Fundamenta Informaticae XXI (2001) 1001–1019
DOI 10.3233/FI-2012-0000
IOS Press

Logical theory of the monoid of languages over a non tally alphabet

Christian Choffrut
LIAFA
CNRS and Université Paris 7 Denis Diderot
France

Serge Grigorieff
LIAFA
CNRS and Université Paris 7 Denis Diderot
France

Abstract. We consider the first-order theory of the monoid $\mathcal{P}(A^*)$ of languages over a finite or infinite alphabet A (with at least two letters) endowed solely with concatenation lifted to sets: no set theoretical predicate or function, no constant. Coding a word u by the submonoid u^* it generates, we prove that the operation $(u^*, v^*) \mapsto (uv)^*$ and the predicate $\{(u^*, X) \mid \varepsilon \in X,\ u \in X\}$ are definable in $\langle \mathcal{P}(A^*); \cdot, = \rangle$. This allows to interpret the second-order theory of $\langle A^*; \cdot, = \rangle$ in the first-order theory of $\langle \mathcal{P}(A^*); \cdot, = \rangle$ and prove the undecidability of the Π_8 fragment of this last theory. These results involve technical difficulties witnessed by the logical complexity of the obtained definitions: the above mentioned predicates are respectively Δ_5 and Δ_7.

1. Introduction

The topic of this paper falls under the following general issue: given a monoid M, investigate the first-order theory of the power set $\mathcal{P}(M)$ endowed solely with concatenation lifted to sets (cf. the definition on top of §2), i.e. with no set theoretical predicate or function and no constant. In particular the complement of a subset is not immediately expressible. In a previous work we studied the case of the additive monoid of subsets of the set $\mathbb{N}$ of nonnegative integers (where $X + Y = \{x + y \mid x \in X,\ y \in Y\}$ for $X, Y \subseteq \mathbb{N}$) and were able to give a fairly complete account of what can and cannot be defined and on the complexity of the logic which is highly undecidable, cf. [1].

Here we consider the case of the free monoids generated by an arbitrary alphabet A which is not necessarily finite but which otherwise contains at least two letters. It can thus be viewed as an extension of the above publication since $\mathbb{N}$ is the free one generator monoid. When passing from one to several generators the situation changes drastically. We were ready to fail in extending all definability results of $\mathbb{N}$ to A^*, thus the difficulties we met did not surprise us. A moment's reflection on the probably most elementary properties expressible in the logic such as $XX = X$ (submonoids), $XM = X$ (right ideals in the sense of semigroups, cf. [3]) and $XY = YX$ (commutation) show that we can expect different solutions to these questions. Indeed, in $\mathbb{N}$ all submonoids are finitely generated and recognizable by

a finite automaton, all ideals are principal and commutation is trivial. None of this holds when there are at least two generators. The most spectacular example is the commutation where maximal subsets commuting with a given finite subset can be not recursively enumerable, [5]. One could draw the false conclusion that the more involved the structure the easier it is to prove the undecidability of the logic. Or that its undecidability would follow from the undecidability for $\mathbb{N}$ by simple transfer. However, we were not able to express the embedding of $\mathbb{N}$ in A^* which means that we had to use a different reduction. Conversely, we cannot recover the undecidability of the theory of $\mathbb{N}$ from that of A^* since the latter makes heavy use of the noncommutativity of A^*.

Without being too technical, let us illustrate the type of difficulties we encountered. Since all ideals of $\mathbb{N}$ are principal it is not difficult to show that the predicate "X is a singleton of $\mathbb{N}$" is expressible in a low level fragment of the logic. However we did not find a way the express this predicate in A^*; all we could do is encode a word u by the language u^*. Then the product of two words u and v is encoded by $(uv)^*$. It just happens that the predicate of all triples $(u^*, v^*, (uv)^*)$ is definable thanks to combinatorial properties of noncommuting words. This allowed us to interpret the second-order theory of $\langle A^*; \cdot, = \rangle$, which is equivalent to the second-order theory of arithmetics, in the first-order theory of $\langle \mathcal{P}(A^*); \cdot, = \rangle$ and to prove the undecidability of the Π_8 fragment of this last theory. This method uses the fact that the alphabet contains more than one letter and is therefore of no use for $\mathbb{N}$.

Taking another road was tempting. Since the subsets u^* are definable, why not identify $\mathbb{N}$ with any cyclic submonoid u^* and reduce the decidability problem for A^* to that of $\mathbb{N}$? But this is elusive since the undecidability of $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ as we proved it in [1] requires the possibility to speak of arbitrary subsets of $\mathbb{N}$, i.e., considering the identification of $\mathbb{N}$ with a cyclic submonoid of A^*, it requires the possibility to speak of subsets not necessarily containing the empty word. But we were only able to define the relation "X is subset of u^*" for those X which contain the empty word.

Now we give a quick review of the manuscript.

In Section 2 we group all the elementary material of logical or algebraic nature.

Section 3 is a combinatorial investigation of the maximal submonoids of a submonoid. It later provides a tool for discriminating submonoids by comparing their minimal generating sets with that of their maximal submonoids.

In Section 4 we introduce the important family of cyclic submonoids which can be identified with their (unique) generator and allow to deal with words. A useful tool is the family of their submonoids which we call "special" and which are the exact equivalent of those introduced in [1].

In Section 5 we show how we can express the cyclic submonoid generated by the product of two words in terms of the two cyclic submonoids generated by these words. This leads us to interpret the second-order theory of arithmetics in the first-order theory $\langle \mathcal{P}(A^*); \cdot, = \rangle$ proving that it sits high in the hierarchy of undecidable sets.

2. Preliminaries

The free monoid generated by the *alphabet* A is denoted A^*. Its elements are *words*. We denote by $|u|$ the length of the word u and by ε the *empty word* of length 0 which is the unit of the free monoid. The operation is the *concatenation* $(u, v) \to uv$.

Given an integer n and subsets $X, Y \subseteq A^*$, we define $XY = \{xy \mid x \in X, y \in Y\}$ and $X^0 = \{\varepsilon\}$ and $X^n = \{x_1 \cdots x_n \mid x_1, \ldots, x_n \in X\} = X \cdots X$ (n copies of X) for $n \geq 1$.

Most of the following is folklore or requires simple proofs. For the sake of completeness we recall these results with some detail.

2.1. Definability of elementary constants and predicates

Proposition 2.1. (Simple definability results in $\langle \mathcal{P}(A^*); \cdot, = \rangle$)
1. The predicate $X = \emptyset$ is Π_1.
2. The predicate $X = \{\varepsilon\}$ is Π_1. We denote it $\mathtt{Triv}(X)$.
3. The predicate $\varepsilon \in X$ is Σ_1. We denote it $\mathtt{In}_\varepsilon(X)$.
4. The predicate $X = A^*$ is $\Sigma_1 \wedge \Pi_1$.

Proof:
1. Observe that $X = \emptyset$ if and only if $\forall Y\ XY = X$. Indeed, $\emptyset Y = \emptyset$ and if $X \neq \emptyset$ then $X\emptyset = \emptyset \neq X$.
2. The set $\{\varepsilon\}$ is the neutral element of $\mathcal{P}(A^*)$ hence is the unique set X satisfying $\forall Y\ XY = YX = Y$.
3. Observe that $\varepsilon \in X$ if and only if $\exists Y\ (Y \neq \emptyset \wedge XY = Y)$. Indeed, if $\varepsilon \in X$ then $XA^* \supseteq \{\varepsilon\}A^* = A^*$ hence $XA^* = A^*$. If $\varepsilon \notin X$ and $Y \neq \emptyset$ then $\min\{|u| \mid u \in XY\} > \min\{|u| \mid u \in Y\}$ hence $XY \neq Y$.
4. Observe that $X = A^*$ if and only if $\varepsilon \in X \ \wedge\ \forall Y\ (\varepsilon \in Y \Rightarrow XY = X)$. Indeed, if $\varepsilon \in Y$ then $A^*Y \supseteq A^*$ hence $A^*Y = A^*$. If $\varepsilon \in X$ and $X \neq A^*$ and $u \notin X$ then $u \in X\{\varepsilon, u\}$ hence $X\{\varepsilon, u\} \neq X$. □

2.2. Submonoids

Definition 2.2. A subset $X \subseteq A^*$ is a *subsemigroup* if it is closed under product, i.e., $XX \subseteq X$. A subsemigroup is a *submonoid* if it contains the empty word ε. Equivalently, a submonoid is a subset which contains the empty word and satisfies the condition $XX = X$.

The *submonoid generated* by X, denoted X^*, is the minimal submonoid containing X, i.e. the union of all X^n, $n \in \mathbb{N}$, namely $X^* = \bigcup_{n \geq 0} X^n$. The subset X is a *generating subset* of X^*.

A trivial consequence of the definition (and of Lemma 2.4 below).

Proposition 2.3. The following relations are Σ_1 definable in the structure $\langle \mathcal{P}(A^*); \cdot, = \rangle$:

$$\mathtt{Mon} = \{M \mid M \text{ is a submonoid}\} \quad , \quad \mathtt{Sub} = \{(M, N) \mid M, N \text{ are submonoids and } M \subseteq N\}$$

The importance of the submonoids of A^* relies on the fact that they provide special cases of two important relations, namely subset inclusion $X \subseteq M$ and membership $x \in X$.

2.3. A restricted case of inclusion

When M is a submonoid and X contains ε, inclusion $X \subseteq M$ is expressible in $\langle \mathcal{P}(A^*); \cdot, = \rangle$.

Lemma 2.4. Let M be a submonoid and $\varepsilon \in X \subseteq A^*$. Then

$$X \subseteq M \iff MX = M \iff XM = M$$

Proof:
If $M = MX$ then $M = MX \supseteq \{\varepsilon\}X = X$. Conversely, if $X \subseteq M$ then $MX \subseteq MM = M$ (since M is a submonoid) and $M \subseteq MX$ (since $\varepsilon \in X$) hence $MX = M$. Idem with equality $M = XM$. □

We can push the previous result to products of two monoids.

Lemma 2.5. Let M, N be two submonoids and $\varepsilon \in X \subseteq A^*$. Then

$$X \subseteq MN \iff MXN = MN$$

Proof:
If $MN = MXN$ then $MN = MXN \supseteq \{\varepsilon\}X\{\varepsilon\} = X$. Conversely, if $X \subseteq MN$ then $MXN \subseteq M(MN)N = MN$ and $MN \subseteq MXN$ (since $\varepsilon \in X$) hence $MXN = MN$. □

2.4. A restricted case of membership

We give a special case of membership. It is technical but this tour de force should be judged by the fact that we can express the predicate $u \in X$ under some hypotheses, without being able to express the inclusion of two subsets or the fact that a subset is a singleton.

Lemma 2.6. Suppose $\varepsilon \in X$ and $\varepsilon \in Z$ and u is a shortest non empty word of Z. The equivalences

$$u \in X \iff (Z \setminus \{u\})\, X = ZX \iff X\,(Z \setminus \{u\}) = XZ \qquad (*)$$

hold in the following two cases: $\begin{cases} \text{(i)} & \text{If } X \text{ is a submonoid of } A^* \\ \text{(ii)} & \text{If } uX \subseteq Z \end{cases}$.

Proof:
Let $T = Z \setminus \{u\}$. Observe that $ZX = TX \cup uX$. We first prove the right to left implication of $(*)$.

Suppose $ZX = TX$. Then $uX \subseteq ZX = TX$. Since $\varepsilon \in X$ we have $u \in uX$ hence $u \in TX$ hence $u = vx$ with $v \in Z$ and $v \neq u$ and $x \in X$. Since u has shortest length in $Z \setminus \{\varepsilon\}$, we have $v = \varepsilon$ or $|v| \geq |u|$. But $v \neq u$ hence $|v| \geq |u|$ contradicts equality $u = vx$. Thus, $v = \varepsilon$ and $x = u$ hence $u \in X$.

We now prove the left to right implication of $(*)$.
1. Assume X is a submonoid of A^*. If $u \in X$ then $uX \subseteq XX = X = \{\varepsilon\}X \subseteq TX$ hence $ZX = TX$.
2. Assume now $uX \subseteq Z$. We first prove that $ZX = TX \cup \{u\}$. Since $\varepsilon \in X$ and $uX \subseteq Z$ we have $u \in Z \subseteq ZX$ hence $ZX \supseteq TX \cup \{u\}$. Also,

$$ZX \;=\; \overbrace{TX \cup uX \;\subseteq\; TX \cup Z}^{\text{since } uX \subseteq Z} \;=\; \overbrace{TX \cup T \cup \{u\} \;=\; TX \cup \{u\}}^{\text{since } \varepsilon \in X \text{ hence } T \subseteq TX}.$$

In particular, if $u \in X$ then $u \in TX$ hence $ZX = TX$. □

Example 2.7. Lemma 2.6 does not cover all cases of membership. For instance, let $M = \{\varepsilon, ab, aba, ab^2\}$, $X = \{\varepsilon, ab, (ab)^2, bab\}$. Then $ab \in X$ but M and X are not submonoids and $ab\,X \not\subseteq M$.

3. Generators and maximal submonoids

3.1. Generators

A remarkable well-known property is the existence of a smallest set of generators (Care: we use a notion of generator specific to language theory).

Definition 3.1. For $X \subseteq A^*$ we let $G(X) = Y \setminus Y^2Y^*$ where $Y = X \setminus \{\varepsilon\}$.

When X is a submonoid we have $YY^* = (X \setminus \{\varepsilon\})X = X \setminus \{\varepsilon\}$. Thus $Y^2Y^* = (YY^*)(YY^*)$ holds and the above condition becomes $G(X) = Y \setminus YY$.

Proposition 3.2. If M is a submonoid of A^* then $G(M)$ generates M and is included in every subset generating M. $G(M)$ is called the *minimum generator* or the *minimum generating set* of M and its elements are called the *generators* of M.

Proof:
Inclusion $G(M)^* \subseteq M$ is trivial. An easy induction on the length of words shows that every element $x \in M$ is in $G(M)^*$. This is trivial if $x = \varepsilon$ or $x \in G(M)$. Suppose $\varepsilon \neq x \notin G(M)$. Then $x = yz$ with $y, z \in M \setminus \{\varepsilon\}$. In particular, $|y|, |z| < |x|$ and, by induction hypothesis, $y, z \in G(M)^*$. A fortiori $x = yz \in G(M)^*$.

Assume by contradiction that there exists a generating set H not containing $G(M)$ and let α be an element in $G(M) \setminus H$. We can assume that $\varepsilon \notin H$. Since $M = H^*$ we have $\alpha = \beta\gamma$ where $\beta \in H$ and $\gamma \in H^* \setminus \{\varepsilon\}$. Then $\alpha \in (M \setminus \{\varepsilon\})(M \setminus \{\varepsilon\})$ which contradicts the definition of $G(M)$. □

Proposition 3.3. Let $M \neq \{\varepsilon\}$ be a submonoid of A^*. Any shortest non empty word $u \in M$ belongs to $G(M)$.

Proof:
Having shortest length, u cannot be a product of two nonempty words in M. □

3.2. Maximal submonoids

Definition 3.4. We write $X \lhd Y$ whenever X is a maximal (with respect to inclusion) proper submonoid of the submonoid Y.

Proposition 3.5. The predicate $\lhd$ is $\Sigma_1 \wedge \Pi_1$ as the trace of a Π_1 predicate on the Σ_1 family of submonoids.

Proof:
Indeed, using Propositions 2.3 and 2.4, $Y \lhd X$ if and only if

$$\mathtt{Sub}(Y, X) \ \wedge \ Y \neq X \ \wedge \ \forall Z \ \big((\mathtt{Mon}(Z) \ \wedge \ Y \subseteq Z \subseteq X) \Rightarrow (Z = Y \ \vee \ Z = X)\big) \quad \square$$

3.3. Maximal submonoids and generators

Proposition 3.6. Let M be a submonoid of A^* with $G(M)$ as minimal generating set.
1. The proper maximal submonoids of M are the sets $M \setminus \{g\}$ where $g \in G(M)$.
2. Every generator of M distinct from g is a generator of $M \setminus \{g\}$ (but there may be other ones). I.e.

$$G(M) \setminus \{g\} \subseteq G(M \setminus \{g\}) \tag{1}$$

Proof:
1. If $g \in G(M)$ then $g \notin (M \setminus \{g\})(M \setminus \{g\})$. Thus, $M \setminus \{g\}$ is a subsemigroup. Since it contains ε it is a submonoid. Finally, since $M \setminus \{g\}$ is obtained by removing only one element to M, it is necessarily a maximal submonoid. Conversely, if N is a maximal submonoid of M then there is at least one generator g of M outside N. Thus, N is a submonoid of $M \setminus \{g\}$. Since N is maximal we have $N = M \setminus \{g\}$.
2. Finally, letting $S = M \setminus \{\varepsilon\}$, for $g \in G(M)$, we have

$$\big(S \setminus (SS)\big) \setminus \{g\} \quad \subseteq \quad (S \setminus \{g\}) \setminus \big((S \setminus \{g\})(S \setminus \{g\})\big)$$

since $(S \setminus SS) \cap (S \setminus \{g\})(S \setminus \{g\}) = \emptyset$ and $(S \setminus SS) \setminus \{g\} \subseteq S \setminus \{g\}$. Thus, every M-generator distinct from g is an $(M \setminus \{g\})$-generator. □

Let us state a practical way to compute $G(M \setminus \{g\})$ by ruling out those elements of $M \setminus \{g\}$ which cannot possibly be generators.

Lemma 3.7. Let M be a submonoid of A^* and $g \in G(M)$. Then

$$G(M \setminus \{g\}) = G\big(G(M) \setminus \{g\} \cup gG(M) \cup G(M)g \cup gG(M)g\big) \tag{2}$$

Proof:
Indeed, let $B = \{a\} \cup C$, $a \notin C$ be an alphabet in one-to-one correspondence with the generators of M where a corresponds to g and let $\varphi : B^* \to M$ be the canonical morphism. Then $\varphi(B^* \setminus \{a\}) = M \setminus \{g\}$. Every word different from the letter a is a product of words in the (finite) set $C \cup aB \cup Ba \cup aBa$. This is checked by a direct computation on the words of length less than or equal to 3. All words u of length greater than or equal to 4 can be written as $u = vw$ with $|v| = 2$ and $|w| \geq 2$ and we argue by induction. Furthermore, it is clear that no word in $C \cup aB \cup Ba \cup aBa$ is a product of two or more elements of this set. This shows that

$$G(B^* \setminus \{a\}) = C \cup aB \cup Ba \cup aBa \tag{3}$$

Equality 2 is a consequence of the inclusion $\varphi(G(B^* \setminus \{a\})) \supseteq G(\varphi(B^* \setminus \{a\})) = G(M \setminus \{g\}))$. □

Remark 3.8. The expression 3 is an illustration, as stated in Proposition 3.6 claim 2, that every generator different from g is a generator of $M \setminus \{g\}$ but there exist other generators. In this particular case, if B^* has k generators then $B^* \setminus \{a\}$ has $4k - 2$ generators.

3.4. Creative rank

The following notions are crucial for the definition to be given in §5.2 of the function $(u^*, v^*) \mapsto (uv)^*$, i.e. a version of concatenation in A^* which makes sense in the monoid $\mathcal{P}(A^*)$ (where there are no words, only sets of words).

Proposition 3.9. Let M be a submonoid of A^* and g a generator of M and h a generator of $M \setminus \{g\}$. The following conditions are equivalent.

(i) h is not a generator of M,

(ii) $\forall Z \ (M \setminus \{g\} = Z \iff M \setminus \{g,h\} \lhd Z \lhd M)$.

Proof:
There are only two sets Z such that $M \setminus \{g,h\} \subsetneq Z \subsetneq M$, namely $M \setminus \{g\}$ and $M \setminus \{h\}$. Also, $M \setminus \{g,h\} \lhd M \setminus \{h\} \lhd M$ holds if and only if $M \setminus \{h\}$ is a submonoid of M. Thus, condition (ii) holds if and only if $M \setminus \{h\}$ is not a submonoid of M if and only (i) holds. □

Proposition 3.9 can be restated in terms of maximal submonoids.

Proposition 3.10. Let M, N, P be submonoids of A^* such that $P \lhd N \lhd M$. There exists a unique set Q such that $Q \neq N$ and $P \subsetneq Q \subsetneq M$. Moreover, the following conditions are equivalent.

(i) Q is not a submonoid,

(ii) $\forall Z \ (N = Z \iff P \lhd Z \lhd M)$.

Definition 3.11. Let M be a submonoid of A^* and $k \in \mathbb{N} \setminus \{0\}$.
1. A generator g of M is *k-creative* (or has creative rank k) if there exists exactly k generators of $M \setminus \{g\}$ which are not generators of M, i.e. $G(M \setminus \{g\}) \setminus G(M)$ has exactly k elements.
2. A maximal submonoid N of M is *k-creative* if $N = M \setminus \{g\}$ with g a *k-creative* generator.
3. We shall write $(\geq \ell)$-creative to mean k-creative for some $k \geq \ell$.

E.g., as observed in Remark 3.8, every maximal proper submonoid of the free monoid generated by k elements is $3k-1$-creative. As a corollary of Proposition 3.9, we get

Proposition 3.12. The following relations are respectively Σ_2 and $\Sigma_2 \wedge \Pi_2$:

$$\begin{aligned} \mathit{Creative}_k^{\geq} &= \{(N,M) \mid N \lhd M \text{ and } N \text{ is } (\geq k)\text{-creative}\} \\ \mathit{Creative}_k^{=} &= \{(N,M) \mid N \lhd M \text{ and } N \text{ is } k\text{-creative}\} \end{aligned}$$

Proof:
Let $\mathtt{Creative}_k^{\geq}(N,M)$ be the formula

$$N \lhd M \ \wedge \ \exists L_1, \ldots, L_k \ \left(\bigwedge_i L_i \lhd N \ \wedge \bigwedge_{i \neq j} L_i \neq L_j \right.$$
$$\left. \wedge \ \bigwedge_i \forall L \ (\mathtt{Mon}(L) \Rightarrow (L = N \Leftrightarrow L_i \subsetneq L \subsetneq M)) \right)$$

Since the family of submonoids is Σ_1 and the predicate $\lhd$ is $\Sigma_1 \wedge \Pi_1$ (cf. Propositions 2.3, 3.5), and inclusion is quantifier free for submonoids, the above formula is Σ_2.

Let $\texttt{Creative}_k^{=}$ be $\texttt{Creative}_k^{\geq} \wedge \neg \texttt{Creative}_{k+1}^{\geq}$. Proposition 3.10 insures that $\texttt{Creative}_k^{\geq}$ and $\texttt{Creative}_k^{=}$ define the two considered relations. □

Proposition 3.3 can be improved.

Proposition 3.13. Let $M \neq \{\varepsilon\}$ be a submonoid of A^*. Every shortest word $u \in M \setminus \{\varepsilon\}$ is a generator of M which is (≥ 2)-creative.

Proof:
Let v be a minimal proper extension of u lying in M. We show that u^2 and vu witness that u is (≥ 2)-creative. It is clear that u^2 and vu are distinct and are not generators of M. Thus, it suffices to prove that u^2 and vu are generators of $M \setminus \{u\}$. Since u has minimal length in $M \setminus \{\varepsilon\}$, equation $u^2 = xy$ has no solution x, y in $M \setminus \{u, \varepsilon\}$. Thus, u^2 is a generator of $M \setminus \{u\}$. To show that vu is a generator of $M \setminus \{u\}$, assume by way of contradiction that $vu = xy$ with $x, y \in M \setminus \{u, \varepsilon\}$. Since u has shortest length in $M \setminus \{\varepsilon\}$ and u is a prefix of v, equation $vu = xy$ implies that u is a prefix of x hence a proper prefix of x (recall $x \neq u$). Since v is a minimal proper extension of u and $vu = xy$, v is a prefix of x. Now, u is a suffix of y since u has shortest length and $vu = xy$. Length consideration then insures that $v = x$ and $u = y$. This last equality contradicts the hypothesis $y \neq u$. □

4. Commutative submonoids

The purpose of this section is to develop the machinery which will allow us in §5.2 to define concatenation as the ternary relation $\{(u^*, v^*, w^*) \mid w = uv\}$ on submonoids.

We shall use the following well-known fact about submonoids of $\langle \mathbb{N}; + \rangle$.

Lemma 4.1. 1. Every submonoid of $\langle \mathbb{N}; + \rangle$ different from $\{0\}$ is of the form $F \cup (a + p\mathbb{N})$ where $p \geq 1$ and $F \subseteq \{0, \ldots, a-1\}$.
2. The submonoids of $\langle \mathbb{N}; + \rangle$ with no minimal supermonoid are the submonoids $p\mathbb{N}$ with $p \geq 1$.

4.1. Commutative versus monogeneous submonoids

We recall the following elementary combinatorial result on words, cf. §1.3 of [6]

Proposition 4.2. Given two words $u, v \in A^*$, the following conditions are equivalent
1. $uv = vu$,
2. there exists $w \in A^*$ such that $u, v \in w^*$,
3. $u^* \cap v^* \neq \{\varepsilon\}$.

This leads to the following very classical definition.

Definition 4.3. A word v is *primitive* if it is not of the form w^n with $n \geq 2$ (in particular, it is nonempty). The *root* of a nonempty word u is the unique primitive word v such that $u = v^n$ for some $n \geq 1$. By convention the empty word is its own root.

Proposition 4.4. 1. Every commutative submonoid of A^* is a submonoid of some unique v^* where v is primitive.
2. A submonoid M of A^* is in $\mathit{Prim} = \{v^* \mid v \in A^*,\ v \text{ is primitive}\}$ if and only if it is maximal in Com.
3. A submonoid M of A^* is in $\mathit{Word} = \{u^* \mid u \in A^*\}$ if and only if it is in Com and has no minimal supermonoid in Com.

Proof:
1. Trivial if $M = \{\varepsilon\}$. Otherwise, assume $M \subseteq u^* \cap v^*$ where u and v are primitive. Then by Proposition 4.2 $u = z^n$ and $v = z^m$ which implies $n = m = 1$ and thus $u = v$.
2. If u is primitive and $u^* \subseteq v^*$ then $u = v^n$ for some integer n which implies $n = 1$ and $v = u$. Conversely if u is nonempty and nonprimitive then $u = v^n$ with $n \geq 2$, thus $v \in v^* \setminus u^*$ and $u^* \subseteq v^*$.
3. Interpret Lemma 4.1 in v^* where v is primitive such that $M \subseteq v^*$. □

Proposition 4.5. The family Com of commutative submonoids is $\Sigma_1 \wedge \Pi_1$.

Proof:
A submonoid P of A^* is in $\mathtt{Com}$ if and only if all its words commute if and only if all its subsets containing ε commute. Using Propositions 2.1, 2.3 and Lemma 2.4, this can be expressed by the following $\Sigma_1 \wedge \Pi_1$ formula $\mathtt{Com}(P)$: $\mathtt{Mon}(P) \wedge \forall U, V\ (PU = PV = P \Rightarrow UV = VU)$. □

Proposition 4.6. 1. The family $\mathit{Word} = \{u^* \mid u \in A^*\}$ is Π_2.
2. The family $\mathit{Prim} = \{v^* \mid v \in A^*,\ v \text{ is primitive}\}$ is Π_2.

Proof:
Claims 2 and 3 of Proposition 4.4 show that

$$\begin{aligned} U \in \mathtt{Prim} &\equiv U \in \mathtt{Com} \wedge \forall P\ ((P \in \mathtt{Com} \wedge PU = P) \Rightarrow P = U) \\ U \in \mathtt{Word} &\equiv U \in \mathtt{Com} \wedge \forall P\ (P \in \mathtt{Com} \Rightarrow \neg(U \lhd P)) \end{aligned}$$

and Propositions 4.5, 3.5 and Lemma 2.4 give the complexity. □

4.2. Letters

Proposition 4.7. The family $\mathcal{R} = \{(A^* \setminus \{a\}, X) \mid a \in A,\ \varepsilon, a \in X\}$ is Δ_2.

Proof:
The generators of A^* are the letters $a \in A$ hence the maximal submonoids of A^* are the $A^* \setminus \{a\}$ for $a \in A$ (cf. Proposition 3.6). To define $\mathcal{R}$, express that $\varepsilon \in X$ and (using Lemma 2.4) $X \not\subseteq L = A^* \setminus \{a\}$:

$$\begin{aligned} (L, X) \in \mathcal{R} &\equiv \varepsilon \in X \wedge \exists Z\ (Z = A^* \wedge L \lhd Z \wedge L \neq LX) \\ &\equiv \varepsilon \in X \wedge \forall Z\ (Z = A^* \Rightarrow (L \lhd Z \wedge L \neq LX)) \end{aligned}$$

Propositions 2.1, 3.5 give the stated complexity. □

Proposition 4.8. The family $\mathit{Letter} = \{a^* \mid a \in A\}$ is Σ_2.

Proof:
A set X is in this family if and only if it is the smallest submonoid containing some given letter a :

$$\mathtt{Mon}(X) \wedge \exists Z, L\ (F(Z,L) \wedge L \neq LX \wedge \forall Y\ ((\mathtt{Mon}(Y) \wedge L \neq LY) \Rightarrow Y = YX))$$

where $F(Z,L)$ is the $\Sigma_1 \wedge \Pi_1$ formula $Z = A^* \ \wedge\ L \lhd Z$. □

4.3. Special commutative submonoids

Definition 4.9. For each integer $n \geq 1$ we pose $S_{u,n} = \{\varepsilon\} \cup u^n u^*$. These submonoids of u^* are called *special commutative*.

E.g., $S_{u,0} = S_{u,1} = u^*$ and $S_{u,2} = \{\varepsilon\} \cup u^2u^* = u^* \setminus \{u\}$ is the largest proper submonoid of u^* since the minimal generating subset of u^* is $\{u\}$, cf. Proposition 3.6 Claim 1.

Definability of the subset $S_{u,n}$ is done in Theorem 4.12 infra.

The following shows how to use $S_{u,n}$'s to test membership.

Lemma 4.10. If $\varepsilon \in X \subseteq u^*$ then $u^n \in X$ if and only if $XS_{u,n} = XS_{u,n+1}$.

Proof:
Since $u^nX \subseteq u^nu^* \subseteq S_{u,n}$ and $S_{u,n} \setminus \{u^n\} = S_{u,n+1}$, apply claim 2 of Lemma 2.6 with $M = S_{u,n}$. □

4.4. Defining each special commutative submonoid

To define each $S_{u,n}$ we carefully investigate its maximal submonoids.

Recall the notation $G(M)$ for the minimal generating set of a submonoid M, cf. Definition 3.2.

Lemma 4.11. Assume $n \geq 1$. Then

$$\begin{aligned} G(S_{u,n}) &= \{u^i \mid n \leq i \leq 2n-1\} \\ G(S_{u,n} \setminus \{u^n\}) \setminus G(S_{u,n}) &= \{u^{2n}, u^{2n+1}\} \\ G(S_{u,n} \setminus \{u^{n+1}\}) \setminus G(S_{u,n}) &= \{u^{2n+1}\} \\ G(S_{u,n} \setminus \{u^{\ell}\}) \setminus G(S_{u,n}) &= \emptyset \qquad \text{for } n+2 \leq \ell \leq 2n-1. \end{aligned}$$

Proof:
The first equation is trivial. The three other equations are immediate consequences of the following routine verifications:
$G(S_{u,n} \setminus \{u^n\}) = G(S_{u,n+1}) = \{u^i \mid n+1 \leq i \leq 2n+1\}$,
$G(S_{u,n} \setminus \{u^{n+1}\}) = \{u^n\} \cup \{u^i \mid n+2 \leq i \leq 2n-1\} \cup \{u^{2n+1}\}$,
$G(S_{u,n} \setminus \{u^{\ell}\}) = \{u^i \mid i \neq \ell \text{ and } n \leq i \leq 2n-1\}$ if $n+2 \leq \ell \leq 2n-1$. □

We now get a definition of the $S_{u,n}$'s for each fixed n.

Theorem 4.12. For each $n \geq 1$, the relation $\{(u^*, S_{u,n}) \mid u \in A^*\}$ is Π_2.

Proof:
Case $n = 1$. Since $S_{u,1} = u^*$ apply Proposition 4.6.
Case $n \geq 2$. Recall $\texttt{Word}$ is Π_2 (cf. Proposition 4.6). Let $G_i(P)$ be a Σ_2 formula expressing that P is a submonoid with at least i distinct maximal submonoids. Let $K(M, N_1, \ldots, N_n)$ be the Σ_2 formula

$$N_1 = M \ \wedge \bigwedge_{i=1}^{i=n-1} (N_{i+1} \lhd N_i \ \wedge \ G_{i+1}(N_{i+1}))$$

Applying Lemma 4.11, we express $(M, N) \in \{(u^*, S_{u,n}) \mid u \in A^*\}$ as follows:

$$\texttt{Word}(M) \ \wedge \ \forall N_1, \ldots, N_n \ (K(M, N_1, \ldots, N_n) \Rightarrow N = N_n) \ . \qquad \square$$

4.5. Derivative of a commutative submonoid

We now want a formula defining the family of $S_{u,n}$'s, $n \geq 1$ and $u \in A^*$.

Definition 4.13. If M is a submonoid of u^* with u^m as shortest nonempty element, we denote by ∂M the submonoid $M \setminus \{u^m\}$ of M.

Lemma 4.11 yields the following result about the creative rank of a maximal submonoid, cf. §3.4.

Proposition 4.14. Let M be a submonoid of u^* different from $\{\varepsilon\}$. Then ∂M is a (≥ 2)-creative maximal submonoid of M.

Proof:
Apply Proposition 3.13. $\square$

Proposition 4.14 leads to the following definition.

Definition 4.15. A submonoid of A^* is *good* if it has a unique maximal proper submonoid which is (≥ 2)-creative.

Proposition 4.14 implies the following variant of Definition 4.15.

Proposition 4.16. A submonoid of A^* is *good* if it has at most one maximal proper submonoid which is (≥ 2)-creative.

Lemma 4.17. For $n \geq 2$, $\partial S_{u,n} = S_{u,n+1}$ is the unique (≥ 2)-creative maximal submonoid of $S_{u,n}$. In particular, $S_{u,n}$ is good.

Proof:
Direct consequence of Lemma 4.11 which shows that u^n is the sole (≥ 2)-creative generator of $S_{u,n}$. $\square$

Proposition 4.18. The following relations are respectively Π_2 and $\Sigma_2 \wedge \Pi_2$:

$$\begin{aligned} \texttt{Good} &= \{(u^*, M) \mid M \text{ is a good submonoid of } u^*\} \\ \texttt{Deriv} &= \{(u^*, M, L) \mid M \text{ is a good submonoid of } u^* \text{ and } L = \partial M\} \end{aligned}$$

Proof:
By Proposition 4.16, we can express $(U, M) \in \texttt{Good}$ as follows:

$$\texttt{Word}(U) \wedge UM = U \wedge \forall P\, \forall Q\, ((\texttt{Creative}_{\geq 2}(P, M) \wedge \texttt{Creative}_{\geq 2}(Q, M)) \Rightarrow P = Q)$$

When M is good, ∂M is the unique (≥ 2)-creative maximal submonoid of M hence $\texttt{Deriv}(U, M, L)$ is definable as $\texttt{Good}(U, M) \wedge \texttt{Creative}_{\geq 2}(L, M)$. Propositions 4.6, 3.12 give the complexities. □

4.6. Defining the family of special commutative submonoids

We are now in a position to define the family of all $S_{u,n}$'s.

Theorem 4.19. The relation $\mathit{SpCom} = \{(u^*, S_{u,n}) \mid u \in A^*,\ n \geq 1\}$ is Π_3.

Proof:
Consider the following formula $\texttt{SpCom}(U, X)$ which, by Lemma 2.6, expresses that $U \in \texttt{Word}$ and X is a submonoid of U and, for any good $M \subseteq U$ with u^m as shortest nonempty element (hence $M = \{u^m\} \cup \partial M$), if $u^m \in X$ then $M \subseteq X$:

$$\texttt{Word}(U) \wedge \texttt{Sub}(X, U) \wedge \forall M\, \forall L\, \big((\texttt{Deriv}(U, M, L) \wedge MX = LX) \Rightarrow XM = X\big)$$

This property is satisfied by $(u^*, S_{u,n})$ since (cf. Lemma 2.6) equality $MX = LX$ implies that the shortest nonempty element of M, say u^m belongs to $S_{u,n}$ thus $u^m u^* \subseteq S_{u,n}$. Conversely, let (U, X) satisfy this property and let u^n be the shortest nonempty element of X. Take M as the good submonoid $S_{u,n}$ whose shortest nonempty element is u^n. By Lemma 2.6 this implies $MX = LX$. Then we get $S_{u,n} \subseteq X$ hence $S_{u,n} = X$. Finally, by Propositions 4.6, 4.18, the complexity of $\texttt{SpCom}$ is Π_3. □

Proposition 4.20. The following predicates are Π_3 :

$$\begin{aligned} \mathit{Succ}_k(U, X, Y) &\equiv (U, X, Y) \in \{(u^*, S_{u,n}, S_{u,n+k}) \mid u \in A^*,\ n \geq 1\} \\ \mathit{Succ}_*(U, X, Y) &\equiv (U, X, Y) \in \{(u^*, S_{u,n}, S_{u,n+k}) \mid u \in A^*,\ n, k \geq 1\} \end{aligned}$$

For $k = 1$ we simply write Succ in place of Succ_1.

Proof:
Observe that
$$\begin{cases} \texttt{Succ}_k(U, X_0, Y) \iff \texttt{SpCom}(U, X_0) \wedge \forall X_1 \ldots \forall X_k \\ \qquad\qquad \Big(\Big(\bigwedge_0^{k-1} \texttt{Deriv}(U, X_i, X_{i+1})\Big) \Rightarrow Y = X_k\Big) \\ \texttt{Succ}_*(U, X, Y) \iff \texttt{SpCom}(U, X) \wedge \texttt{SpCom}(U, Y) \wedge XY = X \end{cases}.$$
Then apply Theorem 4.19 and Proposition 4.18. □

5. First-order theory of concatenation on languages

5.1. Singleton sets up to an epsilon

Though singleton sets are easily obtained in the monoid $\langle \mathbb{N}; + \rangle$, it is not the case with the monoid $\langle A^*; \cdot \rangle$ for a nontally alphabet. In fact, we are not able to get singletons, only pairs $\{\varepsilon, u\}$'s. And this requires all the machinery developed in order to get the $S_{u,n}$'s.

Lemma 5.1. The relation $\mathtt{Single}_{+\varepsilon} = \{(u^*, \{\varepsilon, u\}) \mid u \in A^*\}$ is Π_4.

Proof. By Lemma 4.10, if $\varepsilon \in X$ and $X \subseteq u^*$ then $u^k \in X$ if and only if $S_{u,k+1}X = S_{u,k}X$. Thus, $(U, X) \in \mathtt{Sing}_{+\varepsilon}$ if and only if $\varepsilon \in X \subseteq U$ and $u \in X$ but $u^k \notin X$ for $k \geq 2$, which is expressed by the Π_4 formula

$$\mathtt{Word}(U) \wedge \varepsilon \in X \wedge UX = U \wedge \forall Q\, (\mathtt{Succ}(U, U, Q)) \Rightarrow QX = UX)$$
$$\wedge\ \forall P, Q\, ((P \neq U \wedge \mathtt{Succ}(U, P, Q)) \Rightarrow QX \neq PX) \quad \square$$

5.2. Concatenation on words

Theorem 5.2. The relation $\mathtt{Conc} = \{(u^*, v^*, w^*) \mid uv = w\}$ is Δ_5.

Proof:
We consider three cases.
Case 0: $u = \varepsilon$ or $v = \varepsilon$. Then $uv = w$ is expressed by the Π_2 formula

$$C_\varepsilon(U, V, W) \equiv \mathtt{Word}(U) \wedge \mathtt{Word}(V) \wedge (U = \{\varepsilon\} \vee V = \{\varepsilon\}) \wedge W = UV$$

Case 1: $uv \neq vu$ (hence $u, v \neq \varepsilon$). We claim that $w = uv$ holds if and only if $\{\varepsilon, w\}$ is included in $u^*\, v^*$ but is not included in any of the sets $(u^* \setminus \{u\})\, v^*$, $u^*\, (v^* \setminus \{v\})$.

Suppose $w = uv$. Clearly, $uv \in u^*\, v^*$. By way of contradiction, suppose $uv \in (u^* \setminus \{u\})\, v^*$, i.e. $uv = u^p v^q$ with $p \neq 1$. If $p = 0$ then $uv = v^q$ hence $q \geq 2$ (since $u \neq \varepsilon$) and $u = v^{q-1}$, contradicting the hypothesis $uv \neq vu$. If $p \geq 2$ then $v = u^{p-1}v^q$ and $u^{p-1} \neq \varepsilon$ hence $q = 0$ and $v = u^{p-1}$, again contradicting the hypothesis $uv \neq vu$. Thus, $uv \notin (u^* \setminus \{u\})\, v^*$. Similarly, $uv \notin u^*\, (v \setminus \{v\})$.

Conversely, suppose $w \in u^*\, v^*$ and $w \notin (u^* \setminus \{u\})\, v^*$, $w \notin u^*\, (v \setminus \{v\})$. Condition $w \in u^*\, v^*$ implies $w = u^p v^q$ for some $p, q \geq 0$. Conditions $u^p v^q \notin (u^* \setminus \{u\})\, v^*$ and $u^p v^q \notin u^*\, (v^* \setminus \{v\})$ respectively imply $p \notin \mathbb{N} \setminus \{1\}$ and $q \notin \mathbb{N} \setminus \{1\}$ hence $p = 1$ and $q = 1$. Thus, $w = u^p v^q = uv$.

Consider now the auxiliary predicates

$$G(U, V, W, \widetilde{U}, \widetilde{V}, X) \equiv \mathtt{Deriv}(U, \widetilde{U}) \wedge \mathtt{Deriv}(V, \widetilde{V}) \wedge \mathtt{Sing}_{+\varepsilon}(W, X)$$
$$H(U, V, W, \widetilde{U}, \widetilde{V}, X) \equiv UXV = UV \wedge \widetilde{U}XV \neq \widetilde{U}V \wedge UX\widetilde{V} \neq U\widetilde{V}\,.$$

Clearly, G means that, for some $u, v, w \in A^*$, we have $U = u^*$, $V = v^*$, $W = w^*$ and $\widetilde{U} = u^* \setminus \{u\}$, $\widetilde{V} = v^* \setminus \{v\}$, $X = \{\varepsilon, w\}$. If G holds then (by Lemma 2.5) H expresses that $\{\epsilon, w\} \subseteq u^*v^*$ but $\{\epsilon, w\} \not\subseteq u^* \setminus \{u\}$ and $\{\epsilon, w\} \not\subseteq v^* \setminus \{v\}$.

Formally, we can express $(U, V, W) \in \mathtt{Conc}$ for the present Case 1 by the formulas

$$C_1^\Sigma(U, V, W) \equiv \mathtt{Word}(U) \wedge \mathtt{Word}(V) \wedge UV \neq VU \wedge \mathtt{Word}(W)$$
$$\wedge \exists \widetilde{U}, \widetilde{V}, X\, (G(U, V, W, \widetilde{U}, \widetilde{V}, X) \wedge H(U, V, W, \widetilde{U}, \widetilde{V}, X))$$
$$C_1^\Pi(U, V, W) \equiv \mathtt{Word}(U) \wedge \mathtt{Word}(V) \wedge UV \neq VU \wedge \mathtt{Word}(W)$$
$$\wedge \forall \widetilde{U}, \widetilde{V}, X\, (G(U, V, W, \widetilde{U}, \widetilde{V}, X) \Rightarrow H(U, V, W, \widetilde{U}, \widetilde{V}, X))$$

By Proposition 4.18 and Lemma 5.1, formula C_1^Σ is Σ_5 and C_1^Π is Π_5.
Case 2: $uv = vu$ and $u, v \neq \varepsilon$. The proof consists of taking advantage of the solution in the previous

case. Indeed, loosely speaking the previous case shows that, given two noncommuting words u and v, the concatenation uv is indirectly expressible through the monoid it generates. Let z be a letter of the alphabet which does not commute with u nor v. Such a letter exists, it suffices to take a letter different from the first letter of u since both u and v start with the same letter. Observe that z is expressible up to a permutation of the alphabet. Since $uz \neq zu$ and $vz \neq zv$ hold the products zu and vz are indirectly expressible. Furthermore we have $zuvz \neq vzuz$ because v does not start with z, and thus the product $zuvz$ is again indirectly definable. Finally, it suffices to say that w is the only solution of the equation $zwz = zuvz$ by cancelling out the first and last occurrences of z.

This is implemented by introducing the variables Z, X_{zu}, X_{vz}, X_{zuvz}, X_{uvz} (to represent the sets z^*, $(zu)^*$, $(vz)^*$, $(zuvz)^*$ and $(uvz)^*$) and the auxiliary predicates

$$\begin{array}{rcl} E(U,V,W) & \equiv & \mathtt{Word}(U) \wedge \mathtt{Word}(V) \wedge \mathtt{Word}(W) \\ & & \wedge\, UV = VU \wedge U \neq \{\varepsilon\} \wedge V \neq \{\varepsilon\} \\ F^{\Sigma}(\cdots) & \equiv & D(\cdots) \wedge\ C_1^{\Sigma}(W,Z,X_{uvz}) \\ F^{\Pi}(\cdots) & \equiv & D(\cdots) \Rightarrow\ C_1^{\Pi}(W,Z,X_{uvz}) \\ \text{with}\quad D(\cdots) & \equiv & \mathtt{Letter}(Z)\ \wedge\ C_1^{\Sigma}(Z,U,X_{zu})\ \wedge\ C_1^{\Sigma}(V,Z,X_{vz}) \\ & & \wedge\, C_1^{\Sigma}(X_{zu},X_{vz},X_{zuvz})\ \wedge\ C_1^{\Sigma}(Z,X_{uvz},X_{zuvz}) \end{array}$$

The auxiliary predicates can be interpreted as follows
• C_1^{Σ} and C_1^{Π} have the same meaning as in case 1 (and defines the products of noncommuting words).
• E defines the subsets u^*, v^* and w^* and expresses the property that u and v are two nonempty commuting words.
• D expresses that z is a letter not commuting with u and v and defines the products zu, vz, $zuvz$
• F^{Σ} (resp.F^{Π}) defines existentially (resp. universally) the equality $wz = uvz$ and thus asserts $w = uv$.

Formally, we can express $(U,V,W) \in \mathtt{Conc}$ for Case 3 by the formulas

$$\begin{array}{rcl} C_2^{\Sigma}(U,V,W) & \equiv & E(U,V,W)\ \wedge\ \exists Z, X_{zu}, X_{vz}, X_{zuvz}, X_{uvz}\ \ F^{\Sigma}(\cdots) \\ C_2^{\Pi}(U,V,W) & \equiv & E(U,V,W)\ \wedge\ \forall Z, X_{zu}, X_{vz}, X_{zuvz}, X_{uvz}\ \ F^{\Pi}(\cdots) \end{array}$$

By Proposition 4.8 and the previous case, formula C_2^{Σ} is Σ_5 and C_2^{Π} is Π_5.

To conclude, by gathering the three cases, $\mathtt{Conc}$ can be expressed by the following Σ_5 and Π_5 formulas where the predicate C_ε is as in case 0:

$$\begin{array}{l} C_\varepsilon(U,V,W) \vee C_1^{\Sigma}(U,V,W) \vee C_2^{\Sigma}(U,V,W) \\ C_\varepsilon(U,V,W) \vee C_1^{\Pi}(U,V,W) \vee C_2^{\Pi}(U,V,W) \end{array}$$ □

Proposition 5.3. The relation $\mathtt{Pref} = \{(u^*, v^*) \mid u \text{ is a prefix of } v\}$ is Σ_5.

Proof:
Use $\mathtt{Conc}$ and $\mathtt{Word}$ to transfer to u^*, v^*, w^* the fact that u is a prefix of v if and only if $\exists w\ v = uw$. □

5.3. Special submonoids

Definition 5.4. For each $u \in A^*$ we pose $T_u = \{\varepsilon\} \cup uA^*$. These submonoids are called *special*.

Proposition 5.5. 1. $T_u \setminus \{u\}$ is a maximal submonoid of T_u.
2. The relations $\mathit{Spec} = \{(u^*, T_u) \mid u \in A^*\}$ and $\mathit{Spec}\partial = \{(u^*, T_u \setminus \{u\}) \mid u \in A^*\}$ are Π_6.

Proof:
1. Observe that u is a generator of T_u.
2. Observe that $M = T_u$ if and only if M is a submonoid of A^* and, for every word $v \neq \varepsilon$, we have $v^* \subseteq M$ if and only if u is a prefix of v.

Also, $N = T_u \setminus \{u\}$ if and only if N is a submonoid of A^* and, for every word $v \neq \varepsilon$, we have $v^* \subseteq M$ if and only if u is a proper prefix of v.

Thus, $(U, M) \in \mathtt{Spec}$ and $(U, N) \in \mathtt{Spec}\partial$ are expressed as follows:

$$\mathtt{Word}(U) \wedge \mathtt{Mon}(T) \wedge \forall V\ (\mathtt{Word}(V) \Rightarrow (TV = T \Leftrightarrow \mathtt{Pref}(U, V)))$$
$$\mathtt{Word}(U) \wedge \mathtt{Mon}(T) \wedge \forall V\ (\mathtt{Word}(V) \Rightarrow (TV = T \Leftrightarrow (\mathtt{Pref}(U, V) \wedge U \neq V))$$

which are Π_6 by Proposition 5.3. □

5.4. Membership up to an epsilon

In §2.4 we obtained a restriction of the membership relation which proved very useful subsequently. Now, we are in a position to get the membership relation only constrained by the clause $\varepsilon \in X$.

Theorem 5.6. The relation $\mathit{IsIn} = \{(u^*, X) \mid \{\varepsilon, u\} \subseteq X\}$ is Δ_7.

Proof. Since uX is always included in $T_u = \{\varepsilon\} \cup uA^*$, Lemma 2.6 insures that, if $\varepsilon \in X$ then $u \in X$ if and only if $(T_u \setminus \{u\})\, X = T_u\, X$. Thus, lettting $\mathtt{In}_\varepsilon(X)$ be the Σ_1 formula of Proposition 2.1, and using the predicates $\mathtt{Spec}$ and $\mathtt{Spec}\partial$ of Proposition 5.5, the relation $(U, X) \in \mathtt{IsIn}$ is expressed by the Σ_7 and Π_7 formulas

$$\mathtt{In}_\varepsilon(X)\ \wedge\ \exists P, Q\ (\mathtt{Spec}(U, P)\ \wedge\ \mathtt{Spec}\partial(U, Q)\ \wedge\ QX = PX)$$
$$\mathtt{In}_\varepsilon(X)\ \wedge\ \forall P, Q\ ((\mathtt{Spec}(U, P)\ \wedge\ \mathtt{Spec}\partial(U, Q)) \Rightarrow QX = PX) \qquad \square$$

5.5. Second-order theory of words up to an epsilon

Theorem 5.7. Let $\mathcal{L}$ be the language consisting of the equality predicate and a binary operation $\cdot$. Let $\mathcal{F}^{(2)}_{x_1,\ldots,x_n,X_1,\ldots,X_p}$ be the family of second-order formulas of $\mathcal{L}$ with free first-order variables $x_1, \ldots, x_n$ and free second-order variables $X_1, \ldots, X_p$. Let $\mathcal{F}^{(1)}_{\lambda_1,\ldots,\lambda_n}$ be the family of first-order formulas of $\mathcal{L}$ with free variables $\lambda_1, \ldots, \lambda_n$.

Hint: the second-order formulas are interpreted in the monoid of words whereas the first-order formulas are interpreted in the monoid of languages.

There exists a computable family of maps

$$\mathit{Trad}_{x_1,\ldots,x_n,X_1,\ldots,X_p,\delta}\ :\ \mathcal{F}^{(2)}_{x_1,\ldots,x_n,X_1,\ldots,X_p}\ \longrightarrow\ \mathcal{F}^{(1)}_{\lambda_1,\ldots,\lambda_{n+p}}$$

(with $\delta \subseteq \{X_1, \ldots, X_p\}$) such that, for any $F \in \mathcal{F}^{(2)}_{\mathbf{x},\mathbf{X}}$, $\delta \subseteq \{X_1, \ldots, X_p\}$, $u_1, \ldots, u_n \in A^*$ and $E_1, \ldots, E_p \in \mathcal{P}(A^*)$ such that $\varepsilon \in E_i$ if $X_i \in \delta$ and $\varepsilon \notin E_i$ if $X_i \notin \delta$,

$$\langle A^*; \cdot, = \rangle \models F(u_1, \ldots, u_n, E_1, \ldots, E_p) \iff \langle \mathcal{P}(A^*); \cdot, = \rangle \models \mathit{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)(u_1^*, \ldots, u_n^*, \{\varepsilon\} \cup E_1, \ldots, \{\varepsilon\} \cup E_p)$$

Moreover, if the second-order formula F is Σ_k (resp. Π_k) when first and second-order quantifications are not distinguished then the first-order formula $\mathit{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ is Σ_{6+k} (resp. Π_{6+k}).

Proof:
The Trad maps are defined by induction on F.
Case $x = y$. Then δ is empty and $\mathtt{Trad}_{x,y,\emptyset}(F)$ is $\lambda_1 = \lambda_2$.
Case $x = yz$. Then δ is empty and $\mathtt{Trad}_{x,y,\emptyset}(F)$ is $\mathtt{Conc}(\lambda_1, \lambda_2, \lambda_3)$.
Case $x \in X$. Then $\delta \in \{\emptyset, \{X\}\}$ and we use Theorem 5.6

$$\begin{cases} \mathtt{Trad}_{x,X,\emptyset}(x \in X) & \text{is} \quad \lambda_1 \neq \{\varepsilon\} \wedge \mathtt{IsIn}(\lambda_1, \lambda_2) \\ \mathtt{Trad}_{x,X,\{X\}}(x \in X) & \text{is} \quad \lambda_1 = \{\varepsilon\} \vee \mathtt{IsIn}(\lambda_1, \lambda_2) \end{cases}$$

Case $F(\mathbf{x}, \mathbf{X}) = \neg G(\mathbf{x}, \mathbf{X})$. Then $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ is $\neg \mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(G)$.
Case $F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}) = G(\mathbf{x}, \mathbf{y}, \mathbf{X}, \mathbf{Y}) \wedge H(\mathbf{x}, \mathbf{z}, \mathbf{X}, \mathbf{Z})$. Then

$$\mathtt{Trad}_{\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{X},\mathbf{Y},\mathbf{Z},\delta}(F) \quad \text{is} \quad \mathtt{Trad}_{\mathbf{x},\mathbf{y},\mathbf{X},\mathbf{Y},\delta_1}(G) \wedge \mathtt{Trad}_{\mathbf{x},\mathbf{z},\mathbf{X},\mathbf{Z},\delta_2}(H)$$

where $\delta_1 = \delta \restriction \mathbf{X}, \mathbf{Y}$ and $\delta_2 = \delta \restriction \mathbf{X}, \mathbf{Z}$.
Case $F(\mathbf{x}, \mathbf{X}) = \exists y G(y, \mathbf{x}, \mathbf{X})$. If $\mathtt{Trad}_{y,\mathbf{x},\mathbf{X},\delta}(G)$ is $H(\lambda, \lambda_1, \ldots, \lambda_{n+p})$ then

$$\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F) \quad \text{is} \quad \exists \lambda \, (\mathtt{Word}(\lambda) \wedge H(\lambda, \lambda_1, \ldots, \lambda_{n+p}))$$

Case $F(\mathbf{x}, \mathbf{X}) = \exists Y G(\mathbf{x}, \mathbf{X}, Y)$. Let $\delta \subseteq \mathbf{X}$. If

$$\begin{aligned} \mathtt{Trad}_{\mathbf{x},\mathbf{X},Y,\delta}(G) &\quad \text{is} \quad H_0(\lambda_1, \ldots, \lambda_{n+p}, \lambda) \\ \mathtt{Trad}_{\mathbf{x},\mathbf{X},Y,\delta\cup\{Y\}}(G) &\quad \text{is} \quad H_1(\lambda_1, \ldots, \lambda_{n+p}, \lambda) \end{aligned}$$

then $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ is $\exists \lambda \, (H_0(\lambda, \lambda_1, \ldots, \lambda_{n+p})) \vee H_1(\lambda, \lambda_1, \ldots, \lambda_{n+p}))$.

To conclude, observe that if F is an atomic second-order formula then $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ simply involves equality or the predicate Conc or the predicates IsIn and the constant set $\{\varepsilon\}$. Thus, if F is atomic then $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ can be taken Σ_7 and can be taken Π_7. In particular, if we existentially quantify such an atomic F then the associated Trad formula can be taken Σ_7 if we replace each positive (resp. negative) occurrence of an atomic subformula by a Σ_7 (resp. Π_7) formula. Similarly, if we universally quantify such an atomic F then the associated Trad formula can be taken Π_7.

Since the inductive construction of $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ respects the quantification pattern of F we see that if F is Σ_k (resp. Π_k) – where first and second-order quantifications are not distinguished – then $\mathtt{Trad}_{\mathbf{x},\mathbf{X},\delta}(F)$ can be taken Σ_{6+k} (resp. Π_{6+k}). □

5.6. Undecidability

The following two results essentially date back to Quine, 1946 [9], cf. also [11], 1935, and [2], 1938.

Proposition 5.8. Let $\mathcal{F}^{(1)}_{\text{arith}}$ and $\mathcal{F}^{(1)}_{\text{conc}}$ be the families of closed first-order formulas in the language $\{+, \times, =\}$ of arithmetic and in the language $\{\cdot, =\}$ of concatenation. Let $\mathcal{F}^{(2)}_{\text{arith}}$ and $\mathcal{F}^{(2)}_{\text{conc}}$ be the analogous families of second-order formulas.

For $i = 1, 2$, let $\mathit{Truth}^{(i)}_{\text{arith}}$ be the set of formulas of $\mathcal{F}^{(i)}_{\text{arith}}$ true in the structure $\langle \mathbb{N}; =, +, \times \rangle$.

For A a finite or infinite countable alphabet with at least two letters, let $\mathit{Truth}^{(i)}_{A\text{-words}}$ be the set of formulas of $\mathcal{F}^{(i)}_{\text{conc}}$ true in the structure $\langle A^*; =, \cdot \rangle$.

There exist computable bijections $\mathcal{F}^{(1)}_{\text{conc}} \xrightarrow{\varphi^A_1} \mathcal{F}^{(1)}_{\text{arith}}$ and $\mathcal{F}^{(2)}_{\text{conc}} \xrightarrow{\varphi^A_2} \mathcal{F}^{(2)}_{\text{arith}}$ such that

$$\mathit{Truth}^{(i)}_{A\text{-words}} = (\varphi^A_i)^{-1}\left(\mathit{Truth}^{(i)}_{\text{arith}}\right) \quad \text{for } i = 1, 2.$$

In other words, the first-order (resp. second-order) theory of $\langle A^*; =, \cdot \rangle$ is computably isomorphic with the first-order (resp. second-order) theory of arithmetic, In particular, both are undecidable.

Proof:
By interpreting each one of these two structures into the other, one gets computable injective maps reducing one truth set to the other. Using Myhill's isomorphism theorem (i.e. the computable analog of Cantor-Bernstein's theorem in set theory, cf. [10] Theorem VI page 85, or [8] Theorem III.7.13 page 325), one gets the wanted computable bijective map. □

Proposition 5.9. (Marchenkov, 1982)
The Π_2 fragment of the first-order theory of $\langle A^*; =, \cdot \rangle$ is undecidable

Remark 5.10. The decidability of the Σ_1 fragment of the first-order theory of $\langle A^*; =, \cdot \rangle$ is a corollary of the decidability of diophantine equations in words, a very difficult result due to Makanin, cf. [6].

As a corollary of Theorem 5.7 and Proposition 5.8, we get the following result in which formulas in $\mathcal{F}^{(1)}_{\text{conc}}$ (first-order formulas of concatenation) are considered for the monoid of languages and those in $\mathcal{F}^{(2)}_{\text{conc}}$ (second-order formulas of concatenation) are considered for the monoid of words.

Theorem 5.11. For A a finite or infinite countable alphabet with at least two letters, let $\mathit{Truth}^{(i)}_{A\text{-languages}}$ be the set of formulas of $\mathcal{F}^{(i)}_{\text{conc}}$ true in the structure $\langle \mathcal{P}(A^*); \cdot, = \rangle$. There exists computable bijections $\mathcal{F}^{(1)}_{\text{conc}} \xrightarrow{\psi_A} \mathcal{F}^{(2)}_{\text{arith}}$ and $\mathcal{F}^{(1)}_{\text{conc}} \xrightarrow{\theta_A} \mathcal{F}^{(2)}_{\text{conc}}$ such that

$$\mathit{Truth}^{(1)}_{A\text{-languages}} = (\psi_A)^{-1}\left(\mathit{Truth}^{(2)}_{\text{arith}}\right) = (\theta_A)^{-1}\left(\mathit{Truth}^{(2)}_{A\text{-words}}\right)$$

Thus, the first-order theory of $\langle \mathcal{P}(A^*); \cdot, = \rangle$ is computably isomorphic with the second-order theory of arithmetic and with the second-order theory of concatenation on A^*. In particular, it is undecidable.

As a corollary of Theorem 5.7 and Proposition 5.9, we get

Theorem 5.12. Let A be a finite or infinite alphabet with at least two letters. The Π_8 fragment of the first-order theory of $\langle \mathcal{P}(A^*); \cdot, = \rangle$ is undecidable.

Remark 5.13. In the hierarchy of undecidable sets, $\mathtt{Truth}^{(1)}_{A\text{-languages}}$ is very high. In fact,
- The halting problem for Turing machines is Σ^0_1.
- $\mathtt{Truth}^{(1)}_{\text{arith}}$ and $\mathtt{Truth}^{(1)}_{A\text{-words}}$ are Δ^1_1 and are not Σ^0_n for any $n \in \mathbb{N}$.
- $\mathtt{Truth}^{(2)}_{\text{arith}}$ and $\mathtt{Truth}^{(2)}_{A\text{-words}}$ hence also $\mathtt{Truth}^{(1)}_{A\text{-languages}}$ are Δ^2_1 and not Σ^1_n for any $n \in \mathbb{N}$.

5.7. Definable relations on sets containing ε

As another corollary of Theorem 5.7, we get

Theorem 5.14. Let $\mathcal{R} \subseteq \mathcal{X}_{\varepsilon\in} \times \cdots \times \mathcal{X}_{\varepsilon\in}$ be a family of n-tuples of subsets of A^* which are either empty or contain the empty word. Then $\mathcal{R}$ is first-order definable in $\langle \mathcal{P}(A^*); \cdot, = \rangle$ if and only if it is second-order definable in $\langle A^*; \cdot, = \rangle$.

6. Conclusion

Concerning the decision problem in the monoid $\langle \mathcal{P}(A^*); \cdot, = \rangle$, we have proved the undecidability of the Π_8 fragment of the first-order theory. We leave open the decision problem for smallest fragments.

As for the question "What is definable in the monoid $\langle \mathcal{P}(A^*); \cdot, = \rangle$?", we have proved that, for relations involving only the empty set and sets containing the empty word, first-order definability in the monoid of languages coincides with second-order definability in the monoid of words.

The empty word appears as a stumbling block for definability. In the case of the additive monoid $\langle \mathbb{N}; +, = \rangle$ we were able in [1] to prove that many relations involving sets not containing 0 are not definable. However, the used technique does not extend to the noncommutative case.

References

[1] Christian Choffrut and Serge Grigorieff. Logical theory of the additive monoid of subsets of the set of integers. In *Automata, Universality, Computation. Tribute to Maurice Margenstern* (Andrew Adamatsky editor), p. 39–74. Emergence, Complexity and Computation, vol. 12, Springer, 2014.

[2] Hans Hermes. Semiotik. Eine Theorie der Zeichengestalten als Grundlage fur Untersuchungen von formalisierten Sprachen. *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, new series, n. 5, 22 pp., Leipzig, 1938.

[3] John M. Howie. *Fundamentals of Semigroup Theory*. Clarendon Press, 1995.

[4] Artur Jez and Alexander Okhotin. Equations over sets of natural numbers with addition only. In *STACS*, 577–588, 2009.

[5] Michal Kunc. The power of commuting with finite sets of words, in *22nd Annu. Symp. Theoretical Aspects of Computer Science, STACS 2005*, *LNCS* 3404, 569-580, Springer, 2005.

[6] M. Lothaire. *Combinatorics on Words*. Addison-Wesley, 1983.

[7] S.S. Marchenkov. Undecidability of the $\forall\exists$-positive theory of a free semigroup (in Russian). *Sibirskii Matematicheskii Zhurnal,* 23:196–198, 1982.

[8] Piergiorgio Odifreddi. *Classical recursion theory. Vol. 1: The theory of functions and sets of natural integers*. North Holland, 1989.

[9] Willard V. Quine. Concatenation as a basis of arithmetic, *Journal of Symbolic Logic,* 11(4):105–114, 1946

[10] Hartley Rogers. *Theory of recursive functions and effective computability*. McGraw Hill, 1967.

[11] Alfred Tarski. Der Wahrheitebegriff in den formalisierten Sprachen, *Studia philosophica* (Lwow), 1:261–405, 1935. English translation: The concept of truth in formalized languages, in: Alfred Tarski, *Logic, Semantics, Metamathematics*, p. 152-278. Clarendon Press, 1956.