Finite and Algorithmic Model Theory : Lecture 4

Ehrenfeucht - Fraïssé Games

During the i-th round:

(1) $\exists$ selects a structure (e.g. $\mathfrak{A}$) and its element a_i

(2) $\forall$ responds in the other structure (e.g. $\mathfrak{B}$) and selects b_i

such that there is a partial isomorphism between $\mathfrak{A}\restriction_{\{a_1, a_2, \dots, a_i\}}$ and $\mathfrak{B}\restriction_{\{b_1, b_2, \dots, b_i\}}$

$a_j = c^{\mathfrak{A}}$ iff $b_j = c^{\mathfrak{B}}$

$(a_{j_1}, \dots, a_{j_k}) \in R^{\mathfrak{A}}$ iff $(b_{j_1}, \dots, b_{j_k}) \in R^{\mathfrak{B}}$

$\exists$ wins if $\forall$ cannot make its move.

$\forall$ wins if he survives m rounds.

We write $\mathfrak{A} \equiv_m \mathfrak{B}$ if B∀rt has a winning strategy in m-round E-F game on $\mathfrak{A}$ and $\mathfrak{B}$.

Why E-F games are useful?

(1) Quantifier rank

Definition 3.8 (Quantifier rank). *The* quantifier rank *of a formula* $\mathsf{qr}(\varphi)$ *is its depth of quantifier nesting. That is:*

- *If φ is atomic, then* $\mathsf{qr}(\varphi) = 0$.
- $\mathsf{qr}(\varphi_1 \vee \varphi_2) = \mathsf{qr}(\varphi_1 \wedge \varphi_2) = \max(\mathsf{qr}(\varphi_1), \mathsf{qr}(\varphi_2))$.
- $\mathsf{qr}(\neg\varphi) = \mathsf{qr}(\varphi)$.
- $\mathsf{qr}(\exists x \varphi) = \mathsf{qr}(\forall x \varphi) = \mathsf{qr}(\varphi) + 1$.

We use the notation ~~FO[k]~~ FO_k *for all* FO *formulae of quantifier rank up to* k.

Examples:

* $qr(\exists x \forall y \exists z\, R(x,y,z)) = 3$
* $qr(\exists x\, A(x) \wedge (\forall y\, R(x,y)) \vee (\exists z\, Q(z))) = 2$
* Let $\varphi_0(x,y) = E(x,y)$, $\varphi_{n+1} := \exists z\ \varphi_n(x,z) \wedge \varphi_n(z,y)$

  φ_n has $2^n - 1$ quantifiers but $q_r(\varphi_n) = n$.
* For φ in prenex-normal-form $qr = \#$quantifiers

! **Theorem.** $\mathfrak{A} \equiv_m \mathfrak{B}$ iff for all $\varphi \in FO_m$ we have $\mathfrak{A} \models \varphi \Leftrightarrow \mathfrak{B} \models \varphi$.

Proof: todo.

(2) How to use them? E.g. we want to show that eveness/parity is not FO[$\emptyset$]-definable.

1. By compactness, employ

   $T_{even} = \{\varphi, \lambda_n \mid n \in \mathbb{N}\}$ $T_{odd} = \{\neg\varphi, \lambda_n\}$

   + Skolem theorem to get countable models

   + $\exists$ with $\cong$

2. By zero-one laws

   $\mu_\infty(\text{even})$ doesn't exist!

3. E-F games?

Why E-F games are useful?

(1) Quantifier rank

Definition 3.8 (Quantifier rank). *The* quantifier rank *of a formula* $\mathsf{qr}(\varphi)$ *is its depth of quantifier nesting. That is:*

- *If* φ *is atomic, then* $\mathsf{qr}(\varphi) = 0$.
- $\mathsf{qr}(\varphi_1 \vee \varphi_2) = \mathsf{qr}(\varphi_1 \wedge \varphi_2) = \max(\mathsf{qr}(\varphi_1), \mathsf{qr}(\varphi_2))$.
- $\mathsf{qr}(\neg\varphi) = \mathsf{qr}(\varphi)$.
- $\mathsf{qr}(\exists x\varphi) = \mathsf{qr}(\forall x\varphi) = \mathsf{qr}(\varphi) + 1$.

We use the notation ~~FO[k]~~ FO_k *for all* FO *formulae of quantifier rank up to* k.

Examples:

* $qr(\exists_x \forall_y \exists_z R(x,y,z)) = 3$
* $qr(\exists_x A(x) \wedge (\forall_y R(x,y)) \vee (\exists z\, Q(z))) = 2$
* Let $\varphi_0(x,y) = E(x,y)$, $\varphi_{n+1} := \exists z\ \varphi_n(x,z) \wedge \varphi_n(z,y)$

  φ_n has $2^n - 1$ quantifiers but $qr(\varphi_n) = n$.
* For φ in prenex-normal-form $qr = \#$quantifiers

! Theorem. $\mathfrak{A} \equiv_m \mathfrak{B}$ iff for all $\varphi \in FO_m$ we have $\mathfrak{A} \models \varphi \Leftrightarrow \mathfrak{B} \models \varphi$.

Proof: todo.

(2) How to use them? E.g. we want to show that parity is not FO[$\emptyset$]-definable.

structures are just sets

Ad absurdum, assume that such a formula exists. Call it φ and let $qr(\varphi) = m$.

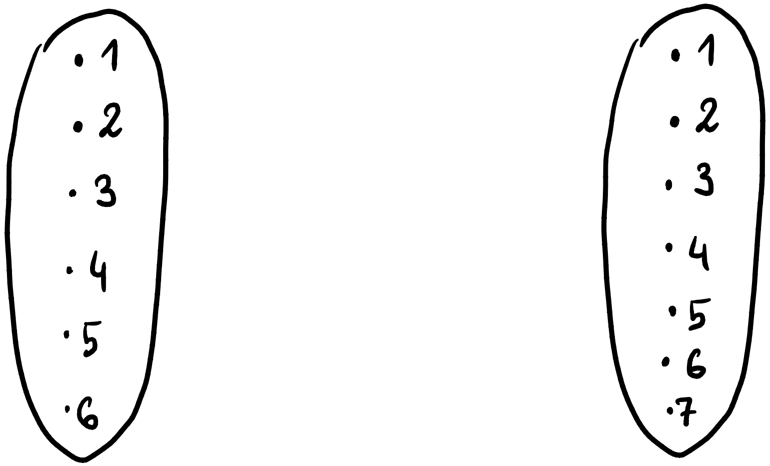
Then find $\mathfrak{A}_m$ and $\mathfrak{B}_m$ such that $\mathfrak{A}_m \equiv_m \mathfrak{B}_m$ and $\underbrace{\mathfrak{A}_m \models \varphi}_{\text{of even size}}$ and $\underbrace{\mathfrak{B} \not\models \varphi.}_{\text{of odd size}}$

A contradiction with the fact that $\varphi \in FO_m$. $\lightning$

(2) How to use them? E.g. we want to show that parity is not FO[$\emptyset$]-definable.

Ad absurdum, assume that such a formula exists. Call it φ and let $qr(\varphi) = m$.
Then find $\mathcal{A}_m$ and $\mathcal{B}_m$ such that $\mathcal{A}_m \equiv_m \mathcal{B}_m$ and $\underbrace{\mathcal{A}_m \models \varphi}_{\text{of even size}}$ and $\underbrace{\mathcal{B} \not\models \varphi}_{\text{of odd size}}$.
A contradiction with the fact that $\varphi \in FO_m$. $\lightning$

Take $\mathcal{A}_m := \{1, 2, ..., 2m\}$ and $\mathcal{B}_m := \{1, 2, 3, ..., 2m+1\}$.

$\mathcal{A}_3 :=$ 

$=: \mathcal{B}_3$

Winning strategy for $\forall$: just take a fresh element each round (or the same if the $\exists$ picks the same)

! Lemma: If $\mathcal{A}$ and $\mathcal{B}$ are sets with $|A| \geq m$ and $|B| \geq m$ then $\mathcal{A} \equiv_m \mathcal{B}$.

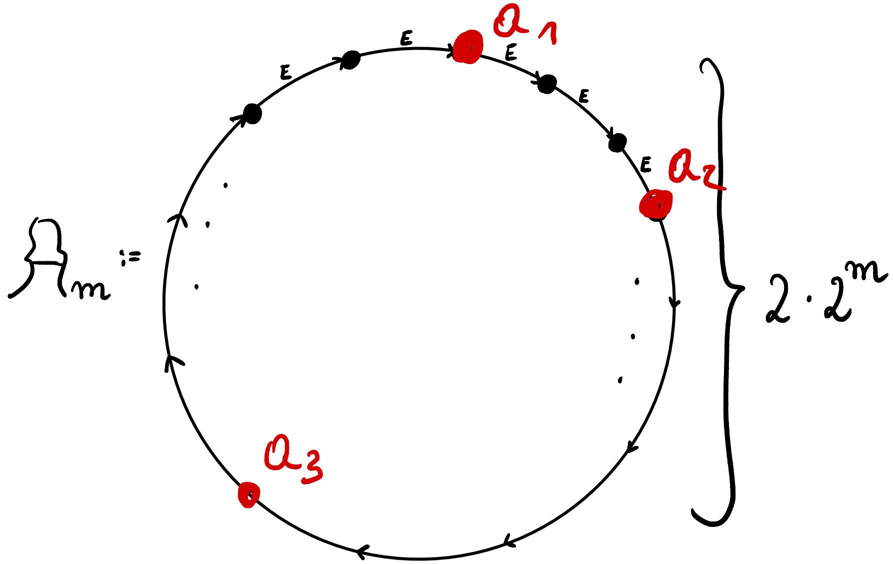
(2') E.g. we want to show that connectivity is not FO[{E}]-definable.

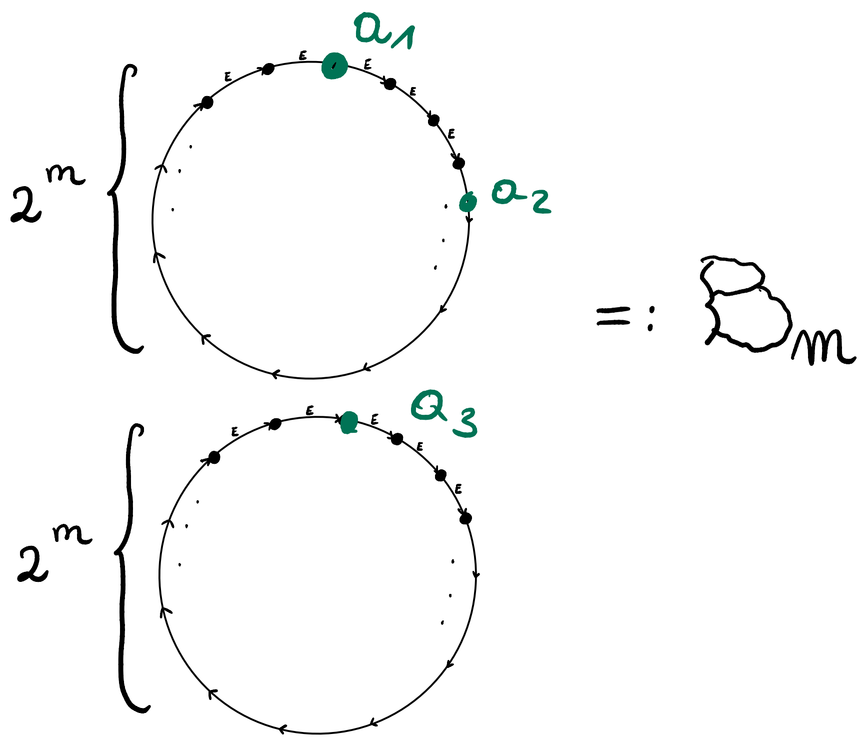
Ad absurdum, assume that such a formula exists. Call it φ and let $qr(\varphi) = m$.

Then find $\mathfrak{A}_m$ and $\mathfrak{B}_m$ such that $\mathfrak{A}_m \equiv_m \mathfrak{B}_m$ and $\underbrace{\mathfrak{A}_m \models \varphi}_{\text{is connected}}$ and $\underbrace{\mathfrak{B} \not\models \varphi}_{\text{is not connected}}$.

A contradiction with the fact that $\varphi \in FO_m$. ↯

Take $\mathfrak{A}_m$ and $\mathfrak{B}_m$ to be:





Game idea: play close if spoiler plays close
otherwise play far.

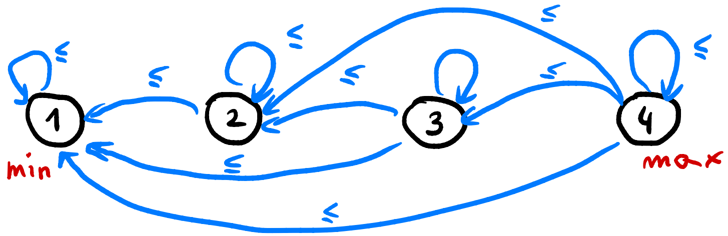
Proof: Next time.

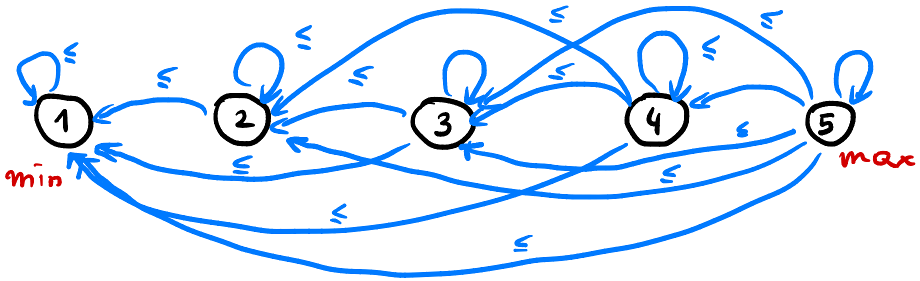
(3) E.g. we want to show that evenness of linear orders is not FO[<]-definable.

Ad absurdum, assume that such a formula exists. Call it φ and let $qr(\varphi) = m$.
Then find $\mathfrak{A}_m$ and $\mathfrak{B}_m$ such that $\mathfrak{A}_m \equiv_m \mathfrak{B}_m$ and $\underbrace{\mathfrak{A}_m \models \varphi}_{\text{is even}}$ and $\underbrace{\mathfrak{B} \not\models \varphi.}_{\text{is odd}}$
A contradiction with the fact that $\varphi \in FO_m$. ↯

Take $\mathfrak{A}_m = 1 - 2 - 3 - \dots - 2m-1 - 2m$ and $\mathfrak{B}_m := 1-2-3-\dots-2m-2m+1$

$\mathfrak{A}_2$





$\mathfrak{B}_2$

To show that $\mathfrak{A}_m \equiv_m \mathfrak{B}_m$, we show a stronger result:

Theorem: If $\mathfrak{A}$ and $\mathfrak{B}$ are $\{min, max, \leq\}$-structures, where $\leq$ is interpreted as a linear order over the domain, min and max are constant symbols interpreted as the first and the last element w.r.t. $\leq$, then $|A| \geq 2^m$ and $|B| \geq 2^m$ implies $\mathfrak{A} \equiv_m \mathfrak{B}$.

!

Theorem: If $\mathfrak{A}$ and $\mathfrak{B}$ are $\{min, max, \leq\}$ - structures, where $\leq$ is interpreted as a linear order over the domain, min and max are constant symbols interpreted as the first and the last element w.r.t. $\leq$, then $|A| \geq 2^m$ and $|B| \geq 2^m$ implies $\mathfrak{A} \equiv_m \mathfrak{B}$.

!

Proof:

Let $\vec{a} = (\underset{min^{\mathfrak{A}}}{a_{-1}}, \underset{max^{\mathfrak{A}}}{a_0}, a_1, \ldots, a_i)$ and $\vec{b} = (\underset{min^{\mathfrak{B}}}{b_{-1}}, \underset{max^{\mathfrak{B}}}{b_0}, b_1, b_2, \ldots, b_i)$

denote the result of the first i rounds of E-F games on $\mathfrak{A}$ and $\mathfrak{B}$.

We will show, inductively, that the duplicator ($\forall$) can survive m rounds by employing a strategy satisfying the following invariant:

For all $l, j \in [-1, i]$:

required to make $a_i \mapsto b_i$ partial isomorphism

(3) $a_j \leq a_l$ iff $b_j \leq b_l$

what if a_l and a_j lie "far"

(2) if $dist(a_j, a_l) \geq 2^{m-i}$ then $dist(b_j, b_l) \geq 2^{m-i}$

a_j and a_l lie "close"

(1) if $dist(a_j, a_l) < 2^{m-i}$ then $dist(a_j, a_l) = dist(b_j, b_l)$.

Recap :

$\vec{a} = (\underset{\min^{\mathfrak{A}}}{a_{-1}}, \underset{\max^{\mathfrak{A}}}{a_0}, a_1, \ldots, a_i)$ $\qquad$ $\vec{b} = (\underset{\min^{\mathfrak{B}}}{b_{-1}}, \underset{\max^{\mathfrak{B}}}{b_0}, b_1, b_2, \ldots, b_i)$

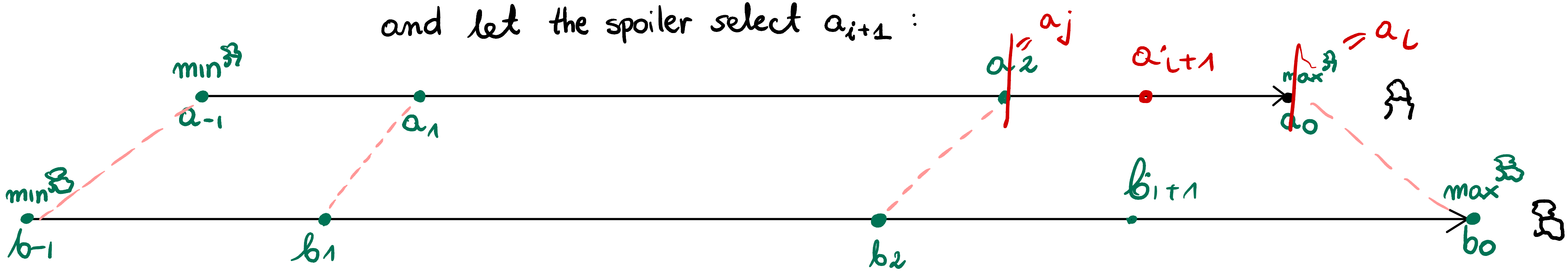
We show inductively that duplicator can play s.t after the i-th round we have :

③ $a_j \le a_\ell$ iff $b_j \le b_\ell$ $\qquad$ ② If $\underbrace{\text{dist}(a_j, a_\ell) \geq 2^{m-i}}_{\text{are "far"}}$ then $\underbrace{\text{dist}(b_j, b_\ell) \geq 2^{m-i}}_{\text{are "far"}}$

① If $\underbrace{\text{dist}(a_j, a_\ell) < 2^{m-i}}_{\text{are "close"}}$ then $\text{dist}\underbrace{(a_j, a_\ell) = \text{dist}(b_j, b_\ell)}_{\text{keep the distance equal if spoiler plays close}}$.

Base case $i = 0$ easy ◡̈ , so assume that $\forall$ survived i-rounds and ① – ③ hold.

and let the spoiler select a_{i+1} :



There are two options :

(1) Spoiler picks a_{i+1} equal to some of $\{a_{-1}, a_0, a_1, \ldots, a_i\}$ ← boring

(2) Spoiler picks a_{i+1} to be precisely between some a_j, a_ℓ.
By ③ we need to respond with $\underline{b_{i+1}}$ between b_j and b_ℓ.

Recap :

$\vec{a} = (\overset{\min^A}{a_{-1}}, \overset{\max^A}{a_0}, a_1, \ldots, a_i)$ $\qquad$ $\vec{b} = (\overset{\min^B}{b_{-1}}, \overset{\max^B}{b_0}, b_1, b_2, \ldots, b_i)$

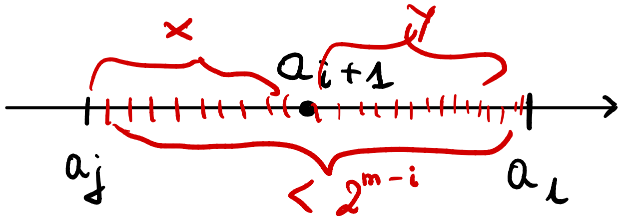
We show inductively that duplicator can play s.t after the i-th round we have :

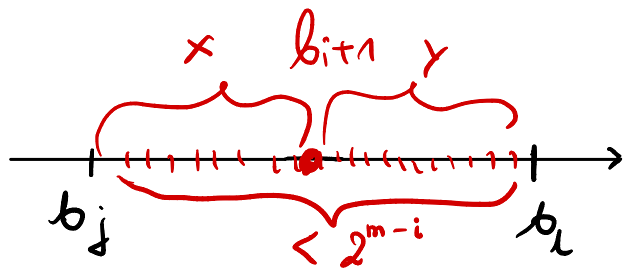
③ $a_j \leq a_l$ iff $b_j \leq b_l$ ② If $\underbrace{\text{dist}(a_j, a_l) \geq 2^{m-i}}_{\text{are "far"}}$ then $\underbrace{\text{dist}(b_j, b_l) \geq 2^{m-i}}_{\text{are "far"}}$

① If $\underbrace{\text{dist}(a_j, a_l) < 2^{m-i}}_{\text{are "close"}}$ then $\underbrace{\text{dist}(a_j, a_l) = \text{dist}(b_j, b_l)}_{\text{keep the distance equal if spoiler plays close}}$.

We need to consider two cases :

1° $\text{dist}(a_j, a_l) < 2^{m-i}$





By ① we know ...

2°

Recap: $\vec{a} = (\overset{\min^{A}}{a_{-1}}, \overset{\max^{A}}{a_0}, a_1, \ldots, a_i)$ $\qquad$ $\vec{b} = (\overset{\min^{B}}{b_{-1}}, \overset{\max^{B}}{b_0}, b_1, b_2, \ldots, b_i)$

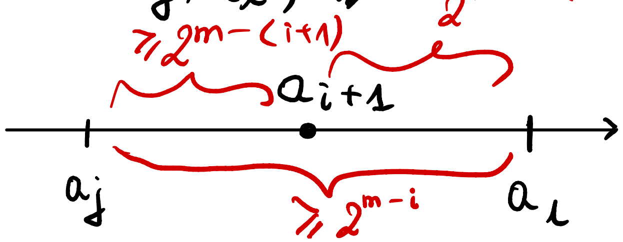
We show inductively that duplicator can play s.t after the i-th round we have:

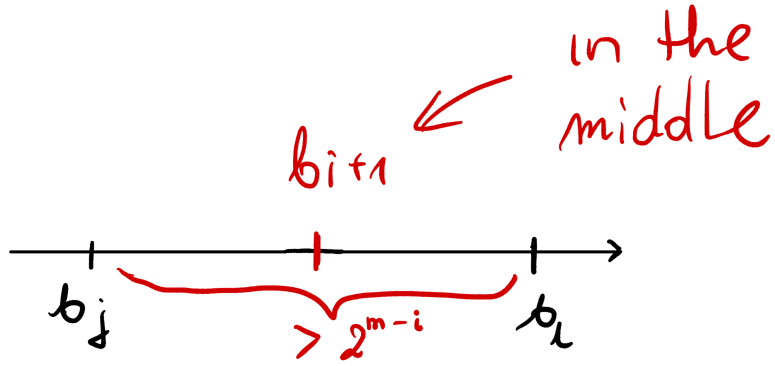
③ $a_j \leq a_\ell$ iff $b_j \leq b_\ell$ $\quad$ ② If $\underbrace{\text{dist}(a_j, a_\ell) \geq 2^{m-(i+1)}}_{\text{are "far"}}$ then $\underbrace{\text{dist}(b_j, b_\ell) \geq 2^{m-(i+1)}}_{\text{are "far"}}$

① If $\underbrace{\text{dist}(a_j, a_\ell) < 2^{m-(i+1)}}_{\text{are "close"}}$ then $\text{dist}(\underbrace{a_j, a_\ell) = \text{dist}(b_j, b_\ell}_{\text{keep the distance equal if spoiler plays close}})$.

We need to consider two cases:

2° $\text{dist}(a_j, a_\ell) \geq \cancel{2^{m-i}}\ 2^{m-(i+1)}$





Is ??? $\geq 2^{m-(i+1)}$?

Yes. Bisection.

No. Equal cutting.

3.2 Definition and Examples of Ehrenfeucht-Fraïssé Games

Ehrenfeucht-Fraïssé games give us a nice tool for describing expressiveness of logics over finite models. In general, games are applicable for both finite and infinite models (at least for FO), but we have seen that in the infinite case we have a number of more powerful tools. In fact, in some model theory texts, Ehrenfeucht-Fraïssé games are only briefly mentioned (or even appear only as exercises), but in the finite case, their applicability makes them a central notion.

The idea of the game – for FO and other logics as well – is almost invariably the same. There are two players, called the *spoiler* and the *duplicator* (or, less imaginatively, player I and player II). The board of the game consists of two structures, say $\mathfrak{A}$ and $\mathfrak{B}$. The goal of the spoiler is to show that these two structures are different; the goal of the duplicator is to show that they are the same.

In the classical Ehrenfeucht-Fraïssé game, the players play a certain number of rounds. Each round consists of the following steps:

1. The spoiler picks a structure ($\mathfrak{A}$ or $\mathfrak{B}$).

2. The spoiler makes a move by picking an element of that structure: either $a \in \mathfrak{A}$ or $b \in \mathfrak{B}$.
3. The duplicator responds by picking an element in the other structure.

An illustration is given in Fig. 3.1. The spoiler's moves are shown as filled circles, and the duplicator's moves as empty circles. In the first round, the spoiler picks $\mathfrak{B}$ and selects $b_1 \in \mathfrak{B}$; the duplicator responds by $a_1 \in \mathfrak{A}$. In the next round, the spoiler changes structures and picks $a_2 \in \mathfrak{A}$; the duplicator responds by $b_2 \in \mathfrak{B}$. In the third round the spoiler plays $b_3 \in \mathfrak{B}$; the response of the duplicator is $a_3 \in \mathfrak{A}$.

Since there is a game, someone must win it. To define the winning condition we need a crucial definition of a *partial isomorphism*. Recall that all finite structures have a relational vocabulary (no function symbols).

Definition 3.5 (Partial isomorphism). *Let $\mathfrak{A}$, $\mathfrak{B}$ be two σ-structures, where σ is relational, and $\vec{a} = (a_1, \ldots, a_n)$ and $\vec{b} = (b_1, \ldots, b_n)$ two tuples in $\mathfrak{A}$ and $\mathfrak{B}$ respectively. Then $(\vec{a}, \vec{b})$ defines a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ if the following conditions hold:*

- *For every $i, j \leq n$,*
$$a_i = a_j \quad \textit{iff} \quad b_i = b_j.$$
- *For every constant symbol c from σ, and every $i \leq n$,*
$$a_i = c^{\mathfrak{A}} \quad \textit{iff} \quad b_i = c^{\mathfrak{B}}.$$
- *For every k-ary relation symbol P from σ and every sequence $(i_1, \ldots, i_k)$ of (not necessarily distinct) numbers from $[1, n]$,*
$$(a_{i_1}, \ldots, a_{i_k}) \in P^{\mathfrak{A}} \quad \textit{iff} \quad (b_{i_1}, \ldots, b_{i_k}) \in P^{\mathfrak{B}}.$$

In the absence of constant symbols, this definition says that the mapping $a_i \mapsto b_i, i \leq n$, is an isomorphism between the substructures of $\mathfrak{A}$ and $\mathfrak{B}$ generated by $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$, respectively.

After n rounds of an Ehrenfeucht-Fraïssé game, we have moves $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$. Let $c_1, \ldots, c_l$ be the constant symbols in σ; then $\vec{c}^{\,\mathfrak{A}}$ denotes $(c_1^{\mathfrak{A}}, \ldots, c_l^{\mathfrak{A}})$ and likewise for $\vec{c}^{\,\mathfrak{B}}$. We say that $(\vec{a}, \vec{b})$ is a *winning position for the duplicator* if
$$((\vec{a}, \vec{c}^{\,\mathfrak{A}}), \quad (\vec{b}, \vec{c}^{\,\mathfrak{B}}))$$
is a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$. In other words, the map that sends each a_i into b_i and each $c_j^{\mathfrak{A}}$ into $c_j^{\mathfrak{B}}$ is an isomorphism between the substructures of $\mathfrak{A}$ and $\mathfrak{B}$ generated by $\{a_1, \ldots, a_n, c_1^{\mathfrak{A}}, \ldots, c_l^{\mathfrak{A}}\}$ and $\{b_1, \ldots, b_n, c_1^{\mathfrak{B}}, \ldots, c_l^{\mathfrak{B}}\}$ respectively.

We say that the duplicator has an *n-round winning strategy in the Ehrenfeucht-Fraïssé game on $\mathfrak{A}$ and $\mathfrak{B}$* if the duplicator can play in a way

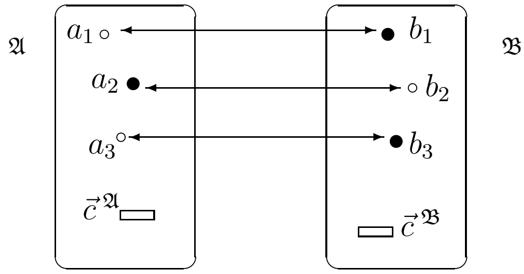


Fig. 3.1. Ehrenfeucht-Fraïssé game

that guarantees a winning position after n rounds, no matter how the spoiler plays. Otherwise, the spoiler has an n-round winning strategy. If the duplicator has an n-round winning strategy, we write $\mathfrak{A} \equiv_n \mathfrak{B}$.

Observe that $\mathfrak{A} \equiv_n \mathfrak{B}$ implies $\mathfrak{A} \equiv_k \mathfrak{B}$ for every $k \leq n$.

Before we connect Ehrenfeucht-Fraïssé games and FO-definability, we give some examples of winning strategies.

Games on Sets

In this example, the vocabulary σ is empty. That is, a structure is just a set. Let $|A|, |B| \geq n$. Then $A \equiv_n B$.

The strategy for the duplicator works as follows. Suppose i rounds have been played, and the position is $((a_1, \ldots, a_i), (b_1, \ldots, b_i))$. Assume the spoiler picks an element $a_{i+1} \in A$. If $a_{i+1} = a_j$ for $j \leq i$, then the duplicator responds with $b_{i+1} = b_j$; otherwise, the duplicator responds with any $b_{j+1} \in B - \{b_1, \ldots, b_i\}$ (which exists since $|B| \geq n$).

Games on Linear Orders

Our next example is a bit more complicated, as we add a binary relation $<$ to σ, to be interpreted as a linear order. Now suppose L_1, L_2 are two linear orders of size at least n (i.e., structures of the form $\langle \{1, \ldots, m\}, < \rangle$, $m \geq n$). Is it true that $L_1 \equiv_n L_2$?

It is very easy to see that the answer is negative even for the case of $n = 2$. Let L_1 contain three elements (say $\{1, 2, 3\}$), and L_2 two elements ($\{1, 2\}$). In the first move, the spoiler plays 2 in L_1. The duplicator has to respond with either 1 or 2 in L_2. Suppose the duplicator responds with $1 \in L_2$; then the spoiler plays $1 \in L_1$ and the duplicator is lost, since he has to respond with an element less than 1 in L_1, and there is no such element. If the duplicator selects $2 \in L_2$ as his first-round move, the spoiler plays $3 \in L_1$, and the duplicator is lost again. Hence, $L_1 \not\equiv_2 L_2$.

However, a winning strategy for the duplicator can be guaranteed if L_1, L_2 are much larger than the number of rounds.

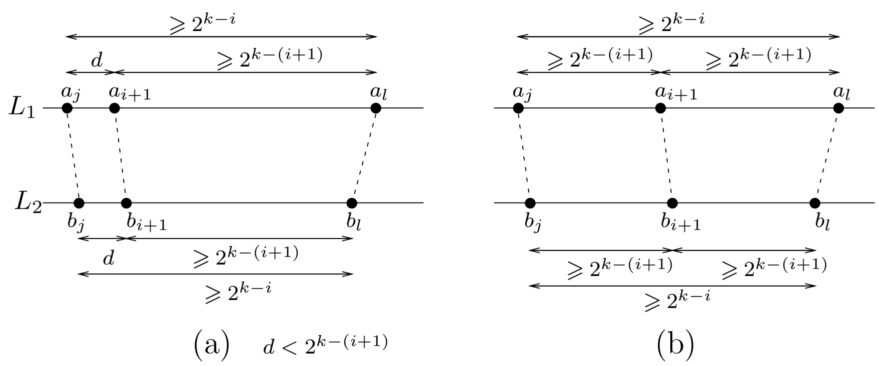


Fig. 3.2. Illustration for the proof of Theorem 3.6

Theorem 3.6. *Let $k > 0$, and let L_1, L_2 be linear orders of length at least 2^k. Then $L_1 \equiv_k L_2$.*

We shall give two different proofs of this result that illustrate two different techniques often used in game proofs.

Theorem 3.6, Proof # 1. The idea of the first proof is as follows. We use induction on the number of rounds of the game, and our induction hypothesis is *stronger* than just the partial isomorphism claim. The reason is that if we simply state that after i rounds we have a partial isomorphism, the induction step will not get off the ground as there are too few assumptions. Hence, we have to make additional assumptions. But if we try to impose too many conditions, there is no guarantee that a game can proceed in a way that preserves them. The main challenge in proofs of this type is to find the right induction hypothesis: the one that is strong enough to imply partial isomorphism, and that has enough conditions to make the inductive proof possible.

We now illustrate this general principle by proving Theorem 3.6. We expand the vocabulary with two new constant symbols $\underline{\min}$ and $\underline{\max}$, to be interpreted as the minimum and the maximum element of a linear ordering, and we prove a stronger fact that $L_1 \equiv_k L_2$ in the expanded vocabulary.

Let L_1 have the universe $\{1, \ldots, n\}$ and L_2 have the universe $\{1, \ldots, m\}$. Assume that the lengths of L_1 and L_2 are at least 2^k; that is, $n, m \geq 2^k+1$. The distance between two elements x, y of the universe, $d(x, y)$, is simply $|x - y|$. We claim that the duplicator can play in such a way that the following holds after each round i. Let $\vec{a} = (a_{-1}, a_0, a_1, \ldots, a_i)$ consist of $a_{-1} = \underline{\min}^{L_1}, a_0 = \underline{\max}^{L_1}$ and the i moves $a_1, \ldots, a_i$ in L_1, and likewise let $\vec{b} = (b_{-1}, b_0, b_1, \ldots, b_i)$ consist of $b_{-1} = \underline{\min}^{L_2}, b_0 = \underline{\max}^{L_2}$ and the i moves in L_2. Then, for $-1 \leq j, l \leq i$:

$$\begin{array}{ll} 1. \text{ if } d(a_j, a_l) < 2^{k-i}, \text{ then } d(b_j, b_l) = d(a_j, a_l). \\ 2. \text{ if } d(a_j, a_l) \geq 2^{k-i}, \text{ then } d(b_j, b_l) \geq 2^{k-i}. \\ 3. \ a_j \leq a_l \iff b_j \leq b_l. \end{array} \tag{3.2}$$

We prove (3.2) by induction; notice that the third condition ensures partial isomorphism, so we do prove an induction statement that says more than just maintaining partial isomorphism.

And now a simple proof: the base case of $i = 0$ is immediate since $d(a_{-1}, a_0), d(b_{-1}, b_0) \geq 2^k$ by assumption. For the induction step, suppose the spoiler is making his $(i+1)$st move in L_1 (the case of L_2 is symmetric). If the spoiler plays one of $a_j, j \leq i$, the response is b_j, and all the conditions are trivially preserved. Otherwise, the spoiler's move falls into an interval, say $a_j < a_{i+1} < a_l$, such that no other previously played moves are in the same interval. By condition 3 of (3.2), this means that the interval between b_j and b_l contains no other elements of $\vec{b}$. There are two cases:

- $d(a_j, a_l) < 2^{k-i}$. Then $d(b_j, b_l) = d(a_j, a_l)$, and the intervals $[a_j, a_l]$ and $[b_j, b_l]$ are isomorphic. Then we simply find b_{i+1} so that $d(a_j, a_{i+1}) = d(b_j, b_{i+1})$ and $d(a_{i+1}, a_l) = d(b_{i+1}, b_l)$. Clearly, this ensures that all the conditions in (3.2) hold.
- $d(a_j, a_l) \geq 2^{k-i}$. In this case $d(b_j, b_l) \geq 2^{k-i}$. We have three possibilities:
  1. $d(a_j, a_{i+1}) < 2^{k-(i+1)}$. Then $d(a_{i+1}, a_l) \geq 2^{k-(i+1)}$, and we can choose b_{i+1} so that $d(b_j, b_{i+1}) = d(a_j, a_{i+1})$ and $d(b_{i+1}, b_l) \geq 2^{k-(i+1)}$. This is illustrated in Fig. 3.2 (a), where d stands for $d(a_j, a_{i+1})$.
  2. $d(a_{i+1}, a_l) < 2^{k-(i+1)}$. This case is similar to the previous one.
  3. $d(a_j, a_{i+1}) \geq 2^{k-(i+1)}$, $d(a_{i+1}, a_l) \geq 2^{k-(i+1)}$. Since $d(b_j, b_l) \geq 2^{k-i}$, by choosing b_{i+1} to be the middle of the interval $[b_j, b_l]$ we ensure that $d(b_j, b_{i+1}) \geq 2^{k-(i+1)}$ and $d(b_{i+1}, b_l) \geq 2^{k-(i+1)}$. This case is illustrated in Fig. 3.2 (b).

Thus, in all the cases, (3.2) is preserved.

This completes the inductive proof; hence we have shown that the duplicator can win a k-round Ehrenfeucht-Fraïssé game on L_1 and L_2. $\square$

3.3 Games and the Expressive Power of FO

And now it is time to see why games are important. For this, we need a crucial definition of quantifier rank.

Definition 3.8 (Quantifier rank). *The* quantifier rank *of a formula* $\mathsf{qr}(\varphi)$ *is its depth of quantifier nesting. That is:*

- *If* φ *is atomic, then* $\mathsf{qr}(\varphi) = 0$.
- $\mathsf{qr}(\varphi_1 \vee \varphi_2) = \mathsf{qr}(\varphi_1 \wedge \varphi_2) = \max(\mathsf{qr}(\varphi_1), \mathsf{qr}(\varphi_2))$.
- $\mathsf{qr}(\neg\varphi) = \mathsf{qr}(\varphi)$.
- $\mathsf{qr}(\exists x\varphi) = \mathsf{qr}(\forall x\varphi) = \mathsf{qr}(\varphi) + 1$.

We use the notation FO$[k]$ *for all* FO *formulae of quantifier rank up to* k.

In general, quantifier rank of a formula is different from the total of number of quantifiers used. For example, we can define a family of formulae by induction: $d_0(x, y) \equiv E(x, y)$, and $d_k \equiv \exists z\ d_{k-1}(x, z) \wedge d_{k-1}(z, y)$. The quantifier rank of d_k is k, but the total number of quantifiers used in d_k is $2^k - 1$. For formulae in the prenex form (i.e., all quantifiers are in front, followed by a quantifier-free formula), quantifier rank is the same as the total number of quantifiers.

Given a set S of FO sentences (over vocabulary σ), we say that two σ-structures $\mathfrak{A}$ and $\mathfrak{B}$ *agree on* S if for every sentence Φ of S, it is the case that $\mathfrak{A} \models \Phi \Leftrightarrow \mathfrak{B} \models \Phi$.

Theorem 3.9 (Ehrenfeucht-Fraïssé). *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be two structures in a relational vocabulary. Then the following are equivalent:*

1. $\mathfrak{A}$ *and* $\mathfrak{B}$ *agree on* FO$[k]$.
2. $\mathfrak{A} \equiv_k \mathfrak{B}$.

We will prove this theorem shortly, but first we discuss how this is useful for proving inexpressibility results.

Characterizing the expressive power of FO via games gives rise to the following methodology for proving inexpressibility results.

Corollary 3.10. *A property* $\mathcal{P}$ *of finite* σ*-structures is* not expressible *in* FO *if for every* $k \in \mathbb{N}$, *there exist two finite* σ*-structures,* $\mathfrak{A}_k$ *and* $\mathfrak{B}_k$, *such that:*

- $\mathfrak{A}_k \equiv_k \mathfrak{B}_k$, *and*
- $\mathfrak{A}_k$ *has property* $\mathcal{P}$, *and* $\mathfrak{B}_k$ *does not.*

Proof. Assume to the contrary that $\mathcal{P}$ is definable by a sentence Φ. Let $k = \mathsf{qr}(\Phi)$, and pick $\mathfrak{A}_k$ and $\mathfrak{B}_k$ as above. Then $\mathfrak{A}_k \equiv_k \mathfrak{B}_k$, and thus if $\mathfrak{A}_k$ has property $\mathcal{P}$, then so does $\mathfrak{B}_k$, which contradicts the assumptions. $\square$

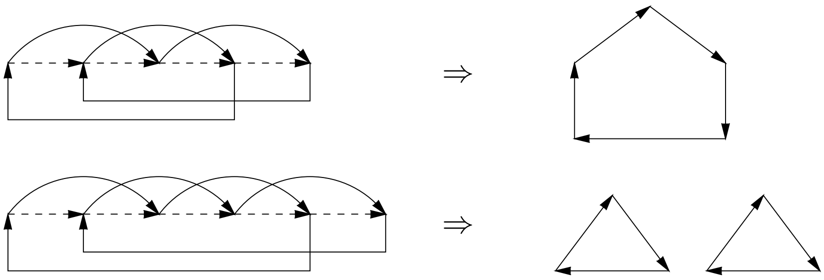


Fig. 3.3. Reduction of parity to connectivity

3.6 More Inexpressibility Results

So far we have used games to prove that EVEN is not expressible in FO, in both ordered and unordered settings. Next, we show inexpressibility of graph connectivity over finite graphs. In Sect. 3.1 we used compactness to show that connectivity of arbitrary graphs is inexpressible, leaving open the possibility that it may be FO-definable over finite graphs. We now show that this cannot happen. It turns out that no new game argument is needed, as the proof uses a reduction from EVEN over linear orders.

Assume that connectivity of finite graphs is definable by an FO sentence Φ, in the vocabulary that consists of one binary relation symbol E. Next, given a linear ordering, we define a directed graph from it as described below. First, from a linear ordering $<$ we define the successor relation

$$\mathrm{succ}(x, y) \ \equiv \ (x < y) \wedge \forall z\big((z \leq x) \vee (z \geq y)\big).$$

Using this, we define an FO formula $\gamma(x, y)$ such that $\gamma(x, y)$ is true iff one of the following holds:

- y is the successor of the successor of x: $\exists z \ \big(\mathrm{succ}(x, z) \wedge \mathrm{succ}(z, y)\big)$, or
- x is the predecessor of the last element, and y is the first element: $\big(\exists z \ (\mathrm{succ}(x, z) \wedge \forall u(u \leq z))\big) \wedge \forall u(y \leq u)$, or
- x is the last element and y is the successor of the first element (the FO formula is similar to the one above).

Thus, $\gamma(x, y)$ defines a new graph on the elements of the linear ordering; the construction is illustrated in Fig. 3.3.

Now observe that the graph defined by γ is connected iff the size of the underlying linear ordering is odd. Hence, taking $\neg\Phi$, and substituting γ for every occurrence of the predicate E, we get a sentence that tests EVEN for linear orderings. Since this is impossible, we obtain the following.

Corollary 3.19. *Connectivity of finite graphs is not* FO*-definable.*