INRIA
INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

On rational graphs

Christophe MORVAN

N°3944

Mai 2000

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--3944--FR+ENG

On rational graphs

Christophe Morvan*

Thème 1 — Réseaux et systèmes
Projet

Rapport de recherche n° 3944 — Mai 2000 — 20 pages

Abstract: Using rationality, like in language theory, we define a family of infinite graphs. This family is a strict extension of the context-free graphs of Muller and Schupp, the equational graphs of Courcelle and the prefix recognizable graphs of Caucal. We give basic properties, as well as an internal and an external characterization of these graphs. We also show that their traces form an AFL of recursive languages, containing the context-free languages.

Key-words: rationality, trace, infinite graphs, rational graphs

(Résumé : tsvp)

This report is an extended version of a paper presented at Fossacs 2000; published in LNCS 1784 (Ed. Jerzy Tiuryn) pp. 252 - 266

* christophe.morvan@irisa.fr

Sur les graphes rationnels

Résumé : En utilisant la rationalité (introduite en théorie des langages), nous définissons une famille de graphes infinis. Cette famille est une extension stricte des graphes *context-free* de Müller et Schupp, des graphes équationnels de Courcelle et des graphes *préfixe-reconnaissables* de Caucal. Nous donnons les propriétés élémentaires de ces graphes, ainsi qu'une carractérisation interne et externe. Nous montrons enfin que les traces des ces graphes forment une famille abstraite de langages contenant les langages algébriques.

Mots-clé : rationalité, trace, graphes infinis, graphes rationnels

1 Introduction

When dealing with computers, infinite graphs are natural objects. They emerge naturally in recursive program schemes or communicating automata, for example. Studying them as families of objects is comparatively recent: Muller and Schupp (in [MS 85]) first captured the structure of the graphs of pushdown automata, then Courcelle (in [Co 90]) defined the set of regular (equational) graphs. More recently Caucal introduced (in [Ca 96]) a characterization of graphs in terms of inverse (rational) substitution from the complete binary tree. Step by step, like Chomsky's languages family, a hierarchy of graph families is built: the *graphs of pushdown automata*, *regular graphs* and *prefix recognizable graphs*.

To define infinite objects conveniently, we have to use finite systems. For infinite graphs, two kinds of finite systems are employed: internal systems or external systems. Roughly speaking an internal characterization is a *machine* producing the arcs of the graph. An external characterization yields the structure of the graph (usually "up to isomorphism"). There is, of course a relationship between internal and external characterization: for example the pushdown automata are an internal characterization of the connected regular graphs of finite degree whereas the *deterministic graph grammars* are an external system for the family of regular graphs.

The purpose of this article is to give both internal and external characterization of a wider family of graphs. Using words for vertices, *rationality* (like in language theory) will provide an internal characterization; it will also give basic results for this family: for example rational graphs will be recognized by transducers; a rational graph is a recursive set; determinism for rational graphs will be decidable. Then *inverse substitution* from the complete binary tree (like in [Ca 96]) will be an external characterization of this family. Strangely this extension will prove to be a slight extension of the *prefix recognizable graphs*: instead of taking the inverse image of the complete binary tree by a rational substitution we will consider the inverse image of the complete binary tree by a linear substitution (*i.e.,* a substitution where the image of each letter is a linear language). Finally properties of the traces of these graphs will be investigated: we will show that the traces of these graphs form an *abstract family of (recursive) languages* containing the context-free languages.

2 Rational graphs

In this section we will define a new family of infinite graphs, namely the set of *rational graphs*. We will state some results for this family and give examples of rational graphs.

2.1 Partial semigroups

This paragraph introduces rationality for partial semigroups and uses this notion to give a natural introduction for rational graphs.
We start by recalling some standards notations: for any set E, its cardinal is denoted by $|E|$; its powerset is denoted by 2^E. Let the set of nonnegative integers be denoted by $\mathbb{N}$. A semigroup S is a set equipped with an operation $\cdot : S \times S \to S$ such that: for all u, v in S there exists w in S such that $\cdot(u, v) = w$ denoted by $u \cdot v = w$ and this operation is associative (*i.e.,* $\forall u, v, w \in S, (u \cdot v) \cdot w = u \cdot (v \cdot w)$. Finally, a monoid M is a semigroup with a (unique) neutral element (denoted ε along these lines) *i.e.,* an element $\varepsilon \in M$ such that for all element u in M $u \cdot \varepsilon = \varepsilon \cdot u = u$.
Now, a *partial semigroup* is a set S equipped with a partial operation $\cdot : S \times S \to S$, with $\mathcal{D} \subseteq S \times S$ the domain of $\cdot$; set $\mathcal{D}$ need not be $S \times S$. Moreover we impose this operation to be *associative* as follows: $[(u, v) \in \mathcal{D} \wedge ((u \cdot v), w) \in \mathcal{D}] \ \Leftrightarrow [(v, w) \in \mathcal{D} \wedge (u, (v \cdot w)) \in \mathcal{D}]$ and in that case, $u \cdot (v \cdot w) = (u \cdot v) \cdot w$. Meaning that if multiplication is defined on the one side, then it is defined on the other side and both agree.
Notice that a partial semigroup S such that $\mathcal{D}$ is $S \times S$ is a semigroup.

Example 2.1. Given two semigroups $(S_1, \cdot_1)$ and $(S_2, \cdot_2)$ such that $S_1 \cap S_2$ is empty. The union $S = S_1 \cup S_2$, with the partial operation $\cdot$ defined as $\cdot_1$ over the elements of S_1 and $\cdot_2$ over the element of S_2, is a partial semigroup.

Taking a new element $\perp$ we complete any partial semigroup S into a semigroup $S \cup \{\perp\}$ by extending its operation $\cdot$ as follows:

$a \cdot b = \perp$ for all $a, b \in S \cup \{\perp\}$ such that $(a,b) \notin \mathcal{D}$.
Also the product $S \times S'$ of two partial semigroups S and S' is a partial semigroup for operation $\cdot$ defined componentwise:
$(a,a') \cdot (b,b') = (a \cdot b, a' \cdot b')$ for all $(a,b) \in \mathcal{D}$ and $(a',b') \in \mathcal{D}'$.
In order to define the rational subsets of a partial semigroup, we have to extend its operation to its subsets:

$$A \cdot B := \{a \cdot b \mid a \in A \wedge b \in B\} \text{ for every } A, B \subseteq S$$

The powerset 2^S of S, is a semigroup for $\cdot$ so defined.
Now, a subset P of a partial semigroup S is a *partial subsemigroup* of S, if P is a partial semigroup for $\cdot$ of S *i.e.,* $P \cdot P$ is a subset of P.
For any subset P of a partial semigroup S, following subset $P^+ = \bigcup_{n \geqslant 1} P^n$ (with $P^1 = P$ and $P^{n+1} = P^n \cdot P$ for every $n \geqslant 1$) is the smallest (for inclusion) partial subsemigroup of S containing P. Set P^+ is called the *partial semigroup generated* by P. In particular $(P^+)^+ = P^+$. Also, S is *finitely generated* if $S = P^+$ for some finite P.
A set $P \subseteq S$ is a *code* if there is no two factorization in P^+ of the same element:
$u_1 \cdots u_m = v_1 \cdots v_n \wedge u_1, \ldots, u_m, v_1, \ldots, v_n \in P \;\Rightarrow\; m = n \wedge \forall\, i \in [1 \cdots n], u_i = v_i$
A partial semigroup S is *free* if there is code P such that $P^+ = S$.
For every $W \subseteq 2^S$, we denote by $\bigcup W = \{a \mid \exists P \in W, a \in P\}$. Operator $+$ commutes with operator $\bigcup$, *i.e.,* $\bigcup(W^+) = (\bigcup W)^+$ for every $W \subseteq 2^S$.

Lemma 2.2. *We have* $\bigcup(W^+) = (\bigcup W)^+$ *for every* $W \subseteq 2^S$

Proof.

$$\begin{aligned}
u \in \bigcup(W^+) &\Leftrightarrow \exists P_1, \ldots, P_n \in W,\ u \in P_1 \cdot P_2 \cdots P_n \\
&\Leftrightarrow \exists P_1, \ldots, P_n \in W \exists u_1 \in P_1, \ldots, u_n \in P_n, u = u_1 \cdot u_2 \cdots u_n \\
&\Leftrightarrow \exists u_1, \ldots, u_n \in \bigcup W, u = u_1 \cdot u_2 \cdots u_n \\
&\Leftrightarrow u \in (\bigcup W)^+
\end{aligned}$$

□

The (left) *residual* $u^{-1}P$ of $P \subseteq S$ by $u \in S$ is following subset:
$u^{-1}P := \{v \in S \mid u \cdot v \in P\}$
and satisfies following basic equality:
$(u \cdot v)^{-1}P = v^{-1}(u^{-1}P)$ for all $u, v \in S$ and $P \subseteq S$.

Definition 2.3. Let $(S, \cdot)$ be a partial semigroup. The family $Rat(S)$ of rational subsets of S is the least family $\mathcal{R}$ of subsets of S satisfying the following conditions:

(i) $\emptyset \in \mathcal{R}$; $\{m\} \in \mathcal{R}$ for all m in S;

(ii) if $A, B \in \mathcal{R}$ then $A \cup B, A \cdot B$ and $A^+ \in \mathcal{R}$.

Remark: Partial associativity is necessary to the definition of A^n. It ensures that A^n is independent of the order of the multiplications. Indeed
In order to generalize well known results for monoids in the case of partial semigroups, and as our purpose is to deal with graphs, we will set some notations and definitions for graphs and automata.
Let P be a subset of S. A (simple oriented labelled) *P-graph* G over V with arcs labelled in P is a subset of $V \times P \times V$. An element (s,a,t) in G is an *arc* of *source* s, *goal* t and *label* a (s and t are *vertices* of G). We denote by $Dom(G)$, $Im(G)$ and V_G the sets respectively of sources, goals and vertices of G. Each (s,a,t) of G is identified with labelled transition $s \underset{G}{\xrightarrow{a}} t$ or simply $s \xrightarrow{a} t$ if G is understood.

A graph G is *deterministic* if distinct arcs with same source have distinct label: $r \xrightarrow{a} s \wedge r \xrightarrow{a} t \;\Rightarrow\; s = t$.
A graph is (source) *complete* if, for every label a, every vertex is source of an arc labelled a: $\forall a \in P$, $\forall s \in V_G$, $\exists t\ s \xrightarrow{a} t$. Set $2^{V \times P^+ \times V}$ of P^+-graphs with vertices in V is a semigroup for *composition relation*: $G \cdot H := \{r \xrightarrow{a \cdot b} t \mid \exists s, r \underset{G}{\xrightarrow{a}} s \wedge s \underset{H}{\xrightarrow{b}} t\}$ for any $G, H \;\subseteq V \times P^+ \times V$. Relation $\underset{G^+}{\xrightarrow{u}}$ denoted by $\underset{G}{\overset{u}{\Longrightarrow}}$ or

simply $\overset{u}{\Longrightarrow}$ if G is understood, is the existence of a *path* in G labelled u in P^+. For any L in S, we denote by $s \overset{L}{\Longrightarrow} t$ that there exists u in L such that $s \overset{u}{\Longrightarrow} t$.
The *trace* (or set of path labels) $L(G, E, F)$ of G from a set E to a set F is the following subset of P^+:

$$L(G,E,F) \ := \ \{u \in S \mid \exists\, s \in E,\ \exists\, t \in F,\ s \underset{G}{\overset{u}{\Longrightarrow}} t\ \}$$

Given $P \subseteq S$, a P-*automaton* A is a P-graph G whose vertices are called *states*, with an *initial state* i and a subset F of *final states*; the automaton *recognizes* subset $L(A)$ of P^+: $L(A) := L(G, \{i\}, F)$. An automaton is finite (resp. deterministic, complete) if its graph is finite (resp. deterministic, complete). This allows to state a standard result for rational subsets.

Proposition 2.4. *Given a subset P of a partial semigroup S, $Rat(P^+)$ is*

(i) the smallest subset of 2^S containing $\emptyset$ and $\{a\}$ for each $a \in P$, and closed for $\cup, \cdot, +$

(ii) the set of subsets recognized by finite P-automata,

(iii) the set of subsets recognized by finite and deterministic P-automata.

Proof. Let us first denote by $Rat(P)$, $Reco(P^+)$, $Reco_{fd}(P^+)$ the sets (i), (ii) and (iii).
• $Rat(P) \subseteq Reco(P^+)$. Set $Reco(P^+)$ contains $\emptyset$ and $\{a\}$ for each $a \in P$. It is closed for $\cup, \cdot$ and $+$ for the same reasons as for monoids. Therefore, $Rat(P)$ which is the smallest subset of 2^S satisfying these properties is a subset of $Reco(P^+)$.
• $Reco(P^+) \subseteq Reco_{fd}(P^+)$. Consider $L \in Reco(P^+)$ and A a finite P-automaton such that $L(A) = L$. As for monoids, we only have to determinize (locally) G (the P-graph associated to A): let V be the set of vertices of G, we construct a graph G' in $2^V \times P \times 2^V$ that will be deterministic. The automaton A' will have G' for graph, state $\{i\}$ (the subset of V reduced to i) as initial state and as set of final states, set $F' := \{Q \mid Q \subseteq V \ \wedge \ Q \cap F \neq \emptyset\}$.
• $Reco_{fd}(P^+) \subseteq Rat(P^+)$. Now, using an induction over the number of states of the automata and lemma 2.2 we show that each element of $Reco_{fd}(P^+)$ is a *finite* union and $+$, of the singletons of P, and thus an element of $Rat(P^+)$.
• Now, $Rat(P) = Reco(P^+) = Reco_{fd}(P^+)$, it remains to show that these sets are $Rat(P^+)$.
By definition $Rat(P) \subseteq Rat(P^+)$, then by induction on the structure of $Rat(P^+)$: singletons of P^+ are concatenation of singletons of P and are therefore in $Rat(P)$, therefore by minimality of $Rat(P^+)$ we have $Rat(P^+) \subseteq Rat(P)$, which concludes this proof. ☐

Let us now consider another family of subsets of a partial semigroup S: the *recognizable subsets* of S.

Definition 2.5. A *partial semigroup morphism* $\varphi : S_1 \to S_2$ where $(S_1, \cdot_1)$ and $(S_2, \cdot_2)$ are partial semigroups (with $\mathcal{D}_i$ domain of $\cdot_i$), is an application such that, for all u, v in S_1: if $(u, v) \in \mathcal{D}_1$ then $(\varphi(u), \varphi(v)) \in \mathcal{D}_2$ and $\varphi(u) \cdot_2 \varphi(v) = \varphi(u \cdot_1 v)$

This definition is a natural extension of semigroups morphisms: if we consider two semigroups and a partial semigroup morphism between them, this morphism is a semigroup morphism.

Definition 2.6. A subset $P \subseteq S$ (S a partial semigroup) is *recognizable* if there exist a finite partial semigroup S', a partial semigroup morphism $\varphi : S \to S'$ and a subset $P' \subseteq S'$ such that $P = \varphi^{-1}(P')$.

Remark: the set of recognizable subsets of a partial semigroup S is denoted by $Rec(S)$. Also, notice that the image of the morphism could, as well, be a finite monoid; in that case the recognizable subsets remain the same.
Now, a graph G is said *path-deterministic* if graph G^+ is deterministic, *i.e.*, if $\underset{G}{\overset{u}{\Longrightarrow}}$ is a function for every u in P^+: $r \overset{u}{\Longrightarrow} s \wedge r \overset{u}{\Longrightarrow} t \ \Rightarrow \ s = t$.

When a graph G is path-deterministic and complete, we denote by su the unique state such that $s \overset{u}{\Longrightarrow} su$; in particular $s(u \cdot v) = (su)v$.
This allows the use of P-automata to characterize recognizable subsets of P^+.

Proposition 2.7. *Given a subset P of a partial semigroup, $Rec(P^+)$ is the set of subsets recognized by the complete path-deterministic P-automata having a finite set of states.*

Proof. Let us first denote by $Reco_{cdd}(P^+)$ the set of subsets of P^+ recognized by a complete path-deterministic P-automata having a finite set of states. We have to prove that $Reco_{cdd}(P^+) = Rec(P^+)$.

• $Reco_{cdd}(P^+) \subseteq Rec(P^+)$. Let us consider $L \in Reco_{cdd}(P^+)$, A a P-automaton such that $L = L(A)$, G be the associated graph and V the set of its vertices. As $|V|$ is finite, the set of fonctions from V to V (denoted V^V) is a finite monoid for composition. We define now a function $h : P^+ \rightarrow V^V$, such that to every u in P is associated the function $\overline{u}$ such that $\overline{u}(p) = pu$ for all p in V. This function is well defined because A is path-deterministic. This function is a partial semigroup morphism because the automaton is complete: [[inutile ? let u and v be two elements of P^+ such that $u \cdot v$ is defined. $u \cdot v \in P^+$ therefore $h(u \cdot v)$ is defined. Also, $h(u)$ and $h(v)$ are defined, for every state $p \in V$, $h(u)(p)$ is defined (A is complete), it is a unique state q of V, as A is complete, $h(v)(q)$ is defined, it is a unique state r of V. This defines a path from p to r labelled $u \cdot v$ and therefore, as A is path-deterministic, $h(u \cdot v)(p) = r = h(u) \circ h(v)(p)$. So h is a partial semigroup morphism,]] now consider finite subset of V^V $P = \{\overline{u} \mid \overline{u}(i) \in F\}$, set L is equal to $h^{-1}(P)$ and therefore recognizable.

• $Rec(P^+) \subseteq Reco_{cdd}(P^+)$. Consider $L \in Rec(P^+)$ and say $L = h^{-1}(N)$ for h a partial semigroup morphism and $N \subseteq S$ with $(S, \cdot_S)$ a finite partial semigroup (and $\mathcal{D}_S$ the domain of $\cdot_S$). Consider set $M := S \cup \{\perp, \varepsilon\}$ with operation $\cdot$ defined as follows:
$a \cdot b = \perp \;\; \forall a, b \in S \cup \{\perp\}, (a, b) \notin \mathcal{D}_S$ and $a \cdot \varepsilon = \varepsilon \cdot a = a \;\; \forall a \in S \cup \{\perp\}$

$(M, \cdot)$ is a partial semigroup. Morphism h can be extended to M (composed with canonical injection of S into M). We still have $L = h^{-1}(N)$, with $N \subseteq M$. We construct the automaton A, using elements of M as vertices: $G := \{(\, p, u, p \cdot h(u)) \mid u \in P \wedge \; p \in M\}$, now $A := (G, \{\varepsilon\}, N)$.
First, does A satisfies the conditions of Proposition 2.7?
Automaton A has a finite set of states (M is a finite monoid), it is complete by construction. In order to establish that A is path-deterministic and it recognizes set L we will show that:
$s \underset{G}{\overset{u}{\Longrightarrow}} t \Rightarrow t = s \cdot h(u)$
By induction on $n \geqslant 1$ for $u \in P^n$.
(Basic) n=1. by construction of G.
(Induction) Suppose for all $i \leqslant n, (u \in P^i)$: $s \underset{G}{\overset{u}{\Longrightarrow}} t \Rightarrow t = s \cdot h(u)$. Now let $s \underset{t}{\overset{au}{\Longrightarrow}}$ with $a \in P$ and $u \in P^{n+1}$. There exists r such that $s \overset{a}{\longrightarrow} r \overset{u}{\Longrightarrow} t$. By basic and induction conditions, $r = s \cdot (a)$ and $t = r \cdot h(u)$. Therefore, $t = s \cdot h(a) \cdot h(u) = s \cdot h(a \cdot u)$.

Thus A is path-deterministic . And also A recognizes L:

$$
\begin{aligned}
L(A) &= \{u \in P^+ \mid \exists s \in L, \varepsilon \overset{u}{\Longrightarrow} h(s)\} \\
&= \{u \in P^+ \mid \exists s \in L, h(s) = h(u)\} \\
&= h^{-1}(h(L)) \\
&= L
\end{aligned}
$$

□

Remark: The hypothesis that the P-automaton is complete is necessary: the identity relation in a free monoid is not a recognizable relation but it can be realized by a P-automaton having a finite set of states and such that $\overset{u}{\Longrightarrow}$ is deterministic for each $u \in P^+$.

Now Propositions 2.4 and 2.7 allows the extension to partial semigroup of well-known results for monoids.

Proposition 2.8. *For any partial semigroups S and S', we have following properties:*

(i) for every $L_1 \in Rec(S)$ and $L_2 \in Rat(S)$, $L_1 \cap L_2 \in Rat(S)$;

(ii) $Rec(S)$ is a boolean algebra;

(iii) $Rec(S) \subseteq Rat(S)$ if S is finitely generated (Mc Knight theorem);

(iv) $Rec(S) = Rat(S)$ if S is finitely generated and free (Kleene theorem);

(v) $L \subseteq Rec(S) \Leftrightarrow \{u^{-1}L \mid u \in S\} finite;$

(vi) $L \in Rec(S \times S') \ \Leftrightarrow\ L = \bigcup_{i=1}^{n}(U_i \times V_i)$
for some $n \in \mathbb{N}$ *and for all* i $U_i \in Rec(S),\ V_i \in Rec(S')$ *(Mezei theorem).*

Proof. **(i).** So let A_1 and A_2 be two S-automata.
Say $A_1 = (G_1, i_1, F_1), A_2 = (G_2, i_2, F_2), L_1 = L(A_1)$ and $L_2 = L(A_2)$. We define $G_1 \times G_2$ to be equal to
$\{(s_1, s_2) \xrightarrow{u} (t_1, t_2) \mid s_1 \xrightarrow[G_1]{u} t_1 \ \wedge\ s_2 \xrightarrow[G_2]{u} t_2\}$.
If G_1 and G_2 are deterministic (resp. path-deterministic, complete), then $G_1 \times G_2$ is deterministic (resp. path-deterministic, complete).
Furthermore if $(i_1, i_2) \underset{G_1 \times G_2}{\overset{u}{\Longrightarrow}} (t_1, t_2)$ then $i_1 \underset{G_1}{\overset{u}{\Longrightarrow}} t_1$ and $i_2 \underset{G_1}{\overset{u}{\Longrightarrow}} t_2$. Therefore we have,
$L(G_1 \times G_2, \{(i_1, i_2)\}, F_1 \times F_2) \subseteq L(A_1) \cap L(A_2)$
$[\ (i_1, i_2) \underset{G_1 \times G_2}{\overset{u}{\Longrightarrow}} (t_1, t_2) \in F_1 \times F_2\ ] \Rightarrow [i_1 \underset{G_1}{\overset{u}{\Longrightarrow}} t_1 \ \wedge\ i_2 \underset{G_1}{\overset{u}{\Longrightarrow}} t_2] \Leftrightarrow [u \in G_1 \ \wedge u \in G_2]$ Conversly, suppose that one of the graph is complet and path-deterministic (for example G_1) and that $u \in L(A_1) \cap L(A_2)$, there is a path in G_2 leading from i_2 to a vertex p in F_2 (labelled $u_1, u_2, \ldots, u_n$ elements of S). As G_1 is complete there is a path from i_1 labelled $u_1, u_2, \ldots, u_n$ leading to a state q, this path is labelled u, thus as G_1 is path-deterministic and $u \in L_1$ $q \in F_1$. Thus there is a path from (i_1, i_2) labelled $u_1, u_2, \ldots, u_n$ leading to $(q, p) \in F_1 \times F_2$. Therefore $L(A_1) \cap L(A_2) \subseteq L(G_1 \times G_2, \{(i_1, i_2)\}, F_1 \times F_2)$, this allows to conclude:
$L(A_1) \cap L(A_2) = L(G_1 \times G_2, \{(i_1, i_2)\}, F_1 \times F_2)$
when either A_1 or A_2 is path-deterministic and complete. This proves that (i) is true, and also that $Rec(S)$ is closed under intersection.
(ii). We have already that $Rec(S)$ is closed under intersection. It remain to show that $\emptyset$ belongs to $Rec(S)$ and that it is closed under complementation. Obviously $\emptyset \in Rec(S)$ furthermore for any path-deterministic and complete S-graph G, we have:
$L(G, i, F) + L(G, i, V_G - F) = M$
So $Rec(S)$ is closed under complementation.
(iii)-(iv). Consider S finitely generated: $M = P^+$ for some finite $P \subseteq S$. Therefore a P-automata with a finite set of states is finite. Hence, by propositions 2.4 and 2.7 $Rec(P^+) \subseteq Rat(P^+)$, *i.e.,* $Rec(S) \subseteq Rat(S)$.
In order to prove that $Rat(S) \subseteq Rec(S)$, it suffices to show that, when P is a code, any deterministic P-graph G is path-deterministic. Suppose that $r \underset{G}{\overset{u}{\Longrightarrow}} s$ and $r \underset{G}{\overset{u}{\Longrightarrow}} t$. We have $r \xrightarrow[G]{u_1} \cdots \xrightarrow[G]{u_m} s$ and $r \xrightarrow[G]{v_1} \cdots \xrightarrow[G]{v_n} t$, with $u_1 \cdots u_m = v_1 \cdots v_n$. As P is a code, $n = m$ and $u_1 = v_1, \ldots, u_n = v_n$, hence, as G is deterministic $s = t$.
(v). ($\Rightarrow$) Let $P \in Rec(S)$: $P = L(G, i, F)$ with G a path-deterministic and complete S-graph (and V_G finite). For each $u, v \in S$ we have:
$u^{-1}L(G, i, F) = \{v \mid u \cdot v \in L(G, i, F)\} = \{v \mid i(u \cdot v) \in F\} = L(G, iu, F)$
Therefore, $u^{-1}L(G, i, F) = L(G, s, F)$ for $i \overset{u}{\Longrightarrow} s$. Hence $\{u^{-1}L \mid u \in S\} \subseteq \{L(G, s, F) \mid s \in V_G\}$ is finite.

($\Leftarrow$) Let $P \subseteq S$ such that $\{u^{-1}P \mid u \in S\}$ is finite. Now we use a new symbol ε (not in S) we complete S into a partial semigroup $S \cup \{\varepsilon\}$ by extending its operation $\cdot$ as follows:
$a \cdot \varepsilon = \varepsilon \cdot a = \varepsilon$ for all $a \in S$.
For all $u \in S$, we define set $\overline{u}$ to be $u^{-1}P$ if $u \notin P$ and $u^{-1}P \cup \{\varepsilon\}$ if $u \in P$. We set $\overline{\varepsilon}$ to be P. This allows the definition of following S-graph:
$G = \{\overline{u} \xrightarrow{u} \overline{u \cdot v} \mid u \in S \cup \{\varepsilon\} \wedge v \in S\}$
V_G is finite, G is complete and path-deterministic: $\overline{u} \overset{v}{\Longrightarrow} \overline{u \cdot v}$. Furthermore the following is true:

$v \in L(G, \overline{\varepsilon}, \{\overline{u} \mid \varepsilon \in \overline{u}\}) \Leftrightarrow \exists u \in S, \varepsilon \in \overline{u} \wedge \overline{\varepsilon} \overset{v}{\Longrightarrow} \overline{u} \Leftrightarrow \varepsilon \in \overline{v} \Leftrightarrow v \in P$

thus $P = L(G, \overline{\varepsilon}, \{\overline{u} \mid \varepsilon \in \overline{u}\}) \in Rec(S)$
(vi).($\Leftarrow$) by (ii) *Rec* is closed by union. Therefore it only remains to show that given P in $Rec(S)$ and P' in $Rec(S')$, $P \times P' \in Rec(S \times S')$.
Now let M and M' be finite partial semigroup; φ and φ' two partial semigroup morphisms such that $P = \varphi^{-1}(N)$ and $P' = \varphi^{-1}(N')$ for $N \subseteq M,\ N' \subseteq M'$. Set $M \times M'$ has a natural partial semigroup structure, we define partial semigroup morphism h as follows:

$h : S \times S' \to M \times M'$ such that $h(u,v) = (\varphi(u), \varphi'(v))$
We have $h^{-1}(N \times N') = (h^{-1}(N) \times h^{-1}(N'))$, thus $P \times P' = h^{-1}(N \times N') \in Rec(S \times S')$.
($\Rightarrow$). Finally, let P be in $Rec(S)$ and $P = L(G,i,F)$ for some path-deterministic and complete S-graph G (with a finite vertex set. By (v), following equivalence on S:
$u \equiv_S v$ if $u^{-1}P = v^{-1}P$, is of finite index. We denote by $[u]_S$ the class of element u. Now $[u]_S = \bigcup\{L(G,i,iv) \mid u \equiv_S v\}$, hence $[u]_S \in Rec(S)$ for all u in S.
Let $R \in Rec(S \times S')$. By (v), R has a finite set E of left residuals:
$E = \{(u,u')^{-1}R \mid u \in S \wedge u' \in S'\}$.
Therefore $\{u^{-1}Dom(R) \mid u \in S\} = \pi_1(E)$ and $\{u'^{-1}Im(R) \mid u' \in S'\} = \pi_2(E)$ (where $Dom(R)$ (resp. $Im(R)$) denotes the set of left (resp. right) hand side elements of relation R) are finite sets. Thus $Dom(R)$ (resp. $Im(R)$) is a recognizable subset of S (resp. S'). So
$R = \bigcup\{[u]_S \times [u']_{S'} \mid (u,u') \in R\}$
is a finite union and each factor is a product of recognizable subsets. □

We simply translated the standards definitions of recognizable and rational subsets of monoids given for example in [Be 79]. An interesting example of a partial semigroup is the subject of these lines: the set of arcs (labelled with an element of a finite set) between elements of a free monoid is a partial semigroup; its rational subsets are the rational graphs.

2.2 Partial semigroups and graphs

In this section, we will consider an important example of partial semigroup: the set of rational graphs. So consider an arbitrary finite set X and denote X^* its associated free monoid. We will consider graphs as subsets of $X^* \times \mathcal{A} \times X^*$ (the set of graphs over X^* with arcs labelled in $\mathcal{A}$). For convenience, set $2^{X^* \times \mathcal{A} \times X^*}$ is denoted $G_{\mathcal{A}}(X^*)$.
Now, with $(u, a_i, v) \cdot_i (u', a_i, v') = (u \cdot u', a_i, v \cdot v')$, set $X^* \times \{a_i\} \times X^*$ (a_i in $\mathcal{A}$) is a monoid. As stated in Example 2.1 the union of these monoids (namely $X^* \times \mathcal{A} \times X^*$) is a partial semigroup. We denote by $\cdot$ the operation in $X^* \times \mathcal{A} \times X^*$ (which is $\cdot_i$ for each $X^* \times \{a_i\} \times X^*$).
Remark: this $\cdot$ operation for graphs is indeed, similar to the synchronization product for transition systems defined by Nivat and Arnold in [AN 88].
We are now able to define the set of rational graphs.

Definition 2.9. The set of rational graphs, denoted $Rat(X^* \times \mathcal{A} \times X^*)$ is the family of rational subsets of $X^* \times \mathcal{A} \times X^*$.

Let us now recall that a *transducer* is a finite automaton over pairs (see for example [Au 88] [Be 79]). A rational relation (*i.e.*, a rational subset of $X^* \times X^*$) is recognized by a rational transducer.
There is a strong relationship between rational graphs and rational relations and to characterize the family of rational graphs in a more practical way we will use *labelled transducers*.

Definition 2.10. A *labelled transducer* $T = \langle Q, I, F, E, L\rangle$ over X, is composed of a finite set of states Q, a set of initial states $I \subseteq Q$, a set of final states $F \subseteq Q$, a finite set of transitions (or edges) $E \subseteq Q \times X^* \times X^* \times Q$ and an application L from F into $2^{\mathcal{A}}$.

Like for P-graphs, transition (p,u,v,q) of transducer T will be denoted by $p \xrightarrow[T]{u/v} q$ or simply $p \xrightarrow{u/v} q$ if T is understood. Now similarly an element $(u,d,v) \in X^* \times \mathcal{A} \times X^*$ is *recognized* by transducer T if there is a path $p_0 \xrightarrow[T]{u_1/v_1} p_1 \cdots p_{n-1} \xrightarrow[T]{u_n/v_n} p_n$ and $p_0 \in I$, $p_n \in F$, $u = u_1 \cdots u_n$, $v = v_1 \cdots v_n$ and $d \in L(p_n)$.
Remark: an illustration of transducer execution will be given in Example 2.12.

Proposition 2.11. *A graph G in $G_{\mathcal{A}}(X^*)$ is rational if and only if it satisfies one of the following equivalent properties:*

(i) G belongs to the smallest subset of $G_{\mathcal{A}}(X^*)$ containing:
$\emptyset$,$\{\varepsilon \xrightarrow{d} \varepsilon\}$,$\{x \xrightarrow{d} \varepsilon\}$ *and* $\{\varepsilon \xrightarrow{d} x\}$, *for all $x \in X$, all $d \in \mathcal{A}$, and closed under $\cup$, $\cdot$ and $+$;*

(ii) G is a finite union of rational relations over each letter:
$G = \bigcup_{d \in \mathcal{A}} R_d$, *for $R_d \in Rat(X^* \times \{d\} \times X^*)$;*

(iii) G is recognized by labelled rational transducer.

Proof. **(i)** This is a direct consequence of Proposition 2.4, (i):
$\{\varepsilon \xrightarrow{d} \varepsilon, x \xrightarrow{d} \varepsilon, \varepsilon \xrightarrow{d} x \mid x \in X \wedge d \in \mathcal{A}\}^+ = X^* \times \mathcal{A} \times X^*$
(ii) If G is a finite union of rational sets, it is then a rational set by definition. Conversely, if G is in $Rat(X^* \times \mathcal{A} \times X^*)$, then set $G \cap X^* \times \{d\} \times X^*$ is in $Rat(X^* \times \{d\} \times X^*)$, according to Proposition 2.8, because $X^* \times \{d\} \times X^*$ is in $Rec(X^* \times \mathcal{A} \times X^*)$.
(iii) This is a direct consequence of (ii), and Proposition 2.4 (ii). □

Proposition 2.11 states that for any graph G in $Rat(X^* \times \mathcal{A} \times X^*)$, relation: $\xrightarrow[G]{d} := \{(u,v) \mid u \xrightarrow[G]{d} v\}$ is rational for each d in $\mathcal{A}$. Therefore we also introduce $\xrightarrow[G]{} := \bigcup_{d \in \mathcal{A}} \xrightarrow[G]{d}$, which is also a rational relation. Naturally we denote by $\xrightarrow[G]{d}(u)$ (resp. $\xrightarrow[G]{}(u)$) the image of word u by relation $\xrightarrow[G]{d}$ (resp. $\xrightarrow[G]{}$) (and similarly for subsets of X). Also for a rational graph G there are possibly many transducers generating it, thus we will denote by $\Theta(G)$ the set of transducers generating G.

We will now give some examples of rational graphs.

Example 2.12. The graph in Figure 2.1 is called the grid. This is a rational graph generated by the transducer on Figure 2.2. Its second order monadic theory is undecidable hence rational graphs have an undecidable second order monadic theory.

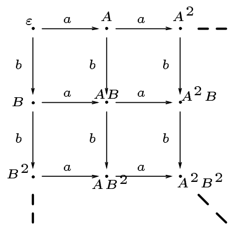


Figure 2.1: The grid, with vertices in $\{A, B\}^*$

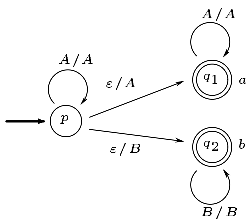


Figure 2.2: Transducer recognizing the grid

Why does the arc (AB, b, AB^2) belong to the graph? Simply because the following path is in the transducer:

$$p \xrightarrow{A/A} p \xrightarrow{\varepsilon/B} q_2 \xrightarrow{B/B} q_2$$

and that b is associated to the final state q_2.

Example 2.13. The graph in Figure 2.3 is another example of rational graph.

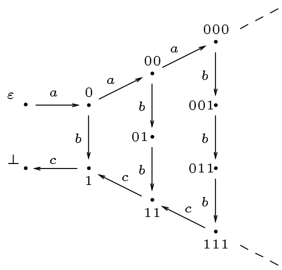


Figure 2.3: The graph $a^n b^n c^n$

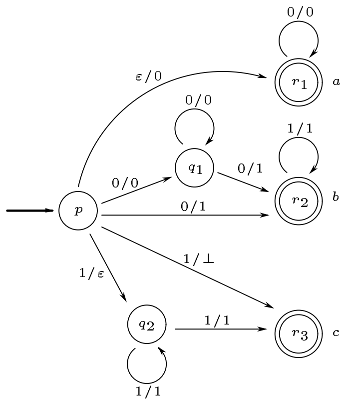
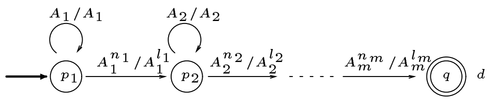


Figure 2.4: Transducer recognizing the graph $a^n b^n c^n$

We finish with a last example showing that the transition graphs of *Petri nets* are rational graphs.

Example 2.14. For more detail on Petri nets the reader may refer to [Re 85]. A Petri net can be seen as a finite set of transitions of this form: $A_1^{n_1} A_2^{n_2} \cdots A_m^{n_m} \xrightarrow{d} A_1^{l_1} A_2^{l_2} \cdots A_m^{l_m}$, with A_i^x representing there are x coins in place A_i (d represents the label (if any) of the transition). Following transducer generates the transition graph associated to the above transition:



Each vertex of the generated graph correspond to a marking of the Petri net. Each arc of the graph represents that a transition has been fired.

2.3 Some results for rational graphs

This section will introduce results for this family of graphs. Some of these results are just a reformulation of known results over rational relations. Others are simple facts on these graphs and their boundary.
The first fact is that this family is an extension of previous families. Simply recall that every *prefix recognizable graph* (defined in [Ca 96]) is a finite union of graphs of the following form :

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

with U, V, W rational sets.
This characterization ensures that *prefix recognizable graphs* are rational graphs. As the regular graphs (defined in [Co 90]) are *prefix recognizable graphs*, they are rational too. Furthermore, the graphs in Examples 2.12 and 2.13 are not *prefix recognizable graphs* thus the inclusion is strict. All this justifies the study of rational graphs.

Let us now translate some well-known results for rational relations, to rational graphs.

Proposition 2.15. *A rational graph G is of finite out-degree if and only if there exists a transducer $T \in \Theta(G)$ such that there exists no cycle in T labelled on the left with the empty word which is not labelled on the right with the empty word. In other words the only cycles labelled on the left ε, are labelled on the right ε.*

Remark: naturally this proposition can be translated to characterize the graphs of finite in-degree, by simply replacing right by left and vice-versa.

Proof. This proposition is a straightforward consequence of Proposition 5.3.a, p40 from [Au 88]. Actually, it is simply replacing the words “finite image” (for relations) with “finite out-degree” (for graphs). $\square$

Proposition 2.16. *Every rational graph is recursive: it is decidable whether an arc (u, d, v) belongs to a rational graph.*

Proof. Given a rational graph G, language $\xrightarrow[G]{d}(u)$ is effectively a rational language; to decide whether (u, d, v) is an arc of G we only have to check if v belongs to $\xrightarrow[G]{d}(u)$. $\square$

Theorem 2.17. *It is decidable whether a rational graph is deterministic (from its transducer).*

Proof. Again this is a simple translation of a rational relation result. This result from Schützenberger stated in [Au 88] (theorem 3.5, p 28) proves the decidability of functionality for rational relations, and of course a rational graph is deterministic if and only if the associated relation on each letter is functional. $\square$

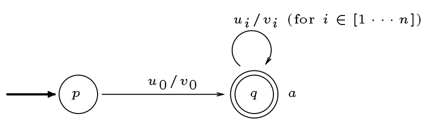
Proposition 2.18. *The inclusion and equality of deterministic rational graphs is decidable.*

Proof. See [Be 79], corollary 1.3, page 95. $\square$

Remark: unfortunately this result ceases to be true for general rational graphs ([Be 79] Theorem 8.4, page 90).
We have already seen that the second order monadic theory of these graphs is undecidable in general. We will now see that it is also the case for the first order theory.

Proposition 2.19. *The first order theory of rational graphs is undecidable.*

Proof. We will prove this proposition by reducing Post's correspondence problem (P.C.P.) to this problem. Let us recall the P.C.P.: given an alphabet X and (u_0, v_0), (u_1, v_1),..., (u_n, v_n) elements of $X^* \times X^*$. Does there exist a sequence $0 \leqslant i_1, i_2, \ldots, i_m \leqslant n$, such that $u_0 u_{i_1} \cdots u_{i_m} = v_0 v_{i_1} \cdots v_{i_m}$? To an instance of P.C.P. (i.e. a family (u_i, v_i)) we associate following transducer:



The resolution of P.C.P. becomes finding a vertex s such that $s \xrightarrow{a} s$ is an arc of the graph generated by the transducer. It is a first order instance, therefore,as P.C.P. is undecidable, the first order theory of rational graphs is not decidable in general. $\square$

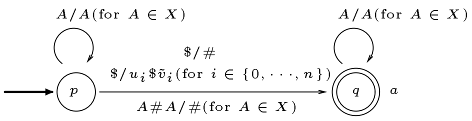
Before giving another negative decision result, let us denote by $\tilde{u}$ the mirror of word u (defined by induction on the length of u: $\tilde{\varepsilon} = \varepsilon$ and $\widetilde{au} = \tilde{u}a$ (for any u with $|u| \geqslant 0$).

Proposition 2.20. *Accessibility is not decidable for rational graphs in general.*

Proof. Once again, we use P.C.P. Using the same notations as earlier define a (word) rewriting system G, using two new symbols $\#$ and $\$$, in the following way:

$$G \begin{cases} \$ & \longrightarrow & u_i \$ \widetilde{v_i} & \forall i \in \{0, \cdots, n\} \\ \$ & \longrightarrow & \# & \\ A\#A & \longrightarrow & \# & \forall A \in X \end{cases}$$

Now "P.C.P. has a solution" is equivalent to the existence of a derivation from $u_0 \$ \widetilde{v_0}$ to $\#$. But, considering the following transducer:



the question becomes: is there a path leading from $u_0 \$ \widetilde{v_0}$ to the vertex $\#$? Answering the last question would allow P.C.P. to be solved in the general case which is a contradiction. Therefore accessibility is undecidable for the rational graphs in general. $\square$

Remark: the transitive closure of a rational graph is, at least, uneffective. If this construction were effective and rational, then accessibility for rational graph would be decidable.
Now we will see a case where accessibility is decidable for rational graphs. A transducer T is *increasing* if every pair (u, v) recognized by T is such that the length of v (denoted by $|v|$) is greater or equal to the length of u : $|v| \geqslant |u|$.

Proposition 2.21. *The accessibility is decidable for any rational graph with an increasing transducer.*

Proof. Let us denote by $T^{\leqslant n}(u)$ following set: $T^{\leqslant n}(u) := \bigcup_{i=0}^{n} T^i(u)$. For all $n \in \mathbb{N}$ this set is rational. Now, let G be a rational graph generated by an increasing transducer T and let u and v be two vertices of G. Let us put $n_0 = |\{w \in X^* \mid |u| \leqslant |w| \leqslant |v|\}| = |X|^{|u|} + \cdots + |X|^{|v|}$. Vertex v is accessible from u if and only if v belongs to $T^{\leqslant n_0}(u)$. Thus accessibility is decidable for rational graphs with an increasing transducer. $\square$

Before stating a technical lemma, recall that $s \overset{u/v}{\Longrightarrow} t$ denotes a path in a transducer from state s to state t labelled u/v. And denote by $s \underset{\neq}{\overset{u/v}{\Longrightarrow}} t$ an elementary path from s to t, *i.e.,* a path with no cycle: no vertex of the path (except, maybe, s and t) occur more than once.

Lemma 2.22. *Let G be a rational graph of finite out-degree. There exists two integers p and q such that for every $(s, a, t) \in G$ we have $|t| \leqslant p.|s| + q$*

Proof. Given T in $\Theta(G)$ (the set of transducers recognizing G), we first remove every transition $s \overset{\varepsilon/\varepsilon}{\longrightarrow} s$ from T. It will still recognize G. Given any path $s \underset{\neq}{\overset{u/v}{\Longrightarrow}} s \in T$, Proposition 2.15 ensures that $|u| \neq 0$, therefore one can define p_0 as follows:

$$p_0 := \max \left\{ \left\lceil \frac{|v|}{|u|} \right\rceil \mid s \underset{\neq}{\overset{u/v}{\Longrightarrow}} s \in T \right\}$$

Also, one can define q in the following way.

$$q := \max\{|v| - |u| \mid s \underset{\neq}{\overset{u/v}{\Longrightarrow}} t \wedge s \in I \wedge t \in F\}$$

Now consider (u, d, v) an arc of G. There exists a path in T labelled u/v. This path can be decomposed into simple cycles on the one hand and into transitions that are not in any cycle on the other hand (notice that a cycle can be "inside" another cycle; it can be moved "outside" because length is the only parameter that matters). Now consider the $u_i/v_i (i \in I)$ labelling the cycles. For all i we have : $|v_i| \leqslant p_0.|u_i|$, therefore we get

$$\sum_{i\in I}|v_i| \leqslant p_0. \sum_{i\in I}|u_i|$$

And, as the u_i's are disjoint factors of u

$$\sum_{i\in I}|v_i| \leqslant p_0.|u|$$

Now the transitions that are not in any cycles are labelled $w_j/z_j (j \in J)$ and we have $\sum_{j\in J}|z_j| \leqslant q + |u|$. Therefore, as $|v| = \sum_{i\in I}|v_i| + \sum_{j\in J}|z_j|$ we get

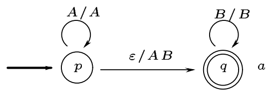
$$|v| \leqslant (p_0 + 1).|u| + q$$

To get the inequality from the lemma we simply have to put $p = p_0 + 1$. □

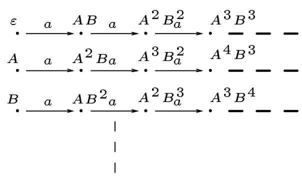
This lemma allows the construction of a graph that is not *structurally* rational.

Example 2.23. Consider an infinite tree in $X^*{\times}\mathcal{A}{\times}X^*$ such that every vertex of depth n has $2^{2^{2^n}}$ sons. This tree is not *strucurally* rational, in other words whatever name you give to the vertices of triplex it is never a rational graph. This is a direct consequence of previous lemma: say n is the length of the root, there are at most $|\mathcal{A}|^{(np^l+p^{l-1}q+\cdots+q)}$ vertices of depth l.

Despite these results the transducers are not able to capture the structure of rational graphs. For example, this transducer:



generates following graph:



The connected component of the empty word, ε, is a straight-line. It is "up to isomorphism" obviously rational, but as a sub-graph of this graph, it is not rational (its vertices form a context-free language). Therefore we need an external ("up to isomorphism") characterization of these graphs. This is the subject of the next section.

3 An external characterization

In this section, we will characterize rational graphs using inverse linear substitutions. Labelled transducers are an internal representation of rational graphs, it clearly depends on the name of the vertices. But often in graph theory, the name of the vertices is not relevant, it carries no information. An external characterization, like the graph grammars for equational graphs, produces graphs without giving names for vertices. It only gives the *structure* of the graph. Inverse linear substitution is an external characterization of rational graphs.

3.1 Graph isomorphism

An external characterization of rational graphs is given "up to isomorphism".
Two graphs G_1 and G_2 in $G_{\mathcal{A}}(X^*)$ are *isomorphic*, if there is a bijection $\psi : V(G_1) \to V(G_2)$ such that: $s_1 \underset{G_1}{\xrightarrow{d}} s_2$(*i.e.,* $(s_1, d, s_2) \in G_1$) if and only if $\psi(s_1) \underset{G_2}{\xrightarrow{d}} \psi(s_2)$.
Two isomorphic graphs have the same structure: they are the same up to a renaming of the vertices.
Now let us consider the equivalence ($\equiv$) generated by graph isomorphism: we say that G_1 is equivalent to G_2 (denoted $G_1 \equiv G_2$) if G_1 and G_2 are isomorphic. This equivalence relation provides us with a partition of $G_{\mathcal{A}}(X^*)$ denoted $Graph_{\mathcal{A}} := G_{\mathcal{A}}(X^*)/\equiv$. This allows the introduction of the set of *structural rational graphs*:

$$GRat_{\mathcal{A}} := \{[G]_{\equiv} \in Graph_{\mathcal{A}} \mid G \in Rat(X^* \times \mathcal{A} \times X^*)\}$$

This set is the set of graphs that are isomorphic to some rational graph, *i.e.,* such that there exists a rational graph with the same structure.
Set $Graph_{\mathcal{A}}$ (and $GRat_{\mathcal{A}}$), that denotes the set of equivalence classes, does not depend on the choice of set X. In other words, we can choose X to be any two letters alphabet with no loss of generality.

Lemma 3.1. *For all subset X' (with at least two elements) of X and all class $[G]_{\equiv}$ of $Graph_{\mathcal{A}}$ ($= G_{\mathcal{A}}(X^*)/\equiv$) there exists G_0 in $G_{\mathcal{A}}(X'^*)$ such that $G_0 \in [G]_{\equiv}$.*

Proof. Consider a and b two letters of X' and $[G]_{\equiv}$ an element of $Graph_{\mathcal{A}}$. Now say $X = \{x_1, x_2, \ldots, x_n\}$, and define the following morphism: $\varphi : X^* \to X'^*$ such that, for all i: $\varphi(x_i) = a^i b$. By definition φ is injective. Furthermore, graph $\varphi(G) = \{(\varphi(u), d, \varphi(v)) | (u, d, v) \in G\}$ is in $G_{\mathcal{A}}(X'^*)$ and, by construction isomorphic to graph G. This is the desired graph. ☐

We now have to characterize the structure of $GRat_{\mathcal{A}}$. This is the goal of the next section.

3.2 Substitution

Recall the definition of the *prefix recognizable graphs* (family REC_{Rat}). This family has been defined as the set of graphs obtained from the complete binary tree by inverse rational substitution, followed by rational restriction. We will use the same process (actually a linear context-free substitution) to obtain the family of rational graphs.
A *substitution* over a free monoid X^* is a morphism $\varphi : \mathcal{A}^* \to 2^{X^*}$, which associates to each letter in $\mathcal{A}$ a language in X^*. Our purpose is to study graphs, starting from the complete binary tree (Λ) labelled $X = \{A, B\}$. To move by inverse arcs, we use a new alphabet : $\overline{X} = \{\overline{A}, \overline{B}\}$ and we say that $x \xrightarrow{\overline{A}} y$ if $y \xrightarrow{A} x$. Given a language L and two vertices x and y, recall that $x \underset{\Lambda}{\overset{L}{\Longrightarrow}} y \Leftrightarrow \exists u \in L, x \underset{\Lambda}{\overset{u}{\Longrightarrow}} y$. Now, given a substitution $\varphi : \mathcal{A}^* \to 2^{(X \cup \overline{X})^*}$, we can define the graph $\varphi^{-1}(\Lambda)$ in the following way:

$$\varphi^{-1}(\Lambda) = \{x \xrightarrow{d} y \mid d \in \mathcal{A} \wedge \ x \underset{\Lambda}{\overset{\varphi(d)}{\Longrightarrow}} y\}$$

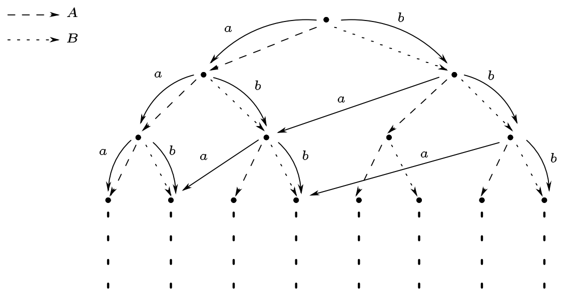
Given a language L, we define now $L_\Lambda = \{s \mid r \underset{\Lambda}{\overset{L}{\Longrightarrow}} s\}$. It allows us to consider the graph $\varphi^{-1}(\Lambda)_{|L_\Lambda}$: it is the image of the complete binary tree by an inverse substitution followed by a restriction; if L is rational, we say a rational restriction.

Example 3.2. Example 2.12 states that the grid is a rational graph. Following substitution: $h(a) = \{\overline{B}^m A B^m | \ m \geqslant 0\}$, $h(b) = \{B\}$ over the complete binary tree on $\{A, B\}$, followed with the restriction to $L = A^* B^*$ produces a graph isomorphic to the grid:

Now, it is well know that there is a close relationship between *linear languages* and rational relations (a linear language is a context-free language generated by a grammar with only, at most, one non-terminal on the right hand side of each rule). And indeed, if we denote the set of linear languages over the alphabet $X \cup \overline{X}$ by $Lin(X \cup \overline{X})$, we have the following proposition.

Proposition 3.3. *The set $GRat_{\mathcal{A}}$ is a subset of the family of the graphs obtained from the complete binary tree (Λ) by an inverse* linear *substitution, followed by a rational restriction:*

$$GRat_{\mathcal{A}} \subseteq \{[\varphi^{-1}(\Lambda)_{|L_\Lambda}]_{\equiv} \mid \forall d \in \mathcal{A}, \varphi(d) \in Lin(X \cup \overline{X}) \wedge L \in Rat(X)\}$$



Proof. First put $E = \{[\varphi^{-1}(\Lambda)_{|L_\Lambda}]_\equiv \mid \forall d \in \mathcal{A}, \varphi(d) \in Lin(X \cup \overline{X}) \wedge L \in Rat(X)\}$.
Let us take a graph $G_0 \in Rat(X^* \times \mathcal{A} \times X^*)$ with $X = \{A, B\}$. Call T_0 a transducer generating G_0, *i.e.,* $T_0 \in \Theta(G_0)$. To construct a substitution corresponding to G_0 we will define a new rational graph in $[G_0]_\equiv$.
We define its transducer T' by substituting in any transition of T_0, every occurrence of A by AA and every occurrence of B by BB. Then to each initial state p, we introduce a new transition (p_0, AB, AB, p). The unique initial state of T' is p_0, the final states of T' are those of T_0. Now, we can easily verify that G'_0 (the rational graph generated by T') is isomorphic to G_0, so $G'_0 \in [G_0]_\equiv$.
Now we construct the following linear grammar $(\mathcal{G})$: for every transition (p, u, v, q) in T' (p and q are states of T') we define the production $q \longrightarrow \widetilde{\overline{u}}pv$; for p_0 we define the production $p_0 \longrightarrow \varepsilon$.
This is a linear grammar. Now for every label d in $\mathcal{A}$ define the linear language:

$$L_d := \bigcup_{p \in \text{ final sates of } T' \text{ producing } d} L(p)$$

where $L(p)$ is the language generated from p by the grammar $\mathcal{G}$.
A triple $(u, d, v) \in X^* \times \mathcal{A} \times X^*$ is an arc in G'_0 if and only if $\widetilde{\overline{u}}v \in L_d$ (by construction).
The language $L := Dom(G'_0) \cup Im(G'_0)$ is by definition a rational language. Furthermore the application $\varphi : \mathcal{A}^* \to 2^{\overline{X}^*}$ such that $\varphi(d) = L_d$ for every d in $\mathcal{A}$, is a linear substitution (notice also that its projections over X^* and $\overline{X}^*$ are rational languages (corresponding to the domain and image of G'_0)). We will show that $G'_0 = \varphi^{-1}(\Lambda)_{|L_\Lambda}$, thus $[G_0]_\equiv$ is equal to $[\varphi^{-1}(\Lambda)_{|L_\Lambda}]_\equiv$ which will allow us to conclude this proof.

First let us prove that $G'_0 \subseteq \varphi^{-1}(\Lambda)_{|L_\Lambda}$.
Suppose $(u, d, v) \in G'_0$. Therefore $\widetilde{\overline{u}}v \in L_d$, *i.e.,* $\widetilde{\overline{u}}v \in \varphi(d)$. Without loss of generality (the name of the vertices of Λ is not important) one can suppose that the root of Λ is ε and then by induction, every vertex in Λ is identified by the unique shortest path from the root to itself. Then we have $u \overset{\widetilde{\overline{u}}}{\underset{\Lambda}{\Longrightarrow}} \varepsilon$ and naturally $\varepsilon \overset{v}{\underset{\Lambda}{\Longrightarrow}} v$, therefore we have : $u \overset{\widetilde{\overline{u}}v}{\underset{\Lambda}{\Longrightarrow}} v$. As u and v are in L_Λ we finally get : $(u, d, v) \in \varphi^{-1}(\Lambda)_{|L_\Lambda}$.

Conversely we will prove that $\varphi^{-1}(\Lambda)_{|L_\Lambda} \subseteq G'_0$.
Suppose now that $(u, d, v) \in \varphi^{-1}(\Lambda)_{|L_\Lambda}$. Therefore $u,\ v \in L_\Lambda$ and there are u_0 and v_0 such that $\widetilde{\overline{u_0}}v_0 \in L_d$ and there is a path $u \overset{\widetilde{\overline{u_0}}v_0}{\underset{\Lambda}{\Longrightarrow}} v$. By definition of L_d, $u_0, v_0 \in L_\Lambda$. The tree structure of Λ implies that there is a vertex w such that $u \overset{\widetilde{\overline{u_0}}}{\underset{\Lambda}{\Longrightarrow}} w$ and $w \overset{v_0}{\underset{\Lambda}{\Longrightarrow}} v$. Therefore (by definition) $w \overset{u_0}{\underset{\Lambda}{\Longrightarrow}} u$ and it implies $u = w.u_0$ and $v = w.v_0$. Now, $u, u_0, v, v_0 \in L_\Lambda$, their length are even integers, therefore the length of w is also even. By construction, every word in L_Λ starts with AB and then is a succession of AA's and BB's. All these facts imply that $w = \varepsilon$. Therefore $u_0 = u$ and $v_0 = v$ so $(u, d, v) \in G'_0$. This yields $\varphi^{-1}(\Lambda)_{|L_\Lambda} \subseteq G'_0$.

Therefore $G'_0 = \varphi^{-1}(\Lambda)_{|L_\Lambda}$ and so $[G_0]_\equiv$ belongs to E which concludes the proof. □

The converse of this result would help us to grab the structure of rational graphs. Unfortunately it is not obvious. Actually the following example illustrate the difficulty of the naive converse of Proposition 3.3.

Example 3.4. Consider $\varphi(a) = \{\overline{B}BA^nB^n | \, n \in \mathbb{N}\}$, it is a linear substitution. Consider $L = BA^*B^*$ and the graph $G = \varphi^{-1}(\Lambda)_{|L_\Lambda}$. Structurally, graph G is rational (it is the star). But the graph naturally associated to G (according to $\varphi(a)$ and L) is $G' = \{(B, a, BA^nB^n) | \, n \in \mathbb{N}\}$, which is not rational.

So there is a deep isomorphism problem to get the converse. Actually, we will try to inject rationality in the "linear language" to achieve a complete characterization of rational graphs.

A natural way to introduce rationality into $Lin(X \cup \overline{X})$ would be to impose the projections over barred and non-barred letters to be rational. The next example shows that again, things are not so nice.

Example 3.5. Consider $\varphi(a) = \{\overline{A}\,\overline{B}BA^nB^m |\ n \geqslant m\} \cup \{\overline{B}BA^nB^m |\ m > n\}$ φ is a linear substitution. Moreover it has rational projections over barred and non-barred letters. Consider $L = BA^*B^*$ and the graph $G = \varphi^{-1}(\Lambda)_{|L_\Lambda}$. Structurally, graph G is rational (it is two stars). But the graph naturally associated to G (according to $\varphi(a)$ and L) is $G' = \{(BA, a, BA^nB^m) |\ n \geqslant m\} \cup \{(B, a, BA^nB^m) |\ m > n\}$, which is not rational (its intersection with the recognizable set $\{BA\} \times \{a\} \times BA^*B^*$ is $\{(BA, a, BA^nB^m) |\ n \geqslant m\}$ which is not rational).

Now consider the set $Ratlin(X \cup \overline{X})$ of linear languages (called rational-linear) over $(X \cup \overline{X})^*$ such that the production of their grammars are of following form: $p \to \overline{u}qv$ (with $\overline{u} \in \overline{X}^*$ and $v \in X^*$) or $p \to \varepsilon$.

Theorem 3.6. *Set $GRat_{\mathcal{A}}$ is precisely the set of graphs obtained from the complete binary tree (Λ) by a* rational-linear *substitution, followed by a rational restriction :*

$$GRat_{\mathcal{A}} = \{[\varphi^{-1}(\Lambda)_{|L_\Lambda}]_{\equiv} \mid \forall d \in \mathcal{A}, \varphi(d) \in Ratlin(X \cup \overline{X}) \wedge L \in Rat(X)\}$$

Proof. Let E be defined as follows:

$$E := \{[\varphi^{-1}(\Lambda)_{|L_\Lambda}]_{\equiv} \mid \forall d \in \mathcal{A}, \varphi(d) \in Ratlin(X \cup \overline{X}) \wedge L \in Rat(X)\}$$

We have to show that $GRat_{\mathcal{A}} = E$.
Step 1: we will first prove that $GRat_{\mathcal{A}} \subseteq E$.
So let us take a graph $G_0 \in Rat(X^* \times \mathcal{A} \times X^*)$. The construction done in the proof of Proposition 3.3 works perfectly. The only thing is to notice that the substitution is a language of *Ratlin* which is obvious.
Step 2: we will now prove that $E \subseteq GRat_{\mathcal{A}}$.
Now consider a graph $G = h^{-1}(\Lambda)|_{L_\Lambda}$ in E.
We will suppose that, for every d in $\mathcal{A}$, $h(d)$ is in $Ratlin(X \cup \overline{X})$ hence every u in $h(d)$ is an element of $\overline{X}^* X^*$. Now fix d in $\mathcal{A}$ and consider a grammar (G_d) generating $h(d)$ (from a non-terminal p_0). Suppose its productions are of the following form:

$$\begin{aligned} p &\longrightarrow \overline{u}qv,\ \overline{u} \in \overline{X}^*, v \in X^* \\ \text{or}\ \ p &\longrightarrow \varepsilon \end{aligned}$$

Now, we construct from each grammar (generating $h(d)$), a transducer T_d (producing the letter d): every non-terminal of the grammar is a state of the transducer. Then for every production $p \longrightarrow \overline{u}qv$ there is a transition $q \xrightarrow{\tilde{u}/v} p$ in T_d. For each non terminal producing the empty word the associated state is an initial state: for all p such that $p \longrightarrow \varepsilon$ then $p \in I(T_d)$. Finally p_0 (the initial non-terminal of G_d) is a final state of the transducer: $F(T_d) = \{p_0\}$.
Consider a word $\overline{u}v \in h(d)$; by construction $(\tilde{u}, v)$ is recognized by T_d. Define T_1 the transducer (with labelled exits) union over all d in $\mathcal{A}$ of the T_d's, and call G_1 the graph generated by T_1. The graph G_1 is rational, and furthermore $G_1 \subseteq h^{-1}(\Lambda)$ (with the same names for the vertices as in step 1 the arcs of G_1 correspond to the paths containing the root of Λ). Now consider the graph G_2:

$$G_2 := \bigcup_{d \in \mathcal{A}} \{(u, d, u) | u \in X^*\} \cdot G_1$$

This is a rational graph (as a finite union and concatenation of rational graphs). Moreover $G_2 = h^{-1}(\Lambda)$:
First, $G_2 \subseteq h^{-1}(\Lambda)$: consider an arc $(u, d, v) \in G_2$. Then there exist u_0, v_0 and u_1 such that $(u, d, v) = (u_1, d, u_1) \cdot (u_0, d, v_0)$, with $(u_0, d, v_0) \in G_1$. Now let us consider the vertices u, v and u_1 in Λ. We have $u = u_1 \cdot u_0$ (as a word) therefore there is a path $u \underset{\Lambda}{\overset{\widetilde{\overline{u}}_0}{\Longrightarrow}} u_1$ and similarly there is a path $u_1 \underset{\Lambda}{\overset{v_0}{\Longrightarrow}} v$. Hence there is a path $u \underset{\Lambda}{\overset{\widetilde{\overline{u}}_0 v_0}{\Longrightarrow}} v$ and, as $\widetilde{\overline{u}}_0 v_0$ is a word in $h(d)$, $(u, d, v) \in h^{-1}(\Lambda)$.
The converse works in exactly the same way. Now we have to produce precisely $h^{-1}(\Lambda)|_{L_\Lambda}$, which is just the intersection of G_2 with $L_\Lambda \times \mathcal{A} \times L_\Lambda$:
$G := G_2 \cap L_\Lambda \times A \times L_\Lambda$. We can then rewrite it in the following way:
$G := \bigcup_{d \in \mathcal{A}} \{G_2 \cap L_\Lambda \times \{d\} \times L_\Lambda\}$. Now, as it is a finite union we just have to prove that $G_2 \cap L_\Lambda \times \{d\} \times L_\Lambda$

is rational. But L_Λ is a rational set over a free monoid, it is therefore recognizable. Hence $L_\Lambda \times \{d\} \times L_\Lambda$ is a recognizable subset of $X^* \times \mathcal{A} \times X^*$ and so its intersection with a rational subset (G_2) of $X^* \times \mathcal{A} \times X^*$ is still rational.
Finally, G equals $h^{-1}(\Lambda)|_{L_\Lambda}$ by construction so we have $E \subseteq GRat_{\mathcal{A}}$ which concludes the proof. □

Now that an external characterization of the rational graphs has been given, the next section will consider the properties of the *traces* of rational graphs.

4 The traces of rational graphs

We have already seen that there is a strong connection between language theory and rational graphs. In this section we will see another connection between graphs and languages, in terms of *traces*.
We first recall that the *trace* of a graph G leading from a vertex set I (of initial states) to a vertex set F (of final states) is the set of all the path labels in the graph, leading from a vertex in the set of initial states to a vertex in the set of final states:

$$L(G, I, F) := \{u \mid \exists s \in I \; \exists t \in F, s \underset{G}{\overset{u}{\Longrightarrow}} t\}$$

In other words the trace of a graph is "the language of its labels". For example the traces of the finite graphs are all rational languages and the traces of *prefix recognizable graphs* are all context-free languages. Notice by the way that the traces of rational graphs contain therefore every context free language.

Proposition 4.1. *The traces of rational graph leading from a rational vertex set to a context free vertex set (or vice-versa) is recursive*

Proof. This simple result is a straightforward consequence of the proof of Proposition 2.16, which states that every rational graph is recursive. Consider a rational graph G. To check whether a word u is in the trace of graph G (from a set I to a set F), it is just to check if the set $S = \underset{G}{\overset{u(|u|)}{\longrightarrow}}(\underset{G}{\overset{u(|u|-1)}{\longrightarrow}}(\cdots \underset{G}{\overset{u(1)}{\longrightarrow}}(I)\cdots))$ intersects set F. Set I is rational, thus its image by a rational transduction is rational too, hence by a simple induction, set S is rational too. Therefore it is decidable whether $S \cap F$ is empty. In the reverse case (the path between a context-free vertex sets and a rational vertex set): the situation is the same because the image of a context-free set by a rational transduction is a context-free set. □

Let us denote by TR the family of the traces of rational graphs leading from a rational vertex set to a rational vertex set: $TR = \{L(G, I, F) |\, G \in Rat(X^* \times \mathcal{A} \times X^*) \wedge I, F \in Rat(X^*)\}$ (notice that we could as well restrict ourselves to a unique initial state and a unique final state). Now we will show that set TR form an *Abstract Family of Languages* (AFL), that is, it satisfies following properties:

- closure for intersection with a rational (regular) language,
- closure under non-erasing (monoid)morphism, and inverse morphism,
- for each $L, L' \in TR$ we have $L \cdot L', L \cap L', L^+, L^* \in TR$.

Proposition 4.2. *The intersection of two elements of TR is an element of TR.*

Proof. Consider two elements L and L' of TR. Say $L = L(G, I, F)$ and $L' = L(G', I, F)$. The language $L \cap L'$ is actually the trace of $G \cdot (\{\$\} \times \mathcal{A} \times \{\$\}) \cdot G'$ (with \$ a new symbol) between $I_G \cdot \{\$\} \cdot I_{G'}$ and $F_G \cdot \{\$\} \cdot F_{G'}$. Hence $L \cap L'$ in an element of TR. □

As rational languages are traces of rational graphs (finite graphs are rational graphs), family TR is closed under intersection with rational languages.
Now let us recall that a finite (*resp.* rational) substitution $\sigma : \mathcal{A}^* \to 2^{\mathcal{A}^*}$ is a morphism such that for each letter d in $\mathcal{A}$ $\sigma(d)$ is a finite (*resp.* rational) subset of $\mathcal{A}^*$. A substitution is *non-erasing* if $\varepsilon \notin \sigma(d)$ for all $d \in \mathcal{A}$.

Proposition 4.3. *Family TR is closed under non-erasing finite substitution.*

Proof. Consider σ a finite substitution, and L a language in *TR*. Say $L = L(G, I, F)$, and T is a transducer in $\Theta(G)$. Now, let d be a letter of $\mathcal{A}$, and w in $\sigma(d)$, we will perform the following construction (there are a finite number of such constructions because $\mathcal{A}$ and each $\sigma(d)$ are finite): we construct a transducer $T_{(d,w)}$ that will produce for each arc (u, d, v) of G a path labelled w between u and v (the construction of such a transducer is explained bellow). Then, naturally language $\sigma(L)$ is the trace of the graph generated by the union of all the constructed transducers:

$$\sigma(L) = L\left(\bigcup_{d\in\mathcal{A}}\left[\bigcup_{w\in\sigma(d)} G(T_{(d,w)})\right], I, F\right)$$

where $G(T)$ is the graph generated by transducer T. Now it only remains to explain the construction of transducer $T_{(d,w)}$, for d in $\mathcal{A}$ and w in $\sigma(d)$. Consider a new symbol $\#$ that is not an element of X and a new symbol $\$$ that is not in $\mathcal{A}$. Now we construct $T_{(d,w)}$ by induction on the length of w.
($|w| = 1$). Then $T_{(d,w)}$ is simply $T_{|d}$ with all exits labelled w (instead of d).
($|w| = 2$). For each final state q producing d in T, we construct two final states of $T_{(d,w)}$: q_1 that will produce $(u, u\#)$ for all (u, v) produced by q; q_2 that will produce $(u\#, v)$ for all (u, v) produced by q. Then we simply label q_1 with $w(1)$ and q_2 with $w(2)$. Doing that for all final state (q) producing d yields transducer $T_{(d,w)}$.
($n \Rightarrow n+1$). Suppose that we can construct $T_{(d,w)}$ for every w such that $|w| \leqslant n$. And consider w with $|w| = n+1$. To construct $T_{(d,w)}$ we first construct a transducer T_1 realizing the substitution $\sigma(d) = \$w(|w|)$ (using ($|w| = 2$)). Then we transform T_1 substituting the word $w(1)w(2)\ldots w(|w|-1)$ to $\$$ (which is possible thanks to the induction hypothesis) which produces $T_{(d,w)}$. □

Following corollary is a direct consequence of this proposition.

Corollary 4.4. *Family TR is closed under non-erasing morphism.*

Notice that the condition "non-erasing" is essential for our proof. A interesting question is whether this condition is necessary. There is a deep graph isomorphism problem to solve to answer this question.

Proposition 4.5. *Assume that L is an element of TR and that σ is a finite substitution over $\mathcal{A}^*$ then $\sigma^{-1}(L)$ is a language of TR.*

Proof. This proposition is a consequence of Elgot and Mezei's theorem, which states that the composition of two rational relations is a rational relation (see for example [Be 79], Theorem 4.4 p 68). To simplify the notations, we will consider the case where $\mathcal{A}$ has two elements (a and b), it would be exactly the same for more letters. Now say $\sigma(a) = \{u_1, u_2, \ldots, u_n\}$ and $\sigma(b) = \{v_1, v_2, \ldots, v_m\}$. Let L_0 be in *TR*. Let G, I and F be respectively a rational graph and two rational sets such that $L_0 = L(G, I, F)$. Let T be a transducer in $\Theta(G)$ and denote by T_a and T_b its rational relations associated to a and b. Now define $T'_a = \bigcup_{i=1}^{n}[T_{u_i(1)}\circ T_{u_i(2)}\circ\cdots\circ T_{u_i(|u_i|)}]$ and $T'_b = \bigcup_{i=1}^{m}[T_{v_i(1)}\circ T_{v_i(2)}\circ\cdots\circ T_{v_i(|v_i|)}]$. These two relations are rational relations, and therefore the associated graph G' (which associates an arc labelled a to each pair of T'_a and an arc labelled b to each pair of T'_b) is rational too. The trace in G' leading from I to F is precisely $\sigma^{-1}(L_0)$ which is therefore an element of *TR*. □

Remark: Note that it is not as straightforward for inverse rational substitution. Actually it seems that it is not true for inverse rational substitution: consider any rational graph with one label (a) and the inverse rational substitution $\sigma(a) = a^*$. The graph image with the same approach would be the transitive closure of the original graph, which is not effectively rational (and might not even be structurally rational) as stated in the remark page 12.
Following corollary is an obvious consequence of proposition 4.5.

Corollary 4.6. *Family TR is closed under inverse morphism.*

Proposition 4.7. *Family TR is closed under concatenation, Kleene plus and star.*

Proof. The argument is more or less the same for the three operations.
Step 1: Consider L_1 and L_2 two elements of *TR*. Suppose now that $L_1 = L(G_1, I_1, F_1)$ and $L_2 = L(G_2, I_2, F_2)$ We can also suppose that $V(G_1) \cap V(G_2) = \emptyset$. Now, for each arc (u, d, v) in G_1, with v in F_1, we want to define new arcs labelled d leading from u to each vertex in I_2. We call $G_{F,I}$ the set of all these arcs. If $G_{F,I}$ is rational then the graph $G_1 \cup G_{F,I} \cup G_2$ is rational. Also the trace of $G_1 \cup G_{F,I} \cup G_2$,

leading from I_1 to F_2, is precisely $L_1 \cdot L_2$. It only remain to show that $G_{F,I}$ is rational.
Let T be a transducer in $\Theta(G_1)$. For each letter d in $\mathcal{A}$ the set $T_d^{-1}(F_1)$ is a rational set. As I_2 is also a rational set, the set $T_d^{-1}(F_1) \times I_2$ is recognizable. Hence $T_d^{-1}(F_1) \times \{d\} \times I_2$ is also recognizable and therefore $G_{F,I} = \bigcup_{d \in \mathcal{A}} (T_d^{-1}(F_1) \times \{d\} \times I_2)$ is rational. This concludes step 1.
Step 2: Consider L in *TR* we want to show that L^+ is in *TR*. It is even simpler: let G be a graph having L for trace between I and F. The graph $G_{I,F} := \bigcup_{d \in \mathcal{A}} (T_d^{-1}(F) \times \{d\} \times I)$ is, for the same reasons as in step 1, a rational set. Similarly $G \cup G_{I,F}$ has L^+ for trace between I and F. Hence L^+ is an element of *TR*.
Step 3: Using the same assumptions as in step 2, we define L^+. Then we have to define a new initial state $\#$ in G (where $\#$ is a new symbol), such that there will be no return to $\#$. Consider set $G_{\#,I} = \bigcup_{d \in \mathcal{A}} (\{\#\} \times \{d\} \times T_d(I))$. Define $F' = F \cup \{\#\}$, we have $L^* = L_{(\{\#\},F')}(G \cup G_{I,F} \cup G_{\#,I})$. It is therefore an element of *TR*. □

As stated earlier, we only have now to summary these results.

Theorem 4.8. *The traces of rational graphs, leading from a rational vertex set to a rational vertex set, form an AFL (Abstract Family of Languages).*

Proof. This result is simply a brief summary of corollaries 4.4, 4.6 and propositions 4.2 and 4.7. □

Now we have an abstract family of languages that contains the context free languages. This AFL is a subset of the recursive languages. It seems that this family is composed of the context sensitive languages.

Conjecture 4.9. *The traces of the rational graphs are precisely the context sensitives languages.*

Notice also that recently graphs of linear bounded machines (which characterize context sensitive languages) have been studied in [KP 99].

5 Conclusion

In this paper, a general family of graphs has been introduced. Rational graphs are a strict extension of previously studied families. It is a well grounded family, related to well known structures of language theory. We have given both an internal and an external characterization, as well as some basic properties.

Unfortunately, or fortunately depending on the point of view, it is a very expressive family. Therefore many decision results are lost. An interesting question is to study restrictions of this family that will retain decision results from former families.

Traces of rational graphs are another aspect of this family. We have shown that it forms an abstract family of recusive languages. An interesting question is to know if these traces are precisely the context sensitive languages.

Rational trees also seem to be an interesting field of research, but this has not been done yet.

Acknowledgements

The author would like to express his gratitude to Didier Caucal for his help along the preparation of this paper.

References

[AN 88] A. Arnold and M. Nivat, *Comportements de processus*, Colloque AFCET “les mathématiques de l’informatique”, pp. 35–68, 1982.

[Au 88] J.-M. Autebert and L. Boasson, *Transductions rationelles*, Ed. Masson, pp. 1–133, 1988.

[Be 79] J. Berstel *Transductions and context-free languages*, Ed. Teubner, pp. 1–278, 1979.

[Ca 96] D. Caucal *On transition graphs having a decidable monadic theory*, LNCS 1099, pp. 194–205, 1996, [Co 90] B. COURCELLE *Graph rewriting: an algebraic and logic approach*, Handbook of TCS, Vol. B, Elsevier, pp. 193–242, 1990.

[GG 66] S. GINSBURG and S. A. GREIBACH *Mappings which preserve context sensitive languages*, Information and control 9, pp. 563–582, 1966.

[HU 79] J. E. HOPCROFT and J. D. ULLMAN *Introduction to automata theory, langages and computation*, Ed. ADDISON-WESLEY pp. 1–284, 1979.

[KP 99] T. KNAPIK and E. PAYET *Synchronization product of Linear Bounded Machines*, in FCT LNCS 1684, pp. 362–373, 1999.

[MS 85] D. MULLER and P. SCHUPP *The theory of ends, pushdown automata, and second-order logic*, TCS 37, pp. 51–75, 1985.

[Ni 68] M. NIVAT *Transduction des langages de Chomsky*, Ann. de l'Inst. FOURIER 18, pp. 339-456, 1968.

[Re 85] W. REISIG *Petri nets*. EATCS Monographs on Theoretical Computer Science, Vol. 4, Springer Verlag, 1985.

[Sc 76] M.P. SCHÜTZENBERGER *Sur les relations rationnelles entre monoïdes libres*, TCS 3, pp. 243–259, 1976.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399