Theoretical Computer Science 290 (2003) 79–115
Theoretical Computer Science
www.elsevier.com/locate/tcs

On infinite transition graphs having a decidable monadic theory ☆

Didier Caucal
IRISA, Campus de Beaulieu, 35042 Rennes, France

Received March 1996; revised July 2000; accepted December 2000
Communicated by W. Thomas

Abstract

We define a family of graphs whose monadic theory is (in linear space) reducible to the monadic theory S2S of the complete ordered binary tree. This family contains strictly the context-free graphs investigated by Muller and Schupp, and also the equational graphs defined by Courcelle. Using words as representations of vertices, we give a complete set of representatives by prefix rewriting of rational languages. This subset of possible representatives is a boolean algebra preserved by transitive closure of arcs and by rational restriction on vertices.

1. Introduction

We consider the satisfaction of properties in structures. The properties are given by monadic second-order sentences, and the structures are labelled directed graphs (see for instance [16] and [18]). Rabin has shown that the complete ordered binary tree Λ has a decidable monadic theory [11]: we can decide whether a given property expressed by a monadic sentence is satisfied by the tree Λ. Later Muller and Schupp have extended this decidability result to the context-free graphs [9] (a context-free graph is a rooted graph of finite degree which has a finite number of non-isomorphic connected components by 'decomposition by distance' from a (any) vertex). These context-free graphs are also the transition graphs of pushdown automata [9]. Finally Courcelle has shown that the monadic theory remains decidable for the equational graphs [6]: an equational graph is a graph generated by a deterministic graph grammar. For rooted

☆ This paper has been partly supported by ESPRIT BRA 6317 (ASMICS) and has been partially presented to ICALP 96.
E-mail address: caucal@irisa.fr (D. Caucal).

PII: S0304-3975(01)00089-5

graphs of finite degree, these equational graphs are the context-free graphs [3]. These decidability results of [9] and [6] are extensions of the definability method used by Rabin.

Another approach is to find transformations on graphs which preserve the decidability of the monadic theory, and to apply these transformations to graphs having a decidable monadic theory (see for instance [17]). A first transformation has been given by Shelah [14] and proved by Stupp [15]: if a graph has a decidable monadic theory then its "tree-graph" (obtained by a version of unravelling) has a decidable monadic theory.

A way to find a transformation f on graphs that preserves the decidability of the monadic theory, is to translate f into an "equivalent" transformation f^* on monadic formulas: for any graph G, $f(G)$ satisfies a sentence φ if and only if G satisfies $f^*(\varphi)$. This method has been applied for instance in [10, 7, 8], and especially in [12] for an extension of the tree-graph transformation. We give here two transformations on graphs which have direct equivalent transformations on monadic formulas: they are based on the fact that the existence of a path labelled in a rational language can be expressed by an equivalent monadic formula. By closure of the binary tree Λ under these two operations, we get a family F of graphs which have a decidable monadic theory as a corollary of Rabin's theorem. We show that this family F is a strict extension of the equational graphs and hence of the context-free graphs as well. By taking words as vertices, we extract a complete subset F_0 of representatives up to isomorphism, such that F_0 remains closed under the two operations defining F, and is a boolean algebra.

2. A family of graphs with a decidable monadic theory

We define two transformations on graphs which can be translated on formulas in such a way that the decidability of the monadic theory is simply preserved. The first transformation is the rational restriction on (arc) labels and the second transformation is the inverse rational substitution on (arc) labels. We start with the complete ordered binary tree which has a decidable monadic theory [11]. By applying to this tree the second transformation followed by the first one, we obtain a family of graphs with a decidable monadic theory, and which is closed under these two transformations.

We take an alphabet (finite set of symbols) T of *terminals* containing at least two symbols a, b. Here a graph is a set of arcs labelled by symbols of T. This means that a *graph* G is a subset of $V \times T \times V$ where V is an arbitrary set. Any triple (s, a, t) of G is a *labelled arc* of *source* s, of *target* t, with *label* a, and is identified with the labelled transition $s \underset{G}{\overset{a}{\longrightarrow}} t$ or directly $s \overset{a}{\longrightarrow} t$ if G is understood. We denote by $V_G := \{s \mid \exists a\, \exists t,\ s \overset{a}{\longrightarrow} t \vee t \overset{a}{\longrightarrow} s\}$ the set of *vertices* of G. A graph is *deterministic* if distinct arcs with the same source have distinct labels: if $r \overset{a}{\longrightarrow} s$ and $r \overset{a}{\longrightarrow} t$ then $s = t$.

And a graph is (source) *complete* if for every label a, every vertex is source of an arc labelled by a.

The existence of a *path* in G from vertex s to vertex t and labelled by a word $w\in T^*$ is denoted by $s \underset{G}{\overset{w}{\Rightarrow}} t$ or directly by $s \overset{w}{\Rightarrow} t$ if G is understood: we have $s \overset{\varepsilon}{\Rightarrow} s$, and $s \overset{aw}{\Rightarrow} t$ if there is some vertex r such that $s \overset{a}{\rightarrow} r$ and $r \overset{w}{\Rightarrow} t$. The label set $L(G,E,F)$ of paths from a set E to a set F is the following language over T:

$$L(G,E,F) := \{w \in T^* \mid \exists s \in E, \exists t \in F,\ s \underset{G}{\overset{w}{\Rightarrow}} t\}.$$

We say that a vertex r is a *root* of G if every vertex is accessible from r: $\forall s\in V_G\,\exists\, w\in T^*\ r \overset{w}{\Rightarrow} s$. And a graph is a *tree* if it has a root r which is target of no arc, and every vertex $s\neq r$ is target of a unique arc.

Recall that the family $Rat(T^*) := \{L(G,E,F) \mid \#G<\infty \wedge E\cup F\subseteq V_G\}$ of path label sets of finite graphs is the family of *rational languages* over T.

Given a binary relation R whose domain is included in V, we consider the graph

$$R(G) := \{s' \overset{a}{\rightarrow} t' \mid \exists s \underset{G}{\overset{a}{\rightarrow}} t, s\,R\,s' \wedge t\,R\,t'\}$$

of the *application* of R to G. We say that $R(G)$ is *isomorphic* to G when R is a bijection from V_G to $V_{R(G)}$. Two deterministic and complete trees on the same label alphabet are isomorphic.

To construct monadic second-order formulas, we take two disjoint denumerable sets: a set of *vertex variables* and a set of *vertex set variables*. *Atomic formulas* have one of the following two forms:

$$x \in X \quad \text{or} \quad x \overset{a}{\rightarrow} y,$$

where X is a vertex set variable, x and y are vertex variables, and $a\in T$.

From the atomic formulas, we construct as usual the *monadic second-order formulas* with the propositional connectives $\neg, \wedge$ and the existential quantifier $\exists$ acting on these two kind of variables. A *sentence* is a formula without free variable. The set $MTh(G)$ of monadic second-order sentences satisfied by a graph G forms the *monadic theory* of G.

Note that two isomorphic graphs satisfy the same sentences: $MTh(R(G)) = MTh(G)$ when R is bijective. Instead of renaming vertices, we consider the *restriction* $G_{|L}$ of G to an arbitrary set $L\subseteq V_G$ as follows:

$$G_{|L} := G\cap (L\times T\times L) = \{(s,s) \mid s\in L\}(G) = \{s \underset{G}{\overset{a}{\rightarrow}} t \mid s,t\in L\}.$$

Analogously we consider the restriction φ_L (resp. $\varphi_{|L}$) of any sentence φ to a set L by imposing that vertex variables (resp. vertex set variables) are interpreted only by vertices in L (resp. by subsets of L). Using a constant $\underline{L}$ for the set of vertices in a given language L, we define by induction over the formulas:

$$
\begin{array}{llll}
(x \in X)_L & = x \in X, & (x \in X)_{|L} & = x \in X, \\
(x \xrightarrow{a} y)_L & = x \xrightarrow{a} y, & (x \xrightarrow{a} y)_{|L} & = x \xrightarrow{a} y, \\
(\neg\varphi)_L & = \neg(\varphi_L), & (\neg\varphi)_{|L} & = \neg(\varphi_{|L}), \\
(\varphi \wedge \psi)_L & = \varphi_L \wedge \psi_L, & (\varphi \wedge \psi)_{|L} & = \varphi_{|L} \wedge \psi_{|L}, \\
(\exists x\ \varphi)_L & = \exists x\ (x \in \underline{L} \wedge \varphi_L), & (\exists x\ \varphi)_{|L} & = \exists x\ (x \in \underline{L} \wedge \varphi_{|L}), \\
(\exists X\ \varphi)_L & = \exists X\ \varphi_L, & (\exists X \varphi)_{|L} & = \exists X\ (X \subseteq \underline{L} \wedge \varphi_{|L}).
\end{array}
$$

These restrictions of graphs and sentences are dual.

Lemma 2.1. *Given a graph G, a set L and a monadic sentence φ, we have*

$$G_{|L} \models \varphi \ \Leftrightarrow\ G \models \varphi_L \ \Leftrightarrow\ G \models \varphi_{|L}.$$

Proof. Note that the vertices of $G_{|L}$ are the vertices of G in L: $V_{G_{|L}} = V_G \cap L$.

For the inductive proof, we refer to formulas $\varphi(X_1,\ldots,X_m,x_1,\ldots,x_n)$ with $m \geqslant 0$ free vertex set variables $X_1,\ldots,X_m$ and $n \geqslant 0$ free variables $x_1,\ldots,x_n$.

(i) By induction on the structure of φ, we have

$$G_{|L} \models \varphi(A_1,\ldots,A_m,a_1,\ldots,a_n) \ \Leftrightarrow\ G \models \varphi_{|L}(A_1,\ldots,A_m,a_1,\ldots,a_n)$$

for any $A_1,\ldots,A_m \subseteq V_G \cap L$ and for any $a_1,\ldots,a_n \in V_G \cap L$.

(ii) Furthermore and by induction on the structure of φ, we have

$$G \models \varphi_L(A_1,\ldots,A_m,a_1,\ldots,a_n) \ \Leftrightarrow\ G \models \varphi_{|L}(A_1 \cap L,\ldots,A_m \cap L,a_1,\ldots,a_n)$$

for any $A_1,\ldots,A_m \subseteq V_G$ and for any $a_1,\ldots,a_n \in V_G \cap L$. $\square$

However for general L, the formulas φ_L and $\varphi_{|L}$ are (by the presence of the symbol $\underline{L}$) not monadic in the original signature. But for the case that L is rational, we show how to transform φ_L (or $\varphi_{|L}$) into an equivalent monadic formula: we have to transform $x \in \underline{L}$ into a monadic formula. This transformation can be reduced to the expression by a monadic formula $[s \overset{L}{\Rightarrow} t]^\bigstar$ for the existence of a path $s \overset{L}{\Rightarrow} t$ from s to t and labelled by a word in $L \in Rat(T^*)$. This formula is defined by induction on the rational structure of L:

$$
\begin{array}{l}
[x \overset{\emptyset}{\Rightarrow} y]^\bigstar : \exists X\ (x \in X \wedge \neg(x \in X)) \text{ i.e. a false formula,} \\
[x \overset{\{a\}}{\Rightarrow} y]^\bigstar : x \xrightarrow{a} y, \\
[x \overset{L+M}{\Rightarrow} y]^\bigstar : [x \overset{L}{\Rightarrow} y]^\bigstar \vee [x \overset{M}{\Rightarrow} y]^\bigstar, \\
[x \overset{L.M}{\Rightarrow} y]^\bigstar : \exists z\ ([x \overset{L}{\Rightarrow} z]^\bigstar \wedge [z \overset{M}{\Rightarrow} y]^\bigstar), \\
[x \overset{L^*}{\Rightarrow} y]^\bigstar : \forall X\ ((x \in X \wedge \forall p\ \forall q((p \in X \wedge [p \overset{L}{\Rightarrow} q]^\bigstar) \Rightarrow q \in X)) \Rightarrow y \in X),
\end{array}
$$

where the transformation for $\overset{L^*}{\Rightarrow}$ is the reflexive and transitive closure $(\overset{L}{\Rightarrow})^*$ of $\overset{L}{\Rightarrow}$: $x \overset{L^*}{\Rightarrow} y$ if and only if every vertex set X containing x and closed by $\overset{L}{\Rightarrow}$ contains y.

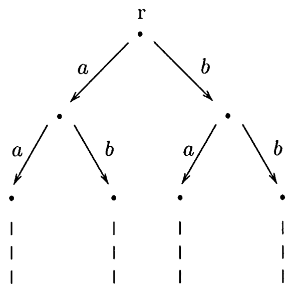
Note that the length $|[x \overset{L}{\Rightarrow} y]^\star|$ of the monadic formula $[x \overset{L}{\Rightarrow} y]^\star$ is linear in the length $|L|$ of the rational expression L.

Now we consider the restriction $G_{\|r,L}$ of a graph G to the vertices accessible from a vertex r by a path labelled in $L \subseteq T^*$:

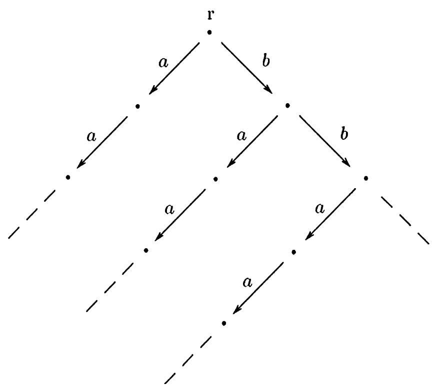
$$G_{\|r,L} := G_{|\,\{s\,|\,r \overset{L}{\Rightarrow} s\}} = \{s \underset{G}{\overset{a}{\rightarrow}} t \mid r \overset{L}{\Rightarrow} s \wedge r \overset{L}{\Rightarrow} t\},$$

in particular r is a root of $G_{\|r,L}$.

For instance taking a deterministic and complete tree Λ on $\{a,b\}$:



with root r i.e. its vertex satisfying the formula $\neg \exists y\ (y \overset{a}{\rightarrow} x \vee y \overset{b}{\rightarrow} x)$, then its restriction $\Lambda_{\|r,b^*a^*}$ is the following graph:



By Lemma 2.1, the restriction of a graph preserves the decidability of the monadic theory if we restrict to the vertices accessible from a fixed (and definable) vertex by a path labelled in a given rational language.

Proposition 2.2. *Given a graph G, a rational language L over T, and a monadic formula $\varphi(x)$ satisfied by a unique vertex r, we have*

$$MTh(G) \text{ decidable } \Rightarrow MTh(G_{\|r,L}) \text{ decidable.}$$

Proof. Let $M := \{s \mid r \overset{L}{\Rightarrow} s\}$ be the set of vertices accessible from r by a path labelled in L. For any sentence ψ, we have

$$\begin{aligned} G_{\|r,L} \models \psi &\Leftrightarrow G_{|M} \models \psi && \text{by definition of } M \\ &\Leftrightarrow G \models \psi_M && \text{by Lemma 2.1} \\ &\Leftrightarrow G \models \exists z\ (\varphi(z) \wedge \psi_{M,z}), \end{aligned}$$

where $\psi_{M,z}$ is the formula obtained from ψ_M by substituting to any $x \in M$ the formula $[z \overset{L}{\Rightarrow} x]^{\star}$. Formally $\psi_{M,z} = \psi^{L,z}$ where $\psi^{L,z}$ is defined by induction on the structure of ψ as follows:

$$\begin{aligned} (x \in X)^{L,z} &= x \in X, & (x \overset{a}{\rightarrow} y)^{L,z} &= x \overset{a}{\rightarrow} y, \\ (\neg\varphi)^{L,z} &= \neg(\varphi^{L,z}), & (\varphi \wedge \psi)^{L,z} &= \varphi^{L,z} \wedge \psi^{L,z}, \\ (\exists x\ \varphi)^{L,z} &= \exists x\ ([z \overset{L}{\Rightarrow} x]^{\star} \wedge \varphi^{L,z}), & (\exists X\ \varphi)^{L,z} &= \exists X\ \varphi^{L,z}. \end{aligned} \qquad \square$$

In particular, the previous graph has a decidable monadic theory because Λ has a decidable monadic theory [11]. We define now a second operation on labels which preserves the decidability of the monadic theory.

To move by inverse arcs, we introduce a new alphabet $\bar{T} := \{\bar{a} \mid a \in T\}$ in bijection with T. Any transition $u \overset{\bar{a}}{\rightarrow} v$ means that $v \overset{a}{\rightarrow} u$ is an arc of G. We extend by composition the existence of a path $\overset{w}{\Rightarrow}$ labelled by a word w in $(T \cup \bar{T})^*$, and we denote by

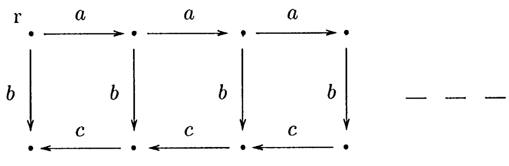
$$\bar{L}(G,E,F) := \{w \in (T \cup \bar{T})^* \mid \exists s \in E,\ \exists t \in F,\ s \underset{G}{\overset{w}{\Rightarrow}} t\},$$

the set of path labels over $T \cup \bar{T}$ from a set E to a set F.

An *extended substitution* h on T^* is a morphism from T^* into the family $2^{(T \cup \bar{T})^*}$ of languages over $T \cup \bar{T}$, *i.e.* satisfying $h(\varepsilon) = \{\varepsilon\}$ and $h(uv) = h(u)h(v)$ for every $u, v \in T^*$. The *inverse substitution* $h^{-1}(G)$ of a graph G according to h is the following graph:

$$h^{-1}(G) := \{s \overset{a}{\rightarrow} t \mid \exists w \in h(a),\ s \underset{G}{\overset{w}{\Rightarrow}} t\}.$$

For instance for $h(a) = \{b\}$, $h(b) = \{b\bar{b}a\}$, $h(c) = \{\bar{a}\bar{b}a\}$ and $h(d) = \emptyset$ for every other d in T, the inverse substitution $h^{-1}(\Lambda_{\|r,b^*a^*})$ of the previous graph is the following graph:



We denote by $\tilde{u}$ the *mirror* of any word u: $\tilde{\varepsilon} = \varepsilon$ and $\widetilde{au} = \tilde{u}a$ for any letter a.

We extend by morphism barred letters to barred words with $\bar{\bar{a}} = a$ for every $a \in T$.

This permits to extend any substitution h into a substitution $\bar{h}$ from $(T \cup \bar{T})^*$ into the family $2^{(T \cup \bar{T})^*}$ by defining for every $a \in T$,

$$\bar{h}(a) := h(a), \quad \bar{h}(\bar{a}) := \widetilde{h(a)}.$$

The set of path (resp. chain) labels of an inverse substitution of a graph is equal to the inverse substitution of the extended path (resp. chain) labels of the graph.

Lemma 2.3. *Given a graph G, a substitution h, and sets E,F, we have*

$$L(h^{-1}(G),E,F) = h^{-1}(\bar{L}(G,E,F)) \quad \textit{and} \quad \bar{L}(h^{-1}(G),E,F) = \bar{h}^{-1}(\bar{L}(G,E,F)).$$

Proof. By induction on the length of $w \in (T \cup \bar{T})^*$, we have

$$s \underset{h^{-1}(G)}{\overset{w}{\Rightarrow}} t \iff s \underset{G}{\overset{\bar{h}(w)}{\Rightarrow}} t \text{ for any } s,t \in V_{h^{-1}(G)}. \qquad \square$$

Similarly to the restriction, we define the substitution φ_h by h of any formula φ. By induction on the structure of any formula, we replace each atomic formula $x \overset{a}{\rightarrow} y$ by the existence of a path $x \overset{h(a)}{\Rightarrow} y$ labelled in $h(a)$:

$$\begin{aligned} (x \in X)_h &= x \in X, & (x \overset{a}{\rightarrow} y)_h &= x \overset{h(a)}{\Rightarrow} y, \\ (\neg\ \varphi)_h &= \neg(\varphi_h), & (\varphi \wedge \psi)_h &= \varphi_h \wedge \psi_h, \\ (\exists x\ \varphi)_h &= \exists x\ \varphi_h, & (\exists X\ \varphi)_h &= \exists X\ \varphi_h. \end{aligned}$$

Note that $(\varphi_L)_h = (\varphi_h)_L$ which is denoted by $\varphi_{L,h}$ and $(\varphi_{|L})_h = (\varphi_h)_{|L}$ which is denoted by $\varphi_{|L,h}$. A graph $h^{-1}(G)$ of vertex set L satisfies a sentence φ if and only if G satisfies $\varphi_{|L,h}$.

Lemma 2.4. *Given a graph G, a substitution h and a monadic sentence φ, we have*

$$h^{-1}(G) \models \varphi \iff G \models \varphi_{L,h} \iff G \models \varphi_{|L,h},$$

where $L = V_{h^{-1}(G)}$ is the set of vertices of $h^{-1}(G)$.

Proof. The second equivalence is shown with Lemma 2.1.

Let $\varphi(X_1,\ldots,X_m,x_1,\ldots,x_n)$ be any (monadic second-order) formula with $m \geqslant 0$ free vertex set variables $X_1,\ldots,X_m$ and $n \geqslant 0$ free variables $x_1,\ldots,x_n$.

By induction on the structure of φ, we have

$$h^{-1}(G) \models \varphi(A_1,\ldots,A_m,a_1,\ldots,a_n) \iff G \models \varphi_{|L,h}(A_1,\ldots,A_m,a_1,\ldots,a_n)$$

for any $A_1,\ldots,A_m \subseteq L$ and for any $a_1,\ldots,a_n \in L$. $\square$

It follows that the inverse according to a substitution h preserves the decidability of the monadic theory when h is rational *i.e.* when $h(a) \in Rat((T \cup \bar{T})^*)$ for any $a \in T$, in other words $h : T^* \rightarrow Rat((T \cup \bar{T})^*)$ is a morphism.

Proposition 2.5. *Given a graph G and a rational substitution h, we have*

$$MTh(G) \textit{ is decidable} \Rightarrow MTh(h^{-1}(G)) \textit{ is decidable}.$$

Proof. Let $L = V_{h^{-1}(G)}$ be the set of vertices of $h^{-1}(G)$ and let $M = \bigcup_{a\in T} h(a)$ be the image of h. For any sentence φ, we have

$$\begin{aligned} h^{-1}(G) \models \varphi &\Leftrightarrow G \models \varphi_{L,h} \text{ by Lemma 2.4} \\ &\Leftrightarrow G \models \varphi^h, \end{aligned}$$

where φ^h is the monadic formula obtained by induction on the structure of any monadic formula φ as follows:

$$\begin{aligned} (x \in X)^h &= x \in X, \quad (x \xrightarrow{a} y)^h = [x \overset{h(a)}{\Rightarrow} y]^\star, \\ (\neg\ \varphi)^h &= \neg(\varphi^h), \quad (\varphi \wedge \psi)^h = \varphi^h \wedge \psi^h, \\ (\exists X\ \varphi)^h &= \exists X\ \varphi^h,\ (\exists x\varphi)^h = \exists x\ (\exists y([x \overset{M}{\Rightarrow} y]^\star \vee [y \overset{M}{\Rightarrow} x]^\star) \wedge \varphi^h), \end{aligned}$$

where $[x \overset{P}{\Rightarrow} y]^\star$ has been yet defined by induction on the rational structure of P in $Rat(T^*)$; and we allow that $P \in Rat((T \cup \bar{T})^*)$ by adding $[x \overset{\{\bar{a}\}}{\Rightarrow} y]^\star : y \xrightarrow{a} x$. $\square$

Let us compose Propositions 2.2 and 2.5.

Proposition 2.6. *Given a graph G with a unique root r, a rational substitution h, and a rational label language $L \in Rat((T \cup \bar{T})^*)$, we have*

$$MTh(G) \text{ is decidable } \Rightarrow MTh(h^{-1}(G)_{|L_G}) \text{ is decidable},$$

where $L_G := \{s \mid r \underset{G}{\overset{L}{\Rightarrow}} s\}$ is the set of vertices accessible in G from r by a path in L.

Proof. Let $Q = V_{h^{-1}(G)}$ be the set of vertices of $h^{-1}(G)$ and let $M = \bigcup_{a\in T} h(a)$ be the image of h. For any sentence φ, we have

$$\begin{aligned} h^{-1}(G)_{|L_G} \models \varphi &\Leftrightarrow\ h^{-1}(G) \models \varphi_{L_G} && \text{by Lemma 2.1} \\ &\Leftrightarrow\ G \models (\varphi_{L_G})_{Q,h} && \text{by Lemma 2.4} \\ &\Leftrightarrow\ G \models \varphi_{L_G \cap Q,h} \\ &\Leftrightarrow\ G \models \exists z\ (\forall y[z \overset{T^*}{\Rightarrow} y]^\star \wedge \varphi^{L,h,z}), \end{aligned}$$

where $\varphi^{L,h,z}$ is the monadic formula obtained by induction on the structure of any monadic formula φ as follows:

$$\begin{aligned} (x \in X)^{L,h,z} &= x \in X, \qquad (x \xrightarrow{a} y)^{L,h,z} = [x \overset{h(a)}{\Rightarrow} y]^\star, \\ (\neg\varphi)^{L,h,z} &= \neg(\varphi^{L,h,z}),\ (\varphi \wedge \psi)^{L,h,z} = \varphi^{L,h,z} \wedge \psi^{L,h,z}, \\ (\exists X\ \varphi)^{L,h,z} &= \exists X\ \varphi^{L,h,z} \\ (\exists x\ \varphi)^{L,h,z} &= \exists x\ ([z \overset{L}{\Rightarrow} x]^\star \wedge \exists y\ ([x \overset{M}{\Rightarrow} y]^\star \vee [y \overset{M}{\Rightarrow} x]^\star) \wedge \varphi^{L,h,z}). \end{aligned}$$

Note that the length of the monadic sentence $\exists z\ (\forall y[z \overset{T^*}{\Rightarrow} y]^\star \wedge \varphi^{L,h,z})$ is linear in the length of φ and in the lengths of regular expressions defining L, h. $\square$

Note that Proposition 2.6 is a corollary of Proposition 3.1 in [6] (the transformation of G to $h^{-1}(G)_{|L_G}$ is a 'noncopying monadic second-order definable transduction').

Furthermore this transformation is a reduction linear in space for the monadic theory.

Remark that $h^{-1}(G_{|L_G}) \subseteq h^{-1}(G)_{|L_G}$ and $h^{-1}(G)_{\|r,h^{-1}(L)} \subseteq h^{-1}(G)_{|L_G}$.

But the restriction to connected components and the inverse substitution commute.

Lemma 2.7. *Given a graph G and a vertex subset L closed by $\underset{G}{\leftrightarrow}$, we have*

$h^{-1}(G_{|L}) = h^{-1}(G)_{|L}$ *for any substitution h.*

Proof. We have

$$h^{-1}(G_{|L}) = h^{-1}(G_{|L})_{|L} \subseteq h^{-1}(G)_{|L}.$$

Let us prove the inverse inclusion. Let $x \underset{h^{-1}(G)_{|L}}{\overset{a}{\longrightarrow}} y$. So $x \underset{h^{-1}(G)}{\overset{a}{\longrightarrow}} y$ with $x, y \in L$.

By definition of $h^{-1}(G)$, there is $w \in h(a)$ such that $x \underset{G}{\overset{w}{\Rightarrow}} y$. So $x \underset{G}{\overset{*}{\leftrightarrow}} y$.

As $x, y \in L$ and by hypothesis on L, we have $x \underset{G_{|L}}{\overset{w}{\Rightarrow}} y$ and hence $x \underset{h^{-1}(G_{|L})}{\overset{a}{\Rightarrow}} y$. $\square$

Another basic property is the composition of inverse substitutions.

Lemma 2.8. *Given a graph G, and substitutions g and h, the composition $g \circ \bar{h}$ defined by $(g \circ \bar{h})(a) := \bar{h}(g(a))$ for every $a \in T$, is a substitution satisfying*

$$g^{-1}(h^{-1}(G)) = \begin{cases} ((g \circ \bar{h})^{-1}(G))_{|V_{h^{-1}(G)}} \\ (g \circ \bar{h})^{-1}(G) \quad \text{if } \varepsilon \notin g(T) \end{cases}$$

Proof. In the proof of Lemma 2.3, we have seen that

$$s \underset{h^{-1}(G)}{\overset{w}{\Rightarrow}} t \;\Leftrightarrow\; s \underset{G}{\overset{\bar{h}(w)}{\Rightarrow}} t \quad \text{for any } w \in (T \cup \bar{T})^* \text{ and any } s, t \in V_{h^{-1}(G)}$$

This remains true for vertices s, t of G when $w \neq \varepsilon$: by induction on $|w| \geqslant 1$,

$$s \underset{h^{-1}(G)}{\overset{w}{\Rightarrow}} t \;\Leftrightarrow\; s \underset{G}{\overset{\bar{h}(w)}{\Rightarrow}} t \quad \text{for any } w \in (T \cup \bar{T})^+ \text{ and any } s, t \in V_G$$

Assuming that $\varepsilon \notin g(T) \;\vee\; s, t \in V_{h^{-1}(G)}$ we have for every $a \in T$:

$$\begin{aligned} s \underset{g^{-1}(h^{-1}(G))}{\overset{a}{\longrightarrow}} t \;&\Leftrightarrow\; s \underset{h^{-1}(G)}{\overset{\bar{g}(a)}{\Rightarrow}} t && \text{by definition of an inverse substitution} \\ &\Leftrightarrow\; s \underset{h^{-1}(G)}{\overset{g(a)}{\Rightarrow}} t && \text{by definition of } \bar{g} \\ &\Leftrightarrow\; s \underset{G}{\overset{\bar{h}(g(a))}{\Rightarrow}} t && \text{for } \varepsilon \notin g(T) \;\vee\; s, t \in V_{h^{-1}(G)} \\ &\Leftrightarrow\; s \underset{(g \circ \bar{h})^{-1}(G)}{\overset{a}{\longrightarrow}} t && \text{by definition of an inverse substitution.} \quad \square \end{aligned}$$

We will study the family REC_{Rat} of graphs obtained by applying to the (up to isomorphism) tree Λ an inverse rational substitution followed by a rational restriction:

$$REC_{Rat} := \left\{ h^{-1}(\Lambda)_{|L_\Lambda} \;\middle|\; \begin{array}{l} \Lambda \text{ is the complete and deterministic tree on } \{a, b\} \\ h : T^* \to Rat(\{a, b, \bar{a}, \bar{b}\}^*) \text{ is a morphism} \\ L \in Rat(\{a, b\}^*) \end{array} \right\}.$$

We consider also the sub-family REC_{Fin} by using only finite substitutions:

$$REC_{Fin} := \left\{ h^{-1}(\Lambda)_{|L_\Lambda} \;\middle|\; \begin{array}{l} \Lambda \text{ is the complete and deterministic tree on } \{a,b\} \\ h : T^* \to Fin(\{a,b,\bar{a},\bar{b}\}^*) \text{ is a morphism} \\ L \in Rat(\{a,b\}^*) \end{array} \right\},$$

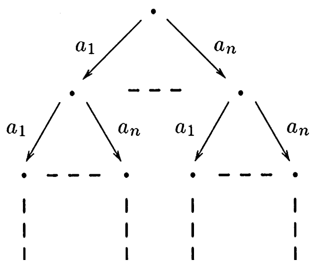
where $Fin(V)$ is the set of finite subsets of a set V.

For instance taking the (up to isomorphism) complete and deterministic tree Λ on $\{a,b\}$ and taking n labels $a_1,\ldots,a_n$, we define the following finite substitution h and the following rational language L:

$$h(a_i) = ab^{i-1}, \quad 1 \leqslant i \leqslant n,$$

$$L = (a + ab + \cdots + ab^{n-1})^*,$$

in order to obtain with $h^{-1}(\Lambda)_{|L_\Lambda}$ the complete and deterministic tree on the alphabet $\{a_1,\ldots,a_n\}$:



Thus any complete and deterministic tree is in REC_{Fin}. More generally, as we now show, REC_{Rat} (resp. REC_{Fin}) can be obtained from the complete and deterministic trees on given $n \geqslant 2$ labels by inverse rational (resp. finite) substitution followed by rational restriction.

Proposition 2.9. *Let $S \subseteq T$ of cardinal $|S| \geqslant 2$. We have*

$$REC_X = \left\{ h^{-1}(\Delta)_{|L_\Delta} \;\middle|\; \begin{array}{l} \Delta \text{ is the complete and deterministic tree on } S \\ h : T^* \to X((S \cup \bar{S})^*) \text{ is a morphism} \\ L \in Rat(S^*) \end{array} \right\},$$

where X is Rat or Fin.

Proof. Let $S = \{a_1,\ldots,a_n\}$ (with $n \geqslant 2$). For the inclusion $\subseteq$, let $h^{-1}(\Lambda)_{|L_\Lambda}$ be in REC_X.

By completion and renaming labels, there is a complete and deterministic tree Δ on S such that

$$\Lambda = g^{-1}(\Delta) \quad \text{with } g(a) = a_1 \text{ and } g(b) = a_2.$$

Note that the trees Λ and Δ have the same root, denoted by r.

Furthermore $M = g(\{a, \dots, ab^{n-1}\}^*) \supseteq g(L)$ and M is the vertex set of the connected component of $g^{-1}(\Lambda)$ containing r. So

$$L_\Delta = \{s \mid r \overset{L}{\underset{\Delta}{\Rightarrow}} s\} = \{s \mid r \overset{g(L)}{\underset{\Delta}{\Rightarrow}} s\} = g(L)_\Delta.$$

Hence

$$\begin{aligned} h^{-1}(\Delta)_{|L_\Delta} &= h^{-1}(g^{-1}(\Delta))_{|g(L)_\Delta} \\ &= (h \circ \bar{g})^{-1}(\Delta)_{|g(L)_\Delta} \qquad \text{by Lemma 2.8} \\ &= (h \circ \bar{g})^{-1}(\Lambda)_{|g(L)_\Lambda}. \end{aligned}$$

For the inclusion $\supseteq$, let $h^{-1}(\Delta)_{|L_\Delta}$ be an element of the right-hand side.

We have seen (above this proposition) that there is a complete and deterministic tree Λ on $\{a, b\}$ such that

$$\Delta = g^{-1}(\Lambda)_{|M_\Lambda} \text{ with } g(a_i) = ab^{i-1}, \quad 1 \leqslant i \leqslant n \text{ and } M = \{a, \dots, ab^{n-1}\}^*.$$

Note that the trees Λ and Δ have the same root r.

Furthermore $M = g(\{a_1, \dots, a_n\}^*) \supseteq g(L)$ and M is the vertex set of the connected component of $g^{-1}(\Lambda)$ containing r. So

$$L_\Delta = \{s \mid r \overset{L}{\underset{\Delta}{\Rightarrow}} s\} = \{s \mid r \overset{L}{\underset{g^{-1}(\Lambda)}{\Rightarrow}} s\} = \{s \mid r \overset{g(L)}{\underset{\Lambda}{\Rightarrow}} s\} = g(L)_\Lambda.$$

Hence

$$\begin{aligned} h^{-1}(\Delta)_{|L_\Delta} &= h^{-1}(g^{-1}(\Lambda)_{|M_\Lambda})_{|g(L)_\Lambda} \\ &= (h^{-1}(g^{-1}(\Lambda))_{|M_\Lambda})_{|g(L)_\Lambda} \qquad \text{by Lemma 2.7} \\ &= (h \circ \bar{g})^{-1}(\Lambda)_{|(M \cap g(L))_\Lambda} \qquad \text{by Lemma 2.8} \\ &= (h \circ \bar{g})^{-1}(\Lambda)_{|g(L)_\Lambda}. \qquad \square \end{aligned}$$

In order to get a family of graphs with a decidable monadic theory, we start with the complete ordered binary tree.

Theorem 2.10 (Rabin [11]). *Any complete and deterministic tree on two labels has a decidable monadic theory.*

Let us apply Proposition 2.6 to this result of Rabin.

Corollary 2.11. *Any graph in REC_{Rat} has a decidable monadic theory.*

Let us show that REC_{Rat} is the closure of the complete and deterministic tree on $\{a, b\}$ by the operations of Propositions 2.2 and 2.5.

Theorem 2.12. *The family REC_{Rat} is closed by rational restriction and by inverse rational substitution.*

This theorem is proved using a complete set of representatives of REC_{Rat}. We will obtain other closure properties and particularly we will deduce that this family contains strictly the graphs generated by deterministic graph grammars.

3. Complete sets of representatives

We show that the rational restrictions on the vertex sets of prefix transition graphs of labelled word rewriting systems form a complete set of representatives of REC_{Fin} (Corollary 3.5). This set of representatives contains the family of context-free graphs of [9]. In fact REC_{Fin} is exactly the class of regular graphs of finite degree (Theorem 3.11) where a regular graph (or equational graph) is a graph generated by a deterministic graph grammar. We show that the rational restrictions on the vertex sets of prefix transition graphs of labelled recognizable relations constitute a complete set of representatives of REC_{Rat} (Corollary 3.5). It follows that REC_{Rat} contains strictly the class of regular graphs (Proposition 3.12).

Finally we extend these sets of representatives to the rationally controlled prefix transition graphs of labelled recognizable relations. This set is also a complete set of representatives of REC_{Rat} (Proposition 3.16). But it is a boolean algebra preserved by inverse rational substitution and by rational restriction on vertices (Theorems 3.17 and 3.19).

We take alphabets $N \subseteq T$ containing the symbols a, b. Usually, words over N represent vertices, and T is the set of arc labels. A representative of the complete deterministic tree labelled on N is a tree Δ_N in $N^* \times N \times N^*$ defined as follows:

$$\Delta_N := \{u \xrightarrow{a} au \mid a \in N \wedge u \in N^*\}.$$

For instance $\Delta_{\{a,b\}}$ is a complete and deterministic tree on $\{a, b\}$. Note that for any $L \subseteq N^*$, the set L_{Δ_N} of vertices accessible from the root ε of Δ_N by a path labelled in L is the *mirror* $\tilde{L} := \{a_n \dots a_1 \mid a_1 \dots a_n \in L\}$ of L: $L_{\Delta_N} = \tilde{L}$. This inversion is due to the fact that we will characterize the inverse rational substitutions of Δ_N by prefix rewriting of relations (instead of suffix rewriting).

The *right closure* $G.N^*$ of a graph G in $N^* \times T \times N^*$ is

$$G.N^* := \{uw \xrightarrow{a} vw \mid u \xrightarrow{a} v \in G \wedge w \in N^*\},$$

the set of prefix transitions of G. For instance

$$\Delta_N = \{\varepsilon \xrightarrow{a} a \mid a \in N\}.N^*.$$

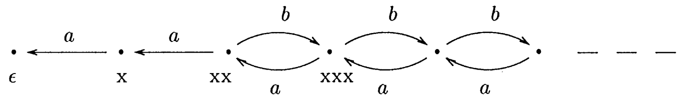
Note that a finite graph G in $N^* \times T \times N^*$ is a labelled rewriting system i.e. a finite set of rules over N and labelled in T; and the right closure of G is the prefix rewriting relation according to G.

Let us give other right closures of finite graphs. For instance the *identity graph*

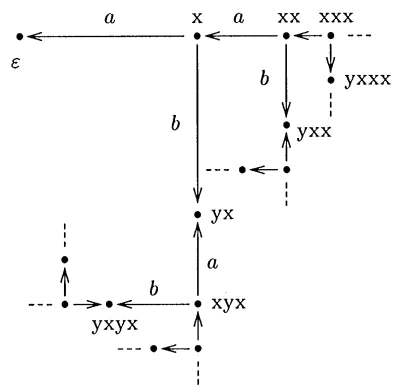
$$\{u \xrightarrow{a} u \mid a \in T \wedge u \in N^*\}$$

is the right closure of the finite graph $\{\varepsilon \xrightarrow{a} \varepsilon \mid a \in T\}$.

Consider also the finite graph $G = \{x \xrightarrow{a} \varepsilon,\ x^2 \xrightarrow{b} x^3\}$ on the non-terminal set $\{x\}$. Its right closure $G.\{x\}^*$ is the following graph:



This graph is connected. This is not the case in general. For instance, consider the graph $G = \{x \xrightarrow{a} \varepsilon, x \xrightarrow{b} yx\}$. Its right closure $G.\{x, y\}^*$ is the infinite replication of the following infinite connected component:



We denote by $U \xrightarrow{a} V := \{u \xrightarrow{a} v \mid u \in U \wedge v \in V\}$ the graph of the transitions from $U \subseteq N^*$ to $V \subseteq N^*$ and labelled by $a \in T$. A *recognizable graph* is a finite union of such graphs $U \xrightarrow{a} V$ where $U, V \in Rat(N^*)$. We denote by $Rec(N^* \times T \times N^*)$ (resp. by $Fin(N^* \times T \times N^*)$) the family of recognizable graphs (resp. finite graphs).

Note that the unlabelled recognizable graphs $\{(u, v) \mid \exists a \in T, u \xrightarrow{a} v \in G\}$ are the *recognizable relations* in $N^* \times N^*$ (by Mezei's theorem). For instance, the *full graph* $\{u \xrightarrow{a} v \mid a \in T \wedge u, v \in N^*\}$ is the recognizable graph $\bigcup \{N^* \xrightarrow{a} N^* \mid a \in T\}$ and is equal to its right closure.

Let us recall some standard and simple properties on rational languages.

Lemma 3.1. *Given $L \in Rat(N^*)$, we have effectively the following properties*:
(a) $[L]_\sim := \{vu \mid uv \in L\} \in Rat(N^*)$;
(b) $[u]_L := \{v \mid v^{-1}L = u^{-1}L\} \in Rat(N^*)$ *and* $\{[u]_L \mid u \in N^*\}$ *is finite*;
(c) $\{U^{-1}L \mid U \subseteq N^*\} \subset Rat(N^*)$ *and is finite*;
(d) *For any $M \subseteq N$, the language $L \cap M^*(N-M)^*$ is a finite union of sets of the form $A.B$ where $A \in Rat(M^*)$ and $B \in Rat((N-M)^*)$.*

The right closures of recognizable graphs in $N^* \times T \times N^*$ are exactly the inverse rational substitutions $h^{-1}(\Delta_N)$ of Δ_N (for morphisms $h: T^* \to Rat((N \cup \bar{N})^*)$), as we now show. The effectiveness claim in the statement and in similar cases below means that in both directions the respective representations can be obtained algorithmically.

Theorem 3.2. *The inverse rational (resp. finite) substitutions $h^{-1}(\Delta_N)$ of Δ_N are effectively the right closures $G.N^*$ of the recognizable graphs (resp. finite graphs) G:*

$$h^{-1}(\Delta_N) = G.N^*$$

with h rational (resp. finite) $\Leftrightarrow$ G recognizable (resp. finite).

Proof. (i) Let G be a recognizable graph, i.e. a finite union of elements in $Rat(N^*) \times T \times Rat(N^*)$. Let $h: T^* \to 2^{(N \cup \bar{N})^*}$ be a morphism defined for every $a \in T$ by

$$h(a) := \bigcup \{ \bar{U}\tilde{V} \mid U \xrightarrow{a} V \in G \}$$

So $h(a) \in Rat((N \cup \bar{N})^*)$. Let us verify that $h^{-1}(\Delta_N) = G.N^*$.

Let us prove that $h^{-1}(\Delta_N) \subseteq G.N^*$.

Let $p \xrightarrow{a} q \in h^{-1}(\Delta_N)$. There is $z \in h(a)$ such that $p \underset{\Delta_N}{\overset{z}{\Rightarrow}} q$.

By definition of $h(a)$, there is $U \xrightarrow{a} V \in G$ with $z \in \bar{U}\tilde{V}$.

Thus there are $u \in U$ and $v \in V$ such that $z = \bar{u}\tilde{v}$.

So there is $w \in N^*$ such that $w \underset{\Delta_N}{\overset{\bar{u}}{\Rightarrow}} p$ and $w \underset{\Delta_N}{\overset{\tilde{v}}{\Rightarrow}} q$.

Hence $p = uw$ and $q = vw$ i.e. $p \xrightarrow{a} q \in (U \xrightarrow{a} V).N^* \subseteq G.N^*$.

Let us prove that $G.N^* \subseteq h^{-1}(\Delta_N)$. Let $p \xrightarrow{a} q \in G.N^*$.

There are $U \xrightarrow{a} V \in G$, $u \in U$, $v \in V$, $w \in N^*$ such that $p = uw$ and $q = vw$.

So $p = uw \underset{\Delta_N}{\overset{\bar{u}}{\Rightarrow}} w \underset{\Delta_N}{\overset{\tilde{v}}{\Rightarrow}} vw = q$. Hence $p \underset{\Delta_N}{\overset{\bar{u}\tilde{v}}{\Rightarrow}} q$ with $\bar{u}\tilde{v} \in \bar{U}\tilde{V} \subseteq h(a)$.

(ii) Let $h: T^* \to Rat((N \cup \bar{N})^*)$ be a morphism.

We will simplify h by removing factors in the following finite set:

$$P := \{ x\bar{x} \mid x \in N \}.$$

The removing of a factor in P is done by the rewriting

$$\begin{aligned} \underset{P \times \{\varepsilon\}}{\longrightarrow} &:= \{ (ux\bar{x}v, uv) \mid u, v \in (N \cup \bar{N})^* \wedge x \in N \} \\ &= (N \cup \bar{N})^*.(P \times \{\varepsilon\}).(N \cup \bar{N})^*. \end{aligned}$$

The *derivation* relation $\underset{P \times \{\varepsilon\}}{\overset{*}{\longrightarrow}}$ is the reflexive and transitive closure of $\underset{P \times \{\varepsilon\}}{\longrightarrow}$.

Given any rational language $L \in Rat((N \cup \bar{N})^*)$, its set $L{\downarrow}P$ of normal forms is the following language:

$$L{\downarrow}P := \{ v \mid \exists\ u \in L,\ u \underset{P \times \{\varepsilon\}}{\overset{*}{\longrightarrow}} v \wedge v \notin (N \cup \bar{N})^* P (N \cup \bar{N})^* \}.$$

It is a rational language. Thus for any $a \in T$, the set $\underline{h}(a)$ of normal forms of $h(a)$ in $\bar{N}^* N^*$:

$$\underline{h}(a) := h(a){\downarrow}P \cap \bar{N}^* N^*$$

is a rational language.

Let us verify that $\underline{h}^{-1}(\Delta_N) = h^{-1}(\Delta_N)$.

As Δ_N is a tree, Δ_N is *co-deterministic i.e.* $p \xrightarrow{x} r \wedge q \xrightarrow{x} r \Rightarrow p = q$.

So we have for any $u, v \in N^*$, any $s, t \in (N \cup \bar{N})^*$ and any $x \in N$,

$$u \underset{\Delta_N}{\overset{sx\bar{x}t}{\Rightarrow}} v \quad \text{iff} \quad u \underset{\Delta_N}{\overset{st}{\Rightarrow}} v,$$

hence for any $z \in (N \cup \bar{N})^*$ and by induction on the length of any derivation $z \underset{P \times \{\varepsilon\}}{\overset{*}{\longrightarrow}} z{\downarrow}P$ from z to its normal form $z{\downarrow}P$ (i.e. $z{\downarrow}P \notin (N \cup \bar{N})^* P (N \cup \bar{N})^*$), we have

$$u \underset{\Delta_N}{\overset{z}{\Rightarrow}} v \text{ iff } u \underset{\Delta_N}{\overset{z\downarrow P}{\Rightarrow}} v.$$

Finally, if $z{\downarrow}P \notin \bar{N}^* N^*$ then $z{\downarrow}P$ has a factor $x\bar{y}$ with $x, y \in N$ and $x \neq y$, hence there is no path in Δ_N labelled by $z{\downarrow}P$.

(iii) By Lemma 3.1(d), for any $a \in T$, there are $n_a \geqslant 0$ and $U_1, \ldots, U_{n_a}$, $V_1, \ldots, V_{n_a} \in Rat(N^*)$ such that

$$\underline{h}(a) = \bar{U}_1.V_1 \cup \cdots \cup \bar{U}_{n_a}.V_{n_a}.$$

So we define

$$G := \bigcup \{U_i \xrightarrow{a} \tilde{V}_i \mid a \in T \wedge 1 \leqslant i \leqslant n_a\}.$$

Finally by (i), we have $G.N^* = \underline{h}^{-1}(\Delta_N)$. $\square$

This theorem implies some direct generalizations of known results. A first consequence follows from the closure by composition of (extended) rational substitutions.

Corollary 3.3. *The class of right closures of recognizable graphs is closed effectively by inverse ε-free rational substitution.*

Proof. Let G be a recognizable graph and g be a rational substitution such that $\varepsilon \notin g(T)$.

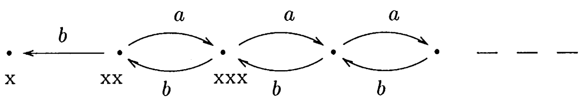
By Theorem 3.2, there is a rational substitution h such that $G.N^* = h^{-1}(\Delta_N)$.

By Lemma 2.8, we have

$$g^{-1}(G.N^*) = g^{-1}(h^{-1}(\Delta_N)) = (g \circ \bar{h})^{-1}(\Delta_N).$$

As $g \circ \bar{h}$ is a rational substitution, we have by Theorem 3.2, $g^{-1}(G.N^*) = H.N^*$ for some recognizable graph H. $\square$

As an example, consider the right closure $G.N^*$ of $G=\{x \xrightarrow{a} \varepsilon, x^2 \xrightarrow{b} x^3\}$ with $N=\{x\}$. Its inverse substitution $h^{-1}(G.N^*)$ by h, defined by $h(a)=\{b\}$ and $h(b)=\{baa\}$, is the following graph:



which is the right closure of $\{x^2 \xrightarrow{a} x^3, x^2 \xrightarrow{b} x\}$.

A consequence of Corollary 3.3 is that the unlabelled right closures of recognizable graphs are preserved by reflexive and transitive closure. More precisely, the *prefix rewriting* $\underset{R}{\mapsto}$ of any binary relation R on N^* is the unlabelled graph $R.N^*$, i.e.

$$\underset{R}{\mapsto} := \{(uw, vw) \mid uRv \wedge w \in N^*\}$$

and its reflexive and transitive closure $\overset{*}{\underset{R}{\mapsto}}$ is the *prefix derivation* relation of R.

Corollary 3.4. *The prefix derivation relation of any recognizable relation is effectively the prefix rewriting of a recognizable relation.*

Proof. Consider a recognizable relation $R \subseteq N^* \times N^*$. We take a terminal $a \in T$ to label the rules of R to get

$$\bar{R} := \{u \xrightarrow{a} v \mid (u,v) \in R\}$$

a recognizable graph. We take the substitution $h(a)=a^+$. By Corollary 3.3, we have

$$\{u \xrightarrow{a} v \mid u \overset{+}{\underset{R}{\mapsto}} v\} = h^{-1}(\bar{R}.N^*) = \bar{S}.N^*$$

for some recognizable graph $\bar{S}$. So $S := \{(u,v) \mid u \xrightarrow{a} v \in \bar{S}\} \cup \{(\varepsilon, \varepsilon)\}$ is a recognizable relation satisfying $\underset{S}{\mapsto} = \overset{*}{\underset{R}{\mapsto}}$. $\square$

In particular for any finite relation, its prefix derivation is a rational transduction [1], and this remains true for any recognizable relation. Thus for the right closure of any recognizable graph, the set of vertices accessible from any rational set is rational; this extends the rationality of the set of words accessible from a given word by prefix derivation of a finite relation [2]. Note that the algorithm given in [Bu] is of exponential complexity. A first algorithm of polynomial complexity has been given in [Ca1] (pages 93,94) to construct a finite automaton recognizing the rational set of words accessible from a given word, and more generally to construct a transducer recognizing the prefix derivation of any finite relation.

For another consequence of Theorem 3.2, we consider the following family:

$$REC_{|Rat} := \{(G.N^*)_{|L} \mid G \in Rec(N^* \times T \times N^*) \wedge L \in Rat(N^*)\}.$$

Using Proposition 2.9, we deduce that this family is a complete set of representatives of REC_{Rat}.

Corollary 3.5. *The set $REC_{|Rat}$ of the rational restrictions on vertices of the right closures of recognizable graphs* (*resp. finite graphs*) *is a complete set of representatives of REC_{Rat}* (*resp. REC_{Fin}*).

Note that Corollary 3.5 is true in particular for $N=\{a,b\}$. By Corollary 2.11, any rational restriction on vertices of the right closure of any recognizable graph has a decidable monadic theory.

Corollary 3.6. *$MTh((G.N^*)_{|L})$ is decidable for any $G \in Rec(N^* \times T \times N^*)$ and for any $L \in Rat(N^*)$.*

A particular case are the pushdown transition graphs (called also context-free graphs) considered in [9]. A *pushdown transition graph* is the graph $(R.N^*)_{|L}$ of the right closure of a pushdown automaton transition relation R in $Q.P \times T \times Q.P^*$, with $N = P \cup Q$ (where P is the set of stack letters disjoint of the set Q of states), and restricted to the set $L = \{s \mid r \underset{R}{\overset{*}{\longmapsto}} s\}$ of vertices accessible from a given axiom $r \in Q.P^*$. By Corollary 3.4, we deduce the well-known fact that L is rational, and it remains to apply Corollary 3.6 to get a principal result of [9] (Theorem 4.4).

Corollary 3.7 (Muller and Schupp [9]). *Any pushdown transition graph has a decidable monadic theory.*

The pushdown automata define up to isomorphism the same accessible prefix transition graphs as the labelled rewriting systems.

Proposition 3.8 (Caucal [3]). *The pushdown transition graphs form effectively a complete set of representatives of the rooted right closures of finite graphs.*

Instead of labelled (word) rewriting systems, we can also use a subclass of context-free term grammars [3]. We will now show that REC_{Rat} contains also the graphs generated by deterministic graph grammars.

We take a ranked set $F = \bigcup_{p \geqslant 1} F_p$ where F_p contains labels of arity p, and such that $F_2 \supseteq T$. A *hyperarc* is a word $as_1 \ldots s_p$ labelled by $a \in F_p$ (of arity p) and joining in order the vertices $s_1, \ldots, s_p$. In particular an arc $s \xrightarrow{a} t$ is the word ast with a of arity 2. A *hypergraph* is a set of hyperarcs, and a *graph* is a set of arcs.

A *graph grammar* R is a finite set of rules of the form $ax_1 \ldots x_p \rightarrow H$ where H is a finite hypergraph and $x_1, \ldots, x_p$ are distinct vertices of H. The labels of the left hand sides of R are in $F - T$ and are the *non-terminals* of R. The other labels in

R are in T. We say that R is *deterministic* if there is only one rule for each non-terminal.

A *rewriting* step $M \underset{R}{\longrightarrow} N$ consists in choosing a non-terminal hyperarc $X = as_1 \ldots s_p$ in M and a rule $ax_1 \ldots x_p \to H$ in R to be applied; the vertices x_i in H indicate how to replace X by H: we have

$$N = (M - X) \cup \{bg(t_1) \ldots g(t_q) \mid bt_1 \ldots t_q \in H\}$$

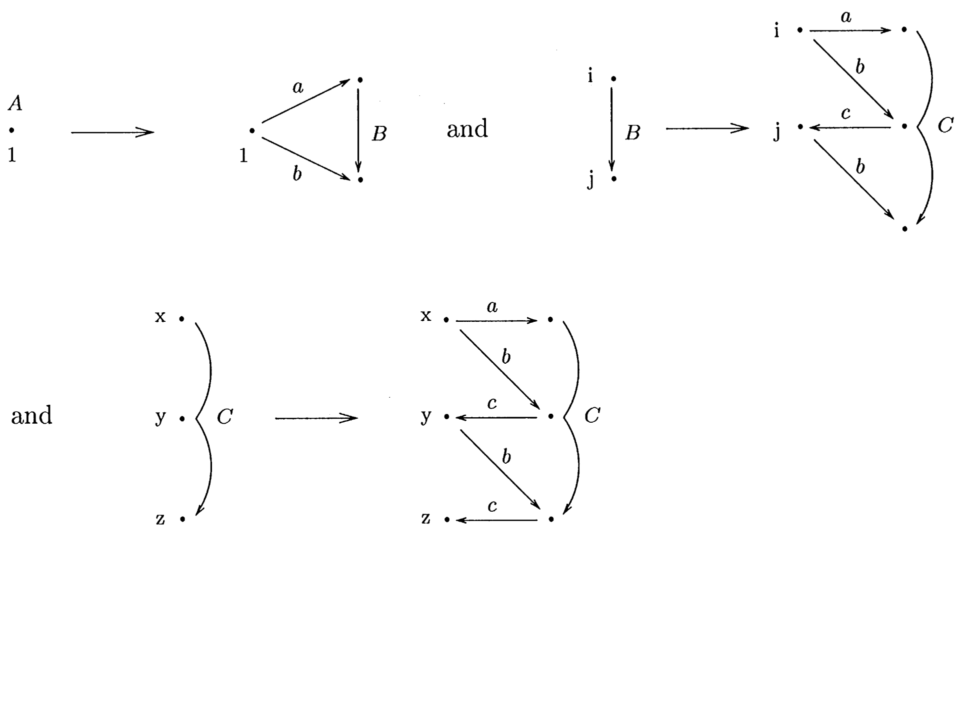
for some matching function g mapping x_i to s_i, and the other vertices of H injectively to vertices outside M; this rewriting step is also denoted by $M \underset{R,X}{\longrightarrow} N$. The rewriting $\underset{R,X}{\longrightarrow}$ of an hyperarc X is extended in an obvious way to the rewriting $\underset{R,E}{\longrightarrow}$ of any set E of non-terminal hyperarcs. A *complete parallel rewriting* $\underset{R}{\Rightarrow}$ is the rewriting according to the set of all non-terminal hyperarcs: $M \underset{R}{\Rightarrow} N$ if $M \underset{R,E}{\longrightarrow} N$ where E is the set of all non-terminal hyperarcs of M. We denote by

$$[H] := \{ast \in H \mid a \in T\},$$

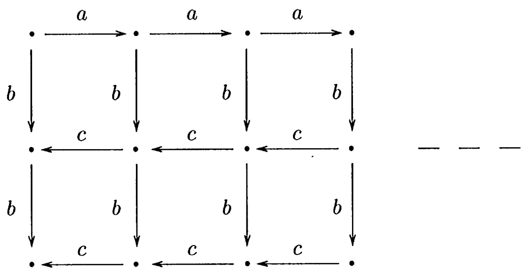
the set of terminal arcs of a hypergraph H. A graph G is *generated by a deterministic graph grammar* R from a hypergraph H if G is isomorphic to a graph of the family $R^{\omega}(H)$ defined as follows:

$$R^{\omega}(H) := \left\{ \bigcup_{n \geqslant 0} [H_n] \mid H = H_0 \underset{R}{\Rightarrow} \cdots \underset{R}{\Rightarrow} H_n \underset{R}{\Rightarrow} H_{n+1} \underset{R}{\Rightarrow} \cdots \right\}.$$

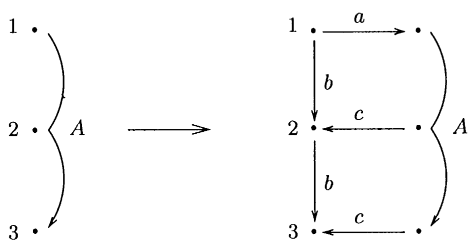
Consider, for instance, the deterministic graph grammar with the following rules:



This grammar generates from the hypergraph $\{A1\}$ the following graph:



Note that this graph can also be generated by the deterministic graph grammar with the rule:

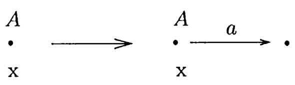


from the hypergraph $\{A123\}$.

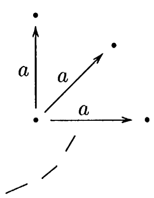
Definition 3.9. A *regular graph* is a graph generated by a deterministic graph grammar from a finite hypergraph.

These graphs are the equational graphs of [6].

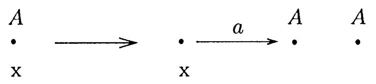
Note that a regular graph may be of *infinite degree*, where a vertex is source or target of an infinite number of arcs. For instance the deterministic graph grammar



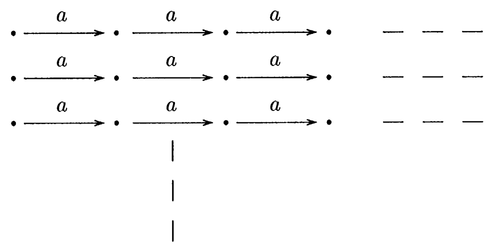
generates from its non-terminal the following graph:



of infinite out-degree. Furthermore, a regular graph may be non-connected. For instance the deterministic graph grammar



generates from its non-terminal the following graph:



with an infinite number of connected components.

Several basic properties of regular graphs are given in [5]. The regular graphs generalize the pushdown transition graphs. In fact the pushdown transition graphs are the rooted graphs of finite degree which can be finitely decomposed by distance from any vertex [9]. As the decomposition is dual to the generation, this implies that any pushdown transition graph is a rooted regular graph of finite degree, as shown by Muller and Schupp. Furthermore, the inverse inclusion remains true, and this correspondence is effective.

Proposition 3.10 (Caucal [3]). *The pushdown transition graphs form effectively a complete set of representatives of the rooted regular graphs of finite degree.*

More precisely every deterministic graph grammar generating a rooted graph G of finite degree, is mapped effectively into a pushdown automaton with an axiom such that its accessible prefix transition graph is isomorphic to G, and the reverse transformation is also effective.

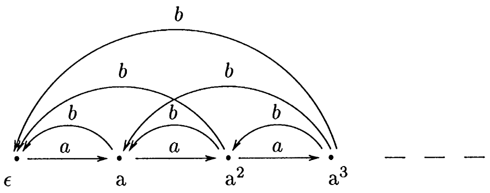
To generalize Proposition 3.10 to all the regular graphs of finite degree, it suffices to take the class of graphs of all the prefix transitions of pushdown automata (or labelled rewriting systems) and to extend this class by restriction to rational vertex sets instead to the rational set of vertices accessible from an axiom.

Theorem 3.11 (Caucal [5]). *REC_{Fin} is effectively the family of regular graphs of finite degree.*

We get the regular graphs of infinite degree with inverse rational substitutions.

Proposition 3.12. *REC_{Rat} contains strictly and effectively the class of regular graphs.*

Proof. (i) For the strict containment, we consider the rational substitution h defined by $h(a) = a$ and $h(b) = \bar{a}^+$, and the rational language $L = a^*$. Then $h^{-1}(\Delta_{\{a,b\}})_{|L} = h^{-1}(\Delta_{\{a\}})_{|L}$ is the following graph:



which is the right closure $G.\{a\}^*$ of the recognizable graph $G = \{\varepsilon \xrightarrow{a} a, a^+ \xrightarrow{b} \varepsilon\}$. By definition this graph is in REC_{Rat} but it is not regular because it has infinitely many vertex out-degrees (a^n is of out-degree $n+1$).

(ii) Let R be a deterministic graph grammar and let K be a finite hypergraph.

We will construct a recognizable graph G on a non-terminal set N, and a rational language $L \subseteq N^*$ such that $(G.N^*)_{|L}$ belongs to $R^\omega(K)$.

Recall that V_H is the set of vertices of any hypergraph H, and that $|X|$ is the length of any word X. We take a new alphabet $V = \{x_1, \ldots, x_m\}$ of *variables* where

$$m := \max\{|X| - 1 \mid X \in Dom(R)\}$$

is the maximum number of vertices needed by each non-terminal hyperarc.

After a possible renaming of vertices, we can assume that for every rule $(X, H) \in R$,

$$\begin{array}{rcll} X & = & X(1)x_1 \ldots x_{|X|-1} & \text{for every } X \in Dom(R) \\ V_H \cap V_{H'} & \subseteq & V_X \cap V_{X'} & \text{for every distinct rules } (X,H),(X',H') \in R. \end{array}$$

Adding a new rule, we can assume that K is restricted to a non-terminal hyperarc.

Let us give some notations and definitions.

We denote by $\bar{V}$ the set of vertices of R *i.e.*

$$\bar{V} := V \cup \bigcup\{V_H \mid H \in Im(R)\}.$$

We take a set $\bar{N}$ of non-terminals defined by

$$\bar{N} := \bigcup_{p=1}^{m} \bar{V}^p.$$

Let $1 \leqslant p \leqslant m$. For every word $u \in \bar{N}^+$, we define

$$u\langle x_1, \ldots, x_p\rangle := \begin{cases} u & \text{if } u \in V, \\ u(x_1, \ldots, x_p) & \text{if } u \notin V, \end{cases}$$

the *right addition* of $(x_1, \ldots, x_p)$ to u when u is not a variable. This operation is extended to any transition: for every non-terminal words $u, v \in \bar{N}^+$ and every terminal

$a \in T$,

$$(u \xrightarrow{a} v)\,\langle x_1,\dots,x_p\rangle := u\langle x_1,\dots,x_p\rangle \xrightarrow{a} v\langle x_1,\dots,x_p\rangle.$$

Finally the right addition is extended by union to any graph labelled in T.

For every $v_1,\dots,v_p \in \bar{V}$, the *substitution* $u[v_1,\dots,v_p]$ in any word $u \in \bar{N}^+$ of the x_i by v_i is the morphism defined on every letter of $\bar{N}$ by

$$\begin{aligned} x_i[v_1,\dots,v_p] &:= v_i, && \forall 1 \leqslant i \leqslant p,\\ s[v_1,\dots,v_p] &:= s, && \forall s \in \bar{V} - \{x_1,\dots,x_p\},\\ (s_1,\dots,s_q)\,[v_1,\dots,v_p] &:= (s_1[v_1,\dots,v_p],\dots,\\ &\qquad s_q[v_1,\dots,v_p]), && \forall q > 1. \end{aligned}$$

The substitution is extended to any transition

$$(u \xrightarrow{a} v)\,[v_1,\dots,v_p] := u[v_1,\dots,v_p] \xrightarrow{a} v[v_1,\dots,v_p]$$

and by union to any graph.

Recall that $[H] := \{ast \in H \mid a \in T\}$ is the set of terminal arcs of any hypergraph H. To every $X \in Dom(R)$, we associate a representative $\bar{G}_{X(1)}$ of $R^{\omega}(X)$.

These graphs $\bar{G}_{X(1)}$ are the least fixpoints of the following equations:

$$\bar{G}_{X(1)} = \Big([H] \cup \bigcup\{\ \bar{G}_{Y(1)}[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\Big)\\ \langle x_1,\dots,x_{|X|-1}\rangle$$

for every rule $(X,H) \in R$.

Let us verify that the set

$$L_{X(1)} := V_{\bar{G}_{X(1)}}$$

of vertices of $\bar{G}_{X(1)}$ is a rational language over $\bar{N}$ effectively obtained from (R,K). Note that

$$\hat{L}_{X(1)} := L_{X(1)} \cap V_X$$

is an effective finite language. So

$$L_{X(1)} = \hat{L}_{X(1)} \cup \bar{L}_{X(1)}(x_1,\dots,x_{|X|-1})$$

with

$$\begin{aligned} \bar{L}_{X(1)} &= \Big(V_{[H]} \cup \bigcup\{L_{Y(1)}[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\Big) - V_X\\ &= (V_{[H]} - V_X) \cup \bigcup\{\hat{L}_{Y(1)}[Y(2),\dots,Y(|Y|)] - V \mid Y \in H \wedge Y(1) \notin T\}\\ &\quad \cup \bigcup\{\bar{L}_{Y(1)}(x_1,\dots,x_{|Y|-1})[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\\ &= \Big(V_{[H]} \cup \bigcup\{\hat{L}_{Y(1)}[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\Big) - V\\ &\quad \cup \bigcup\{\ \bar{L}_{Y(1)}[Y(2),\dots,Y(|Y|)](Y(2),\dots,Y(|Y|)) \mid Y \in H \wedge Y(1) \notin T\}. \end{aligned}$$

We deduce the following left linear grammar:

$$X(1) = \hat{L}_{X(1)} \cup X(1)_{[x_1,\dots,x_{|X|-1}]}.(x_1,\dots,x_{|X|-1}),$$

$$\begin{aligned} X(1)_{[\vec{v}]} = \big(V_{[H]} \cup \bigcup\{\hat{L}_{Y(1)}[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\big) - V \\ \cup\bigcup\{Y(1)_{[Y(2)[\vec{v}],\dots,Y(|Y|)[\vec{v}]]}(Y(2)[\vec{v}],\dots, \\ Y(|Y|)[\vec{v}]) \mid Y \in H \wedge Y(1) \notin T\} \end{aligned}$$

for every rule $(X,H) \in R$ and for every $\vec{v} := v_1,\dots,v_{|X|-1} \in \bar{V}$.

This left linear grammar generates from the non-terminal $X(1)$ the rational language $L_{X(1)}$. It remains to define a recognizable graph $G_{X(1)}$ such that

$$\bar{G}_{X(1)} = (G_{X(1)}.\bar{N}^*)_{|L_{X(1)}}.$$

We take

$$\begin{aligned} G_{X(1)} &= \{u \xrightarrow{a} v \mid \exists w,\ uw \xrightarrow{a} vw \in \bar{G}_{X(1)} \wedge |u| \\ &= \min(2,|uw|) \wedge |v| = \min(2,|vw|)\}. \end{aligned}$$

To construct $G_{X(1)}$ and to verify that $G_{X(1)}$ is recognizable, we restrict the right addition for transitions as follows:

$$(u \xrightarrow{a} v)\,\langle\langle x_1,\dots,x_p\rangle\rangle := \begin{cases} u \xrightarrow{a} v & \text{if } \min(|u|,|v|) \geqslant 2, \\ u\langle x_1,\dots,x_p\rangle \xrightarrow{a} v\langle x_1,\dots,x_p\rangle & \text{otherwise.} \end{cases}$$

Note that $(u \xrightarrow{a} v)\,\langle\langle x_1,\dots,x_p\rangle\rangle = u \xrightarrow{a} v$ iff $\min(|u|,|v|) \geqslant 2 \vee u,v \in V$.

By restriction of the right addition in the equations defining the graphs $\bar{G}_{X(1)}$, we obtain the graphs $G_{X(1)}$ which are the least fixpoints of the following equations:

$$\begin{aligned} G_{X(1)} = &\big([H] \cup \bigcup\{G_{Y(1)}[Y(2),\dots,Y(|Y|)] \mid Y \in H \wedge Y(1) \notin T\}\big) \\ &\langle\langle x_1,\dots,x_{|X|-1}\rangle\rangle \end{aligned}$$

for every rule $(X,H) \in R$.

By left linearity of these equations, these graphs $G_{X(1)}$ are recognizable.

In particular the recognizable graph

$$G := G_{K(1)}[K(2),\dots,K(|K|)]$$

and the rational language

$$L := L_{K(1)}[K(2),\dots,K(|K|)]$$

over

$$N := \bar{N} \cup \{K(2),\dots,K(|K|)\}$$

are appropriate: the right closure $(G.N^*)_{|L}$ of G restricted to L belongs to $R^{\omega}(K)$.

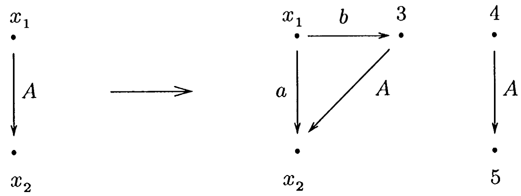
Note that $L_{X(1)}$ is the vertex set of connected components of $G_{X(1)}.\bar{N}^*$:

$$(u \xrightarrow{a} v) \in G_{X(1)} \wedge \{uw, vw\} \cap L_{X(1)} \neq \emptyset \Rightarrow \{uw, vw\} \subseteq L_{X(1)}$$

i.e. $L_{X(1)}$ is closed by $\underset{G_{X(1)}.\bar{N}^*}{\leftrightarrow}$.

In particular L is closed by $\underset{G.N^*}{\leftrightarrow}$: it is the vertex set of connected components of $G.N^*$. $\square$

Example 3.13. Let us apply the proof of Proposition 3.12 to the following deterministic graph grammar R with $K = A12$:



As defined in the proof of Proposition 3.12, we have the following equation:

$$\begin{aligned} G_A = \;& \{x_1 \xrightarrow{a} x_2,\ x_1 \xrightarrow{b} 3\}\,\langle\langle x_1, x_2\rangle\rangle \cup G_A[3, x_2]\,\langle\langle x_1, x_2\rangle\rangle \\ & \cup G_A[4,5]\,\langle\langle x_1, x_2\rangle\rangle. \end{aligned}$$

Its least fixpoint is

$$G_A = \{x_1 \xrightarrow{a} x_2, x_1 \xrightarrow{b} 3(x_1, x_2)\}\ ([3, x_2]\langle\langle x_1, x_2\rangle\rangle + [4,5]\langle\langle x_1, x_2\rangle\rangle)^*,$$

which gives the following recognizable graph $G = G_A[1,2]$:

$$\begin{array}{rlrl}
1 & \xrightarrow{a} 2, & 1 & \xrightarrow{b} 3(1,2), \\
3(1,2) & \xrightarrow{b} 3(3,2)(1,2), & 3(3,2) & \xrightarrow{b} 3(3,2)(3,2), \\
3(3,2)^*(1,2) & \xrightarrow{a} 2, & & \\
4(1,2) & \xrightarrow{a} 5(1,2), & 4(1,2) & \xrightarrow{b} 3(4,5)(1,2), \\
3(4,5) & \xrightarrow{b} 3(3,5)(4,5), & 3(3,5) & \xrightarrow{b} 3(3,5)(3,5), \\
3(3,5)^*(4,5)(1,2) & \xrightarrow{a} 5(1,2), & & \\
4(3,2) & \xrightarrow{a} 5(3,2), & 4(3,2) & \xrightarrow{b} 3(4,5)(3,2), \\
3(3,5)^*(4,5)(3,2) & \xrightarrow{a} 5(3,2), & & \\
4(4,5) & \xrightarrow{a} 5(4,5), & 4(4,5) & \xrightarrow{b} 3(4,5)(4,5), \\
3(3,5)^*(4,5)(4,5) & \xrightarrow{a} 5(4,5), & & \\
4(3,5) & \xrightarrow{a} 5(3,5), & 4(3,5) & \xrightarrow{b} 3(4,5)(3,5), \\
3(3,5)^*(4,5)(3,5) & \xrightarrow{a} 5(3,5), & &
\end{array}$$

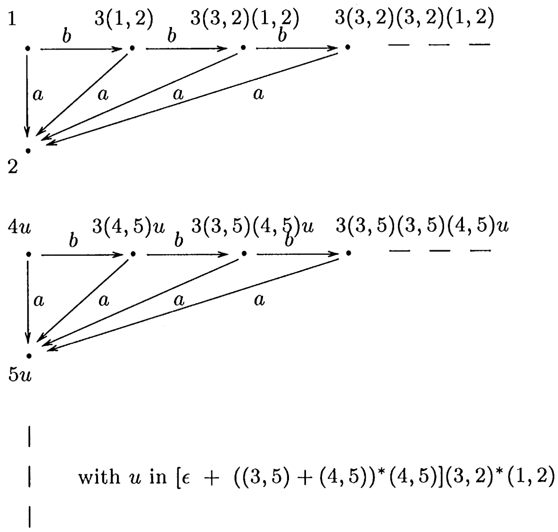
on the set $N := \{1\ ,\ 2\ ,\ 3\ ,\ 4\ ,\ 5\ ,\ (1,2)\ ,\ (3,2)\ ,\ (3,5)\ ,\ (4,5)\}$ of non-terminals. The set L_A of allowed vertices is generated by the following left linear context-free grammar:

$$\begin{aligned}
A &= x_1 + x_2 + A_{[x_1,x_2]}\,(x_1,x_2),\\
A_{[x_1,x_2]} &= 3+4+5+A_{[3,x_2]}\,(3,x_2)+A_{[4,5]}\,(4,5),\\
A_{[3,x_2]} &= 3+4+5+A_{[3,x_2]}\,(3,x_2)+A_{[4,5]}\,(4,5),\\
A_{[4,5]} &= 3+4+5+A_{[3,5]}\,(3,5)+A_{[4,5]}\,(4,5),\\
A_{[3,5]} &= 3+4+5+A_{[3,5]}\,(3,5)+A_{[4,5]}\,(4,5),
\end{aligned}$$

which gives the following rational language $L = L_A[1,2]$:

$$L := 1+2+(3+4+5)\,[\varepsilon + ((3,5)+(4,5))^*(4,5)]\,(3,2)^*(1,2).$$

Then the right closure $(G.N^*)_{|L}$ of G restricted to L is the following graph:



which is a graph generated by R from K. □

Several characterizations of the class of regular graphs as a subset of REC_{Rat} have been given. It remains to apply Corollary 2.11 to get Theorem 7.11 of [6].

Corollary 3.14 (Courcelle [6]). *Any regular graph has a decidable monadic theory.*

Thus Corollaries 3.7 and 3.14 have been obtained by using the following complete set of representatives of REC_{Rat} (see Corollary 3.5):

$$REC_{|Rat} := \ \{(G.N^*)_{|L} \mid G \in Rec(N^* \times T \times N^*) \wedge L \in Rat(N^*)\}.$$

Although $REC_{|Rat}$ is obviously closed by rational restriction on vertices, it is not closed for instance by inverse morphism, nor by union. As an example, consider h with

$h(a) = ba$ and $h(b) = b$. Then with $x, y \in N$, we have

$$\begin{aligned} & h^{-1}(\{\varepsilon \xrightarrow{b} y, y \xrightarrow{a} x, x \xrightarrow{b} xy\}.N^*_{|\{\varepsilon,x,y,xy\}}) \\ & \quad = \{\varepsilon \xrightarrow{a} x, \varepsilon \xrightarrow{b} y, x \xrightarrow{b} xy\} \\ & \quad = \{\varepsilon \xrightarrow{a} x, \varepsilon \xrightarrow{b} y\}.N^*_{|\{\varepsilon,x,y\}} \cup \{x \xrightarrow{b} xy\}.N^*_{|\{x,xy\}} \end{aligned}$$

and this graph is not in $REC_{|Rat}$ otherwise $y \xrightarrow{a} xy$ would be in the graph. A simple extension is to take the family $\bigcup_f REC_{|Rat}$ of finite unions of graphs in $REC_{|Rat}$. But $\bigcup_f REC_{|Rat}$ is not closed by complement: consider

$$(\varepsilon \xrightarrow{a} x^*).N^*_{|x^*} - (x \xrightarrow{a} x^+).N^*_{|x^*} = \varepsilon \xrightarrow{a} x^*$$

and we can show (not verified here) that this graph $\varepsilon \xrightarrow{a} x^*$ is not in $\bigcup_f REC_{|Rat}$.

Now, we give another complete set of representatives of REC_{Rat} which is a boolean algebra, closed by inverse rational substitution and by rational restriction on vertices. Following [**?**], we extend the right closures of recognizable graphs to rational right closures.

Definition 3.15. A *rational right closure of a recognizable graph* is a finite union of graphs

$$(U \xrightarrow{a} V).W := \{uw \xrightarrow{a} vw \mid u \in U \wedge v \in V \wedge w \in W\},$$

where $U, V, W \in Rat(N^*)$.

For instance $(A \xrightarrow{a} BA).(BA)^* \cup (B \xrightarrow{b} AB).A(BA)^*$ is the following graph:

$$\underset{A}{\bullet} \xrightarrow{\quad a \quad} \underset{BA}{\bullet} \xrightarrow{\quad b \quad} \underset{ABA}{\bullet} \xrightarrow{\quad a \quad} \underset{BABA}{\bullet} \xrightarrow{\quad b \quad} \underset{ABABA}{\bullet} - - -$$

Let us verify that the rational right closures of recognizable graphs are also in REC_{Rat}.

Proposition 3.16. *The rational right closures of recognizable graphs form a complete set of representatives of* REC_{Rat}.

Proof. (i) Let G be a regular graph and g be an extended rational substitution *i.e.* g is a morphism $T^* \to Rat((T \cup \bar{T})^*)$. Let us show that $g^{-1}(G) \in REC_{Rat}$.

By the proof of Proposition 3.12, we can construct a recognizable graph H in $N^* \times T \times N^*$ and a rational set L over N such that $H.N^*_{|L}$ is isomorphic to G and L is closed by $\underset{H.N^*}{\longleftrightarrow}$ i.e. L is the vertex set of connected components of $H.N^*$.

By Theorem 3.2, we can construct a rational substitution h such that $h^{-1}(\Delta_N) = H.N^*$. Hence

$$\begin{aligned} g^{-1}(H.N^*_{|L}) &= g^{-1}(H.N^*)_{|L} && \text{by Lemma 2.7} \\ &= g^{-1}(h^{-1}(\Delta_N))_{|L} \\ &= (g \circ \bar{h})^{-1}(\Delta_N)_{|V_{h^{-1}(\Delta_N)} \cap L} && \text{by Lemma 2.8} \\ &= (g \circ \bar{h})^{-1}(\Delta_N)_{|L} \\ &= (g \circ \bar{h})^{-1}(\Delta_N)_{|\tilde{L}_\Delta}. \end{aligned}$$

By Proposition 2.9, $g^{-1}(H.N^*_{|L})$ is in REC_{Rat}. As REC_{Rat} is closed by isomorphism, $g^{-1}(G)$ is also in REC_{Rat}.

(ii) Note that any graph in $REC_{|Rat}$ is a rational closure of a recognizable graph. More generally the family of the rational closures of the recognizable graphs is closed by rational restriction on vertices (see (i) of the proof of Theorem 3.19).

It remains to prove that any rational right closure of a recognizable graph is in REC_{Rat}. Let G be a rational right closure of a recognizable graph in $N^* \times T \times N^*$:

$$G = \bigcup_{i \in I} (U_i \xrightarrow{a_i} V_i).W_i,$$

where I is finite and for every $i \in I$, $U_i, V_i, W_i \in Rat(N^*)$.

To each $a \in T$, we associate a new symbol $\$_a$ and we consider the rational language $L_a := \bigcup\{W_i \mid i \in I \wedge a_i = a\}$ of vertices of the tree Δ_N which must be marked by $\$_a$. We extend N to the following alphabet $M := N \cup \{\$_a \mid a \in T\}$ and we take the following rational language over M:

$$L := N^* \cup \bigcup\{\$_a.L_a \mid a \in T\}.$$

So the tree

$$H := (\Delta_M)_{|L} = \Delta_N \cup \{w \xrightarrow{\$_{a_i}} \$_{a_i}.w \mid i \in I \wedge w \in W_i\}$$

is a regular tree.

We define the following rational extended substitution h:

$$h(a) := \bigcup\{\overline{U_i}\$_a\overline{\$_a}\tilde{V}_i \mid i \in I \wedge a_i = a\} \quad \text{for each } a \in T.$$

Then $G = h^{-1}(H)$ and by (i), G is in REC_{Rat}. $\square$

Contrary to $REC_{|Rat}$ the rational right closures of recognizable graphs are preserved by boolean operations.

Theorem 3.17. *The rational right closures of recognizable graphs form an effective boolean algebra.*

Proof. First we restrict the recognizable relations in such a way that their right closures form a boolean algebra. We say that a graph G is *right-irreducible* if for any transition

$u \xrightarrow{a} v$ in G, u and v are words having different last letters (when they exist):

$$u = \varepsilon \vee v = \varepsilon \vee u(|u|) \neq v(|v|).$$

In particular the recognizable graph:

$$\{N^* \xrightarrow{a} \varepsilon \mid a \in T\} \cup \{\varepsilon \xrightarrow{a} N^* \mid a \in T\}$$
$$\cup \{N^*x \xrightarrow{a} N^*y \mid a \in T \wedge x, y \in N \wedge x \neq y\}$$

is right-irreducible and its right closure is the full graph.

(i) Let us show that the right closures of right-irreducible recognizable graphs form an effective boolean algebra.

For any graphs R and S in $N^* \times T \times N^*$, we have

$$R.N^* \cup S.N^* = (R \cup S).N^*. \tag{1}$$

Assume that R and S are right-irreducible. Note that for any $u \xrightarrow{a} v \in R.N^*$, there is a unique w and a unique $x \xrightarrow{a} y \in R$ such that $u = xw$ and $v = yw$: w is the greatest common suffix of u and v. Hence

$$R.N^* \cap S.N^* = (R \cap S).N^*. \tag{2}$$

So (1) and (2) imply that

$$R.N^* - S.N^* = (R - S).N^*$$

because

$$\begin{aligned}(R - S)N^* \cup (RN^* \cap SN^*) &= (R - S)N^* \cup (R \cap S)N^* \\ &= ((R - S) \cup (R \cap S))N^* = RN^*, \\ (R - S)N^* \cap (RN^* \cap SN^*) &= (R - S)N^* \cap (R \cap S)N^* \\ &= ((R - S) \cap (R \cap S))N^* = \emptyset.\end{aligned}$$

Note that the full graph is also the right closure $R.N^*$ of the following right-irreducible recognizable graph R:

$$\{N^* \xrightarrow{a} \varepsilon \mid a \in T\} \cup \{\varepsilon \xrightarrow{a} N^* \mid a \in T\}$$
$$\cup \{N^*x \xrightarrow{a} N^*y \mid a \in T \wedge x, y \in N \wedge x \neq y\}.$$

Finally the unlabelled recognizable graphs are the recognizable subsets of the product of N^* (Mezei theorem), hence form an effective boolean algebra (otherwise it suffices to apply Lemma 3.1(d)).

Let us extend (i) to rational right closures.

(ii) The rational right closures of right-irreducible recognizable graphs form an effective boolean algebra.

By definition, the rational right closures of recognizable graphs are preserved by (finite) union, and it is the same when the recognizable graphs are right-irreducible.

Let us show the closure by intersection.

Consider an elementary graph $(U \xrightarrow{a} V).W$ such that $U, V, W \in Rat(N^*)$ and $U \xrightarrow{a} V$ is right-irreducible: $\forall u \in U - \{\varepsilon\}$, $\forall v \in V - \{\varepsilon\}$, we have $u(|u|) \neq v(|v|)$.

Note that for any $x \xrightarrow{a} y \in (U \xrightarrow{a} V).W$, there are unique $u \in U, v \in V, w \in W$ such that $x = uw$ and $y = vw$: w is the greatest common suffix of x and y.

Consider another elementary graph $(X \xrightarrow{a} Y).Z$ such that $X, Y, Z \in Rat(N^*)$ and $X \xrightarrow{b} Y$ is right-irreducible. By right irreducibility, the intersection

$$(U \xrightarrow{a} V).W \cap (X \xrightarrow{b} Y).Z := \begin{cases} \emptyset & \text{if } a \neq b, \\ ((U \cap X) \xrightarrow{a} (V \cap Y)).(W \cap Z) & \text{if } a = b. \end{cases}$$

is a rational right closure of a right-irreducible recognizable graph.

The closure by union and the distributivity of the intersection over union imply that the family of rational right closures of right-irreducible recognizable graphs is closed by intersection.

Let us show the closure by complementation.

Consider an elementary graph $(U \xrightarrow{a} V).W$ such that $U, V, W \in Rat(N^*)$ and $U \xrightarrow{a} V$ is right-irreducible.

By closure by intersection, it suffices to show that the complement $\overline{(U \xrightarrow{a} V).W}$ of $(U \xrightarrow{a} V).W$ is a rational closure of a right-irreducible recognizable graph. We have

$$\begin{aligned} &(U \xrightarrow{a} V).W \cup (U \xrightarrow{a} V).(N^* - W) = (U \xrightarrow{a} V).N^*, \\ &(U \xrightarrow{a} V).W \cap (U \xrightarrow{a} V).(N^* - W) = \emptyset \qquad \text{as seen above.} \end{aligned}$$

It follows that

$$\overline{(U \xrightarrow{a} V).W} = \overline{(U \xrightarrow{a} V).N^*} \cup (U \xrightarrow{a} V).(N^* - W).$$

By (i), this implies that $\overline{(U \xrightarrow{a} V).W}$ is a rational right closure of a right-irreducible recognizable graph.

This proves (ii).

To prove the theorem, we will show that any rational right closure of a recognizable graph can be expressed by a rational right closure of a right-irreducible recognizable graph.

(iii) Let us show that any rational right closure of a recognizable graph is equal effectively to a rational right closure of a right-irreducible recognizable graph.

By Lemma 3.1 and using the mirror operation, for any $L \in Rat(N^*)$ and $u \in N^*$,

$$_{\mathrm{L}}[u] := \{v \mid Lv^{-1} = Lu^{-1}\} \in Rat(N^*).$$

We want to express any elementary graph $(U \xrightarrow{a} V).W$ with $U, V, W \in Rat(N^*)$ as a rational right closure of a right-irreducible recognizable graph. Note that

$$\begin{aligned} (U \xrightarrow{a} V).W = &\bigcup \{((Ux^{-1})x \xrightarrow{a} (Vy^{-1})y).W \mid x, y \in N \wedge x \neq y\} \\ &\cup \{(\varepsilon \xrightarrow{a} V).W \mid \varepsilon \in U\} \cup \{(U \xrightarrow{a} \varepsilon).W \mid \varepsilon \in V\} \\ &\cup \{(Ux^{-1} \xrightarrow{a} Vx^{-1}).xW \mid x \in N\}. \end{aligned}$$

By applying the same equality to $(Ux^{-1} \xrightarrow{a} Vx^{-1}).xW$ and by induction, we get that

$$\begin{aligned}(U \xrightarrow{a} V).W = \bigcup\{((U(xw)^{-1})x \xrightarrow{a} (V(yw)^{-1})y).({}_{\mathrm{U}}[w] \cap {}_{\mathrm{V}}[w])W \mid \\ w \in N^* \wedge x, y \in N \wedge x \neq y\} \\ \cup \bigcup\{(\varepsilon \xrightarrow{a} Vw^{-1}).(U \cap {}_{\mathrm{V}}[w])W \mid w \in N^*\} \\ \cup \bigcup\{(Uw^{-1} \xrightarrow{a} \varepsilon).({}_{\mathrm{U}}[w] \cap V)W \mid w \in N^*\}\end{aligned}$$

is a rational right closure of a recognizable graph. □

Let us extend Corollary 3.4.

Proposition 3.18. *The rational right closures of recognizable relations are effectively preserved by transitive closure.*

Proof. Consider a rational right closure of an unlabelled recognizable graph:

$$\bigcup_{i \in I}(U_i \times V_i).W_i \quad \text{where } U_i, V_i, W_i \in Rat(N^*) \text{ for every } i \in I \text{ finite,}$$

i.e. the prefix rewriting $\underset{R,f}{\mapsto}$ of the finite binary relation

$$R = \{(U_i, V_i) \mid i \in I\}$$

on $Rat(N^*)$ and controlled by the mapping $f : R \to Rat(N^*)$ defined by $f(U_i, V_i) = W_i$. In fact we may have $(U_i, V_i) = (U_j, V_j)$ for $i \neq j$, and the domain of the mapping f must be I instead of R.

(i) Let us show that we may assume that for distinct rules (U, V) and (U', V') of R,

$$f(U, V) \cap f(U', V') \neq \emptyset \Rightarrow f(U, V) = f(U', V') \wedge U \neq U' \wedge V \neq V'.$$

First, we prove this implication when $f(U, V) = f(U', V')$ for every $(U, V), (U', V') \in R$.

More precisely, we can transform any finite (resp. and rational) relation

$$R = \{(U_i, V_i) \mid i \in I\} \text{ with } I \text{ finite (resp. and } U_i, V_i \in Rat(N^*))$$

into a finite (resp. and rational) relation

$$[R] = \{(X_i, Y_i) \mid i \in J\} \text{ with } J \text{ finite (resp. and } X_i, Y_i \in Rat(N^*))$$

such that

$$\bigcup_{i \in I} U_i \times V_i = \bigcup_{i \in J} X_i \times Y_i \wedge |J| \leqslant |I|$$

with

$$X_i \neq X_j \wedge Y_i \neq Y_j \quad \text{for every } i \neq j \text{ in } J.$$

This transformation is done by induction on the cardinal of $R : |R| = |I| \geqslant 0$.

Base case : $|I| = 0$. So $[R] = R$ suits.
Inductive case : Let $(U,V) \in R$. We denote by

$$S = R - \{(U,V)\}.$$

By induction hypothesis, we can transform S into

$$[S] = \{(X_i, Y_i) \mid i \in J\}$$

such that

$$\bigcup_{i \in J} X_i \times Y_i = \bigcup\{X \times Y \mid (X,Y) \in R \wedge (X,Y) \neq (U,V)\}$$

and

$$X_i \neq X_j \wedge Y_i \neq Y_j \quad \text{for every } i \neq j \text{ in } J.$$

We have one of the following three cases below.
Case 1: $U \neq X_i \wedge V \neq Y_i$ for every $i \in J$. So $[R] = \{(U,V)\} \cup [S]$ is appropriate.
Case 2: $U = X_i$ for some (unique) $i \in J$. Then the relation

$$R' = \{(U, Y_i \cup V)\} \cup ([S] - \{(X_i, Y_i)\})$$

is of cardinal $|R'| = |[S]| \leqslant |S| < |R|$ and by induction hypothesis $[R] = [R']$ is appropriate.

Case 3: $V = Y_i$ for some (unique) $i \in J$. Then the relation

$$R' = \{(X_i \cup U, V)\} \cup ([S] - \{(X_i, Y_i)\})$$

is of cardinal $|R'| = |[S]| \leqslant |S| < |R|$ and by induction hypothesis $[R] = [R']$ is appropriate.

Let us prove (i). The system (R, f) is of the form

$$R.f = \{(U_i, V_i).W_i \mid i \in I\}.$$

Note that any system reduced to at most one rule satisfies (i).

By induction, it remains to prove that for any system $R.f$ satisfying (i) and for any $U, V, W \in Rat(N^*)$, we can construct a system $S.g$ satisfying (i) and defining the same graph than $R.f \cup (U,V).W$. It suffices to take

$$\begin{aligned} S.g := {} & \{(U_i, V_i).(W_i - W) \mid i \in I\} \\ & \cup \left\{(U,V).\left(W - \bigcup_{i \in I} W_i\right)\right\} \\ & \cup \bigcup_{i \in I} [\{(U,V)\} \cup \{(U_j, V_j) \mid j \in I \wedge W_j = W_i\}].(W_i \cap W). \end{aligned}$$

(ii) We denote by $\underset{R,f}{\longmapsto}_n$ the restriction of the prefix rewriting $\underset{R,f}{\longmapsto}$ which does not rewrite suffixes of length n:

$$u \underset{R,f}{\longmapsto}_n v \text{ if } \exists\ (X,Y) \in R\ \exists x \in X\ \ \exists y \in Y\ \ \exists z \in f(X,Y),$$
$$u = xz \wedge v = yz \wedge |z| \geqslant n.$$

Note that $\underset{R,f}{\longmapsto}_m \subseteq \underset{R,f}{\longmapsto}_n$ for $m \geqslant n$, and $\underset{R,f}{\longmapsto}_0 = \underset{R,f}{\longmapsto}$.

The derivation $\overset{*}{\underset{R,f}{\longmapsto}}_n$ is the reflexive and transitive closure $(\underset{R,f}{\longmapsto}_n)^*$ of the rewriting $\underset{R,f}{\longmapsto}_n$.

Let us verify that

$$x \overset{+}{\underset{R,f}{\longmapsto}} y$$
$$\Leftrightarrow \exists\ (U,V) \in R\ \exists u \in U\ \ \exists v \in V\ \ \exists w \in f(U,V), \quad x \overset{*}{\underset{R,f}{\longmapsto}}_{|w|} uw \wedge vw \overset{*}{\underset{R,f}{\longmapsto}}_{|w|} y.$$

The sufficient condition is due to the closure of $\overset{*}{\underset{R,f}{\longmapsto}}_{|w|}$ by composition, and to the fact that for $(U,V) \in R$, $u \in U$, $v \in V$, $w \in f(U,V)$, we have $uw \underset{R,f}{\longmapsto}_{|w|} vw$.

Let us prove the necessary condition by induction on $n \geqslant 1$ for $x \underset{R,f}{\longmapsto}^n y$.

$n = 1$: there are $(U,V) \in R$, $u \in U$, $v \in V$, $w \in f(U,V)$ such that $x = uw$ and $vw = y$.

$n \Rightarrow n+1$: there is s such that $x \underset{R,f}{\longmapsto}^n s \underset{R,f}{\longmapsto} y$.

By induction hypothesis, there are $(U',V') \in R$, $u' \in U'$, $v' \in V', w' \in f(U',V')$ such that

$$x \overset{*}{\underset{R,f}{\longmapsto}}_{|w'|} u'w' \wedge v'w' \overset{*}{\underset{R,f}{\longmapsto}}_{|w'|} s.$$

In particular there is s' such that $s'w' = s$.

As $s \underset{R,f}{\longmapsto} y$, there are $(P,Q) \in R$, $p \in P$, $q \in Q$, $t \in f(P,Q)$ such that

$$s = pt \quad \text{and} \quad y = qt.$$

Thus $s'w' = pt$ and we distinguish the two cases below.

Case 1: $|w'| \geqslant |t|$. There is h such that $w' = ht$ hence $p = s'h$.

Thus $x \overset{*}{\underset{R,f}{\longmapsto}}_{|t|} pt \wedge qt = y$.

Case 2: $|w'| < |t|$. There is h such that $t = hw'$ hence $s' = ph$.

Thus $x \overset{*}{\underset{R,f}{\longmapsto}}_{|w'|} u'w'$ and $v'w' \overset{*}{\underset{R,f}{\longmapsto}}_{|w'|} s = pt = phw' \overset{*}{\underset{R,f}{\longmapsto}}_{|w'|} qhw' = y$.

(iii) We now extend a polynomial construction [5] of a rational transducer recognizing the prefix derivation of a word rewriting system.

We will construct a graph $G \subseteq F \times E \times F$ on a finite set F of vertices and on a finite set E of labels. Let us define F and E.

We denote by

$$F_1 := Dom(R) \cup \{\{\varepsilon\}\},$$

the set of rational languages of the left hand sides of R, plus $\{\{\varepsilon\}\}$. We denote by

$$F_2 := \{N^*\} \cup \{P_1^{-1} f(U_1,V_1) \cap \cdots \cap P_n^{-1} f(U_n,V_n) \mid n \geqslant 0 \\ \wedge \forall 1 \leqslant i \leqslant n, P_i \subseteq N^* \wedge (U_i,V_i) \in R\},$$

the closure by intersection of the left quotients (by subset of N^*) of the contexts $Im(f)$ of R, plus the language N^*.

By Lemma 3.1 (c), F_2 is finite, and is closed by intersection and by left quotient. Thus the direct product $F := F_1 \times F_2$ of F_1 by F_2 is finite. By Lemma 3.1 (b), the set

$$[N^*]_{\mathrm{F}_2} := \{[y]_{\mathrm{L}} \mid y \in N^* \wedge L \in F_2\}$$

is a finite set of rational languages. Finally we define

$$E := \{P_1^{-1} L_1 \cap \cdots \cap P_n^{-1} L_n \mid n \geqslant 0 \wedge \forall 1 \leqslant i \leqslant n, \\ P_i \subseteq N^* \wedge L_i \in Im(R) \cup [N^*]_{\mathrm{F}_2}\},$$

the closure by intersection of the left quotients of $Im(R) \cup [N^*]_{F_2}$.
Like for F_2, E is finite and is closed by intersection and by left quotient.

We consider now the following graph in $F \times E \times F$:

$$H := \{(P,L) \xrightarrow[{[y]_L \cap P^{-1} V}]{} (U, y^{-1} L \cap f(U,V)) \mid (U,V) \in R \wedge (P,L) \in F \wedge y \in N^*\}$$

of prefix decompositions of the rules of R.

Given any (finite) graph K in $F \times E \times F$, we define a splitting $\langle K \rangle$ of K as follows:

$$\langle K \rangle := \{(P,L) \xrightarrow[{[y]_{\mathrm{L}} \cap (X^{-1}P)^{-1} Y}]{} (Q, y^{-1} L \cap M) \mid (\{\varepsilon\}, N^*) \overset{X}{\underset{K}{\Rightarrow}} \overset{Y}{\underset{K}{\rightarrow}} (Q,M) \\ \wedge (P,L) \in F \wedge y \in N^*\}.$$

Furthermore we define recursively a completion $\bar{K}$ of a graph K in $F \times E \times F$ as being the least graph containing K with $\langle \bar{K} \rangle = \emptyset$, i.e.

$$\bar{K} := \begin{cases} K & \text{if } \langle K \rangle = \emptyset, \\ \overline{K \cup \langle K \rangle} & \text{if } \langle K \rangle \neq \emptyset. \end{cases}$$

Finally we take the completion G of H:

$$G = \bar{H}.$$

We verify that G satisfies the following property $(*)$: for any $U \in Dom(R)$,

$$\exists V \ \exists v \in V, (U,V) \in R \wedge w \in f(U,V) \wedge vw \overset{*}{\underset{R,f}{\mapsto}}{}_{|w|} xw \\ \Leftrightarrow \exists W, (\{\varepsilon\}, N^*) \overset{x}{\underset{G}{\Rightarrow}}{}^{+} (U,W) \wedge w \in W.$$

(iv) To each $(U,M) \in F$, we associate the following rational language:

$$G(U,M) := \{x \mid (\{\varepsilon\}, N^*) \overset{x}{\underset{G}{\Rightarrow}}{}^{+} (U,M)\}.$$

Consider the inverse $R^{-1} := \{(V,U) \mid (U,V) \in R\}$ of R, and the mapping $\bar{f}$ from R^{-1} into $Rat(N^*)$ defined by $\bar{f}(V,U) := f(U,V)$ for every $(U,V) \in R$.

From $(R^{-1}, \bar{f})$ we construct as defined in (iii) a graph $\bar{G}$ in $\bar{F} \times \bar{E} \times \bar{F}$ (where for instance $\bar{F} := \bar{F}_1 \times \bar{F}_2$ with $\bar{F}_1 := Dom(R^{-1}) \cup \{\{\varepsilon\}\} = Im(R) \cup \{\{\varepsilon\}\}$).

We define the following rationally controlled recognizable relation (S,g):

$$S := \{\bar{G}(V,L) \times G(U,M) \mid (U,V) \in R \wedge M \in F_1 \wedge L \in \bar{F}_1\}$$

and $g(\bar{G}(V,L) \times G(U,M)) := L \cap M \cap f(U,V)$.

Let us prove that $\underset{S,g}{\longmapsto} = \overset{+}{\underset{R,f}{\longmapsto}}$.

$\subseteq$: Let $u \underset{S,g}{\longmapsto} v$.

There are $(X,Y) \in S$, $x \in X$, $y \in Y$, $z \in g(X,Y)$ such that $u = xz$ and $v = yz$.

There exist $(U,V) \in R$, $L \in \bar{F}_1$, $M \in F_1$ such that

$$X = \bar{G}(V,L), \quad Y = G(U,M) \text{ and } g(X,Y) = L \cap M \cap f(U,V).$$

Hence $(\{\varepsilon\}, N^*) \overset{x}{\underset{\bar{G}}{\Longrightarrow}}{}^{+} (V,L)$ and $(\{\varepsilon\}, N^*) \overset{y}{\underset{G}{\Longrightarrow}}{}^{+} (U,M)$.

As $z \in M$ and by the direction $\Leftarrow$ of (*) in (iii),

$$\exists V_1 \ \exists v_1 \in V_1, \quad (U,V_1) \in R \wedge z \in f(U,V_1) \wedge v_1 z \overset{*}{\underset{R,f}{\longmapsto}}{}_{|z|} yz.$$

As $z \in L$ and by the direction $\Leftarrow$ of (*) in (iii),

$$\exists U_1 \ \exists u_1 \in U_1, \quad (U_1,V) \in R \wedge z \in f(U_1,V) \wedge u_1 z \overset{*}{\underset{R^{-1},\bar{f}}{\longmapsto}}{}_{|z|} xz.$$

But $z \in f(U,V)$ and by (i), we have $U_1 = U$ and $V_1 = V$.

Furthermore $\underset{R^{-1},\bar{f}}{\longmapsto}{}_{|z|} = (\underset{R,f}{\longmapsto}{}_{|z|})^{-1}$ hence

$$u = xz \overset{*}{\underset{R,f}{\longmapsto}}{}_{|z|} u_1 z \underset{R,f}{\longmapsto}{}_{|z|} v_1 z \overset{*}{\underset{R,f}{\longmapsto}}{}_{|z|} yz = v.$$

$\supseteq$: Let $x \overset{+}{\underset{R,f}{\longmapsto}} y$.

By the direction $\Rightarrow$ of (ii), $\exists (U,V) \in R \ \exists u \in U \ \exists v \in V \ \exists w \in f(U,V)$,

$$x \overset{*}{\underset{R,f}{\longmapsto}}{}_{|w|} uw \wedge vw \overset{*}{\underset{R,f}{\longmapsto}}{}_{|w|} y.$$

So there are $\bar{x}, \bar{y}$ such that $x = \bar{x}w$ and $y = \bar{y}w$.

By the direction $\Rightarrow$ of (*) in (iii),

$$(\{\varepsilon\}, N^*) \overset{\bar{y}}{\underset{G}{\Longrightarrow}}{}^{+} (U,M) \text{ for some } M \in F_1 \text{ with } w \in M,$$

$$(\{\varepsilon\}, N^*) \overset{\bar{x}}{\underset{G}{\Longrightarrow}}{}^{+} (V,L) \text{ for some } L \in \bar{F}_1 \text{ with } w \in L.$$

Let $X = \bar{G}(V,L)$ and $Y = G(U,M)$.

Then $\bar{x} \in X$, $\bar{y} \in Y$ and $w \in L \cap M \cap f(U,V) = g(X,Y)$.

Thus $x = \bar{x}w \underset{S,g}{\mapsto} \bar{y}w = v$.

Note that to get the reflexive closure $\overset{*}{\underset{R,f}{\mapsto}}$ of the transitive closure of $\overset{+}{\underset{R,f}{\mapsto}} = \underset{S,g}{\mapsto}$, it suffices to add to S the rule $(\{\varepsilon\},\{\varepsilon\})$ with $g(\{\varepsilon\},\{\varepsilon\}) = N^*$. $\square$

The proof of Proposition 3.18 is rather long. But it can be strongly reduced by applying a rational marking before an inverse substitution, instead of doing a rational restriction after an inverse substitution [19]. The rational right closures of recognizable relations are also preserved by union and by composition, hence by inverse rational substitution.

Theorem 3.19. *The rational right closures of recognizable graphs are effectively preserved by inverse rational extended substitution on labels, and by rational restriction on vertices.*

Proof. (i) Closure by rational restriction on vertices.

Let U, V, W, L be subsets of N^*. We have

$$
\begin{aligned}
((U \xrightarrow{a} V).W)_{|L} &= \{uw \xrightarrow{a} vw \mid u \in U \wedge v \in V \wedge w \in W \wedge uw, vw \in L\} \\
&= \{uw \xrightarrow{a} vw \mid u \in U \wedge v \in V \wedge w \in W \cap u^{-1}L \cap v^{-1}L\} \\
&= \bigcup \{(u \xrightarrow{a} v).(W \cap u^{-1}L \cap v^{-1}L) \mid u \in U \wedge v \in V\} \\
&= \bigcup \{((U \cap [u]_L) \xrightarrow{a} (V \cap [v]_L)).(W \cap u^{-1}L \cap v^{-1}L) \mid u, v \in N^*\}
\end{aligned}
$$

Hence $((U \xrightarrow{a} V).W)_{|L}$ is a rational right closure of a recognizable graph when U, V, W, L are rational languages.

(ii) Closure by inverse rational extended substitution on labels.

Let G be a rational right closure of a recognizable graph, and let $h : T^* \rightarrow Rat(T \cup \bar{T})^*$ be a morphism. Note that

$$h^{-1}(G) = \bigcup_{a \in T} h_a^{-1}(G),$$

where $h_a : \{a\}^* \rightarrow Rat(T \cup \bar{T})^*$ is the morphism defined by $h_a(a) = h(a)$.

Let $a \in T$. We want to show that $h_a^{-1}(G)$ is effectively a rational right closure of a recognizable graph. By induction on the rational structure of $h(a)$ and by Proposition 3.18, it remains to prove that the rational right closures of recognizable relations are preserved by composition. By distributivity of the composition with respect to the union, it suffices to prove that

$$(U \rightarrow V).W \circ (X \rightarrow Y).Z \quad \text{where } U, V, W, X, Y, Z \in Rat(N^*)$$

is effectively a rational right closure of a recognizable relation.

Recall that for $u \in N^*$ and $L \subseteq N^*$,

$$_L[u] := \{v \mid Lv^{-1} = Lu^{-1}\}.$$

It is easy to verify that

$$\begin{aligned}&(U \to V).W \circ (X \to Y).Z \\ &\quad = \{(U \to Y(X^{-1}V \cap Zw^{-1})).(W \cap {}_Z[w]) \mid w \in W\} \\ &\qquad \cup \{(U(V^{-1}X \cap Wz^{-1}) \to Y).(Z \cap {}_W[z]) \mid z \in Z\}. \qquad \square\end{aligned}$$

Theorem 3.19 with Proposition 3.16 give Theorem 2.12. Let us indicate that for the rational right closure of any recognizable graph, the path language from a vertex to a vertex is context-free. We can extend this property for sets of vertices (not proved here in detail).

Proposition 3.20. *Let G be a rational right closure of a recognizable graph, we have*: $L(G,E,F) \in Alg(T^*)$ *for any* $(E,F) \in Rat(N^*) \times Alg(N^*) \cup Alg(N^*) \times Rat(N^*)$.

We deduce that the language accepted by any pushdown automaton with acceptance by any context-free set of final configurations, is context-free (Theorem 5.5).

Acknowledgements

Thanks are due to Wolfgang Thomas for his support, and in particular for a remark at the origin of Theorem 3.2 from which follows all this article.

References

[1] L. Boasson, M. Nivat, Centers of context-free languages, Internal Report LITP 84–44, 1984.
[2] R. Büchi, Regular canonical systems, Archiv für Mathematische Logik und Grundlagenforschung 6 (1964) 91–111 [reprinted in: S. Mac Lane, D. Siefkes (Eds.), The collected works of J. Richard Büchi, Springer, New York, 1990, pp. 317–337].
[3] D. Caucal, On the regular structure of prefix rewriting, CAAP 90, in: A. Arnold (Ed.), Lecture Notes in Computer Science, Vol. 431, Springer, Berlin, 1990, pp. 87–102 [a full version is in Theoret. Comput. Sci. 106 (1992) 61–86].
[4] D. Caucal, Monadic theory of term rewritings, 7th IEEE Symposium, LICS 92, 1992, pp. 266–273.
[5] D. Caucal, Bisimulation of context-free grammars and of pushdown automata, in: A. Ponse, M. de Rijke, Y. Venema (Eds.), CSLI Vol. 53, Modal Logic and Process Algebra, Stanford, 1995, pp. 85–106.
[6] B. Courcelle, Graph rewriting: an algebraic and logic approach, in: J. Leeuwen (Ed.), Handbook of Theoretical Computer Science, Vol. B, Elsevier, Amsterdam, 1990, pp. 193–242.
[7] B. Courcelle, Monadic second-order definable graph transductions: a survey, Theoret. Comput. Sci. 126 (1994) 53–75.
[8] B. Courcelle, I. Walukiewicz, Monadic second-order logic, graph coverings and unfolding of transition systems, Ann. Pure Appl. Logic 1 (1998) 35–62.
[9] D. Muller, P. Schupp, The theory of ends, pushdown automata, and second-order logic, Theoret. Comput. Sci. 37 (1985) 51–75.

[10] M. Rabin, A simple method for undecidability proofs and some applications, in: Y. Bar-Hillel (Ed.), Logic, Methodology and Philosophy of Science, North-Holland, 1965, pp. 58–68.
[11] M. Rabin, Decidability of second-order theories and automata on infinite trees, Trans. Amer. Math. Soc. 141 (1969) 1–35.
[12] A. Semenov, Decidability of monadic theories, MFCS 84, in: G. Goos, J. Hartmanis (Eds.), Lecture Notes in Computer Science, Vol. 176, Springer, Berlin, 1984, pp. 162–175.
[13] G. Sénizergues, Formal languages and word-rewriting, in: H. Comon, J.-P. Jouannaud (Eds.), French Sring School of Theoretical Computer Science, Lecture Notes in Computer Science, Vol. 909, Springer, Berlin, 1975, pp. 75–94.
[14] S. Shelah, The monadic theory of order, Ann. Math. 102 (1975) 379–419.
[15] J. Stupp, The lattice-model is recursive in the original model, Manuscript, The Hebrew University, 1975.
[16] W. Thomas, Automata on infinite objects, in: J. Leeuwen (Ed.), Handbook of Theoretical Computer Science, Vol. B, Elsevier, Amsterdam, 1990, pp. 135–191.
[17] W. Thomas, On logics, tilings, and automata, ICALP 91, in: L. Albert, B. Monien, R. Artalejo (Eds.), Lecture Notes in Computer Science, Vol. 510, Springer, Berlin, 1991, pp. 191–231.
[18] W. Thomas, Languages, automata, and logic, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, Vol. 3, Springer, Berlin, 1997, pp. 389–456.
[19] D. Caucal, On the transition graphs of Turing machines, MCU 2001, in: M. Margenstern, Y. Rogozhin (Eds.), Lecture Notes in Computer Sciences, Vol. 2055, pp. 177–189, a long version will appear in Theoret. Comput. Sci.