Channel synthesis revisited

Béatrice Bérard[1] and Olivier Carton[2]

[1] Université Pierre & Marie Curie, LIP6/MoVe, CNRS, Paris, France
[2] Université Paris Diderot, LIAFA, CNRS, Paris, France

Abstract. Given a system modeled by a rational relation R, a channel is a pair (E, D) of rational relations that respectively encode and decode binary messages, and such that the composition ERD is the identity relation. This means that the message between E and D has been perfectly transmitted through R. Investigating the links between channels and the growth of rational sets of words, we give new characterizations for relations with channels. In the particular case where the relation is given as a union of functions, we obtain as a consequence the decidability of the synthesis problem with a linear complexity.

1 Introduction

Channel synthesis. The problem of channel synthesis was introduced in [1, 2] as a special case of the general distributed synthesis problem: Given an architecture defined by processes and communication links between them or with the environment, and a specification on the messages transmitted over these links, this general problem aims at deciding the existence of local programs, one for each process, that together meet the specification, whatever the environment does.

In the asynchronous setting, this problem is undecidable for LTL specifications as soon as there are two processes [3]. It was then proposed in [1, 2] to consider two processes modeled by finite transducers, that respectively encode and decode finite binary messages. They communicate asynchronously through a medium, acting as noise over the link between them and also described by a fixed non deterministic finite transducer. Moreover, a particular basic external specification expresses faithful communication: the message received is equal to the message emitted. Such an encoder/decoder pair was called a *channel*. The *channel synthesis problem* then asks if, given the noisy process, the encoder and decoder can be synthesized. This question is related to security properties: when the noisy process describes some protocol, the existence of a channel may lead to possibly illegal communication [4–6]. The problem was proved undecidable (Σ_1^0-complete) for rational relations, but decidable in polynomial time for a rational function. When a channel exists for such a function, it can be effectively computed.

Contribution. We revisit here this notion of channel and show that it has strong links with rational bijections [7], hence it is also related to the growth of languages. Given a language L, the growth function associates with an integer $n \geq 0$

the number of words in L of length less than or equal to n. We introduce the notion of *patterns*, which generate typical languages of exponential growth, and establish some of their properties. Then, combining these properties with results on rational bijections, we prove a new characterization for bounded relations with channels: If R is a bounded rational relation, given as a union of rational functions $h_1 + \cdots + h_n$, then the following conditions are equivalent: (1) R has a channel, (2) at least one of the h_is has a channel, (3) the range of R has an exponential growth. We obtain as a corollary that the channel synthesis problem is decidable in linear time for a finite union of functions. The latter result was already stated in [8] (with a polynomial time complexity), but the proof was not satisfactory. We believe that the notion of exponential growth is central to the study of channels, although it was often implicit in previous works.

2 Definitions and notations

The set of natural numbers is denoted by $\mathbb{N}$ and the set of *words* over a finite alphabet A is denoted by A^*, with ε for the empty word and $A^+ = A^* \setminus \{\varepsilon\}$. The binary alphabet $\{0, 1\}$ is denoted by $\mathbb{B}$. The length of a word u is written $|u|$ and for $n \in \mathbb{N}$, we denote by $A^{\leq n}$ (respectively A^n) the subset of A^* of words of length less than or equal to n (respectively equal to n). A *language* is a subset of A^*. We denote also by $|L|$ the cardinality of a language L.

For two words u and v, v is a *prefix* of u, written $v \preccurlyeq u$, if there is some word w such that $u = vw$. Two words w and w' are called *prefix compatible* if either w is a prefix of w' or w' is a prefix of w. If they are not prefix compatible, there exist two different letters a and b and three words u, v and v' such that $w = uav$ and $w' = ubv'$. The word u is, by definition, the *longest common prefix* of w and w'. The notions of *suffix compatible* words and *longest common suffix* are defined similarly.

A subset X of A^* is a code if any word in X^* admits a unique decomposition over X. A set of two words $X = \{u, v\}$ is not a code if and only if u and v commute ($uv = vu$) ([9]).

Finite automata. A finite automaton, or automaton for short, is a tuple $\mathcal{A} = \langle Q, I, Lab, \Delta, F\rangle$, where Q is a finite set of states, $I \subseteq Q$ is the subset of initial states, Lab is a finite set of labels, $\Delta \subseteq Q \times Lab \times Q$ is a finite transition relation and $F \subseteq S$ is a of final states. Note that Lab can be an alphabet but also a (subset of a) monoid. Given two states $q, q' \in Q$, a *path* from q to q' with *label* u, written as $q \xrightarrow{u} q'$, is a sequence of transitions $q \xrightarrow{a_1} q_1, q_1 \xrightarrow{a_2} q_2, \cdots q_{n-1} \xrightarrow{a_n} q'$, with $a_i \in Lab$ and $q_i \in Q$, for $1 \leq i \leq n-1$ such that $u = a_1 \cdots a_n$. The path is *accepting* if $q \in I$ and $q' \in F$, and the language of $\mathcal{A}$, denoted by $\mathcal{L}(\mathcal{A})$, is the set of labels of accepting paths. A state $q \in Q$ is *useful* if it occurs in some accepting run. Since the accepted language is the same when removing non useful states, we assume in the sequel that the set Q contains only useful states, in which case $\mathcal{A}$ is called *trim*. A regular language over an alphabet A is a subset of A^* accepted by a finite automaton with set of labels $Lab = A$. The regular languages over A are also the rational sets of A^*.

Finite Transducers. A finite transducer (or transducer for short) is a finite automaton $\mathcal{T}$ with set of labels $Lab \subseteq A^* \times B^*$ for two alphabets A and B. A label $(u, v) \in A^* \times B^*$ is also written as $u|v$. The subset $\mathcal{L}(\mathcal{T})$ of $A^* \times B^*$ is a *rational relation* [10] from A^* to B^*. The transducer $\mathcal{T}$ is said to realize the relation $\mathcal{L}(\mathcal{T})$.

Given a rational relation R, we write $R(u) = \{v \in B^* \mid (u, v) \in R\}$ for the image of $u \in A^*$, $R^{-1}(v) = \{u \in A^* \mid (u, v) \in R\}$ for the inverse image of $v \in B^*$, possibly extended to subsets of A^* or B^* respectively, $\mathrm{dom}(R) = \{u \in A^* \mid \exists v \in B^*, (u, v) \in R\}$ for the domain of R and $\mathrm{rg}(R) = \{v \in B^* \mid \exists u \in A^*, (u, v) \in R\}$ for the range of R.

For a subset P of A^*, the identity relation $\{(u, u) \mid u \in P\}$ on $A^* \times A^*$ is denoted by Id_P. The composition of rational relations R_1 on $A^* \times B^*$ and R_2 on $B^* \times C^*$, denoted by $R_1 R_2$ (from left to right) or by $R_2 \circ R_1$ (from right to left), is the rational relation on $A^* \times C^*$ defined by $\{(u, w) \mid \exists v\ (u, v) \in R_1 \wedge (v, w) \in R_2\}$ ([11]). Moreover, the image and inverse image of a regular language by a rational relation are regular languages [10].

The relation R is bounded if there exists $k \in \mathbb{N}$ such that for each word $u \in A^*$, $|R(u)| \leq k$. It is a function if $k = 1$. We often identify a function f with its graph $\{(x, f(x) \mid x \in \mathrm{dom}(f)\}$ and we write $f \subseteq R$ (resp. $f \subseteq f'$) to mean that its graph is contained in R (resp. in the graph of f'). We also write $R + f$ to mean the relation which is the union of R and the graph of f. Since functions play a central role in the rest of the paper, we recall below two powerful and useful results. The first one states that it is always possible to extract a rational function from a rational relation. The second one gives a representation of a bounded rational relation as a union of rational functions.

Theorem 1 (Uniformization [10]). *For any rational relation R, there exists a rational function $f \subseteq R$ with the same domain. Furthermore, a transducer realizing f can be effectively computed from a transducer realizing R.*

Theorem 2 (Bounded relations [12, 13]). *A rational relation R is bounded by k if and only if there exist k rational functions $f_1, \ldots, f_k$ such that $R = f_1 + \cdots + f_k$.*

Growth. The growth of a language L is the function mapping each integer $n \in \mathbb{N}$ to the number of words in L of length less than or equal to n. The growth of L is *polynomial* if it is bounded by some polynomial, that is $|L \cap A^{\leq n}| = O(n^k)$ for some $k \in \mathbb{N}$. For instance, the growth of $P_k = (0^*1)^k 0^*$ is polynomial since $|P_k \cap \mathbb{B}^{\leq n}| = O(n^{k+1})$. The growth of the set $L = (0 + 10)^*$ is not polynomial.

The growth of a set L is *exponential* if it is greater that some exponential, that is if $\theta^n = O(|L \cap A^{\leq n}|)$ for some real number $\theta > 1$. The growth of the set $L = (0 + 10)^*$ is, for instance, exponential. Note that the growth cannot be more than exponential since $|A^{\leq n}| = (|A|^{n+1} - 1)/(|A| - 1)$, for $|A| \geq 2$. The following proposition states that, for rational sets, there is a gap:

Proposition 1 ([7]). *The growth of a rational set of words is either polynomial or exponential. Moreover, for a finite automaton $\mathcal{A}$, the language $\mathcal{L}(\mathcal{A})$ has an*

exponential growth if and only if there exist words $u, v, \bar{v}, w$ with $|v| = |\bar{v}|$ and $v \neq \bar{v}$, and a state q of $\mathcal{A}$ such that $i \xrightarrow{u} q \xrightarrow{v} q \xrightarrow{w} f$ and $q \xrightarrow{\bar{v}} q$ in $\mathcal{A}$ where i is an initial state and f is a final state.

This result suggests the notion of patterns studied in details in Section 4, and defined as tuples of words $(u, v, \bar{v}, w)$ with $|v| = |\bar{v}|$ and $v \neq \bar{v}$.

An automaton with ε-transitions is an automaton in which any transition has either the form $p \xrightarrow{a} q$ for a letter $a \in A$ or the form $p \xrightarrow{\varepsilon} q$. It is well-known [14] that ε-transitions can be removed and that any automaton with ε-transitions is equivalent to an automaton without ε-transition. This latter transformation may however introduce a quadratic blow-up of the number of transitions. The following proposition states that it can be directly checked, without removing ε-transitions, whether the language accepted by an automaton has an exponential growth (the result is more or less part of folklore but a proof is given in [15]).

Proposition 2. *It can be checked in linear time whether the language accepted by an automaton with ε-transitions has an exponential growth.*

The *degree* $\deg(L)$ of a set L with a polynomial growth is the least integer k such that $|L \cap A^{\leq n}| = O(n^k)$. The degree of the set $P_k = (0^*1)^k 0^*$ is, for instance, $\deg(P_k) = k + 1$. Note that a rational set L is finite whenever $\deg(L) = 0$. If L is exponential, we set $\deg(L) = \infty$. The following theorem characterizes the existence of a rational bijection between two rational sets.

Theorem 3 ([7]). *There exists a rational bijection between two rational sets L and L' if and only if either they are both finite, (that is $\deg(L) = \deg(L') = 0$) and they have the same cardinality or they are both infinite and $\deg(L) = \deg(L')$.*

When the languages L and L' are infinite, the relation $\deg(L) = \deg(L')$ should be understood as either their growths are both exponential, that is $\deg(L) = \deg(L') = \infty$, or their growth are both polynomial with the same degree $\deg(L) = \deg(L') < \infty$.

3 Channels

A channel is a way to achieve reliable communication between two processes, an encoder and a decoder, via a noisy medium modeled by a non deterministic transducer with labels in $A^* \times B^*$. The encoder E reads binary input and produces an output in A^*, while the decoder D reads words in B^* and produces a binary word. The pair (E, D) is a channel if the binary message is correctly transmitted:

Definition 1. *Let $R \subseteq A^* \times B^*$ be a rational relation. A* rational channel *(or* channel *for short) for R is a pair (E, D) of rational relations in $\mathbb{B}^* \times A^*$ and $B^* \times \mathbb{B}^*$ respectively such that $ERD = \mathrm{Id}_{\mathbb{B}^*}$.*

As a consequence of this definition, a rational relation R has a channel if and only if R^{-1} has a channel.

Examples. We set here $A = B = \mathbb{B}$.

The prefix relation $\text{Pref} = \{(u, v) \mid u \preccurlyeq v\}$ has a rational channel with $E = \{(0, 00), (1, 11)\}^*(\varepsilon, 01)$ and $D = \{(00, 0), (11, 1)\}^*(01, \varepsilon)$, as illustrated in Figure 1.

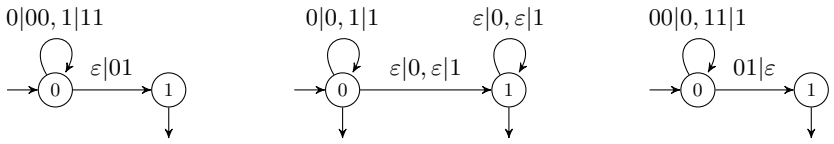

Fig. 1. From left to right: E, Pref and D.

The relation $\text{Diff}_1 = \{(u, v) \mid u \text{ and } v \text{ differ by at most 1 bit}\}$ has a rational channel where 0 and 1 are encoded respectively by 000 and 111 in E. A first decoder D consists in a majority choice (this would be a subcase of channels with substitution in [2]), associating 0 with 000, 001, 010 and 100, and similarly for 1. Another decoder can also be simply the inverse of E, hence a "sub-decoder" of D, ignoring the substitution.

Note that in the definition above, E and D must be rational relations. However, there exist some relations without rational channels but with a channel satisfying the relation $ERD = \text{Id}_{\mathbb{B}^*}$. Let f be a bijection from $\mathbb{B}^*$ onto $\mathbb{N}$ and let R be the relation defined by $R = \{(u, v) \mid |u| = |v|\}$. Then for $E = \{(u, 0^{f(u)}) \mid u \in \mathbb{B}^*\}$ and $D = \{(0^{f(u)}, u) \mid u \in \mathbb{B}^*\}$, the inverse of E, (E, D) satisfies $ERD = \text{Id}_{\mathbb{B}^*}$ but the characterization proved in Section 5 shows that R has no rational channel.

In the sequel, we only consider rational channels. The main result of this paper is the following:

Theorem 4. *Let $R = h_1 + \cdots + h_n$ be a bounded relation where each h_i is a rational function. The following statements are equivalent:*

1. *R has a channel,*
2. *At least one function h_i has a channel,*
3. *$\text{rg}(R)$ has an exponential growth.*

Using the the equivalence between 1 and 3 and Proposition 2, we obtain the decidability of the channel synthesis problem:

Corollary 1. *The channel existence problem for bounded relations, given as a union of functions $h_1 + \cdots + h_n$, is decidable in linear time. When it exists, the channel can be effectively computed.*

4 Patterns

It can be observed in the examples above that, when the encoder/decoder pair (E, D) exists, E can be chosen as a bijection from $\mathbb{B}^*$ onto some language L of the form $u(v + \bar{v})^* w$, with D the inverse of E. In order to generalize this observation, and in relation with Proposition 1, we now introduce the notion of pattern:

Definition 2. *A* pattern *is a 4-tuple $s = (u, v, \bar{v}, w)$ of words such that $|v| = |\bar{v}|$ and $v \neq \bar{v}$. The language associated with s is $L_s = u(v + \bar{v})^* w$. A* sub-pattern *of s is a pattern of the form $(ux, y, \bar{y}, zw)$ where $x, y, \bar{y}, z \in (v + \bar{v})^*$. Two patterns s and s' of the form $s = (u, xv, x\bar{v}, xw)$ and $s' = (ux, vx, \bar{v}x, w)$ (or $s' = (ux, \bar{v}x, vx, w)$) are called* conjugated.

Note that in a pattern $s = (u, v, \bar{v}, w)$, the set $\{v, \bar{v}\}$ is a code hence, the notion of pattern can be seen as the basic element for the canonical channels and encoding states or nodes defined in [5, 6, 2].

If s' is a sub-pattern of s, then the inclusion $L_{s'} \subseteq L_s$ holds. Moreover, if the two patterns s and s' are conjugated, then the languages L_s and $L_{s'}$ are equal. For a pattern $s = (u, v, \bar{v}, w)$, we denote by μ_s the morphism from $\mathbb{B}^*$ to $(v + \bar{v})^*$ which maps 0 to v and 1 to $\bar{v}$. If $t = (x, y, \bar{y}, z)$ is a pattern over the alphabet $\mathbb{B}$, then $(u\mu_s(x), \mu_s(y), \mu_s(\bar{y}), \mu_s(z)w)$ is a sub-pattern of s which we denote by $s \diamond t$. Note that this composition of patterns is associative. Indeed, if $t = (x, y, \bar{y}, z)$ and $t' = (x', y', \bar{y}', z')$ are two patterns over the alphabet $\mathbb{B}$, then $s \diamond (t \diamond t') = (s \diamond t) \diamond t'$.

The proofs of the next lemmas are omitted and can be found in [15].

Lemma 1. *Let s be a pattern and L a rational set of words. There exists a sub-pattern s' of s such that either $L_{s'} \subseteq L$ or $L_{s'} \cap L = \varnothing$.*

Lemma 2. *Let s be a pattern and let L be a rational set accepted by an automaton $\mathcal{A}$. If $L_s \subseteq L$, there exist a sub-pattern $s' = (u', v', \bar{v}', w')$ of s and paths $i \xrightarrow{u'} q \xrightarrow{v'} q \xrightarrow{w'} f$ and $q \xrightarrow{\bar{v}'} q$ in $\mathcal{A}$ where i is an initial state and f is a final state.*

The following proposition shows that if two patterns are not conjugated, one of them can be replaced by one of its sub-pattern to make the associated languages disjoint. In particular, two patterns are conjugated if and only if their associated languages are equal.

Proposition 3. *If the patterns s and s' are not conjugated, then either there exists a sub-pattern s'' of s such that $L_{s''} \cap L_{s'} = \varnothing$ or there exists a sub-pattern s'' of s' such that $L_{s''} \cap L_s = \varnothing$.*

Proof. Let s and s' be the two patterns $(u, v, \bar{v}, w)$ and $(u', v', \bar{v}', w')$. Let k, k', m and m' be the integers defined by $k = |uw|$, $k' = |u'w'|$, $m = |v| = |\bar{v}|$ and $m' = |v'| = |\bar{v}'|$. Let M_s and $M_{s'}$ be the sets $\{|x| \mid x \in L_s\}$ and $\{|x| \mid x \in L_{s'}\}$ of lengths of words in L_s and $L_{s'}$. These two sets M_s and $M_{s'}$ are respectively contained in the sets $k + m\mathbb{N}$ and $k' + m'\mathbb{N}$.

We first suppose that $m \neq m'$ and by symmetry we can assume that $m < m'$. We then consider two sub-cases depending on whether $k \equiv k' \mod m'$ or not. We first suppose that $k \not\equiv k' \mod m'$. The two sets $k + mm'\mathbb{N}$ and $k' + m'\mathbb{N}$ are then disjoint. It follows that the sub-pattern $s'' = (u, v^{m'}, \bar{v}^{m'}, w)$ of s satisfies $L_{s''} \cap L_{s'} = \varnothing$. We now suppose $k \equiv k' \mod m'$. Since $m < m'$, $k + m \not\equiv k' \mod m'$ holds and the two sets $k + m + mm'\mathbb{N}$ and $k' + m'\mathbb{N}$ are then disjoint. It follows that the sub-pattern $s'' = (uv, v^{m'}, \bar{v}^{m'}, w)$ of s satisfies $L_{s''} \cap L_{s'} = \varnothing$.

From now on, we suppose that $m = m'$. If $k \not\equiv k' \mod m$, the two sets M_s and $M_{s'}$ are already disjoint and we can set $s'' = s$. From now on, we also suppose that $k \equiv k' \mod m$.

If uv and u' are not prefix compatible, the sub-pattern $s'' = (uv, v, \bar{v}, w)$ satisfies $L_{s''} \cap L_{s'} = \varnothing$. A similar solution can be found if $u\bar{v}$ and u' are not prefix compatible or if w and $v'w'$ are not suffix compatible. We can now suppose that uv and u' (resp. $u\bar{v}$ and u', u and $u'v'$, u and $u'\bar{v}'$) are prefix compatible and that vw and w' (resp. $\bar{v}w$ and w', w and $v'w'$, w and $\bar{v}'w'$) are suffix compatible.

Since u and $u'v'$ are compatible and u and $u'\bar{v}'$ are also compatible, the word u is a prefix of $u'z$ where z is the longest common prefix of v' and $\bar{v}'$ and thus $|u| \leq |u'| + |z|$. By a similar argument, the word w satisfies $|w| \leq |w'| + |z'|$ where z' is the longest common suffix of v' and $\bar{v}'$. Combining these two relations gives $k = |uw| \leq k' + |zz'| < k' + m$ where the relation $|zz'| < m$ follows from $v' \neq \bar{v}'$. By symmetry the relation $k' \leq k + m$ also holds and this implies $k = k'$ since $k \equiv k' \mod m$.

We now have $m = m'$ and $k = k'$. We can assume by symmetry that $|u| \leq |u'|$ and thus $|w'| \leq |w|$. Since u and u' are prefix compatible, the word u is a prefix of u'. There exists a word x such that $u' = ux$. By symmetry, the word w' is a suffix of w and there exists a word x' such that $w = x'w'$. Since u' and uv are prefix compatible and u' and $u\bar{v}$ are also prefix compatible, the word x is a prefix of v and $\bar{v}$. There exist two words z and $\bar{z}$ such that $v = xz$ and $\bar{v} = x\bar{z}$. By symmetry, the word x' is a suffix of v' and $\bar{v}'$ and there exist two words z' and $\bar{z}'$ such that $v' = z'x'$ and $\bar{v}' = \bar{z}'x'$. Note that $|x| = |x'| = |u'| - |u| = |w| - |w'|$ and that $|z| = |\bar{z}| = |z'| = |\bar{z}'| = m - |x|$.

We first suppose that $x \neq x'$. Let s'' be the sub-pattern $(u'v'^2, v', \bar{v}', w')$ of s'. Any word in $L_{s''}$ starts with $u'v'^2 = uxz'x'z'x'$ whereas any word of L_s is either shorter or starts with a prefix in $u(v + \bar{v})^2 = ux(z + \bar{z})x(z + \bar{z})$. This shows that $L_{s''} \cap L_s = \varnothing$. We now suppose that $x = x'$. If the two sets $\{z, \bar{z}\}$ and $\{z', \bar{z}'\}$ are equal, the two patterns s and s' are conjugated and this a contradiction with the hypothesis. If these two sets are different, we can assume by symmetry that $z \notin \{z', \bar{z}'\}$. The sub-pattern $(uv, v, \bar{v}, w)$ of s satisfies then $L_{s''} \cap L_{s'} = \varnothing$.

5 Channel characterizations

Recall that in both examples of section 3, the encoder and decoder can be chosen as rational bijections. The following characterization generalizes this observation.

Proposition 4. *There is a channel for a relation R if and only there exist two rational sets L_0 and L_1 with exponential growth such that $R \cap (L_0 \times L_1)$ is a bijection between L_0 and L_1.*

Note that it is assumed, in the previous proposition, that both L_0 and L_1 have an exponential growth. It is actually sufficient to assume that only one of them has. Theorem 3 and the fact that $R \cap (L_0 \times L_1)$ is a bijection between L_0 and L_1 ensure that the other one also has an exponential growth.

Proof. We first prove that the condition is sufficient. Suppose that $R \cap (L_0 \times L_1)$ is a bijection between L_0 and L_1. Since L_0 has an exponential growth, there exists, by Theorem 3, a rational bijection E between $\{0,1\}^*$ and L_0. The relation ER is thus a bijection between $\{0,1\}^*$ and L_1. Set $D = (ER)^{-1}$. It then clear than $E\ R\ D = \mathrm{Id}_{\mathbb{B}^*}$.

Suppose now that there are two rational relations E and D such that $ERD = \mathrm{Id}_{\mathbb{B}^*}$. We first claim that there exists a function $E' \subseteq E$ from $\{0,1\}^*$ to A^* and another function $D' \subseteq D$ from B^* to $\{0,1\}^*$ such that $E'RD' = \mathrm{Id}_{\mathbb{B}^*}$. Let K_1 be the rational set $\mathrm{rg}(ER) = \{v \mid \exists u \in \{0,1\}^*\ (u,v) \in ER\}$ and D' the restriction $D' = D \cap (K_1 \times \{0,1\}^*)$. It is clear that $ERD' = \mathrm{Id}_{\mathbb{B}^*}$ and that D' must be functional. Let K_0 be the set $\mathrm{dom}(RD') = \{u \mid \exists v \in \{0,1\}^*\ (u,v) \in RD'\}$ and let E'' be the restriction $E'' = E \cap (\{0,1\}^* \times K_1)$. The relation E'' might not be functional but there exists by Theorem 1 a rational function $E' \subseteq E''$. It is also clear that $E'RD' = \mathrm{Id}_{\mathbb{B}^*}$.

We now suppose that E and D are two functions. Applying the reasoning to E^{-1} and D^{-1} there are two relations $E' \subseteq E$ and $D' \subseteq D$ such that E'^{-1} and D'^{-1} are functions and $E'RD' = \mathrm{Id}_{\mathbb{B}^*}$. Let L_0 and L_1 be the sets $L_0 = \mathrm{dom}(RD') = \{u \mid \exists v \in \{0,1\}^*\ (u,v) \in RD'\}$ and $L_1 = \mathrm{rg}(E'R) = \{v \mid \exists u \in \{0,1\}^*\ (u,v) \in E'R\}$. The relation E' is then a bijection between $\{0,1\}^*$ and L_0 and the relation D' is a bijection between L_1 and $\{0,1\}^*$. It follows that $R \cap (L_0 \times L_1)$ must be a bijection between L_0 and L_1.

For two patterns $s = (u, v, \bar{v}, w)$ and $s' = (u', v', \bar{v}', w')$, we denote by $h_{s,s'}$ the function whose graph is the rational relation $(u,u')((v,v') + (\bar{v},\bar{v}'))^*(w,w')$. Note that this function is a bijection from L_s to $L_{s'}$ and that the inverse function $h_{s,s'}^{-1}$ is actually the function $h_{s',s}$. Let s_0 be the pattern $(\varepsilon, 0, 1, \varepsilon)$. The set L_{s_0} is then the set $\mathbb{B}^*$ and $h_{s_0,s}$ is a bijection from $\mathbb{B}^*$ to L_s. Note finally that if t is a pattern over the alphabet $\mathbb{B}$ the restriction of the function $h_{s,s'}$ to the domain $L_{s \diamond t}$ is the function $h_{s \diamond t, s' \diamond t}$ from $L_{s \diamond t}$ to $L_{s' \diamond t}$.

Lemma 3. *Let h be a rational function such that $\mathrm{rg}(h)$ has an exponential growth. Then there exist two patterns s and s' such that $h_{s,s'} \subseteq h$.*

Proof. Let $\mathcal{T}$ be a transducer realizing the function h. Let $\mathcal{A}$ be the automaton obtained by ignoring the input label of each transition of $\mathcal{T}$ and taking the output label as the label. This automaton accepts the set $\mathrm{rg}(h)$. By Proposition 1 there exists a pattern $s' = (u', v', \bar{v}', w')$ and paths $i \xrightarrow{u'} q \xrightarrow{v'} q \xrightarrow{w'} f$ and $q \xrightarrow{\bar{v}'} q$ in $\mathcal{A}$ where i is an initial state and f is a final state. Since these paths come from

paths in $\mathcal{T}$, there are words u, v_0, $\bar{v}_0$ and w and paths $i \xrightarrow{u|u'} q \xrightarrow{v_0|v'} q \xrightarrow{w|w'} f$ and $q \xrightarrow{\bar{v}_0|\bar{v}'} q$ in $\mathcal{T}$. Note however that the words v_0 and $\bar{v}_0$ may not have the same length. Let v and $\bar{v}$ be the words $v_0\bar{v}_0$ and $\bar{v}_0v_0$. These words have the same length but they are different. Otherwise h maps the single word $uv_0\bar{v}_0$ to the two differents words $u'v'\bar{v}'w'$ and $u'\bar{v}'v'w'$. The function $h_{s,s''}$ where $s = (u, v, \bar{v}, w)$ and $s'' = (u', v'\bar{v}', \bar{v}'v', w')$ satisfies then $h_{s,s''} \subseteq h$.

The proof of Theorem 4 proceeds by induction on the number of functions, so we first establish the result for a single function.

Proposition 5. *If h is a function, then h has a channel if and only if* $\mathrm{rg}(h)$ *has an exponential growth.*

Proof. By Proposition 4, the condition is necessary. Indeed, if h has a channel, there exist two languages L_0 and L_1 with an exponential growth such that $h \cap (L_0 \times L_1)$ is a bijection from L_0 to L_1. The set L_1 is thus contained in $\mathrm{rg}(h)$ and $\mathrm{rg}(h)$ has a exponential growth.

By Lemma 3, there are two patterns s and s' such that $h_{s,s'} \subseteq h$. Since h is a function, $h \cap (L_s \times L_{s'}) = h_{s,s'}$. By Proposition 4, the function h has a channel.

The next lemma, which is one of the key ingredients for the main result, was present in [8] (with an unsatisfactory proof).

Lemma 4. *Let R be a rational relation and let h be a rational function. If R has a channel, then $R + h$ also has a channel.*

Proof. If R has a channel, by Proposition 4, there exist two languages L_0 and L_1 with an exponential growth such that $g = R \cap (L_0 \times L_1)$ is a bijection from L_0 to L_1. Applying Lemma 3 to g, there exist two patterns $s = (u, v, \bar{v}, w)$ and $s' = (u', v', \bar{v}', w')$ such that $h_{s,s'} \subseteq g$. Let L be the domain of the function h. By Lemma 1, s can be replaced by one of its sub-patterns such that either $L_s \cap L = \varnothing$ or $L_s \subseteq L$. Like in the previous proof, $g \cap (L_s \times L_{s'}) = h_{s,s'}$. If $L_s \cap L = \varnothing$ then $(R + h) \cap (L_s \times L_{s'}) = h_{s,s'}$ and the function $h_{s,s'}$ provides a channel for $R + h$.

We now suppose that s satisfies $L_s \subseteq L$. Let $\mathcal{T}$ be a transducer realizing the function h and let $\mathcal{A}$ be the automaton obtained by ignoring the output label of each transition of $\mathcal{T}$ and taking the input label as the label. This automaton accepts the set $L = \mathrm{dom}(h)$. By Lemma 2, s can be replaced by one its sub-patterns such that there are paths $i \xrightarrow{u} q \xrightarrow{v} q \xrightarrow{w} f$ and $q \xrightarrow{\bar{v}} q$ in $\mathcal{A}$ where i is an initial state and f is a final state. Since these paths come from paths in $\mathcal{T}$, we obtain a pattern $s'' = (u'', v'', \bar{v}'', w'')$ such that $i \xrightarrow{u|u''} q \xrightarrow{v|v''} q \xrightarrow{w|w''} f$ and $q \xrightarrow{\bar{v}|\bar{v}''} q$ in $\mathcal{T}$.

We distinguish two cases depending on whether the two patterns s' and s'' are conjugated or not. If they are not conjugated, it can be assumed, without loss of generality, that $L_{s'} \cap L_{s''} = \varnothing$ by Proposition 3. In this latter case, the function $h_{s,s'}$ provides a channel for $R + h$. If they are conjugated, the two functions $h_{s,s'}$ and $h_{s,s''}$ are equal and therefore again provide a channel for $R + h$.

We are now ready to prove Theorem 4.

Proof (Proof of Theorem 4). Let $R = h_1 + \cdots + h_n$ be a bounded relation, where $h_1, \ldots, h_n$ are rational functions. If R has a channel, then by Proposition 4, $\mathrm{rg}(R)$ has an exponential growth, hence (1) implies (3). If $\mathrm{rg}(R)$ has an exponential growth, then one of the images $\mathrm{rg}(h_i)$ has an an exponential growth. By Proposition 5, the function h_i has a channel, hence (3) implies (2). Finally, let us assume that $\mathrm{rg}(h_1)$ has a channel. Using the previous lemma, it can be proved by induction on i, that each relation $h_1 + \cdots + h_i$ has a channel, hence R has a channel. Therefore (2) implies (1) which concludes the proof.

Example of channel synthesis.

We finally illustrate the channel construction of Corollary 1 on an example. With $A = B = \mathbb{B}$, Figure 2 depicts a transducer realizing a relation R which is the union $R = h_1 + h_2$ of two functions h_1 and h_2 from $\mathbb{B}^*$ to $\mathbb{B}^*$. The function h_1 (on the left side) is the morphism which maps the symbols 0 and 1 to 01 and 1 respectively. The function h_2 (on the right side) maps each word $a_1 a_2 \cdots a_n$ to $a_1 a_3 \cdots$ keeping only symbols at odd positions.

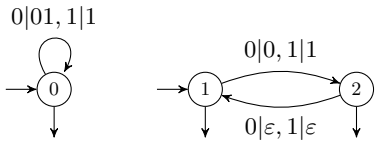

Fig. 2. A union of 2 functions

The function h_1 has obviously a channel since it is one-to-one from $\mathbb{B}^*$ onto $(01 + 1)^*$. We show how the proof of Lemma 4 extracts a channel for the relation R.

1. Since the range $\mathrm{rg}(h_1) = (01+1)^*$ has an exponential growth, we obtain the pattern $s_1 = (\varepsilon, 011, 101, \varepsilon)$ such that $L_{s_1} \subseteq (01 + 1)^*$. The corresponding pattern of the inputs for h_1 is $s = (\varepsilon, 01, 10, \varepsilon)$, hence the function h_{s,s_1} provides a channel for h_1.
2. Examining the part of the transducer realizing h_2, we observe that the input pattern s induces the corresponding output pattern $s_2 = (\varepsilon, 0, 1, \varepsilon)$, but the two patterns s_1 and s_2 satisfy $L_{s_1} \cap L_{s_2} \neq \varnothing$, so h_{s,s_1} does not correspond to a channel for R itself.
3. Since s_1 and s_2 and are not conjugated, it is possible to find a sub-pattern $s_3 = (\varepsilon, 00, 11, \varepsilon)$ of s_2 such that $L_{s_1} \cap L_{s_3} = \varnothing$. The corresponding input pattern for h_2 is then the pattern $s' = (\varepsilon, 0101, 1010, \varepsilon)$ obtained from s.
4. The channel in $R = h_1 + h_2$ is then built from the function h_{s',s_3}. It consists of the pair (E, D) where E is the morphism which maps 0 and 1 to 0101 and 1010 respectively, and D maps 00 and 11 to 0 and 1 respectively.

Besides corollary 1, we also obtain:

Corollary 2. *Let $R = h_1 + \cdots + h_n$ be a relation where each h_i is either a function or the inverse of a function. The following statements are equivalent:*

1. *R has a channel,*
2. *At least one relation h_i has a channel.*

This last result has to be compared with Theorem 4. When R is the finite union of functions and inverses of functions, it may have a domain and a range with exponential growth without having a channel as shown by the following example. Let A and B the alphabet $\mathbb{B}$. Let R be given by $R = h_1 + h_2$ where h_1 is the function from $\mathbb{B}^*$ to $\mathbb{B}^*$ which maps each word to the empty word and h_2 is the inverse of h_1. The domain and the range of R are both equal to $\mathbb{B}^*$ but R has no channel. Hence the decidability result cannot apply in this case.

Proof. Let us recall that a relation R has a channel if and only if R^{-1} has a channel. If one relation h_i has a channel, then R has a channel by Lemma 4.

Suppose now that R has a channel. By Proposition 4, there exist two languages L_0 and L_1 with an exponential growth such that $R' = R \cap (L_0 \times L_1)$ is a bijection between L_0 and L_1. For each $1 \leq i \leq n$, let us denonte by h'_i the restriction $h_i \cap (L_0 \times L_1)$. Since $h'_i \subseteq R'$, and R' is a bijection, each relation h'_i is also a bijection. We may then suppose that each h'_i is a function. It follows then from Thereom 4 that R' has a channel and that R has also a channel.

6 Conclusion

We proved a new characterization of bounded relations with channels, linked to the growth of their image. We conjecture that this characterization could be extended to relations R for which there exists a polynomial P such that for each word u, $|R(u)| \leq P(|u|)$. We also plan to investigate more powerful channels described by (subclasses of) two-way transducers instead of simple transducers.

References

1. Bérard, B., Benattar, G., Lime, D., Mullins, J., Roux, O.H., Sassolas, M.: Channel synthesis for finite transducers. In Dömösi, P., I., S., eds.: Proceedings of the 13th International Conference on Automata and Formal Languages (AFL'11). (August 2011) 79–92
2. Benattar, G., Bérard, B., Lime, D., Mullins, J., Roux, O.H., Sassolas, M.: Channel Synthesis for Finite Transducers. International Journal of Foundations of Computer Science **23**(6) (2012) 1241–1260
3. Schewe, S., Finkbeiner, B.: Synthesis of asynchronous systems. In: Proc. of LOPSTR'06. Volume 4407 of LNCS., Springer (2006) 127–142
4. Millen, J.K.: 20 years of covert channel modeling and analysis. In: Proc. of the 1999 IEEE Symposium on Security and Privacy. (May 1999) 113 –114
5. Hélouet, L., Zeitoun, M., Degorre, A.: Scenarios and Covert channels: another game... In L. de Alfaro, ed.: Proc. of Games in Design and Verification (GDV'04). Volume 119 of ENTCS., Elsevier (2005) 93–116

6. Hélouët, L., Roumy, A.: Covert channel detection using information theory. In Chatzikokolakis, K., Cortier, V., eds.: Proc. of the 8th Int. Workshop on Security Issues in Concurrency (SecCo'10). (August 2010)
7. Maurer, A., Nivat, M.: Rational bijection of rational sets. Acta inf. **13** (1980) 365–378
8. Benattar, G.: Synthèse de systèmes informatiques temporisés non interférents. PhD thesis, Université de Nantes (2011)
9. Lothaire, M.: Combinatorics on words. Volume 17 of Encyclopedia of Mathematics. Addison-Wesley, Reading, MA (1983)
10. Sakarovitch, J.: Elements of automata theory. Cambridge University Press (2009)
11. Elgot, C.C., Mezei, J.E.: On relations defined by generalized finite automata. IBM Journal Res. Develop. **9** (1965) 47–68
12. Weber, A.: Decomposing a k-valued transducer into k unambiguous ones. RAIRO Theoretical Informatics and Applications **30**(5) (1996) 379–413
13. Sakarovitch, J., de Souza, R.: On the decomposition of k-valued rational relations. In Albers, S., Weil, P., eds.: Proceedings of the 25th Symposium on Theoretical Aspects of Computer Science (STACS'08). (2008) 621–632
14. Harrison, M.A.: Introduction to formal language theory. Addison-Wesley (1978)
15. Bérard, B., Carton, O.: Channel synthesis revisited. Technical report, LIP6 (2013) http://pagesperso-systeme.lip6.fr/Beatrice.Berard/PDF/rr-channels-BBOC.pdf.
16. Pin, J.É.: Varieties of Formal Languages. North Oxford, London and Plenum, New-York (1986)

Appendix

A Missing proofs about growth

This part of the appendix is devoted to the proof of Proposition 2. This result is more or less part of folklore but we include its proof for completeness.

An automaton is strongly connected if there is a path between any two states. Let $\mathcal{A}$ be a strongly connected automaton with ε-transitions and for any state q, let K_q be the set of labels of paths from q. The automaton is said to have the *prefix property* if for any state q and any words $w, w' \in K_q$, w and w' are prefix compatible. The next lemmas establish properties of strongly connected automata that allow us to describe a linear algorithm to check whether the growth is polynomial or exponential.

Lemma 5. *A strongly connected automaton has the prefix property if and only if there is at least one of its states q such that for any words $w, w' \in K_q$, w and w' are prefix compatible.*

Proof. Let p be an arbitrary state of a strongly connected automaton $\mathcal{A}$. Suppose that any two words of K_p are prefix compatible. Let q be another of state of $\mathcal{A}$ and let u be the label of a path from p to q. The inclusion $uK_q \subseteq K_p$ implies that any two words of K_q are also prefix compatible.

Recall that a word $w \in A^+$ is primitive if it cannot be written as u^k for some word u and $k > 1$. The primitive root of a word w is the unique primitive word u such that $w = u^k$, with $k \geq 1$.

Lemma 6. *Let $w = a_1 \dots a_n$. Then $a_1 \dots a_k$ is the primitive root of w if and only if the second occurence of w in ww starts in a_{k+1} (the first one starting in a_1).*

Proof. First assume that $u = a_1 \dots a_k$ is the primitive root of w, with $w = u^m$. Then there is an occurrence of w starting at the second occurrence of u in ww (hence at a_{k+1}).

Conversely, if the second occurence of w in ww starts in a_{k+1}, then $w = a_{k+1} \dots a_n a_1 \dots a_k$. Let $u = a_1 \dots a_k$, $v = a_{k+1} \dots a_n$ and $v' = a_1 \dots a_{n-k}$. Since $w = vu$ and $ww = uvuv$, u and v commute and $v = v'$, hence u is the primitive root of w.

The following lemma gives an algorithmic characterization of the prefix property that will allow us to design a linear time algorithm to check it.

Lemma 7. *Let $\mathcal{A}$ be a strongly connected automaton with ε-transitions and let q_0 be one of its states. The automaton $\mathcal{A}$ has the prefix property if and only if there exist a primitive word $w = w_0 \cdots w_{n-1}$ and a function $f : Q \to \{0, \dots, n-1\}$ such that:*

- *$f(q_0) = 0$,*
- *for any transition $p \xrightarrow{\varepsilon} q$ of $\mathcal{A}$, $f(q) \equiv f(p) \mod n$,*

– *for any transition* $p \xrightarrow{a} q$ *of* $\mathcal{A}$, $f(q) \equiv f(p) + 1 \mod n$ *and* $a = w_{f(p)}$.

Furthermore, there exist at most one word w *and one function* f *having these two properties.*

Proof. The result is clear if all transitions of $\mathcal{A}$ are ε-transitions. We now suppose that there is as least one transition $p \xrightarrow{a} q$ in $\mathcal{A}$ labelled by a letter a. It follows that each set K_q is infinite since $\mathcal{A}$ is strongly connected.

We first check that the condition is sufficient. Consider a path from q_0:

$$q_0 \xrightarrow{b_0} q_1 \xrightarrow{b_1} q_2 \xrightarrow{b_2} \cdots \xrightarrow{b_m} q_{m+1}.$$

where each b_i is either ε or a letter. We have then $f(q_0) = 0$ and $f(q_{i+1}) \equiv f(q_i) + |b_i| \mod n$ for any $i \leq m$. It follows that $f(q_i) \equiv |b_0 \cdots b_{i-1}| \mod n$ and thus b_i is either ε or the letter w_j where $j \equiv |b_0 \cdots b_{i-1}| \mod n$ for each $i \leq m+1$. It is then clear that the label of this path is a prefix of w^ω. This shows that any word u in K_{q_0} is a prefix of w^ω.

We have shown in the previous paragraph that if there exist w and f as required, then each word of K_{q_0} is a prefix of w^ω. If w and w' are two different primitive words, then $w^\omega \neq w'^\omega$. Since K_{q_0} is infinite, the word w is unique. We have also shown that if there is a path from q_0 to q with label u then $f(q) \equiv |u| \mod n$. This also shows that f is unique.

We now suppose $\mathcal{A}$ has the prefix property. This means any two words $u, u' \in K_{q_0}$ are prefix compatible. Let w' be the label of a cycle from q_0 to q_0 and let w be the primitive root of w', that is $w' = w^k$ for some $k \geq 1$. Note that any word u from K_{q_0} is a prefix of w^ω. Let n be the length of w. We claim if that there are two paths from q_0 to q with labels u and u', then $|u| \equiv |u'| \mod n$. Let v the label of paths from q to q_0. Both words uvw' and $u'vw'$ are prefixes of w^ω. Since w is not a middle occurrence of w^2, we have $|u| \equiv |u'| \mod n$. This shows the claim. It allows us to define $f(q) = (|u| \bmod n)$ where u is the label of one path from q_0 to m. It is then clear that f satisfies all the required properties.

Lemma 8. *It can be checked in linear time whether a strongly connected automaton has the prefix property.*

Proof. The algorithm performs a first depth-first search to find a simple cycle from q_0 to q_0 of with a non-empty label w'. The length of w' is bounded by the number of states of the automaton. The primitive root $w = w_0 \cdots w_{n-1}$ of w' can be computed in linear time (using lemma 6 above and classical algorithms on strings).

The algorithm performs a second depth-first search to compute the values of the function f and checks whether they satisfy the condition of the previous lemma. Each value $f(q)$ is set by the algorithm at the first visit of q and it is never changed afterwards. The algorithm starts in state q_0 and sets $f(q_0) := 0$. For each traversal of a transition of the automaton, it operates as follows. For an ε-transition $p \xrightarrow{\varepsilon} q$, the algorithm either sets $f(q) := f(p)$ if q has not been visited or checks whether $f(q) = f(p)$ if q has been already visited. For a transition

$p \xrightarrow{a} q$, the algorithm first checks whether $a = w_{f(q)}$. If state q has not been visited, the algorithm sets $f(q) := f(p) + 1 \mod n$. If state q has already been visited, the algorithm checks whether $f(q) \equiv f(p) + 1 \mod n$. If one of these checks fails, then the algorithm return *false*. If the algorithm terminates without failure, it returns *true*.

The following lemma combined with the previous one yields the proof of Proposition 2.

Lemma 9. *The language accepted by a trim automaton with ε-transitions has an exponential growth if and only if one of its strongly connected component has not the prefix property.*

Proof. Let $\mathcal{A}$ be a trim automaton with ε-transitions accepting L and let k be the number of strongly connected components (SCC) of $\mathcal{A}$. For any states p and q, let $L_{p,q}$ be the set of labels of paths from p to q.

Suppose first that each SCC of $\mathcal{A}$ has the prefix property. Each set $L_{p,q}$, for p and q in the same SCC has then a linear growth. Each word w accepted by $\mathcal{A}$ can be uniquely factorized $w = u_0 v_1 u_1 \cdots v_r u_r$ where $r \leq k$ and each u_i is the label of a path made of states from different SCC and v_i is the label of a path made of states from the same SCC. Note that there are finitely many words u_i. It follows that the growth of L is at most n^k.

Suppose now that at least one SCC of $\mathcal{A}$ does not have the prefix property. There exist, for some state q of this SCC, two words v and $\bar{v}$ such that $|v| = |\bar{v}|$, $v \neq \bar{v}$, $q \xrightarrow{v} q$ and $q \xrightarrow{\bar{v}} q$. Since $\mathcal{A}$ is trim, there is a path $i \xrightarrow{u} q$ from an initial state i to q and there is also a path $q \xrightarrow{w} f$ from q to a final state f. It follows that the pattern $s = (u, v, \bar{v}, w)$ satisfies $L_s \subseteq L$ and the set L has an exponential growth.

B Missing proofs about patterns

This section is devoted to the proofs of Lemmas 1 and 2, which use the next lemma.

Lemma 10. *Let $\varphi : \mathbb{B}^* \to M$ be a morphism from $\mathbb{B}^*$ to a finite monoid M. There exist two different words x and $\bar{x}$ over $\mathbb{B}$ such that $|x| = |\bar{x}|$ and $\varphi(x) = \varphi(\bar{x}) = \varphi(x^2)$.*

Proof. We may suppose that φ is onto. Otherwise, M can be replaced by its sub-monoid $\varphi(A^*)$. Consider the quasi-ordering $\leq$ on M defined by $m \leq m'$ if there exists n such that $m = m'n$. This quasi-odering is the Green ordering $\leq_{\mathcal{R}}$ [16, Chap. 3]. Let m be a minimal element for $\leq$. There is an integer k such that $e = m^k$ is an idempotent, that is $e^2 = e$ [16, Prop. 1.6]. Note that e is also minimal for $\leq$ since $e = m^k \leq m$. Let w be a word such that $\varphi(w) = e$. Since e is minimal for $\leq$ there exists two words u and $\bar{u}$ such that $\varphi(w0u) = \varphi(w1\bar{u}) = e$. Let x and $\bar{x}$ be defined by $x = (w0u)^{|w1\bar{u}|}$ and $\bar{x} = (w1\bar{u})^{|w0u|}$. They satisfy the required properties since $|x| = |\bar{x}| = |w0u||w1\bar{u}|$ and $\varphi(x) = \varphi(\bar{x}) = e = e^2$.

Lemma 1. *Let s be a pattern and L a rational set of words. There exists a sub-pattern s' of s such that either $L_{s'} \subseteq L$ or $L_{s'} \cap L = \varnothing$.*

Proof. Let s be the pattern $(u, v, \bar{v}, w)$ and let $\mu_s : \mathbb{B}^* \to A^*$ be the morphism defined by $\mu_s(0) = v$ and $\mu_s(1) = \bar{v}$. Let $\mu : A^* \to M$ be a morphism from A^* into a finite monoid M such that $L = \mu^{-1}(\mu(L))$. Let φ be the morphism $\mu \circ \mu_p$ from $\mathbb{B}^*$ to M. By Lemma 10, there exist two different words x and $\bar{x}$ such that $|x| = |\bar{x}|$ and $\varphi(x) = \varphi(\bar{x}) = e$ for some idempotent e of M, that is $e = e^2$. Let v' and $\bar{v}'$ be the words $\mu_s(x)$ and $\mu_s(\bar{x})$. Note that these two words belong to $(v + \bar{v})^*$ and satisfy $|v'| = |\bar{v}'| = |x||v|$ and $\mu(v') = \mu(\bar{v}') = e$. Let s' be the sub-pattern $s' = (uv', v', \bar{v}', w)$. Note that $\mu(L_{s'})$ is the singleton set $\{\mu(u)e\mu(w)\}$. It follows that $L_{s'} \subseteq L$ if $\mu(u)e\mu(w) \in \mu(L)$ and $L_{s'} \cap L = \varnothing$ otherwise.

Lemma 2. *Let s be a pattern and let L be a rational set accepted by an automaton $\mathcal{A}$. If $L_s \subseteq L$, there exist a sub-pattern $s' = (u', v', \bar{v}', w')$ of s and paths $i \xrightarrow{u'} q \xrightarrow{v'} q \xrightarrow{w'} f$ and $q \xrightarrow{\bar{v}'} q$ in $\mathcal{A}$ where i is an initial state and f is a final state.*

Proof. Let s be the pattern $(u, v, \bar{v}, w)$ and let $\mu_s : \mathbb{B}^* \to A^*$ be the morphism defined by $\mu_s(0) = v$ and $\mu_s(1) = \bar{v}$. Let M be the transition monoid of $\mathcal{A}$ and let $\mu : A^* \to M$ be the canonical morphism from A^* into M. Let φ be the morphism $\mu \circ \mu_s$ from $\mathbb{B}^*$ to M. By Lemma 10, there exist two different words x and $\bar{x}$ such that $|x| = |\bar{x}|$ and $\varphi(x) = \varphi(\bar{x}) = e$ for some idempotent e of M, that is $e = e^2$. Let v' and $\bar{v}'$ be the words $\mu_s(x)$ and $\mu_s(\bar{x})$. Note that these two words belong to $(v + \bar{v})^*$ and satisfy $|v'| = |\bar{v}'| = |x||v|$ and $\mu(v') = \mu(\bar{v}') = e$. Let n be the number of states of $\mathcal{A}$. Since $L_s \subseteq L$ and L is accepted by $\mathcal{A}$, there is an accepting path $i \xrightarrow{uv'^n w} f$. Since n is the number of states, the same state q occurs twice and the path can be factorized $i \xrightarrow{uv'^{n_0}} q \xrightarrow{v'^{n_1}} q \xrightarrow{v'^{n_2} w} f$ where $n = n_0 + n_1 + n_2$. Since $e = \mu(v') = \mu(\bar{v}')$ is an idempotent, there are also two paths $q \xrightarrow{v'} q$ and $q \xrightarrow{\bar{v}'} q$. Let u' and w' be the two words uv'^{n_0} and $v'^{n_2} w$. The sub-pattern $s' = (u', v', \bar{v}', w')$ has the required property.