

A Short Course on Finite Model Theory

Jouko Väänänen

*Department of Mathematics
University of Helsinki
Finland*

Preface

These notes are based on lectures that I first gave at the *Summer School of Logic, Language and Information* in Lisbon in August 1993 and then in the department of mathematics of the University of Helsinki in September 1994. Because of the nature of the lectures, the notes are necessarily sketchy.

Contents

Lecture 1.	<i>Basic Results</i>	3
Lecture 2.	<i>Complexity Theory</i>	17
Lecture 3.	<i>Fixpoint Logic</i>	21
Lecture 4.	<i>Logic with k variables</i>	28
Lecture 5.	<i>Zero-One Laws</i>	35
	<i>Some sources</i>	43

Lecture 1. Basic Results

Classical logic on infinite structures arose from paradoxes of the infinite and from the desire to understand the infinite. Central constructions of classical logic yield infinite structures and most of model theory is based on methods that take infiniteness of structures for granted. In that context finite models are anomalies that deserve only marginal attention.

Finite model theory arose as an independent field of logic from consideration of problems in theoretical computer science. Basic concepts in this field are finite graphs, databases, computations etc. One of the underlying observations behind the interest in finite model theory is that many of the problems of complexity theory and database theory can be formulated as problems of mathematical logic, provided that we limit ourselves to finite structures.

While the objects of study in finite model theory are finite structures, it is often possible to make use of infinite structures in the proofs. We shall see examples of this in these lectures.

Notation

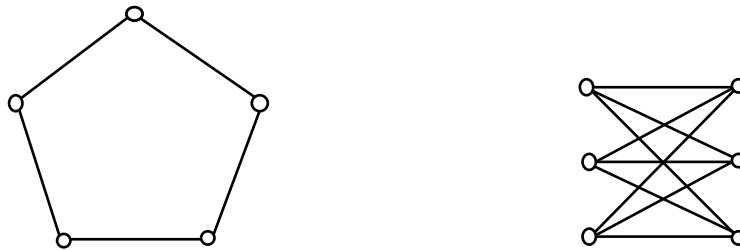
A *vocabulary* is a finite set of relation symbols R_1, \dots, R_n . Each relation symbol has a natural number as its *arity*. An m -ary relation symbol is denoted by $R(x_1, \dots, x_m)$. In some cases we add constant and function symbols to the vocabulary.

If $L = \{R_1, \dots, R_n\}$ is a vocabulary with R_i an m_i -ary relation symbol, then an L -*structure* is an $(n + 1)$ -tuple

$$\mathbf{A} = (A, R_1^{\mathbf{A}}, \dots, R_n^{\mathbf{A}}),$$

where A is a finite set called the *domain* or *universe* of \mathbf{A} , and $R_i^{\mathbf{A}}$ is an m_i -ary relation on A , called the *interpretation* of R_i in \mathbf{A} .

Typical examples of structures are *graphs*. In this case $L = \{E\}$. The binary symbol E denotes the edge-relation of the graph. Here are two well-known graphs:



We assume basic knowledge of *first order logic* FO (also denoted by $L_{\omega\omega}$). *Truth* -relation between structures \mathbf{A} and sentences $\varphi(a_1, \dots, a_n)$ with parameters a_1, \dots, a_n from A is written:

$$\mathbf{A} \models \varphi(a_1, \dots, a_n)$$

as usual. *Elementary equivalence* is denoted by $\mathbf{A} \equiv \mathbf{B}$, and $\mathbf{A} \equiv^n \mathbf{B}$ means that \mathbf{A} and \mathbf{B} satisfy the same sentences φ of FO of quantifier-rank $qr(\varphi) \leq n$.

A *model class* is a class of L -structures closed under \equiv^n , such as the class

$$Mod(\varphi) = \{ \mathbf{A} : \mathbf{A} \text{ is an } L\text{-structure and } \mathbf{A} \models \varphi \}.$$

A model class is *definable* if it is of the form $Mod(\varphi)$ for some $\varphi \in FO$. Examples of definable model classes are the class of all graphs, the class of all groups, the class of all equivalence-relations etc. A property of models is said to be *expressible* in FO (or some other logic) if it determines a definable model class. For example the property of a graph of being *complete* is expressible in FO by the sentence $\forall x \forall y (x \neq y \rightarrow xEy)$.

The *relativisation* $\varphi^{(R)}$ of a formula φ to a unary predicate $R(x)$ is defined by induction as follows

$$\begin{aligned}\varphi^{(R)} &= \varphi && \text{for atomic } \varphi, \\ (\neg \varphi)^{(R)} &= \neg \varphi^{(R)}, \\ (\varphi \wedge \psi)^{(R)} &= \varphi^{(R)} \wedge \psi^{(R)}, \\ (\exists x \varphi)^{(R)} &= \exists x (R(x) \wedge \varphi^{(R)}).\end{aligned}$$

The relativisation $\mathbf{A}^{(R)}$ of an L -structure \mathbf{A} is

$$\mathbf{A}^{(R)} = (R^{\mathbf{A}}, R_1^{(R^{\mathbf{A}})}, \dots, R_n^{(R^{\mathbf{A}})}).$$

The basic fact about relativisation is the equivalence

$$\mathbf{A} \models \varphi^{(R)} \text{ iff } \mathbf{A}^{(R)} \models \varphi.$$

We use $\mathbf{A}|_L$ to denote the *reduct* of the structure \mathbf{A} to the vocabulary L .

1.1 Failure of the Compactness Theorem. *There is a (recursive) set T of sentences so that every proper subset of T has a model but T itself has no models.*

Proof. $T = \{\neg \theta_n : n \in \mathbf{N}\}$, where θ_n says *there are exactly n elements in the universe*. Q.E.D.

Remark. T is universal-existential. Without constant symbols T could not be made universal or existential. With constant symbols, we could let T be

$$\{\neg c_n = c_m : n \neq m\}.$$

Then every finite subset of T has a model but T has no models.

It is only natural that the Compactness Theorem should fail, because its very idea is to generate examples of infinite models.

1.2 Example (An application of the Compactness Theorem in finite model theory). We show that there is no first order sentence φ of the empty vocabulary so that $\mathbf{A} \models \varphi$ iff $|A|$ even. Let $T = \{\varphi^{(P_1)}, \neg \varphi^{(P_2)}\} \cup \{ "P_i(x) \text{ has at least } n \text{ elements"} : n \in \mathbb{N}, i = 1, 2 \}$. Every finite subset of T has a finite model. Hence T has a model \mathbf{A} , which is infinite. Let $\mathbf{A}_1 \prec \mathbf{A}$ be countable. Then $\mathbf{A}_1 \models \varphi^{(P_1)} \wedge \neg \varphi^{(P_2)}$. Let

$$\mathbf{B}_1 = \mathbf{A}_1^{(P_1)} \quad \text{and} \quad \mathbf{B}_2 = \mathbf{A}_1^{(P_2)}.$$

Then $\mathbf{B}_1 \models \varphi$ and $\mathbf{B}_2 \models \neg \varphi$, a contradiction. Q.E.D.

An important phenomenon in finite model theory is that *individual* structures can be characterized up to isomorphism:

1.3 Proposition. (Characterization of finite structures up to isomorphism). For every (finite) \mathbf{A} there is a first order sentence $\theta_{\mathbf{A}}$ so that $\mathbf{B} \models \theta_{\mathbf{A}}$ iff $\mathbf{B} \cong \mathbf{A}$.

Proof. We assume, as always, that the vocabulary L of \mathbf{A} is finite. Let $A = \{a_1, \dots, a_n\}$. Let $\psi(x_1, \dots, x_n)$ be the conjunction of

$$\begin{aligned} &\varphi(x_1, \dots, x_n), \text{ where } \varphi(x_1, \dots, x_n) \text{ is atomic and } \mathbf{A} \models \varphi(a_1, \dots, a_n), \\ &\neg \varphi(x_1, \dots, x_n), \text{ where } \varphi(x_1, \dots, x_n) \text{ is atomic and } \mathbf{A} \not\models \varphi(a_1, \dots, a_n), \\ &x(x = x_1 \vee \dots \vee x = x_n). \end{aligned}$$

Let $\theta_{\mathbf{A}}$ be the sentence $\exists x_1 \dots \exists x_n \psi(x_1, \dots, x_n)$. Clearly $\mathbf{A} \models \theta_{\mathbf{A}}$. If $\mathbf{B} \models \theta_{\mathbf{A}}$, and $\mathbf{B} \models \psi(b_1, \dots, b_n)$, then $b_i \mapsto a_i$ is an isomorphism $\mathbf{B} \cong \mathbf{A}$. Q.E.D.

1.4 Corollary. For any (finite) structures \mathbf{A} and \mathbf{B} we have

$$\mathbf{A} \cong \mathbf{B} \text{ if and only if } \mathbf{A} \models \theta_{\mathbf{B}}. \quad \text{Q.E.D.}$$

It is more interesting to study *classes* of structures than individual structures in finite model theory. Also, it is important to pay attention to quantifier-rank and length of formulas. Note that θ_A above is bigger in size than even \mathbf{A} itself.

The Ehrenfeucht-Fraïssé game

The method of Ehrenfeucht-Fraïssé games is one of the few tools of model theory that survive the passage to finiteness. *The Ehrenfeucht-Fraïssé game* $EF_n(\mathbf{A}, \mathbf{B})$ between two structures \mathbf{A} and \mathbf{B} is defined as follows:

There are two players I and II who play n moves. Each move consists of choosing an element of one of the models:

I	x_1	x_2	\dots	x_n
II	y_1	y_2	\dots	y_n

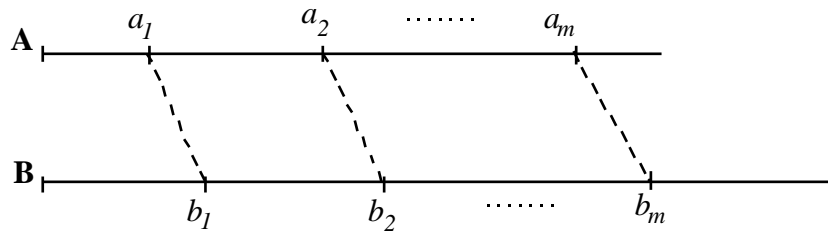
- Rules:
- 1) $x_i \in \mathbf{A}$ and $y_i \in \mathbf{B}$.
 - 2) $x_i \in \mathbf{B}$ and $y_i \in \mathbf{A}$.
 - 3) Player I wins if $\{x_i, y_i\}$ is a partial isomorphism between \mathbf{A} and \mathbf{B} .

1.5 Ehrenfeucht-Fraïssé Theorem. $\mathbf{A} \equiv_n \mathbf{B}$ iff $\mathbf{A} \equiv_n \mathbf{B}$. Q.E.D.

The following is a typical application of Ehrenfeucht-Fraïssé games in finite model theory:

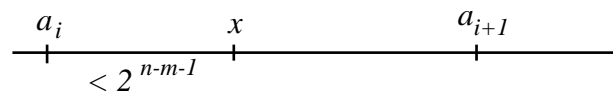
1.6 Proposition (Gurevich 1984). *Suppose \mathbf{A} and \mathbf{B} are linear orders of cardinality $< 2^n$. Then $\mathbf{A} \equiv_n \mathbf{B}$.*

Proof. We may assume **A** and **B** are intervals of the integers. Suppose $(a_1, b_1), \dots, (a_m, b_m)$ have been played:

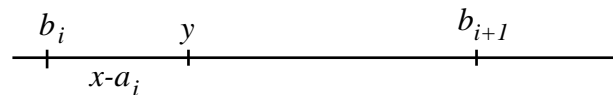


The *strategy* of II: Corresponding closed intervals have the same length unless both are $> 2^{n-m}$ (2^{n-m} , when $m \in \{0, n\}$). Suppose $x \in (a_i, a_{i+1})$.

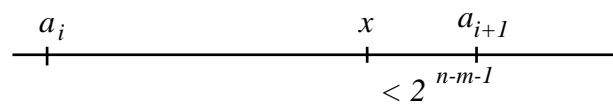
Case 1: x is near the low end of the interval.



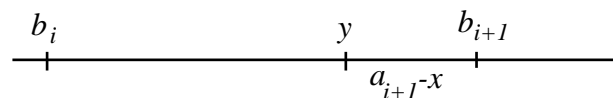
Player II plays y exactly as near to the low end of the interval.



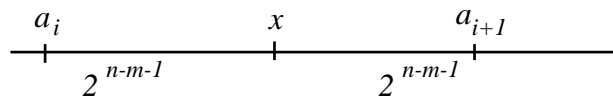
Case 2: x is near the high end of the interval.



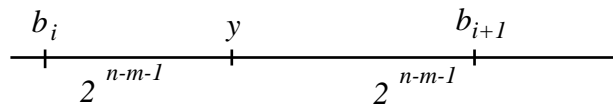
Player II plays y exactly as near to the high end of the interval.



Case 3: x is not near either end of the interval.



Player II plays y so that it is not near either end of the interval.



It should be clear that Player II can maintain this strategy for n moves.

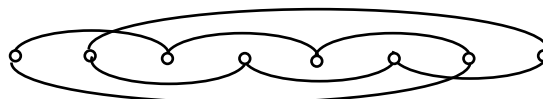
Q.E.D.

1.7 Applications (Gurevich 1984).

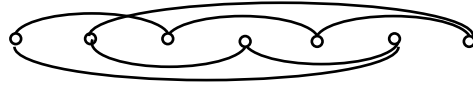
(1) The class of linear orders of even length is not first order definable. ("Even" can be replaced by almost anything.) Indeed, suppose φ defines linear orders of even length. Let $n = qr(\varphi)$. Let \mathbf{A} be a linear order of length 2^n and let \mathbf{B} be of length $2^n + 1$. By 1.6 $\mathbf{A} \equiv \mathbf{B}$. But $\mathbf{A} \models \varphi$ and $\mathbf{B} \not\models \varphi$.

(2) Let $L = \{<\}$. There is no first order L -formula $\varphi(x)$ so that for linear orders \mathbf{A} we have $\mathbf{A} \models \varphi(a)$ iff a is an even element in $<^{\mathbf{A}}$. Indeed, otherwise $\varphi(x) \wedge \exists y(y = x \wedge y < x)$ would contradict (1).

(3) Let $L = \{E\}$. There is no first order L -sentence φ so that for linear orders \mathbf{A} we have $\mathbf{A} \models \varphi$ iff \mathbf{A} is a connected graph. Let $L_1 = \{<\}$. Replace xEy in φ by "y is the successor of the successor of x, or else x is the last but one and y is the first element, or x is the last and y is the second element, and the same with x,y interchanged". Get an L_1 -sentence φ_1 . For a linear order \mathbf{A} , we have $\mathbf{A} \models \varphi_1$ iff the length of \mathbf{A} is odd, contrary to (1). The following two pictures clarify this fact:



even number of elements = not connected



odd number of elements = connected

1.8 Note. The applications of 1.7 can alternatively be obtained by the use of *ultraproducts*: Let \mathbf{A}_n be a linear order of length n . Let D be a non-principal ultrafilter on \mathbf{N} . Let $\mathbf{A} = \prod_D \mathbf{A}_{2n}$ and $\mathbf{B} = \prod_D \mathbf{A}_{2n+1}$. The order-type of \mathbf{A} (and of \mathbf{B}) is $\mathbf{N} + \mathbf{Z}^{\omega} + \mathbf{N}$. Hence $\mathbf{A} \equiv \mathbf{B}$. However, if $\mathbf{A}_n \models \varphi$ iff n is even, then $\mathbf{A} \models \varphi$ and $\mathbf{B} \not\models \varphi$.

Still another alternative is to use the compactness theorem as in 1.2.

1.9 Failure of Beth Definability Theorem (Hájek 1977). *There is a first order L -sentence φ , which defines a unary predicate P implicitly but not explicitly.*

Proof. φ is the conjunction of

$$\begin{aligned} & \text{"< is a linear order",} \\ & x(P(x) \rightarrow y(y = x \rightarrow x < y)), \\ & x \rightarrow y(\text{"y successor of x"} \rightarrow (P(y) \rightarrow \neg P(x))). \end{aligned}$$

Every finite linear order has a unique P with φ . However, if $\varphi \models P(x) \rightarrow \theta(x)$, where $\theta(x)$ is an $\{<\}$ -sentence, then $\theta(x)$ contradicts 1.7. Q.E.D.

Note. The proof gives failure of the so called *Weak Beth Property*, too.

1.10 Failure of Craig Interpolation Theorem (Hájek 1977). *Let $L = \{<, c\}$. There are an $L \setminus \{P\}$ -sentence φ_1 and an $L \setminus \{Q\}$ -sentence φ_2 so that $\varphi_1 \models \varphi_2$ but no L -sentence θ satisfies both $\varphi_1 \models \theta$ and $\theta \models \varphi_2$.*

Proof. Let φ be as in the proof of 1.9. Let φ be obtained from φ by replacing P by Q . Let φ_1 be $\varphi \rightarrow P(c)$ and let φ_2 be $\varphi \rightarrow Q(c)$. Now $\varphi_1 \models \varphi_2$.

Suppose $\varphi_1 \models \theta$ and $\theta \models \varphi_2$ where θ is an L -sentence. Then $\varphi \models P(c) \iff \theta$, whence φ defines P explicitly, contrary to 1.9. Q.E.D.

1.11 Failure of Los-Tarski Preservation Theorem (Tait 1959). *There is a sentence which is preserved by substructures but which is not equivalent to a universal sentence.*

Proof (here by Gurevich 1984). Let ψ be the conjunction of

- (1) " $<$ is a linear order",
- (2) "0 is the least element",
- (3) " $S(x, y)$ (y is the successor of x)
(x is the last element $\iff y = 0$)".

Let φ be the conjunction of (1)-(3) and

- (4) $\exists x \forall y S(x, y) \iff \exists x P(x)$.

Now φ is preserved by submodels, for if $\mathbf{A} \models \varphi$ and $\mathbf{B} \subseteq \mathbf{A}$ such that $\mathbf{B} \models \exists x \forall y S(x, y)$, then $\mathbf{B} = \mathbf{A}$ and hence $\mathbf{B} \models \varphi$. Suppose we have $\mathbf{A} \models \varphi \iff \exists x_1 \dots \exists x_n \theta(x_1, \dots, x_n)$. Let $\mathbf{A} = (\{1, \dots, n+2\}, <, S, \dots)$. Then \mathbf{A} satisfies $\psi \iff \exists x \forall y S(x, y)$, but $\mathbf{A} \not\models \varphi$. Hence there are $a_1, \dots, a_n \in \{1, \dots, n+2\}$ so that $\mathbf{A} \not\models \theta(a_1, \dots, a_n)$. Let $\mathbf{B} = (\{1, \dots, n+2\}, <, S, \{d\})$, where $d \neq a_1, \dots, a_n$. Then $\mathbf{B} \not\models \varphi$. On the other hand, $\mathbf{B} \models \psi \iff \exists x P(x)$, whence $\mathbf{B} \models \varphi$, a contradiction.

Q.E.D.

Note. Kevin Compton (unpublished) has proved that sentences which are preserved by substructures on finite models, are universal on finite models.

In conclusion, first order logic does not have such a special place in finite model theory as it has in classical model theory. But worst is still to come: the set of first order sentences that are valid in finite models, is not recursively enumerable. Hence there can be no *Completeness Theorem* in finite model theory.

Trakhtenbrot's Theorem

A *Turing machine* consists of a *tape*, a finite set of *states* and a finite set of *instructions*. The tape consists of numbered *cells* 1,2,3, ... Each cell contains 0 or 1. *States* are denoted by q_1, \dots, q_n (q_1 is called the *initial state*). *Instructions* are quadruples of one of the following kinds:

$q\alpha\beta q'$, meaning: If in state q reading α , then write β and go to state q' .

$q\alpha Rq'$, meaning: If in state q reading α , then move right and go to state q' .

$q\alpha Lq'$, meaning: If in state q reading α , then move left and go to state q' .

A *configuration* is a sequence $\alpha_1\alpha_2\dots\alpha_{i-1}q\alpha_i\dots\alpha_n$. Its meaning is that the machine is in state q reading α_i in cell i . The sequence $q_1\alpha_1\alpha_2\dots\alpha_n$ is called the *initial configuration* with *input* $\alpha_1\alpha_2\dots\alpha_n$. A *computation* is a sequence I_1, \dots, I_n of configurations so that I_1 is an initial configuration and I_{i+1} obtains from I_i by the application of an instruction. Computation *halts* if no instruction applies to the last configuration. M is *deterministic* (default) if for all q and α there is at most one instruction starting with $q\alpha$. Otherwise M is *non-deterministic*.

1.12 Trakhtenbrot's Theorem (1950). *The set of finitely valid first order sentences is not recursively enumerable.*

We prove this by reducing the *Halting Problem* to the problem of deciding whether a first order sentence has a finite model. With this in mind, let

$$L_0 = \{B_0, B_1, C, succ, 1, N, <\}.$$

We think of L_0 -structures in the following terms. Universe of the model is time, $succ^{(i)}(1)$ means q_i , $B_\alpha(x, t)$ means that cell x contains symbol α at time t , and $C(t, q, x)$ means that at time t machine is in state q reading cell x .

Suppose a machine M is given. We construct a sentence ϕ_M so that

M halts iff φ_M has a model.

This will be the desired reduction of the Halting Problem to the problem under consideration.

Let φ_M be the conjunction of:

- (1) $C(1,1,1)$, meaning: initial state is q_1 ,
- (2) $\forall x (B_0(x,1) \rightarrow \forall t (B_0(x,t) \rightarrow \neg B_1(x,t)))$, meaning: initially tape is blank, and later it contains zeros and ones.
- (3) Axioms of successor function $succ$ and the constants 1 (the first element) and N (the last element) in terms of the linear order $<$,

(4) For every instruction $q_i \alpha \beta q_j$ of M :

$$\forall x \forall t [(C(t, q_i, x) \rightarrow B_\alpha(x, t)) \rightarrow (\exists t' (t' < N \wedge C(t', q_j, x) \rightarrow B_\beta(x, t' + 1)) \wedge \forall y (x < y \rightarrow (B_0(y, t' + 1) \rightarrow B_0(y, t'))))],$$

(5) For every instruction $q_i \alpha R q_j$ of M :

$$\forall x \forall t [(C(t, q_i, x) \rightarrow B_\alpha(x, t)) \rightarrow (\exists t' (t' < N \wedge C(t', q_j, x + 1) \rightarrow B_0(x, t' + 1)) \wedge \forall y (y < x \rightarrow (B_0(y, t') \rightarrow B_0(y, t' + 1)))]],$$

(6) For every instruction $q_i \alpha L q_j$ of M :

$$\forall x \forall t [(C(t, q_i, x + 1) \rightarrow B_\alpha(x + 1, t)) \rightarrow (\exists t' (t' < N \wedge C(t', q_j, x) \rightarrow B_0(x, t' + 1)) \wedge \forall y (y < x \rightarrow (B_0(y, t') \rightarrow B_0(y, t' + 1)))]].$$

Claim. M halts $\iff \varphi_M$ has a model.

Suppose the computation of M is I_1, I_2, \dots, I_k . We define a model of φ_M as follows

$$\begin{aligned} A &= \{1, 2, \dots, k\}, \\ B_\alpha^A(x, t) &\text{ follows the configuration } I_t, \\ C^A(q, t, x) &\text{ follows } I_t. \end{aligned}$$

It is easy to show $\mathbf{A} \models (1)-(6)$.

Claim. φ_M has a model M halts.

Suppose $\mathbf{A} \models \varphi_M$, $|A|=k$. Suppose the computation of M were infinite $I_1, I_2, \dots, I_k, I_{k+1}, \dots$. By following (1)-(6) with I_1, I_2, \dots, I_k one eventually arrives at the contradiction $k < k$.

Let

$$\text{Val} = \{ \#(\varphi) : \varphi \text{ a finitely valid } L \text{-sentence} \},$$

$$\text{Sat} = \{ \#(\varphi) : \varphi \text{ an } L \text{-sentence with a finite model} \}$$

where $\#(\varphi)$ is the Gödel number of φ . We have proved that M halts iff $\#(\varphi_M) \in \text{Sat}$. Thus we have a reduction of the *Halting Problem of Turing Machines* to the problem of *Finite Satisfiability* of L_0 -sentences. Since the Halting Problem is undecidable, so is Sat. Since Sat is trivially recursively enumerable its "complement" Val cannot be recursively enumerable. Q.E.D.

1.13 Consequences of Trakhtenbrot's Theorem.

(1) *There is no effective axiomatic system S so that φ valid $\iff \varphi$ provable in S . This means the total failure of the Completeness Theorem.*

(2) *There is no recursive function f so that if a first order sentence φ has a model, then it has a model of size $f(\varphi)$. This means the failure of the Downward Löwenheim-Skolem Theorem.*

Coding finite structures into words

Suppose $\mathbf{A} = (\{a_1, \dots, a_n\}, R_1, \dots, R_m)$, where R_i n_i -ary. The *code* of \mathbf{A} , $C(\mathbf{A})$, is the binary string

$$\mathbf{A}_0 \# \mathbf{A}_1 \# \dots \# \mathbf{A}_m,$$

where

$\mathbf{A}_0 = n$ in binary,

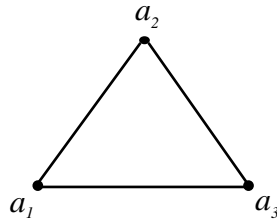
= a new symbol (w.l.o.g.),

\mathbf{A}_i = binary sequence of length n^{n_i} coding R_i .

Note that the length of $C(\mathbf{A})$ is $\log(n) + \sum_{i=1}^m n^{n_i} + m$.

Example. Here is an example of a structure and its code:

Structure \mathbf{A} :



Code $C(\mathbf{A}) = 11\#011101110$

A model class K is *recursive* if the language $\{C(\mathbf{A}) : \mathbf{A} \in K\}$ in the alphabet $\{0,1\}$ is, i.e. if there is a machine which on input $C(\mathbf{A})$ gives output 1 if $\mathbf{A} \in K$ and 0 if $\mathbf{A} \notin K$. K is *recursively enumerable* if there is a machine which on input $C(\mathbf{A})$ halts iff $\mathbf{A} \in K$.

Recall that

M halts iff φ_M has a model.

It is easy (but tedious) to modify the L_0 -sentence φ_M to an $(L_0 \cup L)$ -sentence φ_M with a unary predicate P so that

M halts on input $C(\mathbf{A})$ iff φ_M has a model \mathbf{B} with $\mathbf{B}^{(P)}|_L = \mathbf{A}$,

where L is the vocabulary of \mathbf{A} . A model class K is a *relativized pseudo-elementary class* (or RPC) if there are a sentence φ , and a predicate P such that

$$K = \{\mathbf{A}^{(P)}|_L : \mathbf{A} \models \varphi\}.$$

It is a consequence of the proof of 1.12 that:

1.14 Theorem. *A model class is RPC iff it is recursively enumerable.*

1.15 Corollary. *A model class is RPC co-RPC iff it is recursive.*

Since there are disjoint recursively enumerable sets that cannot be separated by a recursive set, we have:

1.16 Corollary. *There are disjoint RPC-classes that cannot be separated by any RPC co-RPC-class.*

We may conclude that to extend *FO* to a logic with the so called *Many-Sorted Interpolation Theorem*, one has to go beyond recursive model classes. Note that on infinite structures $RPC \text{ co-RPC} = FO$, a consequence of the many-sorted interpolation property.

Lecture 2. Complexity Theory

Let us fix some states of the Turing machine as *accepting states*. Then M *accepts* an input $\alpha_1 \dots \alpha_n$, if there is a computation I_1, \dots, I_k so that $I_1 = q_1 \alpha_1 \dots \alpha_n$ and I_k is accepting. The *time* of the computation I_1, \dots, I_k is k , so that each instruction is thought to take one unit of time. The *space* of I_1, \dots, I_k is the maximum length of I_i . Note, that the space of a computation is always bounded by the length of the input plus the time. A machine M is *polynomial time* (or *polynomial space*) if there is a polynomial $P(x)$ so that for all inputs $\alpha_1 \dots \alpha_n$ the time (or the space) of the computation of M is bounded by $P(n)$. Model class K is *polynomial time* (or *polynomial space*) if there is a polynomial time (or space) machine M that accepts $C(\mathbf{A})$ iff $\mathbf{A} \in K$. We use PTIME (or simply P) and PSPACE to denote the families of polynomial time and, respectively, polynomial space model classes.

PTIME and PSPACE are examples of *complexity classes*. Another important complexity class, LOGSPACE, allows the machine to use $\log(n)$ cells on input of length n , when reading input or writing output is not counted as a use of cells.

A machine M is *non-deterministic polynomial time* if for some polynomial $P(x)$ it is true that, whenever M accepts an input $\alpha_1 \dots \alpha_n$, this happens in time $P(n)$. Here M may be non-deterministic. A model class K is *non-deterministic polynomial time* if there is M as above so that

$$M \text{ accepts } C(\mathbf{A}) \text{ iff } \mathbf{A} \in K.$$

We use NPTIME (or just NP) to denote the family of all non-deterministic polynomial time model classes.

2.1 Fact. $\text{LOGSPACE} \subseteq P \subseteq NP \subseteq \text{PSPACE}$, $\text{LOGSPACE} \subseteq \text{PSPACE}$.

2.2 Open problem. Is it true, that $P \neq NP$? Or is any other inclusion in 2.1 a proper one?

2.3 Proposition. *First order logic is contained in LOGSPACE.*

Proof (by example). We decide $\mathbf{A} \models \exists x \exists y R(x,y)$ in *LOGSPACE*. A counter which runs through numbers from 1 to n takes space $\log(n)$. We need a counter for x and counter for y . A double loop scans through pairs (x,y) and looks up in $C(\mathbf{A})$ whether $R(x,y)$ holds or not. Q.E.D.

A model class is *existential second order definable*, or Σ_1^1 , if it can be defined with a formula $R_1 \dots R_m \varphi$, where φ is FO.

2.4. Fagin's Theorem (1974). $NP = \Sigma_1^1$.

Proof. 1° Easy part: $\Sigma_1^1 \subseteq NP$

Non-determinism allows guessing. For example, the Turing machine:

$$\begin{array}{l} q_1 00 q_2 \\ q_1 01 q_2 \end{array}$$

"guesses" a value for the first cell. If φ is $R_1 \dots R_m \psi$, a non-deterministic program can "guess" values for the matrices of the relations $R_1 \dots R_m$ and then check ψ in *LOGSPACE*.

2° Hard part: $NP \subseteq \Sigma_1^1$.

Suppose that M is a non-deterministic machine. Let

$$K = \{\mathbf{A} : M \text{ accepts } C(\mathbf{A})\}.$$

Choose k so that if M accepts $\alpha_1 \dots \alpha_n$, it happens in time n^k . We use k -sequences of elements of the model to measure time. Given \mathbf{A} , where $A = \{1, \dots, n\}$, we use the sequence

$$\underbrace{(1,1,\dots,1), \dots, (n,n,\dots,n)}_{n^k}.$$

as a model for tape and time. Otherwise we imitate Trakhtenbrot's Theorem. Let $L_1 = L \{B_0(\bar{x}, \bar{t}), B_1(\bar{x}, \bar{t}), C(\bar{t}, q, \bar{x}), succ(\bar{t}, \bar{t}'), \prec, 1, \mathbf{N}\}$, where this time we have $\bar{x} = x_1, \dots, x_k$, $\bar{t} = t_1, \dots, t_k$, B_0 and B_1 are $k+k$ -ary, \prec is $k+k$ -ary and C is $k+1+k$ -ary. By imitating the definition of φ_M we get φ_M so that

M accepts $C(\mathbf{A})$ iff
 there are $B_0, B_1, C, succ, 1, N, \prec$ so that $(\mathbf{A}, B_0, B_1, C, succ, \prec, 1, N) \models \varphi_M$ iff
 $\mathbf{A} \models B_0, B_1, C, \dots \varphi_M$. Q.E.D.

2.5 Corollaries (Fagin 1974).

- (1) $\begin{matrix} 1 \\ 1 \end{matrix} = NP$ co-NP , where $\begin{matrix} 1 \\ 1 \end{matrix} = \text{co-NP}$ (Note: on all structures $\begin{matrix} 1 \\ 1 \end{matrix} = FO$).
- (2) $\begin{matrix} 1 \\ 1 \end{matrix} NP$ co-NP P NP .
- (3) $\begin{matrix} 1 \\ 1 \end{matrix} = \begin{matrix} 1 \\ 1 \end{matrix}$ 3-colorability is $\begin{matrix} 1 \\ 1 \end{matrix}$ (The same holds for Hamiltonicity. Note that if 3-col were $\begin{matrix} 1 \\ 1 \end{matrix}$, then 3-col would be co-NP, whence NP co-NP and so $\begin{matrix} 1 \\ 1 \end{matrix} = \begin{matrix} 1 \\ 1 \end{matrix}$).
- (4) To prove P NP it is enough to construct a 3-colorable graph G so that whenever $n \in \mathbf{N}$ and R_1, \dots, R_m are relations on G , then there are a non-3-colorable H and P_1, \dots, P_m so that

$$(G, R_1, \dots, R_m) \not\models (H, P_1, \dots, P_m).$$

So one approach to proving P NP is to be good in Ehrenfeucht–Fraïssé games!

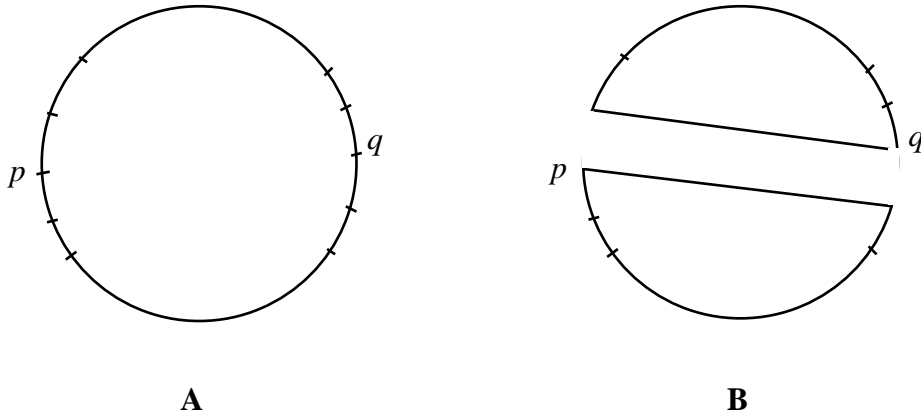
- (5) X is a spectrum iff X is NEXPTIME (= non-deterministic exponential time).

2.6 Theorem (Fagin 75, Hajék 75). Connectivity of graphs is not monadic $\begin{matrix} 1 \\ 1 \end{matrix}$ (i.e. $P_1 \dots P_m \varphi$, where P_1, \dots, P_m are unary.)

Proof (idea only). Suppose $P_1 \dots P_m \varphi$ is a monadic $\frac{1}{1}$ sentence of length k expressing connectedness. Take a large cycle \mathbf{A} of n nodes. Let unary P_1, \dots, P_m be given on \mathbf{A} so that φ . We may color elements of \mathbf{A} according to which predicates P_1, \dots, P_m the element satisfies. Two elements are called *similar* if they have the same color, and their corresponding close neighbours have also the same colors. Take two nodes p and q that are similar and far apart. Then brake \mathbf{A} into a disconnected graph \mathbf{B} as in the picture below. Then use similarity and the fact that p and q are far apart to prove:

$$(\mathbf{A}, P_1, \dots, P_m) \models^k (\mathbf{B}, P_1, \dots, P_m).$$

It follows that \mathbf{B} models $P_1 \dots P_m \varphi$, contradicting non-connectedness of \mathbf{B} .



2.7 Corollary. *Monadic $\frac{1}{1}$ is not closed under complementation .*

Proof. *Non-connectedness is monadic $\frac{1}{1}$.*

Q.E.D.

Lecture 3. Fixpoint Logic

Suppose L is a vocabulary and S is a k -ary predicate symbol no in L . Let $\varphi(\vec{x}, S)$ be a first order formula in which S is positive and $\vec{x} = x_1, \dots, x_k$. Let \mathbf{A} be an L -structure. Define $S^0 = \emptyset$ and $S^{i+1} = \{\vec{a} : \mathbf{A} \models \varphi(\vec{a}, S^i)\}$. Clearly

$$S^0 \subseteq S^1 \subseteq \dots \subseteq S^i \subseteq \dots \subseteq A^k.$$

Eventually $S^{i+1} = S^i$. We denote this S^i by S and call it the *fixpoint* of $\varphi(\vec{x}, S)$ on \mathbf{A} . The set S is called a fixpoint, because for all \vec{a}

$$\mathbf{A} \models S(\vec{a}) \iff \varphi(\vec{a}, S).$$

Example 1. $\varphi(x, y, S) \equiv xEy \wedge \forall z(xEz \rightarrow S(z, y))$. $L = \{E\}$. If G is a graph, and we compute S on G , we get the set of pairs (x, y) for which there is a path from x to y . Thus $x \rightarrow y \iff S(x, y)$ says the graph is connected.

Example 2. $\varphi(x, S) \equiv succ(1, x) \wedge \forall y(z(S(y) \wedge succ(y, z) \wedge succ(z, x)))$. In this case $L = \{succ, 1\}$. If \mathbf{A} is a model of the vocabulary $\{succ, 1\}$ where $succ^{\mathbf{A}}$ is a successor relation with $1^{\mathbf{A}}$ as the least element and the successor of the last element being $1^{\mathbf{A}}$, then S is the set of "even" elements on \mathbf{A} , and $x \rightarrow S(x)$ says " \mathbf{A} has even cardinality".

Fixpoints are examples of *global relations*. A global relation (or a *query*, as they are also called) associates with every structure \mathbf{A} a k -ary relation $R(x_1, \dots, x_k)$ on A . As an example, the fixed point of $\varphi(\vec{x}, S)$ is a global relation which associates with every \mathbf{A} the relation S .

We shall now introduce an elaboration of the fixpoint concept. Rather than looking for the fixpoint of a single formula, we take the *simultaneous fixpoint* of a system of several formulas. Suppose $\varphi_1(\vec{x}, S, R)$ and $\varphi_2(\vec{x}, S, R)$ are positive in S and R . Then we may consider the following simultaneous induction:

$$\begin{aligned}
S^0 &= \text{ , } \\
R^0 &= \text{ , } \\
S^{i+1} &= \{ \bar{a} : \mathbf{A} \models \varphi_1(\bar{a}, S^i, R^i) \} , \\
R^{i+1} &= \{ \bar{a} : \mathbf{A} \models \varphi_2(\bar{a}, S^i, R^i) \} .
\end{aligned}$$

For some i we have $S^{i+1} = S^i$ and $R^{i+1} = R^i$. Then we denote $S = S^i$, $R = R^i$. The pair (S, R) is called the *simultaneous fixpoint* of the formulas $\varphi_1(\bar{x}, S, R)$ and $\varphi_2(\bar{x}, S, R)$, because of the identities:

$$\begin{aligned}
\mathbf{A} \models S(\bar{a}) &\iff \varphi_1(\bar{a}, S, R) , \\
\mathbf{A} \models R(\bar{a}) &\iff \varphi_2(\bar{a}, S, R) .
\end{aligned}$$

The relations S and R are called *multiple fixpoints* because they are members of a simultaneous fixpoint. Naturally, we can do the same for more than two formulas.

3.2 Example. If S and R are fixpoints, then (S, R) is a multiple fixpoint. Indeed, suppose S is the fixpoint of $\varphi_1(\bar{x}, S)$ and R is the fixpoint of $\varphi_2(\bar{x}, R)$. Consider the formulas:

$$(*) \quad \begin{aligned}
\psi_1(\bar{x}, S, R, T) &\iff \varphi_1(\bar{x}, S), \\
\psi_2(\bar{x}, S, R, T) &\iff \varphi_2(\bar{x}, R), \\
\psi_3(\bar{x}, S, R, T) &\iff S(\bar{x}) \wedge R(\bar{x}).
\end{aligned}$$

Let (S, R, T) be the simultaneous fixpoint of $(*)$. Clearly, S and R are the fixpoints we started with, and

$$T = S \wedge R .$$

Hence (S, R) is a multiple fixpoint.

3.3 Definition. *Fixpoint logic FP consists of global relations defined as follows: Suppose*

$$\varphi_i(\bar{x}, S_1, \dots, S_k), \quad i = 1, \dots, k$$

are first order formulas positive in S_1, \dots, S_k . Let (S_1, \dots, S_k) be the simultaneous fixpoint of $\varphi_i(\bar{x}, S_1, \dots, S_k)$ i.e. in any \mathbf{A} :

$$\begin{aligned} S_1(\bar{a}) &= \varphi_1(\bar{a}, S_1, \dots, S_k), \\ &\vdots \\ S_k(\bar{a}) &= \varphi_k(\bar{a}, S_1, \dots, S_k). \end{aligned}$$

Then the multiple fixpoints S_1, \dots, S_k are, by definition, in fixpoint logic.

3.4. Proposition.

- (1) $FO = FP$.
- (2) $FP = PTIME$.
- (3) FP is closed under $\exists, \forall, \neg, \rightarrow$, substitution and relativisation.

Proof. (1) Every $\varphi(\bar{x}) \in FO$ gives rise to a trivial fixpoint $S(\bar{a}) = \varphi(\bar{a})$.

(2) Suppose S is the fixpoint of $\varphi(\bar{x}, S)$ (for simplicity). The following polynomial time algorithm decides $\bar{x} \in S$:

Step 1: Decide $\varphi(\bar{a}, S)$ for each $\bar{a} \in A^k$.
Let S^1 be the set of \bar{a} for which it is true.

Step 2: Decide $\varphi(\bar{a}, S^1)$ for each $\bar{a} \in A^k$.
Let S^2 be the set of \bar{a} for which it is true.

etc until $S^{i+1} = S^i$. There are at most $|A|^k$ steps.

(3) Easy.

Q.E.D.

3.5. Theorem (Immerman 1986, Vardi 1982). *On ordered structures $FP = PTIME$.*

Proof. We imitate the proofs of Fagin's and Trakhtenbrot's theorems. Let M be a polynomial time deterministic machine. Choose k such that on input $\alpha_1 \dots \alpha_n$ machine halts in time n^k . Let $<$ be the assumed order. It gives rise to the lexicographic order of k -tuples. We express the predicates B_0, B_1, C of Fagin's theorem as fixpoints of a set of equations. For example, if M has only the instructions

$$q_1 01q', q_2 10q'$$

ending with q' , one equation is

$$C(\bar{t} + 1, q', \bar{x}) = (B_0(\bar{x}, \bar{t}) \wedge C(\bar{t}, q_1, \bar{x})) \vee (B_1(\bar{x}, \bar{t}) \wedge C(\bar{t}, q_2, \bar{x})).$$

We copy the action of the instructions into such equations. Let q_a be the accepting state of M . Then M accepts $C(\mathbf{A})$ iff $\mathbf{A} \models \bar{t} \bar{x} C(\bar{t}, q_a, \bar{x})$. Q.E.D.

3.6 Corollary. $P=NP$ iff $FP = \frac{1}{1}$ on ordered structures
iff $FP = \frac{1}{1}$ on ordered structures.

Note. On countably infinite acceptable structures $FP = \frac{1}{1}$ (Moschovakis 1974).

3.7 Definition. Suppose $\varphi(\bar{x}, S)$ is positive in S . On a structure \mathbf{A} we define

$$\begin{aligned} |\varphi| &= \text{least } i \text{ such that } S^{i+1} = S, \\ |\bar{a}| &= \text{least } i \text{ such that } \bar{a} \in S^i, \text{ if } \bar{a} \in S \\ &= |\varphi| + 1, \text{ otherwise.} \end{aligned}$$

3.8 Stage-Comparison Theorem (Moschovakis 1974). Suppose $\varphi(\bar{x}, S)$ is positive in S . Then the global relations

$$\begin{aligned} \bar{x} \prec^\varphi \bar{y} & \iff |\bar{x}| < |\bar{y}|, \\ \bar{x} \equiv^\varphi \bar{y} & \iff |\bar{x}| = |\bar{y}| \wedge \bar{x} \in S \end{aligned}$$

and their complements $\prec^\varphi, /^\varphi$ are in FP.

Proof (Here by Leivant 1990). Define also

$$\bar{x} \prec_1^\varphi \bar{y} \quad |\bar{x}| + 1 = |\bar{y}|.$$

The following equations define $\prec, \succ, \prec_1, \not\prec, /$ as multiple fixpoints. We use the shorthand notation $\prec y$ for the set $\{u: u \prec y\}$.

- (1) $x \prec y \quad z(x \prec_1 y),$
- (2) $x \succ y \quad \varphi(x, \prec y),$ meaning " $x \in S^{|x|}$ ",
- (3) $x \prec_1 y \quad \varphi(x, \prec x),$ meaning " $x \in S$ "
 $\neg \varphi(y, \neg \not\prec x),$ meaning " $y \in S^{|x|}$ "
 $[\varphi(y, \prec x) \wedge z(\varphi(z, \neg / x) \wedge \varphi(z, \prec x))],$ meaning
" $y \in S^{|x|+1}$ ", or " $|x|$ is the last stage".
- (4) $x \not\prec y \quad z(x / z \prec_1 y \wedge \varphi(y, \prec) \wedge z \neg \varphi(z, \prec)),$ meaning
" $y \in S^{|x|}$ " or " $S = \emptyset$ ",
- (5) $x / y \quad \neg \varphi(x, \neg \not\prec y),$ meaning " $x \in S^{|y|}$ ".

Once we have the fixpoint $(\prec, \succ, \prec_1, \not\prec, /)$ of these equations it is not hard to prove

$$(\prec, \succ, \prec_1, \not\prec, /) = (\prec^\varphi, \succ^\varphi, \prec_1^\varphi, \not\prec^\varphi, /^\varphi).$$

Hence $\prec^\varphi, \succ^\varphi$ and their complements are in FP. Q.E.D.

3.9 Corollary (Immerman 1982). *FP is closed under negation.*

Proof. The claim follows from the Stage-Comparison Theorem 3.8 and the equivalence $x \in S$ iff $x \in /^\varphi x$. Q.E.D.

3.10 Examples. The following global predicates are in *FP*:

"the graph is non-connected",
 "there is no path from x to y ".

Note. Immerman used the argument behind 3.9 to prove the surprising result that the class of context sensitive languages is closed under complements.

We defined *FP* using formulas $\varphi(\bar{x}, S)$ with S positive. A formula $\varphi(\bar{x}, S)$ is *monotone in S* if $\varphi(\bar{x}, S) \wedge S \subseteq S'$ implies always $\varphi(\bar{x}, S')$. The definition of fixpoint logic can be repeated with monotone formulas instead of positive formulas.

Facts. (1) If a formula is positive in S , then it is also monotone in S .

(2) There is a formula which is monotone in S , but which is not equivalent to a formula that is positive in S (Ajtai-Gurevich 1988).

(3) Fixpoints of monotone formulas are still in *FP*.

Thus monotone fixpoints bring nothing new, on the contrary, Gurevich (1984) showed that we cannot effectively decide whether a formula is monotone or not. This is in sharp contrast to positivity which is trivial to check. If $\varphi(\bar{x}, S)$ is not even monotone, we can still define *inflationary* fixpoints as follows:

$$S^0 = \emptyset, \\ S^{i+1} = S^i \cup \{ \bar{a} : \mathbf{A} \models \varphi(\bar{a}, S^i) \}.$$

Fact. Inflationary fixpoints of first order formulas are in *FP* (Gurevich-Shelah 1986). This follows also from the proof of *the Stage-Comparison Theorem*.

Finally, with any $\varphi(\bar{x}, S)$ we may try the following iteration:

$$\begin{aligned}
S^0 &= \emptyset, \\
S^{i+1} &= \{ \bar{a} : \mathbf{A} \models \varphi(\bar{a}, S^i) \}, \\
S &= S^{i_0} \text{ if there is } i_0 \text{ so that } S^{i_0} = S^{i_0+1} \\
&\quad, \text{ otherwise.}
\end{aligned}$$

We call S the *partial fixpoint* of $\varphi(\bar{x}, S)$. Partial fixpoints give rise to *partial fixpoint logic PFP*.

Facts. (1) $PFP = PSPACE$.

(2) $PFP = PSPACE$ on ordered structures (Vardi 1982).

(3) $FP = PFP = PTIME = PSPACE$ (Abiteboul-Vianu 1991).

Lecture 4. Logic with k variables

A formula can be measured by its:

- length,
- quantifier-rank, or
- number of variables.

In this chapter we focus on the last possibility. By reusing variables where possible one can write interesting and important formulas with very few variables. This corresponds to the programming rule that one should not reserve new memory every time one needs space but rather use the same working area over and over again.

4.1 Example. *Models with a total order \prec .* We show that the property " x has exactly i predecessors", which one normally would write with $i+1$ variables, is in fact definable with just 3 variables:

$$\begin{aligned} \varphi^1(x_1) & \quad x_2(x_1 \prec x_2 \quad x_1 = x_2) \\ \varphi^{i+1}(x_1) & \quad x_2(x_2 \prec x_1 \quad x_3(x_2 \prec x_3 \\ & \quad (x_1 \prec x_3 \quad x_1 = x_3)) \quad x_1(x_1 = x_2 \quad \varphi^i(x_1))). \end{aligned}$$

Now $\varphi^{i+1}(x_1)$ says " x_1 has i predecessors".

4.2 Definition. *First order logic with k variables, FO^k , is defined like FO but only k distinct variables are allowed in any formula. Infinitary logic with k variables, L^k , is defined similarly as FO^k but the infinite disjunction and conjunction are added to its logical operations. (L^k_ω is commonly used for L^k). Furthermore, L^ω is the union $\bigcup_k L^k$.*

4.3 Theorem (Barwise 1977). $FP = L^\omega$.

Proof. Recall that $S(\bar{x}) = S^1(\bar{x}) = S^2(\bar{x}) = S^3(\bar{x}) = \dots$. In each model the disjunction is finite, but its length may change from model to model. Let $\varphi(\bar{x}, S)$

FO^l have S positive, where S is k -ary and $\vec{x} = x_1, \dots, x_k$. Let y_1, \dots, y_k be new variables and

$$\begin{aligned} \Psi^0(\vec{x}) &= \neg x_1 = x_1 \\ \Psi^{i+1}(\vec{x}) &= \text{the result of replacing in } \Phi(\vec{x}, S) \text{ every occurrence of } S(t_1, \dots, t_k) \text{ by} \\ & y_1 \dots y_k (y_1 = t_1 \dots y_k = t_k \\ & x_1 \dots x_k (y_1 = x_1 \dots y_k = x_k \quad \Psi^i(x_1, \dots, x_k))). \end{aligned}$$

Each $\Psi^i(\vec{x})$ has only $l+k$ variables. Moreover $S^i(\vec{x}) = \Psi^i(\vec{x})$. Hence $S(\vec{x}) = \bigwedge_i \Psi^i(\vec{x}) \in L^{l+k}$. Q.E.D.

4.4 Remark. L^ω can express non-recursive properties, hence $L^\omega \neq PTIME$ and $L^\omega \neq FP$.

Proof. Let us work on ordered models: Recall from Example 4.1 the equivalence $\Phi^{i+1}(x_1) \equiv "x_1 \text{ has } i \text{ predecessors}"$. Let $A \subseteq N$ be non-recursive. Recall also that $\Phi^{i+1}(x_1) \in FO^3$. The sentence

$$\bigwedge_{i \in A} x_1 (\Phi^i(x_1) \wedge x_2 (x_2 \prec x_1 \wedge x_2 = x_1))$$

in L^3 is true in a total order of length n iff $n \in A$. Hence it defines a non-recursive model class. Q.E.D.

4.5 Remark. Every class of ordered structures is definable in L^ω .

Proof. Remark 4.4 shows every element of an ordered structure is definable in FO^3 . Every ordered structure is definable in FO^{m+3} where m depends on the vocabulary. Disjunctions of such definitions give all classes of ordered models.

Q.E.D.

This explains why L^ω is only discussed in connection with unordered structures. The advantage of L^ω over FP is that elementary equivalence relative to L^k has a

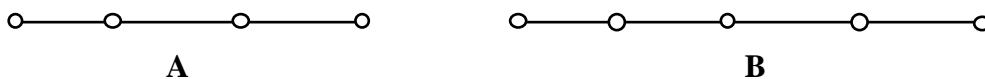
nice criterion. Using this criterion it is possible to show that certain concepts are undefinable in L^0 and hence undefinable in FP .

4.6 Definition. Let \mathbf{A} and \mathbf{B} be L -models. The k -pebble game on \mathbf{A} and \mathbf{B} , $G^k(\mathbf{A}, \mathbf{B})$, has the following rules:

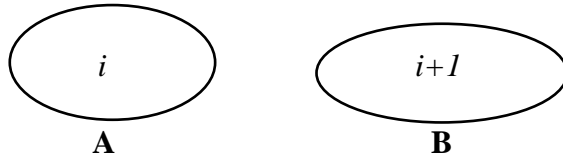
- (1) Players are I and II , and they share k pairs of pebbles. Pebbles in a pair are said to correspond to each other.
- (2) Player I starts by putting a pebble on an element of \mathbf{A} (or \mathbf{B}).
- (3) Whenever I has moved by putting a pebble on an element of \mathbf{A} (or \mathbf{B}), player II takes the corresponding pebble and puts it on an element of \mathbf{B} (or \mathbf{A}).
- (4) Whenever II has moved, player I takes a pebble - either from among the until now unused pebbles or from one of the models - and puts it on an element of \mathbf{A} (or \mathbf{B}).
- (5) This game never ends.
- (6) II wins if at all times the relation determined by the pairs of corresponding pebbles is a partial isomorphism. Otherwise I wins.

4.7 Examples.

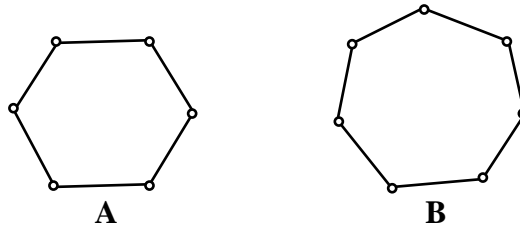
- (1) Let \mathbf{A} and \mathbf{B} be chains of different length. Then II has a winning strategy in $G^2(\mathbf{A}, \mathbf{B})$ but I has in $G^3(\mathbf{A}, \mathbf{B})$.



- (2) Let $L = \{$=$\}$. Let \mathbf{A} be a set with i elements and \mathbf{B} a set with $i + 1$ elements. Then II has a winning strategy in $G^i(\mathbf{A}, \mathbf{B})$ but I has in $G^{i+1}(\mathbf{A}, \mathbf{B})$.



(3) Let \mathbf{A} be a cycle of length n and \mathbf{B} a cycle of length $m \leq n$. II has a winning strategy in $G^2(\mathbf{A}, \mathbf{B})$ but I has in $G^3(\mathbf{A}, \mathbf{B})$.



4.8 Theorem (Barwise 1977, Immermann 1982). *Let \mathbf{A} and \mathbf{B} be L -structures. The following are equivalent*

- (1) $\mathbf{A} \equiv_{FO^k} \mathbf{B}$,
- (2) $\mathbf{A} \equiv_{L^k} \mathbf{B}$,
- (3) II has a winning strategy in $G^k(\mathbf{A}, \mathbf{B})$
(denote this by $\mathbf{A} \equiv_k \mathbf{B}$).

Proof. (1) \Rightarrow (3) Strategy of II : If the pairs (a_i, b_i) , $i = 1, \dots, l$ have been pebbled so far, then for all $\varphi(x_1, \dots, x_l) \in FO^k$ we have

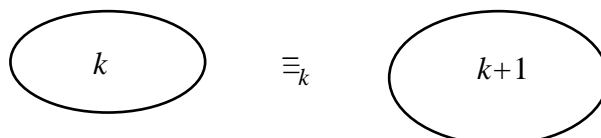
$$(*) \mathbf{A} \models \varphi(a_1, \dots, a_l) \text{ if and only if } \mathbf{B} \models \varphi(b_1, \dots, b_l).$$

This condition holds in the beginning ($l=0$) and it can be seen that Player II can maintain it.

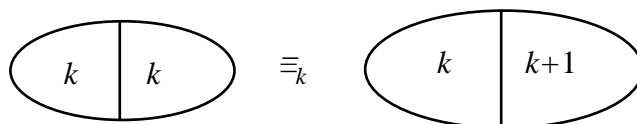
(3) \Rightarrow (2) One uses induction on $\varphi(x_1, \dots, x_l) \in FO^k$ to prove : If II plays his winning strategy and a position (a_i, b_i) , $i = 1, \dots, l$ appears in the game, then condition $(*)$ holds again. Finally, for a sentence φ we let $l=0$. Q.E.D.

4.9 Applications. The following properties of finite structures are not expressible in L^0 and hence not in FP . In each case we choose for each k two models \mathbf{A} and \mathbf{B} so that $\mathbf{A} \equiv_k \mathbf{B}$, \mathbf{A} has the property in question, while \mathbf{B} does not.

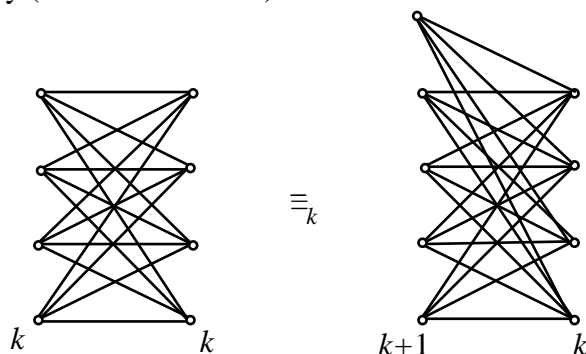
(1) Even cardinality.



(2) Equicardinality. $|P| = |Q|$.



(3) Hamiltonicity (Immermann 1982).



4.10. Corollary (Kolaitis-Vardi 1992). *Let K be an arbitrary model class. The following conditions are equivalent:*

- (1) K is definable in L^k ,
- (2) K is closed under \equiv_k .

Proof. (1) (2) by Theorem 4.8.

(2) (1) follows from Theorem 4.8 and the equivalence $\mathbf{A} \equiv_K \mathbf{B} \iff \{ \varphi \in FO^k : \mathbf{B} \models \varphi \} = \{ \varphi \in FO^k : \mathbf{A} \models \varphi \}$. (A normal-form for L^k). Q.E.D.

4.11 Theorem (Abiteboul-Vianu 1991). *The global relation $R(\bar{x}, \bar{y})$ which on any \mathbf{A} defines the relation*

$$(\mathbf{A}, a_1, \dots, a_k) \equiv_k (\mathbf{A}, b_1, \dots, b_k)$$

i.e. for all $\varphi(x_1, \dots, x_k) \in L^k$

$$\mathbf{A} \models \varphi(a_1, \dots, a_k) \iff \mathbf{A} \models \varphi(b_1, \dots, b_k)$$

is in FP.

Proof. It suffices to show that $\neg R(\bar{x}, \bar{y})$ is in FP. Let $D(x_1, \dots, x_k, y_1, \dots, y_k)$ be the global relation saying that for some atomic $\varphi(x_1, \dots, x_k)$ we have

$$\mathbf{A} \models \varphi(a_1, \dots, a_k) \wedge \mathbf{A} \not\models \varphi(b_1, \dots, b_k).$$

Surely $D(\bar{x}, \bar{y})$ is in FO. We can now define $\neg R(\bar{x}, \bar{y})$ as the solution $S(\bar{x}, \bar{y})$ of the following equation:

$$S(x_1, \dots, x_k, y_1, \dots, y_k) = D(x_1, \dots, x_k, y_1, \dots, y_k) \vee \bigvee_{i=1}^k (x_i = y_i \wedge S(x_1, \dots, x_k, y_1, \dots, y_k)) \vee \bigvee_{i=1}^k (y_i = x_i \wedge S(x_1, \dots, x_k, y_1, \dots, y_k)).$$

Q.E.D.

4.12 Corollary. *The global relation $(\mathbf{A}, a_1, \dots, a_k) \equiv_k (\mathbf{B}, b_1, \dots, b_k)$ is in PTIME. Similarly, the relation $\mathbf{A} \equiv_k \mathbf{B}$ is in PTIME.*

So, although L^k itself can express even non-recursive concepts, elementary equivalence \equiv_k is PTIME. The relations \equiv_k are important polynomial time versions of the NP concept.

Further results about L^k and \equiv_k :

- (1) There is a global relation \prec_k , which linearly orders classes $[\bar{a}]_k$, and the relation \prec_k is in *FP*.
- (2) On every \mathbf{A} , every $[\bar{a}]_k$ is in *FP*.
- (3) Every $\varphi \in L^k$ can be written as $\bigvee_i \varphi_i$ for some $\varphi_i \in FO^k$.
- (4) On k -rigid structures $FP=PTIME$. A structure is k -rigid if there is no permutation of the structure which maps \equiv_k -classes onto \equiv_k -classes.

For details concerning these results, see Dawar 1993, Dawar-Lindell-Weinstein 1994, and Kolaitis-Vardi 1992.

Lecture 5. Zero-One Laws

In this section L is a relational vocabulary and ϕ is an L -sentence. Suppose we choose an L -structure \mathbf{A} at random from among all L -structures of size n . What is the probability that \mathbf{A} satisfies ϕ ? Because no particular n is of interest here, we concentrate on the limit of this probability as $n \rightarrow \infty$.

For simplicity, we consider only graphs. So $L = \{E\}$ and L -structures are limited to graphs. The case of arbitrary structures is entirely similar. Let \mathbf{G}_n be the set of all graphs on $\{1, \dots, n\}$. Note that $|\mathbf{G}_n| = 2^{\binom{n}{2}}$.

The results of this section hold also if *isomorphism types* of structures are considered instead of structures themselves.

5.1. Definition Let \mathbf{P} be a property of graphs. Then we define

$$\mu_n(\mathbf{P}) = \frac{|\{G \in \mathbf{G}_n : G \text{ has } \mathbf{P}\}|}{|\mathbf{G}_n|},$$

$$\mu(\mathbf{P}) = \lim_n \mu_n(\mathbf{P}), \text{ if exists.}$$

Thus $\mu_n(\mathbf{P})$ is the probability that a randomly chosen graph on $\{1, \dots, n\}$ has property \mathbf{P} .

5.2 Example. Suppose \mathbf{P} says "there is an isolated vertex". Then $\mu(\mathbf{P}) = 0$, as the following calculation shows. The isolated vertex can be chosen in n ways and the remaining vertices can form a graph in any way.

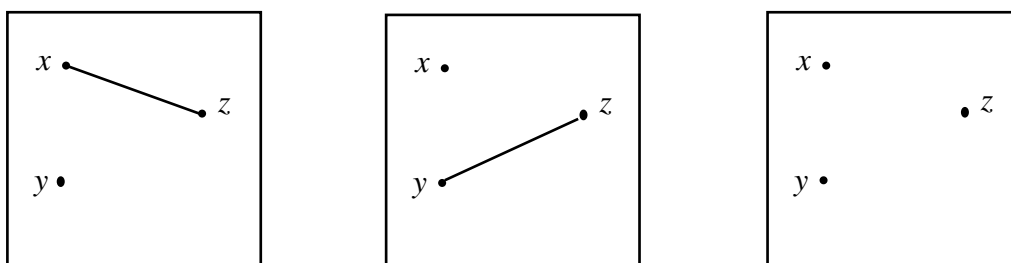
$$\mu_n(\mathbf{P}) = \frac{n \cdot 2^{\binom{n-1}{2}}}{2^{\binom{n}{2}}} = \frac{n}{2^n} \rightarrow 0.$$

Q.E.D.

5.3 Example. Suppose that \mathbf{P} says "the graph is connected". Then $\mu(\mathbf{P}) = 1$ which can be seen as follows. Let Q be the property

$$x \text{ --- } y(x \text{ --- } y \text{ --- } z(xEz \text{ --- } yEz)).$$

Property Q implies connectedness, so it suffices to prove $\mu(Q) = 1$. We shall prove $\mu(\neg Q) = 0$. We count how many pairs (x,y) there are so that one of the following edge patterns is realized with each of the remaining vertices z :



$$\mu_n(\neg Q) = \frac{\binom{n}{2} \cdot 2 \cdot 2^{n-2} + \binom{n}{2} \cdot 2 \cdot 2^{n-2} + \binom{n}{2} \cdot 3^{n-2}}{2^{n-2}} = \frac{n}{2} \cdot \frac{3}{4} \cdot 2^{n-2} = 0$$

Q.E.D.

5.4 Examples.

(1) By 5.3. $\mu(\text{"the graph contains a triangle"}) = 1$.

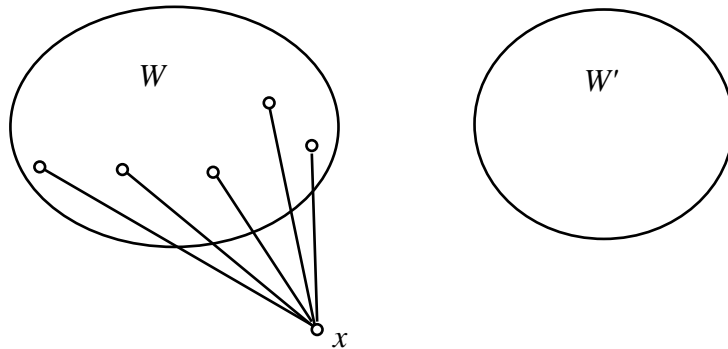
(2) By (1) $\mu(\text{"the graph is acyclic"}) = 0$.

(3) By (1) $\mu(\text{"the graph is 2-colorable"}) = 0$.

(4) μ ("even cardinality") does not exist because μ_n oscillates between 0 and 1.

The following graph property is called the *extension axiom* E_k :

If W and W' are disjoint sets of cardinality k ,
then there is x such that $\forall w \in W (wEx)$ and $\forall w \in W' (\neg wEx)$.



5.5 Proposition (Fagin 1976). $\mu(E_k) = 1$.

Proof. We show $\mu(\neg E_k) = 0$. How can E_k fail? There have to be disjoint k -sets W and W' with no x as above.

The number of ways to choose W and W' : $\binom{n}{k} \binom{n-k}{k}$.

The number of ways to put edges in $W \rightarrow W$: $2^{\binom{2k}{2}}$.

The number of ways to put edges outside $W \rightarrow W$: $2^{\binom{n-2k}{2}}$.

The number of ways to put edges between $W \rightarrow W'$ and its outside: $(2^{2k-1})^{n-2k}$.

$$\begin{aligned} \mu_n(\neg E_k) &= \frac{\binom{n}{k} \binom{n-k}{k} 2^{2k} 2^{n-2k} (2^{2k} - 1)^{n-2k}}{2^n} = \\ &= \binom{n}{k} \binom{n-k}{k} \left(1 - \frac{1}{2^{2k}}\right)^{n-2k} \end{aligned}$$

Remark. Since $\mu(E_k) = 1$, we know that E_k has finite models. This is the easiest way of showing that E_k has finite models at all. It is an example of a probabilistic model construction, which has become more and more important in finite model theory.

In contrast, infinite models of E_k are easy to construct:

5.6 Definition. *The Random Graph R has the vertex set $\{1, 2, 3, \dots\}$ and the edge relation*

$$iEj \text{ iff } i < j \text{ and } p_i | j \text{ (or } j < i \text{ and } p_j | i)$$

Here p_1, p_2, p_3, \dots are the primes $2, 3, 5, 7, \dots$ and $p_i | j$ means p_i divides j . Equivalently, one may just toss coin to decide if iEj holds or not.

5.7. Lemma. $R \models E_k$.

Proof. Suppose $W = \{i_1, \dots, i_k\}$ and $W' = \{j_1, \dots, j_k\}$ are disjoint. Let $z = p_{i_1} \dots p_{i_k}$. Then

$$p_{i_1} | z, \dots, p_{i_k} | z \text{ but } p_{j_1} \nmid z, \dots, p_{j_k} \nmid z.$$

Hence $i_1 E z, \dots, i_k E z$ and $\neg j_1 E z, \dots, \neg j_k E z$. Q.E.D.

5.8 Lemma. (Kolaitis -Vardi 1992). *Suppose H and H' are finite graphs such that $H \models E_k$ and $H' \models E_k$. Then $H \equiv_k H'$.*

Proof. We describe the strategy of II in $G^k(H, H')$. Suppose pairs

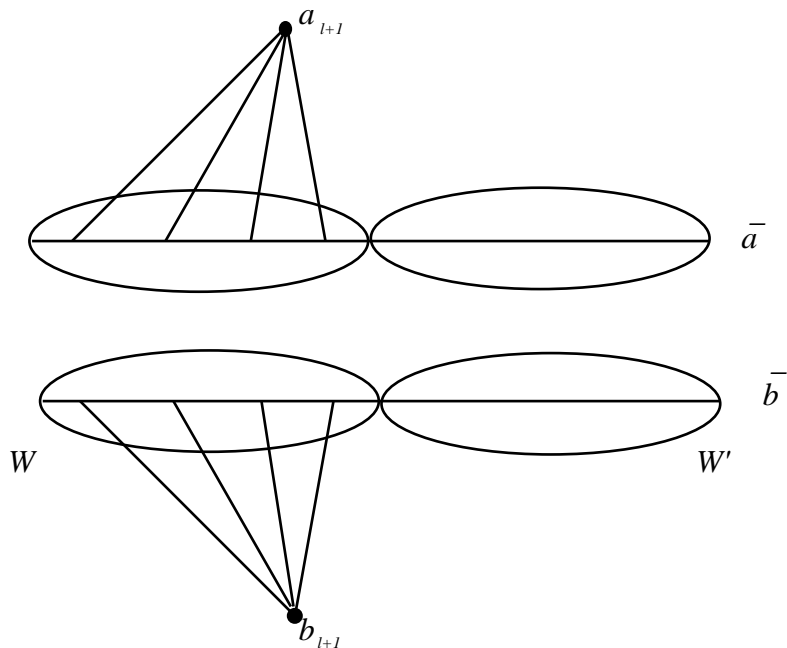
$$(a_1, b_1), \dots, (a_l, b_l), \quad l < k$$

have been pebbled and then I puts a pebble on a_{l+1} . Either $l+1 = k$ or a pebble is taken away from a_{i_0} for some $i_0 \leq l$. Let

$$W = \{b_i : a_i E a_{l+1}\},$$

$$W' = \{b_i : \neg a_i E a_{l+1}\}.$$

Then W and W' are disjoint sets of size k . Since $H \models E_k$ there is $b_{l+1} \in H$ such that $b_i E b_{l+1}$ for $b_i \in W$ and $\neg b_i E b_{l+1}$ for $b_i \in W'$.



Player II pebbles b_{l+1} . Clearly, $\{(a_i, b_i) : i \leq l+1\}$ remains a partial isomorphism.

Q.E.D.

Let $\mathbf{E} = \{E_k : k = 1, 2, 3, \dots\}$ be the first order theory consisting of all extension axioms.

5.9 Lemma. The first order theory \mathbf{E} has exactly one countable model, namely the random model \mathbf{R} .

Proof. We know already that \mathbf{R} is a model of \mathbf{E} . Suppose \mathbf{R}' is another countable model of \mathbf{E} . Recall the proof of 5.8: II could use E_k to make a successful move at a position where k elements had been played. This means that now that II has every E_k at his disposal, he can make the right move no matter how many elements have been played. Hence he can win the infinite Ehrenfeucht-Fraïssé game $EF_\omega(\mathbf{R}, \mathbf{R}')$. On countable structures this implies $\mathbf{R} \equiv \mathbf{R}'$ by a classical result of Carol Karp. Q.E.D.

5.10 Corollary (Gaifman). *The theory \mathbf{E} is complete and decidable.*

Proof. Suppose φ is an L -sentence, $L = \{E\}$. Suppose neither $\mathbf{E} \models \varphi$ nor $\mathbf{E} \models \neg\varphi$. By the Completeness Theorem and the Löwenheim-Skolem Theorem, there are countable models \mathbf{A} and \mathbf{B} of \mathbf{E} with $\mathbf{A} \models \varphi$, $\mathbf{B} \models \neg\varphi$. By Lemma 5.9, $\mathbf{A} \equiv \mathbf{B} \equiv \mathbf{R}$. This is a contradiction. So \mathbf{E} is complete. Every complete axiomatizable first order theory is decidable. Q.E.D.

5.11 Zero-One Law (Glebskii et al 1969, and independently, Fagin 1976 for FO and Kolaitis-Vardi 1992 for L^ω). *If $\varphi \in L^\omega$ then $\mu(\varphi)$ exists and $\mu(\varphi) = 0$ or $\mu(\varphi) = 1$.*

Proof. We assume (w.l.o.g.) $L = \{E\}$. Suppose $\varphi \in L^k$.

Case 1. For some $G \models E_k$ we have $G \models \varphi$. Then

$$G \models (G \models E_k \rightarrow G \models \varphi).$$

Here we used 5.8. So $E_k \models \varphi$. By 5.5. $\mu(E_k) = 1$. Hence $\mu(\varphi) = 1$.

Case 2. For all $G \models E_k$ we have $G \models \neg\varphi$. Since $\mu(E_k) = 1$, necessarily $\mu(\neg\varphi) = 1$ i.e. $\mu(\varphi) = 0$. Q.E.D.

We say that a zero-one law holds for a logic \mathbf{L} if for all $\varphi \in \mathbf{L}$ the limit $\mu(\varphi)$ exists and $\mu(\varphi) = 0$ or $\mu(\varphi) = 1$. We just proved that zero-one law holds for L^ω .

5.12 Corollary. *Zero-one law holds for FO and FP.*

Proof. It suffices to recall that $FO \subseteq L^\omega$ and $FP \subseteq L^\omega$. Q.E.D.

Note. For first order φ we have $\mu(\varphi) = 1$ iff $\mathbf{R} \models \varphi$. For suppose $\mathbf{R} \models \varphi$. Since \mathbf{E} is complete, $\mathbf{E} \models \varphi$, whence $E_k \models \varphi$ for some k . Since $\mu(E_k) = 1$, $\mu(\varphi) = 1$. Similarly, if $\mathbf{R} \models \neg\varphi$, then $\mu(\varphi) = 0$.

5.13 Corollary. *The question whether $\mu(\varphi) = 1$ or $\mu(\varphi) = 0$ can be effectively decided for $\varphi \in FO$.*

Proof. The claim follows from the equivalence $\mu(\varphi) = 1$ iff $\mathbf{R} \models \varphi$, and from the fact that the theory \mathbf{E} of \mathbf{R} is decidable. Q.E.D.

Note. By Trakhtenbrot's Theorem we *cannot* effectively decide whether $\varphi \in FO$ is valid in all finite models. But we *can* decide whether φ is valid in almost all models. Grandjean (1983) showed that this question is PSPACE-complete.

5.14 Definition. *A set A of natural numbers is called a spectrum, if there is a vocabulary L and an L -sentence φ so that A is the set of cardinalities of models of φ . Then A is the spectrum of φ .*

An old problem of logic asks: Is the complement of a spectrum again a spectrum? (Asser 1955). If "no" then $P \neq NP$, so it is a hard problem.

5.15 Corollary (Fagin 1976). *If φ is FO, then the spectrum of φ or the spectrum of $\neg\varphi$ is co-finite.*

Proof. Suppose A is the spectrum of φ and $\mu(\varphi) = 1$. Then there is n_0 so that $\mu_n(\varphi) > 0.5$ for $n \geq n_0$. Hence φ has a model of every size $n \geq n_0$ and $[n_0, \infty) \subseteq A$. If $\mu(\varphi) = 0$, then $\mu(\neg\varphi) = 1$ and the spectrum of $\neg\varphi$ contains $[n_0, \infty)$ for some n_0 . Q.E.D.

The zero-one law is more than an individual result - it is a method. Zero-one laws have been proved for many different logics e.g. fragments of second order logic and for different classes of structures e.g. different classes of graphs and partial orders. Furthermore, zero-one laws have been proved for different probability measures.

Whenever a zero-one law obtains, non-expressibility results follow, e.g. "even cardinality" cannot be expressible in a logic which has a zero-one law. Zero-one laws have similar universal applicability in finite model theory as Compactness Theorem has in infinite model theory.

Some Sources

Chapter 1.

- Gurevich Y. *Toward logic tailored for computer science* in *Computation and Proof Theory* (Ed. Richter et al.) Springer Lecture Notes in Math. 1104(1984), 175-216.
- Trakhtenbrot. B.A. *Impossibility of an algorithm for the decision problem on finite classes*, Doklady Akademii Nauk SSR 70(1950), 569-572.
- Tait, W. *A counterexample to a conjecture of Scott and Suppes*, Journal of Symbolic Logic 24(1959), 15-16.

Chapter 2.

- Fagin R. *Generalized first order spectra and polynomial-time reorganizable sets*, in *Complexity of Computation* (Ed. Karp) SIAM-AMS Proc.7(1974), 43-73.
- Fagin R. *Monadic generalized spectra* Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 21(1975), 89-96.
- Hájek P. *On logics of discovery*, in: Proc. 4th Conf. on Mathematical Foundations of Computer Science, Lecture Notes in Computer Science, Vol. 32 (Springer, Berlin, 1975) 30-45.
- Hájek P. *Generalized quantifiers and finite sets*, in Prace Nauk. Inst. Mat. Politech. Wrocław No. 14 Ser. Konfer. No 1 (1977) 91-104.

Chapter 3.

- Abiteboul S. and Vianu V. *Generic computation and its complexity*, in Proc. of the 23rd AMC Symposium on the Theory of Computing, 1991.
- Ajtai M. and Gurevich Y. *Monotone vs. positive*, Journal of the ACM 34(1987), 1004-1015.
- Immerman N. *Relational queries computable in polynomial time*, Information and Control, 68(1986), 76-98.

- Leivant D. *Inductive definitions over finite structures*, Information and Computation, 89(1990), 95-108.
- Moschovakis Y. *Elementary Induction on Abstract Structures*, North Holland, 1974.
- Vardi M. *The complexity of relational query languages*, in Proc. 14th AMC Symp. on Theory of Computing (1982), 137-146.

Chapter 4.

- Barwise J. *On Moschovakis closure ordinals*, Journal of Symbolic Logic 42(1977), 292-296.
- Dawar A. *Feasible Computation through Model Theory*. PhD thesis, University of Pennsylvania, 1993.
- Dawar A, Lindell S and Weinstein S. *Infinitary logic and inductive definability over finite structures*. Information and Computation, 1994. to appear.
- Immerman N. *Upper and lower bounds for first order expressibility*, Journal of Computer and System Sciences 25(1982), 76-98.
- Kolaitis Ph. and Vardi M. *0-1 laws for infinitary logic*, Information and Computation 98(1992), 258-294.
- Kolaitis Ph. and Vardi M. *Fixpoint logic vs. infinitary logic in finite-model theory*. In Proc. 7th IEEE Symp. on Logic in Computer Science, 1992, 46-57.

Chapter 5.

- Fagin R. *Probabilities on finite models*, Journal of Symbolic Logic 41(1976), 50-58.
- Grandjean, E. *Complexity of the first--order theory of almost all structures*, Information and Control 52(1983), 180-204.
- Kolaitis Ph. and Vardi M. *0-1 laws and decision problems for fragments of second order logic*, Information and Computation 87(1990), 302-338.
- Kolaitis Ph. and Vardi M. *0-1 laws for infinitary logic*, Information and Computation 98(1992), 258-294.
- Spencer J. *Zero-One laws with variable probability*, Journal of Symbolic Logic 1993.