Topics in Logic and Complexity

Handout 5

Anuj Dawar

http://www.cl.cam.ac.uk/teaching/1920/L15

Complexity of First-Order Logic

The problem of deciding whether $\mathbb{A} \models \phi$ for first-order ϕ is in time $O(ln^m)$ and $O(m \log n)$ space.

where n is the size of $\mathbb{A}$, l is the length of ϕ and m is the quantifier rank of ϕ.

We have seen that the problem is PSPACE-complete, even for fixed $\mathbb{A}$.

For each fixed ϕ, the problem is in L.

Is FO contained in an initial segment of P?

Is there a fixed c such that for every first-order ϕ, $\mathrm{Mod}(\phi)$ is decidable in time $O(n^c)$?

If P = PSPACE, then the answer is yes, as the satisfaction relation is then itself decidable in time $O(n^c)$.

Thus, though we expect the answer is no, this would be difficult to prove.

A more uniform version of the question is:

Is there a constant c and a computable function f so that the satisfaction relation for first-order logic is decidable in time $O(f(l)n^c)$?

In this case we say that the satisfaction problem is *fixed-parameter tractable* (FPT) with the formula length as parameter.

Parameterized Problems

Some problems are given a graph G and a positive integer k

Independent Set: does G contain k vertices that are pairwise distinct and non-adjacent?

Dominating Set: does G contain k vertices such that every vertex is among them or adjacent to one of them?

Vertex Cover: does G contain k vertices such that every edge is incident on one of them?

For each fixed value of k, there is a first-order sentence ϕ_k such that $G \models \phi_k$ if, and only if, G contains an independent set of k vertices.
Similarly for dominating set and vertex cover.

Parameterized Complexity

FPT—the class of problems of input size n and *parameter* l which can be solved in time $O(f(l)n^c)$ for some computable function f and constant c.

There is a hierarchy of *intractable* classes.

$$\mathrm{FPT} \subseteq W[1] \subseteq W[2] \subseteq \cdots \subseteq \mathrm{AW}[\star]$$

Vertex Cover is FPT.
Independent Set is $W[1]$-complete.
Dominating Set is $W[2]$-complete.

Parameterized Complexity of First-Order Satisfaction

Writing Π_t for those formulas which, in prenex normal form have t alternating blocks of quantifiers starting with a universal block:

> *The satisfaction problem restricted to Π_t formulas (parameterized by the length of the formula) is hard for the class $W[t]$.*

The satisfaction relation for first-order logic ($\mathbb{A} \models \phi$), parameterized by the length of ϕ is AW[$\star$]-complete.

Thus, if the satisfaction problem for first-order logic were FPT, this would collapse the edifice of parameterized complexity theory.

Restricted Classes

One way to get a handle on the complexity of first-order satisfaction is to consider restricted classes of structures.

Given: a first-order formula ϕ *and a structure* $\mathbb{A} \in \mathcal{C}$
Decide: if $\mathbb{A} \models \phi$

For many interesting classes $\mathcal{C}$, this problem has been shown to be FPT, even for formulas of MSO.

We say that satisfaction of FO (or MSO) is *fixed-parameter tractable on* $\mathcal{C}$.

Words as Relational Structures

For an alphabet $\Sigma = \{a_1, \ldots, a_s\}$ let

$$\sigma_\Sigma = (<, P_{a_1}, \ldots, P_{a_s})$$

where

$<$ *is binary; and* $P_{a_1}, \ldots, P_{a_s}$ *are unary.*

With each $w \in \Sigma^*$ we associate the canonical structure

$$S_w = (\{1, \ldots, n\}, <, P_{a_1}, \ldots, P_{a_s})$$

where

- n is the length of w
- $<$ is the natural linear order on $\{1, \ldots, n\}$.
- $i \in P_a$ if, and only if, the ith symbol in w is a.

Languages Defined by Formulas

The sentence ϕ in the signature σ_Σ defines:

$$\{w \mid S_w \models \phi\}.$$

The class of structures isomorphic to word models is given by:

$$lo(<) \land \forall x \bigvee_{a \in \Sigma} P_a(x) \land \forall x \bigwedge_{a,b \in \Sigma, a \neq b} (P_a(x) \to \neg P_b(x)),$$

where

$lo(<)$ *is the formula that states that* $<$ *is a linear order*

Examples

The set of strings of length 3 or more:

$$\exists x \exists y \exists z (x \neq y \wedge y \neq z \wedge x \neq z).$$

The set of strings which begin with an a:

$$\exists x (P_a(x) \wedge \forall y \, y \geq x)$$

The set of strings of even length:

$$\begin{array}{lll} \exists X & \forall x(\forall y & x \leq y) \rightarrow X(x) \wedge \\ & \forall x \forall y & (x < y \wedge \forall z (z \leq x \vee y \leq z)) \\ & & \rightarrow (X(x) \leftrightarrow \neg X(y)) \wedge \\ & \forall x(\forall y & y \leq x) \rightarrow \neg X(x). \end{array}$$

Examples

$(ab)^*$:

$$\begin{aligned}
&\forall x(\forall y \quad x \leq y) \rightarrow P_a(x) \land \\
&\forall x \forall y \quad (x < y \land \forall z(z \leq x \lor y \leq z)) \\
&\qquad\qquad \rightarrow (P_a(x) \leftrightarrow P_b(y)) \land \\
&\forall x(\forall y \quad y \leq x) \rightarrow P_b(x).
\end{aligned}$$

MSO on Words

Theorem (Büchi-Elgot-Trakhtenbrot)
A language L is defined by a sentence of MSO if, and only if, L is regular.

Recall that a language L is *regular* if:

- it is the set of words matching a *regular expression*; or equivalently
- it is the set of words accepted by some *nondeterministic finite automaton*; or equivalently
- it is the set of words accepted by some *deterministic finite automaton*.

Myhill-Nerode Theorem

Let $\sim$ be an equivalence relation on Σ^*.

We say $\sim$ is *right invariant* if, for all $u, v \in \Sigma^*$,

if $u \sim v$, then for all $w \in \Sigma^*$, $uw \sim vw$.

Theorem (Myhill-Nerode)

The following are equivalent for any language $L \subseteq \Sigma^*$:

- L is regular;
- L is the union of equivalence classes of a right invariant equivalence relation of finite index on Σ^*.

MSO Equivalence

We write $\mathbb{A} \equiv_m^{\mathsf{MSO}} \mathbb{B}$ to denote that, for all MSO sentences ϕ with $\mathrm{qr}(\phi) \leq m$,

$$\mathbb{A} \models \phi \quad \text{if, and only if,} \quad \mathbb{B} \models \phi.$$

We count both first and second order quantifiers towards the rank.

The relation $\equiv_m^{\mathsf{MSO}}$ has finite index for every m.

> *For any m, there are up to logical equivalence, only finitely many formulas with quantifier rank at most m, with at most k free variables.*

Invariance

Suppose u_1, u_2, v_1, v_2 are words over an alphabet Σ such that

$$u_1 \equiv_m^{\mathsf{MSO}} u_2 \quad \text{and} \quad v_1 \equiv_m^{\mathsf{MSO}} v_2$$

then $u_1 v_1 \equiv_m^{\mathsf{MSO}} u_2 v_2$.

Dulpicator has a winning strategy on the game played on the pair of words $u_1 v_1, u_2 v_2$ that is obtained as a composition of its strategies in the games on u_1, u_2 and v_1, v_2.

It follows that $\equiv_m^{\mathsf{MSO}}$ is *right invariant*.

For any MSO sentence ϕ, the language defiend by ϕ is the union of equivalence classes of $\equiv_m^{\mathsf{MSO}}$ where m is the quantifier rank of ϕ.

Regular Expressions to MSO

For the converse, we translate a regular expression r to an MSO sentence ϕ_r.

$r = \emptyset$: $\phi_r = \exists x(x \neq x)$.

$r = \varepsilon$: $\phi_r = \neg\exists x(x = x)$.

$r = a$: $\phi_r = \exists x \forall y(y = x \wedge P_a(x))$.

$r = s + t$: $\phi_r = \phi_s \vee \psi_t$.

$r = st$: $\phi_r = \exists x(\phi_s^{<x} \wedge \phi_t^{\geq x})$,
where $\phi_s^{<x}$ and $\phi_t^{\geq x}$ are obtained from ϕ_s and ϕ_t by relativising the first order quantifiers.
That is, every subformula of ϕ_s of the form $\exists y \psi$ is replaced by $\exists y(y < x \wedge \psi^{<x})$,
and similarly every subformula $\exists y \psi$ of ϕ_t by $\exists y(y \geq x \wedge \psi^{\geq x})$

Kleene Star

$r = s^*$:

$$
\begin{aligned}
\phi_r = \;& \phi_\varepsilon \vee \\
& \exists X \;\; \forall x (X(x) \wedge \forall y (y < x \rightarrow \neg X(y)) \rightarrow \phi_s^{<x}) \wedge \\
& \qquad \forall x (X(x) \wedge \forall y (y \geq x \rightarrow \neg X(y)) \rightarrow \phi_s^{\geq x}) \wedge \\
& \qquad \forall x \forall y \;\; (X < y \wedge X(x) \wedge X(y) \wedge \\
& \qquad\qquad \forall z (x < z \wedge z < y \rightarrow \neg X(z)) \\
& \qquad\qquad \rightarrow \phi_s^{\geq x, < y}),
\end{aligned}
$$

where $\phi_s^{\geq x, < y}$ is obtained from ϕ_s by relativising all first order quantifiers simultaneously with $< y$ and $\geq x$.

First-Order Languages

The class of *star-free* regular expressions is defined by:

- the strings $\emptyset$ and ε are star-free regular expressions;
- for any element $a \in A$, the string a is a star-free regular expression;
- if r and s are star-free regular expressions, then so are (rs), $(r + s)$ and $(\bar{r})$.

A language is defined by a first order sentence *if, and only if,* it is denoted by a star-free regular expression.

Applications

A class of linear orders is definable by a sentence of MSO if, and only if, its set of cardinalities is *eventually periodic*.

Some results on graphs:

> *The class of balanced bipartite graphs is not definable in MSO.*
>
> *The class of Hamiltonian graphs is not definable by a sentence of MSO.*

MSO is FPT on Words

There is a computable function f such that the problem of deciding, given a word w and an MSO sentence ϕ whether,

$$S_w \models \phi$$

can be decided in time $O(f(l)n)$ where l is the length of ϕ and n is the length of w.

The algorithm proceeds by constructing, from ϕ an *automaton* $\mathcal{A}_\phi$ such that the language recognized by $\mathcal{A}_\phi$ is

$$\{w \mid S_w \models \phi\}$$

then running $\mathcal{A}_\phi$ on w.

The automaton $\mathcal{A}_\phi$

The states of $\mathcal{A}_\phi$ are the equivalence classes of $\equiv_m^{\mathsf{MSO}}$ where m is the quantifier rank of ϕ.

We write $\mathsf{Type}_m^{\mathsf{MSO}}(\mathbb{A})$ for the set of all formulas ϕ with $\mathrm{qr}(\phi) \leq m$ such that $\mathbb{A} \models \phi$.

$\mathbb{A} \equiv_m^{\mathsf{MSO}} \mathbb{B}$ is equivalent to

$$\mathsf{Type}_m^{\mathsf{MSO}}(\mathbb{A}) = \mathsf{Type}_m^{\mathsf{MSO}}(\mathbb{B})$$

There is a single formula $\theta_\mathbb{A}$ that characterizes $\mathsf{Type}_m^{\mathsf{MSO}}(\mathbb{A})$.

It turns out that we can compute $\theta_{S_{w \cdot a}}$ from θ_{S_w}.