On Equivalent Representations of Infinite Structures

Arnaud Carayol and Thomas Colcombet
{Arnaud.Carayol, Thomas.Colcombet}@irisa.fr

Irisa, Campus universitaire de Beaulieu,
35042 Rennes Cedex, France

Abstract. According to Barthelman and Blumensath, the following families of infinite graphs are isomorphic: (1) prefix-recognisable graphs, (2) graph solutions of VR equational systems and (3) MS interpretations of regular trees. In this paper, we consider the extension of prefix-recognisable graphs to prefix-recognisable structures and of graphs solutions of VR equational systems to structures solutions of positive quantifier free definable (PQFD) equational systems. We extend Barthelman and Blumensath's result to structures parameterised by infinite graphs by proving that the following families of structures are equivalent: (1) prefix-recognisable structures restricted by a language accepted by an infinite deterministic automaton, (2) solutions of infinite PQFD equational systems and (3) MS interpretations of the unfoldings of infinite deterministic graphs. Furthermore, we show that the addition of a *fuse* operator, that merges several vertices together, to PQFD equational systems does not increase their expressive power.

1 Introduction

The automatic verification of properties on infinite structures is an important technique for proving behavioural properties on programs. A natural encoding of a program behaviour is an infinite directed graph where vertices are states of the machine, and edges mimic the transition steps of the program. Properties on the program can then be expressed as logical formulas referring to this graph (or its unfolding when considering *e.g* temporal logics). The problem of model-checking is then to decide the satisfaction of a formula over the graph. This problem is usually undecidable. However, on certain families of infinite graphs and for some given logics the model-checking problem is decidable.

In this work, we are dealing with monadic second-order (MS) logic: an extension of first-order logic which allows quantification over sets of vertices. The first decidability result for this logic over an infinite graph was provided by Büchi for the infinite semi-line. Rabin extended this result to the infinite binary tree.

With the work of Muller and Schupp on pushdown graphs [MS85], the focus of study shifted from infinite graphs to families of infinite graphs.

Many families of graphs have since been presented with various decidability and structural properties. Those families can be classified according to their representation into three categories.

The equational representation describes an infinite structure as the solution of an equational system. The family of structures (or graphs) obtained in this way depends on the choice of the operators. The most famous examples are hyperedge replacement equational structures (HR) [Cou89] and the vertex replacement equational graphs (VR) [Cou90]. The VR operators also have been extended into vertex replacement with product operators [Col02].

The transformational representation consists in applying some finite sequence of transformations over an already-known structure. Transformations can be the unfolding of graphs [CW98], Shellah-Muchnik-Walukiewicz tree graph construction [Wal96], or logically defined transformations (FO-interpretations, inverse finite or rational mappings [Cau96,Urv02], MS interpretations or general MS-definable transductions [Cou94]).

The internal representation amounts to give an exact description of both the universe and the relations of the structure. The most used universe is the set of words over a given finite alphabet. Relations over words can then be defined by means of many techniques:

Rewriting: Prefix (or suffix) rewriting of words describes the family of push-down graphs [MS85,Cau92]. When the set of rules is recognisable, it leads to prefix-recognisable graphs [Cau96].

Transductions: Relations recognised by synchronised transductions describe the class of automatic graphs [Sén92] and structures [Blu99]. When the transduction is rational, it defines the rational graphs [Mor00].

Structures defined over the universe of closed terms have also been presented [DT90,Blu99,Löd02,Col02].

The above mentioned techniques are not independent from each other. Many connections have been stated in the literature. In our case we are specially involved with the following: the graphs solution of VR equational systems are isomorphic to prefix-recognisable graphs [Bar97] and to MS interpretations of infinite regular trees [Blu01].

To some extent, these classes of graphs are defined upon *finite* objects. In particular, a VR-equational graph is the solution of a *finite* system of equations and prefix-recognisable graph is a rewriting system restricted to the language accepted by a *finite* automaton. These two kinds of graphs are equivalent and can be obtained by MS-definable transduction of the unfolding of a *finite* graph.

We generalise this triple equivalence to structures defined by *infinite* objects. We show that interpretation of infinite systems of PQFD equations (which are a natural extension of VR operators introduced in [Bar97]), *PR*-structures restricted by infinite deterministic automaton and MS-definable transductions of the unfoldings of deterministic infinite graphs are equivalent. Furthermore, this equivalence is effective in the sense that MS-definable transductions link the system of equations, the automaton and the graph.

In [CM02], the authors proved that the addition of a *fuse* operator to VR operators does not increase the expressive power for finite systems of equations. We naturally investigated the infinite counterpart of this result and proved under reasonable technical restrictions that the addition of a *fuse* operator to PQFD operators does not increase their expressive power.

The rest of the paper is divided as follows. Section 2 introduces the basic definitions. Section 3 presents structures defined by equational systems and Section 4 defines unfolding and states the first inclusion and Section 5 introduces PR-systems and states the last two inclusions.

2 Definitions

Relational structures

We define the *global signature* Ξ to be equal to $\bigcup_{n>0} \Xi_n$ where Ξ_n is an infinite set of symbols of *arity* n. For any R in Ξ, $|R|$ designates the arity of R.

A *relational structure* $\mathcal{S}$ is a pair $(\mathcal{U}, Val)$ where $\mathcal{U}$ is an 'at most countable' set called the universe and Val associates to a symbol of arity n a subset of $\mathcal{U}^n$.We will write $R^{\mathcal{S}}$ instead of $Val(R)$. Moreover, we suppose that Val has a finite support (i.e. the set of R such that $Val(R) \neq \emptyset$ is finite). A *signature* Σ of $\mathcal{S}$ is a *finite* set which contains the support of Val.

The *restriction* of a structure $\mathcal{S} = (\mathcal{U}, Val)$ to a universe $\mathcal{U}' \subseteq \mathcal{U}$ is denoted $\mathcal{S}|_{\mathcal{U}'}$ and designates the structure $(\mathcal{U}', Val')$ where $Val'(R) = Val(R) \cap (\mathcal{U}')^{|R|}$.

Two structures $\mathcal{S}$ and $\mathcal{S}'$ of respective universe $\mathcal{U}$ and $\mathcal{U}'$ are *isomorphic* (written $\mathcal{S} \approx \mathcal{S}'$) if there exists a one to one mapping ρ from $\mathcal{U}$ onto $\mathcal{U}'$ such that for any symbol $R \in \Xi$, $R^{\mathcal{S}}(x_1, \ldots, x_n) \Leftrightarrow R^{\mathcal{S}'}(\rho(x_1), \ldots, \rho(x_n))$.

Graphs

A *directed graph* G labelled by a finite set E is a relational structure admitting a signature with binary symbols only (identified with E). The universe is denoted by V and its elements are called vertices. A directed graph is *rooted* if its signatures contain an unary relation *root* which is interpreted as a singleton. By slight abuse, we will use the constant *root* in our formulas. The graph is said to be *deterministic* if for any $x, y, z \in V$ and for any relation $e \in E$, if $e(x, y)$ and $e(x, z)$ then $y = z$.

A *path* π in a graph G labelled by E is a finite sequence $v_1 e_1 \ldots e_{n-1} v_n$ in $(VE)^* V$ such that for all $i \in [1, n-1]$, $e_i(v_i, v_{i+1})$. Furthermore if the graph is rooted, the path is rooted if $v_1 = root$. For any $w \in E^*$, we write $x \stackrel{w}{\Longrightarrow} y$ if there exists a path $v_1 e_1 \ldots e_{n-1} v_n$ between x and y such that $w = e_1 \ldots e_{n-1}$. For $W \subseteq E^*$ a language, $x \stackrel{W}{\Longrightarrow} y$ holds iff for some $w \in W$, $x \stackrel{w}{\Longrightarrow} y$.

Given a graph G labelled by E of universe V and a finite set of *fresh* binary symbols $\mathcal{K} = \{k_1, \ldots, k_n\}$ (i.e $\mathcal{K} \cap E = \emptyset$), the $\mathcal{K}$-copying of G is the graph G' of universe $V \times [0, n]$ and such that for any relation $R \in E$, $R^{G'} = \{((x_1, 0), \ldots, (x_{|R|}, 0)) \mid (x_1, \ldots, x_{|R|}) \in R^G\}$ and for $k_i \in \mathcal{K}$, $k_i^{G'} = \{((x, 0), (x, i)) \,|x \in V\}$.

Example 1. Throughout this paper we illustrate all the techniques for describing structures with one example: the step-ladder graph depicted in Figure 1.

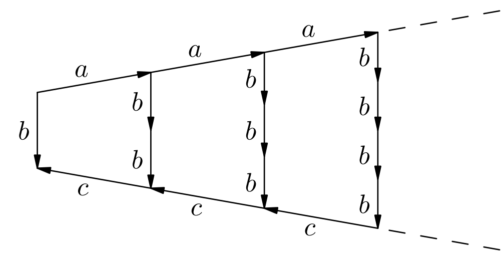

Fig. 1. The step-ladder graph.

Monadic second order logic

In the following, we assume that first order variables are ranged over by $x, y, z\dots$ whereas monadic second order variables are ranged over by $X, Y, Z\dots$ First order variables are interpreted as elements of the universe whereas monadic second order variables are interpreted as subsets of the universe. The atomic predicates of monadic formulas are $x \in X$, $x = y$ and $R(x_1, \dots, x_{|R|})$. Monadic formulas are then inductively defined as $\exists X.\, \Phi$, $\exists x.\, \Phi$, $\neg\Phi$ and $\Phi \vee \Psi$ for Φ and Ψ formulas.

MS formulas have the usual semantic [Tho97]. If $\Phi(x_1, \dots, x_n)$ is an MS formula and if $(u_1, \dots, u_n)$ is a tuple of elements of $\mathcal{U}$, $\mathcal{S} \models \Phi(u_1, \dots, u_n)$ means that $\mathcal{S}$ models Φ when x_i is interpreted by u_i for all $i \in [1, n]$.

A *MS interpretation* $\mathcal{I}$ is given by a MS formula $\delta(x)$ together with a finite set of formulas $(\Phi_R)_{R \in \Sigma}$ where Φ_R has free variables in $\{x_1, \dots, x_{|R|}\}$. $\mathcal{I}$ associates to each structure $\mathcal{S}$ of universe $\mathcal{U}$ the structure $\mathcal{I}(\mathcal{S})$ of universe $\mathcal{U}^{\mathcal{I}(\mathcal{S})} = \{x \in \mathcal{U} \mid \mathcal{S} \models \delta(x)\}$ and such that if $R \in \Sigma$, $R^{\mathcal{I}(\mathcal{S})} = \left\{\bar{x} \in (\mathcal{U}')^{|R|} \mid \mathcal{S} \models \Phi_R(\bar{x})\right\}$ (if $R \notin \Sigma$, $R^{\mathcal{I}(\mathcal{S})} = \emptyset$).

An *MS-definable transduction* $\mathcal{T}r$ [Cou94] is the composition of a $\mathcal{K}$-copying operation and an MS interpretation. This transformation preserves the decidability of the MS theory.

3 Equational systems

In this section we present how to describe infinite structures as solutions of equational systems over a given set of operators. Classical examples of this approach are hyperedge replacement systems [Cou89] or vertex replacement (VR) [Cou90] graphs.

For the rest of this section, we fix a signature Σ. For $\mathcal{V}$ a given set of variable names, we write $B^+(\mathcal{V})$ the set of positive boolean formulas over variables $\mathcal{V}$. Those formulas are built upon predicates of the signature applied to variables in $\mathcal{V}$, of the boolean connectives $\wedge$ and $\vee$, and the constants $\mathbf{t}$ (true) and $\mathbf{f}$ (false). Quantifiers as well as negation are not allowed.

We use the set of symbols $PQFD$ defined by:

$$\begin{aligned} PQFD &= PQFD_0 \cup PQFD_1 \cup PQFD_2 \\ PQFD_0 &= \{\mathbf{one}\} \\ PQFD_1 &= \{\mathbf{pqfd}[(\phi_R)_{R\in\Sigma}] \mid \forall R \in \Sigma,\ \phi_R \in B^+(x_1, \ldots, x_{|R|})\} \\ PQFD_2 &= \{\oplus\} \end{aligned}$$

Symbol in $PQFD_i$ have arity i. The mapping $\underline{\ }$ gives their semantic:

Singleton structure one: $\mathcal{U}^{\underline{\mathbf{one}}} = \{0\}$ and $R^{\underline{\mathbf{one}}} = \emptyset$ for any symbol R ,
Positive quantifier-free definable interpretation pqfd$[(\phi_R)]$**:**
given a relational structure $\mathcal{S}$, $\quad \mathcal{U}^{\underline{\mathbf{pqfd}[(\phi_R)]}(\mathcal{S})} = \mathcal{U}^{\mathcal{S}}$,
and $R^{\underline{\mathbf{pqfd}[(\phi_R)]}(\mathcal{S})}(u_1, \ldots, u_{|R|})$ iff $\mathcal{S} \models \phi_R(x_1 \leftarrow u_1, \ldots, x_{|R|} \leftarrow u_{|R|})$.
Disjoint union $\oplus$**:** given two structures $\mathcal{S}_1$ and $\mathcal{S}_2$,
$\mathcal{U}^{\mathcal{S}_1 \underline{\oplus} \mathcal{S}_2} = \{1\} \times \mathcal{U}^{\mathcal{S}_1} \cup \{2\} \times \mathcal{U}^{\mathcal{S}_2}$ and,
for any symbol R, $R^{\mathcal{S}_1 \underline{\oplus} \mathcal{S}_2} = \{((1, x_1), \ldots, (1, x_{|R|})) \mid R^{\mathcal{S}_1}(x_1, \ldots, x_{|R|})\}$
$\cup \ \{((2, x_1), \ldots, (2, x_{|R|})) \mid R^{\mathcal{S}_2}(x_1, \ldots, x_{|R|})\}$,

A very similar set of operators has already been introduced by Barthelman[Bar97].

Let us emphasize that this set of operators provides a strict and natural extension to relational structures of vertex replacement (VR) graph operators. Let us illustrate how to obtain VR systems with PQFD systems on graphs. The usual definition of VR operators works over coloured directed graphs: directed graphs labelled by a finite set E and extended with a mapping which associates to each vertex a color belonging to some given finite set C of colors. In our case, we can encode such a graph into a structure over the signature $\Sigma = C \cup E$ where symbols in C and E have respective arity 1 and 2 and encode respectively the fact that a vertex has a given color, and the presence of an edge between two nodes. We can now introduce the four VR operators, and their equivalent PQFD expression.

Single vertex constant of color c — simply written $\mathbf{c}$ — represents the graph with one vertex of colour c and no edge.
It can be expressed as $\mathbf{pqfd}[\phi_c](\mathbf{one})$ with $\phi_c(x) = \mathbf{t}$.
Disjoint union — written $\oplus$ as for structures — performs the disjoint union of two coloured graphs.
It can naturally be encoded by the disjoint union of structures $\oplus$.
Renaming color c_1 into color c_2 of a coloured graph G — written $\mathbf{ren}_{c_1,c_2}(\mathrm{G})$ — changes the color mapping in such a way that every vertex of color other than c_1 keeps its original color, and vertices of original color c_1 have new color c_2.
Let us suppose for simplicity that $c_1 \neq c_2$. The renaming operator can be encoded by $\mathbf{pqfd}[(\phi_R)_{R\in\Sigma}](G)$ with $\phi_{c_2} = c_1(x_1) \vee c_2(x_1)$, $\phi_{c_1} = \mathbf{f}$, $\phi_c = c(x_1)$ for $c \in C - \{c_1, c_2\}$, and $\phi_a = a(x_1, x_2)$ for $a \in E$.
Adding edges labelled by a between color c_1 and color c_2 to a graph G — written $\mathbf{add}_{c_1,c_2,a}(G)$ — adds to the coloured graph G all possible edges labelled by a with as origin a vertex of color c_1 and as destination a vertex

of color c_2.
The edge-adding operator can be encoded by $\mathbf{pqfd}[(\phi_R)_{R\in\Sigma}](G)$ with $\phi_a = a(x_1, x_2) \vee (c_1(x_1) \wedge c_2(x_2))$, $\phi_b = b(x_1, x_2)$ for $b \in E - \{a\}$ and $\phi_c = c(x_1)$ for any $c \in C$.

Pqfd operators can be used in equational systems: One can equip structures with the partial order of inclusion defined by $\mathcal{S} \subseteq \mathcal{S}'$ iff $\mathcal{U}^{\mathcal{S}} \subseteq \mathcal{U}^{\mathcal{S}'}$ and $R^{\mathcal{S}} \subseteq R^{\mathcal{S}'}$ for any symbol R. This ordering is a complete partial order (cpo) admitting the only structure of empty universe $\perp$ as smallest element. The semantics of operators is continuous with respect to this cpo. It means that a (even infinite) system of equations using PQFD operators admits a unique smallest solution.

Example 2. Let us illustrate infinite VR systems of equations for producing the graph of Figure 1. Let us first introduce the intermediate coloured graphs X_n presented in Figure 2.

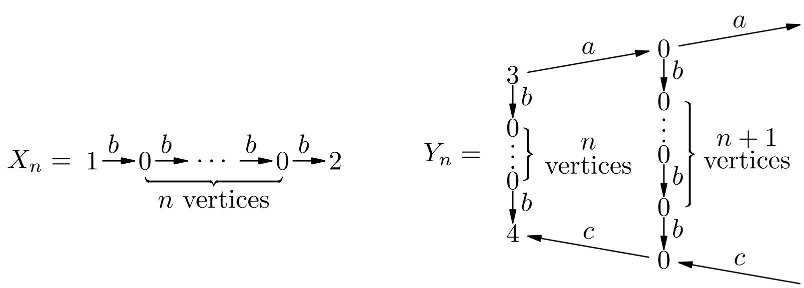

Fig. 2. The graphs X_n and Y_n.

Such graphs can be defined by the following recursive equations:

$$X_0 = \mathbf{add}_{1,2,b}(\mathbf{1} \oplus \mathbf{2}) \quad \text{and} \quad X_{n+1} = \mathbf{ren}_{3,2}(\mathbf{ren}_{2,0}(\mathbf{add}_{2,3,b}(X_n \oplus \mathbf{3}))) \qquad (1)$$

We can now define the Y_n coloured graphs (notice Y_0 is isomorphic to the graph of Figure 1). They satisfy the following equation:

$$Y_n = \mathbf{ren}_{2,4}(\mathbf{ren}_{1,3}(\mathbf{ren}_{4,0}(\mathbf{ren}_{3,0}(\mathbf{add}_{1,3,a}(\mathbf{add}_{4,2,c}(X_n \oplus Y_{n+1})))))) \qquad (2)$$

In fact the coloured graphs X_n and Y_n are the smallest possible graphs satisfying the equations (1) and (2): the step-ladder graph is completely described as the smallest solution of this equational system. Let us notice that, though infinite, this equational system can be represented by an infinite graph as depicted in Figure 3.

This process of encoding the equational system into a rooted graph is general. Formally, a rooted graph $\mathcal{E}$ is a PQFD-equational system if its edges:

– are labelled by $\{\oplus_1, \ \oplus_2\} \cup PQFD_0 \cup PQFD_1$ and

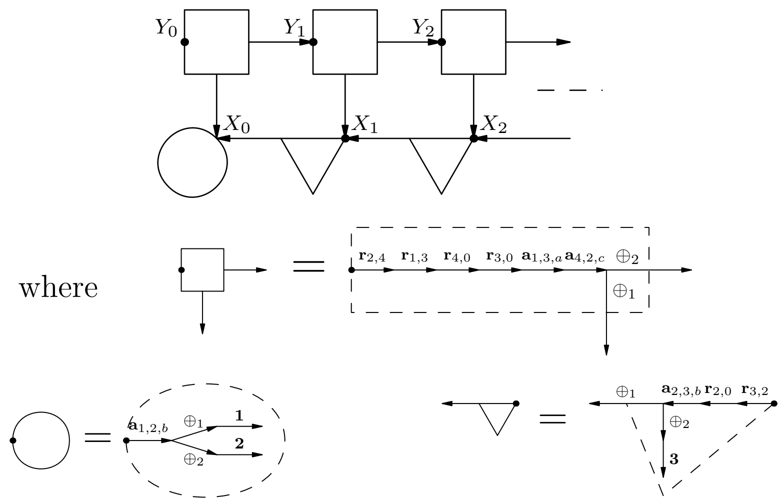

Fig. 3. An infinite VR equational system describing the graph of Figure 1

- for all element x of $\mathcal{U}^{\mathcal{S}}$,
  - if there is an edge labelled by **one** of target x then no edges originate from x.
  - else, either two edges originate from x, and are labelled by respectively $\oplus_1$ and $\oplus_2$,
    or only one edge has origin x, and this edge is labelled by **one** or **pqfd**$[(\phi_R)]$ for some ϕ_R.

The solution of such a system is defined as follows: let $\sigma^{\mathcal{E}}$ be the smallest function from vertices of $\mathcal{E}$ to structures satisfying:

- If $\mathbf{one}^{\mathcal{E}}(x, y)$ then $\sigma^{\mathcal{E}}(x) = \underline{\mathbf{one}}$,
- if $\mathbf{pqfd}[(\phi_R)]^{\mathcal{E}}(x, y)$ then $\sigma^{\mathcal{E}}(x) = \underline{\mathbf{pqfd}[(\phi_R)]}(\sigma^{\mathcal{E}}(y))$,
- and if ${\oplus_1}^{\mathcal{E}}(x, y)$ and ${\oplus_2}^{\mathcal{E}}(x, z)$ then $\sigma^{\mathcal{E}}(x) = \sigma^{\mathcal{E}}(y) \underline{\oplus} \sigma^{\mathcal{E}}(z)$.

then the semantic of the equational system $\mathcal{E}$, written $[\![\mathcal{E}]\!]$ is the graph $\sigma^{\mathcal{E}}(root)$.

We will also make use of another operator: for $p \in \Sigma_1$ a unary symbol, the operator $\mathbf{fuse}_p$ applied to a structure $\mathcal{S}$ keeps the structure unchanged but collapses all the elements x satisfying $p^{\mathcal{S}}(x)$ into a single one. Formally, we define the equivalence relation $\equiv_p$ over $\mathcal{U}^{\mathcal{S}}$ by $x \equiv_p y$ iff $x = y$ or $p^{\mathcal{S}}(x)$ and $p^{\mathcal{S}}(y)$. The classes of equivalence for $\equiv_p$ of an element x is written $[x]_p$. The semantic of $\mathbf{fuse}_p$ is then defined by $\underline{\mathbf{fuse}_p}(\mathcal{S}) = \mathcal{S}'$ with $\mathcal{U}^{\mathcal{S}'} = \{[x]_p \mid x \in \mathcal{U}^{\mathcal{S}}\}$ and for any n-ary symbol R, $R^{\mathcal{S}'} = \{[v_1]_p, \ldots, [v_n]_p \mid R^{\mathcal{S}}(v_1, \ldots, v_n)\}$. The set of operators $PQFD$ increased with **fuse** operators is written $PQFD + F$.

In fact, the cpo used has to be slightly changed for the **fuse** operator to be continuous. Furthermore, the **fuse** operators make it necessary to put some extra restrictions to the system: a $PQFD + F$ equational system is said normalised if there is no predicate $R(y_1, \ldots, y_{|R|})$ such that $y_i = y_j$ for $i \neq j$ in any formula appearing in a **pqfd**[] operator.

4 The transformational approach

Successively applying a finite number of transformations to a relational structure is a second technique for obtaining new relational structures. In this work, we are basically using two such transformations: MS-definable transduction and unfolding.

MS-definable transduction has already been presented. We define here a version of unfolding suitable for deterministic rooted graphs only. Given a deterministic rooted graph G labelled by E with a vertices set V, ρ_G is the function from E^* to V such that $\rho_G(u) = x$ with $root \overset{u}{\Rightarrow} x$ (since the graph is deterministic, there is at most one such x). The unfolding of G is the deterministic rooted graph $Unf(G)$ with a set of vertices $V' = \{u \mid \exists x \in V,\ root \overset{u}{\Rightarrow} x\}$ and such that for all edge symbol a, $a^{Unf(G)}(u, v)$ iff $a^G(\rho_G(u), \rho_G(v))$. The function ρ_G is a morphism of graph and is called the reduction (following the terminology of bisimulation).

We are interested here in transforming a deterministic graph by applying successively an unfolding and a MS-definable transduction.

Example 3. Let G be the graph presented in Figure 4.a with its *root* marked by an unlabelled edge and let $\mathcal{I}$ be the MS interpretation $(\delta, \{\Phi_a, \Phi_b, \Phi_c\})$ with $\delta(x) = \textbf{true}$, $\Phi_a(x_1, x_2) = a(x_1, x_2)$, $\Phi_b(x_1, x_2) = b(x_1, x_2)$ and $\Phi_c(x_1, x_2) = (\exists x_1', x_2'.\, x_1' \overset{b^*}{\Longrightarrow} x_1 \wedge x_2' \overset{b^*}{\Longrightarrow} x_2 \wedge a(x_2', x_1')) \wedge \neg(\exists z.\, a(x_1, z) \vee a(x_2, z))$ where $x \overset{b^*}{\Longrightarrow} y$ is the MS formula stating that there is a path between x and y using only edges labelled by b. $\mathcal{I}(Unf(G))$ is the step-ladder of Figure 1 (Figure 4.b presents the unfolding of G).

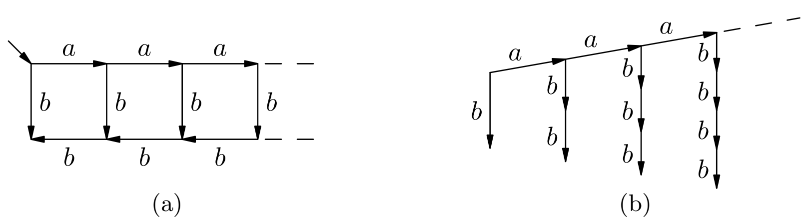

Fig. 4. The graph G (a) and its unfolding (b).

Those two transformations are sufficient for expressing $PQFD + F$ equational systems:

Lemma 1. *Given a $PQFD + F$ equational system $\mathcal{E}$, there exists an MS interpretation $\mathcal{I}$ such that $\mathcal{I}(Unf(\mathcal{E}))$ is isomorphic to $[\![\mathcal{E}]\!]$.*

Proof (sketch). The first remark used in the proof is that $[\![\mathcal{E}]\!] = [\![\mathit{Unf}(\mathcal{E})]\!]$.

For simplicity, let us suppose first that no **fuse** operators are used. Under this hypothesis, the element of $\mathcal{U}^{[\![\mathcal{E}]\!]}$ can be uniquely identified with the **one** operators appearing in $\mathit{Unf}(\mathcal{E})$: the only **one** operator which, if replaced by $\perp$ in $\mathit{Unf}(\mathcal{E})$ makes the vertex disappear from $[\![\mathit{Unf}(\mathcal{E})]\!]$. Let us call ρ this one-to-one mapping which to a vertex of $[\![\mathcal{E}]\!]$ associates the node of $[\![\mathit{Unf}(\mathcal{E})]\!]$ from which the associated **one** edge originates. Given this mapping one can provide MS formulas Φ_R for all symbol R of arity n in the signature such that $\mathit{Unf}(\mathcal{E}) \models \Phi_R(\rho(x_1), \dots, \rho(x_n))$ iff $R^{[\![\mathcal{E}]\!]}(x_1, \dots, x_n)$ holds. Let $\delta(x)$ be $(\exists y,\ \mathbf{one}(x, y))$, then the interpretation $\mathcal{I} = (\delta, (\Phi_R))$ is such that $\mathcal{I}(\mathit{Unf}(\mathcal{E}))$ is isomorphic to $[\![\mathcal{E}]\!]$ (and ρ is the isomorphism).

If **fuse** operators are used, then a similar ρ isomorphism can be provided: the difference is that it maps elements of $\mathcal{U}^{[\![\mathcal{E}]\!]}$ to either **one** operators or **fuse** operators. The intention is that an element of $\mathcal{U}^{[\![\mathcal{E}]\!]}$ is uniquely represented by a **one** operator iff no **fuse** operator 'touched' it in the equational system, in the other case, the element is uniquely represented by the closest to the root **fuse** operator in which it was involved. Apart from this distinction, the same technique is applied for providing the interpretation $\mathcal{I}$.

5 Prefix-recognisable structures

In this section, we focus on the internal representation of structures. Prefix-recognisable graphs have been introduced by Caucal [Cau96]. A possible description of these graphs is by systems of word rewriting. Blumensath [Blu01] extended this definition to relations of arbitrary arity. Those structures, when restricted to binary relations coincide with prefix-recognisable graphs. We give here a similar (and equivalent) definition of prefix-recognisable structures.

For simplicity, we fix a common *infinite* alphabet A. Let R, R' be two relations over A^* of respective arities k and l, we designate by $R \times R'$ the $(k+l)$-ary relation defined by $(R \times R')(u_1, \dots, u_k, v_1, \dots, v_l)$ if and only if $R(u_1, \dots, u_k)$ and $R'(v_1, \dots, v_l)$. Let R be a k-ary relation over A^* and U a language of A^*, we designate by $U \cdot R$ the k-ary relation defined by $(U \cdot R)(uv_1, \dots, uv_k)$ iff $u \in U$ and $R(v_1, \dots, v_k)$. Let R be a k-ary relation and π a permutation of $[1, k]$, $R_\pi(x_1, \dots, x_k)$ iff $R(x_{\pi(1)}, \dots, x_{\pi(k)})$.

Definition 1. *The set of prefix-recognisable (PR) relations over A^* is the smallest set of relations satisfying:*

- *for U a rational subset of A^*, the unary relation U is in PR,*
- *if $R, R' \in PR$ then $R \times R' \in PR$,*
- *for $R, R' \in PR$ of same arity, $R \cup R' \in PR$,*
- *for $R \in PR$ and U a rational subset of A^*, $U \cdot R \in PR$,*
- *for $R \in PR$ and π a permutation of $[1, |R|]$, $R_\pi \in PR$.*

Remark that the definition of each rational language only involves a finite number of letters in A. Thus each relation in PR refers to a *finite* number of letters.

A *PR-structure* is a relational structure of universe A^* with all interpretations in *PR*. Notice that the empty relation is in *PR* as $\emptyset$ is rational.

Prefix-recognisable graphs [Cau96] can be defined as graphs with edges defined by a finite union of relations of the form $U(V \times W)$ (with U, V and W rational languages) and vertices defined by a rational language L. This naturally corresponds to the class of binary *PR*-systems restricted by a finite automaton. We extend this notion of restriction to infinite deterministic automaton.

In this article, we will use the term *automaton* to designate a rooted deterministic graph labelled by a finite subset of A. Moreover, we will assume that this graph comes with a set vertices $Final$. As for finite automaton, we associate to every automaton $\mathcal{A}$ a language $\mathcal{L}_\mathcal{A} \subseteq A^*$ consisting of all words corresponding to the labelling of a path from *root* to an element in $Final$.

A *PR-system* $\mathcal{R}$ is a pair $(\mathcal{S}, \mathcal{A})$ where $\mathcal{S}$ is a *PR*-structure and $\mathcal{A}$ is an automaton. In the following, $\mathcal{R}$ will also designate the structure obtained by restricting $\mathcal{S}$ to $\mathcal{L}_\mathcal{A}$.

Example 4. Our example graph of Figure 1 can be described by a *PR*-system $\mathcal{R} = (\mathcal{S}, \mathcal{A})$. The *PR*-structure $\mathcal{S}$ has three non-empty binary relations a,b and c such that $a^\mathcal{S} = x^*y^*\cdot(\{\varepsilon\}\times(y+z))$, $b^\mathcal{S} = x^*\cdot(\{\varepsilon\} \times x)$ and $c^\mathcal{S} = x^*\cdot(xy^*z \times y^*z)$. The automaton $\mathcal{A}$ is presented in Figure 5.a. Its root is pointed by an unlabelled edge and all its states are in $Final$. The language recognised by $\mathcal{A}$ is the set of prefixes of $\{x^ny^nz \mid n \geq 0\}$. The graph obtained by restricting $\mathcal{S}$ to the language recognised by $\mathcal{A}$ (Figure 5.b) is isomorphic to the step-ladder (Figure 1).

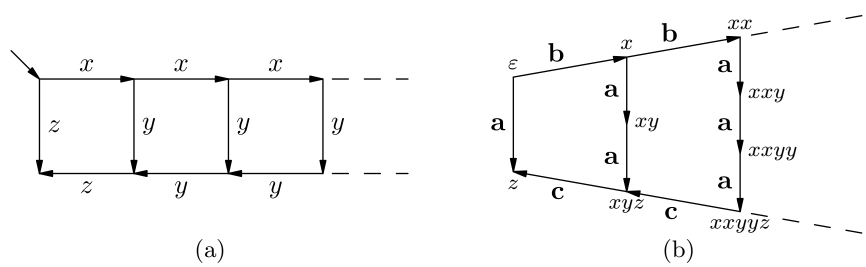

Fig. 5. The automaton $\mathcal{A}$ (a) and the *PR*-system $\mathcal{R}$ (b).

Lemma 2. *For any MS-definable transduction* $\mathcal{T}r$ *and any deterministic graph* G*, there exists a PR-system* $\mathcal{R} = (\mathcal{S}, \mathcal{A})$ *such that* $\mathcal{T}r(Unf(G))$ *is isomorphic to* $\mathcal{R}$ *and* $\mathcal{A}$ *is obtained from* G *by an MS-definable transduction.*

Proof (sketch). Let us first consider the case when $\mathcal{T}r$ is a *non-erasing* MS interpretation $(\mathbf{true}, (\Phi_R)_{R\in\Sigma})$.

For every formula Φ_R, there exists an associated parity automaton $\mathcal{A}_R$. This automaton works on deterministic trees with n distinguished vertices $m_1, \ldots, m_n$ and accepts a tree T iff $T \models \Phi_R(m_1, \ldots, m_n)$. We can always suppose that the states Q of $\mathcal{A}_R$ are enriched with informations about the marks to be seen. More precisely, there exists a mapping ϕ from Q to $2^{[1,n]}$ such that if a node x of T is assigned a state q in a successful run of $\mathcal{A}_R$ then the marks in $T_{/x}$ are exactly the set $\phi(q)$ ($T_{/x}$ designates the subtree of T rooted at x).

We want to attach to every node of $\mathit{Unf}(G)$ the set of transitions of $\mathcal{A}_R$ starting a successful run on $\mathit{Unf}(G)_{/x}$. We denote $\mathcal{M}_R$ the application performing this marking.

- It is well know that this marking can be done before the unfolding (see [Wal96]). There exists a marking $\mathcal{B}_R$ such that $\mathcal{M}_R(\mathit{Unf}(G)) = \mathit{Unf}(\mathcal{B}_R(G))$.
- We prove that B_R can be performed by a MS interpretation on G.

Finally, we defined a n-ary PR-relation R which uses the marking of $\mathcal{M}_R$ to ensure that $R(m_1, \ldots, m_n)$ if and only if there exists an accepting run of $\mathcal{P}_R$ on $\mathit{Unf}(G)$.

The general case when $\mathcal{T}r$ involves both $\mathcal{K}$-copying and restriction is easily deduced from the previous restricted case.

Lemma 3. *For any PR-system $\mathcal{R} = (\mathcal{S}, \mathcal{A})$, there exists a PQFD-system $\mathcal{E}$ such that $\mathcal{E}$ is obtained by an MS definable transduction from $\mathcal{A}$ and $\mathcal{R}$ is isomorphic to $[\![\mathcal{E}]\!]$.*

Proof (sketch). The proof is syntactical. Let $(P_R)_{R\in\Sigma}$ be PR-relations and let $A_1, \ldots, A_k$ be the finite automata accepting the rational languages involved in the PR-expressions describing P's relations and let $\mathcal{A}$ be the automaton restricting the PR-system.

We produce a new equational system working over signature Σ enriched by a new symbol for each state of an automaton A_i. The arity of the symbol is the arity of the relation in which $\mathcal{L}_{A_i}$ is used. The equational system is obtained from $\mathcal{A}$ by replacing each edge labelled by a with a **pqfd**[] operator which simulates simultaneously all a transitions of the A_i's. $\oplus$ operators are used to follow the branching structure of $\mathcal{A}$. **one** operators are used for each Final state of $\mathcal{A}$. **pqfd**[] operations mimic transitions of the $\mathcal{A}$ and the PR-relations.

6 Conclusion

By combining Lemmas 1,2 and 3, we obtain the following theorem:

Theorem 1. *Let $\mathcal{F}$ be a family of deterministic graphs closed by MS-definable transduction, the following classes of structures are isomorphic:*

- *the solutions of a system of equations over the $PQFD + F$ operators represented by a graph in $\mathcal{F}$,*
- *the structures obtained by applying an MS-definable transduction to the unfolding of a graph in $\mathcal{F}$,*
- *the prefix-recognisable structures restricted to the language accepted by a deterministic automaton in $\mathcal{F}$*

An important point of this approach is that according to the second representation, if $\mathcal{F}$ has a decidable MS theory, then it is also the case of the resulting structure.

In fact, we also proved a fourth equivalence with normalised $PQFD + F$ equational systems. We do not know whether unrestricted $PQFD + F$ equational systems have the same solutions.

Acknowledgements. We would like to thank Didier Caucal for his advices.

References

[Bar97] K. Barthelmann. On equational simple graphs. Technical Report 9, Universität Mainz, Institut für Informatik, 1997.

[Blu99] A. Blumensath. Automatic structures. Diploma thesis, RWTH-Aachen, 1999.

[Blu01] A. Blumensath. Prefix-recognisable graphs and monadic second-order logic. Technical Report AIB-06-2001, RWTH Aachen, May 2001.

[Cau92] D. Caucal. On the regular structure of prefix rewriting. *TCS*, 106:61–86, 1992.

[Cau96] D. Caucal. On infinite transition graphs having a decidable monadic theory. In *ICALP'96*, volume 1099 of *LNCS*, pages 194–205, 1996.

[CM02] B. Courcelle and J.A Makowsky. Fusion in relational structures and the verification of mso logic. In *MSCS*, volume 12, pages 203–235, 2002.

[Col02] T. Colcombet. On families of graphs having a decidable first order theory with reachability. In *ICALP'02*, 2002.

[Cou89] B. Courcelle. The monadic second-order logic of graphs ii: infinite graphs of bounded tree width. *Math. Systems Theory*, 21:187–221, 1989.

[Cou90] B. Courcelle. *Handbook of Theoretical Computer Science*, chapter Graph rewriting: an algebraic and logic approach. Elsevier, 1990.

[Cou94] B. Courcelle. Monadic-second order definable graph transductions : a survey. *TCS*, vol. 126:pp. 53–75, 1994.

[CW98] B. Courcelle and I. Walukiewicz. Monadic second-order logic, graph coverings and unfoldings of transition systems. In *Annals of Pure and Applied Logic*, 1998.

[DT90] M. Dauchet and S. Tison. The theory of ground rewrite systems is decidable. In *Fifth Annual IEEE Symposium on Logic in Computer Science*, pages 242–248, 1990.

[Löd02] C. Löding. Ground tree rewriting graphs of bounded tree width. In *Stacs O2*. LNCS, 2002.

[Mor00] C. Morvan. On rational graphs. In J. Tiuryn, editor, *FOSSACS'00*, volume 1784 of *LNCS*, pages 252–266, 2000.

[MS85] D. Muller and P. Schupp. The theory of ends, pushdown automata, and second-order logic. *Theoretical Computer Science*, 37:51–75, 1985.

[Sén92] G. Sénizergues. Definability in weak monadic second-order logic of some infinite graphs. In *Dagstuhl seminar on Automata theory: Infinite computations, Warden, Germany*, volume 28, page 16, 1992.

[Tho97] W. Thomas. Languages, automata, and logic. *Handbook of Formal Language Theory*, 3:389–455, 1997.

[Urv02] T. Urvoy. On abstract families of graphs. In *DLT 02*, 2002.

[Wal96] I. Walukiewicz. Monadic second order logic on tree-like structures. In *STACS'96*, volume 1046 of *LNCS*, pages 401–414, 1996.