Classes of rational graphs

Christophe Morvan
IRISA, Campus de Beaulieu, 35042 Rennes, France
christophe.morvan@irisa.fr

Abstract

The rational graphs are a family of graphs characterized by rational transducers. This family is general and counts many classical infinite graphs. The drawback of this generality is that most of the properties of its elements are undecidable. There is a natural way for solving this problem: to consider sub-families of this family. In this paper we examine several such families: rational trees, commutative rational graphs, monotonous rational graphs. We establish decidability results, like accessibility or first order theory, and discuss on further extensions on these results.

1 Introduction

Finite automata are a very classical model for computation. It has been used intensively to characterize infinite sets of words. Since the nineteen fifties it has been generalized to define sets of infinite words, relation, relational structures, groups or graphs.

These automata have been the building blocks of the study of infinite graphs: in 1960 Büchi, [Büc60], used finite automata to characterize infinite words, and so proving the decidability of monadic second order logic of the integers. This result was extended by Rabin to the complete binary tree [Rab69]. Around the year 1990 Muller and Schupp, then Courcelle and finally Caucal proposed generalizations of Rabin's result based on transformation of the complete binary tree [MS85, Cou90, Cau96].

Finite automata have also been used in 1983 by Hodgson [Hod83], to define relational structures, obtaining the decidability of first order logic. This approach was further utilized by Khoussainov and Nerode who formalized and generalized the notion of automatic structure (and graph) [KN94]. These structures have been considered in a slightly different way by Senizergue, and later on Pelecq, it involved an automatic quotient [Sén92, Pél97]. The author then prolonged this study introducing a general notion of rational graphs [Mor00], this general family had already been defined as asynchronous automatic by Khoussainov and Nerode, but it was not very satisfactory from the logical point of view.

The rational graphs form an expressive family of graphs, the drawback of this richness is that almost every property of this graphs is undecidable. An important property of this family is that its traces are the context-sensitive languages [MS01]. Indeed this result may be extended to automatic graphs [Ris03, MR05].

In this paper we will examine various families of rational graphs. We will present some properties justifying their study. The first families are considered by structural restriction: the rational trees, examined thoroughly in [CM06], and the deterministic rational graphs. In the second part of this paper we consider families defined by restriction on the transducer: automatic graphs, monotonous graphs and commutative rational graphs. For the two latter families very few properties are already established, but each of these family has either a decidable accessibility or a decidable first order theory.

2 Rational graphs

This section recalls the definition of rational graph and some of their simple properties. More details can be found in [Mor00, MR05].

2.1 Definitions

A classical definition of the rational subsets of a monoid $(M, \cdot)$ is the least family containing finite (and empty) subsets of M and closed under union, concatenation and iteration.

Now, let Γ be a finite alphabet and X^* a free monoid. A *graph* G is a subset of $X^* \times \Gamma \times X^*$. If $|\Gamma| = 1$ then it is often omitted in this case the graph is simply a subset of $X^* \times X^*$. An *arc* is a triple: $(u, a, v) \in X^* \times \Gamma \times X^*$ (denoted by $u \xrightarrow[G]{a} v$ or simply $u \xrightarrow{a} v$ if G is understood).

Set $X^* \times \{a_i\} \times X^*$, with $(u, a_i, v) \cdot_i (u', a_i, v') = (uu', a_i, vv')$ (a_i in Γ) is a monoid. We denote by $\cdot$ the operation in $X^* \times \Gamma \times X^*$: $(u, a_i, v) \cdot (u', a_i, v') = (u, a_i, v) \cdot_i (u', a_i, v')$ and $(u, a_i, v) \cdot (u', a_j, v')$ is undefined for i different from j. This operation is similar to the synchronization product for transition systems defined by Arnold and Nivat in [AN82]. Following this definition $X^* \times \Gamma \times X^*$ is not a monoid; still, as shown in [Mor00], we may define rational subset of this set: the set of *rational graphs*, denoted $Rat(X^* \times \Gamma \times X^*)$, is the family of rational subsets of $X^* \times \Gamma \times X^*$.

A transducer is a finite automaton over pairs, see for example [AB88] [Ber79], and a rational relation, that is a rational subset of $X^* \times X^*$, is recognized by a rational transducer. These transducer extends naturally to rational graphs.

Definition 2.1. A *labelled transducer* $T = (Q, I, F, E, L)$ 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^{Γ}.

An arc $u \xrightarrow{a} v$ is *recognized* by a labelled transducer T if there is a path from a state in I to a state f in F labelled (u, v) and that $a \in L(f)$. Labelled transducers are a nice characterization of rational graphs:

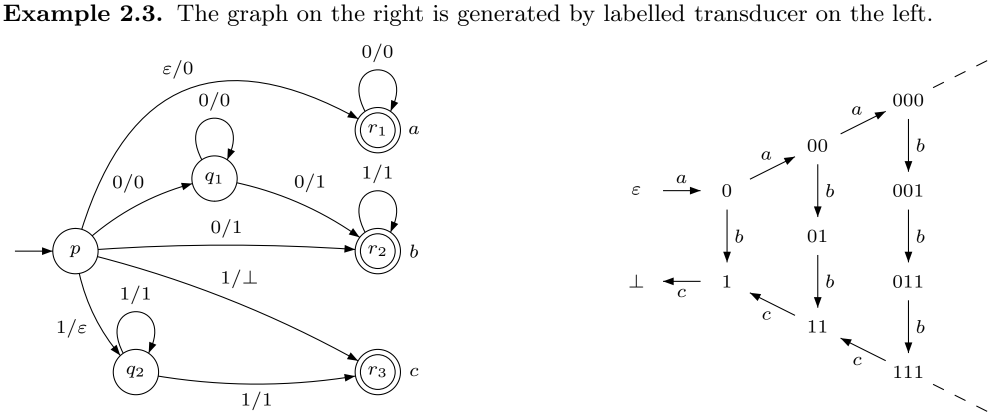
Proposition 2.2. *A graph G in $2^{X^* \times \Gamma \times X^*}$ is rational if and only if it satisfies one of the following equivalent properties:*

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

(ii) G is recognized by labelled rational transducer.

A rational graph G is a rational subset of $X^* \times \Gamma \times X^*$, for each a in Γ we denote by G_a the restriction of G to arcs labelled a (it is a rational graph, and thus defines a rational relation between vertices); let u be a vertex in X^*, we denote by $G_a(u)$ the set of all vertices v such that $u \xrightarrow{a} v$ is an arc of G. We denote by T_a a transducer recognizing relation G_a.

Example 2.3. The graph on the right is generated by labelled transducer on the left.



For example, the arc: $001 \xrightarrow{b} 011$ is recognized by following path: $p \xrightarrow{0/0} q_1 \xrightarrow{0/1} r_2 \xrightarrow{1/1} r_2$. Final state r_2 is associated to letter b therefore the couple $(001, 011)$ is labelled b.

There are very few positive decidability results for rational graphs. We provide here a brief summary of negative results.

Proposition 2.4. *Following properties are undecidable for rational graphs:*

(i) Accessibility (between two vertices);

(ii) Connectedness (of the whole graph);

(iii) Isomorphism (of two graphs);

(iv) First order theory (of a rational graph).

One of the most unexpected result about rational graphs is that the traces (*i.e.*, the set of path label between sets of initial and final vertices) of these graphs is precisely the context sensitive languages.

Theorem 2.5 (Morvan Stirling 01). *The traces of rational graphs between two vertices coincide with the context sensitive languages.*

In the following we will state for which sub-families of rational graph this result remains true.

3 Families of rational graphs

There are several ways to isolate sub-families of graphs. These ways can be divided into two major approaches: structural and internal restrictions. The first ones consist, for example, in considering deterministic graphs, or trees, whereas the second ones involves limitations on the transducer, like synchronicity. In this section we will examine restrictions of both kind: rational trees and deterministic rational graphs are of the first one, whereas automatic, monotonous and commutative rational graphs relies on restriction on the transducers.

3.1 Rational trees

The trees are classical structure in computer science. They can be considered from different points of view. The most simple point of view is simply to consider trees as graphs having certain properties.

One can define formally rational trees as follows:

Definition 3.1. A rational tree is a rational graph satisfying these properties:

(i) it is connected;

(ii) every vertex is the target of at most one arc;

(iii) there is a single vertex which is not the target of any arc (it is called the *root*).

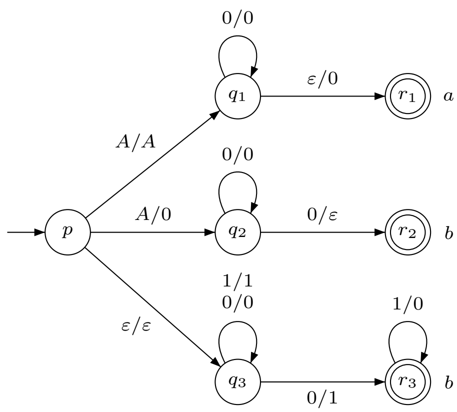
There is a difficulty in studying the rational trees: it is undecidable to know from a transducer is a rational graph is a tree.

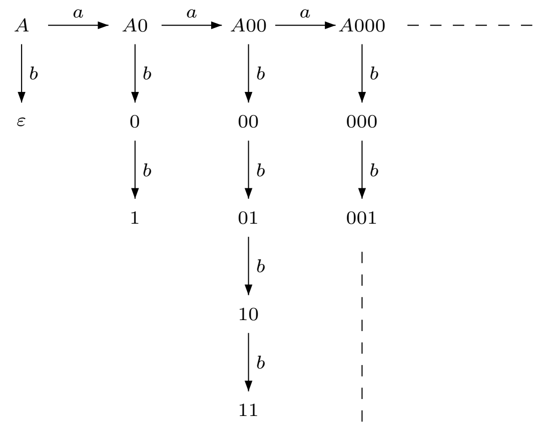
Proposition 3.2. *It is undecidable to know whether a rational graph is a tree.*

Nonetheless provided a rational tree many properties may be computed algorithmically.

Proposition 3.3. *Given any rational tree, accessibility and rational accessibility is decidable for any pair of vertices.*

Example 3.4. We give here an example of rational tree. It is defined by a half-line of a's, and the nth vertex of this line is connected to a segment of 2^n b's.





The encoding of the vertices of this tree relies on the fact that there are 2^n n-tuples over $\{0,1\}$. The transducer on the left performs a binary increment.

It is relatively unexpected that first order theory is decidable for rational trees. It relies only on the structure of the trees.

Proposition 3.5. *Every rational tree has a decidable first-order theory.*

This result is difficult to extend either improving the expressive power of the logic or the class of graphs.

Proposition 3.6.

- *The first order theory with accessibility is undecidable for rational trees*
- *The first order theory of rational directed acyclic graphs is undecidable*

Considering the branch language of rational trees (the language of words labelling branches starting on the root and finishing with a leave) it is not known whether it is all context sensitive languages.

Conjecture 3.7. *There are context sensitive languages which are not branch language of any rational tree.*

3.2 Deterministic rational graphs

Determinism is a simple property of graphs. In the case of finite graphs the expressive power of deterministic graph is the same as non-deterministic ones, this result is no longer true for pushdown-automata (which can be viewed as infinite graphs).

Considering rational graphs determinism coincides with functionality of the underlying transducer, it therefore decidable.

Proposition 3.8. *Determinism is decidable for rational graphs.*

The most obvious interest in the study of deterministic rational graphs lies in the traces. In order to preserve determinacy, one consider only traces leading **from a single vertex** to a finite set of vertices.

The example 2.3 ensures that the traces of deterministic rational graphs contains non context free languages. Using techniques similar to those used for finite graphs we obtain following result.

Proposition 3.9. *The traces of deterministic rational graphs forms a Boolean algebra of context sensitive languages containing non context free languages.*

It is not obvious that all context free languages are traces of deterministic rational graphs.

3.3 Automatic graphs

An interesting family of rational graphs is the set of automatic graphs [KN94, Pél97, BG00].

These graphs are recognized by letter-to-letter transducers with rational terminal functions completing one side of the recognized pairs and assigning a label to the arc. The terminal function associates a set of pairs to each terminal state of the transducer. Then, the relation defined is the set of labels of path ending at a state q, concatenated with pairs of the terminal function's value in q.

Definition 3.10. A graph over $X^* \times \Gamma \times X^*$ is *left-automatic* if it is recognized by a letter-to-letter transducer T with terminal function f taking values in

$$\mathrm{Dif}_{Rat} \quad = \quad (Rat(X^*) \times \{\varepsilon\}) \ \cup \ (\{\varepsilon\} \times Rat(X^*))$$

That is, a left-automatic relation is a finite union of elementary relations of the form $R.S$ where $R \in Rat((X \times X)^*)$ and $S \in \mathrm{Dif}_{Rat}$. Right-automatic relations are defined symmetrically with an initial rational function. A rational relation is *automatic* if it is left-automatic or right-automatic.

As the terminal function is rational, it can be introduced in the transducer adding states and transitions. A *left-synchronized transducer* is a transducer such that each path leading from an initial vertex to a final one can be divided into two parts: the first one contains arcs of the form $p \xrightarrow{A/B} q$ with $A, B \in X$ while the second part contains either arcs of the form $p \xrightarrow{A/\varepsilon} q$ with $A \in X$ or of the form $p \xrightarrow{\varepsilon/B} q$ with $B \in X$ (not both).

Example 3.11. The examples 2.3 and 3.4 are automatic graphs. Notice that there are rational graphs and trees that are not automatic up to isomorphism.

A graph G is *left-automatic* if for each $a \in \Gamma$, the relation G_a is left-automatic.

The restriction on the automatic transducers induce a structure of Boolean algebra which allow the decision of first order theory by definability.

Proposition 3.12 (Hodgson 83). *The first order theory of automatic graphs is decidable.*

This family also has the same traces as rational graphs.

Theorem 3.13 (Rispal 03). *The traces of automatic graphs between finite sets are context sensitive languages.*

Whereas rational graphs may be restricted to finite (yet unbounded) degree to obtain context-sensitive languages between finite sets, the automatic graphs need infinite degree.

3.4 Monotonous graphs

A rational transducer is said monotonous if its transitions u/v, satisfy either $|u| \geqslant |v|$ or $|u| \leqslant |v|$. By extension a rational graph is said monotonous if its transducer is monotonous (for all letters).

This property has an obvious consequence: accessibility is decidable.

Proposition 3.14. *Accessibility for monotonous rational graphs is decidable.*

It is not obvious whether or not first order theory is decidable for monotonous graphs. There are some hints that the answer to this question is positive.

Conjecture 3.15. *The first order theory of monotonous graphs is decidable.*

One can note that every rational graph is equivalent up to bisimulation to a monotonous graph.

Proposition 3.16. *For each rational graph an equivalent bisimilar monotonous rational graph is computable.*

This result implies directly following result.

Corollary 3.17. *The traces of monotonous rational graphs are context sensitive languages.*

3.5 Commutative graphs

An obvious restriction to the expressive power of transducers is to limit the size of the alphabet to only one symbol. In that case, we obtain rational graphs over the natural numbers.

Rational sets on commutative monoid (like $(\mathbb{N}, +)$) have been intensively studied. The most stunning result was demonstrated by Eilenberg and Schützenberger, [ES69].

Theorem 3.18 (Eilenberg, Schützenberger 69). *The family of rational sets in a commutative monoid form a Boolean algebra.*

This result implies directly the decidability of the first order theory of commutative rational graphs.

Corollary 3.19. *First order theory is decidable for commutative graphs.*

Unfortunately it is possible to encode Minsky into commutative graphs, resulting in the undecidability of accessibility.

Proposition 3.20. *Accessibility is undecidable for commutative rational graphs.*

Unlike rational graphs this family does not contain the prefix-recognizable graphs.

Proposition 3.21. *The families of commutative rational graphs and prefix recognizable graphs are incomparable.*

Very few is known about the traces of these graphs. It contains non context-free languages. But like deterministic rational graphs it is not obvious whether it contains all context-free languages.

4 Conclusion

In this paper we have presented several families of rational graphs. Some of these families like rational trees have been deeply investigated. Some others have simply been defined (like monotonous or commutative graphs), these families have promising elementary properties and deserve to be examined more carefully.

The most noticeable aspect of these families is that they share, like rational graphs, a very simple, intuitive and familiar model. All the restrictions proposed in this paper are also natural and correspond to usual restrictions of automaton or graphs.

There is still a direction in which these restrictions have not proved successful: it is to have the decidability of first order theory with accessibility. We have had either or both, but not together. It seems that this is a strong restriction of the transducer whose iteration is rarely computable.

References

[AB88] J.-M. Autebert and L. Boasson, *Transductions rationelles*, Masson, 1988.

[AN82] A. Arnold and M. Nivat, *Comportements de processus*, Colloque AFCET, 1982, pp. 35–68.

[Ber79] J. Berstel, *Transductions and context-free languages*, Teubner, 1979.

[BG00] A. Blumensath and E. Grädel, *Automatic Structures*, Proceedings of 15th IEEE Symposium on Logic in Computer Science LICS 2000, 2000, pp. 51–62.

[Büc60] J. R. Büchi, *On a decision method in restricted second order arithmetic*, ICLMPS, Stanford University press, 1960, pp. 1–11.

[Cau96] D. Caucal, *On transition graphs having a decidable monadic theory*, Icalp 96, LNCS, vol. 1099, 1996, pp. 194–205.

[CM06] A. Carayol and C. Morvan, *On rational trees*, Submited, 2006.

[Cou90] B. Courcelle, *Handbook of theoretical computer science*, ch. Graph rewriting: an algebraic and logic approach, Elsevier, 1990.

[ES69] S. Eilenberg and M.-P. Schützenberger, *Rational sets in commutative monoids*, Journal of Algebra **13** (1969), no. 2, 173–191.

[Hod83] B. R. Hodgson, *Décidabilité par automate fini*, Ann. Sci. Math. Québec **7** (1983), 39–57.

[KN94] B. Khoussainov and A. Nerode, *Automatic presentations of structures*, LCC (D. Leivant, ed.), LNCS, vol. 960, 1994, pp. 367–392.

[Mor00] C. Morvan, *On rational graphs*, Fossacs 00 (J. Tiuryn, ed.), LNCS, vol. 1784, 2000, ETAPS 2000 best theoretical paper Award, pp. 252–266.

[MR05] C. Morvan and C. Rispal, *Families of automata characterizing context-sensitive languages*, Acta Informatica **41** (2005), no. 4-5, 293–314.

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

[MS01] C. Morvan and C. Stirling, *Rational graphs trace context-sensitive languages*, MFCS 01 (A. Pultr and J. Sgall, eds.), LNCS, vol. 2136, 2001, pp. 548–559.

[Pél97] L. Pélecq, *Isomorphismes et automorphismes des graphes context-free, équationnels et automatiques*, Ph.D. thesis, Université de Bordeau I, 1997.

[Rab69] M.O. Rabin, *Decidability of second-order theories and automata on infinite trees*, Trans. Amer. Math. soc. **141** (1969), 1–35.

[Ris03] C. Rispal, *The synchronized graphs trace the context-sensitive languages*, ENTCS, vol. 68, 2003.

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