Topics in Logic and Complexity
Handout 9

Anuj Dawar

MPhil Advanced Computer Science, Lent 2010

Syntax of LFP

- Any relation symbol of arity k is a predicate expression of arity k;
- If R is a relation symbol of arity k, $\mathbf{x}$ is a tuple of variables of length k and ϕ is a formula of LFP in which the symbol R only occurs positively, then

$$\mathbf{lfp}_{R,\mathbf{x}}\phi$$

 is a predicate expression of LFP of arity k.

All occurrences of R and variables in $\mathbf{x}$ in $\mathbf{lfp}_{R,\mathbf{x}}\phi$ are *bound*

Syntax of LFP

- If t_1 and t_2 are terms, then $t_1 = t_2$ is a formula of LFP.
- If P is a predicate expression of LFP of arity k and $\mathbf{t}$ is a tuple of terms of length k, then $P(\mathbf{t})$ is a formula of LFP.
- If ϕ and ψ are formulas of LFP, then so are $\phi \wedge \psi$, and $\neg\phi$.
- If ϕ is a formula of LFP and x is a variable then, $\exists x\phi$ is a formula of LFP.

Semantics of LFP

Let $\mathbb{A} = (A, \mathcal{I})$ be a structure with universe A, and an interpretation $\mathcal{I}$ of a fixed vocabulary σ.

Let ϕ be a formula of LFP, and $\imath$ an interpretation in A of all the free variables (*first or second* order) of ϕ.

To each individual variable x, $\imath$ associates an element of A, and to each k-ary relation symbol R in ϕ that is not in σ, $\imath$ associates a relation $\imath(R) \subseteq A^k$.

$\imath$ is extended to terms t in the usual way.

For constants c, $\imath(c) = \mathcal{I}(c)$.

$\imath(f(t_1, \ldots, t_n)) = \mathcal{I}(f)(\imath(t_1), \ldots, \imath(t_n))$

Semantics of LFP

- If R is a relation symbol in σ, then $\imath(R) = \mathcal{I}(R)$.
- If P is a predicate expression of the form $\mathbf{lfp}_{R,\mathbf{x}}\psi$, then $\imath(P)$ is the relation that is the least fixed point of the monotone operator F on A^k defined by:

$$F(X) = \{\mathbf{a} \in A^k \mid \mathbb{A} \models \phi[\imath\langle X/R, \mathbf{x}/\mathbf{a}\rangle],$$

where $\imath\langle X/R, \mathbf{x}/\mathbf{a}\rangle$ denotes the interpretation $\imath'$ which is just like $\imath$ *except* that $\imath'(R) = X$, and $\imath'(\mathbf{x}) = \mathbf{a}$.

Semantics of LFP

- If ϕ is of the form $t_1 = t_2$, then $\mathbb{A} \models \phi[\imath]$ if, $\imath(t_1) = \imath(t_2)$.
- If ϕ is of the form $R(t_1, \ldots, t_k)$, then $\mathbb{A} \models \phi[\imath]$ if,

$$(\imath(t_1), \ldots, \imath(t_k)) \in \imath(R).$$

- If ϕ is of the form $\psi_1 \wedge \psi_2$, then $\mathbb{A} \models \phi[\imath]$ if, $\mathbb{A} \models \psi_1[\imath]$ *and* $\mathbb{A} \models \psi_2[\imath]$.
- If ϕ is of the form $\neg\psi$ then, $\mathbb{A} \models \phi[\imath]$ if, $\mathbb{A} \not\models \psi[\imath]$.
- If ϕ is of the form $\exists x\psi$, then $\mathbb{A} \models \phi[\imath]$ if there is an $a \in A$ such that $\mathbb{A} \models \psi[\imath\langle x/a\rangle]$.

Transitive Closure

The formula (with free variables u and v)

$$[\theta \quad \equiv \quad \mathbf{lfp}_{T,xy}(x = y \vee \exists z(E(x,z) \wedge T(z,y)))](u,v)$$

defines the *transitive closure* of the relation E.

Thus $\forall u \forall v\ \theta$ defines *connectedness*.

The expressive power of LFP properly extends that of first-order logic.

Greatest Fixed Points

If ϕ is a formula in which the relation symbol R occurs *positively*, then the *greatest fixed point* of the monotone operator F_ϕ defined by ϕ can be defined by the formula:

$$\neg[\mathbf{lfp}_{R,\mathbf{x}}\neg\phi(R/\neg R)](\mathbf{x})$$

where $\phi(R/\neg R)$ denotes the result of replacing all occurrences of R in ϕ by $\neg R$.

Exercise: Verify!.

Simultaneous Inductions

We are given two formulas $\phi_1(S, T, \mathbf{x})$ and $\phi_2(S, T, \mathbf{y})$,
S is k-ary, T is l-ary.

The pair (ϕ_1, ϕ_2) can be seen as defining a map:

$$F : \mathsf{Pow}(A^k) \times \mathsf{Pow}(A^l) \to \mathsf{Pow}(A^k) \times \mathsf{Pow}(A^l)$$

If both formulas are positive in both S and T, then there is a least fixed point.

$$(P_1, P_2)$$

defined by *simultaneous induction* on $\mathbb{A}$.

Simultaneous Inductions

Theorem
For any pair of formulas $\phi_1(S, T)$ and $\phi_2(S, T)$ of LFP, in which the symbols S and T appear only positively, there are formulas ϕ_S and ϕ_T of LFP which, on any structure $\mathbb{A}$ containing at least two elements, define the two relations that are defined on $\mathbb{A}$ by ϕ_1 and ϕ_2 by simultaneous induction.

Proof

Assume $k \leq l$.

We define P, of arity $l + 2$ such that:

> $(c, d, a_1, \ldots, a_l) \in P$ if, and only if, either $c = d$ and $(a_1, \ldots, a_k) \in P_1$ or $c \neq d$ and $(a_1, \ldots, a_l) \in P_2$

For new variables x_1 and x_2 and a new $l + 2$-ary symbol R, define ϕ_1' and ϕ_2' by replacing all occurrences of $S(t_1, \ldots, t_k)$ by:

$$x_1 = x_2 \wedge \exists y_{k+1}, \ldots, \exists y_l R(x_1, x_2, t_1, \ldots, t_k, y_{k+1}, \ldots, y_l),$$

and replacing all occurrences of $T(t_1, \ldots, t_l)$ by:

$$x_1 \neq x_2 \wedge R(x_1, x_2, t_1, \ldots, t_l).$$

Proof

Define ϕ as

$$(x_1 = x_2 \wedge \phi_1') \vee (x_1 \neq x_2 \wedge \phi_2').$$

Then,

$$[\mathbf{lfp}_{R, x_1 x_2 \mathbf{y}} \phi](x, x, \mathbf{y})$$

defines P, so

$$\phi_S \equiv \exists x \exists y_{k+1}, \ldots, \exists y_l [\mathbf{lfp}_{R, x_1 x_2 \mathbf{y}} \phi](x, x, \mathbf{y});$$

and

$$\phi_T \equiv \exists x_1 \exists x_2 (x_1 \neq x_2 \wedge [\mathbf{lfp}_{R, x_1 x_2 \mathbf{y}} \phi](x_1, x_2, \mathbf{y})).$$

Inflationary Fixed Points

We can associtate with any formula $\phi(R, \mathbf{x})$ (even one that is not *monotone* in R) an *inflationary operator*

$$IF_\phi(P) = P \cup F_\phi(P),$$

On any *finite* structure $\mathbb{A}$ the sequence

$$\begin{aligned} IF^0 &= \emptyset \\ IF^{n+1} &= IF_\phi(IF^n) \end{aligned}$$

converges to a limit IF^∞.

If F_ϕ is monotone, then this fixed point is, in fact, the least fixed point of F_ϕ.

IFP

We define the logic IFP with a syntax similar to LFP except, instead of the **lfp** rule, we have

If R is a relation symbol of arity k, $\mathbf{x}$ is a tuple of variables of length k and ϕ is any formula of IFP, then

$$\mathbf{ifp}_{R,\mathbf{x}}\phi$$

is a predicate expression of IFP of arity k.

Semantics: we say that the predicate expression $\mathbf{ifp}_{R,\mathbf{x}}\phi$ denotes the relation that is the limit reached by the iteration of the inflationary operator IF_ϕ.

IFP

If ϕ defines a monotone operator, the relation defined by

$$\mathbf{ifp}_{R,\mathbf{x}}\phi$$

is the least fixed point of ϕ.

Thus, the *expressive power* of IFP is at least as great as that of LFP.

In fact, it is no greater:

Theorem (Gurevich-Shelah)
For every formula of ϕ of LFP, there is a predicate expression ψ of LFP such that, on any finite structure $\mathbb{A}$, ψ defines the same relation as $\mathbf{ifp}_{R,\mathbf{x}}\phi$.

Ranks

Let $\phi(R, \mathbf{x})$ be a formula defining an operator F_ϕ and IF_ϕ be the associated *inflationary* operator given by

$$IF_\phi(S) = S \cup F_\phi(S)$$

In a structure $\mathbb{A}$, we define for each $\mathbf{a} \in A^k$ a *rank* $|\mathbf{a}|_\phi$.

The least n such that $\mathbf{a} \in IF^\alpha$, if there is such an n and ∞ otherwise.

Stage Comparison

We define the two *stage comparison* relations $\preceq$ and $\prec$ by:

$$\mathbf{a} \preceq \mathbf{b} \;\Leftrightarrow\; \mathbf{a} \in IF_\phi^\infty \wedge |\mathbf{a}|_\phi \leq |\mathbf{b}|_\phi;$$

$$\mathbf{a} \prec \mathbf{b} \;\Leftrightarrow\; |\mathbf{a}|_\phi < |\mathbf{b}|_\phi.$$

These two relations can themselves be defined in IFP.

Stage Comparison

$$\mathbf{a} \preceq \mathbf{b} \;\Leftrightarrow\; \mathbf{a} \in IF_\phi(\{\mathbf{a}' \mid \mathbf{a} \prec \mathbf{b}\}).$$

$$\mathbf{a} \prec \mathbf{b} \;\Leftrightarrow\; \mathbf{b} \notin IF_\phi(\{\mathbf{b}' \mid \neg(\mathbf{a} \preceq \mathbf{b}')\}).$$

Together, these give:

$$\mathbf{a} \preceq \mathbf{b} \;\Leftrightarrow\; \mathbf{a} \in IF_\phi(\{\mathbf{a}' \mid \mathbf{b} \notin IF_\phi(\{\mathbf{b}' \mid \neg(\mathbf{a}' \preceq \mathbf{b}')\})\}).$$

This is an inductive definition of $\preceq$.

A similar inductive definition is obtained from $\prec$.

Stage Comparison in LFP

In the inductive definition of $\preceq$:

$$\mathbf{a} \preceq \mathbf{b} \Leftrightarrow \mathbf{a} \in IF_\phi(\{\mathbf{a}' \mid \mathbf{b} \notin IF_\phi(\{\mathbf{b}' \mid \neg(\mathbf{a}' \preceq_\phi \mathbf{b}')\})$$

we can replace the *negative* occurrences of $\mathbf{a} \preceq \mathbf{b}$ with $\neg(\mathbf{b} \prec \mathbf{a})$, and similarly, in the definition of $\prec$ replace negative occurrences of $\prec$ with positive occurrences of $\preceq$

as long as we can define the maximal rank

Maximal Rank

There is a formula $\mu(\mathbf{y})$, which defines the set of tuples of maximal rank.

$$IF_\phi(\{\mathbf{b} \mid \mathbf{b} \preceq \mathbf{a}\}) \subseteq IF_\phi(\{\mathbf{b} \mid \mathbf{b} \prec \mathbf{a}\}).$$

Replace the negative occurrence of $\mathbf{b} \preceq \mathbf{a}$ by $\neg(\mathbf{a} \prec \mathbf{b})$.

Reading List for this Handout

1. Immerman. Chapter 4.
2. Libkin. Sections 10.2 and 10.3.
3. Grädel et al. Secton 2.6.