LGruDat: Logical Foundations of Databases
Lecture 5 — Characterizing FORC-definability, part I

Tadeusz Litak

November 29, 2013

Elementarity (in the finite): overview and statement of the problem

- Recall what we know so far about EC and EC_{Δ}
- Introduce the notation EC^{fin}, remember $EC^{\mathsf{fin}}_{\Delta}$ is trivialized
- Showing that something is EC^{fin}: just write this sentence. But how we can show the failure of being EC^{fin}?
- In homework you did connectedness, but on arbitrary structures

Those few sweetspots

- Sometimes we are lucky on finite structures with no additional apparatus: Libkin's example of evenness in empty signature
- For this purpose, another corollary of completeness proof technique not mentioned so far: the Löwenheim-Skolem Theorem
- In passing: information about the Lindström Theorem
- But what if signature is non-empty? And connectedness does not seem any closer ...

Towards solution: stratification and quantifier rank

- One idea: stratify formulas, e.g. wrt quantifier rank. Introduce the notion of $\mathsf{FORC}[m]$

- Note: do not confuse with quantifier alternation!
- If a property is not definable (on Fin) up to any finite quantifier rank, it cannot be EC (EC^{fin})
- Observe the same applies to FORC**-queries**!

- One more notational convention for satisfaction. Recall a ***query*** ϕ is of the form $\{\overline{\mathbf{v}} \mid \alpha\}$, where $v \in free(\alpha)$ iff $v \in \overline{\mathbf{v}}$. Write

$$\mathfrak{A} \models \phi[\overline{\mathsf{a}}] \text{ iff } \overline{\mathsf{a}} \in \phi(\mathfrak{A})$$

- Syntactic abbreviation: identify a formula α with the query $\{free(\alpha) \mid \alpha\}$ (we assume the order of variables is fixed) and write $\mathfrak{A} \models \alpha[\overline{\mathsf{a}}]$ for $\mathfrak{A} \models \{free(\alpha) \mid \alpha\}[\overline{\mathsf{a}}]$
- So we need a semantic characterization of $\mathfrak{A}, \overline{\mathsf{a}} \equiv_m \mathfrak{B}, \overline{\mathsf{b}}$

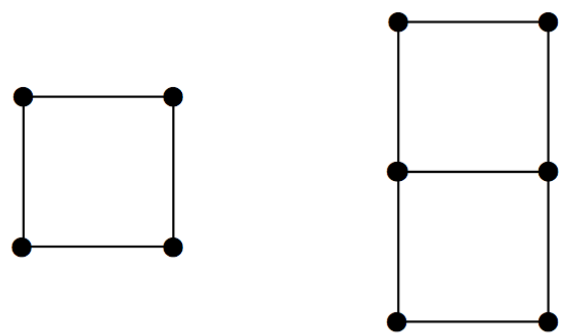
Partial isomorphisms

- Definition of $\mathsf{Part}(\mathfrak{A}, \mathfrak{B})$.
- $\overline{\mathsf{a}} \mapsto \overline{\mathsf{b}} \in \mathsf{Part}(\mathfrak{A}, \mathfrak{B})$ iff quantifier-free equivalent iff atomic equivalent

Ehrenfeucht games: boards and plays

- But what about formulas with quantifiers??
- Example: strict linear $<$ with 2 and 3 elements ...
- The Ehrenfeucht(–Fraïssé) game:
board $\mathsf{G}_m\langle(\mathfrak{A}, \overline{\mathsf{a}}), (\mathfrak{B}, \overline{\mathsf{b}})\rangle$ and ***play of the game***
- **Notation:** $\mathfrak{A}, \overline{\mathsf{a}} \simeq_m \mathfrak{B}, \overline{\mathsf{b}}$ when $\mathsf{G}_m\langle(\mathfrak{A}, \overline{\mathsf{a}}), (\mathfrak{B}, \overline{\mathsf{b}})\rangle$ is winning for duplicator.
- Remark on "Ehrenfeucht" vs. "Ehrenfeucht–Fraïssé"
Ehrenfeucht (1961) later than Fraïssé (1954) but the first to use game-theoretic terms. We'll see Fraïssé's formulation below. And Ehrenfeucht published in English instead of French ...
- Will use "EF-games" below.
Not to confuse with Ebbinghaus–Flum ...

Kolaitis example



A **B**

Lemma 1.

- $f : \mathfrak{A} \cong \mathfrak{B}$ implies $\mathfrak{A}, \bar{\mathsf{a}} \simeq_m \mathfrak{B}, f(\bar{\mathsf{a}})$
- $|\mathsf{A}| < m$ and $\mathfrak{A} \simeq_m \mathfrak{B}$ implies $\mathfrak{A} \cong \mathfrak{B}$
- $\mathfrak{A}, \bar{\mathsf{a}} \simeq_0 \mathfrak{B}, \bar{\mathsf{b}}$ iff $\bar{\mathsf{a}} \mapsto \bar{\mathsf{b}} \in \mathsf{Part}(\mathfrak{A}, \mathfrak{B})$
- $\forall m > 0.$

$$\mathfrak{A}, \bar{\mathsf{a}} \simeq_m \mathfrak{B}, \bar{\mathsf{b}} \text{ iff } \begin{cases} ((\forall \mathsf{a} \in \mathsf{A} \exists \mathsf{b} \in \mathsf{B}.\, \mathfrak{A}, \bar{\mathsf{a}}\mathsf{a} \simeq_{m-1} \mathfrak{B}, \bar{\mathsf{b}}\mathsf{b}) \\ \text{and} \\ (\forall \mathsf{b} \in \mathsf{B} \exists \mathsf{a} \in \mathsf{A}.\, \mathfrak{A}, \bar{\mathsf{a}}\mathsf{a} \simeq_{m-1} \mathfrak{B}, \bar{\mathsf{b}}\mathsf{b})). \end{cases}$$

- $\mathfrak{A}, \bar{\mathsf{a}} \simeq_{m+i} \mathfrak{B}, \bar{\mathsf{b}}$ implies $\mathfrak{A}, \bar{\mathsf{a}} \simeq_m \mathfrak{B}, \bar{\mathsf{b}}$

Play more games

- Empty signature:
  games of length $\leq m$ on sets of cardinality $\geq m$
- Does this still work on linear orders?