to stable $\aleph_0$-categorical, quantifier eliminable structures and the extensions to the unstable case [58].

Major work in the 90's produces increasingly strong connections and applications between model theory and real and complex algebraic geometry and Diophantine geometry. It is difficult to see direct connections of this work with finite model theory. However, there is one link. Zilber proved that the geometry of a strongly minimal set could be classified as trivial, locally modular, or otherwise. He conjectured that all non-locally modular strongly minimal sets were 'essentially' fields. Hrushovski [48] disproved that conjecture. The counterexample is close to results concerning 0-1 laws (Section 4.4). Hrushovski and Zilber reformulated and proved the conjecture for the more restrictive class of Zariski geometries [50]. This abstract result is key to applications to Diophantine Geometry. Space considerations preclude serious consideration of this work here. The collection [16] not only describes Hrushovski's proof of the geometric Mordell-Lang conjecture but provides a good survey of basic stability theory with algebraic applications.

2.4.2 Classification theory in Infinitary Logic

Shelah has led another school which has developed over the last thirty years a model theory for classes defined in infinitary logic. The aim is to prove analogs of e.g. Morley's categoricity theorem for logics beyond first order. At first sight this subject seems far removed from finite model theory; the field is permeated with theorems which depend on extending the axioms of set theory. There are however several important similarities and possible analogies. The compactness theorem fails; therefore results, which in first order logic are easily derived from the compactness theorem, are shown by difficult combinatorial arguments. The hope is that combinatorial arguments which rely on large cardinals can be replaced on finite models by the observation that ω is large with respect to any finite number. Further, Shelah has developed this theory in a formulation which is almost universal algebra. In addition to the many papers of Shelah (e.g., [59, 67, 66]), the survey of Villaveces [75] puts this subject in context. This kind of generality led to the Baldwin-Shelah work on 0-1 laws [9, 3, 2]. Further, it was by exploiting these analogies that Hyttinen laid the groundwork for stability theory for the L^n-theory of finite models that we discuss below.

3 Finite Model Theory

Finite model theory stands at the intersection of complexity theory and model theory. Most of the work thus far in finite model theory is analogous to work in 1950's style model theory of first order logic. That is, the focus is on the study of properties of the logic such as expressed in the following results.

THEOREM 3.1 (Trahtenbrot)
The collection of sentences valid on finite models is not recursively enumerable.

THEOREM 3.2
Compactness fails on the class of finite models.

Various examples due as early as the fifties to Tait, sporadically since then, and more systematically in the 80's by Gurevich and various collaborators show that the preservation theorems of first order model theory almost uniformly fail when restricted to finite model theory.

One of the insights of work in finite model theory is encoded in the following mantra.

First order logic is too strong as it is categorical on finite models.

First order logic is too weak as it doesn't support recursion or counting.

Connections with complexity theory have so profoundly influenced the development of finite model theory as to result in the new field of descriptive complexity theory [52]. The fundamental insight of Fagin and Immerman is that one can characterize the complexity of recognizing (by a Turing machine) members of a class of finite structures by the logic in which the class is described. Here are a few of the basic results. We write **P** (**NP**) for the class of classes of structures decidable in deterministic polynomial time (nondeterministic polynomial time).

THEOREM 3.3
If **K** is a class of finite structures that is defined by a first order sentence then **K** is decided in polynomial time.

THEOREM 3.4 (Fagin)
On finite structures $\mathbf{\Sigma_1^1} = \mathbf{NP}$.

DEFINITION 3.5
FP denotes the logic obtained by adding to first order logic formulas expressing the fixed point of inductive definition over a formula $\phi(\mathbf{x}, X)$ where the relation symbol X occurs positively.

THEOREM 3.6 (Immerman-Vardi)
On *ordered structures* $\mathbf{FP} = \mathbf{P}$.

QUESTION 3.7
Is there a logic which captures polynomial time on arbitrary finite structures?

The study of 0-1 laws represent a second distinguishing feature of finite model theory; we discuss them in Subsection 4.4.

3.1 Abstact and Finite Model Theory

The 'weakness' of first order logic on finite structures has led to the study of various extensions of first order logic. Much of this work follows up another theme of 70's model theory: Generalized Quantifiers. Two examples were: 'There exists uncountably many x such that $\phi(x)$' and 'There exists an uncountable set which is homogeneous for $\phi(x, y)$'.

A general theory of such quantifiers is due to Lindstr—"om. There were many results on the model theory of such logics and a general theory of which logics have the interpolation property. This work was summarised as 'Abstract model theory' in [13]. One application to finite model theory is to replace the 'uncountable' with a function measuring the size of a solution set compared to the size of the universe (see Subsection 4.4). A second is the introduction of counting quantifiers.

The following theorem of Lindström not only characterizes first order logic but provides a proof technique which has played a role in 'collapse theorems' of embedded finite model theory as discussed in Subsection 4.2. The idea is that if a sentence ψ (of an 'extended logic') defines a class which is not defined in first order logic, there are a family of pairs of models A_n, B_n which witness this failure ($A_n \models \psi$ and $B_n \models \neg\psi$) but A_n and B_n are equivalent on sentences of quantifier rank up to n. The compactness theorem allows one to encode this so as to construct a pair of countable models which are elementarily equivalent (and indeed isomorphic!).

THEOREM 3.8 (Lindström)
No proper extension $\mathcal{L}$ of first order logic (**FO**) satisfies both the compactness theorem and the Lowenheim-Skolem theorem.

3.2 Finite variable logic

Finite variable logic is a useful tool for the study of finite models.

DEFINITION 3.9
1. L^n denotes first order logic with only n-variables *free or bound.*
2. $L^n_{\omega_1,\omega}$ denotes $L_{\omega_1,\omega}$ with only n-variables *free or bound.*
3. $L^\omega_{\omega_1,\omega} = \bigcup_n L^n_{\omega_1,\omega}$.

Barwise [12] used this logic to analyze inductive definitions on certain infinite structures. The technique moved to finite model theory with the following result.

THEOREM 3.10 (Kolaitis-Vardi [56])
On finite structures, **FP** is contained in $L^\omega_{\omega_1,\omega}$.

It is more convenient to analyze $L^\omega_{\omega_1,\omega}$ by means of 'pebble games' than to work directly with fixed-point logics. Considerable work has attempted to find a logic which captures polynomial time, e.g. by adding Lindström quantifiers. However, this particular approach is doomed to failure as shown by results of Cai, Furer and Immerman [17] and Hella [42] (see [30]).

THEOREM 3.11 (Hella)
No extension of $L^{\omega}_{\omega_1,\omega}$ (and thus of **FP**) by a finite number of Lindström quantifiers captures polynomial time.

As one of the referees pointed out to me this remark is somewhat misleading. If fact, the weakness of the Lindström quantifiers is the failure of the logics obtained by adding the generalized quantifiers to obey certain natural closure conditions (e.g. vectorization). Dawar proves that if any logic captures **P** then there is one which is an extension of first order by a Lindström quantifier and vectorization in [24]; [41] includes a nice summary of this situation.

3.3 L^n-theories

What classes of finite models admit a structure theory which it is profitable to study? We need a notion of complete theory which is both sufficiently weak – it must admit arbitrarily large finite models – that there is something to study and sufficiently strong that one can say something about the class of models. Our candidate is a complete theory in L^n. The similarity of the models of such a theory is guaranteed because they are equivalent with respect to n-pebble games. Such works as [25, 27, 26, 1, 36, 37, 61] began the development of 'Vaught-style' model theory for finite variable logic.

DEFINITION 3.12

1. An L^n theory is a collecion of L^n sentences, which is closed (within L^n) under logical consequence.
2. For a finite structure $M \models T$, $m \leq n$, and $\mathbf{b} \in M^m$, $\mathrm{tp}_{\mathrm{L}^n}(\mathbf{b}/\emptyset, \mathrm{M}) = \{\phi(\mathrm{x}, \mathbf{a}) : \mathrm{M} \models \phi(\mathbf{b}, \mathbf{a})\}$ is the m-type in L^n, realized by $\mathbf{b}$.
3. $S^n_n(M)$ is the collection of n-ary L^n-types realized in M.
4. The n-size of M is $|S^n_n(M)|$.
5. If T is an L^n theory, which is complete for L^n, all models M of T have the same $S^n_n(M)$, denoted $S^n_n(T)$. Its cardinality is the n-size of T.

Careful consideration of the construction of a Scott sentence justifies the assertion in the section of Definition 3.12 that the n-size of T is well-defined. Moreover,

FACT 3.13
Let T be an L^n-complete L^n-theory.

1. If T has a finite model, then T has finite n-size.
2. If T has finite n-size, then T is axiomatized by a single sentence of L^n.
3. For finite A, B, if $A \equiv_{L^n} B$ then $A \equiv_{L^n_{\omega_1,\omega}} B$

Note that for dense linear order $S^3_3(T)$ is finite but there are no finite models. The *n-size* of an L^n theory [1, 25] has important connections with calculations on the relational machines of Arbiteboul and Vianu [1].

LEMMA 3.14 (Dawar)
For finite A whose L^n-theory is T, there are functions

1. $L_n(m)$ (Lowenheim) such that $|S_n^n(T)| \leq m$ implies there is a model B with $B \equiv_{L^n} A$ and $|B| < L_n(m)$;
2. $H_n(m)$(Hanf) such that $|S_n^n(T)| \leq n$ and there is a model B such that $B \equiv_{L^n} A$ and $|B| > H_n(m)$ implies there are arbitrarily large finite models L^n-equivalent to A.

Dawar asked whether these functions could be recursive. Grohe [36] has shown L_n can not be recursive. Barker [11] extended his methods to refute the Hanf number conjecture. The main tool is a coding of Hilbert's 10th problem.

4 Current Interactions between finite model theory and stability

Our discussion of the interactions of finite model theory with stability theory will cover 4 areas: Stability Theory for L^n, Embedded finite model theory, L^n-axiomatizability of first order theories, and 0-1 laws.

4.1 Stable L^n-theories

Here are some examples of complete L^n-theories; they are all routine except item 3 which relies on deep results of Cherlin, Harrington, and Lachlan [22] and Hrushovski[46] which are discussed in Theorems 2.34 and 4.11.

1. For $n \geq 3$, the vector spaces of dimension at least n over a finite field are axiomatized in L^n and are all L^n-elementarily equivalent.
2. For any complete T, T_n, the L^n sentences of T; for any structure A, the L^n-theory of A.
3. If a first order theory, T, is categorical in all infinite cardinalities, T is axiomatized by a single sentence plus an axiom of infinity. For n large enough to include that sentence, T_n is L^n-complete.
4. The theory of an equivalence relation with one class of each size less than n is a complete L^n-theory.
5. The n-extension axioms (i.e. the L^n-theory of the random graph).
6. Dense linear order without end points.

The first difficulty in developing stability theory in this context is to extend the idea of types in L^n over the empty set (Subsection 3.3) to a notion of type (in a given ambient model M) over an arbitrary subset of M. This was accomplished by Hyttinen; we use the formulation of Djordjević.

Definition 4.1
Let $M \models T$, $A \subseteq M, \mathbf{b} \in M$.

$$\mathrm{tp_{L^n}(\mathbf{b}/A, M) = \{\phi(x, \mathbf{a}) : M \models \phi(c, \mathbf{a})\}}$$

where $\mathbf{a}, \mathbf{x}, \mathbf{b}$ etc. are finite sequences with $\mathbf{b} \in B$, $\mathbf{a} \in M$ and $\phi(\mathbf{x}, \mathbf{b})$ is a substitution instance of a formula in L^n.

For finite $B \subset M$, we write $\mathrm{tp_{L^n}(B;M)}$ or $\mathrm{Th_{L^n}(M,B)}$ for $\mathrm{tp_{L^n}}(\mathbf{b}/\emptyset,\mathrm{M})$ where $\mathbf{b}$ lists B. We use the same notation for infinite B but here we mean the union of all the types of the finite sequences; we permit infinitely many free variables but each formula in the type is in L^n. Even if B is finite the conjunction of the type need not be in L^n. In the absence of amalgamation, we need the parameter M. Finally, let

$$S_n^k(A,M) = \{\mathrm{tp_{L^n}}(\mathbf{b},\mathrm{A},\mathrm{N}) : \mathbf{b} \in \mathrm{N^k} \text{ and } (\mathrm{M},\mathrm{A}) \equiv_{\mathrm{L^n}} (\mathrm{N},\mathrm{A})\}.$$

Our previous Definition 3.12 of $S_n^n(T)$ yields in the present notation: $S_n^n(T) = S_n^n(\emptyset, M)$ (any $M \models T$).

Any complete first order theory has the amalgamation property for elementary embeddings. Thus, in a first order theory the meaning of $S(A,M)$ does not depend on M (we just work in the monster model). Here, unless we explicitly assume amalgamation, we must specify the ambient model M in order to define the notion of type. For, there are complete L^n-theories which do not have the amalgamation property for L^n-elementary embeddings. Poizat [63] constructed complete L^n-theories with exactly one model with cardinality n or $n+1$ but various possibilities for other finite models. He asked whether it was possible for a complete L^n-theory to have more than one but still finitely many finite models. Thomas [73, 72] has constructed for $n \geq 3$ theories T_n which have nonisomorphic models of the same finite cardinality and no larger models. He has similar examples with infinite models. The work of Djordjević [29] discussed below depends on the following strong form of the amalgamation property.

DEFINITION 4.2
T has the *amalgamation property over sets* for L^n if the following holds:
If M_1 and M_2 are models of T and $a_\alpha \in M_1$, $b_\alpha \in M_2$, for $\alpha < \kappa$, and

$$\mathrm{Th_{L^n}}(\mathrm{M_1},(\mathrm{a}_\alpha : \alpha < \kappa)) = \mathrm{Th_{L^n}}(\mathrm{M_2},(\mathrm{b}_\alpha : \alpha < \kappa))$$

then there are $N \succcurlyeq_{L^n} M_1$ and an L^n-elementary embedding $f : M_2 \to N$ such that $f(b_\alpha) = a_\alpha$ for all $\alpha < \kappa$.

Djordjević [29] has shown one of Thomas's examples does not have the amalgamation property over sets. Further examples of complete L^n-theories with only finitely many finite models were constructed by Cherlin [20] and as a by-product of another construction by Gurevich and Shelah [38] - multipedes. Still further examples appear in the work of Grohe [36]; essentially for Diophantine equations with a finite number of solutions he constructs a complete theory T in an appropriate L^n such that the set of cardinalities of finite models of T is determined by the set of solutions of the equation. In each case, in view of Djordjević's Theorem 4.11, such examples are either unstable or don't have the amalgamation property over sets.

There are a number of questions which arise around the amalgamation property.

QUESTION 4.3
Is there an L^n-theory which is complete in first order logic and has only finitely many L^n-types but does not have the L^n-amalgamation property (over sets)?

Rosen has observed that, if not, Theorem 4.11 yields that no $\aleph_0$-categorical, stable theory is finitely axiomatizable. Hyttinen [51] defines the independence property for L^n as follows.

DEFINITION 4.4
An L^n-theory T has the *independence property* if for some L^n-formula ϕ for every $m < \omega$ there exists a model M_m of T and elements $A = \langle a_i : i < m\rangle \in M_m$ for such that for every $X \subset m$ each type

$$p_X = \{\phi(v, a_i) : i \in X\}$$

is consistent.

More precisely, for each $X \subseteq m$, there is a model M_X containing A with p_X realized in M_X such that for each $X, Y \subseteq m$, (M_X, A) and (M_Y, A) are L^n-equivalent. In the absence of amalgamation the various p_X do not have to be simultaneously realizable.

PROBLEM 4.5
Construct a complete theory in L^4 which has the independence property in the sense of Definition 4.4 but such that ϕ does not satisfy the m-extension axiom for sufficiently large m.

In the absence of amalgamation, Hyttinen [51] proves that a stability notion (defined in terms of non-splitting sequences) is equivalent to the conjunction of 'not the order property' and 'not the independence property'. Hyttinen's paper [51] takes place in a somewhat more abstract setting than L^n and a number of results are proved without the amalgamation property. A uniform notation for considering the variants of notions like the order property, independence property, stability in the absence of amalgamation needs to be adapted.

DEFINITION 4.6
An L-structure M is *κ-saturated for L^n* if $|M| \geq \kappa$ and for every $A \subseteq M$ with $|A| < \kappa$ and every $p \in S^n(A, M)$, p is realized in M.

THEOREM 4.7 (Hyttinen, Djordjević)
If T has L^n-amalgamation over sets then for any κ, every model of T can be L^n-embedded in an L^n-κ-saturated model.

Since there are $2^{\aleph_0}$ distinct divisibility types (describable in L^3), if T is the L^3-theory of $(Z, +, 1)$, there is no countable countably L^3-saturated model. The following example illustrates L^n-saturation.

Consider the structure (M, E) where: E is an equivalence relation; there is one class of each finite cardinality; there are infinitely many infinite classes.

M is L-saturated but not L^n-saturated for any n. Fix $m > n$; let $a_0, \ldots a_{m-1}$ enumerate one equivalence class. The L^n-type $p = \{xEa_1 \wedge x \neq a_i : i < m\}$ is omitted in M. Note that p is inconsistent with the L^n-theory of $(M, \mathbf{a})$ but consistent with the first order theory of M. All finite equivalence classes in an L^n-saturated model of T have cardinality at most n.

To define stability, we parallel the most syntactic definition; here is a definition which does not require the amalgamation property.

DEFINITION 4.8
A complete L^n-theory T is *stable* if there is no L^n-formula $\phi(\mathbf{x}, \mathbf{y})$ with the order property.

Thus, T is unstable if in some model M of T there are an L^n-formula ϕ and sequences $\mathbf{a}_i, \mathbf{b}_j$ for $i, j < \omega$ such that

$$\phi(\mathbf{a}_i, \mathbf{b}_j) \text{ iff } i < j.$$

Working entirely in finite models we could talk about arbitrarily long sequences satisfying these conditions. Hyttinen has taken a different definition and there is continuing discussion about the proper framework for this study.

Definition 4.9
A complete L^n-theory T is *ω-stable* if for every $M \models T$ and every infinite $A \subseteq M$, and every m

$$|S^m_{L^n}(A, M)| = |A|$$

.

For the interesting case of L^n-theories with finite models the stability hierarchy collapses

Theorem 4.10
If T is a stable theory with the amalgamation property over sets in n-variable logic and $S^n_n(T)$ is finite then T is *ω-stable*.

This is based on old rank computations by Shelah but one must be careful working with incomplete theories. Various variants have been shown by Djordjević, Hyttinen, Lessmann and myself. There are foundational questions about which definition of stable/ω-stable to take. However, all the possibilities agree if the theory T has the amalgamation property over sets. The hypothesis that $S^n_n(T)$ is finite is essential in the last theorem (consider $(Z, +)$); for this definition of ω-stable, amalgamation is not used but it is required for some more restrictive definitions.

The following theorem unites the methods of stability theory and finite model theory.

Theorem 4.11
(Djordjević) If a stable L^n-theory T with the amalgamation property over sets has only finitely many L^n-types then T can be extended to an L-theory which is $\aleph_0$-categorical and ω-stable. Further, T has infinitely many finite models.

Here is a sketch of the proof [29].
1. Let T^c be the theory of an L^n-saturated model.
2. Prove T^c is $\aleph_0$-categorical and stable.
3. Using ω-stablity of T, show that T^c is ω-stable.
4. The Zilber/Cherlin-Harringon-Lachlan theorem [22] asserts:
An ω-stable, $\aleph_0$-categorical T has the *strong finite model property*:
If $M \models T$ and $M \models \psi$ then some finite substructure of N of M models ψ.
5. Apply this result to ψ – the sentence axiomatizing the L^n theory T – to obtain the finite models.

This theorem is not as widely applicable as one might hope. Djordjević observed that the L^n-theory of vector spaces does not have amalgamation over finite sets. In fact, more generally he has shown a strictly minimal theory in a L^n-theory which has amalgamation over sets, algebraic closure is trivial. Recall algebraic closure is *trivial* if whenever $a \in \mathrm{acl}(\mathrm{B})$, $a \in \mathrm{acl}(\mathrm{b})$ for some single element b of B.

Lessmann and Baldwin [7] (with considerable input by Rosen) have rescued this situation by generalizing the Djordjević analysis to a more general context (by assuming a weaker amalgamation principle). This context includes the case of vector spaces over a finite field. This is the first place where the analogy between finite model theory and infinitary logic has directly born fruit. Lessmann's analysis extends to the study of 'simple' theories in L^n and directly establishes the combinatorial notion of forking in this context.

4.2 Embedded Finite Model Theory

View a database of employment records as a ternary relation containing triples: name, position, salary. Each data base is a finite relational structure. So the study of finite models is an abstract way of considering data bases. But in the background we make queries such as, 'List all employees who make over \$50,000 per year.' (That is, we have access to $(Z, +, <)$.) So a framework which is in some ways more faithful to actual application is that of embedded finite models which are defined as follows.

In the following L and S are disjoint finite relational languages.

DEFINITION 4.12
Let M be an L-structure with domain $\mathbb{U}$. For any (finite) S-structure A with domain contained in $\mathbb{U}$, let $M(A)$ denote the unique $L \cup S$-structure whose reduct to L is M and the interpretation of the predicates in S is just that of A. Then A is an *embedded finite model.*

It is natural to ask whether consideration of this infinite universe increases our expressive power. That it, does it allow us to define additional classes of finite S-structures? We need a few more definitions to formalize this question.

DEFINITION 4.13
Let $\mathbb{U}$ be an infinite set, and let S be a finite relational language.

1. A *query* on $\mathbb{U}$, Q, is a collection of S-structures with domain a finite subset of $\mathbb{U}$.
2. A *generic query* on $\mathbb{U}$ is a query that is closed under S-isomorphism.
3. An *abstract query* Q is a collection of isomorphism types of finite S-structures.

Given an abstract query Q and set $\mathbb{U}$, Q defines a generic query $Q_{\mathbb{U}}$ by considering all finite S-structures embedded in $\mathbb{U}$ whose isomorphism type is in Q.

Given a first-order sentence ϕ in $L \cup S$, the *query defined by* ϕ is the set of finite S-structures A with $A \subset \mathbb{U}$ such that $M(A) \models \phi$. ϕ is generic if the query defined by it is generic. A query on $\mathbb{U}$, Q, is *first-order over* M if there is $\phi \in L \cup S$ that defines Q; Q is a *pure first-order query* if can be defined by a ψ using only symbols from S. An abstract query Q, is first-order over M if there is a first-order $\phi \in L \cup S$ that defines $Q_{\mathbb{U}}$.

There are several versions of *generic collapse to equality* which assert that there is no increase in expressiveness by considering embedded finite models.

THEOREM 4.14 (Hull-Su, Belegradek et al)
Let $\mathbb{U}$ be infinite. If Q is a first-order definable generic query over $\mathbb{U}$, then Q is a pure first-order query.

This result can be proved along the same lines as Lindström's Theorem 3.8. Baldwin and Benedikt [5] extended this result by replacing the ambient structure $\mathbb{U}$, the pure theory of equality, with any stable structure. E.g, the ambient structure could be any abelian group or an algebraically closed field.

THEOREM 4.15
Let $\mathbb{U}$ be infinite and M be a stable L-structure with domain $\mathbb{U}$. If Q is a first-order definable generic query over M, then Q is a pure first-order query.

Again, the final argument is a variant of the Lindström construction. But now technical tools of stability must be used to find a set of indiscernibles (in an appropriate expansion of the base language). The crux is an inductive argument reducing the quantifier complexity of a formula. This argument was generalized back into standard stability theory in [18].

However, many natural candidates for the ambient structure are not stable; in particular they admit a natural linear order. So we make a similar sequence of definitions for *ordered queries* [5]. The proof of the following results in [5] is largely parallel to the proof for the stable case but the basic quantifier reduction argument is different for this case.

THEOREM 4.16
Let $(\mathbb{U}, <)$ be an infinite linear order and M an $L \cup \{<\}$-structure that expands $(\mathbb{U}, <)$, which does not have the independence property. If Q is a first-order definable locally order generic query over M, then Q is an order-definable query.

THEOREM 4.17
Let $\mathbb{U}$ be infinite and M be an L-structure with domain $\mathbb{U}$ that does not have the independence property. If Q is a first-order definable generic query over M, then Q is definable over any dense linear order without endpoints.

A model M satisfying the second sentence of Theorem 4.17 is said to have *generic collapse to order*.

The study of expressibility over embedded finite structures began in connection with a particular database formalism, constraint databases, in [55] (there are other frameworks for studying "mixed" structures, particularly the one presented in [35]). [55] focused on the real ordered group and ordered field. Both Theorem 4.16 and Corollary 4.19 were proved in the special case of the real ordered group in [62]. This result was extended to o-minimal structures in [15] and to quasi-o-minimal structures in [14]. Here are some further corollaries of these methods.

COROLLARY 4.18
Let $\mathbb{U}$ be infinite and M be an L-structure with domain $\mathbb{U}$ that does not have the independence property. If Q is a first-order definable generic query over M, then the abstract ordered query defined by Q (i.e the class of linearly ordered expansions of isomorphism types in Q) is definable by a formula in $\{S, <\}$.

COROLLARY 4.19
Let S contain a single binary relation symbol. If M does not have the independence property, then the following queries are not first-order definable over M: parity, transitive or deterministic transitive closure of a graph; maximal matching in a bipartite graph; Eulerian cycle.

4.3 From finite to infinite: L^n-axiomatizability

In this section we reverse the main line of our discussion and describe the use by Shawn Hedman [40] of methods developed primarily for the study of finite models to deepen our understanding of the axiomatizability of $\aleph_1$-categorical theories. The proof of the existence of a Scott sentence in L^n for a finite structure extends to any theory with finite size; this yields:

FACT 4.20
An $\aleph_0$-categorical first order theory is L^n-axiomatizable (for some n) if and only if it is finitely axiomatizable.

Hedman [40] shows eliminating functions for relations does not affect finite variable axiomatizability (although the exact n shifts a bit). Hedman's deeper results depend on a more subtle analysis of 'trivial' strongly minimal sets than had previously been given. Following Gaifman [33], in any structure M define $d_M(a,b) = 1$ if for some relations symbol R and some sequence $\mathbf{e}$ including a, b, $R(\mathbf{e})$ holds. By induction $d_M(a,b)$ is at most $n+1$ if b is directly connected to a point c with $d_M(a,c) \leq n$.

Two elements a and b are said to be *generic* if they are algebraically independent.

THEOREM 4.21 (Hedman)
In any strongly minimal set, two generic points a, b satisfy one of the following.

Type I. $d_M(a,b) = 1$
Type II. $d_M(a,b) = 2$
Type III. $d_M(a,b) = \infty$

Hedman shows that if a strongly minimal set is not locally modular then it is of Type I. More important for our present aims, he shows:

THEOREM 4.22 (Hedman)
Every trivial strongly minimal set is bidefinable with one of type III.

The connection with axiomatizability comes through Hrushovski's Galois theory for strongly minimal sets. Hrushovski defines the notion of a finite group G being *involved* in T; this means, roughly, G occurs as a Galois group of B/A for $A \subset B$ contained in a model of T. If A contains the algebraic closure of the empty set G is *strongly involved*. In [49], Hrushovski connects this notion with finite axiomatizability.

THEOREM 4.23 (Hrushovski)
If an $\aleph_1$-categorical T is finitely axiomatizable only finitely many simple groups are involved in T.

Hedman proves a partial converse in terms of L^n-axiomatizability.

THEOREM 4.24 (Hedman)
If only finitely many simple groups are involved in a trivial strongly minimal (but not $\aleph_0$-categorical) theory then T is axiomatized with finitely many variables.

He almost obtains this hypothesis for type III theories.

THEOREM 4.25 (Hedman)
If T is a type III trivial strongly minimal theory only finitely many simple groups are strongly involved in T.

By Theorem 4.22, he can apply this result for any trivial (but not $\aleph_0$-categorical) theory; to get involved rather than strongly involved, he works over constants, yielding:

THEOREM 4.26 (Hedman)
If T is a trivial (but not $\aleph_0$-categorical) strongly minimal theory then T is axiomatized with finitely many variables over constants.

This area has a striking connection with classical mathematics.

CONJECTURE 4.27 (A. Robinson)
There is no n such that the complex field is L^n-axiomatizable.

Robinson's conjecture follows easily from the even stronger conjecture of Hedman which arises from consideration of Hrushovski [49].

THEOREM 4.28 (Hrushovski)
If a strongly minimal set is finitely axiomatized then the geometry is locally modular.

CONJECTURE 4.29 (Hedman)
If T is an L^n-axiomatized strongly minimal theory then the geometry is locally modular.

4.4 0-1 *laws*

Along with descriptive complexity, a major new aspect of finite model theory is the ability to construct a first order theory from a collection of finite models via a 0-1 law. Fix a finite relational language L. Let $\mathbf{K_n}$ be a collection of L-structures with universe n. Let P_n be a probability measure on $\mathbf{K_n}$. For any formula ϕ, let

$$P_n(\phi) = \sum\{P_n(B) : B \models \phi, |B| = n\}.$$

For example, $\mathbf{K_n}$ is all graphs with universe n; P_n is the uniform distribution (edge probability 1/2).

DEFINITION 4.30
The *almost sure theory* T for the sequence $(\mathbf{K_n}, P_n)$ is the collection of sentences ϕ such that

$$P(\phi) = \lim_{n\to\infty} P_n(\phi) = 1.$$

The sequence $(\mathbf{K_n}, P_n)$ satisfies the 0-1 law if the almost sure theory is complete.

This area began with the 0-1 law for finite graphs.

THEOREM 4.31 (Glebski et al, Fagin)
The theory of the random graph is almost sure with respect to the uniform distribution as each extension axiom has limit probability 1.

Such probabilists as Erdos, Renyi, Rado and Spencer developed a thriving field investigating the evolution of the random graph for various probability measures. One of the most striking results involves edge probabilities of the form $n^{-\alpha}$ for $0 < \alpha < 1$.

DEFINITION 4.32
Let $\mathbf{K_n}$ be the collection of graphs with universe n. Fix $0 < \alpha < 1$, and for $B \in \mathbf{K_n}$ let $e(B)$ denote the number of edges of B. Let $P_n^\alpha(B) = n^{-\alpha e(B)} \cdot (1 - n^{-\alpha})^{\binom{n}{2} - e(B)}$.

This induces as above a sequence of measures P_n^α on first order formulas.

THEOREM 4.33 (Spencer-Shelah)

- If α is irrational, for each sentence ϕ, $\lim_{n\to\infty} P_n^\alpha(\phi)$ is 0 or 1.
- If α is rational, there are formulas whose probabilities do not converge.

We close our survey with a surprising tie to the construction of homogeneous universal models. Recall that the earliest general constructions of many $\aleph_0$-categorical theories yielded unstable theories. In fact, the question of the existence of a stable but not superstable $\aleph_0$-categorical theory remained open for almost 25 years until solved by Hrushovski (see [44] for the best published account). The proof proceeded by varying the notion of 'strong substructure' in defining the amalgamation class. Hrushovski had used other versions of this construction to discover strongly minimal sets with strikingly new model theoretic properties [48, 47]. It turned out that a minor variant of this construction gave another (and more complete) proof of the model theoretic portion of the Shelah-Spencer theorem. Further, [9, 10] showed that the almost sure theory was stable and that the results went through for any (symmetric) finite relational language L; not just graphs.

THEOREM 4.34 (Baldwin-Shelah)
T^α, the almost sure theory of random L-structures (with respect to the measure in Definition 4.32 is the same as the theory T_α of the homogeneous universal model for a properly chosen class and notion of strong substructure. This theory is complete, stable, and not finitely axiomatizable.

From a model theoretic standpoint one of the most interesting features of this theory is that, while a fairly uniform axiomatization can be given [9], the theory requires two alternations of quantifiers (Π_3^0) for the axiomatization. Simple extension axioms (Π_2^0) of the type that axiomatize the random graph do not suffice. Independently, Baldwin[3] and Shelah [68] have extended this result to show the 0-1 law for expansions of successor. Baldwin [4] has combined three of the notions discussed in this paper: generalized quantifiers, random graphs, and stability to show

THEOREM 4.35
For appropriate functions f determining the interpretation of the Ramsey quantifier the logic $L_{\omega,\omega}(Q_{ram,f})$ satisfies the 0-1 law on graphs with respect to edge probability $n^{-\alpha}$ for irrational α.

References

[1] S. Abiteboul and V. Vianu. Generic computation and its complexity. In *23rd ACM Symposium on Theory of Computing*, pages 212–219, 1991.

[2] J. T. Baldwin. Near model completeness and 0-1-laws. In R. Boppana and James Lynch, editors, *Proceedings of the DIMACS Workshop on Logic and Random Structures: 1995*. American Mathematical Society, 1997.

[3] J.T. Baldwin. Random expansions of geometries. submitted.

[4] J.T. Baldwin. Stability, the finite cover property and 0-1 laws. *Journal of Logic and Computation.* to appear.

[5] J.T. Baldwin and M. Benedikt. Embedded finite models, stability theory, and the impact of order. In *Proceedings of 13th annual IEEE Symposium on Logic in Computer Science*, pages 490–500, 1998.

[6] J.T. Baldwin and A.H. Lachlan. On strongly minimal sets. *Journal of Symbolic Logic*, 36:79–96, 1971.

[7] J.T. Baldwin and O. Lessmann. Amalgamation properties and finite models in L^n-theories. preprint, 200x.

[8] J.T. Baldwin and R. N. McKenzie. Counting models in universal Horn classes. *Algebra Universalis*, 15:359–384, 1982.

[9] J.T. Baldwin and S. Shelah. Randomness and semigenericity. *Transactions of the American Mathematical Society*, 349:1359–1376, 1997.

[10] J.T. Baldwin and Niandong Shi. Stable generic structures. *Annals of Pure and Applied Logic*, 79:1–35, 1996.

[11] R. Barker. A proof that there is no recursive link between the k-size of a model and its cardinality. preprint, 200x.

[12] J. Barwise. On Moschovakis closure ordinals. *Journal of Symbolic Logic*, 42:292–296, 1977.

[13] J. Barwise and S. Feferman, editors. *Model-Theoretic Logics.* Springer-Verlag, 1985.

[14] O. Belagradek, A. Stolboushkin, and M. Tsaitlin. Extended order-generic queries. *Annals of Pure and Applied Logic*, 97:85–125, 1999.

[15] M. Benedikt, G. Dong, L. Libkin, and L. Wong. Relational expressive power of constraint query languages. In *Proceedings of 15th ACM Symposium on Principles of Database Systems*, pages 5–17, 1996.

[16] E. Bouscaren, editor. *Model Theory and Algebraic Geometry : An Introduction to E. Hrushovski's Proof of the Geometric Mordell-Lang Conjecture.* Springer-Verlag, 1999.

[17] J. Cai, M. Furer, and N. Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12:389–410, 1992.

[18] E. Casanovas and M. Zeigler. Stable theories with a new predicate. preprint.

[19] C.C. Chang and H.J Keisler. *Model theory.* North-Holland, 1973.

[20] G.L. Cherlin. Theories categorical in power $n + 2$. manuscript, 1982.

[21] G.L. Cherlin. Large finite structures with few types. In Valeriote Hart, Lachlan, editor, *Algebraic Model Theory*, pages 53–107. Kluwer Academic Publisher, 1997.

[22] G.L. Cherlin, L. Harrington, and A.H. Lachlan. $\aleph_0$-categorical, $\aleph_0$-stable structures. *Annals of Pure and Applied Logic*, 28:103–135, 1985.

[23] G.L. Cherlin and E. Hrushovski. Psuedofinite structures. preprint: www.math.rutgers.edu/ cherlin/Paper.

[24] A. Dawar. Generalized quantifiers and logical reducibilities. *Journal of Logic and Computation*, 5:213–226, 1995.

[25] A. Dawar. Types and indiscernibles in finite models. In J. Makowsky and E.V. Ravve, editors, *Logic Colloquium '95*, volume 11 of *Lecture Notes in Logic*, pages 51–65. Springer-Verlag, 1998.

[26] A. Dawar, S. Lindell, and S. Weinstein. Infinitary logic and inductive definability. *Information and Computation*, 119:160–175, 1995.

[27] A. Dawar, S. Lindell, and S. Weinstein. First order logic, fixed point logic, and linear order. In H. Kleine-Büning, editor, *Computer Science Logic '95*, pages 161–177. Springer-Verlag: LNCS 1092, 1996.

[28] Lou Van den Dries. *Tame Topology and O-Minimal Structures.* London Mathematical Society Lecture Note Series, 248, 1999.

[29] Marko Djordjević. Finite variable logic, stability, and finite models. preprint; to appear JSL, 200x.

[30] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory.* Springer-Verlag, 1995.

[31] A. Ehrenfeucht. There are continuum ω_0-categorical theories. *Bulletin de l'Acadamie Polonaise des sciences math. , astr., et phys.*, XX:425–427, 1972.

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

[33] H. Gaifman. On local and non-local properties. In *Proceedings Herbrand Logic Colloquium '81*, pages 105–135. North Holland, 1982.

[34] W. Glassmire. There are $2^{\aleph_0}$ countably categorical theories. *Bulletin de l'Acadamie Polonaise des sciences math. , astr., et phys.*, XIX:185–190, 1971.

[35] E. Grädel and Y. Gurevich. Metafinite model theory. *Information and Computation*, 140:26–81, 1998.

[36] M. Grohe. Large finite structures with few L^k-types. In *Proceedings 12th IEEE Symp. of Logic in Computer Science*, pages 216–227, 1997.

[37] M. Grohe. Finite variable logics in descriptive complexity. *The Bulletin of Symbolic Logic*, 4:345–398, 1998.

[38] Y. Gurevich and S. Shelah. On finite rigid structures. *The Journal of Symbolic Logic*, 61:549–562, 1996.

[39] B. Hart, Hrushovski E., and C. Laskowski. The uncountable spectra of countable theories. preprint.

[40] S. Hedman. *Finitary Axiomatizations of Categorical Theories.* PhD thesis, University of Illinois at Chicago, 1999.

[41] Hella and Nurmonen. Vectorization hierarchies of some graph quantifiers. to appear Archive for Math. Logic; preprint at www.math.helsinki.fi/ logic/hella/preprints, 200x.

[42] L. Hella. Logical hierarchies in Ptime. *Information and Computation*, 129:1–19, 1996.

[43] W. Henson. Countable homogeneous relational structures and $\aleph_0$-categorical theories. *The Journal of Symbolic Logic*, 37:494–500, 1972.

[44] B. Herwig. Weight ω in stable theories with few types. *J. Symbolic Logic*, 60:353–373, 1995.

[45] W. Hodges. *Model Theory.* Cambridge University Press, 1993.

[46] E. Hrushovski. Totally categorical structures. *Transactions of the American Mathematical Society*, 313:131–159, 1989.

[47] E. Hrushovski. Strongly minimal expansions of algebraically closed fields. *Israel Journal of Mathematics*, 79:129–151, 1992.

[48] E. Hrushovski. A new strongly minimal set. *Annals of Pure and Applied Logic*, 62:147–166, 1993.

[49] E. Hrushovski. Finitely axiomatizable $\aleph_1$-categorical theories. *Journal of Symbolic Logic*, 79:838–844, 1994.

[50] E. Hrushovski and B. Zilber. Zariski geometries. *Bulletin of the American Mathematical Society*, 28:315–324, 1993.

[51] T. Hyttinen. On stability in finite models. preprint.

[52] N. Immerman. *Descriptive Complexity.* Springer-Verlag, Heidelberg, 1999.

[53] B. Jónsson. Universal relational systems. *Mathematica Scandinavica*, 4:193–208, 1956.

[54] B. Jónsson. Homogeneous universal relational systems. *Mathematica Scandinavica*, 8:137–142, 1960.

[55] P. Kanellakis, G. Kuper, and P. Revesz. Constraint query languages. *Journal of Computer and System Sciences*, 51:26–52, 1995.

[56] P. Kolaitis and M. Vardi. Infinitary logics and 0-1 laws. *Information and Computation*, 98:258–294, 1992.

[57] A.H. Lachlan. Stable finitely homogeneous structures: a survey. In Valeriote Hart, Lachlan, editor, *Algebraic Model Theory*, pages 145–161. Kluwer Academic Publisher, 1997.

[58] D. Macpherson. Homogeneous and smoothly approximated structures. In Valeriote Hart, Lachlan, editor, *Algebraic Model Theory*, pages 161–181. Kluwer Academic Publisher, 1997.

[59] Michael Makkai and Saharon Shelah. Categoricity of theories in $l_{\kappa\omega}$, with κ a compact cardinal. *Annals of Pure and Applied Logic*, 47:41–97, 1990.

[60] M. Morley and R.L. Vaught. Homogeneous universal models. *Mathematica Scandinavica*, 11:37–57, 1962.

[61] Martin Otto. *Bounded Variable Logics and Counting.* Springer Verlag, Heidelberg, 1996. Lecture Notes in Logic:9.

[62] J. Paredaens, J. Van den Bussche, and D. Van Gucht. First-order queries on finite structures over the reals. In *Logic In Computer Science '95*, pages 79–87, 1995. Full version to appear in *SIAM J. Comput.*

[63] Bruno Poizat. Deux ou trois choses que je sais de L_n. *The Journal of Symbolic Logic*, 47:641–658, 1982.

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

[65] A. Robinson. Completeness and persistence in the theory of models. *Zeitshrift Math. Logik. and Grundlagen der Math.*, 2:15–26, 1956.

[66] S. Shelah. Classification theory for nonelementary classes. I. the number of uncountable models of $\psi \in l_{\omega_1 \omega}$ part a. *Israel Journal of Mathematics*, 46:212–240, 1983.

[67] S. Shelah. Universal classes: Part 1. In J. Baldwin, editor, *Classification Theory, Chicago 1985*, pages 264–419. Springer-Verlag, 1987. Springer Lecture Notes 1292.

[68] S. Shelah. Zero-one laws with probability varying with decaying distance. Shelah 467, 199?

[69] S. Shelah. *Classification Theory and the Number of Nonisomorphic Models.* North-Holland, 1991. second edition.

[70] C. Steinhorn. O-minimality. 200x.

[71] A. Tarski and R.L. Vaught. Arithmetical extensions of relational systems. *Compositio Mathematica*, 13:81–102, 1931.

[72] Simon Thomas. Theories with finitely many models II. manuscript, 198?

[73] Simon Thomas. Theories with finitely many models. *Journal of Symbolic Logic*, 51:374–376, 1986.

[74] R.L. Vaught. Denumerable models of countable theories. In *Infinitistic Methods, Proc. Symp. Foundations of Math., Warsaw, 1959*, pages 303–321. Oaństwowe Wydawnictwo Naukowe, Warsaw, 1961.

[75] Andrés Villaveces. De łoś al presente: un teorema aún central en teoria de la clasificación. preprint, 1998.

Received October 8, 1999