Finite Model Theory Lecture 8: Ehrenfeucht-Fraisse Games

Spring 2025

Announcements

• Homework 2 due on Friday.

• HW3 will be about EF games: will release it soon.

• Today: we finish EF games.

• Next week: Second order logic, strings, trees, regular languages.

Review: Expressibility of FO

 $\mbox{Goal:}$ describe what properties $\mathrm P$ we can express in FO.

Each finite structure **A** is uniquely identified by a sentence φ^A .

• We cannot find **A**, **B** such that P(A) = 1, P(B) = 0 and **A**, **B** satisfy exactly the same sentences.

Workaround: restrict the quantifier depth k, FO[k]:

• For each k find A_k, B_k such that $P(A_k) = 1, P(B_k) = 0$ and A_k, B_k satisfy exactly the same sentences in FO[k].

Lots of notations!

- **A** ~ **B**: isomorphic
- A ≡ B: elementary equivalent, i.e. satisfy the same FO sentences when A, B are finite, then this implies A ≃ B.
- $\mathbf{A} \equiv_k \mathbf{B}$: equivalent for all sentences in FO[k].
- **A** ~_k **B**: duplicator wins the EF game with k pebbles

Ehrenfeucht-Fraisse theorem:

$$\mathbf{A} \equiv_k \mathbf{B}$$
 iff $\mathbf{A} \sim_k \mathbf{B}$.

$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$

Prove that the Spoiler has a winning strategy with k = 3 pebbles.



Whatever spoiler does at round 3, duplicator wins. Let's try a better strategy for spoiler

5/30

$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E,B)$$
 $E(x,y) = edge, B(x) = "black"$



$$\sigma = (E,B)$$
 $E(x,y) = edge, B(x) = "black"$



$$\sigma = (E,B)$$
 $E(x,y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$



$$\sigma = (E, B)$$
 $E(x, y) = edge, B(x) = "black"$

Prove that the Spoiler has a winning strategy with k = 3 pebbles.



At home: find $\varphi \in FO[3]$ that separates the two trees!

Finite	Model ⁻	Theory

Discussion

• Number of rounds k correspond to quantifier depth FO[k]

When duplicator plays on *A* it corresponds to ∃
When duplicator plays on *B* it corresponds to ∀.

• We will re-examine this shortly, but first let's see two applications.

EVEN and CONNECTIVITY

Review: Linear Order

$$L_n = ([n], <)$$
 (1-2-3-4-5-6

If $m, n \geq 2^k - 1$, then $\mathbf{L}_m \sim_k \mathbf{L}_n$.

Otherwise $\mathbf{L}_m \sim_k \mathbf{L}_n$ iff m = n.

Corollary

EVEN is not expressible over linear orders.

Proof Suppose φ is s.t. $L_n \models \varphi$ iff *n* is even. Choose $m \stackrel{\text{def}}{=} 2^k - 1$ and $n \stackrel{\text{def}}{=} 2^k$, where $k = qr(\varphi)$. Then $L_n \models \varphi$, hence $L_m \models \varphi$, contradiction.

CONNECTIVITY: given a graph G = (V, E), check if it is connected.

Corollary

CONNECTIVITY is not expressible in FO

CONNECTIVITY: given a graph G = (V, E), check if it is connected.

Corollary

CONNECTIVITY is not expressible in FO

Proof Reduction from EVEN. Assume ψ expresses CONNECTIVITY.

CONNECTIVITY: given a graph G = (V, E), check if it is connected.

Corollary

CONNECTIVITY is not expressible in FO

5

Proof Reduction from EVEN. Assume ψ expresses CONNECTIVITY.

Given L_n , construct **G** with edges (i, i+2) and (1, n)



3

CONNECTIVITY: given a graph G = (V, E), check if it is connected.

Corollary

CONNECTIVITY is not expressible in FO

5

Proof Reduction from EVEN. Assume ψ expresses CONNECTIVITY.

Given L_n , construct **G** with edges (i, i+2) and (1, n)



Then n is even iff **G** is connected.

Finite Model Theory	
---------------------	--

3

CONNECTIVITY: given a graph G = (V, E), check if it is connected.

Corollary

CONNECTIVITY is not expressible in FO

5

Proof Reduction from EVEN. Assume ψ expresses CONNECTIVITY.

Given L_n , construct **G** with edges (i, i+2) and (1, n)



 $E(x,y) \stackrel{\text{def}}{=} (\exists z(\operatorname{succ}(x,z) \land \operatorname{succ}(z,y))) \lor (\operatorname{isMin}(x) \land \operatorname{isMax}(y))$

Then n is even iff G is connected.

Finite Model The	ory
------------------	-----

3

Discussion

EF games are a powerful tool for proving inexpressibility. Examples:

• EVEN, CONNECTIVITY.

• Check if a graph is a tree.

• Check if a graph is planar.

Deep Dive into FO[k]

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

12/30

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

• Sentences with constants only, e.g. $R_1(c_1, c_2) \land \neg R_2(c_2, c_2)$.

12/30

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

- Sentences with constants only, e.g. $R_1(c_1, c_2) \land \neg R_2(c_2, c_2)$.
- If σ has no constants, FO[0] is empty, and $\mathbf{A} \equiv_0 \mathbf{B}$ for any \mathbf{A}, \mathbf{B} .

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

- Sentences with constants only, e.g. $R_1(c_1, c_2) \land \neg R_2(c_2, c_2)$.
- If σ has no constants, FO[0] is empty, and $\mathbf{A} \equiv_0 \mathbf{B}$ for any \mathbf{A}, \mathbf{B} .

Can the duplicator lose in 0-rounds?

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

- Sentences with constants only, e.g. $R_1(c_1, c_2) \wedge \neg R_2(c_2, c_2)$.
- If σ has no constants, FO[0] is empty, and $\mathbf{A} \equiv_0 \mathbf{B}$ for any \mathbf{A}, \mathbf{B} .

Can the duplicator lose in 0-rounds?

• YES: e.g. when $\boldsymbol{A} \vDash R_1(c_1, c_2)$ and $\boldsymbol{B} \vDash \neg R_1(c_1, c_2)$.

Vocabulary $\sigma = (R_1, \ldots, R_n, c_1, \ldots, c_m).$

EF games add the constants (a_1, \ldots, a_k) and (b_1, \ldots, b_k) ; unavoidable.

What is FO[0]?

- Sentences with constants only, e.g. $R_1(c_1, c_2) \land \neg R_2(c_2, c_2)$.
- If σ has no constants, FO[0] is empty, and $\mathbf{A} \equiv_0 \mathbf{B}$ for any \mathbf{A}, \mathbf{B} .

Can the duplicator lose in 0-rounds?

- YES: e.g. when $\boldsymbol{A} \models R_1(c_1, c_2)$ and $\boldsymbol{B} \models \neg R_1(c_1, c_2)$.
- If σ has no constants, duplicator always wins in 0 rounds.
Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ...

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

13/30

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.



Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

FO[0] sentences: e.g. $E(s,s) \land \neg E(t,s)$.

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

FO[0] sentences: e.g. $E(s,s) \land \neg E(t,s)$.

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

FO[0] sentences: e.g. $E(s,s) \land \neg E(t,s)$. # FO[0] atoms with 1 variable x: $E(x,x) \land \neg E(x,s), \ldots$

2²²² 32

Counting the Number of Sentences in FO[k]

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

FO[0] sentences: e.g. $E(s,s) \land \neg E(t,s)$. # FO[0] atoms with 1 variable x: $E(x,x) \land \neg E(x,s), \ldots$

Fix *n* Boolean variables: X_1, X_2, \ldots, X_n .

How many Boolean functions exists? E.g. $X_1 \wedge X_2$; or $X_1 \vee \neg X_3 \wedge X_4$, ... Answer: 2^{2^n}

Graphs with source and target: $\sigma = (E, s, t)$.

FO[0] atoms without variables: E(s,s), E(s,t), E(t,s), E(t,t). 2²

- # FO[0] sentences: e.g. $E(s,s) \land \neg E(t,s)$. $2^{2^{2^2}}$ # FO[0] atoms with 1 variable x: $E(x,x) \land \neg E(x,s), \dots$ 3^2
- # FO[0] formulas with 1 variable x:
- # FO[1] sentences:

2^{23*}

Deep Dive into FO[k]

-O[*k*] Types

Hanf's Lemma

Main Takeaway for FO[k]

FO[k] is finite!

Next: let's talk about FO[k] types.

FO[k] Types

A type in logic is everything you can say about an element, or a tuple of elements.

Types in programming languages are very simple special cases.

We will define FO[k] types and use them in the proof of the EF theorem. Later, we define types of other logics.

Deep Dive into FO[*k*]

The EF Theorem

Theorem (Ehrenfeucht-Fraisse)

 $\mathbf{A} \equiv_k \mathbf{B}$ iff $\mathbf{A} \sim_k \mathbf{B}$.

Deep Dive into FO[*k*]

The EF Theorem

Theorem (Ehrenfeucht-Fraisse)

 $\mathbf{A} \equiv_k \mathbf{B}$ iff $\mathbf{A} \sim_k \mathbf{B}$.

Part 1: (we will review) If duplicator wins k rounds then $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$.

If $\mathbf{A} \vDash \varphi$ we use the duplicator's strategy to show $\mathbf{B} \vDash \varphi$.

The EF Theorem

Theorem (Ehrenfeucht-Fraisse)

 $\mathbf{A} \equiv_k \mathbf{B}$ iff $\mathbf{A} \sim_k \mathbf{B}$.

Part 1: (we will review) If duplicator wins k rounds then $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$.

If $\mathbf{A} \vDash \varphi$ we use the duplicator's strategy to show $\mathbf{B} \vDash \varphi$.

Part 2: (we will prove) If A, B satisfy the same FO[k] sentences, then duplicator wins k rounds

Here we use "FO-types" to inform the duplicator how to play.

If duplicator wins k rounds, then $\mathbf{A} \vDash \varphi$ iff $\mathbf{B} \vDash \varphi$.

If duplicator wins k rounds, then $\mathbf{A} \vDash \varphi$ iff $\mathbf{B} \vDash \varphi$.

Base Case k = 0: duplicator wins in 0 rounds. Then the constants in A, B form a partial isomorphism, thus $A \equiv_0 B$.

If duplicator wins k rounds, then $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$.

Induction k > 0: The interesting cases are $\exists x \psi$ and $\forall x \psi$.

If
$$\boldsymbol{A} \vDash \exists x \psi(x)$$
, we prove that $\boldsymbol{B} \vDash \exists x \psi(x)$

If duplicator wins k rounds, then $\mathbf{A} \vDash \varphi$ iff $\mathbf{B} \vDash \varphi$.

Induction k > 0: The interesting cases are $\exists x \psi$ and $\forall x \psi$.

If
$$\mathbf{A} \models \exists x \psi(x)$$
, we prove that $\mathbf{B} \models \exists x \psi(x)$

There exists $a \in A$ such that $A \models \psi(a)$. Use the duplicator's strategy to find $b \in B$ s.t. $B \models \psi(b)$.

¹Formally: $(\mathbf{A}, \mathbf{a}) \sim_{k-1} (\mathbf{B}, \mathbf{b})$; by induction $(\mathbf{A}, \mathbf{a}) \equiv_{k-1} (\mathbf{B}, \mathbf{b})$, thus $\mathbf{B} \models \psi(\mathbf{b})$.

18 / 30

If duplicator wins k rounds, then $\mathbf{A} \vDash \varphi$ iff $\mathbf{B} \vDash \varphi$.

Induction k > 0: The interesting cases are $\exists x \psi$ and $\forall x \psi$.

If
$$\mathbf{A} \models \exists x \psi(x)$$
, we prove that $\mathbf{B} \models \exists x \psi(x)$

There exists $a \in A$ such that $A \models \psi(a)$. Use the duplicator's strategy to find $b \in B$ s.t. $B \models \psi(b)$.

If
$$\mathbf{A} \models \forall x \psi(x)$$
 then $\mathbf{B} \models \forall x \psi(x)$. Proof by contradiction:

If there exists $b \in B$ s.t. $B \models \neg \psi(b)$, Use duplicator to find $a \in A$ s.t. $A \models \neg \psi(a)$. Contradiction.

• Existential quantifier: spoiler plays on **A**, duplicator plays on **B**.

• Universal quantifier: spoiler plays on **B**, duplicator plays on **A**.

If $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$, then duplicator wins k rounds.

If $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$, then duplicator wins k rounds.

Suppose spoiler placed the third pebble a_3 :

 $a_1, a_2, a_3 \in A$ $b_1, b_2 \in B$

If $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$, then duplicator wins k rounds.

Suppose spoiler placed the third pebble a_3 :

 $a_1, a_2, a_3 \in A$ $b_1, b_2 \in B$

 a_1, a_2 and b_1, b_2 form a partial isomorphism.

Where should the duplicator place $b_3 \in B$?

20 / 30

If $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$, then duplicator wins k rounds.

Suppose spoiler placed the third pebble a_3 :

 $a_1, a_2, a_3 \in A$ $b_1, b_2 \in B$

 a_1, a_2 and b_1, b_2 form a partial isomorphism.

Where should the duplicator place $b_3 \in B$?

The partial isomorphism is not sufficient for us to help duplicator find b_3 . We need to strengthen the requirement on the duplicator to:

If $\mathbf{A} \models \varphi$ iff $\mathbf{B} \models \varphi$, then duplicator wins k rounds.

Suppose spoiler placed the third pebble a_3 :

 $a_1, a_2, a_3 \in A$ $b_1, b_2 \in B$

 a_1, a_2 and b_1, b_2 form a partial isomorphism.

Where should the duplicator place $b_3 \in B$?

The partial isomorphism is not sufficient for us to help duplicator find b_3 . We need to strengthen the requirement on the duplicator to:

a, **b** have the same FO[k]-Types

Deep Dive into FO[*k*]

FO[k]-Types

Fix k and m.

Definition

Let **A** be a structure, $\mathbf{a} \stackrel{\text{def}}{=} (a_1, \ldots, a_m) \in A^m$. The rank k m-type of **a** is:

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

• $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$

22 / 30

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

- $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$
- Equivalent definition: an k, m-type is a complete, consistent set of formulas with qr ≤ k and m free variables x₁,..., x_m. Note: may differ for finite or infinite: e.g. end-of-the-line.

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

- $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$
- Equivalent definition: an k, m-type is a complete, consistent set of formulas with qr ≤ k and m free variables x₁,..., x_m. Note: may differ for finite or infinite: e.g. end-of-the-line.

• What is $tp_{k,0}$?

22 / 30

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

- $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$
- Equivalent definition: an k, m-type is a complete, consistent set of formulas with qr ≤ k and m free variables x₁,..., x_m. Note: may differ for finite or infinite: e.g. end-of-the-line.
- What is $tp_{k,0}$? Set of FO[k] sentences true in **A**.

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

- $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$
- Equivalent definition: an k, m-type is a complete, consistent set of formulas with qr ≤ k and m free variables x₁,..., x_m. Note: may differ for finite or infinite: e.g. end-of-the-line.
- What is $tp_{k,0}$? Set of FO[k] sentences true in **A**.
- What is tp_{0,m}??

$$\mathsf{tp}_{k,m}(\boldsymbol{A},\boldsymbol{a}) = \{\varphi(x_1,\ldots,x_m) \in FO[k] \mid \boldsymbol{A} \vDash \varphi(a_1,\ldots,a_m)\}$$

- $tp_{k,m}(\mathbf{A}, \mathbf{a})$ is complete: For all $\varphi \in FO[k]$, either $\varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$ or $\neg \varphi \in tp_{k,m}(\mathbf{A}, \mathbf{a})$
- Equivalent definition: an k, m-type is a complete, consistent set of formulas with qr ≤ k and m free variables x₁,..., x_m. Note: may differ for finite or infinite: e.g. end-of-the-line.
- What is $tp_{k,0}$? Set of FO[k] sentences true in **A**.
- What is $tp_{0,m}$?? Boolean expressions on atoms with constants.

Will prove that $\mathbf{A} \equiv_k \mathbf{B}$ implies $\mathbf{A} \sim_k \mathbf{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$\operatorname{tp}_{k-i,i}(\boldsymbol{A},\boldsymbol{a}) = \operatorname{tp}_{k-i,i}(\boldsymbol{B},\boldsymbol{b})$$

Why is this sufficient?

23 / 30

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$\operatorname{tp}_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = \operatorname{tp}_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Why is this sufficient?

After k rounds, $tp_{0,k}(\mathbf{A}, \mathbf{a}) = tp_{0,k}(\mathbf{B}, \mathbf{b})$, thus \mathbf{a}, \mathbf{b} partial isomorphism.
Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Proof by induction on *i*.

When i = 0: $\mathbf{A} \equiv_k \mathbf{B}$ is same as $tp_{k,0}(\mathbf{A}) = tp_{k,0}(\mathbf{B})$.

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$\operatorname{tp}_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = \operatorname{tp}_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$; spoiler picks $\boldsymbol{a}_{i+1} \in A$.

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(\mathbf{A}, \mathbf{a}) = tp_{k-i,i}(\mathbf{B}, \mathbf{b})$; spoiler picks $a_{i+1} \in A$. We need to help duplicator find "the right" $\mathbf{b} \in B$.

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(\mathbf{A}, \mathbf{a}) = tp_{k-i,i}(\mathbf{B}, \mathbf{b})$; spoiler picks $a_{i+1} \in A$. We need to help duplicator find "the right" $\mathbf{b} \in B$. \mathbf{B} is finite, $B = \{u_1, u_2, \dots, u_n\}$, and assume u_j is "wrong":

 $\mathsf{tp}_{k-i-1,i+1}(\boldsymbol{A},(\boldsymbol{a},a_{i+1})) \neq \mathsf{tp}_{k-i-1,i+1}(\boldsymbol{B},(\boldsymbol{b},u_j))$

23 / 30

Will prove that $\mathbf{A} \equiv_k \mathbf{B}$ implies $\mathbf{A} \sim_k \mathbf{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$\operatorname{tp}_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = \operatorname{tp}_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(\mathbf{A}, \mathbf{a}) = tp_{k-i,i}(\mathbf{B}, \mathbf{b})$; spoiler picks $a_{i+1} \in A$. We need to help duplicator find "the right" $\mathbf{b} \in B$. \mathbf{B} is finite, $B = \{u_1, u_2, \dots, u_n\}$, and assume u_j is "wrong": $\begin{bmatrix} tp_{k-i-1,i+1}(\mathbf{A}, (\mathbf{a}, a_{i+1})) \neq tp_{k-i-1,i+1}(\mathbf{B}, (\mathbf{b}, u_j)) \end{bmatrix}$ There exists $\varphi_j \in FO[k - i - 1]$ s.t.: $\mathbf{A} \models \varphi_j(\mathbf{a}, a_{i+1})$ and $\mathbf{B} \models \neg \varphi_j(\mathbf{b}, u_j)$

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(A, a) = tp_{k-i,i}(B, b)$; spoiler picks $a_{i+1} \in A$. We need to help duplicator find "the right" $b \in B$. **B** is finite, $B = \{u_1, u_2, \dots, u_n\}$, and assume u_j is "wrong": $\begin{array}{c} tp_{k-i-1,i+1}(A, (a, a_{i+1})) \neq tp_{k-i-1,i+1}(B, (b, u_j))) \\ \text{There exists } \varphi_j \in FO[k - i - 1] \text{ s.t.:} \\ A \models \varphi_j(a, a_{i+1}) \text{ and } B \models \neg \varphi_j(b, u_j) \\ \text{Let } \varphi(x_1, \dots, x_i, x_{i+1}) = \varphi_1 \land \dots \land \varphi_n. \text{ Then:} \\ A \models \exists x \varphi(a, x) \text{ and } B \models \neg \exists x \varphi(b, x) \end{array}$

Will prove that $\boldsymbol{A} \equiv_k \boldsymbol{B}$ implies $\boldsymbol{A} \sim_k \boldsymbol{B}$.

We design a strategy for the duplicator that ensures, for i = 0, 1, ..., k:

$$tp_{k-i,i}(\boldsymbol{A}, \boldsymbol{a}) = tp_{k-i,i}(\boldsymbol{B}, \boldsymbol{b})$$

Induction step Assume $tp_{k-i,i}(A, a) = tp_{k-i,i}(B, b)$; spoiler picks $a_{i+1} \in A$. We need to help duplicator find "the right" $b \in B$. B is finite, $B = \{u_1, u_2, ..., u_n\}$, and assume u_j is "wrong": $\begin{bmatrix} tp_{k-i-1,i+1}(A, (a, a_{i+1})) \neq tp_{k-i-1,i+1}(B, (b, u_j)) \end{bmatrix}$ There exists $\varphi_j \in FO[k - i - 1]$ s.t.: $A \models \varphi_j(a, a_{i+1})$ and $B \models \neg \varphi_j(b, u_j)$ Let $\varphi(x_1, ..., x_i, x_{i+1}) = \varphi_1 \land \dots \land \varphi_n$. Then: $A \models \exists x \varphi(a, x)$ and $B \models \neg \exists x \varphi(b, x)$ Contradiction, since $\varphi \in tp_{k-i,j}(A, a) = tp_{k-i,j}(B, b)$

23 / 30



• Types are defined for various logics; they are useful to reason about both formulas and elements in structures.

- Since $tp_{k,m}$ is finite, there is a single formula that captures the type: $\varphi(x_1, \dots, x_m) \stackrel{\text{def}}{=} \bigwedge_{\psi \in tp_{k,m}(A,a)} \psi(x_1, \dots, x_m)$
- A thought experiment: replace a huge database *D*, with the finite collection of FO[k] types that hold in *D*. This is sufficient to answer all queries with ≤ k quantifiers over *D*!

Hanf's Lemma

Describing Winning Strategies

Recall: to prove inexpressibility, we need to show that, for all k, the duplicator has a winning strategy for k rounds: A_k ~_k B_k.

• But difficult to prove that the duplicator has a winning strategy.

• Hanf's lemma: sufficient condition for a winning strategy.

Neighborhoods

 $\sigma = (R_1, \dots, R_m, c_1, \dots, c_s).$ The Gaifman graph² of a **A**: edges (a, b), a, b co-occurring in a relation.

d-neighborhood of $a \in A$: $N(a,d) \stackrel{\text{def}}{=} \{b \in A \mid d(a,b) \leq d\} \cup \{c_1^A, \ldots, c_s^A\}.$

Definition

The *d*-type of *a* is the isomorphism type of the substructure induced by N(a, d), plus the constant *a*.

Definition

A, B are *d*-equivalent if, for each *d*-type, they have the same number of elements of that type.

²A.k.a. primal graph.

Hanf's Lemma

Theorem

Let $d \ge 3^{k-1} - 1$. If \mathbf{A}, \mathbf{B} are d-equivalent, then $\mathbf{A} \sim_k \mathbf{B}$.

The proof exhibits a winning strategy for the duplicator.

We will omit the proof.

28 / 30

Prove that duplicator has winning strategy with k = 3 pebbles.



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(=3^{k-1} - 1)$



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(=3^{k-1} - 1)$



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)?



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)?



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)?



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)? What is N(b, d)?



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)? What is N(b, d)?



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)? What is N(b, d)?

Their types are structures x - x - * - x - x; there are 12 copies in each structure. Therefore, duplicator wins with k = 2 pebbles.



Prove that duplicator has winning strategy with k = 3 pebbles. Let's use Hanf's lemma: k = 2 and $d = 2(= 3^{k-1} - 1)$ What is N(a, d)? What is N(b, d)?

Their types are structures x - x - * - x - x; there are 12 copies in each structure. Therefore, duplicator wins with k = 2 pebbles.



At home: prove that spoiler wins in 3 rounds

29 / 30

- Ehrenfeucht-Fraisse games can be applied to infinite structures as well! Try at home: update Part 2 of the proof of the EF theorem to remove the need for **A**, **B** to be finite.
- EF games generalize to other games to prove inexpressibility results. We will discuss two:
 - ► Inexpressibility for ∃MSO
 - Inexpressibility for logics with recursion.