Theoretical Computer Science 72 (1990) 65-94
North-Holland

FACTORIZATION FORESTS OF FINITE HEIGHT*

Imre SIMON

Instituto de Matemática e Estatística, Universidade de São Paulo, 05508 São Paulo, SP, Brasil

Abstract. We introduce factorization forests. It is shown that the vertex set of a dense factorization forest of finite height satisfies a bounded gap Ramsey- roperty of words. The main result is that every morphism from a free semigroup to a finite se p S admits a Ramseyan factorization forest of height at most $9|S|$. Techniques for constru ctorization forests are developed.

1. Introduction

We introduce factorization forests to describe recursive factorizations of words belonging to a free semigroup. Our aim is to point out that these objects can reflect a considerable amount of structure especially if the global height of the forest is bounded.

We show in particular that factorization forests of finite height can be used to derive well known bounded gap Ramsey-type properties of words. Such properties have many applications; actually their discoveries were always motivated by a need in those applications. It appears, however, that factorization forests might be easier to use in the applications than the corresponding Ramsey-type properties. This is because while using factorization forests we gain a powerful tool: induction on the height of vertices. Another advantage of using factorization forests in the applications is that so doing one can bypass not only the long chains of alternating quantifiers but also the enormous upper bounds, both typical of Ramsey Theory. Last, but not least, our factorization forests are constructed through explicit (and efficient) algorithms which could be exploited to obtain a quick and deterministic localization of monochromatic objects.

Once we have made a case for the possible advantages of factorization forests the natural question to answer is: do they exist? Indeed, we shall show that there are many factorization forests of finite height.

The main result of the paper is that every morphism f from a free semigroup A^+ to a finite semigroup S admits a Ramseyan factorization forest of height at most $9|S|$.

Loosely speaking, this means that every word with at least two letters is decomposed in at least two nonempty factors. The decomposition is such that starting

* Part of this work was done while the author was visiting the Institut de Programmation, Université Pierre et Marie Curie, Paris VI, in 1986 and the Université de Rouen, in 1987. Those visits were partially supported by FAPESP.

from any word we arrive at factors of length one in at most $9|S|$ steps. On the other hand, decomposition of a word in three or more factors is allowed only if f maps every one of the factors to the same idempotent element of S.

The proof of the main result relies heavily on the local structure theory of finite semigroups.

We give now a brief outline of the paper. Factorization forests are defined in Section 2 where the basic terminology is introduced. In Section 3 we deal with the connections of factorization forests to Ramsey Theory for words. An open problem is stated there. In Section 4 we study carefully three examples which are used as building blocks to find the factorization forest of the main result. Two of them correspond to the particular cases of the theorem when S is a group or a rectangular band. The third example corresponds to monotonic colorings of words. Section 5 deals with techniques to construct factorization forests. We study, in particular, codifications, substitutions and restrictions of factorization forests. An associative product of factorization forests is also introduced. All these tools are used in the proof of the main result. Section 6 contains the proof of the main result. Finally, in Section 7, we show an application which motivated our investigation of factorization forests.

2. Factorization forests

For a set A we denote the free semigroup generated by A either by A^+, as usual, or by $\mathscr{F}(A)$. In the second notation, $\mathscr{F}(A)$, we neither use the usual convention of representing the unitary sequence (a) by a nor do we use juxtaposition to represent the product of $\mathscr{F}(A)$. Thus, an element s of $\mathscr{F}(A)$ will be represented by

$$s = (a_1, a_2, \ldots, a_n) \quad (a_i \in A).$$

This double notation for the same object is needed to avoid confusion in the case when A itself is a subset of a free semigroup. The free monoid generated by A will be denoted by A^*.

Let $x \in A^+$. We say that an element $(x_1, x_2, \ldots, x_n)$ of $\mathscr{F}(A^+)$ is a *factorization* of x if $x = x_1 x_2 \ldots x_n$.

Let A be an alphabet. A *factorization forest $F = (X, d)$ over A* consists of a subset X of A^+ together with a function $d: X \to \mathscr{F}(X)$ such that $d(x)$ is a factorization of x, for every $x \in X$.

Thus, a factorization forest is just a description of a recursive factorization of words in X in terms of factors belonging to X itself. The factorization is specified by the decomposition function d. Observe that the alphabet A was not required to be finite: indeed, we shall consider many cases in which it will be infinite.

The terminology is suggested by the fact that from a factorization forest F one can define, by induction on $|x|$, a labeled rooted ordered tree $T(x)$, for every $x \in X$. Indeed, it suffices to take a root v_x, label it $\lambda(v_x) = x$, consider the factorization

$d(x)$ of x, say,

$$d(x) = (x_1, x_2, \dots, x_n),$$

and define $T(x)$ as follows. If $n = 1$ then $T(x)$ is the vertex tree consisting of v_x. If $n > 1$ then $T(x)$ consists of the root v_x together with the ordered set of trees

$$(T_x(x_1), T_x(x_2), \dots, T_x(x_n)),$$

where $T_x(x_i)$ is a copy of the tree $T(x_i)$, disjoint from the other subtrees of v_x.

Observe that factorization in the trees $T(x)$ is uniform in the sense that distinct occurrences of a word x are factored the same way: always according to $d(x)$. After introducing some definitions we shall show that this uniformity does not restrict the applicability of the concept because it can always be achieved.

In view of the above representation, our terminology for factorization forests will be based on the theory of graphs. We call X the *vertex set* of F. A vertex x is *internal* if $|d(x)| \geq 2$; otherwise it is *external* (or a leaf). The set of external (internal) vertices of F is called the *external* (*internal*) *set* of F. The *degree* of an external vertex is 0, that of an internal vertex is $|d(x)|$. Observe that every factorization forest with nonempty vertex set has external vertices. Indeed, any shortest word in X is external.

If x is internal and y is an element of $d(x)$ then we say that y is a *direct descendant* of x and also that x is a *direct ancestor* of y. A *path* in F is a sequence

$$c = (x_0, x_1, \dots, x_k)$$

of words in X, such that, for each $1 \leq i \leq k$, x_i is a direct descendant of x_{i-1}. The natural k is the *length* of the path c and we allow the trivial path of length zero. The *origin* of c is x_0, its *terminus* is x_k. If there is a path with origin x and terminus y then we say that x is an *ancestor* of y and also that y is a *descendant* of x. As usual, paths will be concatenated and decomposed; we shall use the current conventions.

The *height* of x is the length of a longest path with origin in x; it is denoted by $h(x)$. Clearly, $h(x) < |x|$, for every $x \in X$. Alternatively, h can be defined as the function $h : X \to \mathbb{N}$, given by

$$h(x) = \begin{cases} 0 & \text{if } x \text{ is external,} \\ 1 + \max\{h(x_i) \mid 1 \leq i \leq n\} & \text{if } x \text{ is internal.} \end{cases}$$

The *height* of F is

$$h = \sup\{h(x) \mid x \in X\}.$$

In case the set $h(X)$ is infinite, we consider $h = \infty$ and say that F is of *infinite height*; otherwise we say that F is of *finite height*.

In the sequel we shall need some properties of factorization forests. We say that F is *dense* if the length of its external vertices is bounded; otherwise F is *sparse*. We say that F is *alphabetical* if every external vertex of F is a letter in A. We say that F is *complete* if it is alphabetical and has vertex set $X = A^+$.

One of our main objectives is to construct factorization forests restricting the possible values of $d(x)$ when x has degree at least three. This is expressed through the next definition. Let $\mathcal{T} \subseteq \mathcal{F}(A^+)$. We say that F is of *type* $\mathcal{T}$ if $d(x) \in \mathcal{T}$ for every vertex x of degree at least three.

Now we show how to achieve the uniformity of factorizations mentioned earlier. Let $Z \subseteq X \subseteq A^+$ and let $\mathcal{T} \subseteq \mathcal{F}(A^+)$. Let **F** be a family of labeled rooted ordered trees such that:

- every vertex v in every tree in **F** has label $\lambda(v)$ in X;
- **F** contains a tree $T(x)$ of height $H(x)$ with root labeled by x, for every $x \in X$;
- for every external vertex v we have $\lambda(v) \in Z$;
- for every internal vertex v, with direct descendants $(v_1, v_2, \ldots, v_n)$, we have $n \geq 2$ and

$$\lambda(v) = \lambda(v_1)\lambda(v_2)\ldots\lambda(v_n);$$

further, if $n > 2$ then $(\lambda(v_1), \lambda(v_2), \ldots, \lambda(v_n))$ is in $\mathcal{T}$.

Proposition 2.1. *Given A, X, Z, $\mathcal{T}$ and* **F** *as above, there exists a factorization forest $F = (X, d)$ of type $\mathcal{T}$, with external set Z, such that $h(x) \leq H(x)$, for every $x \in X$. In particular, if $\{H(x) \mid x \in X\}$ is bounded then F is of finite height.*

Proof. We initially define the function d together with an auxiliary function $g: X \to \mathbb{N}$. For x in Z we put $g(x) = 0$ and $d(x) = (x)$. For x in $X - Z$ we consider the set $V(x)$ of all vertices in the family **F** with label x. Every vertex in $V(x)$ is internal in its tree; further the set $V(x)$ is not empty since it contains the root of $T(x)$. Then we choose a vertex v in $V(x)$ with minimum height, say h_x, and let $(v_1, v_2, \ldots, v_n)$ be the sequence of direct descendants of v. We put $g(x) = h_x$ and $d(x) = (\lambda(v_1), \lambda(v_2), \ldots, \lambda(v_n))$.

Now, by induction on $|x|$, one can show that $h(x) \leq g(x) \leq H(x)$, for every $x \in X$. Clearly, F is a factorization forest which satisfies the required properties. □

3. Ramseyan factorization forests

We show in this section that the vertex set of every dense factorization forest of finite height satisfies a bounded gap Ramsey-type property. This property has been studied in [2, 9, 3, 8, 13, 11]. Our results, established in this and the next three sections, imply all but one of the related results obtained in these references.

Let $v = (x_1, x_2, \ldots, x_n)$ be in $\mathcal{F}(A^+)$. The *gap size* of v is

$$\text{gap}(v) = \max\{|x_i| \mid 1 \leq i \leq n\}.$$

Let $F = (X, d)$ be a factorization forest over A. For $x \in X$ the *gap size* of x is $\text{gap}(d(x))$.

Proposition 3.1. *Let $F=(X, d)$ be a dense factorization forest of finite height over A and let S be a subset of X which contains arbitrarily long words. Then there exists a $g \in \mathbb{N}$ for which arbitrarily long descendants of words in S have gap size at most g.*

Proof. Let S' be the set of all descendants of words in S. Then, $S \subseteq S'$; hence S' contains arbitrarily long words. Let h be the height of F and let L_f, for $0 \leqslant f \leqslant h$, be the set of words in S' of height at most f. Then $L_0 \subseteq L_1 \subseteq \cdots \subseteq L_h = S'$. Now, F being dense implies that the words in L_0 have bounded length; on the other hand, L_h contains arbitrarily long words. Let then f' be the least integer such that $L_{f'}$ contains arbitrarily long words. Let

$$g = \max\{|x| \mid x \in L_{f'-1}\}.$$

The choice of f' implies that $g < \infty$. Now, any direct descendant of any word in $L_{f'}$ belongs to $L_{f'-1}$. Hence, every word in $L_{f'-1}$ has gap size at most g. The proof is complete since $L_{f'}$ contains arbitrarily long words. □

Now we introduce some terminology inspired by [11]. Let f be a function, $f: A^+ \to E$ or $f: A^* \to E$, and let $v \in \mathcal{F}(A^+)$, say $v = (x_1, x_2, \ldots, x_n)$. We say that v is an (nth) *power modulo f* if for some $e \in E$

$$f(x_1) = f(x_2) = \cdots = f(x_n) = e.$$

If, in addition, for every $1 \leqslant i < j \leqslant n$,

$$f(x_i x_{i+1} \ldots x_j) = e,$$

then we say that the (nth) power v is *Ramseyan*. Function f is *strongly Ramseyan* if for every infinite word $s \in A^\omega$ there exists $g > 0$ such that, for each $k > 0$, some factor of s has a factorization in $\mathcal{F}(A^+)$ which is a Ramseyan kth power modulo f of gap size at most g.

In the next sections we shall construct factorization forests whose internal vertices of degree at least three satisfy a restriction, captured by the next definition. Such a restriction plays the role of monochromatic properties of objects in Ramsey theory.

A factorization forest $F=(X, d)$ is *Ramseyan modulo f* if $d(x)$ is a Ramseyan power modulo f for every vertex x of degree at least three. We say that a function *f admits a Ramseyan factorization forest* if there exists a complete factorization forest which is Ramseyan modulo f.

Proposition 3.2. *Let $f: A^+ \to E$ be a function. If f admits a Ramseyan factorization forest of finite height then f is strongly Ramseyan.*

Proof. Let s be an infinite word in A^ω and let S be the set of finite factors of s. Applying Proposition 3.1 results in a value g which shows that f is strongly Ramseyan. □

In [13] we have exhibited a function $\phi:\{a\}^+\to\{0,1\}$ which is not strongly Ramseyan. Hence, there exist functions, even with finite codomain, which do not admit Ramseyan factorization forests of finite height. Other such examples can be found in [10].

An interesting open question is whether the converse of Proposition 3.2 is true or not. This seems especially interesting for finite E.[1] While we do not know the answer in general we can use our main result to show a partial converse of Proposition 3.2. It is worth pointing out that we do not have a direct proof of this restricted converse.

Assume, for the rest of this section, that the main result holds.

Theorem 3.3. *Let $f: A^+\to S$ be an epimorphism of semigroups such that $f(A)$ is finite. Then the following conditions are equivalent:*

(i) *S is finite,*

(ii) *f admits a Ramseyan factorization forest of finite height,*

(iii) *f is strongly Ramseyan.*

Proof. Theorem 6.1 and Proposition 3.2 guarantee, respectively, that (i) implies (ii) and (ii) implies (iii).

Now we show that (iii) implies (i). Assume that f is strongly Ramseyan. Let B be a minimal subset of A such that $f(B)=f(A)$. The finiteness of $f(A)$ implies that B is finite. On the other hand, $f(B^+)=f(B)^+=f(A)^+=f(A^+)=S$; i.e. the restriction g of f to B^+ is still an epimorphism. Clearly, g is strongly Ramseyan.

We say that a word s in B^+ is *irreducible* if it is a shortest word in $g^{-1}g(s)$. Let W be the set of irreducible words in B^+; then every factor of a word in W is again in W. Assume that the set W is infinite. Then, by König's Lemma, there exists an infinite word s in B^ω such that every finite factor of s is in W. Now, g being strongly Ramseyan, it follows that there exist $x_1, x_2\in B^+$ and an idempotent $e\in S$ such that x_1x_2 is a factor of s and $g(x_1)=g(x_2)=e$. Since x_1x_2 is a factor of s, it is irreducible. However, $g(x_1x_2)=g(x_1)g(x_2)=e^2=e=g(x_1)$. Since $|x_1|<|x_1x_2|$ we conclude that x_1x_2 is not irreducible, a contradiction. Thus, W is finite and, consequently, S itself is finite. The proof of Theorem 3.3 is complete. □

4. Examples of factorization forests

In this section we show that three families of functions admit Ramseyan factorization forests of finite height. Actually, we even exhibit some of the structure, captured by the next definitions, of the forests we construct.

Let $F=(X,d)$ be a factorization forest over A and let S be a subset of X. We say that F is *simple* on S if every path in F contains at most one word in S.

[1] *Note added in proof*: J. Justin constructed a strongly Ramseyan function $f: A^+\to E$, with E infinite, which does not admit a Ramseyan factorization forest of finite height. For finite E, however, the problem is still open.

A family $\{(X_i, C_i) \mid 1 \leq i \leq n\}$ of pairs of subsets of A^+ is a *stratification* for the factorization forest $F = (X, d)$ over A if
- the X_i's are pairwise disjoint subsets of X,
- each C_i is a code over A, such that $C_i \subseteq X_i \subseteq C_i^+$,
- all direct descendants of words in $X_i - C_i$ lie in C_i; i.e. $d(X_i - C_i) \subseteq \mathscr{F}(C_i)$,
- every vertex in X, with degree at least three, belongs to some $X_i - C_i$.

The stratification is *simple* if F is simple on each $X_i - C_i$.

We point out that the above notation is redundant. Indeed, a subset X of A^+ can contain at most one code C such that $X \subseteq C^+$. It contains such a code iff X^* is a free submonoid of A^* and in this case the code C is the unique minimal generator of X^*. The reason for this redundancy is that specifying the codes in the notation will allow simplification of some arguments. Our terminology for codes follows the treatise in [1].

The following proposition will be used to prove simplicity of a stratification.

Proposition 4.1. *Let $\{(X_i, C_i) \mid 1 \leq i \leq n\}$ be a stratification for the factorization forest $F = (X, d)$ over A. If each C_i is disjoint from $A^* C_i^2 A^*$ then the given stratification is simple.*

Proof. Assume, for a contradiction, that, for some $1 \leq i \leq n$, there exist x and y in $X_i - C_i$ such that $x \neq y$ and y is a descendant of x. Let $(t_1, t_2, \ldots, t_m)$ be a path from $x = t_1$ to $y = t_m$. Since $d(t_1) \in \mathscr{F}(C_i)$ it follows that $t_2 \in C_i$. On the other hand, $d(t_m) \in \mathscr{F}(C_i)$ also, say, $d(t_m) = (s_1, s_2, \ldots s_k)$, for some $k \geq 2$ and $s_j \in C_i$. It follows that there exist $u, v \in A^*$ such that $t_2 = u s_1 s_2 v$ and $t_2, s_1, s_2 \in C_i$. But this contradicts the hypothesis that $C_i \cap A^* C_i^2 A^* = \emptyset$. $\square$

4.1. The case of groups

In this subsection we show the following theorem.

Theorem 4.2. *Every (semigroup) morphism $f: A^+ \to G$, with G a finite group, admits a Ramseyan factorization forest of height at most $3|G|$.*

Actually, we shall prove a more precise statement, from which Theorem 4.2 follows immediately.

Proposition 4.3. *Let $f: A^+ \to G$ be a semigroup morphism where G is a finite group with identity e. Define*

$$X_e = f^{-1}(e) \quad \textit{and} \quad C_e = X_e - X_e^2.$$

Then C_e is a biprefix code, $X_e = C_e^+$, and there exists a complete factorization forest over A of height at most $3|G|$, for which $\{(X_e, C_e)\}$ is a stratification.

Proof. We denote the identity e of G by 1 and extend f to A^* by putting $f(1)=1$. We begin showing that C_e is a biprefix code. Let $u, v \in C_e$ and assume that $u = vw$ for some $w \in A^+$. Then, $e = f(u) = f(v)f(w) = f(w)$; hence $w \in X_e$ and $u \in X_e^2$, a contradiction. Thus, C_e is a prefix and by a dual argument it is also a suffix. Altogether, C_e is a biprefix, as claimed. Since e is idempotent it follows that $X_e = C_e^+$.

We shall construct the factorization forest $F = (X, d)$, with vertex set $X = A^+$ and internal set $Y = AA^+$. First we define three functions:

$$\mathrm{int}: X \to 2^G, \qquad \mathrm{ref}: Y \to G, \qquad \mathrm{type}: Y \to \{1, 2, 3\},$$

as follows. We put

$$\mathrm{int}(x) = \{f(u) \in G \mid x = uv \text{ with } u, v \in A^+\};$$

i.e. int (x) is the set of elements of G "visited internally" by x. Let $x \in aA^*$, with $a \in A$; we put

$$\mathrm{ref}(x) = \begin{cases} 1 & \text{if } 1 \in \mathrm{int}(x), \\ f(a) & \text{if } 1 \notin \mathrm{int}(x); \end{cases}$$

ref(x) is an element in int(x) which will be used as a "reference point" for x. Note that ref$(x) = 1$ iff x visits 1 internally. Finally, let

$$\mathrm{type}(x) = \begin{cases} 1 & \text{if } \mathrm{ref}(x) = 1 = f(x), \\ 2 & \text{if } \mathrm{ref}(x) = 1 \neq f(x), \\ 3 & \text{if } \mathrm{ref}(x) \neq 1. \end{cases}$$

Note that type partitions Y in three blocks and type$(x) = 1$ iff $x \in X_e - C_e$. The proof uses the following assertion.

Assertion 4.4. *Let $x \in Y$, $v \in A^+$ and $u, w \in A^*$ be such that $x = uvw$. Then* $|\mathrm{int}(v)| \leq |\mathrm{int}(x)|$. *Further, if* $\mathrm{ref}(x) \notin f(u)\,\mathrm{int}(v)$ *then* $|\mathrm{int}(v)| < |\mathrm{int}(x)|$.

Proof. It is easy to see that $f(u)\,\mathrm{int}(v) \subseteq \mathrm{int}(x)$. Since G is a group, it follows that $|\mathrm{int}(v)| = |f(u)\,\mathrm{int}(v)| \leq |\mathrm{int}(x)|$.

Since $|x| \geq 2$, we have $\mathrm{ref}(x) \in \mathrm{int}(x)$. Thus, if $\mathrm{ref}(x) \notin f(u)\,\mathrm{int}(v)$ then $f(u)\,\mathrm{int}(v) \subset \mathrm{int}(x)$; and, consequently, $|\mathrm{int}(v)| < |\mathrm{int}(x)|$. □

Now we define, by induction on $|x|$, the decomposition $d(x)$ and show simultaneously that if $|x| \geq 2$ and y is in $d(x)$ then

$$|\mathrm{int}(y)| < |\mathrm{int}(x)| \quad \text{or} \quad \mathrm{type}(y) < \mathrm{type}(x). \tag{1}$$

If $|x| = 1$ then we put $d(x) = (x)$ and there is nothing to prove. Assume then $|x| > 1$; we shall consider three cases according to the value of type(x).

Case 1 (type$(x) = 1$): In this case $x \in X_e - C_e$. Thus, there exist $n \geq 2$ and $x_1, x_2, \ldots, x_n \in C_e$, such that $x = x_1x_2 \ldots x_n$. Since $x_i \in C_e$, it follows that, for every $1 \leq i \leq n$, $f(x_i) = 1$ and $1 \notin \mathrm{int}(x_i)$. Hence, $\mathrm{ref}(x) = 1 \notin \mathrm{int}(x_i) = f(x_1x_2 \ldots x_{i-1})\,\mathrm{int}(x_i)$, and, by Assertion 4.4, $|\mathrm{int}(x_i)| < |\mathrm{int}(x)|$, for every i. Thus, it suffices to put $d(x) = (x_1, x_2, \ldots, x_n)$.

Case 2 (type$(x)=2$): Let x_1 be the longest left factor of x for which $f(x_1)=1$ and let $x=x_1x_2$. Then, since ref$(x)=1$ we have $x_1\in A^+$ and since $f(x)\neq 1$ we have $x_2\in A^+$. Further, by the choice of x_1, $1\notin \text{int}(x_2)$. Then, ref$(x)=1\notin \text{int}(x_2)=f(x_1)\,\text{int}(x_2)$, and, by Assertion 4.4, $|\text{int}(x_2)|<|\text{int}(x)|$. Assume now that $1\notin \text{int}(x_1)$; then, ref$(x)=1\notin \text{int}(x_1)=f(1)\text{int}(x_1)$, and, by Assertion 4.4, $|\text{int}(x_1)|<|\text{int}(x)|$. Otherwise, $1\in \text{int}(x_1)$ and $f(x_1)=1$ imply that type$(x_1)=1<2=$ type(x). In any case, both x_1 and x_2 satisfy (1). It suffices to put $d(x)=(x_1, x_2)$.

Case 3 (type$(x)=3$): Let $x=x_1x_2$, with $|x_1|=1$. Then ref$(x)=f(x_1)\neq 1$. Since $x_1\in A$, we have that $|\text{int}(x_1)|=0<|\text{int}(x)|$. On the other hand, if $1\in \text{int}(x_2)$ then type$(x_2)\leqslant 2<3=$ type(x). Otherwise, ref$(x)=f(x_1)\notin f(x_1)\,\text{int}(x_2)$; hence, by Assertion 4.4, $|\text{int}(x_2)|<|\text{int}(x)|$. In any case both x_1 and x_2 satisfy (1). It suffices to put $d(x)=(x_1, x_2)$.

This concludes the definition of d and the verification of (1).

Now, let y be a direct descendant of x. Assertion 4.4 and (1) imply that either $|\text{int}(y)|<|\text{int}(x)|$ or $|\text{int}(y)|=|\text{int}(x)|$ and type$(y)<$ type(x). It follows that $h(x)\leqslant 3|\text{int}(x)|+\text{type}(x)-3$, for every $x\in Y$. Thus, the height of F is at most $3|G|$.

Finally, $\{(X_e, C_e)\}$ is a stratification for F. Indeed, $x\in X_e-C_e$ iff type$(x)=1$; and in this case $d(x)\in \mathscr{F}(C_e)$. Also, if the degree of x is at least three then type$(x)=1$; hence, again $x\in X_e-C_e$. This concludes the proof of Proposition 4.3. □

4.2. *The case of monotonic colorings*

A *coloring* of a set S is a function $f: S\to E$. Function f is a *finite coloring* if E is a finite set. A coloring f of A^+ is *monotonic* if for every $w\in A^+$ and $u_1, v_1, v_2, u_2\in A^*$

$$f(w)=f(u_1v_1wv_2u_2) \text{ implies } f(w)=f(v_1wv_2).$$

In this subsection we show the following theorem.

Theorem 4.5. *Every finite monotonic coloring* $f: A^+\to E$ *of* A^+ *admits a Ramseyan factorization forest of height at most* $3|E|$.

Before the proof we derive a particular case. A *rank* over A is a function $f: A^*\to \mathbb{N}$ which satisfies

$$f(uv)\leqslant \min\{f(u), f(v)\}$$

for every $u, v\in A^*$.

Corollary 4.6. *Every rank function* $f: A^*\to\mathbb{N}$ *admits a Ramseyan factorization forest of height at most* $3(1+f(1))$.

Proof. Note that, by definition, for every $u\in A^*$, we have that $u=u1$; hence, $f(u)\leqslant f(1)$. It follows that the image of f has at most $1+f(1)$ elements. On the other hand, the restriction of f to A^+ is a monotonic coloring of A^+. Indeed, assume that $w \in A^+$ and $u_1, v_1, v_2, u_2 \in A^*$ are such that $f(w) = f(u_1 v_1 w v_2 u_2)$. From the definition of a rank, we have that

$$f(w) = f(u_1 v_1 w v_2 u_2) \leq f(v_1 w v_2 u_2) \leq f(v_1 w v_2) \leq f(w v_2) \leq f(w).$$

Thus, $f(v_1 w v_2) = f(w)$, as required. The corollary now follows from Theorem 4.5. $\square$

We point out that Corollary 4.6 and Proposition 3.2 imply a result of Jacob [8] (see also [11]) stating that every rank function is strongly Ramseyan. Also, the related result in [13] that every monotonic coloring of A^+ is strongly Ramseyan follows from Theorem 4.5 and Proposition 3.2.

We begin the proof by an assertion.

Assertion 4.7. *Let $f: A^+ \to E$ be a monotonic coloring of A^+. Let $e \in E$ and let $C_e = \{t \in f^{-1}(e) \mid t \in A$ or $f(tA^{-1}) \neq e\}$. Then C_e is a prefix code which is disjoint from $A^* C_e^2 A^*$.*

Proof. Let $u, v \in C_e$ and assume that $u = vw$ for some $w \in A^+$. Then, $e = f(u) = f(v)$; and, since f is monotonic and $|v| < |u|$, it follows that $f(uA^{-1}) = e$. This contradicts that $u \in C_e$; hence C_e is a prefix as claimed.

Assume now that $C_e \cap A^* C_e^2 A^* \neq \emptyset$. Then there are $t, s_1, s_2 \in C_e$ and $u, v \in A^*$, such that $t = u s_1 s_2 v$. Since $|s_2| > 0$ and $f(s_1) = f(t)$ it follows from the monotonicity of f that $f(tA^{-1}) = f(t)$. This contradicts that $t \in C_e$; hence, $C_e \cap A^* C_e^2 A^* = \emptyset$. $\square$

Now we prove a result from which Theorem 4.5 follows immediately.

Proposition 4.8. *Let $f: A^+ \to E$ be a finite monotonic coloring of A^+. For $e \in E$ define*

$$C_e = \{t \in f^{-1}(e) \mid t \in A \text{ or } f(tA^{-1}) \neq e\},$$

$$X_e = C_e^+ \cap f^{-1}(e).$$

Then each C_e is a prefix code and there exists a complete factorization forest over A of height at most $3|E|$, for which $\{(X_e, C_e) \mid e \in E\}$ is a simple stratification.

Proof. From Assertion 4.7 we know that each C_e is a prefix code. Observe that, in general, C_e^+ and $f^{-1}(e)$ are incomparable; however, the X_e's are pairwise disjoint since $X_e \subseteq f^{-1}(e)$. Also, for every $x \in A^+$, we have $x \in f^{-1} f(x)$; hence

$$x \in \bigcup X_e \text{ iff } x \in X_{f(x)} \text{ iff } x \in C_{f(x)}^+.$$

We shall construct the factorization forest $F = (X, d)$, with vertex set $X = A^+$ and internal set $Y = AA^+$. First we define two functions

$$\text{int}: X \to 2^E \quad \text{and} \quad \text{type}: Y \to \{1, 2, 3\}$$

as follows. We put

$$\text{int}(x) = \{f(u) \in E \mid u \in A^+ \text{ is a factor of } x\},$$

i.e. int(x) is the set of colors of the nonempty factors of x. We also put

$$\text{type}(x) = \begin{cases} 1 & \text{if } x \in C_{f(x)}, \\ 2 & \text{if } x \in C^+_{f(x)} - C_{f(x)}, \\ 3 & \text{if } x \notin C^+_{f(x)}. \end{cases}$$

Note, for future use, that

$$x \in \bigcup (X_e - C_e) \text{ iff } \text{type}(x) = 2.$$

The proof uses the following assertions.

Assertion 4.9. *Let $u, v \in A^+$ be such that u is a factor of v. Then* $\text{int}(u) \subseteq \text{int}(v)$; *hence,* $|\text{int}(u)| \leq |\text{int}(v)|$. *Further, if $f(u) \neq f(v)$ then* $|\text{int}(u)| < |\text{int}(v)|$.

Proof. The first claim clearly holds. Let $s_1, s_2 \in A^*$ be such that $v = s_1 u s_2$ and assume $|\text{int}(u)| = |\text{int}(v)|$. Since $\text{int}(u) \subseteq \text{int}(v)$ and the coloring is a finite one, it follows that $\text{int}(u) = \text{int}(v)$. In particular, $f(v) \in \text{int}(u)$. Let $w \in A^+$ and $t_1, t_2 \in A^*$ be such that $u = t_1 w t_2$ and $f(w) = f(v)$. Since f is monotonic, $v = s_1 t_1 w t_2 s_2$ implies that $f(v) = f(t_1 w t_2) = f(u)$, a contradiction. $\square$

Assertion 4.10. *Every $x \in A^+$ has a nonempty left factor in $C_{f(x)}$.*

Proof. It suffices to take a shortest nonempty left factor t of x which satisfies $f(t) = f(x)$. Then either $t \in A$ or, by the choice of t, we have that $f(tA^{-1}) \neq f(x)$. In either case $t \in C_{f(x)}$. $\square$

Now we define, by induction on $|x|$, the decomposition $d(x)$ and show simultaneously that if $|x| \geq 2$ and y is in $d(x)$ then

$$|\text{int}(y)| < |\text{int}(x)| \quad \text{or} \quad \text{type}(y) < \text{type}(x). \tag{2}$$

If $|x| = 1$, then we put $d(x) = (x)$ and there is nothing to prove. Assume then $|x| > 1$; we shall consider three cases according to the value of type(x).

Case 1 (type$(x) = 1$): Let $x = x_1 x_2$, with $x_2 \in A$. Since $|x| > 1$ and $x \in C_{f(x)}$ we have $f(x_1) = f(xA^{-1}) \neq f(x)$. Then, by Assertion 4.9, $|\text{int}(x_1)| < |\text{int}(x)|$. As for x_2, we have that $|\text{int}(x_2)| = 1 < 2 = |\{f(x), f(x_1)\}| \leq |\text{int}(x)|$. It suffices to put $d(x) = (x_1, x_2)$.

Case 2 (type$(x) = 2$): There exist $n \geq 2$ and $x_1, x_2, \ldots, x_n \in C_{f(x)}$ such that $x = x_1 x_2 \ldots x_n$. Then, each $x_i \in C_{f(x)}$; hence $f(x_i) = f(x)$. That is, $x_i \in C_{f(x_i)}$ and, consequently, type$(x_i) = 1 < 2 =$ type(x). Thus, (2) is satisfied and it suffices to put $d(x) = (x_1, x_2, \ldots, x_n)$.

Case 3 (type$(x) = 3$): There exists, by Assertion 4.10, a longest left factor x_1 of x, such that $x_1 \in C^+_{f(x)}$. Let x_2 be such that $x = x_1 x_2$. Note that type$(x) = 3$ implies that $x_2 \in A^+$. Now, since x_1 has a left factor of color $f(x)$ and f is a monotonic coloring, it follows that $f(x_1) = f(x)$. Thus, $x_1 \in C^+_{f(x_1)}$ and, consequently, type$(x_1) \leq 2 < 3 =$ type(x). Hence, x_1 satisfies (2). On the other hand, the choice of x_1 implies, via Assertion 4.10, that $f(x_2) \neq f(x)$. Thus, by Assertion 4.9, $|\mathrm{int}(x_2)| < |\mathrm{int}(x)|$, and x_2 satisfies (2). It suffices to put $d(x) = (x_1, x_2)$.

This concludes the definition of d and the verification of (2).

Now, let y be a direct descendant of x. Assertion 4.9 and (2) imply that either $|\mathrm{int}(y)| < |\mathrm{int}(x)|$ or $|\mathrm{int}(y)| = |\mathrm{int}(x)|$ and $\mathrm{type}(y) < \mathrm{type}(x)$. It follows that $h(x) < 3|\mathrm{int}(x)| + \mathrm{type}(x) - 3$, for every $x \in Y$. Thus, the height of F is at most $3|E|$.

Note that the (X_e, C_e)'s do form a stratification for F. This is so because, as we noted earlier, $x \in \bigcup (X_e - C_e)$ iff $\mathrm{type}(x) = 2$ and in this case, by construction, $d(x) \in \mathscr{F}(C_{f(x)})$. Also, if $\mathrm{type}(x)$ is 1 or 3 then the degree of x is two; hence, if x has degree at least three then $\mathrm{type}(x) = 2$ and x belongs to $X_{f(x)} - C_{f(x)}$. The stratification is simple in view of Proposition 4.1 and Assertion 4.7. This completes the proof of Proposition 4.8. □

4.3. *The case of rectangular bands*

In this subsection we show the following theorem.

Theorem 4.11. *Every morphism $f: A^+ \to S$, with S a finite rectangular band, admits a Ramseyan factorization forest of height at most* $3|S|$.

We begin by recalling the definition of a rectangular band. Let I and Λ be nonempty sets. The set $I \times \Lambda$ becomes a semigroup by defining the associative multiplication

$$(i, \lambda)(j, \mu) = (i, \mu) \quad (i, j \in I;\ \lambda, \mu \in \Lambda).$$

A semigroup is called a *rectangular band* if it is isomorphic to $I \times \Lambda$ for appropriate sets I and Λ.

The proof of Theorem 4.11 needs some preparation. We assume, from now on, that S is a finite rectangular band and that $f: A^+ \to S$ is a morphism of semigroups. Initially we list a few properties of rectangular bands.

Assertion 4.12. *Let u, $v \in A^+$ and a, b, c, $d \in A^*$. Then,*

(i) $f(u) = f(uau)$,

(ii) $f(uv) = f(uav)$,

(iii) $f(ua) = f(vb)$ *and* $f(cu) = f(dv)$ *imply* $f(u) = f(v)$.

Proof. Let $f(u) = (i, \lambda)$; the definition of the product of S implies that $f(uau) = (i, \lambda)$, whatever the value of $a \in A^*$. Now, (ii) follows, since $f(uv) = f(uavu \cdot v) = f(ua \cdot vuv) = f(uav)$. Finally, $f(u) = f(uacu) = f(vbdv) = f(v)$, which establishes (iii). □

Now we define, for each $e \in S$,

$$X_e = f^{-1}(e) \quad \text{and} \quad C_e = X_e - X_e^2.$$

atomic actions:
RESET: $\rightarrow$ *process*
PUSH : *item* $\rightarrow$ *process*
POP : $\rightarrow$ *process*
equations for atomic actions:
RESET::s = empty
PUSH(i)::s = push(i, s)
POP::empty = $\emptyset$
POP::push(i, s) = s
where i: *item*, s: *stack*
end process specification PSTACK_0

The process POP pops the current environment when it is not empty, and otherwise returns a *no* result. This is a possible way to take into account exceptions (cf. [8, 6] for different approaches to this question, in the classical algebraic framework). In fact, PSTACK_0 may be obtained by a straightforward translation of the classical algebraic specification STACK_0 of stacks into a process specification. The translation is analyzed in full detail in [37]. In PSTACK_0, it is possible to consider terms such as:

$$(\text{PUSH}(i_1) \parallel \text{PUSH}(i_2) \parallel \text{POP}) :: \text{push}(\ldots)$$

where $\parallel$ is the parallel composition operator defined hereafter. It is thus possible to describe concurrent computations at this stage—as fully discussed in what follows.

2.2. *Semantics of process specifications*

The semantics of a process specification SPEC is provided by translating SPEC into a classical algebraic specification SEM(SPEC). Then, by definition, the semantics of SPEC is the classical semantics of SEM(SPEC), i.e. the class of its hierarchical models (cf. Section 1).

Actually, we could have presented the whole approach by considering the **process spec**... construct as a *macro* returning a classical specification—thus staying within the usual abstract data type framework. However, such a simplification hides the intuition behind processes and concurrent specifications. Also, in this fashion, making use of a classical notion of implementation or of classical composition primitives, the questions arising in the description of concurrent systems would be irrelevant.

We now give intuitive characteristics of the transformation from SPEC to SEM(SPEC). The signature of SEM(SPEC) has the two following parts:

(1) A "constant" part, common to all process specifications. It consists of operators allowing to define processes:

- δ: a constant standing for *deadlock*,
- $+$, $;$, $\parallel$: operators for *nondeterministic*, *sequential* and *parallel* composition,
- the part describing the application of processes to data (operator ::),

Assertion 4.16. *Let $x \in A^2A^*$ be such that x has neither composite factor nor cutpoint. Then there exist $u, v \in A^+$ such that $x = uv$, u has a cutpoint and $|\mathrm{int}(v)| < |\mathrm{int}(x)|$.*

Proof. Let u be the shortest left factor of x such that $\mathrm{int}(u) = \mathrm{int}(x)$ and let v be such that $x = uv$. Now, $|x| > 1$; hence $\mathrm{int}(x) \neq \emptyset$. Since $\mathrm{int}(a) = \emptyset$ for every $a \in A$, it follows that $|u| > 1$. Let $a \in A$ and $u' \in A^+$ be such that $u = u'a$. From the choice of u we conclude that $\mathrm{int}(u') \subset \mathrm{int}(x) = \mathrm{int}(u'a) \supset \emptyset = \mathrm{int}(a)$, i.e. u has a cutpoint. Next, if $v = 1$ then $x = u$ and, by what we have just seen, x has a cutpoint, a contradiction. Thus, $v \in A^+$.

Now we claim that $\mathrm{int}(v) \subset \mathrm{int}(x)$. Assume that this is not the case, then, by Assertion 4.13, $\mathrm{int}(v) = \mathrm{int}(x)$. Then $x = uv$, with $u, v \in A^+$ and

$$\mathrm{int}(u) = \mathrm{int}(uv) = \mathrm{int}(v).$$

We will show that this implies that x has a composite factor. Indeed, let $e = f(vu)$. Then, $e \in \mathrm{int}(x)$; hence, $e \in \mathrm{int}(u) \cap \mathrm{int}(v)$. There exist then $u_1, u_2, v_1, v_2 \in A^+$ such that $u = u_1u_2$, $v = v_1v_2$ and $f(u_2u_1) = e = f(v_2v_1)$. Now, $f(v_1v_2u_1u_2) = f(u_2u_1)$ and $f(v_2v_1) = f(v_1v_2u_1u_2)$ imply, by Assertion 4.12(iii), that $f(v_1) = f(u_2)$. Thus, the factor u_2v_1 of x is composite. This contradicts the hypothesis and this contradiction establishes the claim. Then, $\mathrm{int}(v) \subset \mathrm{int}(x)$, and this completes the proof of Assertion 4.16. □

Now we show that the subsemigroup X_e is free.

Assertion 4.17. *The set C_e is a code of deciphering delay one in either direction which is disjoint from $A^*C_e^2A^*$.*

Proof. First we show that C_e is a code of deciphering delay one (see [1, p. 128]). Let $x, x', y \in C_e$, let $u \in A^*$ and $z \in C_e^*$ be such that $xyu = x'z$. We have to prove that $x = x'$. If $|x| = |x'|$ then we have nothing to do. Assume that $|x'| < |x|$ and let $v \in A^+$ be such that $x = x'v$. Then, $z = vyu$. Consequently $z \neq 1$ and we have $f(x') = f(z) = f(vyu)$ and $f(x') = f(x) = f(x'v)$. Hence, by Assertion 4.12(iii), $f(x') = f(v)$ and $x = x'v$ is composite, a contradiction. Assume, finally, that $|x| < |x'|$ and let $w \in A^+$ be such that $x' = xw$. This implies that $yu = wz$. Since $f(x) = f(y)$, we have $f(xu) = f(yu) = f(wz)$ and $f(x) = f(x') = f(xw)$. Hence, by Assertion 4.12(iii), $f(x) = f(w)$ and $x' = xw$ is composite, a contradiction. Thus, C_e has deciphering delay one. A dual argument shows that it also has delay one in the other direction.

Assume now that $C_e \cap A^*C_e^2A^* \neq \emptyset$. Then there exist $t, s_1, s_2 \in C_e$ and $u, v \in A^*$ such that $t = us_1s_2v$. We shall argue that this is impossible. Indeed, from $e = e^2 = f(t) = f(s_1)$ we have, using Assertion 4.12(i), $f(t) = f(ts_1) = f(us_1s_2vs_1) = f(us_1)$. By a dual argument, $f(t) = f(s_2v)$. Thus, $x = us_1s_2v$ is composite, a contradiction. This completes the proof of Assertion 4.17. □

We are ready to prove Theorem 4.11. Actually, we shall prove the following more precise statement which promptly implies that theorem.

Proposition 4.18. *Let $f: A^+ \to S$ be a morphism of semigroups with S a finite rectangular band. For $e \in S$ define*

$$X_e = f^{-1}(e) \quad \text{and} \quad C_e = X_e - X_e^2.$$

Then each C_e is a code of deciphering delay one in either direction, $X_e = C_e^+$, and there exists a complete factorization forest over A of height at most $3|S|$, for which $\{(X_e, C_e) \mid e \in S\}$ is a simple stratification.

Proof. We already know that $X_e = C_e^+$ and that C_e is a code as required.

We shall construct the factorization forest $F = (X, d)$, with vertex set $X = A^+$ and internal set $Y = AA^+$.

Initially we define two partitions of Y,

$$\text{nat}: Y \to \{a, b, c, d, e\} \quad \text{and} \quad \text{type}: Y \to \{1, 2, 3\},$$

given by

$$\text{nat}(x) = \begin{cases} a & \text{if } x \text{ is composite,} \\ b & \text{if } x \text{ is not composite but has a composite left factor,} \\ c & \text{if } x \text{ has no composite left factor but has such a factor,} \\ d & \text{if } x \text{ has no composite factor but has a cutpoint,} \\ e & \text{if } x \text{ has neither composite factor nor cutpoint;} \end{cases}$$

and

$$\text{type}(x) = \begin{cases} 1 & \text{if nat}(x) \in \{a, d\}, \\ 2 & \text{if nat}(x) \in \{b, e\}, \\ 3 & \text{if nat}(x) = c. \end{cases}$$

We define, by induction on $|x|$, the decomposition $d(x)$ and show simultaneously that if $|x| \geq 2$ and y is in $d(x)$ then

$$|\text{int}(y)| < |\text{int}(x)| \quad \text{or} \quad \text{type}(y) < \text{type}(x). \tag{3}$$

If $|x| = 1$ then we put $d(x) = (x)$ and there is nothing to prove. Assume then $|x| > 1$; we shall consider five cases according the value of nat(x).

Case 1 (nat$(x) = a$): Let $e = f(x)$. Since x is composite, it is in $X_e - C_e$. Since C_e is a code there exists a unique factorization of x in C_e^+, say $x = x_1 x_2 \ldots x_n$. Let $1 \leq i \leq n$. By Assertion 4.13, $\text{int}(x_i) \subseteq \text{int}(x)$. Now, $x_i \in C_e$, hence it is not composite and, by Assertion 4.14, $f(x_i) \notin \text{int}(x_i)$. On the other hand, $f(x_i) = f(x) \in \text{int}(x)$. Thus, $|\text{int}(x_i)| < |\text{int}(x)|$ and (3) is satisfied. It suffices to put $d(x) = (x_1, x_2, \ldots, x_n)$.

Case 2 (nat$(x) = b$): Let x_1 be the longest composite left factor of x and let $x = x_1 x_2$. Since x is not composite it follows that $x_2 \in A^+$. By Assertion 4.15, $|\text{int}(x_2)| < |\text{int}(x)|$. On the other hand, type$(x_1) = 1 < 2 =$ type(x); i.e. both x_1 and x_2 satisfy (3). Thus, it suffices to put $d(x) = (x_1, x_2)$.

Case 3 (nat$(x) = c$): Let v be the longest composite factor of x. Then, $v \in A^+$ and $x = uvw$, for suitable $u, w \in A^*$. Since x has no composite left factor, it follows that

$u \in A^+$. Now, v is also the longest composite right factor of uv. Then, by the dual of Assertion 4.15, we can conclude that $|\text{int}(u)| < |\text{int}(uv)|$. Since $\text{int}(uv) \subseteq \text{int}(x)$, it follows that $|\text{int}(u)| < |\text{int}(x)|$. We also have that $\text{type}(vw) = 2 < 3 = \text{type}(x)$; i.e. both u and vw satisfy (3). Thus, it suffices to put $d(x) = (u, vw)$.

Case 4 ($\text{nat}(x) = d$): Since x has a cutpoint, there exist $x_1, x_2 \in A^+$, such that $x = x_1x_2$ and $|\text{int}(x_1)|, |\text{int}(x_2)| < |\text{int}(x)|$. It suffices to put $d(x) = (x_1, x_2)$.

Case 5 ($\text{nat}(x) = e$): By Assertion 4.16, there exist $u, v \in A^+$ such that $x = uv$, u has a cutpoint and $|\text{int}(v)| < |\text{int}(x)|$. Clearly, u has no composite factor; hence, $\text{type}(u) = 1 < 2 = \text{type}(x)$ and both u and v satisfy (3). It suffices to put $d(x) = (u, v)$.

This concludes the definition of d and the verification of (3).

Now, let y be a direct descendant of x. Assertion 4.13 and (3) imply that either $|\text{int}(y)| < |\text{int}(x)|$ or $|\text{int}(y)| = |\text{int}(x)|$ and $\text{type}(y) < \text{type}(x)$. It follows that $h(x) \leq 3|\text{int}(x)| + \text{type}(x) - 3$, for every $x \in Y$. Thus, the height of F is at most $3|S|$.

Finally, we note that the (X_e, C_e)'s do form a stratification for F. Indeed, we have that $x \in \bigcup (X_e - C_e)$ iff x is composite iff $\text{nat}(x) = a$ and in this case, by construction, $d(x) \in \mathscr{F}(C_e)$. Also, if $\text{nat}(x) \neq a$ then the degree of x is two; hence, if x has degree at least three then $\text{nat}(x) = a$ and x belongs to $X_{f(x)} - C_{f(x)}$. The stratification is simple by Proposition 4.1 and Assertion 4.17, and this completes the proof of Proposition 4.18. □

5. Constructions of factorization forests

In this section we develop some techniques for constructing factorization forests, needed in the proof of the main result. The principal aim is to substitute portions of a factorization forest refining its type while maintaining a close control of its height. This will be done in Subsection 5.3, Lemma 5.4.

5.1. *Coded factorization forests*

A factorization forest over A is said to be *coded* if its external set is a code over A. Coded factorization forests play an important role in the sequel. Proposition 5.2 shows one way to obtain them. The proofs in this subsection are omitted since they can be done by standard methods described in [1].

Let Z be the external set of the factorization forest $F = (X, d)$. We define $d^* : X \to \mathscr{F}(Z)$ by

$$d^*(x) = \begin{cases} (x) & \text{if } x \in Z \\ d^*(x_1) \circ d^*(x_2) \circ \cdots \circ d^*(x_n) & \text{if } x \in X - Z, \end{cases}$$

where we assume that $d(x) = (x_1, x_2, \ldots, x_n)$ and $\circ$ represents concatenation in $\mathscr{F}(Z)$. An induction on the height of vertices shows that d^* is well defined. We say that $d^*(x)$ is the *frond* of x.

The next proposition points out why coded factorization forests are important for us.

Proposition 5.1. *Let $F=(X, d)$ be a factorization forest with external set Z. Then $X \subseteq Z^+$ and, for every $x \in X$, $d^*(x)$ is a factorization of x with components in Z. Further, if F is coded then $d^*(x)$ is the unique such factorization of x. Thus, if F and F′ are coded factorization forests which have identical internal and external sets then they have identical fronds.*

An alternate view of coded factorization forests will be useful. Indeed, they can be seen as faithful representations of alphabetical factorization forests over appropriate alphabets.

Let $\varphi: A'^+ \to A^+$ be an injective morphism of free semigroups. Then φ extends in a natural way to an injective morphism $\varphi: \mathcal{F}(A'^+) \to \mathcal{F}(A^+)$. Let $F'=(X', d')$ be a factorization forest over A'. The image of F' by φ is the factorization forest $\varphi(F')=(\varphi(X'), \varphi d' \varphi^{-1})$ over A.

Proposition 5.2. *Let $F=(X, d)$ be a coded factorization forest over A of height h with external set Z. Then there exists a set A', an injective morphism $\varphi: A'^+ \to A^+$ and an alphabetical factorization forest $F'=(X', d')$ over A' of height h such that $\varphi(A')=Z$ and $\varphi(F')=F$.*

Conversely, let us consider an alphabetical factorization forest $F'=(X', d')$ over A' of height h' and an injective morphism $\varphi: A'^+ \to A^+$. Then, $F=\varphi(F')$ is a coded factorization forest of height h' with external set $\varphi(A' \cap X')$. Further, if $\{(X'_i, C'_i) \mid 1 \leq i \leq n\}$ is a (simple) stratification for F' then $\{(\varphi(X'_i), \varphi(C'_i)) \mid 1 \leq i \leq n\}$ is a (simple) stratification for F.

Coded factorization forests can be used to give an alternate view of a stratification. Let $F=(X, d)$ be a factorization forest over A and let $\{(X_i, C_i) \mid 1 \leq i \leq n\}$ be a stratification for F. Define, for each i, $d_i: X_i \to \mathcal{F}(X_i)$ by

$$d_i(x)=\begin{cases} d(x) & \text{if } x \in X_i - C_i, \\ (x) & \text{if } x \in C_i. \end{cases}$$

Then, for each i, $F_i=(X_i, d_i)$ is a coded factorization forest of height one, with external set C_i. Further, the X_i's are pairwise disjoint and every vertex of degree at least three of F is an internal vertex in some F_i.

5.2. Composition of factorization forests

In this subsection we define an associative product of factorization forests. The aim is to substitute portions of a forest by portions of another one.

Let $F_i=(X_i, d_i)$ be factorization forests over an alphabet A, for $i=1, 2$. Let Y_i and Z_i denote, respectively, the sets of internal and external vertices in X_i. The

We still need to show that the normal forms—the existence and unicity modulo MOD of which have just been established—are of the form given in Proposition 2.2. To this effect, we remark that a ground normal form of sort *process* is made a priori out of operators δ, $+$, $;$, $\|$, $\mathbin{\|\!\_}$ and of α's (ground instances of atomic actions). Now, because of rule (i), a normal form cannot contain $\|$ symbols. Suppose it contains a $\mathbin{\|\!\_}$ symbol in a normal form t. Let ω stand for its leftmost-innermost occurrence. Then $t_{|\omega} = u \| v$, where u can only be built out of symbols α's, δ, $+$ and $;$. But

- if an α is the root of u, then (k) reduces $t_{|\omega}$,
- if δ is the root of u, then (j) reduces $t_{|\omega}$,
- if $+$ is the root of u, then (l) reduces $t_{|\omega}$,
- if $;$ is the root of u, we may write $u = u_1 ; u_2$. Then, by case analysis on u_1, one of the rules (g), (j), (m) reduces $t_{|\omega}$.

In any case, we derive a contradiction. This terminates the proof of Proposition 2.2. □

We obtain the following corollary.

2.3. Corollary. *Assume a given process specification* SPEC *such that the atomic and the composite actions have* locally *consistent and complete definitions. Then* SEM(SPEC) *is hierarchically complete and consistent w.r.t.* **set-of**(DATA) *and* **basic-process**(SPEC).

Proof. The only observer with range *data* is $::$. Completeness w.r.t. DATA derives from the fact that for any normal form p and any d of sort *data*, $p :: d$ is congruent to a term built with the constructors of sort *data* only (using (n), (o), (p), (q), (r) and EQ_A). Consistence w.r.t. DATA then stems from the confluence of the above system modulo MOD. Likewise, completeness and consistence w.r.t. **basic-process**(SPEC) follow from Proposition 2.2 and from the canonicity of the system. □

Note. It follows from Proposition 2.2 that all processes defined so far are *static*, i.e. that their normal form is "frozen" once for all. It has the canonical form shown in Fig. 2 (cf. Section 6.1 for further discussion).

2.4. Observational congruence

The semantics of a process specification SPEC has been defined as the classical semantics of SEM(SPEC), i.e. the class of all its hierarchical models. However, when focussing on their *process* component, not all models are equally interesting. In particular, some of them allow to distinguish between processes that one would wish to assimilate because there is no actual way to observe differences between them. There are at least two possibilities to consider less discriminating models:

Then the factorization forest $F = F_1 * F_2 * \cdots * F_n * F'$ is Ramseyan modulo f and has height $h \leq h' + \sum h_i$. Further, F and F' have identical vertex sets, identical external sets and identical fronds.

If, in addition, $\{(X_{ij}, C_{ij}) \mid 1 \leq j \leq n_i\}$ is a (simple) stratification for F_i, $1 \leq i \leq n$, then $\{(X_{ij}, C_{ij}) \mid 1 \leq i \leq n, 1 \leq j \leq n_i\}$ is a (simple) stratification for F.

Proof. Let $F = (X, d)$ be the factorization forest $F_1 * F_2 * \cdots * F_n * F'$. Let Z and Z' (Y and Y') be the external (internal) sets of F and F', respectively. Since the X_i's are pairwise disjoint subsets of X', since the internal vertices of F are exactly $X_i - C_i$ and since every vertex in $X_i - C_i$ is internal in F' (because (X_i, C_i) belongs to a stratification of F'), it follows that $X = X'$, $Y = Y'$, $Z = Z'$ and

$$d(x) = \begin{cases} d_i(x) & \text{if } x \in X_i - C_i, \\ d'(x) & \text{if } x \in X' - \bigcup (X_i - C_i). \end{cases}$$

We note that, in particular, $d(x) = d'(x)$, for every $x \in \bigcup C_i$.

We claim that, for every $x \in X_i - C_i$,

$$d'(x) = d_i^*(x). \tag{4}$$

Indeed, since (X_i, C_i) belongs to a stratification of F' we have that $d'(x) \in \mathscr{F}(C_i)$. Also, by definition, $d_i^*(x) \in \mathscr{F}(C_i)$. The claim follows since C_i is a code and both $d'(x)$ and $d_i^*(x)$ are factorizations of x.

Our next claim is that, for every $x \in X_i - C_i$, if $d_i^*(x) = (x_1, x_2, \ldots, x_m)$ then

$$d^*(x) = d^*(x_1) \circ d^*(x_2) \circ \cdots \circ d^*(x_m). \tag{5}$$

This follows, by induction on $|x|$, using the observation that $d(x) = d_i(x)$ and remembering that F_i is a coded factorization forest.

Now, by induction on $|x|$, we prove that $d^*(x) = d'^*(x)$. Since $Z = Z'$ the claim holds for external vertices. If $x \in X_i - C_i$, for some i, then, by (4), $d'(x) = d_i^*(x)$. Thus, if $d_i^*(x) = (x_1, x_2, \ldots, x_m)$ then, using (5), and the induction hypothesis, both $d'^*(x)$ and $d^*(x)$ result equal to $d^*(x_1) \circ d^*(x_2) \circ \cdots \circ d^*(x_m)$. On the other hand, if $x \in X - \bigcup (X_i - C_i)$ then $d(x) = d'(x)$ and the result follows from the induction hypothesis.

Now we claim that every vertex of F-degree at least three belongs to some $X_i - C_i$. Indeed, let x be an internal vertex of F in $X - \bigcup (X_i - C_i)$. Then $d(x) = d'(x)$. But, since the (X_i, C_i)'s form a stratification of F', the F'-degree of x is two. Hence, so is its F-degree. This establishes the claim.

Now it is easy to show that F is Ramseyan modulo f. Indeed, let x be a vertex of F-degree at least three. By the previous claim $x \in X_i - C_i$, for some i, and, consequently, $d(x) = d_i(x)$. Thus, $d(x)$ is a Ramseyan power modulo f since, by hypothesis, F_i is Ramseyan modulo f.

An argument similar to the one just given shows that if $\{(X_{ij}, C_{ij}) \mid 1 \leq j \leq n_i\}$, for $1 \leq i \leq n$, is a stratification for F_i then $\{(X_{ij}, C_{ij}) \mid 1 \leq i \leq n, 1 \leq j \leq n_i\}$ is a stratification for F.

The remaining parts of the proof depend on a careful analysis of the paths of F which is based on the next assertion. For a path c in F let $\mathrm{cont}(c)$ denote the set

$$\mathrm{cont}(c) = \{i \mid c \text{ contains some vertex in } X_i - C_i\}.$$

For every $u \in X$, let $r(u)$ denote the length of a longest path in F' with terminus in u or ∞ if F' is of infinite height and there is no such path. Note that $r(u) \leqslant h'$ for every $u \in X$.

Assertion 5.5. *If F contains a path from $u \in X$ to $v \in X$ then either F' contains a path from u to v or there exist i, $1 \leqslant i \leqslant n$, and $w \in X_i - C_i$ such that $v \in X_i$, F' contains a path from u to w and F_i contains a path from w to v.*

Further, if c is a path in F then

$$r(u) + |c| \leqslant h' + \sum_{i \in \mathrm{cont}(c)} h_i,$$

where $u \in X$ is the origin of c.

Proof. Let c be a path in F from u to v. We proceed by induction on $|c|$. Assume initially that u belongs to some $X_j - C_j$. Let c_1 be the longest left segment of c which is a path in F_j. Let c_2 be such that $c = c_1c_2$.

If c_1 terminates in v then the trivial path (u) in F' and $c = c_1$ in F_j, from $w = u$ to v satisfy the first part of the lemma. As for the other part we have that $\mathrm{cont}(c) = \{j\}$ and, since c is a path in F_j, $|c| \leqslant h_j$. Also, $r(u) \leqslant h'$, and this guarantees the second part.

Assume then that c_1 does not terminate in v and let u' be the terminus of c_1. We claim that $u' \in C_j$. Indeed, if $u' \in X_j - C_j$ then $d(u') = d_j(u')$ and u' not being the last vertex of c we conclude that the vertex following u' in c belongs to $d_j(u')$, hence to X_j, contradicting the choice of c_1. Now, the claim implies that u' belongs to $d_j^*(u)$ and that $|c_1| > 0$. Thus, by (4), u' belongs to $d'(u)$ and the first part follows from the induction hypothesis applied to c_2. As for the second part, observe initially that $r(u) + 1 \leqslant r(u')$. Assume now that $j \in \mathrm{cont}(c_2)$. Let c_3 be the shortest left segment of c_2 with terminus in $X_j - C_j$. Applying the induction hypothesis to c_3 we can conclude that there exists a path in F' from u' to some vertex w' in $X_j - C_j$. Then F' contains a nontrivial path from $u \in X_j - C_j$ to $w' \in X_j - C_j$. This contradicts the simplicity of $X_j - C_j$. We conclude that $\mathrm{cont}(c)$ is the disjoint union of $\{j\}$ and $\mathrm{cont}(c_2)$. Applying the induction hypothesis to c_2 we obtain

$$r(u') + |c_2| \leqslant h' + \sum_{i \in \mathrm{cont}(c_2)} h_i.$$

Recalling that $|c_1| \leqslant h_j$ we have

$$r(u) + |c| \leqslant r(u') - 1 + |c_1| + |c_2| \leqslant h' + \sum_{i \in \mathrm{cont}(c)} h_i.$$

This guarantees the second part of the assertion in this case.

Assume now that $u \in X - \bigcup (X_j - C_j)$. If $|c| = 0$ then we have nothing to prove; otherwise let u' be the vertex of c following u and let c' be the segment of c from u' to v. Since $d(u) = d'(u)$ it follows that u' belongs to $d'(u)$ and the first part follows from the induction hypothesis. As for the second part observe that $1 + r(u) \leq r(u')$. Also, $|c'| = |c| - 1$ and, by the induction hypothesis,

$$r(u') + |c'| \leq h' + \sum_{i \in \mathrm{cont}(c')} h_i.$$

Clearly, $\mathrm{cont}(c) = \mathrm{cont}(c')$; hence,

$$r(u) + |c| \leq r(u') - 1 + 1 + |c'| \leq h' + \sum_{i \in \mathrm{cont}(c)} h_i.$$

This concludes the proof of Assertion 5.5 □

Assume now that, for each i, $\{(X_{ij}, C_{ij}) \mid 1 \leq j \leq n_i\}$ is a simple stratification for F_i. We claim that each $X_{ij} - C_{ij}$ is simple in F. Indeed, let c be a path in F from $u \in X_{ij} - C_{ij}$ to $v \in X_{ij} - C_{ij}$; we shall show that $u = v$. Clearly c is a path in F from $u \in X_i - C_i$ to $v \in X_i - C_i$. We apply Assertion 5.5. If F' contains a path from u to v then in view of the simplicity of $X_i - C_i$ in F' we conclude that $u = v$. If there exist a j and $w \in X_j - C_j$ such that F' contains a path c', from u to w, and F_j contains a path c'', from w to v, then $v \in X_i - C_i$ and $v \in X_j$ imply that $i = j$. Thus, $w \in X_i - C_i$ and, in view of the simplicity of $X_i - C_i$ in F', we conclude that $u = w$. Then, c'' is a path in F_i from $u \in X_{ij} - C_{ij}$ to $v \in X_{ij} - C_{ij}$. In view of the simplicity of $X_{ij} - C_{ij}$ in F_i we conclude that $u = v$. Altogether, the stratification $\{(X_{ij}, C_{ij}) \mid 1 \leq i \leq n, 1 \leq j \leq n_i\}$ of F is simple.

Finally, Assertion 5.5 clearly implies that the height h of F is at most $h' + \sum h_i$. The proof of the Substitution Lemma is complete. □

5.4. *Restricting factorization forests*

In this final subsection we study restrictions of factorization forests.

Proposition 5.6. *Let $F = (X, d)$ be a factorization forest over A with height function $h : X \to \mathbb{N}$ and with external set Z. Let X' be a subset of X such that $d(X') \subseteq \mathscr{F}(X')$, i.e. the subset X' of X is closed under descendants. If $d' = d \mid X'$ is the restriction of d to X' then $F' = (X', d')$ is a factorization forest of height function $h' = h \mid X'$ and with external set $Z' = Z \cap X'$.*

If, further, $\{(X_i, C_i) \mid 1 \leq i \leq n\}$ is a (simple) stratification for F then $\{(X_i \cap X', C_i \cap X') \mid 1 \leq i \leq n\}$ is a (simple) stratification for F'.

The verification of this proposition is left to the reader.

6. Proof of the main result

We state the main result of the paper which will be proved in this section.

Theorem 6.1. *Every morphism $f: A^+ \to S$, from a free semigroup to a finite one, admits a Ramseyan factorization forest of height at most $9|S|$.*

Actually, we shall prove the following more precise statement from which Theorem 6.1 follows immediately.

Proposition 6.2. *Let $f: A^+ \to S$ be a morphism with S a finite semigroup. There exist a complete factorization forest F over A of height at most $9|S|$ and codes $R_e \subseteq f^{-1}(e)$, one for each idempotent e in S, such that*

$$\{(R_e^+, R_e) \mid e = e^2 \in S\}$$

is a stratification for F.

The desired factorization forest F will be obtained through a number of applications of the Substitution Lemma. We proceed to give a short description of this construction. We assume the reader is familiar with the ideal theory for semigroups and shall use [7, 12] as the sources for the results needed.

Initially we enumerate the regular $\mathscr{D}$-classes in S:

$$D_1, D_2, \ldots, D_n.$$

Then in each regular $\mathscr{D}$-class D_p we enumerate the group $\mathscr{H}$-classes:

$$H_{p1}, H_{p2}, \ldots, H_{pk_p}.$$

We denote by e_{pq} the identity of H_{pq}.

The construction of F is inspired on the Rees–Suschkewitsch Theorem characterizing completely 0-simple semigroups, even though we do not need all the precision of that result. The construction is represented in Fig. 1, where each internal vertex corresponds to an application of the Substitution Lemma while each external vertex corresponds to a factorization forest obtained through the explicit constructions in Section 4.

First a factorization forest F' is built together with a simple stratification indexed by the regular $\mathscr{D}$-classes of S. This is done in Lemma 6.7.

Then, for each $\mathscr{D}$-class D_p, a factorization forest F'_p is built together with a simple stratification indexed by the group $\mathscr{H}$-classes contained in D_p. This is done in Lemma 6.8.

Finally, for each group $\mathscr{H}$-class H_{pq}, a factorization forest F_{pq} is built together with a stratification with a unique block associated to the idempotent of H_{pq}. The forest F_{pq} results Ramseyan modulo f. This is done in Lemma 6.10.

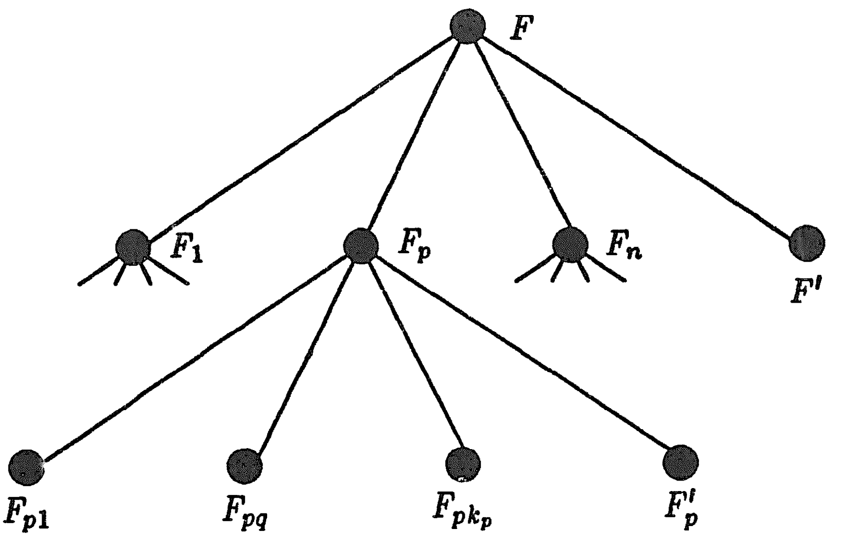

Fig. 1. Synthesis of F.

By applying the Substitution Lemma for each regular $\mathscr{D}$-class D_p, we obtain a factorization forest F_p, Ramseyan modulo f, given by

$$F_p = F_{p1} * F_{p2} * \cdots * F_{pk_p} * F'_p.$$

The union of the stratifications for F_{pq} shall constitute a stratification for F_p.

One more application of the Substitution Lemma yields the desired factorization forest:

$$F = F_1 * F_2 * \cdots * F_n * F'.$$

Forest F will result Ramseyan modulo f and the union of the stratifications for F_p shall constitute a stratification for F.

These constructions are interconnected through a family of codes over A. We shall have codes

$$P_p \subseteq f^{-1}(D_p), \qquad Q_{pq} \subseteq P_p^+ \cap f^{-1}(H_{pq}), \qquad R_{pq} \subseteq Q_{pq}^+ \cap f^{-1}(e_{pq}),$$

associated, respectively, with the regular $\mathscr{D}$-classes of S, with the group $\mathscr{H}$-classes of S and with the idempotents in S.

Before beginning the proof itself we review some facts about the local structure theory of finite semigroups which are not valid for semigroups in general. Of course, $\mathscr{J} = \mathscr{D}$ in a finite semigroup, hence every $\mathscr{J}$-class contains exactly one $\mathscr{D}$-class. The next four lemmas contain the facts we need.

Lemma 6.3. *Let S be a finite semigroup and let a, $b \in S$. Then, $a \mathscr{D} ab$ iff $a \mathscr{R} ab$. Dually, $b \mathscr{D} ab$ iff $b \mathscr{L} ab$.*

Proof. Since $\mathscr{R} \subseteq \mathscr{D}$ it suffices to prove that $a \mathscr{D} ab$ implies $a \mathscr{R} ab$. Now, $a \mathscr{D} ab$ implies $a \mathscr{J} ab$; hence, there exist u, $v \in S^1$ such that $a = uabv$. Then, for every m,

(ii) *for any* axiom $M = N$ *of* SEM(SPEC)^{process}, *for any ground substitution* σ, *and for any* $(q, d) \in ((T_{SPEC})^{data} \times (T_{SPEC})^{process})$,

$$[\phi(M\sigma) \| \phi(q)] :: \phi(d) \equiv_{\mathrm{DATA}_{\mathrm{IMPL}}} [\phi(N\sigma) \| \phi(q)] :: \phi(d).$$

The contexts $[X \| \phi(q)] :: \phi(d)$ correspond to the experiments inherited from the specification that might allow to differentiate between processes of the specification. Theorem 3.3 then means that for any processes p and p' that are equal in the specification, the previous experiments should not be able to discriminate between their implementations $\phi(p)$ and $\phi(p')$. We feel that this captures in an accurate way the intuition behind the implementation of processes. Notice that some experiments from IMPL itself may allow to discriminate between $\phi(p)$ and $\phi(p')$; for instance, a context $[X \| \underline{q}] :: \underline{d}$, where $\underline{q}$ is not equal to some $\phi(q)$ or $\underline{d}$ is not equal to some $\phi(d)$.

Proof of Theorem 3.3. First, rephrasing Definition 3.2 via Birkhoff's theorem, one has SPEC $\rightarrow^{\phi}$ IMPL iff for any axiom $M = N$ of SEM(SPEC) and for any ground substitution σ,

$$\mathrm{FR}^{\phi}(\mathrm{IMPL}) \models^{\mathrm{obs}} M\sigma = N\sigma.$$

On $\mathrm{FR}^{\phi}[\mathbf{Init}(\mathrm{IMPL})]$ we define the relation $\leftrightarrow^{\mathrm{obs}(\phi)}$:
- for t, t' of sort *data*, $t \leftrightarrow^{\mathrm{obs}(\phi)} t'$ iff $t \equiv_{\mathrm{DATA}_{\mathrm{IMPL}}} t'$,
- for p, p' of sort *process*, $p \leftrightarrow^{\mathrm{obs}(\phi)} p'$ iff

$$\forall (q, d) \in (T_{SPEC})^{data} \times (T_{SPEC})^{process},$$
$$(p \| \phi(q)) :: \phi(d) \equiv_{\mathrm{DATA}_{\mathrm{IMPL}}} (p' \| \phi(q)) :: \phi(d).$$

The natural observational congruence $\equiv^{\mathrm{obs}}_{\mathrm{FR}^{\phi}[\mathbf{Init}(\mathrm{IMPL})]}$ on $\mathrm{FR}^{\phi}[\mathbf{Init}(\mathrm{IMPL})]$, w.r.t. to the processes and data of SPEC, contains $\leftrightarrow^{\mathrm{obs}(\phi)}$ by application of Theorem 2.5. Thus, to prove the identity of the two relations, we simply show that the latter is a *congruence* for SPEC.

First, $\leftrightarrow^{\mathrm{obs}(\phi)}$ is an equivalence relation. Then the fact that $\leftrightarrow^{\mathrm{obs}(\phi)}$ is a SPEC-congruence is also deduced from Theorem 2.5. As an example, let us prove that it acts gracefully w.r.t. a context $K[X] = (X \| \pi)$, for some $\pi \in \mathrm{FR}^{\phi}[\mathbf{Init}(\mathrm{IMPL})]^{process}$. Suppose that $p \leftrightarrow^{\mathrm{obs}(\phi)} p'$ (both being of sort *process*). We have $\pi = \phi(\bar{\pi})$ for some $\bar{\pi} \in (T_{SPEC})^{process}$. Then, $\forall (q, d) \in (T_{SPEC})^{data} \times (T_{SPEC})^{process}$,

$$((p \| \pi) \| \phi(q)) :: \phi(d) = ((p \| \phi(\bar{\pi})) \| \phi(q)) :: \phi(d) = (p \| (\phi(\bar{\pi}) \| \phi(q))) :: \phi(d),$$

$$((p' \| \pi) \| \phi(q)) :: \phi(d) = ((p' \| \phi(\bar{\pi})) \| \phi(q)) :: \phi(d) = (p' \| (\phi(\bar{\pi}) \| \phi(q))) :: \phi(d),$$

and thus $K[p] \leftrightarrow^{\mathrm{obs}(\phi)} K[p']$.

The other identities are proved in a similar way.

We now remark that

$$\mathrm{FR}^{\phi}(\mathrm{IMPL}) \models^{\mathrm{obs}} M\sigma = N\sigma \Leftrightarrow \mathrm{FR}^{\phi}(\mathbf{Init}[\mathrm{IMPL}]) \models^{\mathrm{obs}} M\sigma = N\sigma.$$

We apply Proposition 4.8 and call $F' = (X', d')$ the resulting factorization forest. The crucial part of the proof is to show that $X_D = C_D$ for each irregular $\mathscr{D}$-class of S. To see this it is sufficient to show that $X_D - C_D = \emptyset$, so assume there exists a shortest word $x \in X_D - C_D$. Then, $x = x_1 x_2 \ldots x_m$ for some $m > 1$ and $x_i \in C_D$. Now, $X_D \subseteq f^{-1}(D)$ implies that $f(x), f(x_2) \in D$. By Lemma 6.4, $f(x_1 x_2) \in D$. Hence, $D^2 \cap D \neq \emptyset$ and, by Lemma 6.5, the $\mathscr{D}$-class D is regular. This contradiction establishes the claim.

Thus, we can eliminate from the stratification of F' the pairs (X_D, C_D), which correspond to irregular $\mathscr{D}$-classes; the remaining pairs still form a stratification of F'. Now, it suffices to rename the sets X_D and C_D and observe that the height of F' is at most $3|S/\mathscr{D}| \leqslant 3|S|$. $\square$

Lemma 6.8. *Let $1 \leqslant p \leqslant n$ and let $P_p \subseteq f^{-1}(D_p)$ be a code. There exist a coded factorization forest F'_p over A of height at most $3|D_p|$, with vertex set $P_p^+ \cap f^{-1}(D_p)$, with external set P_p, and codes $Q_{pq} \subseteq P_p^+ \cap f^{-1}(H_{pq})$, one for each group $\mathscr{H}$-class H_{pq} contained in D_p, such that $\{(Q_{pq}^+, Q_{pq}) \mid 1 \leqslant q \leqslant k_p\}$ is a simple stratification for F'_p.*

Proof. To alleviate the notation we shall write D and P instead of D_p and P_p. Let $\hat{\mathscr{R}} : S \to S/\mathscr{R}$ and $\hat{\mathscr{L}} : S \to S/\mathscr{L}$ denote the canonical projections corresponding to $\mathscr{R}$ and $\mathscr{L}$, respectively. Let I and Λ be the set of $\mathscr{R}$ and $\mathscr{L}$-classes of S contained in D; we shall consider the rectangular band $I \times \Lambda$. Finally, let B be an alphabet and let $\varphi : B^+ \to A^+$ be an injective morphism such that $\varphi(B) = P$. We shall use the morphism $g : B^+ \to I \times \Lambda$ defined by

$$g(b) = (\hat{\mathscr{R}} f\varphi(b), \hat{\mathscr{L}} f\varphi(b)),$$

for every $b \in B$. We alert the reader that, in general, $g \neq (\hat{\mathscr{R}} f\varphi, \hat{\mathscr{L}} f\varphi)$, because the product of elements of S in D might lie outside of D. The functions just defined are represented in Fig. 2.

Now we apply Proposition 4.18 to the function g and call F the resulting factorization forest. Before proceeding, note that, in view of the egg-box picture of D, we can conclude that $|I \times \Lambda| \leqslant |D|$; hence the height of F is at most $3|D|$. We consider the factorization forest $\varphi(F)$; note that the vertex set of $\varphi(F)$ is P^+.

We claim that the subset $X' = P^+ \cap f^{-1}(D)$ of P^+ satisfies the hypothesis of Proposition 5.6, i.e. X' is closed under descendants. Indeed, let $u_1, u_2, \ldots, u_k \in P$

$$\begin{tikzcd}
B^+ \arrow[r, "\varphi"] \arrow[d, "g"'] & A^+ \arrow[r, "f"] & S \arrow[dl, "\hat{\mathcal{R}}"'] \arrow[dr, "\hat{\mathcal{L}}"] & \\
I \times \Lambda & S/\mathcal{R} & & S/\mathcal{L}
\end{tikzcd}$$

Fig. 2. Objects in lemma.

be such that $u = u_1u_2 \dots u_k \in f^{-1}(D)$. Then, $\varphi(F)$ being a coded factorization forest, all (direct) descendants of u are of the form $u_iu_{i+1} \dots u_j$, for some $1 \leqslant i \leqslant j \leqslant k$. Now, since each u_l belongs to $f^{-1}(D)$, it follows, by Lemma 6.4, that $u_iu_{i+1} \dots u_j \in f^{-1}(D)$, as required.

We apply Proposition 5.6 to $\varphi(F)$ and X' and call F'_p the resulting factorization forest. Forest F'_p has height at most $3|D|$; its vertex set is X' and its external set is P, as required. Further, it has a simple stratification given by

$$\{(\varphi(C_e)^+ \cap f^{-1}(D), \varphi(C_e) \cap f^{-1}(D)), | e \in I \times \Lambda\}, \tag{6}$$

where C_e is defined in Proposition 4.18.

The core of the proof is given by the next assertion which states that if $e \in I \times \Lambda$ corresponds to a group $\mathscr{H}$-class, then the respective block of the stratification is the free semigroup generated by its code; otherwise it consists just of its code and can be eliminated.

Assertion 6.9. *Let $e = (i, \lambda) \in I \times \Lambda$ and let $H = i \cap \lambda$ be the $\mathscr{H}$-class which corresponds to e. Let us define $Q_H = \varphi(C_e) \cap f^{-1}(D)$. Then Q_H is a code contained in $P^+ \cap f^{-1}(H)$ and*

$$\varphi(C_e)^+ \cap f^{-1}(D) = \begin{cases} Q_H^+ & \text{if } H \text{ is a group } \mathscr{H}\text{-class,} \\ Q_H & \text{if } H \text{ is not a group } \mathscr{H}\text{-class.} \end{cases}$$

Proof. Initially we prove that $Q_H \subseteq f^{-1}(H)$. Let $u \in Q_H$, then there exist $k \geqslant 1$, and $b_j \in B$, for $1 \leqslant j \leqslant k$, such that $x = b_1b_2 \dots b_k \in C_e$, and $u = \varphi(x) \in f^{-1}(D)$. Let $i' = \hat{\mathscr{R}}f\varphi(b_1)$ and $\lambda' = \hat{\mathscr{L}}f\varphi(b_k)$. From the definition of g and that of the product of $I \times \Lambda$ we conclude that $g(x) = (i', \lambda')$. Now, $x \in C_e$ implies that $g(x) = e = (i, \lambda)$; hence $i = i'$ and $\lambda = \lambda'$. It remains to be proved that $\hat{\mathscr{R}}f(u) = i$ and $\hat{\mathscr{L}}f(u) = \lambda$; which implies that

$$u \in f^{-1}(i) \cap f^{-1}(\lambda) = f^{-1}(i \cap \lambda) = f^{-1}(H),$$

as required.

It suffices to prove that $\hat{\mathscr{R}}f(u) = i$, the other equality following by duality. Note that $\varphi(b_1) \in P \subseteq f^{-1}(D)$, hence $f\varphi(b_1) \in D$. If $k = 1$ then $f(u) = f\varphi(b_1)$, otherwise, $f\varphi(b_1) \in D$ and $f\varphi(b_1b_2 \dots b_k) = f(u) \in D$ imply, by Lemma 6.3, that $f\varphi(b_1) \,\mathscr{R}\, f(u)$. In any case, $f(u) \,\mathscr{R}\, f\varphi(b_1)$; hence, $\hat{\mathscr{R}}f(u) = \hat{\mathscr{R}}f\varphi(b_1) = i' = i$. This proves the claim.

Thus, we proved that $Q_H \subseteq f^{-1}(H)$. Since C_e is a code, φ is an injective morphism of free semigroups, and $Q_H \subseteq \varphi(C_e)$, it follows that Q_H is a code contained in $\varphi(C_e) \subseteq \varphi(B^+) = \varphi(B)^+ = P^+$. Hence, Q_H is a code contained in $P^+ \cap f^{-1}(H)$.

Now we prove that

$$\varphi(C_e)^+ \cap f^{-1}(D) = Q_H^+ \cap f^{-1}(H).$$

Since $f^{-1}(H) \subseteq f^{-1}(D)$ and $Q_H \subseteq \varphi(C_e)$, it follows that $\varphi(C_e)^+ \cap f^{-1}(D) \supseteq Q_H^+ \cap f^{-1}(H)$. To see the reverse inclusion, let $u \in \varphi(C_e)^+ \cap f^{-1}(D)$. Then there exist $k \geqslant 1$, and $x_j \in C_e$, for $1 \leq j \leq k$, such that $u = \varphi(x_1 x_2 \ldots x_k) \in f^{-1}(D)$. Since each x_j is a product of one or more b's in B and each $\varphi(b) \in P \subseteq f^{-1}(D)$, it follows, by Lemma 6.4, that $\varphi(x_j) \in f^{-1}(D)$, for every $1 \leq j \leq k$. Thus, for every $1 \leq j \leq k$, we have $\varphi(x_j) \in Q_H$, and, consequently, $\varphi(x_j) \in f^{-1}(H)$. Recall that $u = \varphi(x_1 x_2 \ldots x_k) \in f^{-1}(D)$; hence, using Lemmas 6.4 and 6.6, we can conclude, by induction on k, that $u \in f^{-1}(H)$. Also, $u \in Q_H^+$, hence $u \in Q_H^+ \cap f^{-1}(H)$, as claimed.

Finally, assume that H is a group $\mathscr{H}$-class. Since $Q_H \subseteq f^{-1}(H)$ and H is a subsemigroup of S, it follows that $Q_H^+ \subseteq f^{-1}(H)$. Thus, in this case, $Q_H^+ \cap f^{-1}(H) = Q_H^+$; hence $\varphi(C_e)^+ \cap f^{-1}(D) = Q_H^+$, as required.

On the other hand, assume that H is not a group $\mathscr{H}$-class. We claim that $Q_H^+ \cap f^{-1}(H) = Q_H$. Since $Q_H \subseteq f^{-1}(H)$ it is clear that $Q_H \subseteq Q_H^+ \cap f^{-1}(H)$. To see the reverse inclusion, let $u \in Q_H^+ \cap f^{-1}(H)$. Then there exist $k \geq 1$, and $u_j \in Q_H$, for $1 \leq j \leq k$, such that $u = u_1 u_2 \ldots u_k \in f^{-1}(H)$. Assume that $k > 1$. We claim that $u_1 u_2 \in f^{-1}(D)$. If $k = 2$ we have nothing to prove, otherwise, the claim follows from Lemma 6.4, since $u_2 \in f^{-1}(H) \subseteq f^{-1}(D)$. But, by Lemma 6.6, u_1, $u_2 \in f^{-1}(H)$ and $u_1 u_2 \in f^{-1}(D)$ imply that H is a group $\mathscr{H}$-class, a contradiction. Thus, $k = 1$ and $u \in Q_H$. It follows that $\varphi(C_e)^+ \cap f^{-1}(D) = Q_H$, as required. This completes the proof of Assertion 6.9. □

In view of Assertion 6.9 the blocks of the stratification (6) are either of the form (Q_H^+, Q_H) or (Q_H, Q_H), depending whether the corresponding $\mathscr{H}$-class H is a group or not. Clearly, the blocks of the second form can be eliminated; and renaming the remaining ones according to the enumeration of the group $\mathscr{H}$-classes contained in D, we achieve the proof of Lemma 6.8. □

Lemma 6.10. *Let $1 \leq p \leq n$, $1 \leq q \leq k_p$ and let $Q_{pq} \subseteq f^{-1}(H_{pq})$ be a code. There exist a coded factorization forest F_{pq} over A of height at most $3|H_{pq}|$, with vertex set Q_{pq}^+, with external set Q_{pq}, and a code $R_{pq} \subseteq Q_{pq}^+ \cap f^{-1}(e_{pq})$ such that $\{(R_{pq}^+, R_{pq})\}$ is a stratification for F_{pq}.*

Proof. To alleviate the notation we shall write H and Q instead of H_{pq} and Q_{pq}. Let B be an alphabet and let $\varphi : B^+ \to A^+$ be an injective morphism such that $\varphi(B) = Q$. We consider the morphism $g = f\varphi$. Since $Q \subseteq f^{-1}(H)$ and H is a subsemigroup of S, it follows that $Q^+ \subseteq f^{-1}(H)$. This implies that $g(B^+) \subseteq H$ and, consequently, we can consider H as the codomain of g. Now we apply Proposition 4.3 to g and call F the resulting factorization forest, whose height is at most $3|H|$. Let $F_{pq} = \varphi(F)$ and let $R_{pq} = \varphi(C_e)$, where C_e is the code defined in Proposition 4.3. Clearly, $R_{pq} \subseteq Q^+ \cap f^{-1}(e_{pq})$, where e_{pq} is the identity of H. Now, Proposition 5.2 implies that F_{pq} satisfies all the requirements of Lemma 6.10. □

We are ready to prove Proposition 6.2.

Proof of Proposition 6.2. Initially we apply Lemma 6.7. For each $1 \leq p \leq n$ we apply Lemma 6.8. Finally, for each $1 \leq p \leq n$ and $1 \leq q \leq k_p$, we apply Lemma 6.10.

Let us define, for each $1 \leq p \leq n$, the factorization forest $F_p = F_{p1} * F_{p2} * \cdots * F_{pk_p} * F'_p$. The properties of F_{pq} and of the simple stratification of F'_p allow the application of the Substitution Lemma to the above data. It results that F_p has a stratification

$$\{(R^+_{pq}, R_{pq}) \mid 1 \leq q \leq k_p\}.$$

Besides, the height h_p of F_p satisfies

$$h_p \leq 3|D_p| + \sum_{q=1}^{k_p} 3|H_{pq}| \leq 6|D_p|.$$

The second inequality is justified by the egg-box picture of a $\mathscr{D}$-class from which we can conclude that the summation is at most $3|D_p|$.

Now we define the factorization forest $F = F_1 * F_2 * \cdots * F_q * F'$. The properties of F_p and of the simple stratification of F' allow the application of the Substitution Lemma to the above data. It results that F is a complete factorization forest which has a stratification

$$\{(R^+_{pq}, R_{pq}) \mid 1 \leq p \leq n, 1 \leq q \leq k_p\}.$$

Besides, the height h of F satisfies

$$h \leq 3|S| + \sum_{p=1}^{n} 6|D_p| \leq 9|S|.$$

The second inequality is justified by the fact that the regular $\mathscr{D}$-classes of S constitute a partition of a subset of S from which we can conclude that the summation is at most $6|S|$. The proof of Proposition 6.2 is complete. □

We point out that Theorem 6.1 and Proposition 3.2 imply a result in [9] (see also [3]) stating that every morphism from a free semigroup to a finite one is strongly Ramseyan. The particular case for the morphism $\mathrm{alph}: A^+ \to 2^A$, given by

$$\mathrm{alph}(x) = \{a \in A \mid a \text{ occurs in } x\},$$

easily implies a result in [2] which has also been proved in [3, 8]. One statement of this result says that if

$$\mathbb{N} = B_1 \cup B_2 \cup \cdots \cup B_n$$

is a partition of $\mathbb{N}$ in finitely many sets then one of the sets B_j is piecewise syndetic (see [4, p. 34]).

In conclusion we remark that given a morphism $f: A^+ \to S$ and a code C over A, using Proposition 6.2, one can construct a Ramseyan factorization forest modulo f of height at most $9|S|$ with vertex set C^+ and with external set C, much in the same way as was done in Lemma 6.10.

7. An application

In this section we describe an application of the main theorem. A particular case of Proposition 7.1 plays a role in an alternative proof for a deep theorem of Hashiguchi [5, 6] characterizing recognizable limited $\mathcal{M}$-subsets of a free monoid.

Consider a morphism $f: A^+ \to S$, where S is a finite semigroup. An *f-decomposition* of $x \in A^2A^*$ is a factorization $(x_1, x_2, \ldots, x_n)$ of x, where $n \geq 2$, each $x_i \in A^+$, and either $n = 2$ or there exists an idempotent $e \in S$ such that $f(x_i) = e$, for $1 \leq i \leq n$. Of course, $x = x_1x_2 \ldots x_n$. We associate a *cost function* c to f, $c: A^+ \to \mathbb{N}$, defined as follows: $c(a) = 1$, for every $a \in A$, and, for $x \in A^2A^*$,

$$c(x) = \min\{c(x_1) + c(x_n) \mid x = x_1x_2 \ldots x_n \text{ is an } f\text{-decomposition of } x\}.$$

Next we show that $c(x)$ is bounded by a constant which depends only on $|S|$.

Proposition 7.1. *For any morphism $f: A^+ \to S$, with S a finite semigroup, $c(x) \leq 512^{|S|}$, for every x in A^+, where c is the cost function of f.*

Proof. Let us apply Theorem 6.1 and let $F = (A^+, d)$ be the resulting factorization forest. Let $h: A^+ \to \mathbb{N}$ be the height function of F. Proceeding by induction on the height of $x \in A^+$ we can conclude that $c(x) \leq 2^{h(x)}$. This is so because $d(x)$ is an f-decomposition of x, for every $x \in A^2A^*$. The proof is completed by the observation that $h(x) \leq 9|S|$ for every x. □

Acknowledgment

The existence of this work is mostly due to the spectacular hospitality of the people at the Laboratoire d'Informatique Théorique et Programmation in Paris. I thank especially Christian Choffrut, Dominique Perrin and Jacques Sakarovitch. I am also grateful to the Centro de Computação da Unicamp for permission to use their laser printer.

References

[1] J. Berstel and D. Perrin, *Theory of Codes* (Academic Press, New York, 1985).

[2] T.C. Brown, Van der Waerden's theorem on arithmetic progressions, *Notices Amer. Math. Soc.* **16** (1969) 245.

[3] T.C. Brown, An interesting combinatorial method in the theory of locally finite semigroups, *Pacific J. Math.* **36** (1971) 285-289.

[4] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory* (Princeton University Press, Princeton, 1981).

[5] K. Hashiguchi, Limitedness theorem on finite automata with distance functions, *J. Comput. System Sci.* **24** (1982) 233-244.

[6] K. Hashiguchi, Improved limitedness theorems on finite automata with distance functions, manuscript, 1986.

[7] J.M. Howie, *An Introduction to Semigroup Theory* (Academic Press, London, 1976).
[8] G. Jacob, Un théorème de factorisation des produits d'endomorphismes de K^N, *J. Algebra* **63** (1980) 389-412.
[9] J. Justin, Généralisation du théorème de van der Waerden sur les semi-groupes répétitifs, *J. Combin. Theory Ser. A* **12** (1972) 357-367.
[10] J. Justin and G. Pirillo, Two combinatorial properties of partitions of the free semigroup into finitely many parts, *Discrete Math.* **52** (1984) 299-303.
[11] J. Justin and G. Pirillo, On a natural extension of Jacob's ranks, *J. Combin. Theory Ser. A* **43** (1986) 205-218.
[12] G. Lallement, *Semigroups and Combinatorial Applications* (Wiley, New York, 1979).
[13] I. Simon, Word Ramsey theorems, in: B. Bollobás, ed., *Graph Theory and Combinatorics* (Academic Press, London, 1984).