Introduction to Descriptive Complexity

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 600

1. Inductive Definitions
   - Least Fixed Point
   - The Depth of Inductive Definitions
   - Iterating First Order Formulas

Subsection 1

Least Fixed Point

Example

- A useful way to increase the power of first-order logic is to add the power to define new relations by induction.

  Example: A relation that is not first-order expressible, but can be defined inductively, is transitive closure.

  Recall the vocabulary $\tau_g = \langle E^2, s, t \rangle$ of graphs.

  We define the reflexive, transitive closure E^* of E as follows.

  Let R be a binary relation variable.

  Consider the formula

$$\varphi_{4,1}(R, x, y) \equiv x = y \vee \exists z(E(x, z) \wedge R(z, y)).$$

  The formula $\varphi_{4,1}$ formalizes an inductive definition of E^*.

  This may be more suggestively written as

$$E^*(x, y) \equiv x = y \vee \exists z(E(x, z) \wedge E^*(z, y)).$$

Example (Monotonicity)

- For any structure $\mathcal{A}$ with vocabulary τ_g, $\varphi_{4,1}$ induces a map from binary relations on the universe of $\mathcal{A}$ to binary relations on the universe of $\mathcal{A}$,

$$\varphi_{4,1}^{\mathcal{A}}(R) = \{\langle a, b\rangle : \mathcal{A} \vDash \varphi_{4,1}(R, a, b)\}.$$

- Such a map is called **monotone** if for all R, S,

$$R \subseteq S \Rightarrow \varphi^{\mathcal{A}}(R) \subseteq \varphi^{\mathcal{A}}(S).$$

- Note that the relation symbol R appears only positively in $\varphi_{4,1}$, i.e., within an even number of negation symbols.
- It follows that $\varphi_{4,1}^{\mathcal{A}}$ is monotone.

Example (Least Fixed Point)

- Let $(\varphi_{4,1}^{\mathcal{A}})^r$ denote $\varphi_{4,1}^{\mathcal{A}}$ iterated r times.
- If $\mathcal{A}$ any graph, and $r \geq 0$, observe that:

$$\begin{array}{rcl} (\varphi_{4,1}^{\mathcal{A}})(\varnothing) & = & \{\langle a,b\rangle \in |\mathcal{A}|^2 : \text{distance}(a,b) \leq 0\}; \\ (\varphi_{4,1}^{\mathcal{A}})^2(\varnothing) & = & \{\langle a,b\rangle \in |\mathcal{A}|^2 : \text{distance}(a,b) \leq 1\}; \\ & \vdots & \\ (\varphi_{4,1}^{\mathcal{A}})^r(\varnothing) & = & \{\langle a,b\rangle \in |\mathcal{A}|^2 : \text{distance}(a,b) \leq r-1\}; \\ & \vdots & \end{array}$$

- Thus, for $n = \|\mathcal{A}\|$,

$$(\varphi_{4,1}^{\mathcal{A}})^n(\varnothing) = E^* = \text{the least fixed point of } \varphi_{4,1}^{\mathcal{A}}.$$

- That is, $(\varphi_{4,1}^{\mathcal{A}})^n(\varnothing)$ is the minimal relation T, with $\varphi_{4,1}^{\mathcal{A}}(T) = T$.

Knaster-Tarski Theorem (Finite Version)

Knaster-Tarski Theorem (Finite Version)

Let R be a new relation symbol of arity k. 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 $(\varphi^{\mathcal{A}})^r(\varnothing)$, where r is minimal, such that

$$(\varphi^{\mathcal{A}})^r(\varnothing) = (\varphi^{\mathcal{A}})^{r+1}(\varnothing).$$

Furthermore, letting $n = \|\mathcal{A}\|$, we have $r \leq n^k$.

- Consider the sequence

$$\varnothing \subseteq (\varphi^{\mathcal{A}})(\varnothing) \subseteq (\varphi^{\mathcal{A}})^2(\varnothing) \subseteq (\varphi^{\mathcal{A}})^3(\varnothing) \subseteq \cdots.$$

The containment follows because $\varphi^{\mathcal{A}}$ is monotone.

Knaster-Tarski Theorem (Cont'd)

- Suppose $(\varphi^{\mathcal{A}})^{i+1}(\varnothing)$ strictly contains $(\varphi^{\mathcal{A}})^{i}(\varnothing)$.
  Then it must contain at least one new k-tuple from $|\mathcal{A}|$.
  But there are at most n^k such k-tuples.
  So, for some $r \leq n^k$, $(\varphi^{\mathcal{A}})^{r}(\varnothing) = (\varphi^{\mathcal{A}})^{r+1}(\varnothing)$.
  This shows that $(\varphi^{\mathcal{A}})^{r}(\varnothing)$ is a fixed point of $\varphi^{\mathcal{A}}$.
  Let S be any other fixed point of $\varphi^{\mathcal{A}}$.
  We show by induction that $(\varphi^{\mathcal{A}})^{i}(\varnothing) \subseteq S$, for all i.
  The base case is that, $(\varphi^{\mathcal{A}})^{0}(\varnothing) = \varnothing \subseteq S$.
  Inductively, suppose that $(\varphi^{\mathcal{A}})^{i}(\varnothing) \subseteq S$.
  Since $\varphi^{\mathcal{A}}$ is monotone,

$$(\varphi^{\mathcal{A}})^{i+1}(\varnothing) = \varphi^{\mathcal{A}}((\varphi^{\mathcal{A}})^{i}(\varnothing)) \subseteq \varphi^{\mathcal{A}}(S) = S.$$

  Thus, $(\varphi^{\mathcal{A}})^{r}(\varnothing) \subseteq S$.
  So $(\varphi^{\mathcal{A}})^{r}(\varnothing)$ is the least fixed point of $\varphi^{\mathcal{A}}$.

Least Fixed Point Operator

- If R occurs only positively in φ, i.e., within an even number of negation symbols, then φ is monotone.
- The theorem tells us that any R-positive formula $\varphi(R^k, x_1, \ldots, x_k)$ determines a least fixed point relation.
- We denote this least fixed point by

$$(\mathrm{LFP}_{R^k x_1 \ldots x_k} \varphi).$$

- The least fixed point operator (LFP), thus, formalizes the definition of new relations by induction.
- The subscript "$R^k x_1 \ldots x_k$" explicitly tells us which relation and domain variables we are taking the fixed point with respect to.
- When the choice of variables is clear, these subscripts may be omitted.

Example

- Recall the formula

$$\varphi_{4,1}(R,x,y) \equiv x = y \vee \exists z(E(x,z) \wedge R(z,y)).$$

The expression

$$(\text{LFP}_{Rxy}\varphi_{4,1})$$

denotes the reflexive, transitive closure of the edge relation E.

Thus, the boolean query REACH is expressible as

$$\text{REACH} \equiv (\text{LFP}_{Rxy}\varphi_{4,1})(s,t).$$

Language of 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.
For $\varphi(R^k, x_1, \ldots, x_k)$ an R^k-positive formula in FO(LFP),

$$(\mathrm{LFP}_{R^k x_1 \ldots x_k} \varphi)$$

may be used as a new k-ary relation symbol.
Then, in a given structure,

$$(\mathrm{LFP}_{R^k x_1 \ldots x_k} \varphi)$$

denotes the least fixed point of the relation expressed by φ.

Example

- We defined boolean query REACH_a to be the set of graphs having an alternating path from s to t.
- We now give a first-order inductive definition of the alternating path property P_a.
- Consider the P^2-positive formula

$$\begin{aligned} \varphi_{ap} \quad \equiv \quad & x = y \vee [(\exists z)(E(x,z) \wedge P(z,y)) \wedge \\ & (A(x) \rightarrow (\forall z)(E(x,z) \rightarrow P(z,y)))]. \end{aligned}$$

- Thus, we have $P_a = (\text{LFP}_{Pxy}\varphi_{ap})$.
- Moreover,

$$\text{REACH}_a = (\text{LFP}_{Pxy}\varphi_{ap})(s,t).$$

FO(LFP) and Polynomial-Time Boolean Queries

- Recall that REACH_a is complete for P via first-order reductions.
- It follows from the preceding example and the following proposition that FO(LFP) contains all polynomial time boolean queries.

Proposition

FO(LFP) is closed under first-order reductions.

- Let Q be a k-ary first-order query and $\Phi \in \text{FO(LFP)}$.

  We must show that $\widehat{Q}(\Phi) \in \text{FO(LFP)}$.

  This follows from the observation that

$$\widehat{Q}(\text{LFP}_{R^a x_1 \dots x_a} \alpha) \equiv (\text{LFP}_{R^{ka} x_1^1 \dots x_1^k \dots x_a^1 \dots x_a^k} \widehat{Q}(\alpha)).$$

Notational Convention

- Suppose $\varphi(S)$ is a formula in which occurs the n-ary relation symbol S, perhaps among other relation symbols.
- Suppose U is an $(m+n)$-ary relation symbol.
- Suppose, also, that $\overline{t}$ is an m-tuple of constants (or, more generally, of terms).
- Then the notation

  $$\varphi(\{\overline{u} : U(\overline{t}, \overline{u})\})$$

  denotes the result of replacing each occurrence of $S(\overline{v})$ in $\varphi(S)$ by $U(\overline{t}, \overline{v})$.

Simultaneous Induction

- Sometimes it is convenient to use *simultaneous induction* to define several relations.
- Let S and T be new relation symbols, with arities r_0 and r_1.
- Fix variables $\overline{y}$ and $\overline{x}$, such that

$$|\overline{y}| = r_0 = \text{arity}(S) \quad \text{and} \quad |\overline{x}| = r_1 = \text{arity}(T).$$

- Consider first order formulas

$$\psi(\overline{y}, S, T) \quad \text{and} \quad \varphi(\overline{x}, S, T),$$

  which are positive in S and T.
- For any structure $\mathcal{A}$, we define the relations I_0^ω and I_1^ω, by simultaneous induction.

Simultaneous Induction (Cont'd)

- The definition proceeds as follows.

$$\begin{aligned} &I_0^0 = I_1^0 = \varnothing; \\ &\overline{a} \in I_0^n \iff \mathcal{A} \vDash \psi(\overline{a}, I_0^{n-1}, I_1^{n-1}); \\ &\overline{b} \in I_1^n \iff \mathcal{A} \vDash \varphi(\overline{b}, I_0^{n-1}, I_1^{n-1}); \\ &I_b^\omega = \bigcup_{n=1}^\infty I_b^n, \ \ b = 0, 1. \end{aligned}$$

- We show that both I_0^ω and I_1^ω are expressible in FO(LFP).
- For simplicity, suppose there exist distinct constants $c_0 \neq c_1$.
- Choose sequences $\overline{c}$ and $\overline{d}$ of elements of $|\mathcal{A}|$ of the same length as $\overline{x}$ and $\overline{y}$, respectively.
- We define a single new relation U of arity $1 + r_0 + r_1$, such that:
  - $U(c_0, \overline{y}, \overline{c})$ refers to $S(\overline{y})$;
  - $U(c_1, \overline{d}, \overline{x})$ refers to $T(\overline{x})$.

Simultaneous Induction (Cont'd)

- Define a formula

$$\begin{array}{rcl}\chi(z,\overline{y},\overline{x},U) & \equiv & (z=c_0\wedge\psi(\overline{y},\{\overline{y}':U(c_0,\overline{y}',\overline{c})\},\{\overline{x}':U(c_1,\overline{d},\overline{x}')\})) \\ & & \vee(z=c_1\wedge\varphi(\overline{x},\{\overline{y}':U(c_0,\overline{y}',\overline{c})\},\{\overline{x}':U(c_1,\overline{d},\overline{x}')\}))).\end{array}$$

- We claim that, for every n,

$$\begin{array}{rcl}\overline{y}\in I_0^n & \Leftrightarrow & (c_0,\overline{y},\overline{c})\in I_\chi^n, \\ \overline{x}\in I_1^n & \Leftrightarrow & (c_1,\overline{d},\overline{x})\in I_\chi^n.\end{array}$$

- We use induction on n simultaneously for both equivalences, e.g.:

$$\begin{array}{rcl}\overline{y}\in I_0^n & \Leftrightarrow & \psi(\overline{y},I_0^{n-1},I_1^{n-1}) \quad \text{(by definition)} \\ & \Leftrightarrow & \chi(c_0,\overline{y},\overline{c},\{(c_0,\overline{y}',\overline{c}):\overline{y}'\in I_0^{n-1}\}\cup\{(c_1,\overline{d},\overline{x}'):\overline{x}'\in I_1^{n-1}\}) \\ & & \text{(by the definition of } \chi\text{)} \\ & \Leftrightarrow & \chi(c_0,\overline{y},\overline{c},I_\chi^{n-1}) \quad \text{(by the induction hypothesis)} \\ & \Leftrightarrow & (c_0,\overline{y},\overline{c})\in I_\chi^n. \quad \text{(by definition of } I_\chi^n\text{)}\end{array}$$

FO(LFP)=P

Theorem

Over finite, ordered structures, FO(LFP) = P.

($\subseteq$) Let $\mathcal{A}$ be an input structure, with $n = \|\mathcal{A}\|$.

Let $(\mathrm{LFP}_{Rx_1 \dots x_k}\varphi)$ be a fixed-point formula.

We know that this fixed point evaluated on $\mathcal{A}$ is $(\varphi^{\mathcal{A}})^{n^k}(\varnothing)$.

This amounts to evaluating the first-order query φ at most n^k times.

We saw that first-order queries may be evaluated in L.

Thus, they are easily in P.

($\supseteq$) FO(LFP) includes query REACH_a.

We know REACH_a is complete for P via first order reductions.

We also know FO(LFP) is closed under first-order reductions.

So FO(LFP) includes all polynomial-time queries.

Remark: The Role of Ordering

- The use of ordering in the theorem is required in the proof that REACH_a is complete via $\leq_{\text{fo}}$.
- Stripped of its numeric relations, including ordering, FO(LFP) does not describe all polynomial-time properties.
- E.g., we will see that it cannot even express the parity of its universe.

Normal Form in FO(LFP)

Corollary

Let φ be any formula in the language FO(LFP). There exists a first-order formula ψ and a tuple of constants $\overline{c}$, such that over finite, ordered structures,

$$\varphi \equiv (\mathrm{LFP}\psi)(\overline{c}).$$

- The completeness of REACH_a for P means that every polynomial-time query is expressible as $\widehat{Q}(\mathrm{REACH}_a)$ for some first-order query Q.
  In a previous example we saw that $\mathrm{REACH}_a = (\mathrm{LFP}\varphi_{ap})(s,t)$.
  Thus, an arbitrary polynomial-time query is expressible as

  $$\widehat{Q}(\mathrm{REACH}_a) = (\mathrm{LFP}\widehat{Q}(\varphi_{ap}))\widehat{Q}(s,t).$$

  Now recall the first-order reductions used in a previous theorem.
  They replace the constants s and t by k-tuples in 0 and max.
  So the form of the displayed equation is as claimed.

Subsection 2

The Depth of Inductive Definitions

The Depth of Inductive Definitions

- The number of iterations until an inductive definition closes is called its depth.

Definition

Let $\varphi(R, x_1, \ldots, x_k)$ be an R-positive formula, where R is a relation symbol of arity k, and let $\mathcal{A}$ be a structure of size n. Define the **depth of** φ **in** $\mathcal{A}$, in symbols $|\varphi^{\mathcal{A}}|$, to be the minimum r such that

$$\mathcal{A} \models (\varphi^r(\varnothing) \leftrightarrow \varphi^{r+1}(\varnothing)).$$

We saw that $|\varphi^{\mathcal{A}}| \leq n^k$. Define the **depth of** φ as a function of n to be the maximum depth of φ in $\mathcal{A}$ for any structure $\mathcal{A}$ of size n,

$$|\varphi|(n) = \max_{\|\mathcal{A}\|=n} \{|\varphi^{\mathcal{A}}|\}.$$

Remarks

- The inductive definition $\varphi_{4,1}$ given by

$$\varphi_{4,1}(R,x,y) \equiv x = y \vee \exists z(E(x,z) \wedge R(z,y))$$

has depth $|\varphi|(n) = n$.

- However, the following alternate inductive definition of E^* has depth $|\varphi_{tc}|(n) = \lceil \log n \rceil + 1$:

$$\varphi_{tc}(R,x,y) \equiv x = y \vee E(x,y) \vee \exists z(R(x,z) \wedge R(z,y)).$$

- In computer science, the depth of an inductive definition corresponds to the depth of the stack needed to evaluate a recursive definition.
- This is the same as the depth of nesting of recursive calls.
- We will see that this also corresponds to the parallel time needed to evaluate such a recursive definition.

The Language IND$[f(n)]$

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 $O[f(n)]$ are included.

- Recall that REACH is expressible as $(\text{LFP}_{Rxy}\varphi_{tc})$.
- It follows, by the definition, that

$$\text{REACH} \in \text{IND}[\log n].$$

- Note also that,

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

NL and IND[log(n)]

Proposition

NL $\subseteq$ IND[$\log n$].

- The statement is a consequence of the following facts:
  - REACH $\in$ IND[$\log n$];
  - REACH is complete for NL via first-order reductions;
  - IND[$\log n$] is closed under first-order reductions.

Numerical Relations and Inductive Depth

- It can be shown that the numeric relations BIT, PLUS and TIMES are all definable in $\text{IND}(\text{woBIT})[\log n]$, that is, via first-order inductive definitions that use only the numeric relation $\leq$.
- So the descriptive class $\text{IND}[\log n]$ is somewhat more robust than $\text{IND}[0] = \text{FO}$ and has a more general definition.
- It turns out that, generally, the more powerful the language, the less important exactly which numeric relations are included.

Proposition

Relation BIT is definable by a depth $\log n$ induction just from $\leq$,

$$\text{BIT} \in \text{IND}(\text{woBIT})[\log n].$$

- This is proved by showing that, first PLUS and then MULT are definable in $\text{IND}(\text{woBIT})[\log n]$.

  Then use the fact that BIT is definable in terms of PLUS and TIMES.

Subsection 3

Iterating First Order Formulas

Normal Form for Inductive Definitions

- Recall the notation:
  - $(\forall x.M)\psi$ meaning $(\forall x)M \to \psi$;
  - $(\exists x.M)\psi$ meaning $(\exists x)M \wedge \psi$.

Lemma

Let φ be an R-positive first-order formula. Then φ can be written in the following form,

$$\varphi(R, x_1, \ldots, x_k) \equiv (Q_1 z_1.M_1)\cdots(Q_s z_s.M_s)(\exists x_1 \ldots x_k.M_{s+1})R(x_1, \ldots, x_k),$$

where the M_i's are quantifier-free formulas in which R does not occur.

- By induction on the complexity of φ.

  Assume that all negations have been pushed all the way inside.

Normal Form for Inductive Definitions (Cont'd)

- There are two base cases.
  - Suppose, first, $\varphi \equiv R(v_1, \ldots, v_k)$.
    Then
    $$\begin{array}{rcl} \varphi & \equiv & (\exists z_1, \ldots, z_k.M_1)(\exists x_1, \ldots, x_k.M_2)R(x_1, \ldots, x_k), \\ M_1 & \equiv & z_1 = v_1 \wedge \cdots \wedge z_k = v_k; \\ M_2 & \equiv & x_1 = z_1 \wedge \cdots \wedge x_k = z_k. \end{array}$$
  - Suppose, next, φ is quantifier free and R does not occur in φ.
    Then
    $$\varphi \equiv (\forall z.\neg\varphi)(\exists x_1, \ldots, x_k.x_1 \neq x_1)R(x_1, \ldots, x_k).$$
- In the inductive cases $\varphi = (\exists v)\psi$ and $\varphi = (\forall v)\psi$, we simply put the new quantifier $(\exists v)$ in front of the quantifier block for ψ.

Normal Form for Inductive Definitions (Cont'd)

- The remaining cases for $\wedge$ and $\vee$ are similar to each other.

  Suppose that $\varphi = \alpha \wedge \beta$ and

$$\begin{array}{rcl} \alpha & \equiv & (Q_1 y_1.N_1)\cdots(Q_t y_t.N_t)(\exists x_1 \ldots x_k.N_{t+1})R(x_1,\ldots,x_k); \\ \beta & \equiv & (Q_1 z_1.M_1)\cdots(Q_s z_s.M_s)(\exists x_1 \ldots x_k.M_{s+1})R(x_1,\ldots,x_k), \end{array}$$

  where we may assume that the y's and z's are disjoint.

  Let

$$\begin{array}{rcl} QB_1 & \equiv & (Q_1 y_1.N_1')\cdots(Q_t y_t.N_t'), \\ QB_2 & \equiv & (Q_1 z_1.M_1')\cdots(Q_s z_s.M_s'), \end{array}$$

  where, $N_i' \equiv N_i \vee b = 1$ and $M_i' \equiv M_i \vee b = 0$.

Normal Form for Inductive Definitions (Cont'd)

- Let $\psi(\overline{u}/\overline{x})$ denote the formula ψ with variables $u_1, \ldots, u_k$ substituted for $x_1, \ldots, x_k$.

  Define the quantifier-free formulas,

$$\begin{aligned} S &\equiv (b = 0 \wedge N_{t+1}(\overline{u}/\overline{x})) \vee (b = 1 \wedge M_{s+1}(\overline{u}/\overline{x})), \\ T &\equiv (u_1 = x_1 \wedge \cdots \wedge u_k = x_k). \end{aligned}$$

  Recall that $\text{bool}(b)$ means that $b = 0$ or $b = 1$.

  We can now write

$$\varphi \equiv (\forall b.\text{bool}(b))(QB_1)(QB_2)(\exists \overline{u}.S)(\exists \overline{x}.T)R(x_1, \ldots, x_k).$$

Remarks on Requantification

- In $(Q_1 z_1.M_1)\cdots(Q_s z_s.M_s)(\exists x_1 \ldots x_k.M_{s+1})R(x_1,\ldots,x_k)$, the requantification of the x_i's means that these variables may occur free in $M_1,\ldots,M_s$, but they are bound in M_{s+1} and $R(x_1,\ldots,x_k)$.
- The same variables may now be requantified.
- Let us write QB to denote the quantifier block

$$(Q_1 z_1.M_1)\cdots(Q_s z_s.M_s)(\exists x_1 \ldots x_k.M_{s+1}).$$

- Thus, in particular, for any structure $\mathcal{A}$, and any $r \in \mathbb{N}$,

$$\mathcal{A} \vDash ((\varphi^{\mathcal{A}})^r(\varnothing)) \leftrightarrow ([QB]^r\textbf{false}).$$

- Here $[QB]^r$ means QB literally repeated r times.
- It follows that, if $t = |\varphi|(n)$ and $\mathcal{A}$ is any structure of size n, then

$$\mathcal{A} \vDash (\text{LFP}\varphi) \leftrightarrow ([QB]^t\textbf{false}).$$

Example: Normal Form of the Transitive Closure

- Recall the definition of transitive closure:

$$\varphi_{tc}(R,x,y) \equiv x = y \vee E(x,y) \vee (\exists z)(R(x,z) \wedge R(z,y)).$$

- First, code the base case using a dummy universal quantification,

$$\begin{array}{rcl} \varphi_{tc}(R,x,y) & \equiv & (\forall z.M_1)(\exists z)(R(x,y) \wedge R(z,y)), \\ M_1 & \equiv & \neg(x = y \vee E(x,y)). \end{array}$$

- There are no free occurrences of z within the scope of $(\forall z.M_1)$.
- Next, use universal quantification to replace the two occurrences of R with a single one:

$$\begin{array}{rcl} \varphi_{tc}(R,x,y) & \equiv & (\forall z.M_1)(\exists z)(\forall uv.M_2)R(u,v), \\ M_2 & \equiv & (u = x \wedge v = z) \vee (u = z \wedge v = y). \end{array}$$

- Finally, requantify x and y:

$$\begin{array}{rcl} \varphi_{tc}(R,x,y) & \equiv & (\forall z.M_1)(\exists z)(\forall uv.M_2)(\exists xy.M_3)R(x,y), \\ M_3 & \equiv & (x = u \wedge y = v). \end{array}$$

Syntactic Uniformity of the Definition of REACH

- Define the quantifier block,

$$QB_{tc} \equiv (\forall z.M_1)(\exists z)(\forall uv.M_2)(\exists xy.M_3).$$

- We showed that an application of the operator φ_{tc} corresponds exactly to the writing of QB_{tc}, $\varphi_{tc}(R,x,y) \equiv [QB_{tc}]R(x,y)$.
- It follows that for any r,

$$\varphi_{tc}^r(\varnothing) \equiv [QB_{tc}]^r(\mathbf{false}).$$

- We have thus demonstrated a syntactic uniformity for the inductive definition of REACH.
- For any structure $\mathcal{A} \in \mathrm{STRUC}[\tau_g]$,

$$\begin{aligned} \mathcal{A} \in \mathrm{REACH} \quad &\Leftrightarrow \quad \mathcal{A} \vDash (\mathrm{LFP}\varphi_{tc})(s,t) \\ &\Leftrightarrow \quad \mathcal{A} \vDash ([QB_{tc}]^{\lceil 1+\log \|\mathcal{A}\| \rceil})\mathbf{false}(s/x,t/y). \end{aligned}$$

Introducing the Class FO$[t(n)]$

- We now define FO$[t(n)]$ to be the set of properties defined by quantifier blocks iterated $t(n)$ times.
- This is the same as being iterated $O(t(n))$ times since a quantifier block may be any constant size.
- Such expressions grow as a function of the size of their inputs.
- On the other hand, the number of variables is a fixed constant.

The Class $\mathrm{FO}[t(n)]$

Definition

A set $S \subseteq \mathrm{STRUC}[\tau]$ is a member of $\mathrm{FO}[t(n)]$ if there exist:

- Quantifier free formulas M_i, $0 \le i \le k$, from $\mathcal{L}(\tau)$;
- A tuple $\overline{c}$ of constants;
- A quantifier block $QB = [(Q_1 x_1.M_1)\cdots(Q_k x_k.M_k)]$,

such that, for all $\mathcal{A} \in \mathrm{STRUC}[\tau]$,

$$\mathcal{A} \in S \iff \mathcal{A} \models ([QB]^{t(\|\mathcal{A}\|)} M_0)(\overline{c}/\overline{x}).$$

- The reason for the substitution of constants is that the quantifier block QB may contain some free variables that must be substituted for to build a sentence.

Inductive Definitions and Number of Iterations

- Combining the normal form lemma with the preceding definition, we obtain

Lemma

For all $t(n)$ and all classes of finite structures,

$$\mathrm{IND}[t(n)] \subseteq \mathrm{FO}[t(n)].$$

- A converse of this lemma also holds.
- The proof is described in the next chapter.