Undecidable Extensions of Monadic Second-Order Logic over Words Using Cardinality Set Predicates relative to Recursive Sets.

Bruno Courcelle

Institut Universitaire de France & Université Bordeaux 1, LaBRI

References: B. C. and J. Engelfriet: *Graph structure and monadic second-order logic, A language theoretic approach*, book to be published by *Cambridge University Press*,

Monadic Second-order Logic is a favourite language to many researchers because:

1) It is **decidable** on many infinite structures and classes of finite and infinite structures;

2) On finite structures, there are **fixed-parameter tractable** model-checking algorithms on words, trees and graphs for structural parameters: tree-width, clique-width, rank-width;

Also:

3) MSOL replaces finite automata and transducers in the extension to graphs of formal language theory.
For a class of finite graphs and some logic, 2 problems:
Decidability? Time complexity of model-checking?

<table>
<thead>
<tr>
<th>Logic</th>
<th>Decidability</th>
<th>Model-checking</th>
</tr>
</thead>
<tbody>
<tr>
<td>FO, all graphs</td>
<td>Undecidable</td>
<td>Polynomial-time</td>
</tr>
<tr>
<td>MSOL2, tree-width &lt; k</td>
<td>Decidable</td>
<td>Linear-time</td>
</tr>
<tr>
<td>MSOL2, unbounded tree-width</td>
<td>Undecidable</td>
<td>Not FPT (*)</td>
</tr>
</tbody>
</table>

MSOL2 = MSOL over incidence graphs, allows edge set quantification.

(*) by S. Kreutzer, LICS 2010: exact statement is very technical.

**Remarks:** The positive results (decidability, linear-time) are based on finite automata. The negative results for MSOL2 are based on large grid minors. However, there is no implication between them.
How to improve the positive results?

No possible improvement based on waiving the conditions on tree-width for MSOL2. (But possible with \textit{bounded clique-width} for MSOL without edge set quantifications. Nothing new for words and trees.)

Keeping bounded tree-width, can one enrich the language?
Cardinality set predicates

If $R$ is a recursive set of integers, we let $\text{Card}_R(X)$ be the set predicate that means:

the cardinality of $X$ belongs to $R$.

Notation: language $\text{MSOL}^+ R$.

Validity is decidable on a single finite structure.

The constructions of finite automata for decidability and model-checking only work if $R$ is a semi-linear set (e.g., even numbers; more generally “equivalent to $p$ modulo $q$ if $> r$”).
Proposition: The satisfiability problem for MSOL+R is undecidable on words if $R = \{2^n, 3^n / n > 0\}$.

Open questions:

Cases of $R = \{2^n / n > 0\}$ or $R = \text{prime numbers}$.

Method: Using grids, as for the result for MSOL2.
**Fact:** If a set of graphs $L$ contains square grids $G_{n \times n}$ for infinitely many integers $n$, then the satisfiability problem for MSOL is **undecidable** on $L$.

**Proof:** A finite computation sequence of a Turing machine $T$ is a sequence of words that one can put on a large enough square grid. The existence of a finite terminating computation can be expressed by:

there is a large enough square grid and a choice of letters (from a fixed finite alphabet) for its vertices that encode a terminating computation. This is MSO expressible.
Reduction to this case.

If \( L \) is a set of finite structures, if there exist MSO\( \text{+}R \) formulas that define arbitrarily large square grids in (some of) the structures of \( L \), then the satisfiability problem for MSOL\( \text{+}R \) is undecidable on \( L \).

We will do that for words over 1 letter: \( D_n = ([n], \leq) \).
We will specify “large” sets \( A, B, C \) and bijections \( f : A \times B \rightarrow C \) (pairing functions).
From that one gets a rectangular grid with vertex set \( C \).
We let for $p, q$:

$A = \{ 2^k / 2 \leq k \leq p \}.$

$B = \{ 3^m / 1 \leq m \leq q \}.$

$C = \{ 2^k 3^m / 2 \leq k \leq p, 1 \leq m \leq q \}.$

Then a number $x > 3$ is a power of 2

$\iff \quad \text{Card}_R([x]) \& \text{Even}([x]) \quad \text{("even cardinality")}.$

A number $x > 3$ is a power of 3

$\iff \quad \text{Card}_R([x]) \& \text{Odd}([x]).$

Now:

$2^a + 2^b$ is a power of 2 $\iff a = b.$

$3^a + 3^b + 3^c$ is a power of 3 $\iff a = b = c.$
With these remarks:

Card(Y) is power of 2 & y is power of 3 & the union of any three intervals Xa has cardinality = power of 3

→ z = y.2^k for some k.

Similarly, we express that z = x.3^m, x power of 2.

→ z = x. y with x = 2^k, y = 3^m

→ Definition of arbitrary large rectangular and square grids
Similar constructions:

**Undecidable extensions of MSOL** (on words or trees):

Eq(X,Y) to mean Card(X) = Card(Y).

Auto(X) to that the graph induced on X has a nontrivial automorphism.
Related results:

1) Assuming “the linear case of Schinzel’s Hypothesis” about sequences of consecutive prime numbers:

The MSO-theory of \((\mathbb{N}, \leq, \text{Prime}(.))\) is decidable, (Bateman et al. 1993), hence the MSO-satisfiability problem for the structures \(([n], \leq, \text{Prime}(.))\) is also, which does not imply the decidability of MSOL+\text{Prime} on the structures \(([n], \leq)\) (because bijections are not MSO-definable)
2) The MSO-theories of \((\mathbb{N}, \leq, 2^{\text{some}}(.))\) and of \((\mathbb{N}, \leq, 3^{\text{some}}(.))\) are decidable (Elgot & Rabin 1966), but this does not imply that the third one \((\mathbb{N}, \leq, 2^{\text{some}}(.), 3^{\text{some}}(.))\) is. (Semenov)

**Question**: Is the MSO-satisfiability problem of the structures \(([n], \leq, 2^{\text{some}} v 3^{\text{some}}(.))\) is decidable?

**Open question**: For which recursive sets \(R\) is the satisfiability problem for MSOL+\(R\) decidable on words?

What about \(R = \{2^n \mid n > 0\}\) and \(R = \) prime numbers