LGruDat: Logical Foundations of Databases
Lecture 6 — EF games continued

Tadeusz Litak

December 9, 2013

Fraïssé's algebraic formulation

- Recall: $\forall m > 0$.

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

- The notion of m-isomorphism $\mathfrak{A} \cong_m \mathfrak{B}$ with $(I_j)_{j \leq m}$ witnessing
- (forth) : extending the domain, (back) : extending the range

Play more games

- Empty signature:
  games of length $\leq m$ on sets of cardinality $\geq m$
- Does this still work on linear orders? Our transitive example again
- What if the cardinality of both $\mathfrak{A}$ and $\mathfrak{B}$ is at least 2^m?

Proof by composition

- Note that a winning strategy can be always assumed to pair $\min^{\mathfrak{A}}$ with $\min^{\mathfrak{B}}$ and $\max^{\mathfrak{A}}$ with $\max^{\mathfrak{B}}$
- Observation: whenever $a \in A$ and $b \in B$ are s.t. $\mathfrak{A}^{\leq a} \simeq_m \mathfrak{B}^{\leq b}$, $\mathfrak{A}^{\geq a} \simeq_m \mathfrak{B}^{\geq b}$, it holds that $\mathfrak{A} \simeq_m \mathfrak{B}$

- Now the main step in the inductive proof uses neatly most facts established so far above:
  - using the back-and-forth condition in the inductive step
  - isomorphism for "smaller half" if the spoiler played something closer than 2^{k-1} to either end
  - the above fact about composition ...
  - ... and the induction hypothesis, of course
- Discuss the alternative proof via invariance?

Easy direction of EF

Theorem 1. $\mathfrak{A}, \bar{\mathsf{a}} \simeq_m \mathfrak{B}, \bar{\mathsf{b}}$ implies $\mathfrak{A}, \bar{\mathsf{a}} \equiv_m \mathfrak{B}, \bar{\mathsf{b}}$

Proof

- Base step: already shown
- Inductive step: assume $\mathfrak{A}, \bar{\mathsf{a}} \simeq_{m+1} \mathfrak{B}, \bar{\mathsf{b}}$ and $\mathfrak{A} \vDash \exists v.\alpha[\bar{\mathsf{a}}]$. This means exists $\mathsf{a} \in \mathsf{A}.\mathfrak{A} \vDash \alpha[\bar{\mathsf{a}}\mathsf{a}]$. Now use (forth) to find $\mathfrak{A}, \bar{\mathsf{a}}\mathsf{a} \simeq_m \mathfrak{B}, \bar{\mathsf{b}}\mathsf{b}$ and use IH.

  Assume $\mathfrak{A} \vDash \forall v.\alpha[\bar{\mathsf{a}}]$. This means for_all $\mathsf{a} \in \mathsf{A}.\mathfrak{A} \vDash \alpha[\bar{\mathsf{a}}\mathsf{a}]$. Now pick any $\mathsf{b} \in \mathfrak{B}$. By (back) , there always is a suitable $\mathsf{a} \in \mathsf{A}$ s.t. $\mathfrak{A}, \bar{\mathsf{a}}\mathsf{a} \simeq_m \mathfrak{B}, \bar{\mathsf{b}}\mathsf{b}$. IH yields that $\mathfrak{B} \vDash \forall v.\alpha[\bar{\mathsf{b}}]$.

  Transfer of equivalence via boolean connectives is automatic.

Finite types

- For the converse, recall the notion of $\mathsf{FORC}[m]$
- For a fixed $\bar{\mathbf{v}}$, (up to logical equivalence) only finitely many $\alpha \in \mathsf{FORC}[m]$ with $free(\alpha) \subseteq \bar{\mathbf{v}}$
- Assume $\bar{\mathsf{a}} = \mathsf{a}_1 \ldots \mathsf{a}_{\ell(\bar{\mathsf{a}})}$. Then $\mathsf{tp}^m_{\bar{\mathsf{a}}} := \{\alpha \in \mathsf{FORC}[m] \mid \mathsf{A} \vDash \alpha[\bar{\mathsf{a}}]\}$ is only superficially infinite (assuming the indices of free variables $x_1 \ldots x_{\ell(\bar{\mathsf{a}})}$ form an initial interval $(1, \ldots, \ell(\bar{\mathsf{a}}))$ of $\mathbb{N}$)

- Define

$$\nabla^0_{\bar{a}} := \bigwedge\{\alpha \in \mathsf{tp}^0_{\bar{a}}, \alpha \text{ atomic or negated atomic}\}$$

$$\nabla^{m+1}_{\bar{a}} := (\bigwedge_{a \in A} \exists x_{\ell(\bar{a})+1}.\nabla^m_{\bar{a}a}) \wedge (\forall x_{\ell(\bar{a})+1}. \bigvee_{a \in A} \nabla^m_{\bar{a}a})$$

  Obviously, $\mathfrak{A} \vDash \nabla^m_{\bar{a}}[\bar{a}]$ and $\nabla^m_{\bar{a}} \in \mathsf{FORC}[m]$. Hence,

$$\mathfrak{A}, \bar{a} \equiv_m \mathfrak{B}, \bar{b} \text{ implies } \models^{unr}_{loc} \nabla^m_{\bar{a}} \leftrightarrow \nabla^m_{\bar{b}} \text{ and } \mathfrak{B} \vDash \nabla^m_a[\bar{b}]$$

- Our usual assumption of finiteness of Σ guarantees $\nabla^m_{\bar{a}}$ is a finite formula for arbitrary m and $\bar{a}$:
  - Whenever $\mathfrak{A}$ is finite, all conjunctions and disjunctions are finite
  - Actually, even for unrestricted $\mathfrak{A}$ it's finite up to logical equivalence using the above observation
- $\nabla^m_{\bar{a}}$ is called m-Hintikka formula of $\bar{a}$
- ***Theorem***: $\mathfrak{B} \vDash \nabla^m_{\bar{a}}[\bar{b}]$ implies $\mathfrak{A}, \bar{a} \simeq_m \mathfrak{B}, \bar{b}$

Proof.

- Base step: $\mathfrak{B} \vDash \nabla^0_{\bar{a}}[\bar{b}]$ easily entails $f(a_k) := b_k$ is a partial isomorphism
- Inductive step: assume $\mathfrak{B} \vDash \nabla^{m+1}_{\bar{a}}[\bar{b}]$ and assume spoiler picks $a \in A$. Then
  - $\mathfrak{A} \vDash \nabla^m_{\bar{a}a}[\bar{a}a]$, hence $\mathfrak{A} \vDash \exists x_{\ell(\bar{a})+1}.\nabla^m_{\bar{a}a}[\bar{a}]$.
  - This is a conjunct of $\nabla^{m+1}_{\bar{a}}$, hence $\mathfrak{B} \vDash \exists x_{\ell(\bar{a})+1}.\nabla^m_{\bar{a}a}[\bar{b}]$
  - Pick any witness b for this existential sentence: $\mathfrak{B} \vDash \nabla^m_{\bar{a}a}[\bar{b}b]$
  - By IH, $\mathfrak{A}, \bar{a}a \simeq_m \mathfrak{B}, \bar{b}b$. We proved the forth condition
- Now assume spoiler picks $b \in B$. Then
  - $\mathfrak{B} \vDash \forall x_{\ell(\bar{a})+1}. \bigvee_{a \in A} \nabla^m_{\bar{a}a}[\bar{b}]$
  - In particular, $\mathfrak{B} \vDash \bigvee_{a \in A} \nabla^m_{\bar{a}a}[\bar{b}b]$
  - Use the satisfaction clause for disjunction, pick the witnessing $a \in A$ and use IH

□

Corollaries

- Undefinability of evenness on linear orders
- Undefinability of connectedness as a corollary of the above
- Possibly a more complicated example: trees