Closure Properties of Synchronized Relations

María Emilia Descotte
LaBRI, Université de Bordeaux, France

Diego Figueira
CNRS & LaBRI, Université de Bordeaux, France

Santiago Figueira
CONICET & Universidad de Buenos Aires, Argentina

Abstract
A standard approach to define $k$-ary word relations over a finite alphabet $A$ is through $k$-tape finite state automata that recognize regular languages $L$ over $\{1, \ldots, k\} \times A$, where $(i, a)$ is interpreted as reading letter $a$ from tape $i$. Accordingly, a word $w \in L$ denotes the tuple $(u_1, \ldots, u_k) \in (A^*)^k$ in which $u_i$ is the projection of $w$ onto $i$-labelled letters. While this formalism defines the well-studied class of rational relations, enforcing restrictions on the reading regime from the tapes, which we call synchronization, yields various sub-classes of relations. Such synchronization restrictions are imposed through regular properties on the projection of the language $L$ onto $\{1, \ldots, k\}$. In this way, for each regular language $C \subseteq \{1, \ldots, k\}^*$, one obtains a class $Rel(C)$ of relations. Synchronous, Recognizable, and Length-preserving rational relations are all examples of classes that can be defined in this way.

We study basic properties of these classes of relations, in terms of closure under intersection, complement, concatenation, Kleene star and projection. We characterize the classes with each closure property. For the binary case ($k = 2$) this yields effective procedures.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory

Keywords and phrases synchronized word relations, rational, closure, characterization, intersection, complement, Kleene star, concatenation

Digital Object Identifier 10.4230/LIPIcs.STACS.2019.22

Acknowledgements Work supported by ANR project DELTA, grant ANR-16-CE40-0007, grant PICT-2016-0215, and LIA INFINIS.

1 Introduction
We study relations of finite words, that is, sets $R \subseteq (A^*)^k$ for a finite alphabet $A$ and $k \in \mathbb{N}$, where $(A^*)^k$ is the cartesian product of $k$ copies of $A^*$. The study of these relations dates back to the works of Büchi, Elgot, Mezei, and Nivat in the 1960s [11, 15, 24], with much subsequent work done later (e.g., [7, 13]). Most of the investigations focused on extending the standard notion of regularity from languages to relations. This effort has followed the long-standing tradition of using equational, operational, and descriptive formalisms – that is, finite monoids, automata, and regular expressions – for describing relations, and gave rise to three different classes of relations: Recognizable, Automatic (a.k.a. Regular [7] or Synchronous [20, 13]), and Rational.

The above classes of relations can be seen as three particular examples of a much larger (in fact infinite) range of possibilities, where relations are described by special languages over extended alphabets, called synchronizing languages [18]. Intuitively, the idea is to describe a $k$-ary relation by means of a $k$-tape automaton with $k$ heads, one for each tape, which can move independently of one another. In the basic framework of synchronized relations, one lets each head of the automaton either move right or stay in the same position. In addition, one can constrain the possible sequences of head motions by a suitable
regular language \( C \subseteq \{1, \ldots, k\}^* \). In this way, each regular language \( C \subseteq \{1, \ldots, k\}^* \) induces a class of \( k \)-ary relations, denoted \( \text{Rel}(C) \), which is contained in the class \( \text{Rational} \) (due to Nivat’s Theorem [24]). For example, on binary relations, the classes \( \text{Recognizable} \), \( \text{Automatic} \), and \( \text{Rational} \) are captured, respectively, by the languages \( C_{\text{Rec}} = \{1\}^* \cdot \{2\}^* \), \( C_{\text{Aut}} = \{1\}^* \cdot \{1\}^* \cup \{1\}^* \cdot \{2\}^* \), and \( C_{\text{Rat}} = \{1, 2\}^* \). Roughly speaking, any other class that can be defined through the ‘tape behavior’ of a multi-tape automaton will be also captured by this framework. Other examples include length-preserving, or \( \alpha \)-synchronous relations [12]. However, it should be noted that other well-known subclasses of rational relations, such as deterministic or functional relations, are not captured by the notion of synchronization. In general, the correspondence between a language \( C \subseteq \{1, \ldots, k\}^* \) and the induced class \( \text{Rel}(C) \) of synchronized relations is not one-to-one: it may happen that different languages \( C, D \) induce the same class of synchronized relations. The problem of when two classes of synchronized relations coincide, and when one is contained in the other has been only recently solved for the case of binary relations [14], while the case for arbitrary \( k \)-ary relations remains open. In this work we identify, among the infinitely many synchronized classes of relations, which are those with good closure properties, in terms of paradigmatic operations such as intersection, complement, concatenation, projection, or Kleene star.

Motivation

The motivation for identifying and studying well-behaved classes of word relations, besides its intrinsic interest within formal language theory, stems from various areas. One motivation comes from verification of safety and liveness properties of parameterized systems, where relations describe transitions [1, 10, 22, 26]. Another one arises from the study of Automatic Structures [8], where word languages and relations are used to describe infinite structures, and good closure properties are necessary to obtain effective model checking of logics. Another example is the study of formal models underlying IBM’s tools for text extraction into a relational model [16]; where several classes of relations emerge (some outside \( \text{Rational} \)) with differing closure properties. Yet another comes from graph databases, which are actively studied as a suitable model for RDF data, social networks data, and others [2]. Paths in graph databases are described by their labels and hence they are abstracted as finite words. These paths need to be compared, for instance, for their degree of similarity, edit distance, or other relations [3, 5, 23]. As a concrete link with the present work we consider CRPQs – a basic query language for graph-structured data. As it was shown in [4], allowing rational relations in CRPQs turns the query evaluation problem undecidable. There have therefore been efforts towards finding subclasses of \( \text{Rational} \) relations that preserve decidability for CRPQs (e.g. [5, 17, 6]), often exploiting an effective closure under intersection on the underlying subclass of relations. Part of our motivation for studying closure under intersection stems from our ambition, as future work, to characterize all synchronized classes of relations that can be added to CRPQs while preserving decidability.

Contribution

Our main contribution is a characterization for each of the studied closure properties, the main results can be summarized as follows.

\[ \text{Theorem.} \quad \text{For every regular} \ C \subseteq 2^*, \text{it is decidable whether} \ \text{Rel}(C) \ \text{is closed under intersection, complement, concatenation, Kleene star and projection.} \]
While some of the characterizations we give are for arbitrary arity relations, we were only able to show decidability for binary arity. Indeed, the decidability of these characterizations relies, crucially, on the decidability of testing for inclusion between synchronized classes, which has only been shown for binary relations [14].

We do not include closure under union since it can be easily seen that all classes defined in this way are closed under union. The most involved result is closure under intersection. The main property we will prove is that $\text{Rel}(C)$ is closed under intersection if, and only if, $\text{Rel}(C) \subseteq \text{Rel}(D)$ for some $D$ whose Parikh-image is injective (i.e., there are no two distinct words of $D$ with the same Parikh-image). Further, we show that this can be tested, and such a language $D$ can be effectively constructed, whenever possible. In the same vein, we obtain that $\text{Rel}(C)$ is closed under complement if, and only if, $\text{Rel}(C) = \text{Rel}(D)$ for some $D$ with a bijective Parikh-image. (Observe that closure under complement implies closure under intersection in view of the fact that all classes are closed under union.)

Related work

The formalization of the framework to describe synchronized classes of relations has been introduced only recently [18]. As mentioned, the problem of containment between classes of relations has been addressed in [14] for the binary case. The formalism of synchronizations has been also extended beyond rational relations by means of semi-linear constraints [17] in the context of querying graph databases.

The paper [9] studies relations with origin information, as induced by non-deterministic (one-way) finite state transducers. Origin information can be seen as a way to describe a synchronization between input and output words – somehow in the same spirit of our synchronization languages – and was exploited to recover decidability of the equivalence problem for transducers. The paper [19] pursues further this principle by studying “distortions” of the origin information, called resynchronizations. The paper [27] studies the uniformization problem for synchronized relations.

Organization

After a preliminary Section 2, we show the main result characterizing closure under intersection in Section 3. In Section 4 we study closure under complement and another variant that we call “relativized complement”. In Section 5 we give characterizations for closure under concatenation, Kleene star and projection. We conclude with Section 6.

2 Preliminaries

We denote by $\mathbb{N}$ the set of non-negative integers, $A, B$ denote arbitrary finite alphabets and for $k \in \mathbb{N}, k \geq 1$, $\Sigma_k$ denotes the $k$-letter alphabet $\{1, \ldots, k\}$. For a word $w \in A^*$, $|w|$ is its length, and $|w|_a$ is the number of occurrences of symbol $a$ in $w$.

Regular languages

We use standard notation for regular expressions without complement, namely, for expressions built up from the empty set, the empty word $\varepsilon$ and the symbols $a \in A$, using the operations $\cdot$, $\cup$, and $^*$. For economy of space and clarity we use the abbreviated notation $(\cdot)^n$, $(\cdot)^{<n}$, $(\cdot)^{\geq n}$, $(\cdot)^{\leq n}$, $(\cdot)^{\geq n}$, and $(\cdot)^{\leq n}$ – the last two being shorthands for $(\cdot)^n$, $(\cdot)^n$ respectively. We abuse the notation $(\cdot)^k$ to also denote the cartesian product of $k$ copies of the same set (typically $(A^*)^k$) when there is no risk of confusion. We also identify regular expressions with
the defined languages; for example, we may write \( abbc \in a \cdot b \cdot 2 \cdot (c \cup d)^* \), \( b(ab)^* = (ba)^* b \) and \( \{a, b\}^* \cdot c = (a \cup b)^* \cdot c \). The star-height of a regular expression is the maximum number of nested Kleene stars \( (\cdot)^* \). Given \( u = a_1 \cdots a_n \in A^* \) and \( v = b_1 \cdots b_n \in B^* \), we write \( u \odot v \) for the word \( (a_1, b_1) \cdots (a_n, b_n) \in (A \times B)^* \). Similarly, given \( U \subseteq A^* \), \( V \subseteq B^* \), we write \( U \odot V \subseteq (A \times B)^* \) for the set \( \{u \odot v : u \in U, v \in V, \|u\| = \|v\|\} \). Given two languages \( L, L' \) over \( A \), we write \( L \subseteq_{\text{reg}} L' \) to denote that \( L \) is a regular subset of \( L' \).

A regular expression \( C \subseteq 2^* \) is concat-star, if it is of the form

\[
C = C_1^* u_1 C_2^* u_2 \cdots C_n^* u_n,
\]

for \( n \in \mathbb{N} \), words \( u_1, \ldots, u_n \), and regular expressions \( C_1, \ldots, C_n \) where none of the \( C_i \)'s describes the empty language. The \( C_i \)'s from (\( \ast \)) are called components of the concat-star. A concat-star expression like (\( \ast \)) is smooth if either \( n \leq 2 \) or there are no \( t, s \in 2 \) and \( 1 \leq i < n \) such that \( C_i \subseteq t^i \), \( C_{i+1} \subseteq s^i \). We say that a regular language \( L \) is concat-star (resp. smooth) if it admits a concat-star (resp. smooth) expression.

**Parikh-images and linear sets**

The Parikh-image of \( w \in 2^* \) is the pair associating each symbol of \( 2 \) to its number of occurrences in \( w \), i.e., \( \pi(w) = (|w|_1, |w|_2) \). We naturally extend this to languages \( L \subseteq 2^* \) by letting \( \pi(L) \overset{\text{def}}{=} \{\pi(w) : w \in L\} \subseteq \mathbb{N}^2 \). A language \( C \subseteq 2^* \) is Parikh-injective if for every \( u, v \in C \), if \( \pi(u) = \pi(v) \) then \( u = v \); it is Parikh-surjective if \( \pi(C) = \mathbb{N}^2 \); and it is Parikh-bijective if it is both Parikh-injective and -surjective. We will use the product order \((\leq, \mathbb{N}^2)\), defined by \((n, m) \leq (n', m')\) iff \( n \leq n' \) and \( m \leq m' \). Given a vector \( \bar{x} \in \mathbb{N}^2 \) and a set \( X = \{\bar{x}_1, \ldots, \bar{x}_n\} \subseteq \mathbb{N}^2 \) (in our case, the Parikh-image of words from \( 2^* \)), we define the linear set generated by \( X \) and \( \bar{x} \) as \( \langle \bar{x}, X \rangle = \{\bar{x} + \alpha_1 \cdot \bar{x}_1 + \cdots + \alpha_n \bar{x}_n : \alpha_i \in \mathbb{N}\} \). For economy of space we write \( \langle X \rangle \) as short for \( \langle \bar{0}, X \rangle \), where \( \bar{0} = (0, 0) \). Note that, in particular, \( \langle \bar{0} \rangle = \{\bar{0}\} \). A semi-linear set is a finite union of linear sets. The following fact will be useful in the next section.

**Lemma 1.** For every semi-linear set \( V \subseteq \mathbb{N}^2 \) there exists a Parikh-injective language \( C \subseteq_{\text{reg}} 2^* \) such that \( \pi(C) = V \).

Two sets of vectors \( X, Y \subseteq \mathbb{N}^2 \) are independent if \( \bar{0} \not\in X \cap Y \) and \( \langle X \rangle \cap \langle Y \rangle = \{\bar{0}\} \); otherwise they are dependent. We say that two languages over \( 2 \) are Parikh-independent (resp. Parikh-dependent) if their Parikh-images are. We abuse notation and say that \( \bar{x} \) and \( \bar{y} \) are (in)dependent whenever \( \langle \bar{x} \rangle \) and \( \langle \bar{y} \rangle \) are (in)dependent, and likewise for words. We will need the following simple observation later.

**Observation 2.** If \( u \) and \( v \) are Parikh-independent, for every \( s, t, s', t' \in \mathbb{N} \), if \( \pi(u^s v^{s'}) = \pi(u^{t} v^{t'}) \), then \( s' = t \) and \( t' = s \).

Indeed, we have that \( s \cdot \pi(u) + s' \cdot \pi(v) = t' \cdot \pi(u) + t \cdot \pi(v) \). Let us assume that \( s' \leq t \) (the case in which is \( \geq \) is similar). Then \( t' \leq s \) and so we have \( (s - t') \cdot \pi(u) = (t - s') \cdot \pi(v) \) which implies \( s - t' = 0 = t - s' \) since \( u \) and \( v \) are Parikh-independent. Then \( s' = t \) and \( t' = s \).

### 2.1 Synchronized relations

A synchronization of a tuple \((w_1, \ldots, w_k)\) of words over \( A \) is a word over \( A \times A \) such that the projection onto \( A \) of positions labeled by \( i \) is exactly \( w_i \), for \( i = 1, \ldots, k \). For example, the words \((1, a)(1, b)(2, a)\) and \((1, a)(2, a)(1, b)\) are two possible synchronizations of the same pair \((ab, a)\). Every word \( w \in (A \times A)^* \) is a synchronization of a unique tuple \((w_1, \ldots, w_k)\)
of words over $A$, where for all $i \in \{1, \ldots, k\}$, $i^{[w]} \otimes w_i$ is the projection of $w$ onto the alphabet $\{i\} \times A$. We denote such tuple $(w_1, \ldots, w_k)$ by $[w]_k$ and extend the notation to languages $L \subseteq (k \times A)^*$ by denoting the unique $k$-ary relation synchronized by $L$ as $[L]_k \overset{\text{def}}{=} \{[w]_k : w \in L\}$. In our previous example, $[(1, a)(1, b)(2, a)]_2 = [(1, a)(2, a)(1, b)]_2 = (ab, a)$, and $\{((1, a)(2, a), (1, a)(2, b), (1, b)(2, a), (1, b)(2, b))^*\}_2$ is the equal-length relation on the alphabet $\{a, b\}$.

In this setup, we define classes of relations by restricting the set of admitted synchronizations. One way of doing so is to fix a language $C \subseteq_{\text{reg}} k^*$, called control language, and let $L$ vary over all regular languages over $k \times A$ whose projections onto $k$ are in $C$. Thus, given $k \in \mathbb{N}$ and $C \subseteq_{\text{reg}} k^*$, we define the class of $k$-ary $C$-controlled relations as

$$\text{REL}_k(C) \overset{\text{def}}{=} \{([L]_k, A) : L \subseteq_{\text{reg}} C \otimes A^*, A \text{ is a finite alphabet}\}.$$ 

Whenever $k$ is clear from the context, we write $[w]_k$, $[L]_k$ and REL$(C)$. For economy of space, we write $C =_{\text{Rel}} D$ as short for $\text{REL}(C) = \text{REL}(D)$, and we say that $C$ is $\text{Rel}$-equivalent to $D$. Similarly, we write $C \subseteq_{\text{Rel}} D$ as short for $\text{REL}(C) \subseteq \text{REL}(D)$ and we say that $C$ is $\text{Rel}$-contained in $D$. The definition makes explicit the alphabet used for each relation, in contrast to previous definitions of synchronized classes [18, 14]. The reason for this is that in particular we study closure under complement, which requires the alphabet to be specified. However, we observe that synchronized classes are closed under taking super-alphabets, and thus the alphabet can be often disregarded. We then write $R \subseteq \text{REL}(C)$ to denote $(R, A) \in \text{REL}(C)$ for some $A$.

**Observation 3.** If $(R, A) \in \text{REL}(C)$ then $(R, A') \in \text{REL}(C)$ for every $A \subseteq A'$. If $(R, A) \in \text{REL}(C)$ then $(R, A_R) \in \text{REL}(C)$, where $A_R \subseteq A$ is the set of symbols present in $R$.

Clearly, $C \subseteq_{\text{reg}} D \subseteq_{\text{reg}} k^*$ implies $C \subseteq_{\text{Rel}} D$, but the converse does not hold: $\text{REL}_2(C_{\text{Rec}}) = \text{Recognizable} \subseteq \text{Automatic} = \text{REL}_2(C_{\text{Aut}})$, but $C_{\text{Rec}} \subsetneq C_{\text{Aut}}$. Moreover, different control languages may induce the same class of synchronized relations. For any two regular $C, D \subseteq_{\text{reg}} k^*$ it is decidable to test whether $C \subseteq_{\text{Rel}} D$ in the case $k = 2$ [14], but for arbitrary $k$-ary relations the decidability of the class containment problem is open. Henceforward, Rational will denote the class REL($2^*$) of rational relations.

We restate some properties from [14] that we will use throughout (the proofs in [14] are for the case $k = 2$ but they can be easily generalized to arbitrary $k$). We will use the notation $R \cdot S$ to denote the usual concatenation of relations, more specifically, given $R, S \subseteq (A^*)^k$, $R \cdot S = \{(u \cdot u', v \cdot v') : (u, v) \in R \text{ and } (u', v') \in S\}$.

**Lemma 4** (Lemma 2 of [14]). For every $C, D, C', D' \subseteq_{\text{reg}} k^*$,

1. if $R \in \text{REL}(C \cdot D)$, there are $R_1, \ldots, R_n \in \text{REL}(C)$, $R'_1, \ldots, R'_n \in \text{REL}(D)$ such that $R = \bigcup_i R_i \cdot R'_i$;
2. if $R \in \text{REL}(C^*)$, there are $R_1, \ldots, R_n \in \text{REL}(C)$ and $I \subseteq \{1, \ldots, n\}^*$ such that $R = \bigcup_{i \in I} R_i$;
3. For every $R \in \text{REL}(C' \cup D)$, there are $R_1 \in \text{REL}(C)$, $R_2 \in \text{REL}(D)$ such that $R = R_1 \cup R_2$;
4. if $C \subseteq D$, then $C \subseteq_{\text{Rel}} D$;
5. if $C \subseteq_{\text{Rel}} D$ and $C' \subseteq_{\text{Rel}} D'$, then $C \cdot C' \subseteq_{\text{Rel}} D \cdot D'$ and $C \cup C' \subseteq_{\text{Rel}} D \cup D'$;
6. if $C \subseteq_{\text{Rel}} D$, then $C^* \subseteq_{\text{Rel}} D^*$;
7. for every partition $I, J$ of $\{1, \ldots, k\}$ such that $C \subseteq I^*$ and $D \subseteq J^*$, we have $C \cdot D =_{\text{Rel}} D \cdot C$;
8. if $C$ is finite, then $C \cdot D =_{\text{Rel}} D \cdot C$;
9. if $C \subseteq_{\text{Rel}} D$ then $\pi(C) \subseteq \pi(D)$; moreover, if $C$ is finite, the converse also holds.
The following decomposition lemma, which is an immediate consequence of [14, Proposition 3 plus Lemma 2 P7] and basic properties from Lemma 4, will be used throughout.

**Lemma 5.** Every $C \subseteq_{reg} A^*$ is effectively REL-equivalent to a finite union of smooth languages, i.e. given $C \subseteq_{reg} A^*$, one can compute a finite set of smooth languages such that $C$ is REL-equivalent to their union.

In addition to these, our characterization results make use of the following easy properties of relations controlled by Parikh-injective and Parikh-bijective languages.

**Lemma 6.** For any $C \subseteq_{reg} A^*$ and $L, M \subseteq_{reg} C \otimes A^*$,

1. if $C$ is Parikh-injective, and $w, w' \in C \otimes A^*$, then $\|w\| = \|w'\|$ implies $w = w'$.
2. $[L] \cup [M] = [L \cup M]$.
3. if $C$ is Parikh-injective then $[L] \cap [M] = [L \cap M]$ and $[L] \setminus [M] = [L \setminus M]$.
4. if $C$ is Parikh-bijective then $(A^*)^k \setminus [L] = [(C \otimes A^*) \setminus L]$.
5. if $C$ is Parikh-surjective then $1^* \cdots k^* \subseteq_{rel} C$.

**Proof.** The first two items follow immediately from definitions.

3. $[L \cap M] \subseteq [L] \cap [M]$ is always true. For the other containment, let $(w_1, \ldots, w_k) \in [L] \cap [M]$, then there exist $w \in L, w' \in M$ such that $\|w\| = \|w'\| = (w_1, \ldots, w_k)$. Since $C$ is Parikh-injective, by item 1, $w = w' \in L \cap M$ synchronizes $(w_1, \ldots, w_k)$ which concludes the proof.

$[L] \setminus [M] \subseteq [L \setminus M]$ is always true. For the other containment, let $w \in L \setminus M$. Then $\|w\| \in [L]$. By way of contradiction, suppose that $\|w\| \in [M]$. In this case, there exists $w' \in M$ such that $\|w\| = \|w'\|$. Since $C$ is Parikh-injective, by item 1, $M \not\ni w = w' \in M$ which is a contradiction.

4. For $\subseteq$, note that, since $C$ is Parikh-surjective, $(A^*)^k = [C \otimes A^*]$, and so the result follows from the previous item.

5. We make use of closure under componentwise letter-to-letter relations (cf. Lemma 8 of Section 2.2). Suppose $C \subseteq_{reg} A^*$ is Parikh-surjective, and let $R \in REL(1^* \cdots k^*)$. As an immediate consequence of Mezei’s theorem, we have the following:

**Claim 7.** For every $k$, $REL(1^* \cdots k^*) = \{\bigcup_{i \in I} L_{i,1} \times \cdots \times L_{i,k} : I$ is finite and $L_{i,j} \subseteq_{reg} A^*$ for some finite alphabet $A\}$.

Then $R = \bigcup_{i \in I} L_{i,1} \times \cdots \times L_{i,k}$ for a finite $I$ and regular languages $L_{i,j}$. For any $i \in I$ and $j \in k$ consider $T_{i,j}$ as the letter-to-letter relation $T_{i,j} = \{(u, v) : \|u\| = \|v\|$ and $v \in L_{i,j}\} \in REL((12)^*)$. Note that, by Parikh-surjectivity, $U = (A^*)^k = [C \otimes A^*] \in REL(C)$ and therefore $U \circ (T_{i,1}, \ldots, T_{i,k}) = L_{i,1} \times \cdots \times L_{i,k}$. Then, by closure under union and componentwise letter-to-letter relations (Lemma 8), it follows that $R = \bigcup_{i \in I} U \circ (T_{i,1}, \ldots, T_{i,k}) \in REL(C)$.

### 2.2 Universal closure properties

There are some closure properties which are shared by all classes of synchronized relations, that is, by every $REL(C)$ with $C \subseteq_{reg} A^*$. We highlight the most salient ones.

An **alphabetical morphism** between two finite alphabets $A, B$ is a morphism $h : A^* \to B^*$ between the free monoids such that $h(a) \in B$ for every $a \in A$. Its application is extended to any relation $R \subseteq (A^*)^k$ as follows $h(R) = \{(h(u_1), \ldots, h(u_k)) : (u_1, \ldots, u_k) \in R\} \subseteq (B^*)^k$; and its inverse is applied to $S \subseteq (B^*)^k$ as $h^{-1}(S) = \{(u_1, \ldots, u_k) : (h(u_1), \ldots, h(u_k)) \in S\} \subseteq (A^*)^k$.

A letter-to-letter relation is one from $REL((12)^*)$.

We define the following closure properties over classes $C$ of $k$-ary relations.
\begin{itemize}
  \item $\mathcal{C}$ is closed under union if for all $(R, A), (S, A) \in \mathcal{C}$, $(R \cup S, A) \in \mathcal{C}$;
  \item $\mathcal{C}$ is closed under (inverse) alphabetic morphisms if for all $(R, A) \in \mathcal{C}$ and $h : A^* \to B^*$ (resp. $g : B^* \to A^*$) an alphabetic morphism, $(h(R), B) \in \mathcal{C}$ (resp. $(g^{-1}(R), B) \in \mathcal{C}$);
  \item $\mathcal{C}$ is closed under componentwise letter-to-letter relations if for every $(R, A) \in \mathcal{C}$ and $(T_1, A), \ldots, (T_k, A) \in \text{REL}((12)^*)$ the following relation over the alphabet $A$ is also in $\mathcal{C}$: $R \circ (T_1, \ldots, T_k) \overset{\text{def}}{=} \{(u_1, \ldots, u_k) : \text{there is } (v_1, \ldots, v_k) \in R \text{ s.t. } (v_i, u_i) \in T_i \text{ for every } i\}$.
  \item $\mathcal{C}$ is closed under recognizable projections if for all $(R, A) \in \mathcal{C}$ and $(S, A) \in \text{REL}(1^* \cdot \cdots \cdot k^*)$, $(R \cap S, A) \in \mathcal{C}$.
\end{itemize}

\textbf{Lemma 8.} For every $k \in \mathbb{N}$ and $C \subseteq_{\text{reg}} k^*$, $\text{REL}(C)$ is closed under union, alphabetic morphisms, inverse alphabetic morphisms, componentwise letter-to-letter relations, and recognizable projections.

\textbf{Proof.} Closure under union follows from the fact that if $L, L' \subseteq_{\text{reg}} C \otimes A^*$, then $L \cup L' \subseteq_{\text{reg}} C \otimes A^*$ and $[L] \cup [L'] = [L \cup L']$ (Lemma 6). Closure under letter-to-letter relations follows from the fact that, given $L \subseteq_{\text{reg}} C \otimes A^*$ and $k$ letter-to-letter relations $T_1, \ldots, T_k$ over $A$, there exists $L' \subseteq_{\text{reg}} C \otimes A^*$ such that $[L'] = [L] \circ (T_1, \ldots, T_k)$ (one can build an automaton recognizing such language from the automata for $L, T_1, \ldots, T_k$). Since any (inverse) alphabetic morphism can be implemented as a letter-to-letter relation, it follows that $\text{REL}(C)$ is closed under (inverse) alphabetic morphisms. Finally, closure under recognizable projections follows from closure under letter-to-letter relations and closure under union, since for every $R \in \text{REL}(C)$ and $S = \bigcup_{i \in I} L_{i,1} \times \cdots \times L_{i,k} \in \text{REL}(1^* \cdot \cdots \cdot k^*)$ (recall Claim 7) we have that $R \cap S = \bigcup_{i \in I} R \circ (T_{i,1}, \ldots, T_{i,k})$ for $T_{i,j} = \{(w, w) : w \in L_{i,j}\}$.

\section{Closure under intersection}

We say that a class $\mathcal{C}$ of $k$-ary relations is \textit{closed under intersection} if for all $(R, A), (S, A) \in \mathcal{C}$, $(R \cap S, A) \in \mathcal{C}$. In this section we show a decidable characterization of the languages $C \subseteq_{\text{reg}} 2^*$ for which $\text{REL}(C)$ is closed under intersection. Further, for $C \subseteq_{\text{reg}} 2^*$, if $\text{REL}(C)$ is closed under intersection, it is \textit{effectively closed}, that is, for every $R, S \in \text{REL}(C)$ over an alphabet $A$, one can compute $R \cap S$ as a synchronized relation, that is, as some $L \subseteq_{\text{reg}} (2 \times A)^*$ so that $[L] = R \cap S$. The main result is the following.

\textbf{Theorem 9.} For every $C \subseteq_{\text{reg}} 2^*$, $\text{REL}(C)$ is closed under intersection if, and only if, $C \subseteq_{\text{REL}} D$ for some Parikh-injective $D \subseteq_{\text{reg}} 2^*$.

At the end of this section we give an effective procedure to decide, given $C \subseteq_{\text{reg}} 2^*$, whether $\text{REL}(C)$ is closed under intersection. Decidability can be seen as the fact that the set of languages $C \subseteq_{\text{reg}} 2^*$ for which there is a Parikh-injective language $D \subseteq_{\text{reg}} 2^*$ such that $C \subseteq_{\text{REL}} D$ is both computably enumerable and co-computably enumerable. While showing that it is c.e. is straightforward, proving co-c.e. involves all the developments of this section. Concretely, we define some \textit{bad conditions} that characterize all languages $C$ such that $\text{REL}(C)$ is \textit{not} closed under intersection, and in this way we obtain that the set of languages $C \subseteq_{\text{reg}} 2^*$ which satisfy any of the bad conditions is c.e.

We will start by giving a sufficient condition for $\text{REL}(C)$ to be closed under intersection. The following simple lemma (which was already proved in [18]) follows from Lemma 6.

\textbf{Lemma 10.} If $C \subseteq_{\text{reg}} 2^*$ is Parikh-injective, then $\text{REL}(C)$ is closed under intersection.
This lemma implies that any language which is $\Rel$-equivalent to a Parikh-injective one gives rise to a closed under intersection class. A natural question is whether the converse holds but it doesn’t seem to. For instance, if $C = 1^*2^* \cup (12)^*$, $\Rel(C)$ is closed under intersection but it seems unlikely that $C$ is $\Rel$-equivalent to a Parikh-injective language.

Another sufficient condition for $\Rel(C)$ to be closed under intersection is that $C = \Rel D \cup X$ for some Parikh-injective $D, X \subseteq_{reg} 2^*$ such that $X \subseteq_{\Rel} 1^*2^*$ (in fact, it can be seen that injectivity of $X$ is not really necessary). We will prove that this condition is also necessary, and thus we will have another characterization of closure under intersection. This is not obvious and we will prove a stronger statement, which we present below (Theorem 12). Also, in particular, we will show that if $\Rel(C)$ is closed under intersection, we can compute a Parikh-injective $D \subseteq_{reg} 2^*$ such that $C \subseteq_{\Rel} D$, which allows us in turn to compute the intersection of two relations in $\Rel(C)$ as a synchronized relation.

For $C \subseteq_{reg} 2^*$, we denote by $\Rel(C)^\cap$ the closure under intersection of $\Rel(C)$, i.e., the smallest class of relations containing $\Rel(C)$ and being closed under intersection. We present three properties on $C \subseteq_{reg} 2^*$ that we call the bad conditions, which will characterize the languages such that $\Rel(C)$ is not closed under intersection.

**Bad conditions**

For $C \subseteq_{reg} 2^*$, consider the following properties:

- **(A)** There exist $u_1, u_2, v, z \in 2^*$ such that
  1. $u_i$ and $v$ are Parikh-independent for $i = 1, 2$,
  2. $\pi(u_i) \geq (1, 1)$ for some $i$,
  3. $\{u_1, u_2\}$ and $\{v\}$ are Parikh-dependent,
  4. $u_1^*u_2^*z \subseteq_{\Rel} C$ and $v^*z \subseteq_{\Rel} C$.

- **(B)** There exist $u, v, z \in 2^*$ such that
  1. $u$ and $v$ are Parikh-independent,
  2. $\pi(u) \geq (1, 1)$ or $\pi(v) \geq (1, 1)$,
  3. $u^*v^*z \subseteq_{\Rel} C$ and $v^*u^*z \subseteq_{\Rel} C$.

- **(C)** There exist $u, v, w, z \in 2^*$ such that
  1. $u \in 1^* \setminus \{\varepsilon\}$, $w \in 2^* \setminus \{\varepsilon\}$,
  2. $\pi(v) \geq (1, 1)$,
  3. $u^*v^*w^*z \subseteq_{\Rel} C$ or $w^*v^*u^*z \subseteq_{\Rel} C$.

For example, $1^*(12)^*(122)^*$ satisfies A for $u_1 = 1, u_2 = 122$, $v = 12, z = \varepsilon$; $1^*(12)^*1^*$ satisfies B for $u = 1, v = 12, z = \varepsilon$; and $1^*(12)^*2^*$ satisfies C for $u = 1, v = 12, w = 2, z = \varepsilon$.

**Observation 11.** The bad conditions are $\subseteq_{\Rel}$-upward closed: If $C \subseteq_{\Rel} D$ and $C$ satisfies property A (resp. B, C), then $D$ also satisfies property A (resp. B, C).

We can now present the characterization theorem.

**Theorem 12.** For $C \subseteq_{reg} 2^*$, the following are equivalent:
1. $\Rel(C)$ is closed under intersection (i.e. $\Rel(C)^\cap = \Rel(C)$);
2. $\Rel(C)^\cap$ is definable (i.e. there exists $D \subseteq_{reg} 2^*$ such that $\Rel(C)^\cap = \Rel(D)$);
3. $\Rel(C)^\cap \subseteq \text{Rational}$.
4. for all \( R, S \in \text{Rel}(C) \), \( R \cap S \in \text{Rational} \);
5. \( C \) does not satisfy any of the bad conditions;
6. there exist \( D, X \subseteq_{\text{reg}} 2^* \) Parikh-injective such that \( C =_{\text{Rel}} D \cup X \) and \( X \subseteq_{\text{Rel}} 1^* 2^* \);
7. there exists \( D \subseteq_{\text{reg}} 2^* \) Parikh-injective such that \( C \subseteq_{\text{Rel}} D \).

From 1 \( \Rightarrow \) 7 and transitivity of \( \subseteq_{\text{Rel}} \), closure under intersection is \( \subseteq_{\text{Rel}} \)-downward closed:

\[ \text{Lemma 15.} \]

\[ \text{Corollary 13.} \] For \( C, D \subseteq_{\text{reg}} 2^* \), if \( C \subseteq_{\text{Rel}} D \) and \( \text{Rel}(D) \) is closed under intersection, then \( \text{Rel}(C) \) is closed under intersection.

We first explain the main proof strategy for obtaining Theorem 12, and present the three key technical results we will need to prove (Propositions 14, 16 and 17).

**Proof idea of Theorem 12**

The proof strategy is by showing 1 \( \Rightarrow \) 2 \( \Rightarrow \) 3 \( \Rightarrow \) 4 \( \Rightarrow \) 5 \( \Rightarrow \) 6 \( \Rightarrow \) 1 on the one hand, and 6 \( \Rightarrow \) 7 \( \Rightarrow \) 3 on the other hand. First observe that 1 \( \Rightarrow \) 2 \( \Rightarrow \) 3 \( \Rightarrow \) 4 are trivial. We next prove 6 \( \Rightarrow \) 1, 7 \( \Rightarrow \) 3 and 6 \( \Rightarrow \) 7.

For 6 \( \Rightarrow \) 1, suppose that \( C =_{\text{Rel}} D \cup X \) for some Parikh-injective languages \( D, X \) such that \( X \subseteq_{\text{Rel}} 1^* 2^* \). Let \( R, S \in \text{Rel}(C) \). Then, by item 3 of Lemma 4, there exist \( R_1, S_1 \in \text{Rel}(D) \), \( R_2, S_2 \in \text{Rel}(X) \) such that \( R = R_1 \cup R_2 \) and \( S = S_1 \cup S_2 \). Note that:

- \( R_1 \cap S_1 \subseteq \text{Rel}(D) \subseteq \text{Rel}(C) \) by Lemma 10 applied to \( D \);
- \( R_2 \cap S_2 \subseteq \text{Rel}(X) \subseteq \text{Rel}(C) \) by Lemma 10 applied to \( X \); and
- \( R_1 \cap S_2, R_2 \cap S_1 \in \text{Rel}(D) \) by closure under recognizable projections (Lemma 8).

It only remains to observe that \( R \cap S = (R_1 \cap S_1) \cup (R_1 \cap S_2) \cup (R_2 \cap S_1) \cup (R_2 \cap S_2) \) and obtain that \( R \cap S \in \text{Rel}(C) \) due to closure under union (Lemma 8).

On the other hand, 7 \( \Rightarrow \) 3 can be derived from 1 \( \Rightarrow \) 3. Indeed, suppose that \( C \subseteq_{\text{Rel}} D \) for some Parikh-injective language \( D \). By Lemma 10, \( \text{Rel}(D) \) is closed under intersection and so, by 1 \( \Rightarrow \) 3, \( \text{Rel}(C) \cap \subseteq \text{Rel}(D) \cap \subseteq \text{Rational} \).

For 6 \( \Rightarrow \) 7, suppose that \( C =_{\text{Rel}} D \cup X \) for some Parikh-injective languages \( D, X \) such that \( X \subseteq_{\text{Rel}} 1^* 2^* \). By Lemma 1, closure under complement of semi-linear sets and Parikh’s Theorem [25], it follows that there exists \( \hat{D} \subseteq_{\text{reg}} 2^* \) Parikh-injective such that \( \pi(\hat{D}) = \mathbb{N}^2 \setminus \pi(D) \). Note that \( D \cup \hat{D} \) is Parikh-bijective. Since \( D, \hat{D} \) is Parikh-surjective, by Lemma 6, item 5, \( X \subseteq_{\text{Rel}} 1^* 2^* \subseteq \text{Rel}(D) \cup \hat{D} \) and so, by Lemma 4, item 3 plus closure under union of \( D \cup \hat{D} \), we have \( C =_{\text{Rel}} D \cup X \subseteq_{\text{Rel}} D \cup \hat{D} \).

The main difficulty will lie on the proofs of 4 \( \Rightarrow \) 5 and 5 \( \Rightarrow \) 6. For 4 \( \Rightarrow \) 5, we will prove the contrapositive statement:

\[ \text{Proposition 14.} \] If \( C \subseteq_{\text{reg}} 2^* \) satisfies any of the bad conditions, then there exist \( R, S \in \text{Rel}(C) \) such that \( R \cap S \notin \text{Rational} \).

To prove 5 \( \Rightarrow \) 6, we define some basic regular languages over \( 2 \) that we call basic injective. A language \( C \subseteq 2^* \) is basic injective if it can be expressed as \( u^* v^* z \) for \( u, v, z \in 2^* \) such that if \( u, v \neq \varepsilon \), then \( u \) and \( v \) are Parikh-independent. In particular this implies the following.

\[ \text{Lemma 15.} \] Every basic injective language is Parikh-injective.

\[ \text{Proof.} \] Let \( C = u^* v^* z \) be basic injective. The cases in which \( u \) and/or \( v \) are empty are straightforward. We will then assume that \( u \) and \( v \) are Parikh-independent. Suppose then that \( \pi(u^* v^* z) = \pi(u'^* v^* z) \) for some \( r, s, r', s' \in \mathbb{N} \). By Observation 2, \( r = r' \) or \( s = s' \) which concludes the proof.\]
22:10 Closure Properties of Synchronized Relations

Note that singleton sets and languages of the form $u^*z$ for $u$ an arbitrary word are basic injective. The interest of basic injective languages stems from the fact that we can prove the following two results, from which it is not hard to get $5 \Rightarrow 6$.

- **Proposition 16.** If $C \subseteq_{\text{reg}} 2^*$ does not satisfy any of the bad conditions, then $C$ is $\text{REL}$-equivalent to a finite union of basic injective languages.

- **Proposition 17.** If $C$ is a finite union of basic injective languages that are not $\text{REL}$- contained in $1^*2^*$ and $C$ does not satisfy any of the bad conditions, then $C$ is $\text{REL}$-equivalent to a Parikh-injective regular language.

To show $5 \Rightarrow 6$ from the two statements above, suppose that $C$ does not satisfy any of the bad conditions. By Proposition 16, $C =_{\text{REL}} X' \cup D'$, for $X' = \bigcup_{i \in I} X_i$ and $D' = \bigcup_{j \in J} D_j$, where $I, J$ are finite, for every $i \in I$, $X_i$ is basic injective and $X_i \subseteq_{\text{Rel}} 1^*2^*$, and for every $j \in J$, $D_j$ is basic injective and $D_j \not\subseteq_{\text{Rel}} 1^*2^*$. Note that, from the definition of basic injective plus [14, Proposition 7] plus basic properties from Lemma 4, it follows readily that for each $i \in I$, there exist $i, s, t, s' \in N$ such that $X_i$ is $\text{REL}$-equivalent to $1^i2^i2^i2^i$. Therefore $X'$ is $\text{REL}$-equivalent to a Parikh-injective language $X$ such that $X \subseteq_{\text{Rel}} 1^*2^*$. On the other hand, by Observation 11, since $D' \subseteq_{\text{Rel}} C$ and $C$ does not satisfy any of the bad condition, neither does $D'$. Hence, by Proposition 17, $D'$ is $\text{REL}$-equivalent to a Parikh-injective regular language $D$. Thus $C =_{\text{Rel}} X \cup D$ which concludes the proof.

We dedicate the rest of the section to prove Propositions 14, 16 and 17.

**Proof idea of Proposition 14**

We show the proof idea for condition A. The other two conditions follow a similar proof strategy. Suppose that condition A holds, and consider the 3-letter alphabet $\mathbb{A} = \{a_1, a_2, c\}$. Let $R, S$ be the following relations in $(\mathbb{A}^*)^2$:

$$R = \{(a_1^i \otimes a_1^j) \cdot (a_2^i \otimes a_2^j) \cdot z \otimes c^j\}, \quad S = \{(v^* \otimes \{a_1, a_2\}^*) \cdot z \otimes c^j\},$$

note that $R, S \in \text{REL}(C)$ by condition A.4. It is not hard to show that $|R \cap S| = \infty$ due to condition A.3. We show that $R \cap S \not\in \text{Rational}$. By means of contradiction, suppose there is an automaton over the alphabet $2 \times \mathbb{A}$ such that the language recognized by this automaton synchronizes $R \cap S$. Since the language is infinite, there is a non-trivial cycle $q_0 \xrightarrow{w_1} q \xrightarrow{w_2} q \xrightarrow{w_3} q_f$ inside some accepting run. By a pumping argument, it can be seen that: 1) $[w_2]$ is necessarily of the form $(a_1^i, a_1^j)$ for some $i, s, t$ partly due to A.2; 2) $(s, t) \in \langle\langle \pi(u_j)\rangle\rangle$ for some $j$; and 3) $(s, t) \in \langle\langle \pi(v)\rangle\rangle$. Since 2) plus 3) are in contradiction with A.1, it follows that $R \cap S \not\in \text{Rational}$. ▲

**Proof idea of Proposition 16**

It can be seen that one can reduce to the case in which $C$ is of the form $w_1^* \cdots w_n^* z$ with $w_i$ and $w_{i+1}$ Parikh-independent for all $i = 1, \ldots, n - 1$. For this kind of languages, if $n \leq 2$ the result follows trivially since they are already basic injective. A straightforward case inspection shows that if $n \geq 3$ then at least one of the bad conditions holds. ▲

**Proof idea of Proposition 17**

In order to prove Proposition 17 we show the following stronger statement, which gives a characterization of closure under intersection based on the decomposition into basic injective languages. We denote the commutative closure of a language $C \subseteq_{\text{reg}} 2^*$ by $[C]_e = \{w \in 2^* : \pi(w) \in \pi(C)\}$. 

Lemma 18. Given a finite set of basic injective languages \( \{B_i\} \) that are not REL-contained in \( 1^* 2^* \), the following are equivalent:

1. \( \text{REL}( \bigcup_i B_i ) \) is closed under intersection;
2. for all \( R, S \in \text{REL}( \bigcup_i B_i ) \), \( R \cap S \in \text{Rational} \);
3. \( \bigcup_i B_i \) does not satisfy any of the bad conditions;
4. for every \( i, j \), \( B_i \cup B_j \) does not satisfy any of the bad conditions;
5. for every \( i, j \), \( B_i \cap [B_j]_\pi \) is regular and \( B_i \cap [B_j]_\pi \subseteq \text{REL} B_j \);  
6. \( \bigcup_i B_i =_{\text{REL}} C \) for some Parikh-injective \( C \subseteq_{\text{reg}} 2^* \).

Proposition 17 follows from Lemma 18 since it is its implication \( 3 \Rightarrow 6 \). In order to give a proof for Lemma 18, we first define the following property, which is at the core of the next lemmas. A pair of languages \( B_1, B_2 \), is said to verify the dichotomy property if either

- \( B_1 \cap B_2 \) satisfies one of the bad conditions; or
- \( B_1 \cap [B_2]_\pi \) is regular and \( B_1 \cap [B_2]_\pi \subseteq_{\text{REL}} B_2 \).

Note that \( B_1 \cap [B_2]_\pi \) may not be regular in general, for example if \( B_1 = 1^* 2^* \) and \( B_2 = (12)^* \).

The main ingredient to prove Lemma 18 is given by the following statement.

Lemma 19. Every pair of basic injective languages \( B_1, B_2 \) such that \( B_1, B_2 \not\subseteq_{\text{REL}} 1^* 2^* \) satisfies the dichotomy property.

Proof of Lemma 18. \( 1 \Rightarrow 2 \) is trivial; \( 2 \Rightarrow 3 \) follows from the contrapositive of Proposition 14; \( 3 \Rightarrow 4 \) holds by Observation 11; and \( 4 \Rightarrow 5 \) follows from Lemma 19. For \( 5 \Rightarrow 6 \), we proceed by induction on the number of basic injective languages in \( \{B_i\} \). The base case is the empty language, which is (vacuously) Parikh-injective. For the inductive step, consider a union \( B \cup \bigcup_i B_i \). First observe that, by Lemma 15, \( B \) is Parikh-injective. By inductive hypothesis, there exists \( D \subseteq_{\text{reg}} 2^* \) Parikh-injective such that \( \bigcup_i B_i =_{\text{REL}} D \). Also, since \( B \cap [\bigcup_i B_i]_\pi = \bigcup_i (B \cap [B_i]_\pi) \) by hypothesis both \( B \cap [\bigcup_i B_i]_\pi \) and \( B \setminus [\bigcup_i B_i]_\pi \) are regular, and \( B \cap [\bigcup_i B_i]_\pi \subseteq_{\text{REL}} \bigcup_i B_i \). Now it only remains to observe that \( (B \setminus [\bigcup_i B_i]_\pi) \cup D \) is Parikh-injective and REL-equivalent to \( B \cup \bigcup_i B_i \). Finally, \( 6 \Rightarrow 1 \) follows from Lemma 10.

Decidability

We finally discuss briefly the decidability procedure to test whether a class \( \text{REL}(C) \) is closed under intersection.

Proposition 20. It is decidable wether a given \( C \subseteq_{\text{reg}} 2^* \) is such that \( \text{REL}(C) \) is closed under intersection.

Proof idea. It follows by the equivalence \( 1 \Leftrightarrow 5 \Leftrightarrow 7 \) of Theorem 12, together with the fact that the set of languages \( C \subseteq_{\text{reg}} 2^* \) for which there is a Parikh-injective language \( D \subseteq_{\text{reg}} 2^* \) such that \( C \subseteq_{\text{REL}} D \) is computably enumerable; and the fact that the set of languages \( C \subseteq_{\text{reg}} 2^* \) which satisfy any of the bad conditions is computably enumerable.

Note that whenever \( \text{REL}(C) \) is closed under intersection, it is effectively so: given \( L_1, L_2 \subseteq_{\text{reg}} C \otimes A^* \) it is possible to compute \( L \subseteq_{\text{reg}} (2 \times A)^* \) with \( [L] = [L_1] \cap [L_2] \). Indeed, by the previous proposition we can compute some Parikh-injective \( D \) such that \( C \subseteq_{\text{REL}} D \). By the results of [14], one can compute \( L_1', L_2' \subseteq_{\text{reg}} D \otimes A^* \) such that \( [L_1] = [L_1'] \) and \( [L_2] = [L_2'] \); and thus \( L = L_1' \cap L_2' \) is such that \( [L] = [L_1'] \cap [L_2'] \) due to injectivity of \( D \) and Lemma 6.
4 Closure under complement

We say that a class $C$ of $k$-ary relations is closed under complement if for every $(R, A) \in C$, \(((A^*)^k \setminus R, A) \in C$. For every $\Rel_k(C)$ and alphabet $A$, note that there is a unique largest relation $(U, A) \in \Rel_k(C)$ that contains all relations $(R, A) \in \Rel_k(C)$; this is $U = [C \otimes A^*]_k$. Thus, a natural alternative definition for complement could take $U$, instead of $(A^*)^k$, as the universe. We say that $\Rel_k(C)$ is closed under relativized complement if for all $(R, A) \in \Rel_k(C)$ we have $([C \otimes A^*]_k \setminus R, A) \in \Rel_k(C)$. In this section, we give effective characterizations of the languages $C \subseteq \Reg^*$ for which $\Rel(C)$ is closed under complement and relativized complement.

Relativized complement

We show that closure under relativized complement, perhaps surprisingly, is equivalent to closure under intersection, and therefore it is decidable whether $\Rel(C)$ is closed under relativized complement for a given $C \subseteq \Reg^*$.

Proposition 21. For $C \subseteq \Reg^*$, $\Rel(C)$ is closed under relativized complement if, and only if, $\Rel(C)$ is closed under intersection.

Proof. For the left-to-right direction, let $(R, A), (S, A) \in \Rel(C)$. Recall that $\Rel(C)$ is always closed under union and note that $R \cup S = [C \otimes A^*] \setminus (([C \otimes A^*] \setminus R) \cup ([C \otimes A^*] \setminus S))$, and therefore $(R \cap S, A) \in \Rel(C)$. For the right-to-left direction, let $L \subseteq \Reg C \otimes A^*$. We want to check that $[C \otimes A^*] \setminus [L] \in \Rel(C)$. By the characterization of the previous section (Theorem 12, implication 1 $\Rightarrow$ 6) we can assume that $C = D \cup X$, for $X \subseteq \Reg 1^*2^*$ and $X, D$ Parikh-injective. Then,

$$[C \otimes A^*] \setminus [L] = [(D \cup X) \otimes A^*] \setminus [L] = ([[D \otimes A^*] \cup (X \otimes A^*)] \setminus [L]) = (\underbrace{[[D \otimes A^*] \cup [X \otimes A^*]]}_{R} \setminus [L]) \cup (\underbrace{[X \otimes A^*] \setminus [L]}_{S}).$$

(by Lemma 6, item 3)

Since $R, S \in \Rel(C)$, by Lemma 8, $R \cup S \in \Rel(C)$, and thus $[C \otimes A^*] \setminus [L] \in \Rel(C).$ ▲

Note that if $C \subseteq \Reg^*$ is Parikh-surjective, then $[C \otimes A^*] = (A^*)^2$, and hence closure under relativized complement and closure under complement coincide. Thus, by Proposition 21:

Observation 22. If $C \subseteq \Reg^*$ is Parikh-surjective, then $\Rel(C)$ is closed under complement if, and only if, $\Rel(C)$ is closed under intersection.

Complement

Let $\Rel(C)^c$ be the closure under complement of $\Rel(C)$, i.e., the smallest class closed under complement containing $\Rel(C)$. The following lemma gives sufficient conditions for our characterization.

Lemma 23. For any $C \subseteq \Reg^*$,
1. if $C$ is Parikh-bijective, then $\Rel(C)$ is closed under complement;
2. if $\Rel(C)$ is closed under complement, then $C$ is Parikh-surjective.

Proof. For item 1, let $L \subseteq \Reg C \otimes A^*$. By item 4 of Lemma 6, $(A^*)^2 \setminus [L] = [C \otimes A^* \setminus L] \in \Rel(C)$ which concludes the proof.
For item 2, let \( L = C \otimes \{a\}^* \). Then \( (\{a\}^* \setminus \{L\} \cup \{a\}) \in \text{Rel}(C) \) and so there exists \( L' \subseteq \text{reg} C \otimes \{a\}^* \) such that \( [L'] = ([a]^*)^2 \setminus \{L\} \). Then \( [L \cup L'] = [L] \cup [L'] = ([a]^*)^2 \). Therefore, the Parikh-image of the projection of \( L \cup L' \) onto the first component must be \( \mathbb{N}^k \) and so \( C \) is Parikh-surjective since both \( L \) and \( L' \) (and hence \( L \cup L' \)) are \( \subseteq_{\text{reg}} C \otimes \{a\}^* \). ▶

From Lemma 23 plus Observation 22, we have that \( \text{Rel}(C) \) is closed under complement if, and only if, \( \text{Rel}(C) \) is closed under intersection and \( C \) is Parikh-surjective. At the end of this section, we will use this to prove that closure under complement is a decidable property.

We now give a characterization for closure under complement without referring to closure under intersection.

**Theorem 24.** For \( C \subseteq_{\text{reg}} 2^* \), the following are equivalent:
1. there exists \( D \subseteq_{\text{reg}} 2^* \) Parikh-bijection such that \( C =_{\text{Rel}} D \);
2. \( \text{Rel}(C) \) is closed under complement (i.e. \( \text{Rel}(C)^c = \text{Rel}(C) \));
3. \( \text{Rel}(C)^c \) is definable (i.e. there is \( D \subseteq_{\text{reg}} 2^* \) such that \( \text{Rel}(C)^c = \text{Rel}(D) \)).

Before proving the above theorem, we observe that we cannot obtain the third and fourth equivalent statements that we have in Theorem 12.

**Lemma 25.** There is \( C \subseteq_{\text{reg}} 2^* \) with \( \text{Rel}(C)^c \subseteq \text{Rational} \) but \( \text{Rel}(C)^c \) not definable.

**Proof.** Consider any language which is Parikh-injective but not Parikh-surjective, e.g. \( C = (12)^* \). Then, by item 2 of Lemma 23, plus Theorem 24, we have that \( \text{Rel}(C)^c \) is not definable. The result is then an immediate consequence of the following:

▶ **Claim.** If \( C \subseteq_{\text{reg}} 2^* \) is Parikh-injective, then \( \text{Rel}(C)^c \subseteq \text{Rational} \).

Indeed, by Parikh’s Theorem [25], \( \pi(C) \) is a semi-linear set and then so is \( \mathbb{N}^2 \setminus \pi(C) \) (see for example [21]). By Lemma 1, \( \mathbb{N}^2 \setminus \pi(C) = \pi(D) \) for some Parikh-injective language \( D \). It follows then that \( C \cup D \) is Parikh-bijection and so, by Lemma 23, item 1, \( \text{Rel}(C \cup D) \) is closed under complement. Then \( \text{Rel}(C)^c \subseteq \text{Rel}(C \cup D) \subseteq \text{Rational} \).

**Proof of Theorem 24.**

1. \( \Rightarrow \) 2 follows from item 1 of Lemma 23; and \( 2 \Rightarrow 3 \) is trivial.

2. \( \Rightarrow \) 1: Suppose that \( \text{Rel}(C) \) is closed under complement. By Lemma 8, \( \text{Rel}(C) \) is also closed under union and so under intersection. Therefore, by Theorem 12, there exist Parikh-injective languages \( D, X \subseteq_{\text{reg}} 2^* \) such that \( X \subseteq_{\text{Rational}} 1^*2^* \) and \( C =_{\text{Rel}} D \cup X \). It follows then that \( [D \otimes A^*] \in \text{Rel}(C) \) and so \( R = (A^*)^2 \setminus [D \otimes A^*] \in \text{Rel}(C) \). Let \( L \subseteq_{\text{reg}} (D \cup X) \otimes A^* \) be such that \( [L] = R \). By definition of \( R \), we get that \( L \subseteq_{\text{reg}} X \otimes A^* \) and so \( X' = \{u : \exists v \text{ such that } u \otimes v \in L \} \subseteq_{\text{reg}} X \). Besides, also by definition of \( R \), \( \pi(X') = \mathbb{N}^2 \setminus \pi(D) \) and so \( D \cup X' \) is Parikh-bijection. It only remains to observe that \( C =_{\text{Rel}} D \cup X' \): \( \supseteq \) is trivial and \( \subseteq \) follows from the fact that \( 1^*2^* \) is \( \text{Rel} \)-contained in any Parikh-surjective language (Lemma 6, item 5) and so \( X \subseteq_{\text{Rel}} D \cup X' \).

3. \( \Rightarrow \) 2: Let \( D \subseteq_{\text{reg}} 2^* \) such that \( \text{Rel}(C)^c = \text{Rel}(D) \). Since \( \text{Rel}(D) \) is closed under complement, by 2 \( \Rightarrow \) 1, we can assume wlog that \( D \) is Parikh-bijection. By means of contradiction, suppose that \( \text{Rel}(C) \) is not closed under complement. Therefore, by Observation 22, either \( C \) is not Parikh-surjective or \( \text{Rel}(C) \) is not closed under intersection. We show that in both cases we arrive to a contradiction. If \( \text{Rel}(C) \) is not closed under intersection, by Theorem 12 (implication \( \neg 1 \Rightarrow \neg 4 \)), there are \( R, S \in \text{Rel}(C) \) such that \( R \cap S \notin \text{Rational} \); but since \( R \cap S = (A^*)^2 \setminus ((A^*)^2 \setminus R \cup (A^*)^2 \setminus S) \in \text{Rel}(C)^c = \text{Rel}(D) \subseteq \text{Rational} \) (recall that \( \text{Rel}(D) \) is closed under union by Lemma 8), we have a contradiction. On the other hand, if \( C \) is not Parikh-surjective, there exists \( \pi \in \pi(D) \setminus \pi(C) \). Let \( u \in D \).  

STACS 2019
be such that \( \pi(u) = \emptyset \) and let us consider the singleton relation \( R = \{ [a \otimes a^n]\} \). It is clear that \((R, \{a, b\}) \in \text{Rel}(D) = \text{Rel}(C)\). Then, either \((R, \{a, b\})\) or its complement \(((a, b)^* \setminus R, \{a, b\})\) should be in \(\text{Rel}(C)\). But it is easy to see that both relations contain a tuple with Parikh-image \( x : [a \otimes a^n] \in R \) and \([u \otimes b^n] \in (a, b)^* \setminus R\). Since \( x \not\in \pi(C) \), none of the relations is in \(\text{Rel}(C)\), which is a contradiction. \(\blacktriangleleft\)

Decidability

From Observation 22 and item 2 of Lemma 23, decidability of closure under complement follows immediately: \(\text{Rel}(C)\) is closed under complement if, and only if, \(\text{Rel}(C)\) is closed under intersection and \(C\) is Parikh-surjective. The former is decidable due to Proposition 20, and the latter is decidable through Parikh’s Theorem, since universality for semi-linear sets is decidable (see, e.g., [21]).

> Proposition 26. Given \( C \subseteq \text{reg} \cdot 2^* \), testing whether \(\text{Rel}(C)\) is closed under complement is decidable.

5 Closure under concatenation, Kleene star, and projection

In this section, we characterize languages \( C \subseteq \text{reg} \cdot k^* \) such that \(\text{Rel}(C)\) is closed under concatenation, Kleene star, and projection.

- \( C \) is closed under concatenation if for all \( R, S \in C \), \( R \cdot S \in C \), where \( \cdot \) is the component-wise concatenation operation (e.g., \( \{(a, ab), (b, a)\} \cdot \{(b, c)\} = \{(ab, abc), (bb, ac)\} \)).
- \( C \) is closed under Kleene star if for all \( R \in \mathcal{C} \), \( R^* \in C \) for \( R^* = \bigcup_{i \in \mathbb{N}} R(i) \), where \( R(0) = \{ (\varepsilon, \ldots, \varepsilon) \} \), and \( R(i+1) = R \cdot R(i) \).
- \( C \) is closed under projection if for all \( (R, A) \subseteq C \) and \( K \subseteq k^* \), \( (R|_K, A) \subseteq C \), where \( R|_K \subseteq A^* \) is the projection of \( R \) onto the components in \( K \) (with \( \varepsilon \) in the other components). For example, for \( R = \{(aa, ab, b), (a, bbb, aab), (aa, ab, ba)\} \) and \( K = \{1, 2\} \) we have \( R|_K = \{(aa, ab, \varepsilon), (a, bbb, \varepsilon)\} \).

We now give characterizations for closure under concatenation and Kleene star. As we show, closure under concatenation is in fact a necessary condition for closure under Kleene star.

> Proposition 27. For every \( C, C_1, C_2, C_3 \subseteq \text{reg} \cdot k^* \),

1. \( C_1 \cdot C_2 \subseteq \text{Rel}(C_3) \) iff for every \( R_1 \in \text{Rel}(C_1) \), \( R_2 \in \text{Rel}(C_2) \) we have \( R_1 \cdot R_2 \in \text{Rel}(C_3) \);
2. \( \text{Rel}(C) \) is closed under concatenation iff \( C \subseteq \text{Rel} \cdot C \);
3. if \( \text{Rel}(C) \) is closed under Kleene star, then it is closed under concatenation; and
4. \( \text{Rel}(C) \) is closed under Kleene star iff \( C^* \subseteq \text{Rel} \cdot C \).

Proof sketch. For the left-to-right direction of item 1, let \( L_1 \subseteq \text{reg} \cdot C_1 \otimes A^* \) and \( L_2 \subseteq \text{reg} \cdot C_2 \otimes A^* \). Then we only have to observe that \( [L_1]\cdot[L_2] = [L_1 \cdot L_2] \in \text{Rel}(C_1 \cdot C_2) \subseteq \text{Rel}(C_3) \) as we wanted. The right-to-left direction follows from Lemma 8 together with property 1 of Lemma 4. Note that item 2 is a particular case of item 1.

We now turn to item 3. For simplicity assume \( k = 2 \). Suppose \( \text{Rel}(C) \) is closed under Kleene star, and take arbitrary \( R_1, R_2 \in \text{Rel}(C) \) over an alphabet \( A \). Define \( R_i^t \) over the alphabet \( \overline{A} \times \{ \overline{\text{lt}_i}, \text{lt}_i \} \) as the result of replacing every pair \((a_1 \cdots a_n, b_1 \cdots b_m) \in R_i \) with \(((a_1, \overline{\text{lt}_i}) \cdots (a_{n-1}, \overline{\text{lt}_i})(a_n, \text{lt}_i), (b_1, \overline{\text{lt}_i}) \cdots (b_{m-1}, \overline{\text{lt}_i})(b_m, \text{lt}_i)) \). Intuitively, \( \text{lt}_i \) marks the last symbols of tuples from \( R_i \). It is easy to see that \( R_1^t, R_2^t \in \text{Rel}(C) \) using closure under componentwise letter-to-letter relations. Observe that \( R_1^t \cdot R_2^t \subseteq (R_1^t \cup R_2^t)^* \) and, by closure under union and Kleene star, that \((R_1^t \cup R_2^t)^* \subseteq \text{Rel}(C) \). Let \( L \subseteq \text{reg} \cdot \overline{A} \times \{ \overline{\text{lt}_1}, \overline{\text{lt}_2} \} \). Then we only have to observe that \( [L] \cdot [L] = [L \cdot L] \subseteq \text{Rel}(L) \subseteq \text{Rel}(C) \) as we wanted. The right-to-left direction follows from Lemma 8 together with property 1 of Lemma 4. Note that item 2 is a particular case of item 1.
We discuss the decidability of paradigmatic problems within REL. For the converse, first observe that the containment problem within REL reduces to the emptiness problem for automata: for every k ∈ ℕ and C ⊆ REL k, REL k(C) is closed under projection iff REL k |K(C) |K ⊆ REL k(C) for every K ⊆ K.

Decidability

For the binary case, by previous results [14], it is decidable to test whether a synchronized class is included in another, and thus the characterizations for Kleene star and concatenation are decidable. We leave the general case as an open question.

6 Concluding remarks and future work

We discuss the decidability of paradigmatic problems within REL(C). First, note that the emptiness problem for relations reduces to the emptiness problem for automata: [L] = ∅ if, and only if, L = ∅ – and thus the emptiness problem is always decidable. Further, by the results we have shown together with Lemma 6 we obtain the following.

Lemma 29. For C ⊆ REL 2*, if REL(C) is closed under...
- intersection, then equivalence and containment problems within REL(C) are decidable;
- complement, then the universality problem within REL(C) is decidable;
- Op, then the Op operation within REL(C) is computable, for Op ∈ {intersection, complement, concatenation, Kleene star, projection}.

Proof of Lemma 29. Given L, M ⊆ REL C ⊆ A, the containment problem between [L] and [M] amounts to checking if [L] \ [M] is empty. Since REL(C) is closed under intersection, by Theorem 12, there exists a Parikh-injective language D such that C ⊆ REL D. Moreover, our decidability proof, shows that we can effectively compute such language D. Therefore, by the results on [14], we can effectively construct L′, M′ ⊆ REL D ⊆ A such that [L] = [L′], and [M] = [M′]. Then, by Lemma 6, item 3, [L] \ [M] = [L′] \ [M′] = [L′ \ M′] and so the containment problem within REL(C) reduces to the emptiness problem within REL(D). The equivalence problem obviously reduces to the containment problem.

The universality problem for ([L], A) amounts to checking whether (A*) k \ [L] is empty. Since REL(C) is closed under complement, by Theorem 24, there exists a Parikh-bijective language D such that C = REL D. As before, we can effectively compute such language D,
and therefore, by the results on [14], we can effectively construct \( L' \subseteq_{\text{reg}} D \otimes \mathbb{A}^* \) such that \([L'] = [L]\). By Lemma 6, item 4, we thus obtain \((\mathbb{A}^*)^k \backslash [L] = (\mathbb{A}^*)^k \backslash [L'] = ([D \otimes \mathbb{A}^*] \backslash [L'])\) and so the containment problem within \( \text{Rel}(C) \) reduces to the emptiness problem within \( \text{Rel}(D) \).

We prove the last item only for intersection; similar (or simpler) arguments can be used for all the other operations. Given \( L, M \subseteq_{\text{reg}} C \otimes \mathbb{A}^* \), with a similar argument than the one used in the previous item, we can effectively construct a Parikh-injective language \( D \) and \( L', M' \subseteq_{\text{reg}} D \otimes \mathbb{A}^* \) such that \([L] = [L']\), and \([M] = [M']\). Then, by Lemma 6, item 3, \([L] \cap [M] = [L'] \cap [M'] = [L' \cap M']\) and the result follows.

One can then conclude that classes of synchronized binary relations are generally “well-behaved”: a) it is decidable to test whether a class is closed under Boolean connectives; b) every synchronized class closed under intersection (resp. complement, etc.), is effectively closed under intersection (resp. complement, etc.); c) every synchronized class which is closed under Boolean connectives has decidable paradigmatic problems (in the sense of Lemma 29); d) at least for the binary case, the characterizations for Kleene star and concatenation are decidable.

We leave as future work the question of whether it is decidable to test if \( \text{Rel}(C) \) is closed under Kleene star, concatenation and projection when \( C \subseteq_{\text{reg}} k \). We also leave open the characterization for closure under complement and intersection for \( k \)-ary relations. Although it is conceivable that the same characterization for closure under intersection holds for arbitrary arity relations, we were not able to show it – the main issue is that it is not clear how to generalize the bad conditions to a \( k \)-ary alphabet, nor what would be the analog of item 6 in Theorem 12.

\[ \blacktriangleright \textbf{Conjecture 30. } \text{For every } k \in \mathbb{N} \text{ and } C \subseteq_{\text{reg}} k^*, \text{Rel}(C) \text{ is closed under intersection if, and only if, } C \subseteq_{\text{rel}} D \text{ for some Parikh-injective } D \subseteq_{\text{reg}} k^*. \]

\section*{References}


