TOWARDS PARAMETRIZING WORD EQUATIONS

H. ABDULRAB¹, P. GORALČÍK² AND G.S. MAKANIN³

Abstract. Classically, in order to resolve an equation $u \approx v$ over a free monoid X^* , we reduce it by a suitable family $\mathcal F$ of substitutions to a family of equations $uf \approx vf$, $f \in \mathcal F$, each involving less variables than $u \approx v$, and then combine solutions of $uf \approx vf$ into solutions and $uf \approx vf$ into solutions of $uf \approx vf$ i

Résumé. De façon classique, on résout une équation $u \approx v$ dans le monoïde libre X^* en la réduisant par une famille convenable \mathcal{F} de substitutions en une famille d'équations $uf \approx vf$, $f \in \mathcal{F}$, chacune en moins de variables que $u \approx v$, et ensuite en combinant des solutions des $uf \approx vf$ pour obtenir des solutions de $u \approx v$. Le problème qui se pose alors est d'obtenir \mathcal{F} sous une forme commode paramétrisée. La méthode que nous proposons est basée sur la paramétrisation des traces des chemins dans le graphe des équations premières associé à $u \approx v$. Nous effectuons une telle paramétrisation dans le cas où les équations premières dans le graphe contiennent au plus trois variables.

Mathematics Subject Classification. 68R15, 20M05.

Introduction

We try to address the long standing and non solved problem in word equations of getting a "good presentation" of all the solutions of an equation. The early attempts in this direction aimed at how to systematically generate solutions. Lentin

Keywords and phrases: Equation, free monoid, parametrization, universal family.

¹ LIR/INSA de Rouen, BP. 08, 76131 Mont-Saint-Aignan Cedex, France.

 $^{^2}$ LIFAR, Université de Rouen, place E. Blondel, 76821 Mont-Saint-Aignan Cedex, France.

 $^{^3}$ Steklov Mathematical Institute, Vavilova 42, 117966 Moscow GSP-1, Russia.

[3] and Plotkin [7] presented a simple semi-algorithm to get all most general unifiers in case of solvable equations, Jaffar [1] did the same for the case of generalized equations as defined by Makanin [5].

However, what we really dream about is to describe the solutions in a closed form, by means of expressions involving variables, handy to work with.

Things would be nice indeed if the solutions of an arbitrary word equation could be represented – as it happens sometimes – by a finite number of parametric formulas, involving a finite number of constants, of word variables, and of natural parameters ranging over the set of naturals \mathbb{N} . As Hmelevskii [2] has pointed out, however, this type of parametrization would not work beyond the classe of the equations in at most three variables.

Still, the idea of parametrizing equations (i.e. the solutions thereof) can be saved if we admit parameters ranging not only over \mathbb{N} or over a finite cartesian power of \mathbb{N} , but over the free monoid $(\mathbb{N}^m)^*$ of strings of m-tuples of naturals, or over the free monoid of strings of strings of naturals $(\mathbb{N}^*)^*$, and so on.

Typically, such parameters appear when we try to enumerate the elements of a monoid M generated by a parametric family of elements of M. For example, if M is generated by a sequence c_1, \ldots, c_m of generators then it can very conveniently be parametrized by the morphism $\Phi: (\mathbf{N}^m)^* \to M$ such that $\Phi(n_1, \ldots, n_m) = c_1^{n_1} \cdots c_m^{n_m}$ for all $(n_1, \ldots, n_m) \in \mathbb{N}^m$.

This is, roughly, the idea we are going to apply, along the lines of the research programme proposed by Makanin [6], to families of elementary substitutions read out from the so called graph of prime equations associated to the equation to be solved. The graph in question will be treated piecewise, by dividing it into easily parametrisable parts called realms, in order to obtain a parametric family of substitutions taking the equation to equations involving less variables.

We demonstrate the practical usefulness of this method by applying it to the class of the equations without constants whose graphe of prime equations involves at most three variables. To confirm the method in general is a matter of further research.

In order to make the paper practically self contained, we bring all the necessary background in Sections 1 through 3. Section 4 describes the realms in the graph of prime equations in three variables and Section 5 paves the way towards parametrizing a family of substitutions efficiently reducing an equation with a three-variable prime equation.

1. Equations and solutions

The free monoid of words over a set A of symbols will be denoted by A^* , its empty word by 1, the length of $w \in A^*$ by |w|, the number of occurrences of $a \in A$ in w by $|w|_a$, and the set of symbols occurring in w by alph(w).

The monoid A^* is cancellable, $sut = svt \Rightarrow u = v$, prefix ordered, $u \leq v \Leftrightarrow \exists w \in A^* (v = uw)$, and equidivisible, $u \leq w$ and $v \leq w \Rightarrow u \leq v$ or u = v or $v \leq u$.

A morphism $f: A^* \to B^*$ of the free monoid A^* into another free monoid B^* is completely determined by an assignment of word values in B^* to the symbols of A, written as $(a \mapsto af)_{a \in A}$.

In particular, any endomorphism f of A^* which leaves all but finitely many symbols fixed is called a *substitution* and is easily described by just giving the arrows $a \mapsto af$ for $a \in Dom(f) = \{a \in A \mid af \neq a\}$.

The *composite* of a pair of morphisms $f: A^* \to B^*$ and $g: B^* \to C^*$ will be written right-hand, as $fg: A^* \to C^*$.

A morphism $g: A^* \to B^*$ divides on the left a morphism $f: A^* \to C^*$ if f = gh for some morphism $h: B^* \to C^*$.

In order to be able to formulate arbitrary word equations, we fix a countably infinite set X of word variables and build up the free monoid X^* of words called terms

Every term u, together with its alphabet ordered into a sequence, $alph(u) = \{x_1, \ldots, x_n\}$, defines a function $\varphi : (X^*)^n \to X^*$ such that $\varphi(u_1, \ldots, u_n) = f(u)$, where $f \in End(X^*)$ is the substitution $(x_i \mapsto u_i)_{1 \le i \le n}$. We have then $\varphi(x_1, \ldots, x_n) = u$, which is another way to write the term u.

An equation E is defined as a pair (u,v) of terms $u,v \in X^*$, often written as a single word $u \approx v$, with the special symbol \approx as separator. The alphabet of $E = (u \approx v)$ is then defined as alph(E) = alph(uv), and the length of E as |E| = |uv| (the symbol \approx does not count).

A pair of equations $E_1 = (u_1 \approx v_1)$ and $E_2 = (u_2 \approx v_2)$ can be multiplied into $E = E_1 E_2 = (u_1 u_2 \approx v_1 v_2)$. By abuse, we then call E_1 a prefix equation and E_2 a suffix equation of E.

An equation $E=(u\approx v)$ is called *trivial* if its terms are equal (some authors say 'graphically equal' or 'coinciding'), written u=v. Also, more generally, an equation $E=(u\approx v)$ is called *commutatively trivial* (or *balanced*) if $|u|_x=|v|_x$ for every $x\in X$, and, *semilattice trivial* if alph(u)=alph(v).

Given an arbitrary free monoid A^* , a solution of $E=(u\approx v)$ in A^* is an arbitrary monoid morphism $f:X^*\to A^*$ taking the terms of E to equal words uf=vf. In particular, an endomorphism $f\in End(X^*)$ is a solution of E if the image $Ef=(uf\approx vf)$ of E by f is a trivial equation. Obviously, transposing the terms of $E=(u\approx v)$ we get of $E'=(v\approx u)$ with exactly the same solutions.

The word variables only can take up words for values. This semantic restriction is important, because special properties of free monoids, like cancellation or equidivisibility, may enter into play. Without these strong properties only very little could be said about word equations.

One immediate consequence of the cancellation property is that for any $s,t,u,v\in X^*$, the equations $u\approx v$ and $sut\approx svt$ have the same solutions. This enables us to replace, as far as solutions are concerned, any equation E by its left cancelled form lc(E) obtained from E by cancelling its maximal trivial prefix equation. An equation E will be called left cancelled if E=lc(E).

Every nontrivial equation of the form $x_1 \cdots x_m \approx 1$, where $x_1, \cdots, x_m \in X$, has for solutions in A^* all morphisms $f: X^* \to A^*$ such that $x_1 f = \ldots = x_m f = 1$.

This permits us to restrict our attention only to the left cancelled equations with both of its terms non-void. Every such equation E is of the form $xu \approx yv$, where $u, v \in X^*$, $x, y \in X$, and $x \neq y$. The two distinct variables x and y are the so called pivots of E, written $pivot(E) = \{x, y\}$.

If $E=(u\approx v)$ is cancelled but not semilattice trivial then we can always multiply it on the right by a suitable term t, so as to get a left cancelled and semilattice trivial equation $F=(ut\approx vt)$. By the cancellation property again, F has the same solutions as E.

All told, we need only consider the following more restricted set of equations.

Definition 1. The symbol \mathcal{E} will designate throughout the set of all left cancelled semilattice trivial equations over X^* .

Note that \mathcal{E} is a monoid under the product of equations, with $1 \approx 1$ the identity. Here comes one of our key notions.

Definition 2. An equation $E \in \mathcal{E}$ is a prime equation if no non-trivial proper prefix of E is in \mathcal{E} . Put otherwise, if $E = E_1 E_2$ for $E_1 \in \mathcal{E}$ and $E_1 \neq (1 \approx 1)$ then $E_2 = (1 \approx 1)$.

Clearly, every non-trivial equation $E \in \mathcal{E}$ has a unique prime prefix equation.

To conclude with, call F similar to $E = (u \approx v)$, written $E \cong F$, if $F = (uh \approx vh)$ for some automorphism h of X^* . The solutions f of F then correspond, in an obvious way, bijectively to the solutions hf of E.

2. Universality

Any solution $f: X^* \to A^*$ of a given equation E generates a host of new solutions of E of the form fg, with $g: A^* \to B^*$ an arbitrary morphism of free monoids. In this way, a relatively small family of solutions of E may determine all the rest of them.

Definition 3. A family \mathcal{F} of endomorphisms of the free monoid X^* is called universal for an equation E if every solution of E is divided on the left by a member of \mathcal{F} . If, moreover, every $f \in \mathcal{F}$ is a solution of E then \mathcal{F} is called a universal family of solutions of E and the equation E is said to be universally solved by \mathcal{F} .

Remark. A universal family of solutions of E is called a *general solution* of E if it is minimal, in the sense that no proper subfamily of \mathcal{F} is universal. However, we are not concerned here with this property.

Here are some examples of universal solutions.

Proposition 1. Every equation $x \approx w$ such that $x \in X \setminus alph(w)$ is universally solved by the substitution $x \mapsto w$.

Proposition 2. Every nontrivial equation $x^m \approx x^n$, with $x \in V$ and $m \neq n$, is universally solved by the substitution $x \mapsto 1$.

Proposition 3. The equation $E = (xz \approx zy)$ is universally solved by the union of

$$\{(x \mapsto zx, y \mapsto xz, z \mapsto (zx)^m z) \mid m \ge 0\}$$

and

$$\{(x \mapsto x^r, y \mapsto x^r, z \mapsto x^s) \mid r, s \ge 0\}$$

Proposition 4 (Hmelevskii). The equation $E = (yxz \approx zyx)$ is universally solved by

$$\{(x \mapsto (xy)^i x, y \mapsto (yx)^j y, z \mapsto (yx)^k) \mid i, j, k \ge 0\}$$

together with

$$\{(x\mapsto 1, y\mapsto 1, z\mapsto x^k)\mid k\geq 0\}.$$

Let us now examine the notion of universal family of endomorphisms. If well chosen, a universal family \mathcal{F} may take E to a family of equations $E\mathcal{F} = \{Ef \mid f \in \mathcal{F}\}$ which are easier to solve (e.g. involving less variables or in some other way simpler). Moreover, this kind of reduction can be iterated, according to the following mechanism of cascade composition.

Proposition 5. Let $\mathcal{F} \subseteq End(X^*)$ be a universal family of endomorphisms for E, and, for every $f \in \mathcal{F}$, let \mathcal{G}_f be a universal family of endomorphisms for Ef. Then the family

$$\mathcal{G} = \bigcup \{ f \mathcal{G}_f \mid f \in \mathcal{F} \}$$

is a universal family of endomorphisms for E. If, moreover, every \mathcal{G}_f is a universal family of solutions of Ef then \mathcal{G} is a universal family of solutions of E.

Thus, finding a suitable universal family of endomorphisms for E may be seen as a simpler problem than resolution, a step towards a universal family of solutions of E

The following well known result (cf. [4]), though of limited practical use, nicely illustrates the idea of reduction via universal family.

Theorem 2.1 (Defect theorem). Let E = (u, v) be a non trivial equation with $alph(E) = \{x_1, \ldots, x_n\}$. Then the family \mathcal{F} of all substitutions f such that Dom(f) = alph(E) and $alph(xf) \subseteq \{x_1, \ldots, x_{n-1}\}$ for $x \in alph(E)$ is universal for E.

The family \mathcal{F} of this theorem becomes intractable as soon as E involves three or more variables. However, if $alph(E) = \{x, y\}$ then the family is

$$\mathcal{F} = \{ f_{ij} = (x \mapsto x^i, y \mapsto x^j) \mid i, j \ge 0 \}$$

and its members take E to

$$Ef_{ij} = (x^{i|u|_x + j|u|_y} \approx x^{i|v|_x + j|v|_y}),$$

which is trivial if and only if

$$i(|u|_x - |v|_x) = j(|u|_y - |v|_y).$$

Proposition 6. Every equation $E = (u \approx v)$ with $alph(E) = \{x, y\}$ is universally solved by the family of substitutions

$$\{(x \mapsto x^i, y \mapsto x^j) \mid i(|u|_x - |v|_x) = j(|u|_y - |v|_y), i, j \ge 0\}$$

For the equations involving more than two variables, however, more elaborate methods are needed.

3. PIVOTINGS AND DELETIONS, REALMS

A classical way to treat an equation $E \in \mathcal{E}$ is to transform it into lc(Ef) by either a pivoting $f = (x \mapsto yx)$ or by a deletion $f = (x \mapsto 1)$, where $x, y \in pivot(E)$.

The pivotings and deletions represented by labelled arrows $E \xrightarrow{f} F$ from E to F = lc(Ef) structure \mathcal{E} into a graph that will be denoted by \mathcal{E}_{PD} . The subgraph on \mathcal{E} defined by the pivotings only will be denoted by \mathcal{E}_{P} .

In \mathcal{E}_{PD} , exactly four labelled arrows issue from every non-trivial equation $E \in \mathcal{E}$ with $pivot(E) = \{x, y\}$. Their four labels

$$\{(x \mapsto yx), (y \mapsto xy), (x \mapsto 1), (y \mapsto 1)\}$$

may without danger be abbreviated to yx, xy, x, y, respectively.

The four substitutions labelling the arrows issued from E turn out to be a universal family for E. Moreover, dividing out a solution $g: X^* \to A^*$ of E by a member f of the family may take it to a solution of Ef reduced in terms of the norm defined by

$$||g||_E = \sum_{x \in alph(E)} |xg| + |alph(E)|.$$

Proposition 7. For every non-trivial equation $E \in \mathcal{E}$ and for every solution $g: X^* \to A^*$ of E, there exists an arrow $E \xrightarrow{f} F$ and a solution $h: X^* \to A^*$ of F such that g = fh and $||h||_F < ||g||_E$.

Proof. If $E=(xu\approx yv)$ then xg and yg are prefixes of (xu)g=(yv)g, thus by equidivisibility, one is a prefix of the other one. Assume that xg=(yg)w, for some $w\in A^*$. Now, if $w\neq 1\neq yg$ then we have g=fh for $f=(x\mapsto yx)$ and $h:X^*\to A^*$ such that xh=w and zh=zg for $z\neq x$; if w=1 or yg=1 then g=fh for $f=(y\mapsto 1)$ and h=g.

In both cases, h is a solution of Ef, thus also of F = lc(Ef). Moreover, we have $||h||_F < ||g||_E$, because in the first case |xh| = |xg| - |yg| and alph(F) = alph(E), and in the second $alph(F) = alph(E) \setminus \{y\}$.

Every path

$$E_0 \stackrel{f_1}{\rightarrow} E_1 \stackrel{f_2}{\rightarrow} \dots \stackrel{f_k}{\rightarrow} E_k$$

in \mathcal{E}_{PD} defines a composite substitution $f_1 \cdots f_k$ (the identity if k = 0), called the trace of the path.

Theorem 3.1. For any equation $E \in \mathcal{E}$, the traces of the paths from E to the trivial equation $1 \approx 1$ form a universal family of solutions of E.

Proof. Clearly, every trace of a path from E to $1 \approx 1$ is a solution of E. To prove the universality of the family, we proceed by the induction on norm.

The assertion is true for $1 \approx 1$. Consider a non-trivial $E \in \mathcal{E}$ and a solution $g: X^* \to A^*$ of E, assuming the assertion true for all $F \in \mathcal{E}$ admitting a solution $h: X^* \to A^*$ such that $||h||_F < ||g||_E$. The arrow $E \xrightarrow{f} F$ and the solution h of F of Proposition 7 is such. By the hypothesis, there is a path from F to $1 \approx 1$ whose trace $f_1 \cdots f_k$ divides h, say, $h = f_1 \cdots f_k \cdot h'$. Now, we have g = fh, whence $g = f \cdot f_1 \cdots f_k \cdot h'$ and $f \cdot f_1 \cdots f_k$ is a trace of a path from E to $1 \approx 1$.

We will explore the graph \mathcal{E}_{PD} with the aid of the following graph-theoretical notion.

Definition 4. A subgraph \mathcal{R} of a graph \mathcal{G} is a realm in \mathcal{G} if

- 1. \mathcal{R} is a union of strongly connected components of \mathcal{G} ;
- 2. \mathcal{R} contains a vertex from which all vertices of \mathcal{R} are reachable in \mathcal{R} ;
- 3. and every path in G conecting two vertices of R is entirely contained in R.

An escape arrow from \mathcal{R} is an arrow from a vertex in \mathcal{R} to a vertex outside \mathcal{R} .

An escape path from \mathcal{R} starts at a vertex in \mathcal{R} and is terminated by an escape arrow from \mathcal{R} .

The following is a straightforward corollary of Theorem 3.1.

Proposition 8. Let \mathcal{R} be a realm in \mathcal{E}_{PD} such that $E \in \mathcal{R}$ and $T \notin \mathcal{R}$. Then the traces of the escape paths from \mathcal{R} starting at E form a universal family of substitutions for E.

Some realms in \mathcal{E}_{PD} can be defined in terms of pivotings only.

Proposition 9. Every realm \mathcal{R} in \mathcal{E}_P is a realm in \mathcal{E}_{PD} .

Proof. If \mathcal{R} is a realm in \mathcal{E}_P then, for some $E \in \mathcal{R}$, all $F \in \mathcal{R}$ are accessible by a path of pivotings from E, whence alph(F) = alph(E). But then every path in \mathcal{E}_P connecting two vertices of \mathcal{R} consists of pivotings only, hence all its vertices belong to \mathcal{R} .

The equations of the graph of pivotings \mathcal{E}_P can be divided into the classes [E] of equations having the same prime prefix equation as E. We will call [E] the cap of E and write $[E] = u \square v$ when the prime prefix equation of E is $u \approx v$. The variables in alph(uv) then will be referred to as the $cap\ variables$.

Proposition 10. If $F \in [E]$ then pivot(F) = pivot(E) and [lc(Ff)] = [lc(Ef)] for any pivoting f defined at E.

Put otherwise, we have a factorization $E \mapsto [E]$ of the graph of pivotings \mathcal{E}_P onto the graph of cap pivotings $[\mathcal{E}_P]$, defined by the arrows $[E] \xrightarrow{f} [F]$ for $E \xrightarrow{f} F$ in \mathcal{E}_P .

The use we are going to make of this factorization appears clearly in the following proposition:

Proposition 11. For any realm S in $[\mathcal{E}_P]$, its inverse image under the factorization, $\mathcal{R} = \{F \in \mathcal{E}_P \mid [F] \in S\}$, is a realm in \mathcal{E}_{PD} . For $E \in \mathcal{R}$, the traces of the escape paths from \mathcal{R} starting at E are exactly the traces of the escape paths from S starting at E plus the traces of the paths from E to all E composed with a deletion defined at E.

In this way, the search for universal families may essentially be carried out within the quotient graph $[\mathcal{E}_P]$, with a particular attention to its strong components as building blocs of realms and to the way they are related by the escape arrows.

The task is, however, for the whole graph $[\mathcal{E}_P]$, very technical and difficult. For this reason we will have it carried out only for a relatively small portion of $[\mathcal{E}_P]$. Those who envisage to build a cathedral are adwised first to try to construct a cage for birds.

4. Caps in three variables

In this section we will apply the methods we have developed so far to the caps involving at most three cap variables. Given such a cap, we can always rename the three variables into x, y and z. Moreover, permuting the latter we can give every cap a more restricted form.

Proposition 12. Every cap [E] in two variables is isomorphic to one of the following three forms:

$$xy\Box yx$$
, $x^{m+2}y\Box yx$, $x^{m+2}y\Box y^{n+2}x$.

Every cap [F] in three variables is isomorphic to a cap of one of the following two forms:

$$x\varphi(x,y)z\Box y\psi(y,z)x$$
 (y-cap),
 $x\varphi(x,y)z\Box z\chi(y,z)x$ (z-cap),

where φ and χ contain y and ψ contains z.

Proof. Renaming the variables, we can bring the left term of [F] to the form $x\varphi(x,y)z$. The right term then has one of the three forms $y\psi(y,z)x$, $z\chi(y,z)x$, or $z\tau(x,z)y$. However, the cap $x\varphi(x,y)z\Box z\tau(x,z)y$ can be taken to $x\tau(y,x)z\Box y\varphi(y,z)x$ by $(x\mapsto y\mapsto z\mapsto x)$.

Our next task will be to subdivide the caps of the form specified by the above proposition into a well founded escape hierarchy of particular realms admitting a straightforward path enumeration. Each one of the realm forms we will have to introduce will be described as a function $\mathcal{R}_{\alpha}(m, n, \ldots, \varphi, \psi, \ldots)$ indexed by an ordinal number α and possibly depending on some natural parameters m, n, \ldots ranging over the non-negative integers, and, on some word parameters φ, ψ, \ldots ranging over $\{x, y, z\}^*$. For every index α , the easy task of checking up the graph

structure of the α -th realm will be left to the reader. The only difficulty one may have – that of justifying the escape arrows – is resolved by noting that the alledged escape arrows from \mathcal{R}_{α} always go to realms of strictly less indices and that every \mathcal{R}_{β} with $\beta < \alpha$ is disjoint from \mathcal{R}_{α} . Here come the realms.

Proposition 13. $\mathcal{R}_0 = \mathcal{C}(xy \Box yx)$ is formed by an x-loop and an y-loop, hence it has no escape arrow.

$$\mathcal{R}_1(m) = \mathcal{C}(x^{m+2}y\Box yx)$$
 is an y-loop escaped by

$$x^{m+2}y\Box yx \xrightarrow{yx} xy\Box yx \in \mathcal{R}_0.$$

$$\mathcal{R}_2(m,n) = \{x^{m+2}y \Box y^{n+2}x\}$$
 is a singleton escaped by

$$x^{m+2}y\Box y^{n+2}x \xrightarrow{yx} xy\Box y^{n+2}x \in \mathcal{R}_1(x \leftrightarrow y),$$

$$x^{m+2}y\Box y^{n+2}x \xrightarrow{xy} x^{m+2}y\Box yx \in \mathcal{R}_1.$$

Proposition 14. $\mathcal{R}_3 = \mathcal{C}(xyz\Box yzx)$ is formed by the two cycles

$$xyz\Box zyx \xrightarrow{zx} xyz\Box yzx \xrightarrow{yx} xyz\Box zyx,$$

$$xyz\Box zyx \xrightarrow{xz} yxz\Box zyx \xrightarrow{yz} xyz\Box zyx$$
,

meeting at $xyz\Box zyx$, and by the two y-loops

$$xyz\Box yzx \xrightarrow{xy} xyz\Box yzx, \ yxz\Box zyx \xrightarrow{zy} yxz\Box zyx.$$

There are no escape arrows.

Proposition 15. $\mathcal{R}_4 = \mathcal{C}(xyyz \Box yzx)$ is formed by the infinite y-path

$$\{xyx^kyz\Box yzx \mid k>0\}.$$

connected by the infinitely many x-paths

$$\{xyx^kyz\Box yzx\xrightarrow{yx}xy(yx)^kyz\Box zyx\xrightarrow{zx}xyyz\Box yzx\mid k\geq 0\}$$

to the infinitely many z-cycles

$$\{\sigma(x,y)z\Box zyx\mid \sigma\sim (xy)^{k+1}y\},\ k>0.$$

The escape arrows start at the z-caps distinct from $xy(yx)^kyz\Box zyx$, and are of the form

$$yy \cdots \Box zyx \xrightarrow{zy} yz \Box zy \in \mathcal{R}_0(x \leftrightarrow z),$$
$$xyx \cdots \Box zyx \xrightarrow{zx} xyz \Box yzx \in \mathcal{R}_3,$$
$$yxy \cdots \Box zyx \xrightarrow{zy} yxz \Box zyx \in \mathcal{R}_3.$$

Proposition 16. $\mathcal{R}_5 = \mathcal{C}(xyz\Box yzzx)$ is formed by the two cycles

$$xyz\Box yzzx\xrightarrow{yx} xyz\Box zzyx\xrightarrow{zx} xyz\Box zyzx\xrightarrow{zx} xyz\Box yzzx,$$

$$yxz\Box xzzy \xrightarrow{xy} yxz\Box zzxy \xrightarrow{zy} yxz\Box zxzy \xrightarrow{zy} yxz\Box xzzy$$

the two cross-arrows

$$xyz\Box zzyx \xrightarrow{xz} yxz\Box zxzy, \ yxz\Box zzxy \xrightarrow{yz} xyz\Box zyzx,$$

and the two loops

$$xyz\Box yzzx \xrightarrow{xy} xyz\Box yzzx, \ xyz\Box yzzx \xrightarrow{yx} xyz\Box yzzx.$$

It is escaped by

$$xyz \Box zyzx \xrightarrow{xz} yxz \Box zyx \in \mathcal{R}_3,$$
$$yxz \Box zxzy \xrightarrow{yz} xyz \Box zxy \in \mathcal{R}_3(x \mapsto z \mapsto y \mapsto x).$$

Proposition 17. Let $\sigma = \sigma(y, z)$ be a word not conjugate to $(yz)^{k+2}$ and such that $|\sigma|_y, |\sigma|_z \geq 2$, but y^3 , z^3 are not factors of σ^2 . Then $\mathcal{R}_6(\sigma) = \mathcal{C}(xyz\Box\sigma(y, z)x)$ is an x-cycle escaped by the arrows of the form

$$xyz \square zzy \cdots x \xrightarrow{xz} yxz \square zxzy \in \mathcal{R}_{5}(x \leftrightarrow y),$$

$$xyz \square zyz \cdots x \xrightarrow{xz} yxz \square zyx \in \mathcal{R}_{3},$$

$$xyz \square zyyz \cdots x \xrightarrow{xz} yxz \square zyyx \in \mathcal{R}_{4}(x \leftrightarrow z),$$

$$xyz \square y^{2}z \cdots x \xrightarrow{xy} xy \square yx \in \mathcal{R}_{0},$$

$$xyz \square yzy \cdots x \xrightarrow{xy} xyz \square yzx \in \mathcal{R}_{3},$$

$$xyz \square yzzy \cdots x \xrightarrow{xy} xyz \square yzzx \in \mathcal{R}_{5}.$$

Note that in Proposition 17 the conditions stipulated for σ make $\mathcal{R}_6(\sigma)$ disjoint from $\mathcal{R}_4(x \leftrightarrow z)$.

Proposition 18. $\mathcal{R}_7 = \mathcal{C}(xyyz\Box yzzx)$ is an infinite realm divided into two disjoint isomorphic parts, \mathcal{H} and $\mathcal{H}(y \leftrightarrow z)$. The part \mathcal{H} is formed by the x-arrow

$$xyyz\Box zyzx \xrightarrow{zx} xyyz\Box yzzx,$$

by the infinite y-path

$$\{xyx^kyz\Box yzzx, k \ge 0\},\$$

by the infinitely many x-paths

$$\{xyx^kyz\Box yzzx\xrightarrow{yx}xy(yx)^kyz\Box zzyx\xrightarrow{zx}xyyz\Box zyzx,\ k\geq 0\},$$

and by the infinitely many z-arrows

$$\{xy(yx)^kyz\Box zzyx \xrightarrow{xz} y(yx)^{k+1}z\Box zxzy, \ k \ge 0\}$$

The two parts of \mathcal{R}_7 are cross-connected by the infinitely many arrows

$$\{y(yx)^{k+1}z\Box zxzy \xrightarrow{zy} yzyx\Box xzzy, \ k \ge 0\},$$

and by their $(y \leftrightarrow z)$ -images

$$\{yxyz\Box z(zx)^{k+1}y\xrightarrow{yz}xyyz\Box zyzx,\ k\geq 0\}$$

 \mathcal{R}_7 is escaped by the arrows

$$xyyz\Box zyzx \xrightarrow{xz} yyxz\Box zyx \in \mathcal{R}_4,$$

$$yyxz\square zxzy \xrightarrow{yz} yxyz\square zxy \in \mathcal{R}_5(x \mapsto z \mapsto y \mapsto x),$$
$$y(yx)^{k+2}z\square zxzy \xrightarrow{yz} (yx)^{k+2}yz\square zxy \in \mathcal{R}_6(x \mapsto z \mapsto y \mapsto x),$$

for k > 0, and by their $(y \leftrightarrow z)$ -images.

Proposition 19. $\mathcal{R}_8(n) = \mathcal{C}(xyz\Box yz^{n+3}x)$ is formed by a y-loop at $xyz\Box yz^{n+3}x$ and by an x-cycle

$$\{xyz\Box z^iyz^{s-k+3}x\mid 0\leq i\leq n+3\}\cdot$$

It is escaped by

$$xyz\Box z^{i+3}yz^{n-i}x \xrightarrow{xz} yxz\Box (zx)^{i+2}zy \in \mathcal{R}_6(x \leftrightarrow y), \ 0 \le i \le n,$$
$$xyz\Box z^2yz^{n+1}x \xrightarrow{xz} yxz\Box zxzy \in \mathcal{R}_5(x \leftrightarrow y),$$
$$xyz\Box zyz^{n+2}x \xrightarrow{xz} yxz\Box zyx \in \mathcal{R}_3.$$

Proposition 20. Let $\sigma = \sigma(y, z)$ be a word such that $|\sigma|_y, |\sigma|_z \ge 2$ and y^3 or z^3 is a factor of σ^2 . Then $\mathcal{R}_9(\sigma) = \mathcal{C}(xyz\Box\sigma(y,z)x)$ is an x-cycle escaped by the arrows of the form

$$xyz\Box z^{r+3}y\cdots x\xrightarrow{xz}yxz\Box z(xz)^{r+2}y\in\mathcal{R}_{6}(x\leftrightarrow y),$$

$$xyz\Box zzy\cdots x\xrightarrow{xz}yxz\Box zxzy\in\mathcal{R}_{5}(x\leftrightarrow y),$$

$$xyz\Box zyz\cdots x\xrightarrow{xz}yxz\Box zyx\in\mathcal{R}_{3},$$

$$xyz\Box zyyz\cdots x\xrightarrow{xz}yxz\Box zyyx\in\mathcal{R}_{4}(x\leftrightarrow z),$$

$$xyz\Box zy^{s+3}z\cdots x\xrightarrow{xz}yxz\Box zy^{s+3}x\in\mathcal{R}_{7},$$

$$xyz\Box y^{s+2}z\cdots x\xrightarrow{xy}xy\Box yx\in\mathcal{R}_{0},$$

$$xyz\Box yzy\cdots x\xrightarrow{xy}xyz\Box yzx\in\mathcal{R}_{3},$$

$$xyz\Box yzzy\cdots x\xrightarrow{xy}xyz\Box yzzx\in\mathcal{R}_{5},$$

$$xyz\Box yzzy\cdots x\xrightarrow{xy}xyz\Box yzzx\in\mathcal{R}_{5},$$

$$xyz\Box yz^{r+3}y\cdots x\xrightarrow{xy}xyz\Box yz^{r+3}x\in\mathcal{R}_{8}.$$

Proposition 21. Every cap of the form $x^{m+1}y\varphi(x,y)z\Box yzx$, where if m=0 then φ contains x but is not of the form $x^{k+1}y$, defines the realm

$$\mathcal{R}_{10}(\varphi) = \{ x^{m+1} \varphi(x, x^k y) z \Box y z x \mid k \ge 0 \}$$

which is an infinite y-path if φ contains y and a y-loop otherwise, and is escaped by

$$x^{m+1}y\varphi(x,x^ky)z\Box yzx\xrightarrow{yx}(xy)^{m+1}\varphi(yx,(yx)^ky)z\Box zyx$$
 to $(\mathcal{R}_6\cup\mathcal{R}_9)(x\leftrightarrow z)$.

Note that in Proposition 21 the conditions stipulated for φ make $\mathcal{R}_{10}(\varphi)$ disjoint from \mathcal{R}_4 .

Proposition 22. $\mathcal{R}_{11}(r,s) = \mathcal{C}(xy^{r+1}z\Box zy^{s+1}x)$, where $r \leq s$ and $r+s \geq 2$, is a union of an x-cycle and of an z-cycle meeting at $xy^{r+1}z\Box zy^{s+1}x$, escaped by

$$xy^{r+1}z\Box y^{i+2}zy^{s-i-1}x\xrightarrow{xy}xy\Box yx\in\mathcal{R}_0,\ 0\leq i\leq s-1,$$

$$xy^{r+1}z\Box yzy^sx \xrightarrow{xy} (xy)^{r+3}z\Box yzx \in \mathcal{R}_3 \cup \mathcal{R}_4 \cup \mathcal{R}_{10},$$

and by the arrows obtained from these by transposing x and z as well as r and s.

Note that in Proposition 22 the conditions stipulated for r, s make $\mathcal{R}_{11}(r,s)$ disjoint from $\mathcal{R}_4(x \leftrightarrow z)$.

Proposition 23. $\mathcal{R}_{12}(\varphi, \psi) = \{xyx\varphi(x, y)z\Box zyz\psi(y, z)x\}$ is a singleton realm escaped by

$$xyx\varphi(x,y)z\Box zyz\psi(y,z)x \xrightarrow{zx} xyz\Box yz\psi(y,z)zx \in \mathcal{R}_5 \cup \mathcal{R}_6 \cup \mathcal{R}_9$$

and by its $(x \leftrightarrow z)$ -image (with φ and ψ exchanged).

Proposition 24. $\mathcal{R}_{13}(\psi) = \{xyxz \Box zzy\psi(y,z)x\}$ is a singleton realm escaped by

$$xyxz\Box zzy\psi(y,z)x \xrightarrow{zx} xyz\Box zy\psi(y,z)zx \in \mathcal{R}_5 \cup \mathcal{R}_6 \cup \mathcal{R}_8 \cup \mathcal{R}_9$$

$$xyxz\Box zzy\psi(y,z)x \xrightarrow{xz} yxxz\Box zxzy \in \mathcal{R}_7(x \leftrightarrow y).$$

Proposition 25. $\mathcal{R}_{14}(m,\tau,\sigma) = \{x^{m+2}y\tau(x,y)z\Box zyzy\sigma(y,z)x\}$ is a singleton realm escaped by

$$x^{m+2}y\tau(x,y)z\Box zyzy\sigma(y,z)x \xrightarrow{xz} x^{m+1}y\tau(x,y)xz\Box zyx$$

to $(\mathcal{R}_6 \cup \mathcal{R}_9)(x \leftrightarrow z)$ if τ contains y, else to $(\mathcal{R}_5 \cup \mathcal{R}_8)(x \leftrightarrow z)$, and by

$$x^{m+2}y\tau(x,y)z\Box zyzy\sigma(y,z)x\xrightarrow{zx}x(zx)^{m+1}y\Box yzy\sigma(y,z)zx$$

to $\mathcal{R}_{12}(y \leftrightarrow z)$.

Proposition 26. $\mathcal{R}_{15}(s) = \mathcal{C}(xyyz\Box yz^{s+3}x)$ is a realm formed by an x-cycle,

$$\{xyyz \Box z^j yz^{s-j+3}x \mid 0 \le j \le s+3\},\$$

by an infinite y-path,

$$\{xyx^ryz\Box yz^{s+3}x\mid r\geq 0\},\$$

and by infinitely many x-paths

$$\{xyx^{r+1}yz\Box yz^{s+3}x\xrightarrow{yx}xy(yx)^{r+1}yz\Box z^{s+3}yx\xrightarrow{zx}xyyz\Box z^{s+2}yzx\mid r\geq 0\}\cdot$$

 $\mathcal{R}_{15}(s)$ is escaped by the z-arrows from its z-caps, namely,

$$xyyz\Box zyz^{s+2}x \xrightarrow{xz} yyxz\Box zyx \in \mathcal{R}_4,$$

$$xyyz\Box z^2yz^{s+1}x \xrightarrow{xz} yyxz\Box zxzy \in \mathcal{R}_{13}(x \leftrightarrow y),$$
$$xyyz\Box z^{k+3}yz^{s-k}x \xrightarrow{xz} yyxz\Box z(xz)^{k+2}y \in \mathcal{R}_{12}(x \leftrightarrow y), \ 0 \le k \le s,$$
$$xy(yx)^{r+1}yz\Box z^{s+3}yx \xrightarrow{xz} y(yx)^{r+2}z\Box z(xz)^{s+2}y \in \mathcal{R}_{14}(x \leftrightarrow y).$$

Proposition 27. $\mathcal{R}_{16}(\sigma) = \mathcal{C}(xyyz\Box\sigma(y,z)x)$, where $|\sigma|_y, |\sigma|_z \geq 2$, is an x-cycle escaped by (not necessarily all) arrows of the form

$$xyyz \Box z^{r+3}y \cdots x \xrightarrow{xz} yyxz \Box z(xz)^{r+2}y \in \mathcal{R}_{14}(x \leftrightarrow y),$$

$$xyyz \Box zzy \cdots x \xrightarrow{xz} yyxz \Box zxzy \in \mathcal{R}_{13}(x \leftrightarrow y),$$

$$xyyz \Box zyz \cdots x \xrightarrow{xz} yyxz \Box zyx \in \mathcal{R}_{4},$$

$$xyyz \Box zy^{s+2}z \cdots x \xrightarrow{xz} yyxz \Box zy^{s+2}x \in \mathcal{R}_{11},$$

$$xyyz \Box y^{s+2}z \cdots x \xrightarrow{xy} xy \Box yx \in \mathcal{R}_{0},$$

$$xyyz \Box yzy \cdots x \xrightarrow{xy} (xy)^{2}z \Box yzx \in \mathcal{R}_{4},$$

$$xyyz \Box yzzy \cdots x \xrightarrow{xy} (xy)^{2}z \Box yzzx \in \mathcal{R}_{7},$$

$$xyyz \Box yz^{r+3}y \cdots x \xrightarrow{xy} (xy)^{2}z \Box yzzx \in \mathcal{R}_{15}.$$

Proposition 28. $\mathcal{R}_{17}(\psi) = \{x^{m+2}y\varphi(x,y)z\Box zyyz\psi(y,z)x\}$, where ψ is not of the form $(yz)^ky$ if φ is empty, is a singleton realm escaped by

$$x^{m+2}y\varphi(x,y)z\Box zyyz\psi(y,z)x\xrightarrow{zx}x(zx)^{m+1}y\Box yyz\psi(y,z)zx\in(\mathcal{R}_{13}\cup\mathcal{R}_{14})(y\leftrightarrow z),$$

$$x^{m+2}y\varphi(x,y)z\Box zyyz\psi(y,z)x\xrightarrow{xz}x^{m+1}yxz\Box zyyx\in (\mathcal{R}_7\cup\mathcal{R}_{15}\cup\mathcal{R}_{16})(x\leftrightarrow z).$$

Note that in Proposition 28 the condition stipulated for ψ make $\mathcal{R}_{17}(\varphi)$ disjoint from $\mathcal{R}_7 \cup \mathcal{R}_{15}$.

Proposition 29. Every y-cap of the form $x^{m+1}yx\varphi(x,y)z\Box yz^{s+2}x$, where $\varphi \neq x^ky$ if m=0, defines the realm

$$\mathcal{R}_{18}(m, s, \varphi) = \{x^{m+1}yx\varphi(x, x^ky)z\Box yz^{s+2}x \mid k \ge 0\},\$$

which is an infinite y-path if φ contains y and a y-loop otherwise, and is escaped by

$$x^{m+1}yx\varphi(x,x^ky)z\Box yz^{s+2}x\xrightarrow{yx}(xy)^{m+1}yx\varphi(yx,(yx)^ky)z\Box z^{s+2}yx,\ k\geq 0,$$

to
$$(\mathcal{R}_{17} \cup \mathcal{R}_{17})(x \leftrightarrow z)$$
.

Note that in Proposition 29 the conditions stipulated for φ make $\mathcal{R}_{18}(\varphi)$ disjoint from $\mathcal{R}_7 \cup \mathcal{R}_{15}$.

Proposition 30. $\mathcal{R}_{19}(n) = \{xyxz \Box z^{n+3}y\psi(y,z)x\}$ is a singleton realm escaped by

$$xyxz\Box z^{n+3}y\psi(y,z)x \xrightarrow{zx} xyz\Box z^{n+2}y\psi(y,z)zx \in \mathcal{R}_6 \cup \mathcal{R}_8 \cup \mathcal{R}_9,$$
$$xyxz\Box z^{n+3}y\psi(y,z)x \xrightarrow{xz} yxxz\Box z(xz)^{n+2}y \in \mathcal{R}_{16}(x \leftrightarrow y).$$

Proposition 31. $\mathcal{R}_{20}(m,\sigma) = \mathcal{C}(xy^{m+3}z\Box\sigma(y,z)x)$, where $|\sigma|_z \geq 2$, is an x-cycle escaped by the arrows of the form

$$xy^{m+3}z\Box z^{r+3}y\cdots x\xrightarrow{xz}y^{m+3}xz\Box z(xz)^{r+2}y\in\mathcal{R}_{14}(x\leftrightarrow y),$$

$$xy^{m+3}z\Box zzy\cdots x\xrightarrow{xz}y^{m+3}xz\Box zxzy\in\mathcal{R}_{19}(x\to z\to y\to x),$$

$$xy^{m+3}z\Box zy^{s+1}z\cdots x\xrightarrow{xz}y^{m+3}xz\Box zy^{s+1}x\in\mathcal{R}_{11}\cup\mathcal{R}_{11}(x\leftrightarrow z),$$

$$xy^{m+3}z\Box y^{s+2}z\cdots x\xrightarrow{xy}xy\Box yx\in\mathcal{R}_{0},$$

$$xy^{m+3}z\Box yzy\cdots x\xrightarrow{xy}(xy)^{m+3}z\Box yzx\in\mathcal{R}_{10},$$

$$xy^{m+3}z\Box yz^{r+2}\cdots x\xrightarrow{xy}(xy)^{m+3}z\Box yz^{s+2}x\in\mathcal{R}_{18}.$$

Proposition 32. $\mathcal{R}_{21}(r, s, \varphi, \psi) = \{xy^{r+2}x\varphi(x, y)z\Box zy^{s+1}z\psi(y, z)x\}$ is a singleton realm escaped by

$$xy^{r+2}x\varphi(x,y)z\Box zy^{s+1}z\psi(y,z)x\xrightarrow{zx}xy^{r+2}z\Box y^{s+1}z\psi(y,z)zx$$

to $\mathcal{R}_{16} \cup \mathcal{R}_{20}$ if s > 0 or ψ contains y, else to $\mathcal{R}_7 \cup \mathcal{R}_{15} \cup \mathcal{R}_{20}$, and by

$$xy^{r+2}x\varphi(x,y)z\Box zy^{s+1}z\psi(y,z)x\xrightarrow{xz}y^{r+2}x\varphi(x,y)xz\Box zy^{s+1}x$$

to
$$(\mathcal{R}_6 \cup \mathcal{R}_9 \cup \mathcal{R}_{16} \cup \mathcal{R}_{20})(x \leftrightarrow z)$$
.

Proposition 33. $\mathcal{R}_{22}(t,\varphi,\psi) = \{xyxx\varphi(x,y)z\Box z^{t+2}y\psi(y,z)x\}$ is a singleton realm escaped by

$$xyxx\varphi(x,y)z\Box z^{t+2}y\psi(y,z)x \xrightarrow{zx} xyz\Box z^{t+1}y\psi(y,z)zx$$

to $\mathcal{R}_5 \cup \mathcal{R}_6 \cup \mathcal{R}_8 \cup \mathcal{R}_9$, and by

$$xyxx\varphi(x,y)z\Box z^{t+2}y\psi(y,z)x \xrightarrow{xz} yxx\varphi(x,y)xz\Box z(xz)^{t+1}y$$

to $(\mathcal{R}_{20} \cup \mathcal{R}_{21})(x \leftrightarrow y)$.

Proposition 34. $\mathcal{R}_{23}(m,r,\varphi,\psi) = \{x^{m+2}y\varphi(x,y)z\Box zy^{r+3}z\psi(y,z)x\}$ is a singleton realm escaped by

$$x^{m+2}y\varphi(x,y)z\Box zy^{r+3}z\psi(y,z)x\xrightarrow{zx}x(zx)^{m+1}y\Box y^{r+3}z\psi(y,z)zx\in(\mathcal{R}_{14}\cup\mathcal{R}_{19})(y\leftrightarrow z),$$

$$x^{m+2}y\varphi(x,y)z\Box zy^{r+3}z\psi(y,z)x\xrightarrow{xz}x^{m+1}y\varphi(x,y)xz\Box zy^{r+3}x\in\mathcal{R}_{20}(x\leftrightarrow z).$$

Proposition 35. $\mathcal{R}_{24}(m, s, \varphi) = \{x^{m+1}yy\varphi(x, y)z\Box yz^{s+2}x \mid k \geq 0\}, \text{ where } \varphi \text{ contains } x \text{ if } m = 0, \text{ is a singleton realm escaped by}$

$$x^{m+1}yy\varphi(x,y)z\Box yz^{s+2}x \xrightarrow{yx} (xy)^{m+1}y\varphi(yx,y)z\Box z^{s+2}yx \in \mathcal{R}_{17} \cup \mathcal{R}_{23},$$

$$x^{m+1}yy\varphi(x,y)z\Box yz^{s+2}x\xrightarrow{xy}x^{m+1}yxy\varphi(x,xy)z\Box yz^{s+2}x\in\mathcal{R}_{18}.$$

Proposition 36. $\mathcal{R}_{25}(m, s, \varphi, \psi) = \{x^{m+1}y\varphi(x, y)z\Box yz^{s+1}y\psi(y, z)x\}$, where φ contains x if m = 0, is a singleton realm escaped by

$$x^{m+1}y\varphi(x,y)z\Box yz^{s+1}y\psi(y,z)x\xrightarrow{yx}(xy)^{m+1}\varphi(yx,y)z\Box z^{s+1}y\psi(y,z)yx$$

to $\mathcal{R}_6 \cup \mathcal{R}_9 \cup \mathcal{R}_{12} \cup \mathcal{R}_{14} \cup \mathcal{R}_{16} \cup \mathcal{R}_{17} \cup \mathcal{R}_{20} \cup \mathcal{R}_{21} \cup \mathcal{R}_{23}$, and by

$$x^{m+1}y\varphi(x,y)z\Box yz^{s+1}y\psi(y,z)x\xrightarrow{xy}x^{m+1}y\varphi(x,xy)z\Box yz^{s+1}x$$

to $\mathcal{R}_{10} \cup \mathcal{R}_{18}$.

Proposition 37. $\mathcal{R}_{26+r}(m,r,\varphi,\psi) = \{x^{m+1}y\varphi(x,y)z\Box y^{r+2}z\psi(y,z)x\}$, where φ contains x if m=0, is a singleton realm escaped by

$$x^{m+1}y\varphi(x,y)z\Box y^{r+2}z\psi(y,z)x\xrightarrow{yx}(xy)^{m+1}\varphi(yx,y)z\Box y^{r+1}z\psi(y,z)yx$$

to \mathcal{R}_{25} if r=0, else to \mathcal{R}_{26+r-1} , and by

$$x^{m+1}y\varphi(x,y)z\Box y^{r+2}z\psi(y,z)x\xrightarrow{xy}x^{m+1}y\Box yx\in\mathcal{R}_0\cup\mathcal{R}_1.$$

The following realm forms are indexed from the first infinite ordinal ω on.

Proposition 38. $\mathcal{R}_{\omega+s}(m,s,\varphi,\psi) = \{x^{m+2}y\varphi(x,y)z\Box z^{s+2}y\psi(y,z)x\}$ is a singleton realm escaped by

$$x^{m+2}y\varphi(x,y)z\Box z^{s+2}y\psi(y,z)x \xrightarrow{zx} x(zx)^{m+1}y\Box z^{s+1}y\psi(y,z)zx$$

to $\mathcal{R}_{25}(y \leftrightarrow z)$ if s = 0, else to $\mathcal{R}_{26+s-1}(y \leftrightarrow z)$, and by its $(x \leftrightarrow z)$ -image.

Here we stop introducing new realms and sum up what we have done. Let us call the realms with the indices 3, 4, 5, 6, 7, 8, 9, 11, 15, 16, 20, together with their isomorphic copies, the realms of the *first kind* and all the remaining ones the realms of the *second kind*.

Theorem 4.1. The realms of the first kind are non-trivial strong components (i.e. containing at least two caps) covering all caps with the first term of the form $xy^{m+1}z$, hence by the symmetry $(x \leftrightarrow z)$ also all z-caps with the second term $zy^{m+1}x$, and all caps of the form

$$xyx^{k+1}yz\Box yz^{s+1}x$$
 or $xy(yx)^{k+1}yz\Box z^{s+1}yx$,

where $0 \le s \le 1$, contained in \mathcal{R}_4 for s = 0 and in \mathcal{R}_7 for s = 0.

The realms of the second kind are unions of singleton strong components and they cover all the remaining of caps.

Proof. For $[E_{m,n}] = xy^{m+1}z\Box z^ryz^sx$, where $r + s = n \ge 1$, we have $[E_{0,1}] \in \mathcal{R}_3$, $[E_{0,2}] \in \mathcal{R}_5$, and $[E_{0,n}] \in \mathcal{R}_8$ for $n \ge 3$, $[E_{1,1}] \in \mathcal{R}_4$, $[E_{1,2}] \in \mathcal{R}_7$, and $[E_{1,n}] \in \mathcal{R}_{15}$ for $n \ge 3$, $[E_{m,1}] \in \mathcal{R}_{11}(x \leftrightarrow z)$ for $m \ge 2$, and $[E_{m,n}] \in \mathcal{R}_{20}$ for $m, n \ge 2$.

For $[E]_{m,\sigma} = xy^{m+1}z\Box\sigma(y,z)x$, where $|\sigma|_y \geq 2$, we have $[E]_{0,\sigma} \in \mathcal{R}_4$ if $\sigma \sim (yz)^{k+1}y$ and $[E]_{0,\sigma} \in \mathcal{R}_6 \cup \mathcal{R}_9$ otherwise, $[E]_{0,\sigma} \in \mathcal{R}_{11} \cup \mathcal{R}_{16}$, and $[E]_{m,\sigma} \in \mathcal{R}_{20}$ for $m \geq 2$.

The y-caps with the first term non linear in x are covered by the realms with the indices 10, 18, 24, 25, and 26 + r for $r \ge 0$.

The z-caps with the first term non linear in x and the second one non linear in z are covered by the realms with the indices 12, 13, 14, 17, 19, 21, 22, 23, and $\omega + s$ for $s \geq 0$, and by their $(x \leftrightarrow z)$ -copies.

5. Parametrization of Path Traces

For our pour poses, we can define a parametrization of a set S as a calculable surjective function $\Phi:D\to S$ whose domain is one of the sets defined recursively by stipulating that

- the set \mathbb{N} of naturals is a domain;
- if C, D are domains then $C \times D$ and D^* are domains.

The number of stars intervening in the construction of the domain D of Φ , the star-depth of D, is a sort of complexity measure of the parametrization.

Let S be a subgraph of $[\mathcal{E}_{\mathcal{P}}]$, and let $[E_0]$, [E], $[F] \in \mathcal{S}$. Denote by

• $Tr_{\mathcal{S}}([E], [F])$ the set of the traces of the \mathcal{S} -paths from [E] to [F];

- $Tr_{\mathcal{S}}^{(0)}([E], [F])$ the set of the traces of the \mathcal{S} -paths from [E] to [F] with all intermediate vertices different from $[E_0]$;
- $Tr_{\mathcal{S}}([E], -)$ the set of the traces of the \mathcal{S} -paths from [E] to some cap of \mathcal{S} .

As a rule, the subscript S will be suppressed when clear from the context (e.g. when it denotes a realm currently dealt with).

Lemma 1. We have the relation

$$Tr([E], [F]) = Tr^{(0)}([E], [F]) \cup Tr^{(0)}([E], [E_0])Tr([E_0], [E_0])Tr^{(0)}([E_0], [F]),$$

where $Tr([E_0], [E_0])$ is a monoid generated by $Tr(0)([E_0], [E_0])$ (possibly reduced to $\{\Delta\}$, for the identity substitution Δ).

For a realm \mathcal{R} and a cap $[E] \in \mathcal{R}$, every escape path from \mathcal{R} starting at [E] has a beginning (initial segment) in $Tr_{\mathcal{R}}([E], -)$.

For most of the realms described in the preceding section, the parametrization of $Tr^{(0)}([E], [F])$, for $[E_0]$ suitably chosen and $[E] \neq [F]$, poses no problem and will be left to the reader. We will concentrate on the parametrization of $Tr^{(0)}([E_0], [E_0])$, easily convertible into a parametrization of $Tr([E_0], [E_0])$.

Proposition 39. Every realm \mathcal{R}_{α} with $\alpha \in \{12, 13, 14, 17, 19, 21, 22, 23, 25\} \cup \{26+r \mid r \geq 0\} \cup \{\omega+s \mid s \geq 0\}$ is formed by a single cap $[E_0]$ with $Tr([E_0], [E_0]) = \{\Delta\}$.

Proposition 40. Every realm \mathcal{R}_{α} with $\alpha \in \{1, 10, 18, 24\}$ is either a y-loop or an infinite y-path and we have

$$Tr([E],-)=\{(y\mapsto x^ky)\mid k\geq 0\}$$

for every $[E] \in \mathcal{R}$.

Proposition 41. Every realm \mathcal{R}_{α} with $\alpha \in \{6, 9, 16, 20\}$ is, for some $m \in \mathbb{N}$ and $w \in \{y, z\}^*$, an x-cycle of the form

$$\{[E_i] = xy^{m+1}z \square v_i u_i x \mid w = u_i v_i, |u_i| = i, 0 \le i < |w| - 1\},\$$

parametrized by

$$Tr^{(0)}([E_i], [E_0]) = \{(x \mapsto v_i x)\} \qquad for \ i \ge 1,$$

$$Tr([E_0], [E_0]) = \{(x \mapsto w^k x) \mid k \ge 0\}$$

$$Tr^{(0)}([E_0], [E_j]) = \{(x \mapsto u_j x)\} \qquad for \ j \ge 1,$$

$$Tr^{(0)}([E_i], [E_j]) = \{(x \mapsto tx)\} \qquad for \ u_j = u_i t,$$

$$Tr^{(0)}([E_i], [E_j]) = \emptyset \qquad for \ i > j \ge 1.$$

Consequently,

$$Tr([E_i], [E_j]) = \{(x \mapsto v_i w^k u_j) \mid k \ge 0\} \cup Tr^{(0)}([E_i], [E_j]).$$

Lemma 2. A monoid M generated by a pair of substitutions $(x \mapsto yvx)$ and $(y \mapsto xuy)$, where $x, y \notin alph(uv)$, can be parametrized by the homomorphism

$$\Phi_{xu,yv}: (\mathbb{N} \times \mathbb{N})^* \to M: (m,n) \mapsto (x \mapsto yvx)^m (y \mapsto xuy)^n.$$

Proposition 42. Every realm \mathcal{R}_{α} with $\alpha \in \{8,15\}$ contains, for some $m \in \mathbb{N}$ $m \in \{0,1\}$ and $n \in \mathbb{N}$, the cap $[E_0] = xy^{m+1}z\Box yz^{n+3}x$ an x-cycle of the form for which

$$Tr^{(0)}([E_0], [E_0]) = \{(x \mapsto yz^{n+3}x), (y \mapsto xy)\},\$$

hence

$$Tr([E_0], [E_0]) = \Phi_{x, yz^{n+3}}((\mathbb{N} \times \mathbb{N})^*).$$

Proposition 43. The realm $\mathcal{R}_{11}(r,s)$, $r \leq s$, $r+s \geq 2$, is formed of an x-cycle and of an z-cycle meeting at $[E_0] = xy^{r+1}z\Box zy^{s+1}x$, hence

$$Tr^{(0)}([E_0], [E_0]) = \{(x \mapsto zy^{s+1}x), (z \mapsto xy^{r+1}z)\},$$

 $Tr([E_0], [E_0]) = \Phi_{xy^{r+1}, zy^{s+1}}((\mathbb{N} \times \mathbb{N})^*).$

Proposition 44. The realm \mathcal{R}_5 is divided into two disjoint parts, an x-cycle \mathcal{H} with a y-loop at $[E_0] = xyz\Box yzzx$ and an y-cycle $\mathcal{H}(x \leftrightarrow y)$ with a y-loop at $[F_0] = yxz\Box xzzy$, connected by two z-arrows in such a way that

$$Tr^{(0)}([E_0], [F_0]) = \{g_{x,y} = (x \mapsto yx)(z \mapsto xz)(y \mapsto zy)\},\$$

$$Tr^{(0)}([F_0], [E_0]) = \{g_{y,x} = (x \leftrightarrow y)g_{x,y}(x \leftrightarrow y)\}.$$

Consequently, we have the parametrization homomorphisms

$$\Phi_{u.xzz}: (\mathbb{N} \times \mathbb{N})^* \to Tr_{\mathcal{H}}([E_0], [E_0])$$

and

$$\Psi: ((\mathbb{N} \times \mathbb{N})^* \times (\mathbb{N} \times \mathbb{N})^*))^* \to Tr([E_0], [E_0]): (\boldsymbol{m}, \boldsymbol{n}) \mapsto \Phi_{x,yzz}(\boldsymbol{m}) q_{x,y} \Phi_{y,xzz}(\boldsymbol{n}) q_{y,x}$$

(where m and n are strings of pairs of naturals).

Proposition 45. The realm \mathcal{R}_7 is divided into two disjoint parts \mathcal{H} and $\mathcal{H}(x \leftrightarrow y)$, containing the caps $[E_0] = xyyz \Box yzzx$ and $[F_0] = xzzy \Box zyyx$, connected in such a way that

$$Tr^{(0)}([E_0],[F_0])=\{h_{x,y}=(x\mapsto yx)(z\mapsto xz)(y\mapsto zy)(x\mapsto yx)\},$$

$$Tr^{(0)}([F_0], [E_0]) = \{h_{y,x} = (y \leftrightarrow z)h_{x,y}(y \leftrightarrow z)\}.$$

Consequently, we have the parametrization homomorphisms

$$\Phi_{u,xzz}: (\mathbb{N} \times \mathbb{N})^* \to Tr_{\mathcal{H}}([E_0], [E_0])$$

and

$$\Psi: ((\mathbb{N} \times \mathbb{N})^* \times (\mathbb{N} \times \mathbb{N})^*))^* \to Tr([E_0], [E_0]): (\boldsymbol{m}, \boldsymbol{n}) \mapsto \Phi_{x,yzz}(\boldsymbol{m}) h_{x,y} \Phi_{y,xzz}(\boldsymbol{n}) h_{y,x}$$

(where m and n are strings of pairs of naturals).

Proposition 46. For the cap $[E_0] = xyyz \Box yzzx$ of \mathcal{R}_4 we have

$$Tr^{(0)}([E_0], [E_0]) = \{ f_{k,m} = (y \mapsto x^k y)(x \mapsto yx)(z \mapsto (xy(yx)^k y)^m z)(x \mapsto zx) \},$$

whence the parametrization homomorphism

$$\Omega: (\mathbb{N} \times \mathbb{N})^* \to Tr([E_0], [E_0]) : (k, m) \mapsto f_{k,m}.$$

We come to a conclusion. Let E be an equation with [E] in three cap variables, and let $\mathcal{R}[E]$ denote the realm of caps reachable from [E] in \mathcal{E}_P .

The parametrizations obtained in Proposition 39–46 can be combined, for any $[F] \in \mathcal{R}[E]$, into a (cascade composite) parametrization of Tr([E], [F]). By Proposition 11, the family

$$\mathcal{F} = \bigcup \{ Tr([E], [F])(x \mapsto 1) \mid [F] \in \mathcal{R}[E], \ x \in pivot(F) \}$$

is universal for E. Moreover, every Ef involves less variables than E, hence $\mathcal F$ operates a reduction of E to a parametric family of simpler equations. The fact that $\mathcal F$ is parametric is crucial for the treatment of the family of equations $\{Ef\mid f\in\mathcal F\}$. The latter should be divided into caps and the caps treated by the method we have outlined here.

Of course, the work we have done here for the caps in three variables must be continued, in order to see whether the parametrization for the caps in four and more variables goes as smoothly. The goal is to get a general theorem on the existence of an algorithm calculating a parametrization of a universal family of solutions for an arbitrary equation (with or without constants). Such an algorithm will certainly be rather complex to be of practical use. However, it may help establish decidability results concerning various properties of solutions to a given equation.

References

- [1] J. Jaffar, Minimal and Complete Word Unification. J. ACM 37 (1990) 47-85.
- [2] Yu.I. Hmelevskiĭ, Equations in free semigroups. Trudy Mat. Inst. Steklova 107 (1971); English Translation in Proc. Steklov Inst. Math. 107 (1971) 1976.
- [3] A. Lentin, Équations dans les monoïdes libres. Gautier-Villars, Paris (1972).
- [4] M. Lothaire, Combinatorics on Words. Addison-Wesley (1983).
- [5] G.S. Makanin, The problem of solvability of equations in a free semigroup. Mat. Sbornik 103 (1977) 147-236 (in Russian); English Translation in Math. USSR Sbornik 32 (1977) 128-198.

[6] G.S. Makanin, On general solution of equations in free semigroups, in Proc. of IWWERT'91, edited by H. Abdulrab and J.P. Pécuchet. Springer, Lecture Notes in Comput. Sci. 677, 1-5.
[7] G. Plotkin, Building-in Equational Theories. Machine intelligence 7 (1972) 73-90.

Communicated by Ch. Choffrut. Received May, 2000. Accepted October, 2001.

To access this journal online: www.edpsciences.org