Lecture 6d: Using Green's Relations

We now use Green's relations to prove two important theorems: Schutzenberger's theorem (which we have already seen in Lecture 5) and Simon's Factorization Forest Theorem. The material here is drawn from Thomas Colcombet's survey article [2], which the reader is strongly encouraged to read.

Schutzenberger's Theorem

First we show that for any finite monoid, if each of its regular $\mathcal{H}$-classes is trivial (contains a single element) then as a matter of fact all its $\mathcal{H}$-classes are trivial. We say that such a monoid is $\mathcal{H}$-trivial.

Proposition 1 *Let $(M, ., 1)$ be a finite monoid such that every regular $\mathcal{H}$-class is trivial. Then all its $\mathcal{H}$-classes are trivial.*

Proof: We observe that for such monoids, if N is the idempotent power of y then $y^N = y^N y$. This is because, $y^N = y^N y y^{N-1}$, so that $y^N \mathcal{J} y^N y$ and since $y^N y \leqslant_L y^N$ and $y^N \leqslant_R y^N$, we also have $y^N \mathcal{H} y^N y$. As every regular $\mathcal{H}$-class is trivial and y^N is an idempotent we have $y^N y = y^N$.

Let H be an $\mathcal{H}$-class and let $s, t \in H$. Therefore $s = xt$ and $t = sy$ for some x, y. So, $s = xsy$ and hence $s = x^N s y^N$ for any N. Choosing N to be the idempotent power of y we get $s = x^N s y^N = x^N s y^N y = sy = t$. ■

It turns out a similar result holds for $\mathcal{R}, \mathcal{L}$ and $\mathcal{J}$ classes and we leave that as an exercise to the interested reader.

Exercise Show that if every regular $\mathcal{R}$-class (respectively regular $\mathcal{L}$-class, regular $\mathcal{J}$-class) is trivial in a finite monoid then every $\mathcal{R}$-class (respectively $\mathcal{L}$-class, $\mathcal{J}$-class) is trivial.

As a consequence of the previous proposition we have the following result.

Proposition 2 *A monoid $(M, ., 1)$ is aperiodic if and only if it is $\mathcal{H}$-trivial.*

Proof: If M is aperiodic then it contains no groups and consequently the $\mathcal{H}$-class of any idempotent must be trivial. Applying the previous proposition the monoid must be $\mathcal{H}$-trivial. Conversely, if it $\mathcal{H}$-trivial it contains no nontrivial groups and hence is aperiodic. ■

We are now ready to reprove Schutzenberger's theorem using Green's relations.

Theorem 3 *Every language recognized by an aperiodic monoid is a star-free regular language.*

Proof: It suffices to show that for any M is an aperiodic monoid and morphism $h : \Sigma^* \longrightarrow M$ and each $s \in M$, $h^{-1}(s)$ is a star-free regular language.

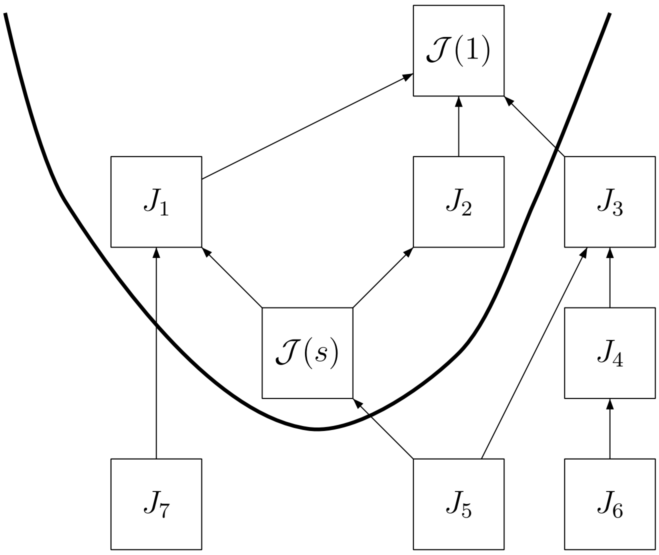
The proof proceeds by induction on $\leqslant_J$ (which if you recall, generalizes the idea of subwords) establishing that if $h^{-1}(t)$ is star-free for all t with $s <_J t$, then $h^{-1}(s)$ is also a star-free language. [However, the key combinatorial steps, remain the same as we shall see]

For the base case, we note that $\mathcal{J}(1)$ is the maximum element under $\leqslant_J$. Further we claim that for any monoid $\mathcal{J}(1) = \mathcal{H}(1)$. This follows from the fact $x \leqslant_L 1$ and $x \leqslant_R 1$ so that if $x\mathcal{J}1$ then $x\mathcal{H}1$. Now, using Proposition 2 we conclude that $\mathcal{J}(1) = \{1\}$ and $h^{-1}(1) = \{a|h(a) = 1\}^*$ is a star-free language.

The inductive step is carried out in 3 steps. Firstly, we show that $h^{-1}(\mathcal{J}(s))$ is a star-free language, and then that $h^{-1}(\mathcal{R}(s))$ (and $h^{-1}(\mathcal{L}(s))$ is also a star-free language. The final step is simple: $h^{-1}(\mathcal{H}(s)) = h^{-1}(\mathcal{L}(s)) \cap h^{-1}(\mathcal{R}(s))$ is therefore star-free and by Proposition 2, $\mathcal{H}(s) = \{s\}$, completing the inductive step.

Claim 1: $h^{-1}(\mathcal{J}(s))$ is star-free.

Consider the forbidden ideal $\mathcal{F}(s)$. In the language of Green's relations, this set can be defined to be the union of all the $\mathcal{J}$ classes J such that $\mathcal{J}(s) \nleqslant_J J$. Clearly, $\overline{\mathcal{F}(s)} = \{t \mid s \leqslant_J t\}$. The part below the curve in the following figure is $\mathcal{F}(s)$.



We show that $h^{-1}(\mathcal{F}(s))$ is a star-free language. Once we prove this, $h^{-1}(\mathcal{J}(s))$ can be expressed as

$$\overline{h^{-1}(\mathcal{F}(s)) \cup \bigcup_{s<_J t} h^{-1}(t)}$$

which is star-free (using the induction hypothesis), completing the proof of the claim.

Since $\mathcal{F}(s)$ is an ideal, $\Sigma^* a \Sigma^* \subseteq h^{-1}(\mathcal{F}(s))$ whenever $h(a) \in \mathcal{F}(s)$. Thus, the star-free language $L_0 = \bigcup_{h(a)\in\mathcal{F}(s)} \Sigma^* a \Sigma^*$ is contained in $h^{-1}(\mathcal{F}(s))$. Let $L_1 = h^{-1}(\mathcal{F}(s)) \backslash L_0$.

If $w \in L_1$ then any shortest subword u of w that also belongs to $h^{-1}(\mathcal{F}(s))$ must be of length at least 2 (otherwise $w \in L_0$). Let $u = avb$ and $h(v) = t$. We have

$$\{h(a), t, h(b), h(a).t, t.h(b)\} \subseteq \overline{\mathcal{F}(s)}$$

Since t is outside $\mathcal{F}(s)$, it is either $s\mathcal{J}t$ or $s <_J t$. Suppose $s\mathcal{J}t$. We already have, $h(a).t \leqslant_L t$ and $s \leqslant_J h(a).t$ (since the latter is in $\overline{\mathcal{F}(s)}$). Therefore, $h(a).t \ \mathcal{L}\ t$ and since $\mathcal{L}$ is a right congruence, we have $h(a).t.h(b) \ \mathcal{L}\ t.h(b)$. This is a contradiction since the l.h.s belongs to an ideal and the r.h.s. does not and we conclude that $s <_J t$. Then the language $\Sigma^*.a.h^{-1}(t).b.\Sigma^*$, which is star-free by the induction hypothesis, is contained in $h^{-1}(\mathcal{F}(s))$ and includes the word w.

For each w in L_1 we may pick such a star-free language and still we will only have finitely many! Thus we may write

$$\begin{aligned} h^{-1}(\mathcal{F}(s)) \quad = \quad L_0 \ \cup \ \bigcup\{\Sigma^*.a.h^{-1}(t).b.\Sigma^* \mid s <_J t, h(a).t \notin \mathcal{F}(s), t.h(b) \notin \mathcal{F}(s),& \\ h(a).t.h(b) \in \mathcal{F}(s)\}& \end{aligned}$$

This completes the proof the Claim 1. [We note that the argument is quite similar to the one used to establish that arbitrary ideals define star-free languages in the first proof of Schutzenberger's Theorem.]

Claim 2: $h^{-1}(\mathcal{R}(s))$ is star-free.

Let $h(w) \in \mathcal{R}(s)$ and let u be the shortest prefix of w such that $h(u) \in \mathcal{R}(s)$. If $u = \epsilon$ then $1\mathcal{R}s$ and we are in the base case. So, w.l.o.g. we may assume that $u = va$ and $h(v) = t$. Since we may write $w = uau'$ we have $h(w) = t.h(a).h(u')$. This means that $s \leqslant_R t.h(a)$ and $s \leqslant_R t$. So, either $s <_J t$ or $s\mathcal{J}t$.

If $s\mathcal{J}t$, then we also have $s\mathcal{R}t$ (since $s \leqslant_R t$ as well) which contradicts the minimality of u. So, we conclude the $s <_J t$. By the induction hypothesis, the language $L_{t,a} = h^{-1}(t).a.\Sigma^*$ is star-free and contains the word w.

But is it contained in $h^{-1}(\mathcal{R}(s))$? There is no real reason for this to hold. However the language $L_{t,a} \cap h^{-1}(\mathcal{J}(s))$ is contained in $h^{-1}(\mathcal{R}(s))$ as we now show (and contains w, since $h(w) \in \mathcal{R}(s) \subseteq \mathcal{J}(s)$). It is also star-free using Claim 1 and the fact that $L_{t,a}$ is a star-free language.

Let $x \in L_{t,a} \cap h^{-1}(\mathcal{J}(s))$. Therefore, $x = yaz$ with $h(y) = t$ and $t.h(a)\ \mathcal{R}\ s$. Consequently, $h(x) \leqslant_R s$. But since $h(x)\mathcal{J}s$, we conclude that $h(x)\mathcal{R}s$ as required.

Now we may pick a language $L_{t,a}$ for each $w \in h^{-1}(\mathcal{R}(s))$ as described above and still end up with just a finite collection so that we have

$$h^{-1}(\mathcal{R}(s)) \ = \ h^{-1}(\mathcal{J}(s)) \ \cap \ \bigcup\{L_{t,a} \mid s <_J t, t.h(a)\mathcal{R}s\}$$

Thus , $h^{-1}(\mathcal{R}(s))$ is a star-free language and this completes the proof of Claim 2. [We note the similarity between this argument and the proof that $xM\backslash F(x)$ is star-free in the first proof of Schutzenberger's theorem]

A similar proof establishes that $h^{-1}(\mathcal{L}(s))$ is also a star-free language hence completing the proof of Schutzenberger's theorem. [We observe that though the two key combinatorial ideas used here are the same as in the first proof, the use of the $\leqslant_J$ and identities on Green's relations significantly simplifies the inductive structure of the proof.] ■

The Factorization Forest Theorem

Let $(M, ., 1)$ be a monoid and let $m_1, m_2, \ldots, m_n$ be a sequence of elements from M. We would like to compute the product $m_1 m_2 \ldots m_n$. This can be done in many ways – for instance, with $n = 4$ we could multiply it as $(((m_1.m_2).m_3).m_4)$ or as $((m_1.m_2).(m_3.m_4))$ and so on. Each such expression yields a unique tree whose internal nodes are labelled . and whose leaves are labelled by $m_1, \ldots, m_n$ from left to right. We would like to minimize the height of such a tree, for instance among the two examples given above, we prefer the latter as it has height 2 (while the former has height 3). Since all this has really nothing to do with monoids (and just depends on the associativity of .), one can't really do better than a height of $log(n)$. However, we plan to generalize the expressions/trees now, making the problem (and its solution) all the more interesting.

The version of trees we have can be thought of as follows: each leaf is labelled by an element of M, each internal node has two children and it is labelled by $m.m'$ where m and m' are the labels of its two children. Each node denotes in some sense the effort/time to compute this product. (The height of the tree denotes in some sense, the number of steps needed to compute the entire product provided we can carry out "independent" products in parallel.) Now, if we take the view that if e is an idempotent then evaluating an expression of the form $e.e.e \ldots e$ can be considered *atomic* (needing just the same effort/time as computing a product), since the answer is just e, this leads to dramatic improvements in the height of the tree.

This generalized version of the expression tree is called a *factorization tree.* Formall, a factorization tree for $m_1, \ldots m_n$ is simply a tree where

1. The leaves of the tree when read from left to right gives the sequence $m_1, m_2, \ldots m_n$.

2. Each internal node with two children is labelled by $m.m'$ where m and m' are the labels of its children.

3. Each internal node with three or more children is labelled by an idempotent e and each of its children is also labelled by e.

Clearly the root of a factorization tree for $m_1, m_2 \ldots, m_n$ is labelled by $m_1 m_2 \ldots m_n$.

A remarkable result of I.Simon says that there is a constant K, that depends only on the size of M such that any sequence $m_1, \ldots m_n$ has a factorization tree of height at most K (in particular, it is independent of n). The proof below is from [2]. The bound proved below is not optimal and for an optimal construction the reader is referred to [**?**].

Theorem 4 *(Simon's Factorization Forest Theorem) Let* $(M, ., 1)$ *be a finite monoid. Every sequence* $m_1, m_2, \ldots, m_n$ *over* M *admits a factorization tree of height at most* $5|M|$.

Proof: We prove by induction w.r.t. the pre-order $\leqslant_J$, the following statement (which is stronger than the requirement of the theorem).

> If $m_1.m_2.\ldots.m_n = m$ then the sequence $m_1, m_2, \ldots, m_n$ has a factorization tree whose height is bounded by $5.|J_{\geqslant m}|$ where $J_{\geqslant m} = \{t | m \leqslant_J t\}$

For the base case, suppose m belongs to the maximum class w.r.t $\mathcal{J}$, then $m\mathcal{J}1$. But, as observed in the proof of Schutzenberger's theorem, $\mathcal{J}(1) = \mathcal{H}(1)$. Further, this also means that for every $1 \leqslant i \leqslant j \leqslant n$, $m_i m_{i+1} \ldots m_j \in \mathcal{H}(1)$.

We say that a sequence $m_1 m_2 \ldots m_n$ is *$\mathcal{H}$-smooth* (similarly *$\mathcal{R}$-smooth or $\mathcal{J}$-smooth*) if the entire set $\{m_i m_{i+1} \ldots m_j \mid 1 \leqslant i \leqslant j \leqslant n\}$ is contained within a single $\mathcal{H}$-class (single $\mathcal{R}$-class or $\mathcal{J}$-class respectively).

Thus, we have just observed that if $m_1 m_2 \ldots m_n \mathcal{J} 1$ then it is a $\mathcal{H}$-smooth sequence. We claim the following (whose proof is presented later) to complete the proof of the base case:

Claim 3: If $m_1, m_2 \ldots, m_n$ is a $\mathcal{H}$-smooth sequence and $m_1 m_2 \ldots m_n \in H$ then it has a factorization forest of height at most $3|H| - 1$.

For the inductive case, we construct a sequence of indices $i_1 < i_2 \ldots < i_r$ as follows: i_1 is the smallest index such that $m_1 m_2 \ldots m_{i_1} \mathcal{J} m$ (it exists since n is a candidate). We let i_{j+1} to be the smallest index such that $m_{i_j+1} \ldots m_{i_{j+1}} \mathcal{J} m$. Therefore we may write $m_1, m_2 \ldots m_n$ as $w_1, m_{i_1}, w_2, m_{i_2} \ldots w_r, m_{i_r}, w_{r+1}$ where $w_j = m_{i_{j-1}+1}, m_{i_{j-1}+2}, \ldots, m_{i_j - 1}$.

Let the product of the sequence w_{i_j} be c_j for $1 \leqslant j \leqslant r+1$ and let $b_j = c_j.m_{i_j}$ for $1 \leqslant j \leqslant r$. Then, by construction, $b_j \mathcal{J} m$ for each $1 \leqslant j \leqslant r$. Further, since $b_1 b_2 \ldots b_r c_{r+1} = m$, it follows that $b_i.b_{i+1} \ldots b_j \mathcal{J} m$. Thus, $b_1, b_2, \ldots b_r$ is a $\mathcal{J}$-smooth sequence. We now use the following claim, whose proof is provided later, to conclude the existence of a factorization tree T of height at most $4|\mathcal{J}(m)| - 1$ for $b_1, b_2, \ldots b_r$.

Claim 4: If $m_1, m_2 \ldots, m_n$ is a $\mathcal{J}$-smooth sequence and $m_1 m_2 \ldots m_n \in J$ then it has a factorization tree of height at most $4|J| - 1$.

This allows us to construct a factorization tree for $m_1, \ldots, m_n$ as described in Figure 1, where T_j is a factorization tree for w_j.

We observe that $m <_J c_j$ for each $1 \leqslant j \leqslant r+1$. This follows from the fact that $c_1.m_{i_1} \ldots c_{r+1} = m$ and that c_i is not in the same $\mathcal{J}$-class as m. So, by the induction hypothesis, the height of each T_j is not more than $5|J_{\geqslant c_j}|$. But $J_{\geqslant c_j} \subseteq J_{\geqslant m} \backslash \mathcal{J}(m)$ and so the height of each T_j is not more than $5.|J_{\geqslant m}| - 5.|\mathcal{J}(m)|$.

The overall height of the above tree is therefore not more than $5.|J_{\geqslant m}| - 5.|\mathcal{J}(m)| + 1 + (4|\mathcal{J}(m)| - 1) + 1$ which is not more than $5.|J_{\geqslant m}|$ as required, completing the proof of the theorem.

We now provide the proofs to the two claims used in this proof.

Proof of Claim 4: Let $\sigma = m_1, m_2, \ldots m_n$ be a $\mathcal{J}$-smooth sequence and $m_1 m_2 \ldots m_n = m$. Every $\mathcal{R}$-class inside the $\mathcal{J}$-class of m has the same size, say k. For any $\mathcal{J}$-smooth sequence $m_1, m_2, \ldots m_n$, we define its R-width ($R_w(m_1, m_2, \ldots m_n)$) to be the size of $\{\mathcal{R}(m_i \ldots m_j) | 1 \leqslant i \leqslant j \leqslant n\}$. By $\mathcal{J}$-smoothness, $\mathcal{R}(m_i \ldots m_j) = \mathcal{R}(m_i)$ and so we could have define R-width

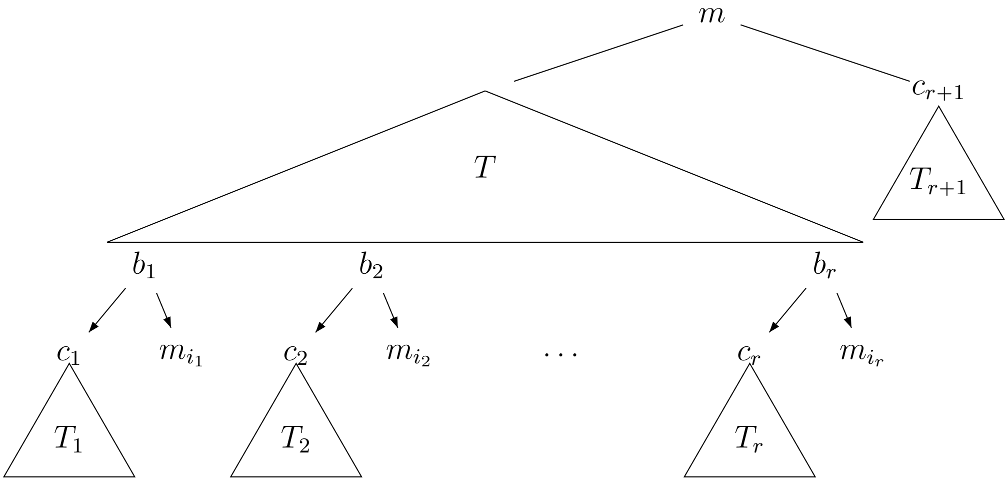


Figure 1: The general case and the case of $\mathcal{J}$-smooth factorizations

to be the size of $\{\mathcal{R}(m_i) | 1 \leqslant i \leqslant n\}$. We prove by induction on the size of the R-width of σ that it has a factorization tree of height at most $4.k.R_w(\sigma) - 1$.

If the R-width is 1 then by the following Claim, whose proof is provided later, we have a factorization tree of the requisite size.

Claim 5: If $m_1, m_2 \dots, m_n$ is a $\mathcal{R}$-smooth sequence and $m_1 m_2 \dots m_n \in R$ then it has a factorization forest of height at most $3|R| - 1$.

Let R be the $\mathcal{R}$-class of m. We identify the subsequence of all positions $i_1 < i_2 < \dots i_r$ such that $\mathcal{R}(m_{i_j}) = R$. (Clearly $i_1 = 1$.) We may then write $m_1, m_2, \dots, m_n$ as $m_1 w_1 m_{i_2} w_2 \dots m_{i_r} w_r$. By the observation about $\mathcal{J}$-smoothness, $\mathcal{R}(m_{i_j} w_j) = \mathcal{R}(m_{i_j}) = R$. Thus, writing b_j for $m_{i_j} w_j$, we conclude that the sequence $b_1 b_2 \dots b_r$ is $\mathcal{R}$-smooth and by Claim 5 has a factorization tree T of height $3.k - 1$.

Further, each w_j has strictly smaller R-width than $m_1, m_2, \dots, m_n$. Thus, by induction hypothesis, it has a factorization tree of height at most $4.k.(R_w(\sigma) - 1) - 1$. Thus, the overall height of the factorization tree obtained by combining these with T is bounded by $4.k.(R_w(\sigma) - 1) - 1 + 1 + (3k - 1)$ which simplifies to $4.k.R_w(\sigma) - k - 1$ and hence not more than $4.k.R_w(\sigma) - 1$. This completes the proof of Claim 4.

Proof of Claim 5: Let $\sigma = m_1, m_2, \dots, m_n$ be a given $\mathcal{R}$-smooth sequence and let $m_1 m_2 \dots m_n = m$. The proof proceeds by induction on the size of the set $\{\mathcal{H}(m_i) | 1 \leqslant i \leqslant n\}$, which we denote by $H_w(\sigma)$.

Let the size of any H-class within $\mathcal{R}(m)$ be k. We prove that σ has a factorization forest of size $3.k.H_w(\sigma) - 1$ which suffices to prove the Claim.

For the basis note that if $H_w(\sigma) = 1$ then $m_i \mathcal{H} m_j$ for all $1 \leqslant i, j \leqslant n$. Further $m_i m_{i+1} \dots m_j \leqslant_L m_i$ and $m_i m_{i+1} \dots m_j \leqslant_R m_j$ and therefore $m_i m_{i+1} \dots m_j \ \mathcal{H} \ m_i$ and the given sequence is actually $\mathcal{H}$-smooth. Thus we can use Claim 3 to obtain a factorization

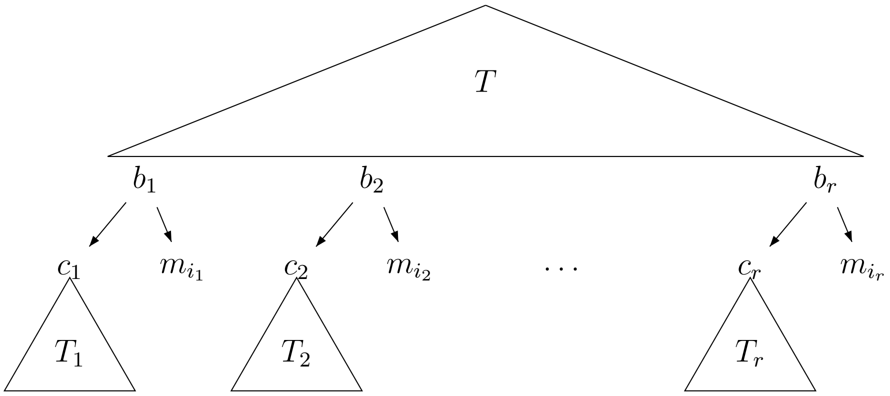


Figure 2: $\mathcal{R}$-smooth Factorizations

tree of the required size.

For the inductive case, let H be the $\mathcal{H}$-class of m_n. Let $i_1 < i_2 \ldots < i_r$ be the set of positions in the sequence such that $m_{i_j} \mathcal{H} m_n$. We can write $\sigma = w_1 m_{i_1} w_2 m_{i_2} \ldots w_r m_{i_r}$. We write c_j for the product of the sequence w_j and b_j for $c_j.m_{i_j}$. Clearly $b_i b_{i+1} \ldots b_j \leqslant_L m_{i_j}$, further since the given sequence is $\mathcal{R}$-smooth, $m_{i_j} \leqslant_R b_i b_{i+1} \ldots b_j$ so that $b_i b_{i+1} \ldots b_j \mathcal{H} m_{i_j}$. Thus, the sequence $b_1, b_2, \ldots, b_r$ is a $\mathcal{H}$-smooth sequence and we may use Claim 3 to conclude that it has a factorization tree T of size $3|H| - 1$.

Moreover, by construction, none of the sequences w_i contain any element of H and thus $H_w(w_j) < H_w(\sigma)$ for each j with $1 \leqslant j \leqslant r$. Thus, by the induction hypothesis, each w_j has a factorization tree T_j of height $3.k.(H_w(m_1, \ldots, m_n) - 1) - 1$.

We then combine these with the tree T as shown in Figure 2 to obtain a factorization tree for σ. The height of the resulting tree is bounded by $3.k.(H_w(\sigma) - 1) - 1 + 1 + (3.k - 1)$. Thus, it is bounded by $3.k.H_w(\sigma) - 1$ as required. This completes the proof of Claim 5.

Proof of Claim 3: Let $\sigma = m_1, m_2, \ldots, m_n$ be a $\mathcal{H}$-smooth sequence, with $m_1 m_2 \ldots m_n = m$ and let H be the $\mathcal{H}$-class of m. We first observe that if $n \geqslant 2$ then, by $\mathcal{H}$-smoothness, $H^2 \cap H \neq \emptyset$ and thus H is regular and hence it is a group. In what follows we assume that H is a group (and $n \geqslant 2$).

Let $S(\sigma)$ be the set $\{m_1 \ldots m_j \mid 1 \leqslant j \leqslant n\}$. The proof proceeds by induction on the size of this set and establishes that σ has a factorization tree of height $3.|S(\sigma)| - 1$ (which is bounded from above by $3.|H| - 1$ as required by the Claim)

For the basis, suppose this set is a singleton. Then $m_1.m_2 = m_1$ which means $m_2 = e$ where e is the identity of the group H. Similarly, $m_i = e$ for each $i > 1$. Thus, we have a factorization tree of height 2 (one to combine the idempotents and another to combine with m_1) for σ.

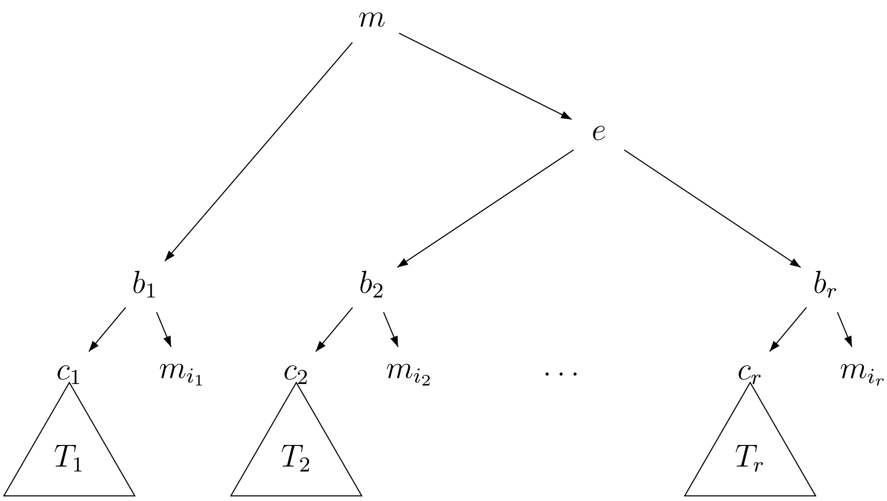


Figure 3: $\mathcal{H}$-smooth Factorizations

For the inductive step, let $i_1 < i_2 < \dots i_r$ be the subsequence of all those positions such that $m_1.m_2.\dots.m_{i_j} = m$. As usual we write σ as $w_1, m_{i_1}, w_2, m_{i_2}, \dots, w_{i_r}, m_{i_r}$. We claim that $|S(w_i)| < |S(\sigma)|$. For w_1, $S(w_1) \subseteq S(\sigma)\backslash\{m\}$ and thus $|S(w_1)| < |S(\sigma)|$. For $j > 1$, we note that if $t \in S(w_j)$ then $mt \in S(\sigma)$ since $w_1 m_{i_1} w_2 m_{i_2} \dots m_{i_{j-1}} = m$. Thus $mS(w_j) \subseteq S(\sigma)$ and once again, by construction $m \notin mS(w_j)$. Thus $|mS(w_j)| < |S(\sigma)|$ for all $2 \leqslant j \leqslant r$. But since H is a group $|S(w_j)| = |mS(w_j)|$ and thus we may apply the induction hypothesis to each w_j to obtain a factorization tree T_j for w_j of height $3.|S(w_j)| - 1$.

Now if $r = 1$ we combine the tree T_1 with m_n to obtain a tree of height at most $(3.|S(w_1)| - 1) + 1 \ < \ 3.|S(\sigma) - 1|$ as required.

Otherwise, let the product of w_j be c_j and let $c_j.m_{i_j} = b_j$. Therefore $b_1 = b_1.b_2 = \dots = b_1.b_2 \dots b_r$ and since H is a group this means that $b_2 = b_3 = \dots b_r = e$ where e is the identity of H. Thus we may contruct a factorization tree for σ by combining the trees T_j as described in Figure 3.

The height of this factorization is 3 plus the maximum height of the T_j's. This is bounded by $3 + 3.(|S(\sigma)| - 1) - 1 \ = \ 3.|S(\sigma)| - 1$ as required. This completes the proof of Claim 3 and hence the proof of the theorem. ■

Just in case you wondered why the theorem is called the Factorization *Forest* theorem, it is because it constructs a whole collection of trees (a forest), one for each sequence over the monoid (in such a way that the maximum height of the entire family is bounded by a function that depends only on the the size of the monoid).

Well-typed Regular Expressions

As an application of the factorization forest theorem we show how to construct regular expressions that are *consistent* w.r.t. a morphism recognizing the (regular) language.

Definition 5 *Let $h : (\Sigma^*, ., \epsilon) \longrightarrow (M, ., 1)$ be a morphism. A regular expression E (over the alphabet Σ) is said to be well-typed w.r.t. h if for each sub-expression E' of E and for each w, w' such that $w, w' \in L(E')$, we have $h(w) = h(w')$. The set of sub-expressions is inductively defined as follows:*

1. *if $Sub(\epsilon) = \{\epsilon\}$ and $Sub(a) = \{a\}$.*
2. *if $E = E_1 + E_2$ or $E = E_1.E_2$ then $Sub(E) = \{E\} \cup Sub\{E_1\} \cup Sub\{E_2\}$.*
3. *if $E = E_1^+$ then $Sub(E) = \{E\} \cup Sub(E_1)$.*

Notice, that in any well-typed regular expression E^+, $h(L(E^+)) = h(L(E)) = \{e\}$ for some idempotent $e \in M$.

Theorem 6 *Let $h : (\Sigma^*, ., \epsilon) \longrightarrow (M, ., 1)$ be a morphism into a finite monoid. For any $s \in M$, $h^{-1}(s)$ has a well-typed regular expression w.r.t. h. Consequently, every language recognised by h is a finite union of languages with well-typed regular expressions.*

Proof: Given any word $w = a_1 a_2 \ldots a_n$ over Σ, we shall refer to a factorization tree for the sequence $h(a_1), h(a_2), \ldots, h(a_n)$ as a factorization tree for w. By the factorization forest theorem such a factorization tree of height at most $5|M|$ exists for any w.

We construct a well-typed regular expression E_s^i, for each $s \in M$ and $i \leqslant 5|M|$, such that

$$L(E_s^i) \;=\; \{w \mid h(w) = s, \text{w has a factorization tree of height} \leqslant i\}$$

This is done by induction on i. We have

$$E_s^0 \;=\; \sum\{a \mid a \in \Sigma \cup \{\epsilon\}, h(a) = s\}$$

Further, if s is not an idempotent then

$$E_s^{i+1} \;=\; E_s^i \;+\; \sum\{E_u^i.E_v^i \mid u.v = s\}$$

and if s is an idempotent then

$$E_s^{i+1} \;=\; E_s^i \;+\; \sum\{E_u^i.E_v^i \mid u.v = s\} \;+\; (E_s^i)^+$$

The proof that $L(E_s^i)$ is the desired language is easy to establish by induction on i using the definition of factorization trees and is left as an exercise to the reader. Finally, using the factorization forest theorem we conclude that $E_s^{5|M|}$ is a well-typed regular expression for the language $h^{-1}(s)$. ■

For a number of other applications of Green's relations see [2] and for other applications of the Factorization Forest theorem see [1].

References

[1] M. Bojanczyk: “Factorization Forests”, *Proceedings of DLT 2009*, Springer LNCS 5583 (2009) 1-17.

[2] T. Colcombet: “Green’s Relations and their Use in Automata Theory”, *Proceedings of LATA 2011*, Spring LNCS 6638 (2011) 1-21.

[3] J.E.Pin: *Mathematical Foundations of Automata Theory*, MPRI Lecture Notes.