

Monadic second-order logic for graphs.
Algorithmic and language theoretical applications

Exercises, open questions, appendix

Bruno Courcelle

Reference : Graph structure and monadic second-order logic,
book to be published by Cambridge University Press, readable on :
<http://www.labri.fr/perso/courcell/ActSci.html>

Exercises : 1. Equational sets of graphs

- 1.1) Prove that $\{a^n b^n c^n \mid n > 0\}$ and the set of *square words* (ww) are HR-equational.
- 1.2) Construct HR equation systems for outerplanar and Halin graphs.
- 1.3) Construct an HR equation system for series-parallel graphs having an even number of vertices.
- 1.4) Construct a VR equation system for trees having a number of nodes multiple of 3.
- 1.5) Construct a VR equation system for cographs having an even number of edges.
- 1.6) Prove that the non-context-free language $\{a^n \mid n=2^p \text{ for some } p \geq 0\}$ is HR-equational for some appropriate algebra extending the monoid of words.
- 1.7) Complete the proof of the algebraic characterization of tree-width : transform a tree-decomposition into a term of the HR algebra defining the same graph.

Exercises : 2. Recognizable sets

2.1) Construct congruences proving that the set of connected graphs is **HR**- and **VR**-recognizable.

2.2) Prove that the image of a recognizable language under an alphabetical homomorphism **h** (replacement of letter **a** by **h(a)**) is recognizable, by constructing a congruence for the image from one for the given language.

2.3) Prove the *Filtering Theorem*.

Hint : Let (L_1, \dots, L_n) be the least solution of a system **S** in $P(\mathbf{M})$ and **h** : $\mathbf{M} \rightarrow \mathbf{A}$ be a homomorphism with **A** finite. Construct a system **S'** with unknowns $x_{i,a}$ for all $i = 1, \dots, n$ and **a** in **A**, such that the component of the least solution of **S'** corresponding to $x_{i,a}$ is $L_i \cap h^{-1}(a)$.

2.4) Prove “inductively” that every series-parallel graph is **3**-colorable.

Exercises : 3. Monadic second-order logic

- 3.1) Write an MS-sentence expressing that the considered graph is a tree.
- 3.2) Write an MS-formula with free variables x, y, z expressing that, in a tree (undirected and unrooted), if one takes x as root, then y is an ancestor of z .
- 3.3) Write an MS_2 -sentence expressing that a graph has a Hamiltonian cycle.
- 3.4) A nonempty word over alphabet $\{a, b\}$ can (also) be considered as a directed path given with unary relations lab_a and lab_b representing the sets of occurrences of letters a and b . Prove that every regular language over $\{a, b\}$ is MS-definable.
- 3.5) A complete bipartite graph $K_{n,m}$ has a Hamiltonian cycle iff $n=m$. Construct such a graph “over” any word in $\{a, b\}^+$ having at least one a and at least one b . Deduce from 4) that Hamiltonicity is not MS-expressible.
- 3.6) Write an MS_2 -sentence expressing that a graph has a spanning tree of degree ≤ 3 . Show as in 3.5) that this property is not MS-expressible.

Exercises : 4. Monadic second-order transductions

- 4.1) Construct an MS-transduction that associates with a directed graph G the *directed acyclic graph* D of its strongly connected components. (The vertices of D are chosen among those of G).
- 4.2) Let G be undirected. For each k , $G(k)$ is the graph with same vertices and an edge $x—y$ iff x and y are at distance at most k . Define an MS-transduction that transforms G into $G(k)$.
- 4.3) If G has clique-width $\leq d$, then $G(k)$ has clique-width $\leq f(k,d)$ for some function f . Try to prove this (without looking at the next section). No such function exists for tree-width.
- 4.4) Prove that the transformation of a graph G into its *incidence graph* $\text{Inc}(G)$ is a MS-transduction on trees, and also on graphs of degree $\leq d$, for each fixed d .

A few open questions

Question 1 : What should be the clique-width or rank-width of hypergraphs (or relational structures) ?

Question 2 : Which graph operations, quantifier-free definable or not, yield extensions of the signatures VR, HR that are *equivalent* to them, i.e., that define the same recognizable and equational sets ?

(some answers already given by A. Blumensath, B.C., P. Weil)

Or that yield larger classes of equational sets for which MS logic is decidable ?

Question 3 : Is it true that the decidability of the MS (and not of the C_2 MS) satisfiability problem for a set of graphs implies bounded clique-width, as conjectured by Seese ?

More important (personal opinion) :

Question 4 : What about Question 3 for sets of hypergraphs or relational structures ?

Other questions (and answers) are in the proceedings of ICALP 2008

Appendix : The “Büchi” style proof of the Recognizability Theorem.

The case of terms in $T(F)$, for F *finite* signature.

Logical representation of terms : Nodes of the syntactic tree, *son* binary relation, labels saying “my label is f ” and “I am the i -th brother”.

Construction: automaton $A(\varphi)$ for sentence φ that defines $\text{Models}(\varphi) \cap T(F)$ by induction on sentences (without first-order variables and \forall).

For the case $\exists X. \varphi(X)$, we need $A(\varphi(X))$, and more generally $A(\varphi(X_1, \dots, X_n))$: f in F is replaced by $(f, (w_1, \dots, w_n))$ in $F \times \{0, 1\}^n$ of same arity, a term t in $T(F \times \{0, 1\}^n)$ encodes $\text{pr}_0(t)$ in $T(F)$ and an assignment $v(t) : \{X_1, \dots, X_n\} \rightarrow P(\text{Nodes}(t))$

(if u is an occurrence of $(f, (w_1, \dots, w_n))$, then $w_i = 1$ iff $u \in X_i$)

We construct $A(\varphi(X_1, \dots, X_n))$ that defines

$$\{ t \in T(Fx\{0,1\}^n) \mid (\text{pr}_0(t), v(t)) \models \varphi \}$$

Then : $L(A(\exists X_{n+1} . \varphi(X_1, \dots, X_{n+1}))) = \text{pr}(L(A(\varphi(X_1, \dots, X_{n+1}))))$

where pr is the projection that eliminates the last Boolean.

One obtains a *nondeterministic* automaton ; the case of negation needs determinization. The case of atomic formulas is easy.

The number of states is an *h -iterated exponential*, where h = maximum nesting of negations. This is not avoidable.

The case of graphs : VR_k -recognizability

We fix k , a bound on clique-width of the considered graphs and

F_k = the binary operation \oplus and $O(k^2)$ unary operations (*add-edg_{a,b}*, *relab_{a,b}*)

We use a single constant **1** defining one vertex labelled by 1.

(1 can be changed into another label).

A term t in $T(F_k \cup \{1\})$ defines a graph $G(t)$ with vertex set $V_{G(t)} =$
the set of occurrences of **1** in t . For representing assignments

$$v: \{X_1, \dots, X_n\} \rightarrow P(V_{G(t)})$$

we replace **1** by the constants $(1, (w_1, \dots, w_n))$ where $w_i \in \{0, 1\}$.

A term t in $T(F_k \cup (\{1\} \times \{0, 1\}^n))$ encodes the graph $G(\text{pr}_0(t))$ and
some assignment $v(t)$.

As in the case of terms, we can construct for all relevant k, n and φ a finite deterministic automaton $A(\varphi(X_1, \dots, X_n))$ that defines

$$\{ t \in T(F_k \cup (\{1\} \times \{0,1\}^n)) \mid (G(\text{pr}_0(t)), v(t)) \models \varphi \}$$

Inductive steps :

$\varphi \wedge \psi, \varphi \vee \psi$: product of two automata

$\neg \varphi$: exchange of accepting / non-accepting states of a **deterministic** automaton

$\exists X_{n+1} \varphi$: “projection” **pr** : makes automata **non-deterministic**

Basic cases : atomic formulas.

Example : the automaton $A(\text{edg}(X_1, X_2))$ with k^2+k+3 states

Graph labels are in $[k] = \{1, \dots, k\}$.

States : 0, Ok, $a(1)$, $a(2)$, ab , Error, for a, b in $[k]$, $a \neq b$

Meanings of states (at occurrence u in term t that defines a graph G)

0 : $X_1 = \emptyset$, $X_2 = \emptyset$

Ok : $X_1 = \{v\}$, $X_2 = \{w\}$, $\text{edg}(v, w)$ true in $G(t/u)$ defined by subterm t/u

$a(1)$: $X_1 = \{v\}$, $X_2 = \emptyset$, v has label a in $G(t/u)$

$a(2)$: $X_1 = \emptyset$, $X_2 = \{w\}$, w has label a in $G(t/u)$

ab : $X_1 = \{v\}$, $X_2 = \{w\}$, v has label a , w has label b

and $\text{edg}(v, w)$ not true, all that in $G(t/u)$,

Error : all other cases.

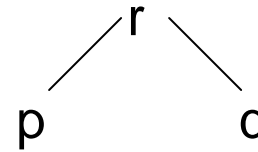
Accepting state : Ok

Transition rules

For the constant **1** :

(1,00) \rightarrow 0 ; **(1,10)** \rightarrow **1(1)** ; **(1,01)** \rightarrow **1(2)** , **(1,11)** \rightarrow Error

For the binary operation \oplus :



If $p = 0$ then $r = q$

If $q = 0$ then $r = p$

If $p = a(1)$ and $q = b(2)$, $a \neq b$ then $r = ab$

If $p = b(2)$ and $q = a(1)$, $a \neq b$ then $r = ab$

Otherwise $r = \text{Error}$

For unary operations *add-edge*_{a,b}

r
|
p

If $p = ab$ then $r = Ok$ else $r = p$

For unary operations *relab*_a \longrightarrow *b*

If $p = a(i)$ where $i = 1$ or 2 then $r = b(i)$

If $p = ac$ where $c \neq a$ and $c \neq b$ then $r = bc$

If $p = ca$ where $c \neq a$ and $c \neq b$ then $r = cb$

If $p = 0$ or Ok or $Error\ c(i)$ or cd or dc where $c \neq a$
then $r = p$

We have **only** proved that for all k and MS sentence φ

1) $\{ t \in T(F_k \cup \{1\}) \mid G(t) \models \varphi \}$ is $T(F_k \cup \{1\})$ -recognizable

2) The set of models of φ of clique-width $\leq k$, is VR_k -recognizable

This is **NOT** the **Recognizability Theorem** : it needs a type-preserving, locally finite congruence on ALL graphs (of unbounded clique-width).

VR_k -recognizability for each k *does not* imply VR -recognizability.

The set of duplicated square grids $G_{m \times m} \oplus G_{m \times m}$ is not VR -recognizable but each of its restrictions to graphs of clique-width at most k is finite hence VR_k -recognizable (by the proved special case).

However this proof is sufficient for **FPT algorithms** and for the **Filtering Theorem** because in these cases only graphs of clique-width at most some k are concerned.