Rational relations having a rational trace on each finite intersection of rational relations

Christian Choffrut *
http://www.liafa.jussieu.fr/∼cc
cc@liafa.jussieu.fr

Serge Grigorieff *
http://www.liafa.jussieu.fr/∼seg
seg@liafa.jussieu.fr

May 18, 2012

Abstract

We consider the family of rational relations on words, i.e., relations recognized by multitape automata, which have rational trace on any finite intersection of rational relations. We prove that this family consists of finite unions of relations which are of two types: stars of tuples possibly with an extra prefix and suffix and relations with all but one singleton set components.

Contents

*LIAFA, Université Paris 7 & CNRS, 2, pl. Jussieu 75251 Paris Cedex 05

1 Introduction

We are concerned with the family of k-ary relations on words which are recognized by k-tape automata as defined in the model of Rabin-Scott [9] and later of Elgot and Mezei [4] now known as rational relations. A big departure from the single tape case is the fact that this family is not closed under complement and intersection. There exist subfamilies of binary relations which are Boolean algebras, in increasing order the recognizable relations, [4], the *regular prefix* relations in the terminology of [1], also known as *special* relations in [6] and the synchronous relations, [2]. However, the diagonal $\{(u, u) \mid u \in A^*\}$ fails to have a rational intersection with all rational relations though it is very low in the hierarchy: it is not recognizable but it is special and therefore synchronous.

Characterizing the rational relations whose intersection with an arbitrary relation is rational seems out of reach in the current state of the theory. However, we are able to characterize the hereditary counterpart, namely the rational relations, all rational subrelations of which have a rational intersection with an arbitrary rational relation. In that case we say that the relation has the **HI** (for hereditary intersection) property.

The characterization is an elaboration of the idea that there are only two reasons to enjoy the property above, namely to be degenerate, i.e., all components are singleton sets but one which is an arbitrary rational language or to be an encoding of a direct product of unary free monoids. The precise statement (using Definitions 3.1, 4.2) is in Theorem 5.6.

The paper is organized as follows. The preliminaries recall all the basic definitions on direct products of free monoids, the main closure properties of the families of rational and recognizable relations and a minimal survival kit of combinatorics of words. All these definitions allow us to state the problem precisely. Section 3 and 4 respectively introduce two families of relations which enjoy the property under investigation. Section 5 is devoted to the converse and thus to prove the main result.

2 Preliminaries

The free monoid generated by the set A, called the *alphabet*, is denoted by A^*. Its elements are *words* and the neutral element, denoted 1 is the *empty* word. On a direct product of k free monoids $A_1^* \times \cdots \times A_k^*$, the operation is the componentwise concatenation. Its neutral element is also denoted 1. In the sequel, all alphabets are finite.

In all our statements, unless otherwise stated, the symbol M denotes a product $A_1^* \times \cdots \times A_k^*$ of $k \geq 1$ free monoids.

The results of the literature that we use hold most of the time for more general monoids, but we did not bother to optimize the statements. The subsets of a direct product of $k \geq 2$ free monoids is usually called a *relation*. The term *language* is reserved for the free monoid, i.e., when $k = 1$.

Since we are dealing with direct products of free monoids, we use a graphical convention about tuples of words in order to distinguish between the components of the tuples and the elements of a sequence of tuples. We write $\mathbf{u}$ for a k-tuple $(u_1, \ldots, u_k)$ and $\mathbf{u_1}, \ldots, \mathbf{u_m}$ for a sequence of tuples of words where $\mathbf{u_i} = (u_{i,1}, \ldots, u_{i,k})$.

2.1 Rational and recognizable subsets

The definitions of this section are standard. We recall them for the purpose of selfcontainment. The reader who wishes to deepen his familiarity with these concepts is referred to the numerous textbooks on the topic, in particular Sakarovitch [10].

Definition 2.1. The family $\mathrm{Rat}(M)$ of *rational* subsets of the monoid M is the least family of relations containing the finite relations and closed under

- (i) set union,
- (ii) *set concatenation* (also simply called *product*): if $R, S \subseteq M$ their concatenation RS is the set $\{xy \mid x \in R, y \in S\}$,
- (iii) and *Kleene star*: if $R \subseteq M$, its Kleene star R^* is the submonoid of M generated by R, i.e. $R^* = \bigcup_{i \geq 0} R^{(i)}$ where $R^{(0)} = \{1\}$ and $R^{(i+1)} = R\, R^{(i)}$.

Whenever $k = 1$ (M is a free monoid) or all alphabets are unary (M is isomorphic to $\mathbb{N}^k$) $\mathrm{Rat}(M)$ is a Boolean algebra. The motivation of the present study is the observation that except for these two cases, $\mathrm{Rat}(M)$ is not closed under complement and intersection. The classical example is as follows: consider $R_1, R_2 \subseteq \{a,b\}^* \times \{c\}^*$ where $R_1 = (a,c)^*(b,1)^*$ and $R_2 = (a,1)^*(b,c)^*$. Then their intersection is the relation $\{a^n b^n, c^n) \mid n \geq 0\} \subseteq \{a,b\}^* \times \{c\}^*$. It is therefore natural to inquire about the family of rational relations R for which $R \cap S$ is rational for all rational relations S. A similar question could be asked about context-free languages of a free monoid: which context-free languages have a context-free intersection with all context-free languages? These two questions are tightly connected and we doubt that they could be answered easily in the next future. Intuitively, the difficulty stems from the fact that it is not easy to tailor a rational relation which extracts by intersection a subset of R with some specific properties. We overcome this issue by choosing an arbitrary subset of R in the first place.

More precisely, we say that a rational relation R has the *hereditary intersection property*, abbreviated **HI**, whenever the following holds

(**HI**) *for all rational relations $S \subseteq R$, for all rational relations T, the intersection $S \cap T$ is rational.*

In this paper we characterize these rational relations.

We now recall the second family of relations of importance.

Definition 2.2. The family $\mathrm{Rec}(M)$ of *recognizable* subsets of the monoid M consists of all subsets $X \subseteq M$ for which there exists a morphism φ from M onto a finite monoid F such that $X = \varphi^{-1}\varphi(X)$ holds.

The following property, due to Elgot & Mezei [4], characterizes the recognizable relations of the direct product of free monoids.

Proposition 2.3. *A relation included in the direct product $A_1^* \times \cdots \times A_k^*$ is recognizable if and only if it is a finite union of products $X_1 \times \cdots \times X_k$ where $X_i \in \mathrm{Rec}(A_i^*)$ for $i = 1, \ldots, k$.*

How do these two families compare? When $k > 1$, $\mathrm{Rec}(M)$ is properly included in $\mathrm{Rat}(M)$. They coincide when $k = 1$, which is Kleene's Theorem. Besides the obvious closure properties of $\mathrm{Rat}(M)$ stemming from the definition, there are a few general closure properties. Let us recall some that we use in this paper. Particularly useful is a weak intersection property which is item 2. The proposition is stated for arbitrary monoids.

Proposition 2.4. *Let M, N be two arbitrary monoids.*
1. The family $\mathrm{Rec}(M)$ is a Boolean algebra: if X, Y are in $\mathrm{Rec}(M)$ then so are $X \cup Y$, $X \cap Y$ and $M \setminus X$.
2. If M is finitely generated, the intersection of a rational and of a recognizable subset is rational: $R \in \mathrm{Rat}(M)$ and $X \in \mathrm{Rec}(M)$ implies $R \cap X \in Rat(M)$.
3. If $\varphi : M \to N$ is a morphism and $R \in \mathrm{Rat}(M)$ then $\varphi(R) \in \mathrm{Rat}(N)$.

We will also need the following technical result.

Proposition 2.5. *Let $M = A_1^* \times \cdots \times A_k^*$. If $R \in \mathrm{Rat}(M)$ and $x, y \in M$ then $x^{-1}Ry^{-1} = \{m \in M \mid xmy \in R\}$ is in $\mathrm{Rat}(M)$.*

2.2 Star-chain and loop sets

Definition 2.6. A subset of M is *star-chain rational* if, for some m, if it is of the form

$$\mathbf{x}_0 U_1^* \mathbf{x}_1 \ldots U_m^* \mathbf{x}_m$$

where the U_i's are rational subsets of M and the $\mathbf{x}_j$'s are elements of M.

The following result is trivial by structural induction on rational relations of a monoid. Its importance is due to the fact that it implies that intersection with star-chain rational relations suffices to guarantee the **HI** property of a relation.

Proposition 2.7. *Any rational set is a finite union of star-chain rational sets.*

The simplest example of star-chain rational set is when all U_i's are singleton sets.

Definition 2.8. Let $m \geq 1$. A subset of M is an m-loop if, for some elements $\mathbf{x}_0, \ldots, \mathbf{x}_m, \mathbf{u}_1, \ldots, \mathbf{u}_m$ in M, it is of the form

$$\mathbf{x}_0 \mathbf{u}_1^* \mathbf{x}_1 \ldots \mathbf{x}_{m-1} \mathbf{u}_m^* \mathbf{x}_m \ .$$

2.3 Combinatorics of words in a nutshell

We recall some notations and properties about words. The *length* of a word $u \in A^*$ is denoted by $|u|$. The word v is a *factor* of u if there exist x and y such that $u = xvy$. A factor can have several *occurrences* in u, e.g., ba has two occurrences in $ababa$. The word u is *primitive* if the condition $u = w^n$ implies $n = 1$. All nonempty words u are powers of a unique primitive word, called its *primitive root* and denoted by $\rho(u)$.

We recall two classical results about commutation and conjugacy of words, cf. Lentin & Schützenberger [7] and Lothaire [8].

Proposition 2.9 (Commutation)**.** *Let $u, v \in A^* \setminus \{1\}$. Then $uv = vu$ if and only if $\rho(u) = \rho(v)$, i.e. there exists $x \in A^* \setminus \{1\}$ and $\alpha, \beta \in \mathbb{N} \setminus \{0\}$ such that $u = x^\alpha$, $v = x^\beta$.*

Two words u, v are *conjugate* if they satisfy any of the conditions of the next lemma. It is stated in a nonusual form which is more suitable for the proof of Lemma 2.12.

Proposition 2.10 (Conjugacy)**.** *For all $u, v \in A^* \setminus \{1\}$ and $w \in A^*$, the following properties are equivalent.*

(i) $\rho(u)w = w\rho(v)$.

(ii) *Equation $u^i w v^j = u^m w v^n$ has some non trivial solution $(i, j) \neq (m, n)$.*

(iii) *There exist two unique words $x, y \in A^*$ and $\alpha, \beta, \gamma \in \mathbb{N}$ such that xy is primitive, $u = (xy)^\alpha$, $v = (yx)^\beta$ and $w = (xy)^\gamma x$.*

The next result is a simple application of Proposition 2.10 to be used to prove Lemma 4.5.

Lemma 2.11. *Let $x, y, u \in A^* \setminus \{1\}$ and $w \in A^*$. If $|x^m| = |u^a|$ and $|y^n| = |u^b|$ with $a, b \geq 1$ and $x^m w y^n$ is a factor of some u^r then $\rho(x)w = w\rho(y)$.*

2.4 Intersection of 2-loop and 3-loop sets

The rational relations are not closed under intersection. As an application to the elementary combinatorial properties of words just recalled, we state general conditions under which the intersection of two rational relations is not rational. Technical though they may look, the next two lemmas are the crux for our main Theorem 5.6.

Lemma 2.12. *Consider a 2-loop relation $\mathbf{x}\mathbf{u}^*\mathbf{z}\mathbf{v}^*\mathbf{t}$ where $\mathbf{x}$, $\mathbf{z}$, $\mathbf{t}$ are in M and $\mathbf{u}$, $\mathbf{v}$ are in $M \setminus \{1\}$. Assume that on some component i we have $u_i \neq 1$, $v_i \neq 1$ and $\rho(u_i)z_i \neq z_i\rho(v_i)$, whereas on some component $j \neq i$ we have $u_j \neq 1$ or $v_j \neq 1$. Then there exists a 3-loop set T such that $\mathbf{x}\mathbf{u}^*\mathbf{z}\mathbf{v}^*\mathbf{t} \cap T$ is not rational.*

Lemma 2.13. *Consider a 3-loop relation $R = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{w}^*\mathbf{t}$ where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are in M and $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are in $M \setminus \{1\}$. Assume that, for some components $i \neq j$, at least one of the following conditions holds:*

(a)	$u_i \neq 1$,	$v_i \neq 1$,	$\rho(u_i)y_i \neq y_i\rho(v_i)$	*and*	$w_j \neq 1$
(b)	$v_i \neq 1$,	$w_i \neq 1$,	$\rho(v_i)z_i \neq z_i\rho(w_i)$	*and*	$u_j \neq 1$
(c)	$u_i \neq 1$,	$w_i \neq 1$,	$\rho(u_i)y_iz_i \neq y_iz_i\rho(w_i)$	*and*	$v_j \neq 1$

Then there exists a 2-loop subset S of R and a 3-loop set T such that $S \cap T$ is not rational.

Proof of Lemma 2.12. The conclusion of the Lemma holds if and only if it holds relative to the projections onto the two components i, j. Thus, it suffices to consider the case $k = 2$, i.e., to work in the monoid $A_1^* \times A_2^*$. Also, without loss of generality we may assume $\mathbf{x} = \mathbf{t} = 1$. Finally, to simplify the notations we suppose $i = 1$, $j = 2$.

Suppose $u_2 \neq 1$ (the case $v_2 \neq 1$ is similar) and consider the 3-loop set

$$T \ = (u_1, 1)^*(z_1, 1)(v_1, u_2)^*(1, z_2)(1, v_2)^* \ = \{(u_1^\alpha z_1 v_1^\beta, u_2^\beta z_2 v_2^\gamma) \mid \alpha, \beta, \gamma \in \mathbb{N}\}.$$

An element of $\mathbf{u}^*\mathbf{z}\mathbf{v}^*$, say $(u_1^i z_1 v_1^j, u_2^i z_2 v_2^j)$ with $i, j \in \mathbb{N}$, is in T if and only if we have

$$u_1^\alpha z_1 v_1^\beta \quad = \quad u_1^i z_1 v_1^j \tag{1}$$
$$u_2^\beta z_2 v_2^\gamma \quad = \quad u_2^i z_2 v_2^j \tag{2}$$

We argue by case study.

Case $v_2 = 1$. Then T is a 2-loop. Since $v_2 = 1$, equation (2) yields $\beta = i$. Using Proposition 2.10, the hypothesis $\rho(u_1)z_1 \neq z_1\rho(v_1)$ implies that equation (1) is equivalent to $\alpha = i$ and $\beta = j$. Thus, $\alpha = \beta = i = j$ and $\mathbf{u}^*\mathbf{z}\mathbf{v}^* \cap T = \{\mathbf{u}^i\mathbf{z}\mathbf{v}^i \mid i \in \mathbb{N}\}$ is not rational.

Case $v_2 \neq 1$ and $\rho(u_2)z_2 \neq z_2\rho(v_2)$. Then equation (2) yields $\beta = i$ and $\gamma = j$. Also, as above, equation (1) yields $\alpha = i$ and $\beta = j$. Again, $i = j$ and $\mathbf{u}^*\mathbf{z}\mathbf{v}^* \cap T = \{\mathbf{u}^i\mathbf{z}\mathbf{v}^i \mid i \in \mathbb{N}\}$ is not rational.

Case $v_2 \neq 1$ and $\rho(u_2)z_2 = z_2\rho(v_2)$. Then Proposition 2.10 insures that $u_2 = (\lambda\mu)^a$, $z_2 = (\lambda\mu)^b\lambda$ and $v_2 = (\mu\lambda)^c$ for some $\lambda, \mu \in A_2^*$ and $a, b, c \in \mathbb{N}$. We consider now the 2-loop relation

$$\begin{aligned} T' \ &= (u_1, (\lambda\mu)^{a+c})^*(z_1, (\lambda\mu)^b\lambda)(v_1, 1)^* \\ &= \{(u_1^\alpha z_1 v_1^\beta, (\lambda\mu)^{\alpha(a+c)+b}\lambda) \mid \alpha, \beta \in \mathbb{N}\} \ . \end{aligned}$$

An element $\mathbf{u}^i\mathbf{z}\mathbf{v}^j = (u_1^i z_1 v_1^j, u_2^i z_2 v_2^j)$ is in T' if and only if

$$u_1^\alpha z_1 v_1^\beta \quad = \quad u_1^i z_1 v_1^j \tag{3}$$
$$\alpha(a + c) + b \quad = \quad ia + b + jc \tag{4}$$

Again, equation (3) yields $\alpha = i$ and $\beta = j$. Which shows that equation (4) reduces to $\alpha c = jc$ and yields $\alpha = j$ since $c \neq 0$ (recall $v_2 \neq 1$). Thus, $i = j$ and $\mathbf{u}^*\mathbf{z}\mathbf{v}^* \cap T' = \{\mathbf{u}^i\mathbf{z}\mathbf{v}^i \mid i \in \mathbb{N}\}$ is not rational. □

Proof of Lemma 2.13. Let $R = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{w}^*\mathbf{t}$. We cannot reduce to the case $k = 2$. Indeed, let π_{ij} be the projection of M onto its components of rank i and

j for all $1 \le i, j \le k$. Then it is not the case that the condition $\pi_{ij}(S) \subseteq \pi_{ij}(R)$ for all $1 \le i, j \le k$ implies $S \subseteq R$. To simplify the notations we assume $i = 1$.

First we treat the case where condition (a) holds: $u_1, v_1 \neq 1$, $\rho(u_1)y_1 \neq y_1\rho(v_1)$ and $w_j \neq 1$ for some $j \neq 1$. We argue by case study on the values of w_1, the u_ℓ's and the v_ℓ's for $\ell \neq 1$. In each case, we find some 2-loop relation $S \subseteq R$ which fulfills the hypothesis of Lemma 2.12.

Case $u_\ell \neq 1$ or $v_\ell \neq 1$ for some $\ell \neq 1$. Then $S = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{t}$ is a 2-loop subset of R which fulfills the hypothesis of Lemma 2.12 (with ℓ in place of j).

Case $w_1 \neq 1$. Observe that $\rho(u_1)y_1 \neq y_1\rho(v_1)$ implies $\rho(u_1)y_1z_1 \neq y_1z_1\rho(w_1)$ or $\rho(v_1)z_1 \neq z_1\rho(w_1)$. Indeed, if both equalities were true we would have $\rho(u_1)y_1z_1 = y_1\rho(v_1)z_1$. Canceling out the suffix z_1 to both handsides yields $\rho(u_1)y_1 = y_1\rho(v_1)$, contradicting the hypothesis on $\mathbf{u}$, $\mathbf{v}$. So one of these equalities fails hence one of the two subsets $\mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{z}\mathbf{w}^*\mathbf{t}$ and $\mathbf{x}\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{w}^*\mathbf{t}$ is a 2-loop subset of R which fulfills the hypothesis of Lemma 2.12 (with $i = 1$ and the j given by the assumed condition (a)).

Case $u_\ell = v_\ell = 1$ for all $\ell \neq 1$ and $w_1 = 1$. Decompose any k-tuple $\mathbf{a} \in M$ as $\mathbf{a} = \mathbf{a}'\mathbf{a}'' = \mathbf{a}''\mathbf{a}'$ where $\mathbf{a}'$ is the k-tuple having all components equal to 1 except the first one equal to the first component a_1 of $\mathbf{a}$. Observe that $\mathbf{a}'$ and $\mathbf{b}''$ commute for all $\mathbf{a}$, $\mathbf{b}$ in M hence $(\mathbf{a}'\mathbf{b}'')^* \subseteq \mathbf{a}'^*\mathbf{b}''^*$. The hypothesis insures that $\mathbf{v}'' = \mathbf{w}' = 1$ hence $\mathbf{v}\mathbf{w} = \mathbf{v}'\mathbf{w}''$. Consider $S = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{z}''(\mathbf{v}'\mathbf{w}'')^*\mathbf{z}'\mathbf{t}$. We have

$$S \subseteq \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{z}''\mathbf{v}'^*\mathbf{w}''^*\mathbf{z}'\mathbf{t} = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}'^*\mathbf{z}''\mathbf{z}'\mathbf{w}''^*\mathbf{t} = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{w}^*\mathbf{t} = R$$

By definition of $\mathbf{z}''$, $\mathbf{w}''$, $\mathbf{v}'$, we have $z_1'' = w_1'' = 1$ and $v_1' = v_1$, hence, using condition (a) we get $\rho(u_1)y_1z_1'' = \rho(u_1)y_1 \neq y_1\rho(v_1) = y_1z_1''\rho(v_1'w_1'')$. Also, again by condition (a), we have $u_1 \neq 1$, $v_1'w_1'' = v_1 \neq 1$ and, for the j in condition (a), $v_j'w_j'' = w_j'' \neq 1$. Thus, the subset S of R is a 2-loop which fulfills the hypothesis of Lemma 2.12.

The case of condition (b) reduces to the previous case by taking the mirror image. Concerning the last case with condition (c), observe similarly as above that $\rho(u_1)y_1z_1 \neq y_1z_1\rho(w_1)$ implies $\rho(u_1)y_1 \neq y_1\rho(v_1)$ or $\rho(v_1)z_1 \neq z_1\rho(w_1)$. Let $S = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{t}$ if $\rho(u_1)y_1 \neq y_1\rho(v_1)$ and $S = \mathbf{x}\mathbf{y}\mathbf{v}^*\mathbf{z}\mathbf{w}^*\mathbf{t}$ if $\rho(v_1)z_1 \neq z_1\rho(w_1)$. Since we have $v_j' \neq 1$, we see that S is a 2-loop included in R which fulfills the hypothesis of Lemma 2.12. □

3 Degenerate relations

This section introduces the first family of rational relations which have the **HI** property.

Definition 3.1. $R \subseteq A_1^* \times \cdots \times A_k^*$ is *degenerate* if at most one component of R is not a singleton set. I.e., there exists $\ell \in \{1, \ldots, k\}$ such that, for all $j \neq \ell$, the projection of R on A_j^* is a singleton.

Proposition 3.2. *A degenerate relation is rational if and only if its unique non singleton component, say the i-th component, is a rational subset of A_i^*.*

By Proposition 2.3, degenerate rational relations are a very special case of recognizable relation with which they share some properties. Actually, they have stronger stability properties than recognizable relations.

Proposition 3.3. *A rational subrelation of a degenerate relation is itself degenerate. In particular, finite unions of rational degenerate relations have the* **HI** *property.*

The following easy Lemma is used in our main theorem to rule out a trivial condition for a rational relation to have the **HI** property. It characterizes the star-chain relations which are degenerate.

Lemma 3.4. *Let $R = \mathbf{x_0}U_1^*\mathbf{x_1}U_2^* \ldots \mathbf{x_{m-1}}U_m^*\mathbf{x_m}$ where the U_ℓ's are subsets of $M = A_1^* \times \cdots \times A_k^*$ and the $\mathbf{x}_\ell = (x_{\ell,1}, \ldots, x_{\ell,k})$'s are in M.*

Then R is a finite union of degenerate relations if and only if it is degenerate if and only if there exists $i \in \{1, \ldots, k\}$ such that for every $p = 1, \ldots, m$ and every $(u_1, \ldots, u_k) \in U_p$, it holds: $u_j = 1$ if $j \neq i$.

4 Quasi-tally relations

This section presents the second family of rational relations which enjoy the **HI** property. The intuition is as follows: if a relation R is included in a subset of the form $u_1^* \times \cdots \times u_k^*$ where $u_1 \in A_1^*, \ldots, u_k \in A_k^*$ then it can be viewed in a natural way as a subset of $\mathbb{N}^k$. Furthermore, for all other relations $S \subseteq A_1^* \times \cdots \times A_k^*$ the intersection $R \cap S$ can also be viewed as the intersection of two subsets of $\mathbb{N}^k$. It then suffices to resort to the fact that rational subsets of $\mathbb{N}^k$ are closed under the Boolean operations, [5] or [3].

4.1 Tally and quasi-tally sets

By Proposition 2.9, any two arbitrary elements of a subset $X \subseteq A_1^* \times \cdots \times A_k^*$ commute if and only if for some $u_1 \in A_1^*, \ldots, u_k \in A_k^*$, X is included in the commutative submonoid $u_1^* \times \cdots \times u_k^*$, i.e., if it is included in a k-loop of the form $\mathbf{u}_1^* \cdots \mathbf{u}_k^*$ where $\mathbf{u}_i$ is a k-tuple having all components equal to 1 except the i-th component. This leads to the following definition.

Definition 4.1. Let $M = A_1^* \times \cdots \times A_k^*$ and $R \subseteq M$.
1. R is *strictly tally* if it is included in $u_1^* \times \cdots \times u_k^*$ for some $(u_1, \ldots, u_k) \in M$.
2. R is *tally* if it is a finite union of strictly tally relations.

We introduce a slight variant of tally relations which behave in the same way relative to the intersection property.

Definition 4.2. Let $M = A_1^* \times \cdots \times A_k^*$ and $R \subseteq M$.
1. R is *strictly quasi-tally* if it is of the form $\mathbf{x}T\mathbf{y}$ for some $\mathbf{x}, \mathbf{y} \in M$ and some strictly tally relation T.
2. R is *quasi-tally* if it is a finite union of strictly quasi-tally relations.

Observe that all 1-loop sets (cf. Definition 2.8) are strictly quasi-tally sets. The converse is false: for instance, with $k = 2$, $A_1 = \{a\}$, $A_2 = \{b\}$, there are two degrees of freedom in the set $a^* \times b^* = \{a^i, b^j) \mid i, j \in \mathbb{N}\}$ whereas there is only one in the 1-loop relation $\{(a, b)\}^* = \{(a^i, b^i) \mid i \in \mathbb{N}\}$. However, every strictly quasi-tally set in $A_1^* \times \cdots \times A_k^*$ is included in a k-loop of the form $\mathbf{x}\mathbf{u_1}^* \cdots \mathbf{u_k}^*\mathbf{y}$ where $\mathbf{u_1}, \ldots, \mathbf{u_k}$ commute as discussed at the beginning of this section.

We focus our attention on the quasi-tally relations that are rational. It should not come as a surprise that they are defined in terms of rational subsets of $\mathbb{N}^k$.

Proposition 4.3. *Let $\mathbf{x} = (x_1, \ldots, x_k)$, $\mathbf{y} = (y_1, \ldots, y_k)$ and $(u_1, \ldots, u_k)$ be three elements of $A_1^* \times \cdots \times A_k^*$. A strictly quasi-tally relation R included in $\mathbf{x}(u_1^* \times \cdots \times u_k^*)\mathbf{y}$ is rational if and only if*

$$\{(i_1, \ldots, i_k) \in \mathbb{N}^k \mid \mathbf{x}(u_1^{i_1}, \cdots, u_k^{i_k})\mathbf{y} \in R\}$$

is rational in the commutative monoid $\mathbb{N}^k$.

Proof. Let $\varphi : \mathbb{N}^k \to M$ be such that $\varphi(i_1, \ldots, i_k) = (u_1^{i_1}, \ldots, u_k^{i_k})$. Since φ is a morphism, if $P \subseteq \mathbb{N}^k$ is in $\mathrm{Rat}(\mathbb{N}^k)$ then Point 3 of Proposition 2.4 insures that $\varphi(P)$ is in $\mathrm{Rat}(M)$. Hence $R = \mathbf{x}\varphi(P)\mathbf{y}$ is also in $\mathrm{Rat}(M)$. Conversely, if $R \in \mathrm{Rat}(M)$ then Proposition 2.5 insures that $S = \mathbf{x}^{-1}R\mathbf{y}^{-1}$ is rational. Since φ is an isomorphism between the monoids $\mathbb{N}^k$ and $u_1^* \times \cdots \times u_k^*$, we have $P = \varphi^{-1}(S) \in \mathrm{Rat}(\mathbb{N}^k)$. $\square$

4.2 Closure properties of quasi-tally relations

Quasi-tally rational relations have remarkable closure properties.

Proposition 4.4. *Let $M = A_1^* \times \cdots \times A_k^*$. The intersection and difference of a rational quasi-tally relation with a rational relation (non necessarily quasi-tally) are quasi-tally rational.*

The complement of a quasi-tally rational relation is rational.

Observe that the complement of a tally relation is never tally except if all A_i's are singleton alphabets (in which case M is isomorphic to $\mathbb{N}^k$).

Proof. Consider first the intersection. Since a quasi-tally relation is a finite union of strictly quasi-tally relations and intersection distributes over union, we suppose R is strictly quasi-tally rational: $R \subseteq \mathbf{x}U\mathbf{y}$ for some $\mathbf{x}, \mathbf{y}$ in M and $U = u_1^* \times \cdots \times u_k^*$. By Proposition 2.3, U is recognizable . Consider $S \in \mathrm{Rat}(M)$. Then $R' = \mathbf{x}^{-1}R\mathbf{y}^{-1} \subseteq U$ is rational by Proposition 2.5. The relation $S' = \mathbf{x}^{-1}S\mathbf{y}^{-1} \cap U$ is rational as an intersection of a rational and a recognizable relation by Proposition 2.4. Clearly, $R \cap S = \mathbf{x}(R' \cap S')\mathbf{y}$. To prove that $R \cap S$ is rational, it suffices to prove that $R' \cap S'$ is rational. Now, R' and S' are rational relations included in the monoid $u_1^* \times \cdots \times u_k^*$. By Proposition 4.3, they correspond to rational relations P, Q in the monoid $\mathbb{N}^k$. Since $\mathrm{Rat}(\mathbb{N}^k)$ is a Boolean algebra, $P \cap Q$ is rational in $\mathbb{N}^k$. Finally we conclude by observing that $P \cap Q$ corresponds to $R' \cap S'$.

Concerning the set difference, since union left distributes over set difference, we again reduce to R strictly quasi-tally rational and argue in a similar way.

For the complement, observe that $M \setminus R = (M \setminus U) \cup (U \setminus R)$. The first term is recognizable as is U hence is rational (cf. Proposition 2.4). The second term is the difference of two quasi-tally rational relations hence is rational. $\square$

4.3 Quasi-tally star-chain languages

Lemma 3.4 characterized the star-chain relations which are degenerate. Here we characterize those that are quasi-tally. Since this property holds if and only if it holds componentwise, it suffices to state it in the case of languages.

Lemma 4.5. *Let $x_0, \ldots, x_m$ be words in A^* and let $U_1, \ldots, U_m$ be non empty subsets of A^*. Consider the star-chain language $L = x_0 U_1^* x_1 U_2^* \ldots x_{m-1} U_m^* x_m$. The following conditions are equivalent*

i) L is strictly quasi-tally

ii) L is quasi-tally

iii) the following conditions are both satisfied:

(a) *$\rho(u) = \rho(v)$ for $p = 1, \ldots, m$ and all elements $u, v \in U_p \setminus \{1\}$,*

(b) *$\rho(u)\ x_p = x_p\ \rho(v)$ for $1 \le p < m$ and all $u \in U_p \setminus \{1\}$ and $v \in U_{p+1} \setminus \{1\}$.*

Proof. i) implies ii). Trivial.
iii) implies i). Let ρ_i be the common root of the elements in $U_i \setminus \{1\}$. Using the equality $\rho_i^* x_i = x_i \rho_{i+1}^*$ a simple induction on $1 \le i \le m$ shows that $x_0 U_1^* x_1 U_2^* \ldots x_{i-1} U_i^* \subseteq x_0 x_1 \ldots x_{i-1} \rho_i^*$. Then $L \subseteq x_0 x_1 \ldots x_{m-1} \rho_m^* x_m$ as claimed.
ii) implies iii). Suppose L is quasi-tally, say $L \subseteq (z_1 w_1^* t_1) \cup \ldots \cup (z_n w_n^* t_n)$. Let e be the least common multiple of the $|w_i|$'s.
Condition (a). Let u, v be in $U_p \setminus \{1\}$ and choose a so that

$$|x_0 \cdots x_{p-1} u^a| > \max(|z_1|, \ldots, |z_n|), \quad |v^a x_p \cdots x_m| > \max(|t_1|, \ldots, |t_n|)\ .$$

Then $(x_0 \cdots x_{p-1}) u^{a+e} v^{e+a} (x_p \cdots x_m)$ is in L hence in some $z_\ell w_\ell^* t_\ell$. By choice of a, we see that $u^e v^e$ is a factor of some w_ℓ^N. Using Lemma 2.11, we conclude that $\rho(u) = \rho(v)$.
Condition (b). Let $u \in U_p \setminus \{1\}$, $v \in U_{p+1} \setminus \{1\}$ and choose a as above. Then $(x_0 \cdots x_{p-1}) u^{a+e} x_p v^{e+a} (x_{p+1} \cdots x_m)$ is in L hence in some $z_\ell w_\ell^* t_\ell$. By choice of a, we see that $u^e x_p v^e$ is a factor of some w_ℓ^N. Using Lemma 2.11, we conclude that $\rho(u) x_p = x_p \rho(v)$. $\square$

5 Hereditary rational intersection property

Before proving our main results in this section we recall the central property.

Definition 5.1. A rational subset R has the *hereditary rational intersection property* if every rational set S included in R has a rational intersection with every rational relation.

Definition 5.2. Let $n \geq 1$. A rational subset R has the *rational n-intersection property* if $R \cap T_1 \cap \ldots \cap T_n$ is rational when $T_1, \ldots, T_n$ are rational.

This property which, for $n \geq 2$, is formally stronger than **HI** happens to coincide with it.

Proposition 5.3. *The following conditions are equivalent*

(1) *R has the hereditary rational intersection property,*

(2) *For all $n \geq 1$, R has the rational n-intersection property,*

(3) *R has the rational 2-intersection property.*

Proof. (1) $\Rightarrow$ (2). By induction of n. It is clear for $n = 1$. Consider now $R \cap T_1 \cap \ldots \cap T_{n+1} = (R \cap T_1 \cap \ldots \cap T_n) \cap T_{n+1}$. By induction hypothesis, $R \cap T_1 \cap \ldots \cap T_n$ is a rational subset of R and thus its intersection with T_{n+1} is again rational.
(2) $\Rightarrow$ (3). Trivial
(3) $\Rightarrow$ (1). If $S \subseteq R$ is rational then for all rational relations T the relation $R \cap S \cap T = S \cap T$ is rational. □

For 2-loop relations, we can strengthen the above equivalence.

Proposition 5.4. *For a 2-loop relation, the following conditions are equivalent*

(1) *$\mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{t}$ has the hereditary rational intersection property,*

(2) *$\mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{t}$ has the rational 1-intersection property,*

(3) *The intersection of $\mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{t}$ with every 3-loop rational set is rational.*

Proof. The sole non trivial implication is (3) $\Rightarrow$ (1). If $R = \mathbf{x}\mathbf{u}^*\mathbf{y}\mathbf{v}^*\mathbf{t}$ is quasitally or degenerate then it satisfies the **HI** property. Suppose R is neither quasitally nor degenerate. A simple application of Lemma 4.5 shows that, on some component i we have $u_i, v_i \neq 1$ and $\rho(u_i)y_i \neq y_i\rho(v_i)$. Since R is not degenerate, on some other component $j \neq i$ we have $u_j \neq 1$ or $v_j \neq 1$. Thus, R satisfies the hypothesis of Lemma 2.12 hence there is a 3-loop relation contradicting the **HI** condition. □

Theorem 5.5. *A star-chain rational subset of $A_1^* \times \cdots \times A_k^*$ has the hereditary rational intersection property if and only if it is degenerate or strictly quasi-tally.*

Proof. $\Leftarrow$ implication. Use Propositions 4.4 and 3.3.
For the $\Rightarrow$ implication, suppose R is neither degenerate nor strictly quasi-tally and is of the form

$$R = \mathbf{x_0} U_1^* \mathbf{x_1} U_2^* \ldots \mathbf{x_{m-1}} U_m^* \mathbf{x_m} \ . \tag{5}$$

We show that there exists some 3-loop subset R' of R satisfying the hypothesis of Lemma 2.13. Fix some i such that the i-th projection of R is not quasi-tally. Let X be the set of p's in expression (5) such that the projection of U_p on A_i^* is not reduced to $\{1\}$. Observe that the projection of R on A_i^* is a star-chain language L where the star sets are the projections of the U_p's with $p \in X$. Since L is not quasi-tally, Lemma 4.5 insures that there exist $p, q \in X$ and $\mathbf{u} \in U_p$, $\mathbf{v} \in U_q$ such that $p \leq q$ and, letting $\mathbf{z} = \mathbf{x_p} \cdots \mathbf{x_{q-1}}$ (which is equal to 1 by convention in case $p = q$), we have $u_i \neq 1$, $v_i \neq 1$ and $\rho(u_i)\ z_i \neq z_i\ \rho(v_i)$.

Now, since R is not degenerate, there exists some r among $1, \ldots, m$ and some $\mathbf{w} \in U_r$ such that $w_j \neq 1$ for some $j \neq i$ (cf. Lemma 3.4).

Define a 3-loop relation R' included in R as follows:

$$R' = \begin{cases} \mathbf{x_0} \cdots \mathbf{x_{p-1}} \mathbf{u}^* \mathbf{x_p} \cdots \mathbf{x_{q-1}} \mathbf{v}^* \mathbf{x_q} \cdots \mathbf{x_{r-1}} \mathbf{w}^* \mathbf{x_r} \cdots \mathbf{x_m} & \text{if } q \leq r \qquad (\alpha) \\ \mathbf{x_0} \cdots \mathbf{x_{r-1}} \mathbf{w}^* \mathbf{x_r} \cdots \mathbf{x_{p-1}} \mathbf{u}^* \mathbf{x_p} \cdots \mathbf{x_{q-1}} \mathbf{v}^* \mathbf{x_q} \cdots \mathbf{x_m} & \text{if } r \leq p \qquad (\beta) \\ \mathbf{x_0} \cdots \mathbf{x_{p-1}} \mathbf{u}^* \mathbf{x_p} \cdots \mathbf{x_{r-1}} \mathbf{w}^* \mathbf{x_r} \cdots \mathbf{x_{q-1}} \mathbf{v}^* \mathbf{x_q} \cdots \mathbf{x_m} & \text{if } p \leq r \leq q \qquad (\gamma) \end{cases}$$

Observe that in case (α) (resp. (β), (γ)) the set R' is a 3-loop relation included in R which satisfies condition (a) (resp. (b), (c)) of Lemma 2.13. We conclude by applying Lemma 2.13. □

Theorem 5.6. *Let R be rational subset of $A_1^* \times \cdots \times A_k^*$. The following conditions are equivalent.*

1. *R is the union of a quasi-tally relation and finitely many degenerate relations,*
2. *R has the hereditary rational intersection property,*
3. *For all 2-loop subsets S of R and all 3-loop relations T, the intersection $S \cap T$ is rational.*

Proof. (2) ⇒ (3) is trivial. For (1) ⇒ (2), use Propositions 4.4 and 3.3.
(3) ⇒ (1). By Proposition 2.7, R is a finite union of star-chain rational relations. If condition (3) holds for R then it holds for each star-chain rational relation in the union. Thus, the implication follows from Theorem 5.5, which completes the proof. □

References

[1] D. Angluin and D. N. Hoover. Regular prefix relations. *Mathematical Systems Theory*, 17(3):167–191, 1984.

[2] S. Eilenberg, C. C. Elgot, and J.C. Shepherdson. Sets recognized by n-tape automata. *Journal of Algebra*, 13(4):447–464, 1969.

[3] S. Eilenberg and M.-P. Schützenbeger. Rational sets in commutative monoids. *Journal of Algebra*, 13(2):173–191, 1969.

[4] C. C. Elgot and J. E. Mezei. On Relations Defined by Generalized Finite Automata. *IBM Journal of Research and Development*, 9(1):47–68, 1965.

[5] S. Ginsburg and Spanier. Bounded ALGOL-like languages. *Transactions of the American Mathematical Society*, 113(1):333–368, 1964.

[6] H. Läuchli and C. Savioz. Monadic second order definable relations on the binary tree. *Journal of Symbolic Logic*, 52(1):219–226, 1987.

[7] A. Lentin and M.-P. Schützenberger. A combinatorial problem in the theory of free monoids. In R. C. Bose and T. E. Bowlings, editors, *Combinatorial Mathematics*, pages 112–144. North Carolina Press, Chapel Hill, N. C., 1967.

[8] Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1983.

[9] M. Rabin and D. Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3(1):114–125, 1959.

[10] J. Sakarovitch. *Elements of Automata Theory*. Cambridge University Press, 2009. (English translation of *Eléments de théorie des automates*, Vuibert, Paris, 2003).