Fragments of ESO and linear-time computation

Joel Rybicki

Finite model theory seminar

20 April 2012

Outline

- Existential second-order logic
- Random access machines
- ESO and polynomial non-deterministic time
- Classifying graph problems
  - Positive results
  - Negative results
- Conclusions
- Open problems

Existential second-order logic

- Extend FO with second-order existential quantification of relation and function symbols
- Let $\sigma \cap \tau = \emptyset$ and $\psi \in \mathrm{FO}^{\sigma\cup\tau}$. Then

  $$\langle M, \tau \rangle \models \exists \sigma \psi$$

  iff there exists a model $\mathcal{M}^*$ such that

  $$\mathcal{M}^* = \langle M, \tau, \sigma \rangle \models \psi$$

Fragments of ESO

Definition

Let $k, d \in \mathbb{N}_+$ and τ be a vocabulary. We say that

$$\varphi \in \mathrm{ESO}^\tau[\#k, \forall d]$$

if it holds that

- $\varphi \equiv \exists \sigma \forall x_0 \cdots \forall x_{d-1} \psi$
- $\#\sigma \leq k$
- $\psi \in \mathrm{FO}^{\sigma \cup \tau}$ is quantifier-free

We focus on $\mathrm{ESO}[\#1, \forall 1]$.

Skolemization

Theorem (Skolem normal-form)

Every ESO*-formula* $\varphi \equiv \exists \sigma \psi$ *where* $\psi \in \mathrm{FO}$ *is logically equivalent to an* ESO*-formula*

$$\varphi^* \equiv \exists f_0 \cdots \exists f_n \forall x_0 \cdots \forall x_m \psi'$$

where ψ' *is a quantifier-free* FO*-formula and* $f_0, \ldots, f_n$ *are function symbols.*

Skolemization 2

Lemma

Let σ be a vocabulary such that $\#\sigma \leq 1$. Any formula

$$\varphi \equiv \exists\sigma\forall x_0\exists x_1 \cdots \exists x_m \psi$$

where ψ is quantifier-free first-order formula can be expressed in $\mathrm{ESO}[\#1, \forall 1]$.

Random access machines

Let τ be a finite vocabulary. A τ-machine decides a property of a finite τ-structure:

- Input: A finite τ-structure $\mathcal{M}$ where $\mathrm{Dom}(\mathcal{M}) = [n]$.
- Output: Accept or reject.

Solves decision problems such as

“Is the input graph $\mathcal{G} = \langle G, E^{\mathcal{G}} \rangle$ Hamiltonian?”

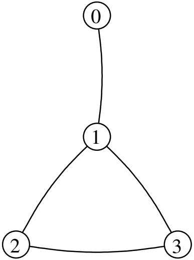
RAM: Memory

Turing machines use tape, τ-RAMs use *registers*.

- Cardinality register N initialized to n.
- Input register s for each $s \in \tau$.
- Accumulators $A, B_0, \ldots, B_{\#r}$ for accessing input.
- Working memory: $R_0, R_1, \ldots$

Example input

$$\mathcal{G} = \langle \{0, 1, 2, 3\}, E^{\mathcal{G}} \rangle$$



$i \diagup j$	0	1	2	3
0	0	1	0	0
1	1	0	1	1
2	0	1	0	1
3	0	1	1	0

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$
5. $A \leftarrow R_A$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$
5. $A \leftarrow R_A$
6. $B_i \leftarrow A$ where $i \in [\#\tau + 1]$
7. $R_A \leftarrow B_i$ where $i \in [\#\tau + 1]$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$
5. $A \leftarrow R_A$
6. $B_i \leftarrow A$ where $i \in [\#\tau + 1]$
7. $R_A \leftarrow B_i$ where $i \in [\#\tau + 1]$
8. $\mathrm{guess}(A)$ — guess an integer bounded by $O(n)$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$
5. $A \leftarrow R_A$
6. $B_i \leftarrow A$ where $i \in [\#\tau + 1]$
7. $R_A \leftarrow B_i$ where $i \in [\#\tau + 1]$
8. $\mathrm{guess}(A)$ — guess an integer bounded by $O(n)$
9. if $A = B_i$ then jump to I_a else jump to I_b, where $a, b \in [k + 1]$

RAM: Instructions

A τ-RAM has a sequence $\langle I_0, \ldots, I_k \rangle$ of instructions:

1. $A \leftarrow N$
2. $A \leftarrow s[B_0, \ldots, B_q]$ where $q = \#s$
3. $A \leftarrow 0$
4. $A \leftarrow A + 1$
5. $A \leftarrow R_A$
6. $B_i \leftarrow A$ where $i \in [\#\tau + 1]$
7. $R_A \leftarrow B_i$ where $i \in [\#\tau + 1]$
8. $\mathrm{guess}(A)$ — guess an integer bounded by $O(n)$
9. if $A = B_i$ then jump to I_a else jump to I_b, where $a, b \in [k+1]$
10. accept
11. reject

Time complexity

Definition

A τ-problem $P \subseteq \mathrm{Str}(\tau)$ belongs to $\mathrm{NTIME}[T(n)]$ if

- there exists a τ-NRAM that recognizes P
- for all finite τ-structures the running time is bounded by $O\big(T(n)\big)$

A refinement of Fagin's theorem

Theorem (Fagin, 1973)

$$\mathrm{ESO} \equiv \mathrm{NP}.$$

A refinement of Fagin's theorem

Theorem (Fagin, 1973)

$$\mathrm{ESO} \equiv \mathrm{NP}.$$

Theorem (Grandjean and Olive, 2004)

Let τ be a vocabulary and $d \in \mathbb{N}_+$, then

$$\mathrm{ESO}[\#d, \forall d] \equiv \mathrm{NTIME}[n^d].$$

Degree of the polynomial = number of FO-variables!

A proof?

Theorem (Grandjean and Olive, 2004)

Let $d \in \mathbb{N}_+$*, then*

$$\mathrm{ESO}[\#d, \forall d] \equiv \mathrm{NTIME}[n^d].$$

How to prove the case for $d = 1$?

- Similar to proof of Fagin's theorem
- Proving $\mathrm{ESO}[\#1, \forall 1] \subseteq \mathrm{NTIME}[n]$ is straightforward
- Showing $\mathrm{NTIME}[n] \subseteq \mathrm{ESO}[\#1, \forall 1]$ is not..

$\mathrm{ESO}[\#d, \forall d] \subseteq \mathrm{NTIME}[n]$

Let $\varphi \equiv \exists \sigma \forall x_0 \cdots \forall x_{d-1} \psi$ and suppose $\mathcal{M} = \langle [n], \tau \rangle$ is the input structure.

- Guess interpretations for symbols in $s \in \sigma$
  - For function symbols, guess $n^{\#s}$ values (the image)
  - For relation symbols, guess $n^{\#s}$ bits

$\mathrm{ESO}[\#d, \forall d] \subseteq \mathrm{NTIME}[n]$

Let $\varphi \equiv \exists\sigma\forall x_0 \cdots \forall x_{d-1}\psi$ and suppose $\mathcal{M} = \langle [n], \tau \rangle$ is the input structure.

- Guess interpretations for symbols in $s \in \sigma$
  - For function symbols, guess $n^{\#s}$ values (the image)
  - For relation symbols, guess $n^{\#s}$ bits
- Iteratively check that $\langle [n], \tau, \sigma \rangle \models \forall x_0 \cdots \forall x_{d-1}\psi$
  - Check n^d assignments for symbols $x_0, \ldots, x_{d-1}$

$\mathrm{ESO}[\#d, \forall d] \subseteq \mathrm{NTIME}[n]$

Let $\varphi \equiv \exists\sigma\forall x_0 \cdots \forall x_{d-1}\psi$ and suppose $\mathcal{M} = \langle [n], \tau \rangle$ is the input structure.

- Guess interpretations for symbols in $s \in \sigma$
  - For function symbols, guess $n^{\#s}$ values (the image)
  - For relation symbols, guess $n^{\#s}$ bits
- Iteratively check that $\langle [n], \tau, \sigma \rangle \models \forall x_0 \cdots \forall x_{d-1}\psi$
  - Check n^d assignments for symbols $x_0, \ldots, x_{d-1}$
- Running time is $O(|\sigma| n^d) + O(n^d)$

$\mathrm{NTIME}[n] \subseteq \mathrm{ESO}[\#1, \forall 1]$

Theorem

$$\mathrm{ESO}[\#1, \forall 1] \equiv \mathrm{NTIME}[n].$$

We need..

- A formula that simulates a τ-NRAM with $O(n)$ running time
- We need to refer to cn time steps using n elements
- A linear order over the domain
  - Easy if $d \geq 3$.. non-trivial if $d = 1$.

A linear order using a single FO-variable

Theorem (Grandjean 1990)

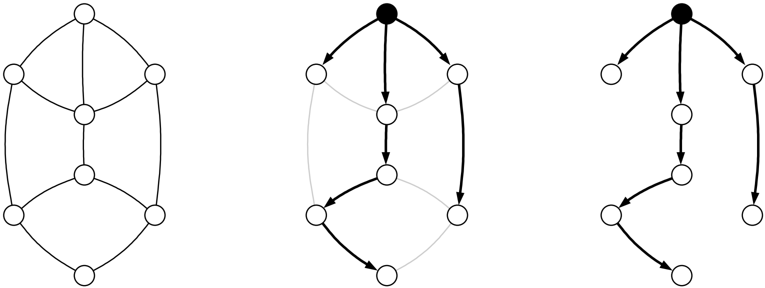
The class of structures definable in $\mathrm{ESO}[\#1, \forall 1]$ is not enlarged by the addition of a second-order quantifier $\exists_{ord} <$ which states that the symbol $<$ is a linear order over the domain.

- Construct a v-formula $\Phi \equiv \bigwedge_{1 \leq i \leq 20} \varphi_i$
- Defines a so-called arithmetical n-structure using only one universal quantification
- Use ESO-quantification $\exists v$ and replace $x_i < x_j$ with a formula $\mathrm{order}(x_i.x_j)$ that gives a linear order

Summary of the story so far

- A family of logics that capture polynomial non-deterministic time on RAMs
- Properties definable in $\mathrm{ESO}[\#1, \forall 1]$ can be checked in $O(n)$ time where n is the cardinality of the input structure
- NRAMs are slightly more powerful than Turing machines but a natural model of computation

Classifying graph problems



Graphs as structures

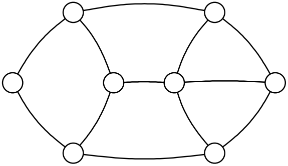
Definition

A relational structure $\mathcal{G} = \langle [n], E \rangle$ is a graph if

- The relation symbol E is binary, i.e., $\#E = 2$
- The interpretation of E is anti-reflexive and symmetric

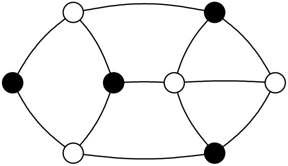
The complexity of graph problems

Many properties are very easy in $\mathrm{ESO}[\#1, \forall\mathbf{2}]$



The complexity of graph problems

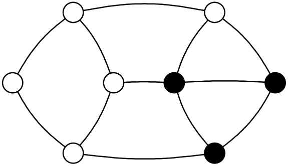
Many properties are very easy in $\mathrm{ESO}[\#1, \forall\mathbf{2}]$



- **independent sets**

The complexity of graph problems

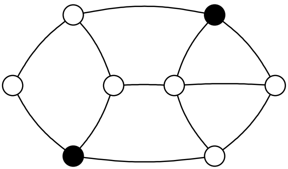
Many properties are very easy in $\mathrm{ESO}[\#1, \forall\mathbf{2}]$



- independent sets, **cliques**

The complexity of graph problems

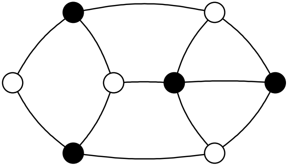
Many properties are very easy in $\mathrm{ESO}[\#1, \forall\mathbf{2}]$



- independent sets, cliques, **dominating sets**

The complexity of graph problems

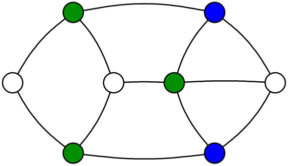
Many properties are very easy in $\mathrm{ESO}[\#1, \forall \mathbf{2}]$



- independent sets, cliques, dominating sets, **vertex covers**

The complexity of graph problems

Many properties are very easy in $\mathrm{ESO}[\#1, \forall\mathbf{2}]$



- independent sets, cliques, dominating sets, vertex covers, k**-colouring**

The complexity of graph problems

What about $\mathrm{ESO}[\#1, \forall\mathbf{1}]$?

The complexity of graph problems

What about $\mathrm{ESO}[\#1, \forall\mathbf{1}]$?

- It has been shown (Durand et al., 2004) that if the $\#\tau \leq 1$, then

$$\mathrm{ESO}^{\tau}[\#1, \forall 1] \equiv \mathrm{ESO}^{\tau}[\#1, \forall 2].$$

The complexity of graph problems

What about $\mathrm{ESO}[\#1, \forall \mathbf{1}]$?

- It has been shown (Durand et al., 2004) that if the $\#\tau \leq 1$, then

$$\mathrm{ESO}^{\tau}[\#1, \forall 1] \equiv \mathrm{ESO}^{\tau}[\#1, \forall 2].$$

- But $\#\tau = 2$ for graphs!

Nondeterministic vertex-linear time

Definition (Vertex-linear time)

The class NVLIN is the set of graph problems solvable in linear time on a NRAM.

- A graph property definable in $\mathrm{ESO}[\#1, \forall 1]$ can be decided in vertex-linear time
- If we have $\omega(n)$ edges, then the problem is solvable in sublinear time in the size of the graph

Positive results

- Recall that we can use the $\exists_{\text{ord}}$ operator in $\mathrm{ESO}[\#1, \forall 1]$
- The same theorem gives us a successor function S and the first element $\mathbf{0}$
- Denote $\exists_{\text{ord}}(<, S, \mathbf{0})$ as the quantification of the linear order and a corresponding successor function

Hamiltonicity of graphs

Definition

A *Hamiltonian cycle* of graph $\mathcal{G}$ is a cycle that spans $\mathcal{G}$.

The following sentence defines the existence of a Hamiltonian cycle:

$$\underset{\text{ord}}{\exists}(<, S)\forall x\Big(E\big(x, S(x)\big)\Big)$$

Dominating sets

Definition

A *dominating set* of $\mathcal{G} = \langle V, E \rangle$ is a subset $D \subseteq V$ such that for all $v \in V$ either $v \in D$ or there exists $(u, v) \in E$ such that $u \in D$.

The decision problem:

- Input: A graph $\mathcal{G}$ and $k \in \mathbb{N}_+$
- Output: A bit indicating if $\mathcal{G}$ has a dominating set of size at most k
- To encode k, add a unary relation K such that $|K| = k$

Dominating sets

$$\begin{aligned}\exists D \exists h \forall x \Big[&\exists y [x = h(y)] \\ &\wedge D(x) \to K\big(h(x)\big) \\ &\wedge \Big(D(x) \vee \exists z \big(D(z) \wedge E(x,z) \big) \Big) \Big]\end{aligned}$$

There is a set D and a function h such that

- h is a bijection
- $|D| \leq |K|$
- D is a dominating set

Lower bounds

- Class NVLIN contains NP-complete problems
  - Hamiltonian cycles
  - Dominating sets
  - Spanning tree isomorphism

Lower bounds

- Class NVLIN contains NP-complete problems
  - Hamiltonian cycles
  - Dominating sets
  - Spanning tree isomorphism
- At the same time, many problems of low deterministic complexity reside outside NVLIN
  - is the graph a tree (cycle-free)?
  - is the graph Eulerian?

Lower bounds

- Class NVLIN contains NP-complete problems
  - Hamiltonian cycles
  - Dominating sets
  - Spanning tree isomorphism
- At the same time, many problems of low deterministic complexity reside outside NVLIN
  - is the graph a tree (cycle-free)?
  - is the graph Eulerian?
- Reason: A linear time NRAM cannot check all possible edges!

The lower bound lemma

Lemma

Let P *be a graph problem. If there exists an infinite set*

$$A = \{(\mathcal{G}_n, F_n) \colon \mathcal{G}_n = \langle V, E \rangle, |V| = n, F_n \cap E = \emptyset\},$$

such that for any $(\mathcal{G}_n, F_n) \in A$ *the following properties hold:*

- $\mathcal{G}_n \in P$
- $|F_n| \in \Omega(n^2)$
- $\mathcal{G} + e \notin P$ *for any* $e \in F_n$*,*

then $P \notin \mathrm{NVLIN}$*.*

Negative results

Many properties cannot be checked in vertex-linear time..

- $\mathcal{G}$ is a tree
- $\mathcal{G}$ is *not* a tree
- $\mathcal{G}$ does not have a Hamiltonian cycle
- $\mathcal{G}$ graph has a vertex cover of size at most k
- See Grandjean and Olive, 2004 for more.

A corollary

Corollary

The class NVLIN *is not closed under complement.*

Proof:

- Hamiltonicity is in NVLIN
- Non-Hamiltonicity is not in NVLIN

Conclusions

- All formulas we have seen are of the form

$$\exists \sigma \forall x \varphi$$

  where $\varphi \in \mathrm{FO}^{\sigma \cup \tau}$ is quantifier-free.

- Positive results for non-deterministic linear time via $\mathrm{ESO}[\#1, \forall 1]$
- Non-expressibilty results for $\mathrm{ESO}[\#1, \forall 1]$ via NRAMs
- Several NP-complete problems are of low non-deterministic complexity..
- ..while many problems in P cannot be solved in non-deterministic linear time

Open problems

- The problems studied are solvable either in $O(n)$ time or require $\Omega(n^2)$ time.
- Are there *natural* graph problems that reside between these two?

Open problems 2

- Are there strict hierarchies between different logics $\mathrm{ESO}[\#k, \forall d]$?
- Conjecture: If $\#\tau \leq 1$, then $\mathrm{ESO}[\#1, \forall d] \equiv \mathrm{ESO}[\forall 1]$ holds.
  - It is known that $\mathrm{ESO}[\#d, \forall d] \equiv \mathrm{ESO}[\forall d] \equiv \mathrm{ESO}[\mathrm{Var}\ d]$

Open problems 2

- Are there strict hierarchies between different logics $\mathrm{ESO}[\#k, \forall d]$?
- Conjecture: If $\#\tau \leq 1$, then $\mathrm{ESO}[\#1, \forall d] \equiv \mathrm{ESO}[\forall 1]$ holds.
  - It is known that $\mathrm{ESO}[\#d, \forall d] \equiv \mathrm{ESO}[\forall d] \equiv \mathrm{ESO}[\mathrm{Var}\ d]$
- Corollary: If the conjecture holds, for all k and τ, $\mathrm{ESO}^{\tau}[\#k, \forall^*] \subsetneq \mathrm{ESO}^{\tau}[\#k+1, \forall^*]$.

References

- Etienne Grandjean (1990). *First-order spectra with one variable*.
- Etienne Grandjean and Frédéric Olive (2004). *Graph properties checkable in linear time in the number of vertices*.
- Arnaud, Durand, Etienne Grandjean, and Frédéric Olive (2004). *New results on arity vs. number of variables*.