Lecture 5: Schutzenberger's Theorem

In this lecture we shall prove the theorem of Schutzenberger that relates languages recognized by aperiodic monoids with regular languages that can be described via star-free regular expressions. The proof described below is an adaptation of the proof given by Nick Pippenger in [1] (and Pippenger indicates that his proof follows the original presentation by Schutzenberger in [2]).

Theorem 1 *(Schutzenberger) Every aperiodic language can described via star-free expressions.*

The proof of this theorem proceeds by induction on the size of the monoid recognizing the language L. To start with we set out a few lemmas that help us to carry out the inductive argument. In this lecture every monoid we consider will be aperiodic.

Lemma 2 *For any x in an aperiod monoid M, if $x = pxq$ then $x = px$ and $x = xq$.*

Proof: Clearly $x = p^i x q^i$ for each i. But, since M is aperiodic, there is an N such that $p^N = p^{N+1}$. Thus, $x = p^N x q^N = p^{N+1} x q^N = px$. Similarly $x = xq$. ■

One easy consequence of this lemma is that if $e = ab$ then $e = a$ and $e = b$ for the identity e of an aperiodic monoid M.

We say that a subset I of M is an *ideal* if $IM \subseteq I$ and $MI \subseteq I$. Thus, an ideal is a subset that is closed w.r.t. multiplication (on both sides) by the elements of the monoid. Ideals are interesting subsets as one can define quotient monoids via ideals.

Definition 3 *Let M be an monoid and let I be an ideal of M. Then, there is a natural monoid M/I whose elements are $M - I \cup \{i\}$ and whose multiplication operation . is defined as follows:*

- $x.i = i.x = i.i = i$
- $x.y = i$, *if* $x.y \in I$
- $x.y$ *is the same as* $x.y$ *in* M *otherwise.*

It is easy to check M/I is a monoid. There is the obvious morphism η_I from M to M/I which is identity on the elements of $M - I$ and maps every element of I to i. Note that if a language L is recognized by a morphism h as $h^{-1}(I)$ in M then the same is recognised by $\eta \circ h$ to M/I as the pre-image of $\{i\}$. We shall simply write h to denote the composed map $\eta \circ h$ to M/I.

Thus, if I is an ideal of size more than 2 then any language recognized as the preimage of I can be recognised via a smaller monoid (M/I). More generally,

Lemma 4 *Let M be an finite aperiodic monoid, I be an ideal and let either $I \subseteq X$ or $X \cap I = \emptyset$. Then, any language L recognized as the preimage of X is recognized via the monoid M/I. In particular, if I has atleast 2 elements then L is recognized by a smaller aperiodic monoid.*

Definition 5 *With each element x of a monoid M we can associate a interesting ideal $F(x)$, called the forbidding ideal of x.*

$$F(x) = \{y \mid \forall p, q.pyq \neq x\}$$

It consists of all the elements that cannot "divide" x or cannot generate x via multiplication.

It is easy to check that $F(x)$ is an ideal for any $x \in M$.

Lemma 6 *Let h be a morphism from Σ^* to M. Then, $h^{-1}(x) = (\eta_{F(x)} \circ h)^{-1}(x)$. Thus, if $F(x)$ has at least 2 elements then the language recognized as $h^{-1}(x)$ can be recoginized using a smaller monoid.*

This follows from the fact that $x \notin F(x)$ and $F(x)$ is an ideal.

We are now in a position to describe the main ideas behind the proof. The proof, understandably, proceeds by induction on the size of the monoid M. If M is the trivial monoid, then the only languages recognised via M are $\emptyset$ and Σ^* and clearly both are star-free languages.

For the induction step, consider any language L recognized via some monoid M. Firstly, for any $X = \{x_1, x_2, \ldots, x_k\}$, $h^{-1}(X)$ is the union of the sets $h^{-1}(x_1)$, $h^{-1}(x_2)$, …, $h^{-1}(x_k)$. Thus, it suffices to show that $h^{-1}(x)$ can be expressed as a star-free expression involving languages definable using aperiodic monoids smaller than M.

If $F(x)$ has at least two elements, this would just be an application of Lemma 6. Otherwise, we need to do a lot of hard work. The idea is to show that we can find a collection Y of elements in M such that $h^{-1}(x)$ can be described as a star-free expression involving $h^{-1}(y)$, $y \in Y$, and further, for each $y \in Y$, $F(y)$ strictly contains $F(x)$. Once we do this, notice that we can always focus our attention on $h^{-1}(x)$ for only those elements with $|F(x)| > 1$ and complete our proof using Lemma 6.

Observe that $F(e) = M \setminus \{e\}$ for an aperiodic monoid (this follows from Lemma 2). Thus, $h^{-1}(e)$ poses any problems only in the case that M has fewer than 2 elements and we leave that as an exercise and assume henceforth that we are only interested in $h^{-1}(x)$ for $x \neq e$.

As a first step in that direction, we show that languages recognised by any ideal of M, even those of size 1, can be reduced to star-free expressions involving languages definable via smaller monoids.

Lemma 7 *If $L = h^{-1}(I)$ for some ideal I in M then L can be expressed using star-free expressions involving languages which are recognized by smaller aperiodic monoids.*

Proof: If the ideal $I = \emptyset$ then $L = \emptyset$ and it can be described by the star-free expression $\emptyset$.

Otherwise $|I| \geq 1$. (When $|I| > 1$, we can appeal to Lemma 7 to complete the proof. However, the following argument does not distinguish between this case and when $|I| = 1$.) Let $A = \{a \mid h(a) \in I\}$. Clearly, for each $a \in A$, the expression $E_a = \overline{\emptyset}.a.\overline{\emptyset}$ defines a language L_a contained in L (since I is an ideal). Thus, we could focus our attention on writing star-free expressions to cover words in $L \setminus \bigcup_{a \in A} L_a$.

Pick any word w in L. Consider a minimal substring of u of w that is in L.

If $u = \epsilon$ then the identity of M is in I which means that $M = I$ and thus $L = \Sigma^*$. If $u = a$ for some $a \in \Sigma$ then $w \in L_a$ and we need to do nothing in this case.

Thus we only need to consider the case when $u = avb$ for some v. Let $h(v) = y$. Note that, by the minimality of u, none of y, $h(a)y$ or $yh(b)$ can be in I.

Note that, since I is an ideal, $h(w_1).a.y.b.h(w_2) \in I$ for each $w_1, w_2 \in \Sigma^*$. Thus, $\overline{\emptyset}.a.h^{-1}(y).b.\overline{\emptyset} \subseteq L$. Note that this language contains the word w. If we show that $h^{-1}(y)$ may be described via a smaller monoid, this would take us one step closer to a star-free expression for L, since we have now managed to cover the word w.

Next we show that $F(y)$ has at least two elements, thus establishing that $h^{-1}(y)$ can be accepted via a smaller monoid $(M/F(y))$. Since $y \notin I$ and I is an ideal, $I \subseteq F(y)$ and since I is nonempty, there is at least one element in $F(y) \cap I$. We now show that there is at least one other element in $F(y)$.

Consider $h(a)y$. If $h(a)y \notin F(y)$ then there must be p, q such that $y = ph(a)yq$. Thus $y = ph(a)y$ and multiplying both sides by $h(b)$ we get $yh(b) = ph(a)yh(b)$ which is in I since $h(a)yh(b) \in I$. But this contradicts the minimality of u (since then $vb \in L$). Thus, $h(a)y \in F(y) \setminus I$. Thus $F(y)$ has at least two elements.

Finally, even though there are infinitely many w's outside of $\bigcup_{a \in A} L_a$ in L, the monoid M is finite and so is the alphabet Σ and thus we only have finitely many choices for triples of the form (a, y, b). Thus, we can describe all of $L \setminus \bigcup_{a \in A} L_a$ as a finite union of languages of the form $\overline{\emptyset}.a.h^{-1}(y).b.\overline{\emptyset}$, with $|F(y)| > 1$. This completes the proof of this lemma.

■

The other key idea behind Schutzenberger's proof is the following technical lemma :

Lemma 8 $x = (xM \cap Mx) \setminus F(x)$

Proof: Let $y \in (xM \cap Mx) \setminus F(x)$. Thus $y = px$ and $y = xq$ and $x = rys$. Then, by Lemma 2, $y = xq = rysq$. Thus $y = ry$. Similarly, $y = px = prys$ and thus $y = ys$. Thus $y = rys = x$. ■

We say "language defined by x" to mean $h^{-1}(x)$. Next we show that $h^{-1}(x)$ can be expressed as a star-free expression involving languages defined by other elements for which the forbidding set is larger than $F(x)$, for any $x \neq e$. (When $x = e$, $F(x) = M - \{e\}$ and assuming that $|M| > 2$ we can use Lemma 7.)

Lemma 9 *Let $x \in M$, $x \neq e$, then there is a subset $Y \subseteq M$ such that, $\forall y \in Y$. $F(y)$ strictly contains $F(x)$ and $h^{-1}(x)$ can be expressed as a star-free expression involving $h^{-1}(y)$, $y \in Y$ and other languages definable via smaller aperiodic monoids.*

Proof: We know that $x = (xM \cap Mx) \setminus F(x)$. Note that $h^{-1}((xM \cap Mx) \setminus F(x)) = h^{-1}(xM \setminus F(x)) \cap h^{-1}(Mx \setminus F(x))$ and $h^{-1}(xM \setminus F(x)) = h^{-1}(xM) \setminus h^{-1}(F(x))$.

$F(x)$ is an ideal and so, using Lemma 7, $h^{-1}(F(x))$ can be expressed via smaller monoids. We show that for each $w \in h^{-1}(xM \setminus F(x))$ one can find a letter a and an element y such that $h^{-1}(y).a.\overline{\emptyset} \setminus h^{-1}(F(x)) \subseteq h^{-1}(xM \setminus F(x))$, $w \in h^{-1}(y).a.\overline{\emptyset}$, where $F(y)$ has more elements than $F(x)$. This combined with the fact that $h^{-1}(F(x))$ is recognized by smaller monoids (by Lemma 7) gives the inductive argument.

Let $w \in h^{-1}(xM \setminus F(x))$. Let u be the shortest prefix of w such that $h(u) \in (xM \setminus F(x))$. If $u = \epsilon$, then then $e = xd$ for some $d \in M$ where e is the identity of M. This contradicts the aperiodicity of M unless $x = e$. But by assumption $x \neq e$. Thus, we may assume that u is not ϵ.

Thus $u = va$ for some v such that $h(v) = y \notin (xM \setminus F(x))$. We claim that $h^{-1}(y).a.\overline{\emptyset} \subseteq h^{-1}(xM)$. In proof, $h(avw') = h(va)h(w') = xm.d = xm'$. Thus it is in xM. Thus, $h^{-1}(y).a.\emptyset \setminus h^{-1}(F(x))$ is a subset of $h^{-1}(xM \setminus F(x))$ containing w.

Next we show that $F(y)$ has more elements than $F(x)$. First of all $y \notin F(x)$. Otherwise, $yh(a)$ will also be in $F(x)$ leading to a contradiction. Thus $F(x) \subseteq F(y)$. We now show that $yh(a)$ is also in $F(y)$. Suppose $yh(a)$ is not in $F(y)$. Then $y = pyh(a)q$. This means that $y = py$ and $y = yh(a)q$. Thus $y = xdq$ (since $yh(a) \in xM$) and thus $y \in xM$. But we already showed that $y \notin F(x)$ and thus $y \in xM \setminus F(x)$. This contradicts the minimality of u. Thus it must be the case that $yh(a) \in F(y)$ and thus $F(y)$ has at least one more element than $F(x)$.

Thus, we have picked an arbitrary element w of $h^{-1}(xM \setminus F(x))$ and shown that there is a star-free regular expression involving $h^{-1}(y)$ for some y with $F(y)$ strictly containing $F(x)$, that describes a language containing w and which is contained in $h^{-1}(xM \setminus F(x)) \cup h^{-1}(F(x))$. Since M and Σ are finite sets, it follows that we can write a star-free regular expression involving some elements $y_1, \ldots y_k$ of M, where $F(y_i)$ strictly contains $F(x)$ for each i, that describes a language containing $h^{-1}(xM \setminus F(x))$ and which is contained in $h^{-1}(xM \setminus F(x)) \cup h^{-1}(F(x))$. The proof follows from Lemma 7 and the fact that star-free expressions may use the boolean operations.

A similar argument works for $h^{-1}(Mx \setminus F(x))$ and this completes the proof of this lemma. ■

Finally, we can establish Schutzenberger's theorem.

Proof: As indicated earlier if M is the trivial monoid, it can only describe $\emptyset$ and Σ^* and both of these are star-free languages. There is only one aperiod monoid with $|M| = 2$ and we leave that case as an easy exercise.

For the induction step, since star-free languages are closed under union, it suffices to consider $h^{-1}(x)$ for some $x \in M$. If $x = e$, since $|M| > 2$ we apply Lemma 7 and complete the proof from the induction hypothesis. Otherwise, applying Lemma 9 twice, we may conclude that there is a set $Y \subseteq M$ such that $F(y)$ contains two more elements than $F(x)$ for each $y \in Y$ and further $h^{-1}(x)$ can be expressed as a star-free expression involving $h^{-1}(y)$, $y \in Y$ and other languages definable using small aperiodic monoids. But then $h^{-1}(y)$ is recognizable via a smaller monoid $M/F(y)$ for each $y \in Y$. Thus $h^{-1}(x)$ can be described via a star-free expression that only involves languages definable via smaller aperiodic monoids and we may now apply the induction hypothesis to conclude that $h^{-1}(x)$ is an aperiodic language. ■

References

[1] Nick Pippenger: *Theories of Computability*, Cambridge University Press, 1997.

[2] M.P.Schutzenberger: "On Finite Monoids Having only Trivial Subgroups", *Information and Control* **8** (1965) 190-194.