Lecture 5

What have we done so far?

We prepared the grounds to formulate and prove Theorem A: The 0-1 law for $\mathcal{L}_{\infty,\omega}$.

- We defined the logic $\mathcal{L}_{\infty,\omega}$.
  It's model theory is discussed in detail in
  J. Keisler, *Model theory for infinitary logic: Logic with countable conjunctions and finite quantifiers*, North Holland, 1971

- We defined the logic $\mathcal{L}^{\omega}_{\infty,\omega}$.
  The logic $\mathcal{L}_{\infty,\omega}$ is too expressive for the study of finite structures. So we restricted the overall number of variables.

- We introduced pebble games $PG^n_k(\mathfrak{A}, \mathfrak{B})$ and $PG^\omega_k(\mathfrak{A}, \mathfrak{B})$.

- We practiced playing the game on simple graphs and on graphs satisfying the extension axioms $EA_{n,m}$ for all $m \leq n \leq k$.

Pebble games, IV

Let $\mathfrak{A}_0$ and $\mathfrak{A}_1$ be two τ-structures.

We say $\mathfrak{A}_0 \sim_k^{\infty,\omega} \mathfrak{A}_1$ if they agree on all sentences of $\mathcal{L}^k_{\infty,\omega}$.

We say $\mathfrak{A}_0 \sim_{k,n}^{\infty,\omega} \mathfrak{A}_1$ if they agree on all sentences of $\mathcal{L}^k_{\infty,\omega}$ of quantifier rank n.

Theorem 1 *(Theorem 2 of Lecture 4)*

(i) $\mathfrak{A}_0 \equiv_k^n \mathfrak{A}_1$ *iff* $\mathfrak{A}_0 \sim_{k,n}^{\infty,\omega} \mathfrak{A}_1$

(ii) $\mathfrak{A}_0 \equiv_k^\infty \mathfrak{A}_1$ *iff* $\mathfrak{A}_0 \sim_k^{\infty,\omega} \mathfrak{A}_1$

Before we prove this theorem, we give some applications.

Disjoint unions of structures

Let $\mathfrak{A}_0$ and $\mathfrak{A}_1$ be two τ-structures. We denote the disjoint union by $\mathfrak{A}_0 \sqcup \mathfrak{A}_1$.

Theorem 2

Let $\mathfrak{A}_0$, $\mathfrak{A}_1$ $\mathfrak{B}_0$, $\mathfrak{B}_1$ be τ-structures such that

$$\mathfrak{A}_0 \equiv_k^n \mathfrak{B}_0$$

and

$$\mathfrak{A}_1 \equiv_k^n \mathfrak{B}_1.$$

Then

$$\mathfrak{A}_0 \sqcup \mathfrak{A}_1 \equiv_k^n \mathfrak{B}_0 \sqcup \mathfrak{B}_1$$

Proof: Exercise!

Exercise: Prove the analogue theorem for graph complement, for join and cartesian product of two graphs.

What about graphs as two-sorted structures, line graphs?

Proposition 3
$EVEN$ is not definable in $\mathcal{L}^k_{\infty,\omega}$ with equality only, for any $k \in \mathbb{N}$.

Proof:

We show that sets of cardinality bigger than k
are not distinguishable. Q.E.D.

Proposition 4
$\exists^{\geq k+1} x(x = x)$,
$\exists^{\leq k} x(x = x)$ and
$\exists^{=k} x(x = x)$
are not definable in $\mathcal{L}^k_{\infty,\omega}$ with equality only.

Proof:

We show that sets of cardinality bigger or equal than k
are not distinguishable. Q.E.D.

Proposition 5
The class of linear orders is not definable in $\mathcal{L}^2_{\infty,\omega}$, but is definable in FOL^3.

Proof:

Let Ord_n be a finite linear order with n elements.
Let $C_n^{dir}(s,t)$ be directd graph obtained from C_n^{dir} by adding two vertices s,t (source and target), such that s has edges pointing to all the vertices, and t has edges with all vertices pointing to t.
We show that, for $n \geq 3$, the directed graph $C_n^{dir}(s,t)$ is not distinguishable in $\mathcal{L}^2_{\infty,\omega}$ from Ord_n. Q.E.D.

Proposition 6

The class of 2-colorable graphs **IS** *definable in* $\mathcal{L}^k_{\infty,\omega}$*, for some* k*, but the class of 3-colorable graphs* **IS NOT** *definable in* $\mathcal{L}^k_{\infty,\omega}$*, for any* $k \in \mathbb{N}$*.*

Proof:

The undirected circle C_n on n vertices is 2-colorable iff n is even.
Now a graph is 2-colorable iff it has no even cycle. But the existence of even cycles is definable in $\mathcal{L}^k_{\infty,\omega}$, for some k.

Exercise: Find the smallest k!

The non-definability of 3-colorability is due to

A. Dawar, A restricted second order logic for finite structures,
Logic and Computational Complexity, Springer LNCS 960 (1994), pp. 393-413. Q.E.D.

Proposition 7

The class of HAM of graphs which have a hamiltonian circuit, is not definable in $\mathcal{L}^k_{\infty,\omega}$, for any $k \in \mathbb{N}$.

Proof: **Exercise!**

Hint: $K_{m,n}$ is in HAM iff $m = n$. $K_{m,n} = E_m \bowtie E_n$. But for m, n much larger than k, E_m and E_n cannot be distinguished in $\mathcal{L}^k_{\infty,\omega}$.

Problem:

Is the class $PLANAR$ of planar graphs, definable in $\mathcal{L}^k_{\infty,\omega}$, for any $k \in \mathbb{N}$?

Hint: We try a negative answer:

Look at K_5 and change an edge preserving the degrees but making it planar. The replace each edge by very long paths.

Theorem 8
On ordered $\tau^{<}$-structures, $\tau^{<} = \tau \cup R_{ord}$,
any class of structures $K \in \mathbf{PSpace}$ is definable in $\mathcal{L}^{\omega}_{\infty,\omega}$.

This remains without proof.

See Libkin's book.

Exercise:

Pick your favorite $\mathbf{PSpace}$-complete problem, e.g. HEX, and write down the formula which defines it, without simulating the proof of the theorem.

Pebble games for $\mathbf{SOL}$

- For $\mathbf{SOL}^m$ the players can also pick relations of arity $\leq m$.
- The notions of partial isomorphisms now includes the relations already picked.
- Formulate the analogues of Theorem 1.
- Try to play the game for $\mathbf{SOL}^1 = \mathbf{MSOL}$.
- Try to play the game for $\mathbf{SOL}^2$.

Recall:

0-1 Laws for asymptotic probabilities

Theorem A: (Kolaitis and Vardi, 1992)
(generalizing a long sequence of earlier papers since 1964)

For $\mathcal{P}$ definable in infinitary logic with finitely many variables, $\mathcal{L}^{\omega}_{\infty,\omega}$, either $\mu_{\mathcal{P}} = 0$ or $\mu_{\mathcal{P}} = 1$.

This works also for any constant probability p and $\mu^{p}_{\mathcal{P}}$.

Theorem B: (Shelah and Spencer, 1988)

For $\alpha \in [0,1]$ irrational, $\mathcal{P}$ definable in FOL, either $\mu^{n^{-\alpha}}_{\mathcal{P}} = 0$ or $\mu^{n^{-\alpha}}_{\mathcal{P}} = 1$.

For all rational α there are counterexamples.

Strategy for the Proof of Theorem A

- Use extension axioms EA and their probabilities.

  $\mu(EA_k) = 1$ for each $k \in \mathbb{N}$.

- Use winning strategies for extension axioms EA.

  Assume G_0, G_1 are two graphs in which for all $m \leq n \leq k \in \mathbb{N}$ the exioms $EA_{n,m}$ hold. Then $G_0 \equiv_k^{\infty} G_1$.

- Use Theorem 1.

  Conclude that $G_0 \sim_k^{\infty,\omega} G_1$.

The last steps of the proof of Theorem A

- Let $\Phi \in \mathcal{L}^k_{\infty,\omega}(\tau)$.
- We can assume that Φ has arbitrarily large finite models which also satisfy EA_k, otherwise, $\mu(\Phi) \leq \mu(\neg EA_k) = 0$ and we are done.

  Let $G_0 \models (EA_k \wedge \Phi)$ be of size at least $2k$.
- Assume $G_1 \models EA_k$ and has at least $2k$ elements.

  Then we have $G_0 \sim_k^{\infty,\omega} G_1$,
  and therefore, $G_1 \models \Phi$.
- Now we conclude that $\mu(\Phi) \geq \mu(EA_k) \geq 1$.

Q.E.D.

Lecture 5 (continued)

Last lecture:

- We showed how to use Theorem 1 linking Pebble Games and Logics.
- We proved various non-definability results.
- We looked briefly at games picking also relations.
- We proved Theorem A.

Strategy for a proof of Theorem 1

Linking logics and games.

- Assume $\phi \in \mathcal{L}^k_{\infty,\omega}(\tau)$ and $\mathfrak{A}_0 \models \phi$ but $\mathfrak{A}_1 \models \neg\phi$.

  We build a winning strategy for player I (the spoiler) for the k-pebble game.

- Given a τ-structure $\mathfrak{A}$, we construct a sentence $\sigma_{\mathfrak{A}} \in \phi \in \mathcal{L}^k_{\infty,\omega}(\tau)$, such that, whenever $\mathfrak{B} \models \sigma_{\mathfrak{A}}$, then $\mathfrak{B} \equiv^{\infty}_{k} \mathfrak{A}$.

  $\sigma_{\mathfrak{A}}$ is called the **Scott sentence of** $\mathfrak{A}$.

- We conclude $\mathfrak{B} \equiv^{\infty}_{k} \mathfrak{A}$ iff $\mathfrak{B} \sim^{\infty,\omega}_{k} \mathfrak{A}$.

We can do the same for the finite versions of the game.

Quantifier rank, I

We write, say, an **MSOL**-formula ϕ as a tree:

$$\exists X_1 \forall x_2 \, (x_2 \in X_1 \rightarrow \exists x_3 E(x_2, x_3))$$

$$\exists X_1$$

$$\forall x_2$$

$$\rightarrow$$

$$x_2 \in X_1 \qquad\qquad \exists x_3$$

$$E(x_2, x_3)$$

The quantifier rank is biggest number of quantifiers one can find along a path in this tree.

Here it is 3.

Quantifier rank of a formula, II

- For formulas in prenex normal form
  the quantifier rank equals
  the number of quantifiers.

- If we reuse variables, the quantifier rank can be smaller than the number of quantifiers used in prenex normal form.

$$\forall x_1 \, (\exists x_2 E(x_1, x_2) \wedge \exists x_2 \neg E(x_1, x_2))$$

Quantifier rank 2

$$\forall x_1 \exists x_2 \exists x_3 \, (E(x_1, x_2) \wedge \neg E(x_1, x_3))$$

Quantifier rank 3

Quantifier rank of a formula, III

- In the **FOL**-case and **MSOL**-case the quantifiers of elements and the set quantifiers our counted the same way.
- For **SOL**-formulas we may want to give the second order quantifiers a different weight, say using the arity.

How many non-equivalent formulas? FOL atomic case

Assume we have (first order) variables

$$x_1, x_2, \ldots, x_v$$

This gives $\binom{v}{2} + \binom{v}{1} = O(v^2)$ many instances of $x_i = x_j$ with $i \leq j$.

For a r-ary relation symbol R we get r^v many instances of $R(x_{j_1}, x_{j_2}, \ldots, x_{j_r})$.

If we allow $c_1, c_2, \ldots, c_{v'}$ constants the numbers become $O((v+v')^2)$ and $r^{v+v'}$ respectively.

Proposition:
For a fixed finite relational vocabulary τ with constants and v first order variables, there are a finite number of atomic formulas $\alpha_{\tau,v}^{FOL}$.

How many non-equivalent formulas? $MSOL$ atomic case

Assume we have first and second order variables

$$x_1, x_2, \ldots, x_{v_1}, U_1, U_2, \ldots, U_{v_2}$$

This gives
$O(v_1^2)$ many instances of $x_i = x_j$ with $i \leq j$
and $v_1 \cdot v_2$ many instances of $x_i \in U_j$.

For a r-ary relation symbol R we get r^v many instances of $R(x_{j_1}, x_{j_2}, \ldots, x_{j_r})$.
If we allow $c_1, c_2, \ldots, c_{v_3}$ constants the numbers become $\binom{v_1+v_3}{2}$, $(v_1+v_3)v_2$ and $r^{v_1+v_3}$ respectively.

Proposition:
For a fixed finite relational vocabulary τ with constants and v first order variables, there are a finite number of atomic formulas $\alpha_{\tau,v}^{MSOL}$.

How many non-equivalent formulas? Quantifierfree case

For quantifierfree formulas we only count formulas in CNF.

There are $2^{\alpha^{FOL}_{\tau,v}}$, resp. $2^{\alpha^{MSOL}_{\tau,v}}$ many disjunctions

$$\bigvee_{j=1}^{2^{\alpha^{FOL}_{\tau,v}}} (\neg)^{\nu(j)} A_j$$

where A_j ranges over atomic formulas.

Hence we have (at most) $2^{2^{\alpha^{FOL}_{\tau,v}}}$ many formulas in CNF.

Proposition:
For a fixed finite relational vocabulary τ with constants and v first order variables, there are a finite number of atomic formulas $\beta^{FOL}_{\tau,v}$ and $\beta^{MSOL}_{\tau,v}$, respectively.

How many non-equivalent formulas?
Quantifiers I: PNF

Counting quantified formulas is a bit more tricky.
We can assume that the formulas are in

Prenex Normal Form

But then variables are NOT reused.

So for each CNF formula with v variables there are $3^v \cdot v!$ many quantifier prefixes
($\exists, \forall$, not quantified).

This gives at most

$$3^v \cdot v! \cdot \beta\tau, v^{FOL}$$

many prenex normal form formulas.

How many formulas are there ?
Quantifiers II: quantifier rank

Theorem:
For each τ and $v = v_1 + v_2$ many variables

$$x_1, x_2, \ldots, x_{v_1}, U_1, U_2, \ldots, U_{v_2}$$

there are only $\gamma^{MSOL}_{\tau,v,q}$ many formulas of quantifier rank q.

Proof: We estimate this number by induction over q for $MSOL$.

For $q = 0$ we have at most γ many formulas with $\gamma_0 = \beta\tau, v^{MSOL}$.

Treating them as atomic formulas we have $2v$ many ways of adding one quantifier, and hence at most

$$\gamma^{MSOL}_{\tau,v,q+1} = \gamma_{q+1} = 2^{2^{2v \cdot \eta q}}$$

many formulas of rank $q + 1$.

How many formulas are there ?
Quantifiers II: quantifier rank

Exercise:

(i) Count the formulas of $\mathbf{SOL}$.

(ii) Count the non-equivalent formulas of $\mathcal{L}^k_{\infty,\omega}(\tau)$.

How many non-equivalent formulas
are there really?

Exact estimates to the best of our knowledge unknown.

From distinguishing formulas to winning strategies, I

$\phi \in \mathbf{MSOL}^k$ in Prenex Normal Form

Ehrenfeucht-Fraïssé Theorem, I

Theorem 1:(Easy part)

Assume there is a $\mathbf{MSOL}(\tau)$-sentence ϕ with k variables and quantifier depth n in Prenex Normal Form such that $\mathcal{A}_0 \models \phi$ and $\mathcal{A}_1 \models \neg\phi$.

Then I has a winning strategy for the k-pebble n-moves game on $\mathcal{A}_0$ and $\mathcal{A}_1$.

Ehrenfeucht-Fraïssé Theorem, II

We first assume that there infinitely many pebbles.

We write ϕ and $\neg\phi$ in Prenex Normal Form:

$$\phi = \exists X_1 \exists x_2 \forall X_3 \exists x_4 \ldots \exists x_{n-1} \exists X_n$$
$$B(X_1, x_2, \ldots, x_{n-1}, X_n)$$
$$\neg\phi = \forall X_1 \forall x_2 \exists X_3 \forall x_4 \ldots \forall x_{n-1} \forall X_n$$
$$\neg B(X_1, x_2, \ldots, x_{n-1}, X_n)$$

where B is without quantifiers.

We can read from the quantifier prefix a winning strategy.

Ehrenfeucht-Fraïssé Theorem, III

Assume $\mathcal{A}_0 \models \phi$ and $\mathcal{A}_1 \models \neg\phi$.
Player I follows the existential quantifiers.

Player I picks in $\mathcal{A}_0$ a set A_1 such that

$$\mathcal{A}_0, A_1^0 \models \exists x_2 \forall X_3 \exists x_4 \ldots \exists x_{n-1} \exists X_n B(A_1^0, x_2, \ldots, x_{n-1}, X_n)$$

Whatever player II picks as A_1^1

$$\mathcal{A}_1, A_1^1 \models \forall x_2 \exists X_3 \forall x_4 \ldots \forall x_{n-1} \forall X_n \neg B(A_1^1, x_2, \ldots, x_{n-1}, X_n)$$

Ehrenfeucht-Fraïssé Theorem, III (continued)

Next player I picks an element a_2^0 in $\mathcal{A}_0$ such that

$$\mathcal{A}_0, A_1^0, a_2^0 \models \forall X_3 \exists x_4 \ldots \exists x_{n-1} \exists X_n B(A_1^0, a_2^0, \ldots, x_{n-1}, X_n)$$

Whatever player II picks as a_2^1

$$\mathcal{A}_1, A_1^1, a_2^1 \models \exists X_3 \forall x_4 \ldots \forall x_{n-1} \forall X_n \neg B(A_1^1, a_2^1, \ldots, x_{n-1}, X_n)$$

Now player I picks in $\mathcal{A}_1$ a set A_3^1 such that

$$\mathcal{A}_1, A_1^1, a_2^1, A_3^1 \models \forall x_4 \ldots \forall x_{n-1} \forall X_n \neg B(A_1^1, a_2^1, A_3^1, \ldots, x_{n-1}, X_n)$$

and so on..........

Ehrenfeucht-Fraïssé Theorem, IV

Finally the outcome is from $\mathcal{A}_0$

$$A_1^0, a_2^0, A_3^0, \ldots, a_{n-1}^0, A_n^0$$

and from $\mathcal{A}_1$

$$A_1^1, a_2^1, A_3^1, \ldots, a_{n-1}^1, A_n^1$$

such that

$$\mathcal{A}_0 \models B(A_1^0, a_2^0, A_3^0, \ldots, a_{n-1}^0, A_n^0)$$

and

$$\mathcal{A}_1 \models \neg B(A_1^1, a_2^1, A_3^1, \ldots, a_{n-1}^1, A_n^1)$$

which shows that player I wins, as this cannot be a local isomorphism

We need a **Lemma** on local isomorphisms and quantifierfree formulas.

The easy case: What remains?

- We have to discuss what we do if the formulas are not in Prenex Normal Form.
- We have to discuss the case for $\mathcal{L}_{\infty,\omega}$.
- The easy part also works for $\mathbf{SOL}$ in all its variations.