Regular graphs and the spectra of two-variable logic with counting

Eryk Kopczyński, Tony Tan

University of Warsaw,
Hasselt University and Transnational University of Limburg

Sep 19, 2013

Example

Consider models of the following formula ϕ_1:

$$\forall x\ P_1(x) \to \exists^{=2} y\ R(x,y) \wedge P_2(y)$$

$$\wedge$$

$$\forall y\ P_2(y) \to \exists^{=3} x\ R(x,y) \wedge P_1(x)$$

In the models of ϕ_1,
how many elements satisfy P_1 and P_2, respectively?

The main problem

For a formula ϕ over signature σ with unary relational symbols P_1, $\ldots$, P_d, define

$$\Psi(\phi) = \{(n_1, \ldots, n_d) : \phi \text{ has a model where } {=}n_i \text{ elements satisfy } P_i\}$$

For example $\Psi(\phi_1) = \{(n_1, n_2) : 2n_1 = 3n_2\}$

Problem: Characterize the set $\Psi(\phi)$ for the given formula ϕ, or the possible sets for formulae of given class

Spectra

This can be seen as a generalization of the well known **spectrum** (Scholz' 52):

$$\Psi(\phi) = \{(n_1, \ldots, n_d) : \phi \text{ has a model where } {=}n_i \text{ elements satisfy } P_i\}$$

$$\mathrm{spec}(\phi) = \{n : \phi \text{ has a model of size } n\}$$

- Ψ counts elements in many dimensions
- Ψ allows an arbitrary number of elements which do not satisfy any P_i (thus the membership problem is potentially undecidable)

Too many variables: images

Three variables allow simulating a Turing machine. Thus, for $S \in \mathbb{N}^d$, the following conditions are equivalent:

- S is recursively enumerable,
- $S = \Psi(\phi)$ for some $\phi \in FO_3$.

Too many variables: spectra

Let $\textbf{Spec}_k = \{\mathrm{spec}(\phi) : \phi \in FO_k\}$. Then:

- $\textbf{Spec}_k \subseteq \textbf{NTIME}(2^{(2k-1)n})$
- $\textbf{NTIME}(2^{mn}) \subseteq \textbf{Spec}_{3m}$

Corollary: $\textbf{Spec}_k \subsetneq \textbf{Spec}_{6k}$

Two variable logic with counting

Since three variables allow simulating a Turing machine, we consider **two variable logic with counting** FO_2C.

$$\forall x \, \exists^{\geq 3} y \; \big(R_1(x,y) \wedge \exists^{\leq 3} x \; R_2(x,y)\big)$$

This is a logic with good properties:

- Decidable (Grädel, Otto, Rosen '97)
- Related to modal logic

Two variable logic with counting: Example 1

$$\phi_1' :$$

$$\forall x\ P_1(x) \to \exists^{=2} y\ R(x,y) \land P_2(y)$$

$$\land$$

$$\forall y\ P_2(y) \to \exists^{=3} x\ R(x,y) \land P_1(x)$$

$$\land$$

$$\forall x\ (P_1(x) \land \neg P_2(x)) \lor (P_2(x) \land \neg P_1(x))$$

$\mathrm{spec}(\phi_1')$ is the set of possible sizes of a $2,3$-regular bipartite graph.

$$\Psi(\phi_1') = \{(n_1, n_2) : 2n_1 = 3n_2\}$$

$$\mathrm{spec}(\phi_1') = \{n : 5|n\}$$

Two variable logic with counting: Example 2

$$\phi_2:$$

$$\forall x\ \exists^{=5} y\ R(x,y) \wedge R(y,x)$$

$\mathrm{spec}(\phi_2)$ is the set of possible sizes of a 5-regular graph.

$$\mathrm{spec}(\phi_2) = \{n : n \neq 2, n \neq 4, 2|n\}$$

Two variable logic with counting: Example 3

$$\phi_3:$$

$$\forall x P_1(x) \to \exists^{=2} y R(x,y) \wedge \exists^{=5} y R(y,x)$$

$$\wedge$$

$$\forall x P_2(x) \to \exists^{=4} y R(x,y) \wedge \exists^{=3} y R(y,x)$$

$$\wedge$$

$$\forall x (P_1(x) \wedge \neg P_2(x)) \vee (P_2(x) \wedge \neg P_1(x))$$

This formula represents a directed graph with n_1 vertices with out-degree 2 and in-degree 5, and n_2 vertices with out-degree 4 and in-degree 3.

$$\Psi(\phi_3) = \{(n_1, n_2) : 3n_1 = n_2, n_1 \neq 1\}$$

Two variable logic with counting: Example 4

$$\phi_4 :$$

$$\forall x \bigvee_i^{=1} P_i(x)$$

$$\wedge$$

$$\forall x \exists^{=1} y R(x,y)$$

$$\wedge$$

$$\forall x \forall y R(x,y) \rightarrow \bigwedge_i (P_i(x) \rightarrow \neg P_i(y))$$

$$\Psi(\phi_4) = \{(n_1, n_2, n_3) : n_1 \leq n_2 + n_3, n_2 \leq n_3 + n_1, n_3 \leq n_1 + n_2\}$$

The main result

Theorem

Let ϕ be a formula of two-variable logic with counting. Then $\Psi(\phi)$ is definable in Presburger arithmetic (i.e., is a semilinear set).

Note that for each semilinear set S it is easy to construct a FO_2C formula ϕ such that $S = \Psi(\phi)$.

Corollary

A set of positive integers is a spectrum of a FO_2C formula iff it is eventually periodic.

Corollary

FO_2C spectra (and images) are closed under complement.

Proof: simplify the universe

Let ϕ be a FO_2C formula.
We can assume that:

- ϕ is over a signature including only unary relations $\mathcal{P} = \{P_1, \ldots, P_d\}$ and binary relations $\mathcal{R} = \{R_1, \ldots, R_l\}$
- for each two elements x, y, either $x = y$ or $R_i(x, y)$ for exactly one relation R_i
- For each relation $R \in \mathcal{R}$ there is a reverse relation $\overleftarrow{R} \in \mathcal{R}$, such that $R_i(x, y)$ iff $\overleftarrow{R_i}(y, x)$.

Proof: use modal logic

We transform ϕ into a formula of QMLC (quantified modal logic with counting).

$$\mathbf{MLC}\text{: } \psi ::= \neg\psi \mid P \mid \psi_1 \wedge \psi_2 \mid \Diamond_R^k \psi$$

$a \models \Diamond_R^k \psi$ iff there are at least k elements b such that $R(a, b)$ and $b \models \psi$

$$\mathbf{QMLC}\text{: } \phi ::= \neg\phi | \phi_1 \wedge \phi_2 | \exists^k \psi$$

where $\exists^k \psi$ (where $\psi \in$ MLC) means that there are at least k elements a such that $a \models \psi$

Proof: matrices

Let $\mathbb{B} = \{= 0, = 1, = 2, \ldots, \geq 0, \geq 1, \geq 2, \ldots\}$
Let $C \in \mathbb{B}^{l \times m}$, $D \in \mathbb{B}^{l \times n}$

Theorem

There is a Presburger formula $BiREG\text{-}COMP_{C,D}(X_1, \ldots, X_m, Y_1, \ldots, Y_n)$ such that $BiREG\text{-}COMP_{C,D}(M_1, \ldots, M_m, N_1, \ldots, N_n)$ holds iff there exists a complete bipartite graph $(U, V, E_1, \ldots, E_l)$ such that:

- $U = U_1 \cup \ldots \cup U_m$, $V = V_1 \cup \ldots \cup V_n$,
- $|U_i| = M_i$, $|V_i| = N_i$
- *each element in U_i has $C_{i,j}$ edges of type E_j*
- *each element in V_i has $D_{i,j}$ edges of type E_j*

Proof: types and functions

Each element a of the universe has a *type* (the set of MLC subformulae of ϕ which are satisfied in a). Let $\mathcal{T}$ be the set of all types.

Let $X_{T,f}$ be a variable (intuitively, the number of elements of type T whose number of edges to other types is given by a function $f : \mathcal{R} \times \mathcal{T} \to \mathbb{B}$; we consider only functions consistent with the semantics of T).

For each two types T_1, T_2 we use the previous Theorem to generate Presburger formulas to verify whether $X_{T_1,f}$ and $X_{T_2,f}$ are consistent. We also need another theorem for the case where $T_1 = T_2$.

Future research

How can we extend FO_2C while still keeping decidability?

For example, what about the logic $FO_2C(<)$, which has an access to a total order on the universe?

We know that $\Psi(\phi)$ for $\phi \in FO_2C(<)$ include reachability sets of Petri nets (so no longer semilinear, but still decidable – Kosaraju '82)

Summary

- Why three variables is too much
- Why two variables and counting is not too much
- Regular graphs
- Presburger formulae and semilinearity
- What about $FO_2C(<)$?

Thank you!