Finite and Algorithmic Model Theory

Lecture 4 (Dresden 02.11.22, Short version with errors)

Lecturer: Bartosz "Bart" Bednarczyk

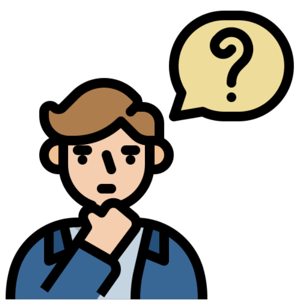
Technische Universität Dresden & Uniwersytet Wrocławski

Today's agenda

Goal: Provide a game-theoretic framework for proving FO-inexpressivity (also in the finite!).

1. Quantifier rank of FO sentences.
2. Quantifier rank $\not\approx$ #variable.
3. Definition of Ehrenfeucht-Fraïssé games (proof omitted)
4. Showcase 1: Games on sets (FO[$\emptyset$]-nondefinability of "even" strikes back)
5. Showcase 2: Games on linear orders ("even" is not FO[$\{<\}$]-definable)
6. Logical reductions, e.g. "even" $\notin$ FO[$\{<\}$] $\Longrightarrow$ "connectivity" $\notin$ FO[$\{\mathrm{E}\}$]

Lecture based on chapters 3.1, 3.2, 3.6 of [Libkin's FMT Book]



Feel free to ask questions and interrupt me!

Don't be shy! If needed send me an email (bartosz.bednarczyk@cs.uni.wroc.pl) or approach me after the lecture!

Reminder: this is an advanced lecture. Target: people that had fun learning logic during BSc studies!

Measuring complexity of a formula: quantifier rank

The quantifier rank $\mathrm{qr}(\varphi)$ of φ is its depth of quantifier nesting.

- $\mathrm{qr}(\varphi) := 0$ for atomic φ
- $\mathrm{qr}(\neg\varphi) := \mathrm{qr}(\varphi)$
- $\mathrm{qr}(\varphi \oplus \varphi') := \max(\mathrm{qr}(\varphi), \mathrm{qr}(\varphi'))$ for $\oplus \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$
- $\mathrm{qr}(\exists x\ \varphi) = \mathrm{qr}(\forall x\ \varphi) := \mathrm{qr}(\varphi) + 1$

Examples:

$$\mathrm{qr}(\exists\varphi\forall y\forall z\ \mathrm{R}(x, y, z)) = 3$$

$$\mathrm{qr}(\exists x\, [\mathrm{A}(x) \wedge (\forall y \mathrm{R}(y)) \vee (\exists z \top)]) = 2$$

for φ in PNF $\mathrm{qr}(\varphi) = \#$quantifiers.

Quantifier rank can be exponentially smaller than the total number of quantifiers.

$\varphi_0(x, y) := \mathrm{E}(x, y), \quad \varphi_{n+1}(x, y) := \exists z\ (\varphi_n(x, z) \wedge \varphi_n(z, y)) \quad \rightsquigarrow\ \mathrm{qr}(\varphi_n) = n$ but φ_n has $2^n - 1$ quants.

Formulae with bounded quantifier rank

Let τ be a *finite* signature, and let $m \in \mathbb{N}$. $\mathrm{FO}_m[\tau]$ is set of all FO formulae over τ with q.r. $\leq m$.

Notation: $\mathfrak{A} \equiv^{\tau}_{m} \mathfrak{B}$ iff $\mathfrak{A}$ and $\mathfrak{B}$ satisfy precisely the same $\mathrm{FO}_m[\tau]$ sentences (τ often omitted).

Lemma (Finiteness of $\mathrm{FO}_m[\tau]$ with $\leq k$ variables)

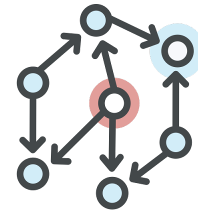
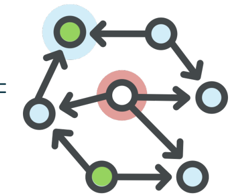
The set of all $\mathrm{FO}_m[\tau]$ formulae with at most k free variables is finite up to logical equivalence.

Proof

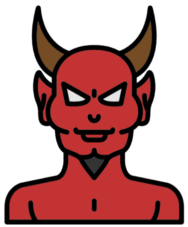
Idea: characterise $\mathrm{FO}_0[\tau]$ with a “truth table” of equality between constants/variables + induction!

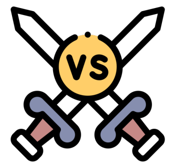
Ehrenfeucht-Fraïssé games

- Duration: m rounds.
- Playground: two τ-structures $\mathfrak{A}$ and $\mathfrak{B}$.

$\mathfrak{A} :=$  $\mathfrak{B} :=$ 

- Two players: Spoil∃r (D∃vil/∃loise/∃ve/Player I) vs Duplic∀tor (∀ngel/∀belard/∀dam/Player II)





Goal of $\forall$: $\mathfrak{A}, \mathfrak{B}$ "look the same".

Goal of $\exists$: pinpoint the difference.

- During the i-th round:

1. $\exists$ selects a structure (say $\mathfrak{A}$) and picks an element (say $a_i \in A$)

2. $\forall$ replies with an element (say $b_i \in B$) in the other structure (in this case $\mathfrak{B}$)

so that $(a_1 \mapsto b_1, \ldots, a_i \mapsto b_i)$ is a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

- $\exists$ wins if $\forall$ cannot reply with a suitable element. $\forall$ wins if he survives m rounds.

Theorem (Fraïssé 1950 & Ehrenfeucht 1961)

$\forall$ has a winning strategy in m-round Ehrenfeucht-Fraïssé game on τ-structures $\mathfrak{A}$ and $\mathfrak{B}$ iff $\mathfrak{A} \equiv_m^\tau \mathfrak{B}$.

Playing Ehrenfeucht-Fraïssé games on sets

Consider an 3-round play of E-F game on sets $\mathfrak{A} := \{1, 2, 3\}$, $\mathfrak{B} := \{a, b, c, d\}$.

$\mathfrak{A} :=$

$1 \mapsto d,\ 2 \mapsto b,\ 3 \mapsto c$

Result: $\forall$ wins, so $\mathfrak{A} \equiv_3 \mathfrak{B}$.

$\mathfrak{B} :=$

Following the strategy "always reply with a fresh element", $\forall$ wins any m-round game on sets of size $\geq m$.

Lemma (Even is not expressible in FO[$\emptyset$])

Proof Assume that such a φ exists. Let $m := \mathrm{qr}(\varphi)$. Let $\mathfrak{A}$ (resp. $\mathfrak{B}$) be an $2m$ (resp. $2m+1$) element set.

By definition, we clearly have $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$.

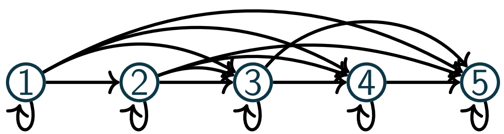
As we already noticed $\forall$ has the winning strategy in any m-round E-F game. Thus $\mathfrak{A} \equiv_m \mathfrak{B}$ holds.

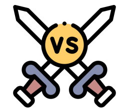
By collecting the inferred information, we conclude $\mathfrak{B} \models \varphi$. A contradiction!

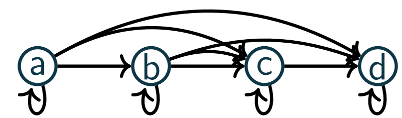
General proof scheme for showing that $\mathcal{P}$ is not FO[τ]-definable with Ehrenfeucht-Fraïssé games

ad absurdum φ exists $\rightsquigarrow$ q.r. of φ $\rightsquigarrow$ craft τ-structures $\mathfrak{A} \models \varphi$, $\mathfrak{B} \not\models \varphi$ $\rightsquigarrow$ play qr(φ)-round game $\rightsquigarrow$ E-F theorem $\rightsquigarrow$ infer $\mathfrak{B} \models \varphi$ contradiction!

Playing Ehrenfeucht-Fraïssé games on linear orders

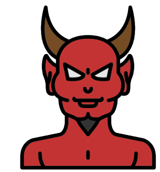
$\mathfrak{A} :=$ 



$\mathfrak{B} :=$ 

- Who has the winning strategy in 2 rounds?

- In 3 rounds? more?



Lemma (Even is not expressible in FO$[\{\leq\}]$)

Proof Suppose that φ exists. Let $m := \mathrm{qr}(\varphi)$. Let $\mathfrak{A}$ (resp. $\mathfrak{B}$) be linear orders of size 2^m (resp. 2^m+1).

By definition, we clearly have $\mathfrak{A} \models \varphi$ and $\mathfrak{B} \not\models \varphi$.

What remains to be done is to show that $\forall$ has the winning strategy in any m-round E-F game.

Thus $\mathfrak{A} \equiv_m \mathfrak{B}$ holds. By collecting the inferred information, we conclude $\mathfrak{B} \models \varphi$. A contradiction!

---

ad absurdum φ exists — q.r. of φ — craft τ-structures $\mathfrak{A} \models \varphi$, $\mathfrak{B} \not\models \varphi$ — play $\mathrm{qr}(\varphi)$-round game — E-F theorem — infer $\mathfrak{B} \models \varphi$ contradiction!

Super Lemma About Linear Orders: I

Lemma (Sufficiently large linear orders look similar)

Any linearly ordered[a] $\{\leq\}$-structures $\mathfrak{A}, \mathfrak{B}$ of cardinality $\geq 2^m$ satisfy $\mathfrak{A} \equiv_m^{\{\leq\}} \mathfrak{B}$.

[a] We assume that $\mathfrak{A}, \mathfrak{B}$ interpret $\leq$ as a linear order over the domain

induction

- Let $\overline{a} := (a_{-1}, a_0, \ldots, a_i)$ and $\overline{b} := (b_{-1}, b_0, \ldots, b_i)$ be the history of the play after i-rounds.
- Dummy (-1)-th and 0-th rounds of the game: select min/max elements of $\mathfrak{A}, \mathfrak{B}$.

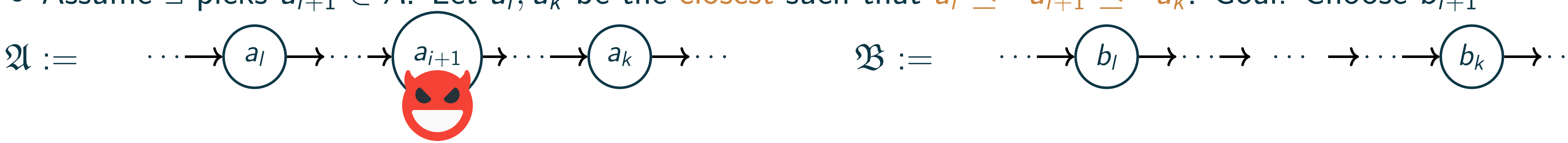
This establishes an invariant that any freshly selected element is between some previously selected ones.

- We play as $\forall$: we want to guarantee that after the i-th round we have for all $l, k \leq i$:

1. $a_k \leq^{\mathfrak{A}} a_l$ iff $b_k \leq^{\mathfrak{B}} b_l$ (maintain the partial isomorphism).
2. If $\mathrm{dist}(a_k, a_l) \geq 2^{m-i}$ then $\mathrm{dist}(b_k, b_l) \geq 2^{m-i}$ ("play far if $\exists$ plays far").
3. If $\mathrm{dist}(a_k, a_l) < 2^{m-i}$ then $\mathrm{dist}(a_k, a_l) = \mathrm{dist}(b_k, b_l)$ ("play close if $\exists$ plays close").

$\forall$ should preserve these conditions

- Assume $\exists$ picks $a_{i+1} \in A$. Let a_l, a_k be the closest such that $a_l \leq^{\mathfrak{A}} a_{i+1} \leq^{\mathfrak{A}} a_k$. Goal: Choose b_{i+1}



Super Lemma About Linear Orders: II

Recall that $\exists$ picked $a_{i+1} \in A$ and a_l, a_k are the closest such that $a_l \leq^{\mathfrak{A}} a_{i+1} \leq^{\mathfrak{A}} a_k$.

$$\mathfrak{A} := \cdots \to a_l \to \cdots \to a_{i+1} \to \cdots \to a_k \to \cdots \qquad \mathfrak{B} := \cdots \to b_l \to \cdots \to \cdots \to \cdots \to b_k \to \cdots$$

Inductive assumption for all $l, k \leq i$:

1. $a_k \leq^{\mathfrak{A}} a_l$ iff $b_k \leq^{\mathfrak{B}} b_l$ (maintain the partial isomorphism).
2. If $\mathrm{dist}(a_k, a_l) \geq 2^{m-i}$ then $\mathrm{dist}(b_k, b_l) \geq 2^{m-i}$ ("play far if $\exists$ plays far").
3. If $\mathrm{dist}(a_k, a_l) < 2^{m-i}$ then $\mathrm{dist}(a_k, a_l) = \mathrm{dist}(b_k, b_l)$ ("play close if $\exists$ plays close").

$\forall$ should find a suitable b_{i+1}

Case I: $\mathrm{dist}(a_l, a_k) < 2^{m-i}$

$$\mathfrak{A} := \cdots \to \overbrace{a_l \to \cdots \to a_{i+1} \to \cdots \to a_k}^{\mathrm{dist}(a_l, a_k) < 2^{m-i}} \to \cdots \qquad \mathfrak{B} := \cdots \to \overbrace{b_l \to \cdots \to \cdots \to b_k}^{\text{by assump: } \mathrm{dist}(b_l, b_k) = \mathrm{dist}(a_l, a_k)} \to \cdots$$

Then by ind. ass. $\mathrm{dist}(a_l, a_k) = \mathrm{dist}(b_l, b_k)$, and hence $[a_l, a_k] \cong [b_l, b_k]$.

Pick b_{i+1} such that $b_l \leq^{\mathfrak{B}} b_{i+1} \leq^{\mathfrak{A}} b_l$. $\mathrm{dist}(a_l, a_{i+1}) = \mathrm{dist}(b_l, b_{i+1})$, and $\mathrm{dist}(a_k, a_{i+1}) = \mathrm{dist}(b_k, b_{i+1})$.

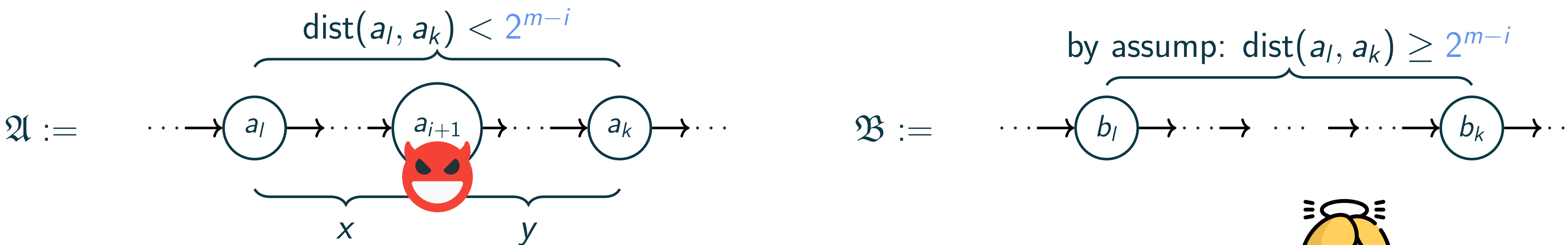
Super Lemma About Linear Orders: III

Inductive assumption for all $l, k \leq i$:

1. $a_k \leq^{\mathfrak{A}} a_l$ iff $b_k \leq^{\mathfrak{B}} b_l$ (maintain the partial isomorphism).
2. If $\mathrm{dist}(a_k, a_l) \geq 2^{m-i}$ then $\mathrm{dist}(b_k, b_l) \geq 2^{m-i}$ ("play far if $\exists$ plays far").
3. If $\mathrm{dist}(a_k, a_l) < 2^{m-i}$ then $\mathrm{dist}(a_k, a_l) = \mathrm{dist}(b_k, b_l)$ ("play close if $\exists$ plays close").

$\forall$ should find a suitable b_{i+1}

Case II: $\mathrm{dist}(a_l, a_k) \geq 2^{m-i}$



$\forall$ wins!

Then by ind. ass. $\mathrm{dist}(b_l, b_k) \geq 2^{m-i}$. We have three cases.

- $x \geq 2^{m-i-1}$ and $y \geq 2^{m-i-1} \rightsquigarrow$ Take b_{i+1} to the middle between b_l and b_k.
- $x < 2^{m-i-1}$ and $y \geq 2^{m-i-1} \rightsquigarrow$ b_{i+1} is the unique node to the right of b_l so that $\mathrm{dist}(b_l, b_{i+1}) = x$.
- $x \geq 2^{m-i-1}$ and $y < 2^{m-i-1} \rightsquigarrow$ b_{i+1} is the unique node to the left of b_k so that $\mathrm{dist}(b_{i+1}, b_k) = y$.

More about Ehrenfeucht-Fraïssé games

There is an alternative approach to the previous proof by composing winning strategies. Key lemma:

Lemma (Composition lemma)

Let $\mathfrak{A}, \mathfrak{B}$ be linearly-ordered, with $a \in A, b \in B$ s.t. $\mathfrak{A}^{\leq a} \equiv_m \mathfrak{B}^{\leq b}$ and $\mathfrak{A}^{\geq a} \equiv_m \mathfrak{B}^{\geq b}$. Then $\mathfrak{A} \equiv_m \mathfrak{B}$.

We can compose strategies under:

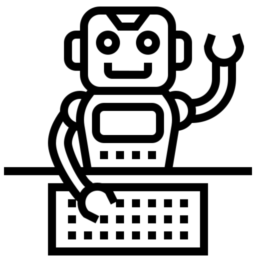
1. Disjoint unions. Consult a lecture by Anuj Dawar 9:50-19:20 [Youtube].
2. Ordered sums. as well as Thm. 3.6, Proof #2 (p. 30–31) and Ex. 3.15 from [Libkin's book].
3. Products.

Algorithmic approach to Ehrenfeucht-Fraïssé games: Can we make E-F games computable?

Input: First-Order φ over τ, finite structures $\mathfrak{A}, \mathfrak{B}$ and $m \in \mathbb{N}$.

Output: Is this the case that Duplication has the winning strategy in m-round E-F game?

Is this problem decidable?: **YES!** and $\mathrm{PS{\small PACE}}$-complete, c.f. [Pezzoli 1998]



A lot of open problems, e.g. "how difficult is to solve the above problem when $\mathfrak{A}$, $\mathfrak{B}$ are trees?"

Consult excellent slides by [Angelo Montanari] for more!

Logical Reductions

If $\mathcal{P}$ is not expressible, show that $\mathcal{P}'$ is not. Use case: "even" $\notin$ FO[$\{\leq\}$] implies "connectivity" $\notin$ FO[$\{\leq\}$]

- Suppose $\varphi \in$ FO[$\{\mathrm{E}\}$] defines connectivity.
- From $\leq$ we can define the succ. relation:

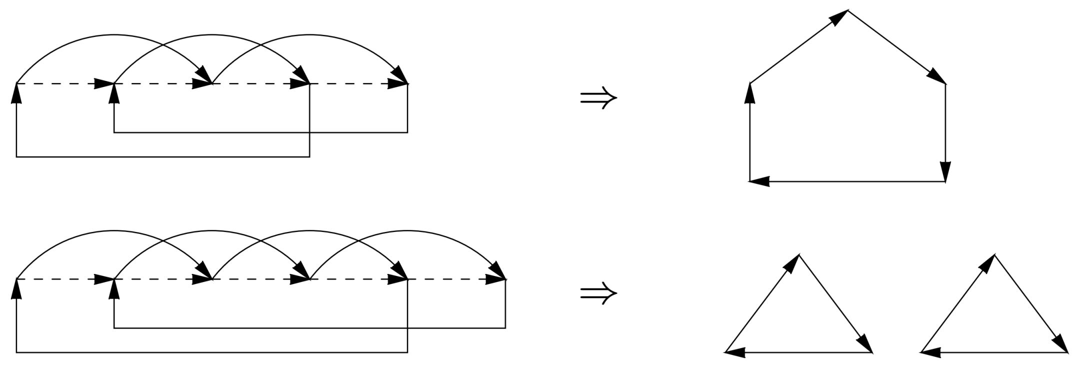
$$\mathrm{succ}(x, y) := (x < y) \land \forall z\ ((z \leq x) \lor (y \leq z))$$

- Prepare $\gamma(x, y)$ that holds if

1. y is the succ of succ of x, or

2. x is sec-to-last and y is the first w.r.t $\leq$, or

3. x is the last one and y is the second w.r.t $\leq$.



Reduction of parity to connectivity

- Note: γ defines a graph on the elements of the linear order!
- Observation: graph defined by γ is connected iff the underlying linear order is odd.

Conclusion: $\neg\varphi[\mathrm{E}/\gamma]$ defined even. A contradiction!

Playing Ehrenfeucht-Fraïssé games is quite difficult. Can we simplify them?

Yes, with a notion of locality. Next 2–3 lectures!

Copyright of used icons and pictures

1. Universities/DeciGUT/ERC logos downloaded from the corresponding institutional webpages.
2. Idea icon created by Vectors Market — Flaticon `flaticon.com/free-icons/idea`.
3. Head icons created by Eucalyp — Flaticon `flaticon.com/free-icons/head`
4. Graph icons created by SBTS2018 — Flaticon `flaticon.com/free-icons/graph`
5. Angel icons created by Freepik — Flaticon `flaticon.com/free-icons/angel`
6. Devil icons created by Freepik and Pixel perfect — Flaticon `flaticon.com/free-icons/devil`
7. VS icons created by Freepik — Flaticon `flaticon.com/free-icons/vs`
8. Robot icon created by Eucalyp — Flaticon `flaticon.com/free-icons/robot`.
9. Figure 3.3. from Libkin's Book `homepages.inf.ed.ac.uk/libkin/fmt/fmt.pdf`
10. Warning icon created by Freepik - Flaticon `flaticon.com/free-icons/warning`.