Logic, descriptive complexity and theory of databases. Lecture 3

Lecturer: Luc Segoufin
Scribe: Jérémie Dimino

October 4, 2010

Exercise from last week:

Example 1. *Connectivity is in ESO.*

Proof. Follows immediately from:

$$\begin{aligned} &\exists T, O \quad O \text{ is a linear order} \\ &\qquad \wedge T(i,x,y) \text{ iff there is a path of length } i \text{ from } x \text{ to } y\text{:} \\ &\qquad\quad \exists u \quad \text{“}u \text{ is the minimum of } O\text{”} \wedge \forall x, y E(x,y) \leftrightarrow T(0,x,y) \\ &\qquad\quad \forall x, y, u \quad (\exists z E(x,z) \wedge T(u,z,y)) \leftrightarrow T(u+1,x,y) \end{aligned}$$

□

1 Expressive power of FO

1.1 Ehrenfeucht-Fraïssé games (EF-games)

Definition 1 (partial isomorphism)**.** *Given:*

- *$\mathcal{A}$, $\mathcal{B}$ two σ-structures.*
- $\overline{a} = (a_1, \ldots, a_n) \subseteq domain(\mathcal{A})$
- $\overline{b} = (b_1, \ldots, b_n) \subseteq domain(\mathcal{B})$

$h : \overline{a} \to \overline{b}$ *(where $h(a_i) = b_i$) is a partial isomorphism iff:*

$$\begin{cases} a_i = a_j \textit{ iff } b_i = b_j \\ \textit{and} \\ R(a_{i_1}, \ldots, a_{i_k}) \textit{ iff } R(b_{i_1}, \ldots, b_{i_k}) \textit{ for all relations } R \in \sigma \\ a_i = c \textit{ iff } b_i = c \textit{ for all constants } c \in \sigma \end{cases}$$

We define the EF-game $EF_n(\mathcal{A}, \mathcal{B})$ by the game where:

- the board is the disjoint union of the two structures $\mathcal{A}$ and $\mathcal{B}$,
- there are two players: the Duplicator and the Spoiler,
- there are n rounds where round i is defined as follow:
  - The Spoiler selects either $a_i \in \mathcal{A}$ or $b_i \in \mathcal{B}$,
  - the duplicator respond in the opposite structure by choosing $b_i \in \mathcal{B}$ in the first case or $a_i \in \mathcal{A}$ in the second case.

In the end $(a_1, \dots, a_n)$ is selected in $\mathcal{A}$ and $(b_1, \dots, b_n)$ is selected in $\mathcal{B}$. The Duplicator wins the game if there is a partial isomorphism $h : \overline{a} \to \overline{b}$.

Definition 2. *The Duplicator has a winning strategy in $EF_n(\mathcal{A}, \mathcal{B})$ if no matter what the Spoiler plays, he wins. In this case we write $\mathcal{A} \equiv_n \mathcal{B}$.*

Definition 3. *Similarly we define $\mathcal{A}, \bar{a} \equiv_n \mathcal{B}, \bar{b}$, where $\bar{a}$ and $\bar{b}$ are tuples of elements of $\mathcal{A}$ and $\mathcal{B}$ respectively, if Duplicator has a winning strategy in $EF_n(\mathcal{A}', \mathcal{B}')$, where $\mathcal{A}'$ is $\mathcal{A}$ with new constants denoting $\bar{a}$ (similarly for $\mathcal{B}'$).*

Example 2. *Take $\sigma = \emptyset$ and S_n the set of size n. Then it is easy to see that we have for all n:*

$$S_n \equiv_n S_{n+1}$$

Example 3. *Take $\sigma = \{E\}$ (a graph) and C_n the circle of size n. Then we have for all n:*

$$C_{2^n} \equiv_n C_{2^n+1}$$

Proof: fix an orientation for both cycles. This define a cyclic order between the elements of the cycles. Play while maintaining the following invariant at round i:

- *The elements $a_1, \cdots, a_i$ appear in the same cyclic order as the elements of $b_1, \cdots b_i$.*
- *If a_α is close to a_β then the distance between b_α and b_β is the same as the distance between a_α and a_β. Here "close" means less than 2^{n-i}.*
- *If b_α is close to b_β then the distance between a_α and a_β is the same as the distance between b_α and b_β. Here "close" means less than 2^{n-i}.*

Example 4. *Take $\sigma = \{<, P_a\}$ (words with one letter) and W_n the word of length n. Then we have for all n:*

$$W_{2^n} \equiv_n W_{2^n+1}$$

Proof: same as for Example 3

Definition 4 (quantifier rank)**.** *Given three first-order formula φ, φ_1 and φ_2, we define q_r inductively as follow:*

- $q_r(\varphi) = 0$ *if φ has no quantifier*
- $q_r(\varphi_1 \wedge \varphi_2) = max\{q_r(\varphi_1), q_r(\varphi_2)\}$
- $q_r(\varphi_1 \vee \varphi_2) = max\{q_r(\varphi_1), q_r(\varphi_2)\}$
- $q_r(\exists x.\varphi) = q_r(\varphi) + 1$
- $q_r(\forall x.\varphi) = q_r(\varphi) + 1$
- $q_r(\neg\varphi) = q_r(\varphi)$

Theorem 1. *We have:*

- $\mathcal{A} \equiv_n \mathcal{B}$ *iff* $\forall \varphi, q_r(\varphi) \leq n \Rightarrow (\mathcal{A} \vDash \varphi \Leftrightarrow \mathcal{B} \vDash \varphi)$
- $\mathcal{A}, \overline{a} \equiv_n \mathcal{B}, \overline{b}$ *iff* $\forall \varphi(\overline{x}), q_r(\varphi) \leq n \Rightarrow (\mathcal{A} \vDash \varphi(\overline{a}) \Leftrightarrow \mathcal{B} \vDash \varphi(\overline{b})$

Proof. : ($\Rightarrow$) assume $\exists \varphi \quad q_r(\varphi) \leq n$, $\mathcal{A} \vDash \varphi$ and $\mathcal{B} \vDash \neg\varphi$. We show that Spoiler has a winning strategy.

- if $\varphi = \varphi_1 \vee \varphi_2$ then for $\delta = 1$ or 2, $\mathcal{A} \vDash \varphi_\delta$ and $\mathcal{B} \vDash \neg\varphi_\delta$ and we continue with a smaller formula

- if $\varphi = \varphi_1 \wedge \varphi_2$ then for $\delta = 1$ or 2, $\mathcal{A} \vDash \varphi_\delta$ and $\mathcal{B} \vDash \neg\varphi_\delta$ and we continue with a smaller formula
- if $\varphi = \neg\varphi_1$ then we swap the role of $\mathcal{A}$ and $\mathcal{B}$ and continue with a smaller formula
- if φ is $\exists x.\varphi_1$ then the Spoiler selects a in $\mathcal{A}$ such that $\mathcal{A} \vDash \varphi_1(a)$. No matter what the Duplicator selects, we have $\mathcal{B} \vDash \neg\varphi_1(b)$.

In the end we have selected $a_1 \dots a_i$ and $b_1 \dots b_i$ with $i \leq n$ and we have $\mathcal{A} \vDash \Psi(a_1 \dots a_i)$ and $\mathcal{B} \vDash \neg\Psi(b_1 \dots b_i)$ where Ψ is an atom. This contradict the existence of an partial isomorphism and therefore Spoiler wins.

($\Leftarrow$) Harder, not done during the lecture, not used in this class. $\square$

Corollary 1. *Let P be a property of σ-structures. $P \notin FO$ if $\forall n \in \mathbb{N}.\exists \mathcal{A}_n, \mathcal{B}_n$ such that:*

1 $\mathcal{A}_n \equiv_n \mathcal{B}_n$

2 $\mathcal{A}_n \vDash P$ *and* $\mathcal{B}_n \vDash \neg P$

Proof. Assume the right-hand side. If $P \in FO$, let φ define P. Let n be $q_r(\varphi)$.

- $\mathcal{A}_n \equiv_n \mathcal{B}_n \Rightarrow \mathcal{A}_n \vDash \varphi \Leftrightarrow \mathcal{B}_n \vDash \varphi$
- But φ defines P and second item implies $\mathcal{A}_n \vDash \varphi$ and $\mathcal{B}_n \vDash \neg\varphi$

Contradiction hence $P \notin FO$. $\square$

<u>Applications:</u>

- parity of sets is not in FO:

  Take $\mathcal{A}_n = S_{2n}$ and $\mathcal{B}_n = S_{2n+1}$. We have seen that $S_{2n} \equiv_n S_{2n+1}$. Hence the result follows from Corollary 1.
- $(aa)^* \notin FO$

  Take $\mathcal{A}_n = W_{2^n} \in (aa)^*$ and $\mathcal{B}_n = W_{2^n+1} \notin (aa)^*$. We have seen that $W_{2^n} \equiv_n W_{2^n+1}$. Hence the result follows from Corollary 1.
- 2-colorability is not in FO

  A cycle is 2-colorable iff it has even length. Take $\mathcal{A}_n = C_{2^n}$, $\mathcal{B}_n = C_{2^n+1}$. We have seen $C_{2^n} \equiv_n C_{2^n+1}$. Hence the result follows from Corollary 1.

Exercice 1. *Show that:*

- *planarity is not in FO*
- *halmitonicity is not in FO*
- *connectivity is not in FO*
- *being a tree is not in FO*

1.2 Locality:

σ relational (no functions).

Let I be a structure.

Definition 5. *$G(I)$ is the Gaifman graph of I, defined as follow:*

- *the vertices are the elements of I*
- *there is an edge (a, b) iff a and b appear in the same tuple of a relation of I.*

Definition 6. *In I, $d(a, b)$ is the distance between a and b in $G(I)$.*

$N_k^I(\overline{a})$ is the k-neighborhood of $\overline{a}$ in I, where $\overline{a}$ is a tuple of elements of I and $k \in \mathbb{N}$, is defined as the substructure of I induced by the elements of I at distance less than k from $\overline{a}$ follow: $N_k^I(\overline{a}) := I/_{\{b\ :\ d(b,\overline{a}) \leq k\}}$.

We write $N_k^I(\overline{a}) \sim N_k^J(\overline{b})$ to say that the two neighborhoods are isomorphic. That is there is an isomorphism h:

$$\begin{aligned} \exists h : N_k^I(\overline{a}) &\to N_k^J(\overline{b}) \\ \overline{a} &\mapsto \overline{b} \end{aligned}$$

Definition 7. *I and J are two σ-structures. $\overline{a} \in I$ and $\overline{b} \in J$.*

$$\begin{array}{rl} & (I, \overline{a}) \rightleftharpoons_k (J, \overline{b}) \\ & \textit{iff} \\ \exists h\ \textit{(bijection)}\ : I & \to J \\ \overline{a} & \mapsto \overline{b} \\ \forall c & N_k^I(\overline{a}c) \sim N_k^J(\overline{b}h(c)) \end{array}$$

Theorem 2 (Hanf locality)**.** *$\forall \varphi \in FO, \exists k$ such that $\forall I, J$*

$$(I, \overline{a}) \rightleftharpoons_k (J, \overline{b}) \qquad \textit{implies} \qquad I \vDash \varphi(\overline{a}) \Leftrightarrow J \vDash \varphi(\overline{b}) \ .$$

Proof. Next week. $\square$

Corollary 2. *Let P be a property of σ-structures. $P \notin FO$ if $\forall k \in \mathbb{N}.\exists \mathcal{A}_k, \mathcal{B}_k$ such that:*

1 $\mathcal{A}_k \rightleftharpoons_k \mathcal{B}_k$

2 $\mathcal{A}_k \vDash P$ and $\mathcal{B}_k \vDash \neg P$

Proof. If P is in FO, let φ defining it. Let k be the number for φ given by Theorem 2. If we have $\mathcal{A}_k \rightleftharpoons_k \mathcal{B}_k$ then by Theorem 2 we have $\mathcal{A} \vDash \varphi \Leftrightarrow \mathcal{B} \vDash \varphi$. Hence we cannot have $\mathcal{A}_k \vDash P$ and $\mathcal{B}_k \vDash \neg P$. $\square$

Application: connectivity is not in FO.

Take $\mathcal{A} = C_{2k+2}$, the cycle of length $2k + 2$ and $\mathcal{B} = C_{k+1} \cup C_{k+1}$. It is easy to see that $\mathcal{A}_k \rightleftharpoons_k \mathcal{B}_k$.