Chapter 4

First Order Fixed-Point Logic

Evergetis

March 20, 2023

National Technical University of Athens,
National and Kapodistrian University of Athens

Μεταπτυχιακό Πρόγραμμα «Αλγόριθμοι, Λογική και Διακριτά Μαθηματικά» – 2016

Table of contents

Least Fixed Point

Introduction

Do we get more expressibility by adding something extra to first-order logic, without having to jump all the way to second-order logic? **Yes!**

- Define new relations by induction!

e.g. Recall the vocabulary $\tau_g = \langle E^2, s, t\rangle$ of graphs. We define formally the reflexive, transitive closure of E^* of E as follows:

$$\varphi(R, x, y) \quad \equiv \quad x = y \ \vee \ \exists z\,(E(x, z) \wedge R(z, y)) \tag{1}$$

or written differently

$$E^*(x, y) \quad \equiv \quad x = y \ \vee \ \exists z\,(E(x, z) \wedge E^*(z, y)) \tag{2}$$

For any structure $\mathcal{A}$ with vocabulary τ_g, formula (1) *induces* a map from binary relations on the universe of $\mathcal{A}$ to binary relations on the universe of $\mathcal{A}$,

$$\varphi^{\mathcal{A}}(R) = \left\{\langle a, b\rangle \mid \mathcal{A} \vDash \varphi(R, a, b)\right\}$$

We observe that, for $\mathcal{A}$ any graph and $r \geq 0$:

$$(\varphi^{\mathcal{A}})(\emptyset) = \left\{\langle a, b\rangle \in |\mathcal{A}|^2 \mid \text{distance}(a, b) \leq 0\right\}$$
$$(\varphi^{\mathcal{A}})^2(\emptyset) = \left\{\langle a, b\rangle \in |\mathcal{A}|^2 \mid \text{distance}(a, b) \leq 1\right\}$$

and in general

$$(\varphi^{\mathcal{A}})^r(\emptyset) = \left\{\langle a, b\rangle \in |\mathcal{A}|^2 \mid \text{distance}(a, b) \leq r - 1\right\}.$$

Thus, for cardinality $n = \|\mathcal{A}\|$, then $(\varphi^{\mathcal{A}})^n(\emptyset) = E^* =$ the least fixed point of $\varphi^{\mathcal{A}}$ i.e. the minimal relation T such that $\varphi^{\mathcal{A}}(T) = T$.

Least Fixed Point

Following comes the finite version of the Knaster-Tarski Theorem.

Theorem

Let R be a new relation symbol of arity k, and let $\varphi(R, x_1, \ldots, x_k)$ be a monotone first-order formula. Then for any finite structure $\mathcal{A}$, the ***least fixed point*** *of $\varphi^{\mathcal{A}}$ exists. It is equal to $(\varphi^{\mathcal{A}})^r(\emptyset)$ where r is minimal so that $(\varphi^{\mathcal{A}})^r(\emptyset) = (\varphi^{\mathcal{A}})^{r+1}(\emptyset)$. Furthermore, letting $n = \|\mathcal{A}\|$, we have $r \leq n^k$.*

Proof.

Consider the sequence

$$\emptyset \subseteq (\varphi^{\mathcal{A}})(\emptyset) \subseteq (\varphi^{\mathcal{A}})^2(\emptyset) \subseteq (\varphi^{\mathcal{A}})^3(\emptyset) \subseteq \dots$$

The containment follows because $\varphi^{\mathcal{A}}$ is monotone. If $(\varphi^{\mathcal{A}})^{i+1}(\emptyset)$ strictly contains $(\varphi^{\mathcal{A}})^i(\emptyset)$, then it must contain at least one new k-tuple from $|\mathcal{A}|$. Since there are at most n^k such k-tuples, for some $r \leq n^k$, $(\varphi^{\mathcal{A}})^r(\emptyset) = (\varphi^{\mathcal{A}})^{r+1}(\emptyset)$, i.e. $(\varphi^{\mathcal{A}})^r(\emptyset)$ is a fixed point of $\varphi^{\mathcal{A}}$.

Let S be any other fixed point of $\varphi^{\mathcal{A}}$. We show inductively that $(\varphi^{\mathcal{A}})^i(\emptyset) \subseteq S$ for all i.

- IB: $(\varphi^{\mathcal{A}})^0(\emptyset) = \emptyset \subseteq S$
- IH: $(\varphi^{\mathcal{A}})^i(\emptyset) \subseteq S$
- IS: $(\varphi^{\mathcal{A}})^{i+1}(\emptyset) = \varphi^{\mathcal{A}}((\varphi^{\mathcal{A}})^i(\emptyset)) \subseteq \varphi^{\mathcal{A}}(S) = S$.

Thus, $(\varphi^{\mathcal{A}})^r(\emptyset) \subseteq S$ and $(\varphi^{\mathcal{A}})^r(\emptyset)$ is the least fixed point of $\varphi^{\mathcal{A}}$ as claimed. □

- Notation: We write $(\text{LFP}_{R^k x_1 \ldots x_k} \varphi)$ to denote the least fixed point of any R-positive formula $\varphi(R^k, x_1, \ldots, x_k)$, although the subscript may be ommited when the choice of variables is clear.

- $(\text{LFP}_{Rxy\, \varphi_{(1)}})$ denotes the reflexive, transitive closure of the edge relation E. Thus boolean query REACH is expressible as:

$$\text{REACH} \quad \equiv \quad (\text{LFP}_{Rxy\, \varphi_{(1)}})(s, t)$$

First-Order Inductive Definitions

Definition

Define FO(LFP), the **language of first-order inductive definitions**, by adding a least fixed point operator (LFP) to first-order logic. If $\varphi(R^k, x_1, \ldots, x_k)$ is an R^k-positive formula in FO(LFP), then $(\mathrm{LFP}_{R^k x_1 \ldots x_k} \varphi)$ may be used as a new k-ary relation symbol denoting the least fixed point of φ.

REACH_α expressed with FO(LFP)

- REACH_α is defined to be the set of graphs having alternating path from s to t.
- Let's give a first-order inductive definition of the alteranting path property P_α,

$$\varphi_{\alpha p} \quad \equiv \quad x = y \ \vee \ [(\exists z)\,(E(x,z) \wedge P(z,y)) \wedge \\ (A(x) \rightarrow (\forall z)\,(E(x,z) \rightarrow P(z,y)))]$$

- Thus,

$$P_\alpha = (\text{LFP}_{Pxy\,\varphi_{\alpha p}}) \quad \text{and} \quad \text{REACH}_\alpha = (\text{LFP}_{Pxy\,\varphi_{\alpha p}})(s,t).$$

Expressive power of FO(LFP)

FO(LFP)=P

Theorem

Over finite, ordered structures,

$$FO(LFP) = \mathbf{P}$$

Proof.

- ($\subseteq$): Let $\mathcal{A}$ an input structure, let $n = \|\mathcal{A}\|$, and let $(\text{LFP}_{R^k x_1 \ldots x_k} \varphi)$ be a fixed-point formula. By the previous theorem we know that this fixed point evaluated on $\mathcal{A}$ is $(\varphi_{\mathcal{A}})^{n^k}(\emptyset)$, which amounts to evaluating the first-order query φ at most n^k times. We also know FO $\subseteq$ **L**, thus it's in **P** as well.
- ($\supseteq$): Since FO(LFP) includes query REACH$_\alpha$, which is complete for **P** via first-order reductions, and FO(LFP) is closed under first-order reductions, FO(LFP) includes all polynomial-time queries.

□

Connections to Parallelism

Definition

Depth $|\varphi^{\mathcal{A}}|$ of a formula φ in a structure $\mathcal{A}$ of size n is the minimum r such that

$$\mathcal{A} \models \left(\varphi^r(\emptyset) \leftrightarrow \varphi^{r+1}(\emptyset)\right)$$

- Intuitively, it is the number of iterations until an inductive definition closes and as we saw it is upper bounded by n^k
- It also corresponds to the depth of recursive calls in the stack when it comes to evaluate recursive definitions.

Definition

Let IND[$f(n)$] be the sublanguage of FO(LFP) in which only fixed points of first-order formulas φ for which $|\varphi|$ is $\mathcal{O}[f(n)]$ are included. Thus,

$$\text{FO(LFP)} = \bigcup_{k=1}^{\infty} \text{IND}[n^k]$$

- We know that REACH $\in$ IND[$logn$], REACH is complete for **NL** via first-order reuctions, and IND[$logn$] is closed under first-order reductions. All that imply

$$\mathbf{NL} \subseteq \text{IND[logn]}$$

Spoiler alert:

$$\text{IND[logn]} = \mathbf{AC}^1$$

Questions?

Thank you!