Schützenberger's Aperiodic Monoid Characterization of Star-Free Languages

Nov 2019

Prelude

- A star-free language is one that can be described by a regular expression constructed from the letters of the alphabet, the empty set symbol, all boolean operators, but no Kleene star.
- They can also be characterized logically as languages definable in $FO[<]$.
- and as languages definable in linear temporal logic
- Marcel-Paul Schützenberger characterized star-free languages as those with aperiodic syntactic monoids

Monoids

Definition : A 'monoid' is a set $M \neq \emptyset$ equipped with a binary operation $\cdot : M \times M \to M$ such that

- $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in M$
- $\exists \mathbb{1} \in M$ such that $a \cdot \mathbb{1} = a = \mathbb{1} \cdot a$ for all $a \in M$

Here, $\mathbb{1}$ is called the 'identity' element of M. Also, the identity must be unique.
We denote the monoid and it's operation along with it's identity by a triplet $(M, \cdot, \mathbb{1})$

Monoids : Examples

- $(\mathbb{Z}, +, 0)$ i.e the set of integers with integer addition, and 0 as the additive identity.
- $(\mathbb{N}, \cdot, 1)$ i.e the set of positive integers with integer multiplication, and 1 as the multiplicative identity.
- For any $n \in \mathbb{N}$, $(\mathbb{Z}_n, +, \bar{0})$ is a finite monoid, where $\mathbb{Z}_n$ is the set of residue classes of integers modulo n, $+$ is addition integers modulo n, and $\bar{0}$ is the residue class of zero.
- $(A^*, \cdot, \epsilon)$, where A is any alphabet, $\cdot$ is the concatenation of strings, and ϵ is the empty string, is a monoid.

Idempotent element

An element m in a monoid M, is called an 'idempotent element' if $m^2 := m \cdot m = m$.

Proposition 1. :

Every element in a finite monoid has an idempotent power.

proof : Let $m \in M$ be arbitrary. Then $m^n \in M$, for all $n \in \mathbb{N}$, and since $\mid M \mid < \infty$, we know $\exists i, p \in \mathbb{N}$ such that, $m^{i+p} = m^i$. In fact, $m^{i+rp} = m^i$, $\forall r \in \mathbb{N}$.
Thus taking $k = rp$ such that $k \geq i$, we get
$(m^k)^2 = m^{2k} = m^{k+rp} = m^{k-i} \cdot m^{i+rp} = m^{k-i} \cdot m^i = m^k$. □

Corollary : $\exists\ \omega \in \mathbb{N}$ such that m^ω is idempotent $\forall m \in M$.

proof : $\forall\ m \in M\ \exists\ k_m$ such that m^{k_m} is idempotent, and take $\omega = LCM\{k_m : m \in M\}$. □

We call smallest such ω is called the 'exponent' of M

Green's relations

Let M be a monoid. We define four relations on M as:

i) $s \leqslant_R t$ iff $s = tu$ for some $u \in M$
ii) $s \leqslant_L t$ iff $s = ut$ for some $u \in M$
iii) $s \leqslant_J t$ iff $s = utv$ for some $u \in M$
iv) $s \leqslant_H t$ iff $s \leqslant_R t$ and $s \leqslant_L t$

Equivalently,

i) $s \leqslant_R t$ iff $sM \subseteq tM$
ii) $s \leqslant_L t$ iff $Ms \subseteq Mt$
iii) $s \leqslant_J t$ iff $MsM \subseteq MtM$
iv) $s \leqslant_H t$ iff $s \leqslant_R t$ and $s \leqslant_L t$

Green's relation's

We define the equivalence relations :

i) $s\mathcal{R}t$ iff $sM = tM$

ii) $s\mathcal{L}t$ iff $Ms = Mt$

iii) $s\mathcal{J}t$ iff $MsM = MtM$

iv) $s\mathcal{H}t$ iff $s\mathcal{R}M$ and $s\mathcal{L}t$

v) $s\mathcal{D}t$ iff $\exists u \in M$ such that $s\mathcal{R}u$ and $u\mathcal{L}t$
iff $\exists u \in M$ such that $s\mathcal{L}v$ and $v\mathcal{R}t$

Green's relations

Theorem 1. :

In a finite monoid, the *Green's* relations $\mathcal{J}$ and $\mathcal{D}$ are equal. Furthermore,

i) $s \leqslant_J sm \implies s\mathcal{R}(sm)$

ii) $s \leqslant_J ms \implies s\mathcal{L}(ms)$

iii) $s\mathcal{J}t \wedge s \leqslant_R t \implies s\mathcal{R}t$

iv) $s\mathcal{J}t \wedge s \leqslant_L t \implies s\mathcal{L}t$

v) $\exists u, v \in M \ (s = usv) \implies (us)\mathcal{H}s\mathcal{H}(sv)$

Green's Lemma

Theorem. :

Let $s, t \in M$ such that $s\mathcal{R}t$. Let $s = tp$ and $t = sq$ then, the maps, $x \mapsto xp$ and $x \mapsto xq$ are bijections from $\mathcal{L}(t)$ onto $\mathcal{L}(s)$, and from $\mathcal{L}(s)$ onto $\mathcal{L}(t)$, resp.
Moreover, these bijections preserve $\mathcal{H}$-classes and are inverse to one another.

Ordered monoids

Definition :

An ordered monoid $(M, \leqslant)$ is a monoid M with an order relation $\leqslant$, such that $x \leqslant y$ implies $uxv \leqslant uyv$, $\forall\ u, v \in M$.

Definition :

An upper set in an ordered monoid $(M, \leqslant)$ is a subset $P \subseteq M$ such that $u \in P$ and $u \leqslant v$ implies $v \in P$.

Definition :

Given an upper set P in an ordered monoid $(M, \leqslant)$, the 'syntactic order' relation on M is defined as,
$u \leqslant_P v$ iff $xuy \in P \implies xvy \in P$ for all $x, y \in M$

Ordered monoids : Examples

- Any monoid $(M, \cdot, \mathbb{1})$ can be equipped with the equality order relation $(=)$ to get ordered monoid $(M, =)$.
- The natural order on positive integers is compatible with addition. Thus $(\mathbb{N}, +, 0, \leqslant)$ is an ordered monoid.

Homomorphisms

Definitions :

Given two monoids $(M, \cdot, id_M)$ and $(N, \times, id_N)$,
a monoid homomorphism (also called 'morphism') is a map $\varphi : M \to N$ such that $\varphi(m_1 \cdot m_2) = \varphi(m_1) \times \varphi(m_2)$, for all $m_1, m_2 \in M$.

Note :

- $(\varphi(M), \times, \varphi(id_M))$ is a monoid.
- $\varphi(id_M) = id_N$

Recognisable sets

Definition :

Given a monoid M a subset $L \subseteq M$ is said to be 'recognisable' if $\exists$ a finite monoid, N, a morphism $\varphi : M \to N$, and a set $X \subseteq N$ such that, $L = \varphi^{-1}(X)$. We say that the pair (φ, X) recognises L.

Definition(Recognisable language)**:**

Given an alphabet A, a language $L \subseteq A^*$ is called a 'recognisable language' if it is a recognisable set in the monoid $(A^*, \cdot, \epsilon)$.

Transition monoid

- Given a DFA, $\mathcal{A} \equiv (Q, A, \delta, s, F)$, for each $w \in A^*$, we define $f_w : Q \to Q$ as $q \mapsto \hat{\delta}(q, w)$. Then $(F_{\mathcal{A}}, \circ, id)$ is a finite monoid called the 'transition monoid' of $\mathcal{A}$, where $F_{\mathcal{A}} := \{f_w : w \in A^*\}$, $\circ$ is the composition of maps, and id is the identity map.
- The morphism-set pair $(\varphi, \varphi(\mathscr{L}(A)))$ recognises $\mathscr{L}(\mathcal{A})$ where φ is defined as $w \mapsto f_w$, $\forall\ w \in A^*$.
- Given a language $L \subseteq A^*$ recognised by a (finite) monoid M by the pair (φ, X), $\mathcal{A} := (M, A, \delta, id, X)$ is a DFA, where $\delta(m, a) := m \cdot \varphi(a)$, $\forall m \in M, a \in A$, and $L = \mathscr{L}(\mathcal{A})$.

Thus, given an alphabet A the set of regular languages over A is precisely the set of recognisable languages over A.

Syntactic monoid

Definition :

Given a subset $X \subseteq M$ where M is a monoid, the 'syntactic congruence' of X is defined as the relation on M as :
$u \cong_X v$ iff $xuy \in X \iff xvy \in X$ for all $x, y \in M$.

Note that $\cong_X$ is an equivalence relation on M.

Definition :

The syntactic monoid of $X \subseteq M$ is defined as the monoid $M/\cong_X$ i.e the set of equivalence classes of the syntactic congruence of X. The 'ordered syntactic monoid' of X is the ordered monoid $(M/\cong_X, \leqslant_X)$

Syntactic monoid of a language

Two monoids M and N are said to be 'isomorphic' if $\exists$ an 'isomorphism' i.e a bijective morphism, $\varphi : M \to N$.

Proposition 2. :

Given an alphabet A, the syntactic monoid of a recognisable language is isomorphic to the transition monoid of it's minimal DFA.

Aperiodic monoid

Definition :

A finite monoid M is said to be 'aperiodic', if $\forall\ m \in M$ $\exists n \in \mathbb{N}$ such that $m^n = m^{n+1}$.

E.g : Let $M := \{0, 1, a, b\}$ and define

- $0 \cdot m := 0 =: m \cdot 0$ and $1 \cdot m := m =: m \cdot 1\ \forall m \in M$,
- $a^2 := 0 =: b^2$
- $a \cdot b := 1 =: b \cdot a$

Then $(M, \cdot, 1)$ is an aperiodic monoid since $m^2 = m^3$ for all $m \in M$.

Aperiodic monoids

Proposition 3. :

Let M be a finite monoid. Then the following are equivalent,

1) M is aperiodic.
2) $\exists\ n \in \mathbb{N}$ such that $\forall\ m \in M,\ m^n = m^{n+1}$.

proof :

To see that 1) $\implies$ 2), assume M is aperiodic. Then take $n = max\{n_m : m \in M\}$ such that $\forall m \in M,\ m^{n_m} = m^{n_m+1}$, so $m^n = m^{n_m} \cdot m^{n-n_m} = m^{n_m+1} \cdot m^{n-n_m} = m^{n+1}, \forall\ m \in M$
The converse, 2) $\implies$ 1) is trivial □

Aperiodic monoids

Proposition 4. :

A finite ordered monoid $(M, \leqslant)$ is aperiodic iff $\forall\ m \in M$, $\exists\ n \in \mathbb{N}$ such that $m^{n+1} \leqslant m^n$.

proof :

If M is aperiodic, then $\forall m \in M$, $\exists\ n \in \mathbb{N}$ such that $m^{n+1} = m^n \implies m^{n+1} \leqslant m^n$.
Conversely, suppose $\forall m \in M$, $\exists n \in \mathbb{N}$ such that $m^{n+1} \leqslant m^n$. Then taking ω as a multiple of the exponent of M such that $\omega \geq n$ we get, $m^\omega = m^{2\omega} \leqslant m^{2\omega-1} \leqslant \ldots \leqslant m^{\omega+1} \leqslant m^\omega$, so that $m^{\omega+1} = m^\omega\ \forall\ m \in M$. □

Aperiodic monoids

Lemma 1. :

Let $L_1, L_2 \subseteq A^*$ be recognisable languages and $L := L_1 L_2$. If M_1, M_2, and M are the ordered syntactic monoids recognising L_1, L_2, and L respectively. Then,
M_1 and M_2 are aperiodic $\implies$ M is aperiodic.

Lemma 2. :

A finite monoid M is aperiodic $\implies$ M is $\mathcal{H}$-trivial.

Lemma 3. :

Let M be an aperiodic monoid and let $m \in M$. Then $\{m\} = (mM \cap Mm) \setminus J_m$, where $J_m := \{s \in M : m \notin .MsM\}$

Simplification lemma :

Lemma :

Let M be an aperiodic monoid and let $p, q, r \in M$. Then $pqr = q \implies pq = q = qr$.

proof :

Let n be the exponent of M. Since $pqr = q$, $p^n q r^n = q$. And since M is aperiodic, $p^n = p^{n+1}$ and hence $q = p^n q r^n = p^{n+1} q r^n = p(p^n q r^n) = pq$.
And similarly, $qr = q$ $\square$

Star-Free language

Definition :

Given an alphabet A, the set of 'star-free' languages in A is the smallest set $\mathscr{R} \subseteq 2^{A^*}$ such that,

a) $\varnothing \in \mathscr{R}$, $\{\epsilon\} \in \mathscr{R}$, and $\{a\} \in \mathscr{R}$, $\forall\ a \in A$.

b) $S, T \in \mathscr{R} \implies A^* \setminus S \in \mathscr{R}$, $S \cup T \in \mathscr{R}$, and $S \cdot T \in \mathscr{R}$

Notation :

$L_1 + L_2$ denote $L_1 \cup L_2$, 0 denote $\varnothing$, 1 denote $\{\epsilon\}$, u denote $\{u\}$ for all $u \in A^*$, L^c denote $A^* \setminus L$, and $L_1 L_2$ denote $L_1 \cdot L_2$

Thus, a star-free language is one that can be described by the letters in $A \cup \{0, 1\}$ and operations $\{+, ^c\}$

Star-Free languages : Examples

- Any finite language $L \subseteq A^*$ is star-free, since $L = \sum_{i=1}^{n} \left(\prod_{j=1}^{m_i} a_{ij} \right)$, where $L = \{a_1, ..., an\}$ and $a_i = a_{i1}...a_{im_i}$, where $a_{ij} \in A \; \forall j \leq m_i, \; i \leq n$
- For any alphabet A, A^* is star-free, since $A^* = \varnothing^c = 0^c$.
- $\forall B \subseteq A$, A^*BA^* is star-free.
- Also, B^* is star-free, since $B^* = \left(\sum_{a \in A \setminus B} A^*aA^* \right)^c$.
- $A = \{a, b\}$, then $(ab)^*$ is star-free. Since $(ab)^* = \left(b0^c + 0^c a + 0^c aa0^c + 0^c bb0^c \right)^c$

Schützenberger's theorem

Theorem :

A language is star-free iff it's syntactic monoid is aperiodic.

Proof of Schützenberger's theorem

Let A be an alphabet and define $\mathscr{A}(A)$ as the set of recognisable languages over A, whose syntactic monoids are aperiodic. Thus,

- $\varnothing, \{\epsilon\}, \{a\} \in \mathscr{A}(A)$, $\forall\ a \in A$.
- $\mathscr{A}(A)$ is closed under complementation.
- $\mathscr{A}(A)$ is closed under finite intersection, and by previous property, is closed under finite union.
- $\mathscr{A}(A)$ is closed under finite product.

Therefore, $\mathscr{A}(A)$ contains all star-free languages over A.

Proof of Schützenberger's theorem

For the converse, let $\varphi : A^* \to M$ be a monoid morphism such that, M is an aperiodic monoid.
We now claim that $\varphi^{-1}(P)$ is star-free, $\forall\ P \subseteq M$.
But since $\varphi^{-1}(P) = \sum_{m \in P} \varphi^{-1}(m)$ and $P \subseteq M$ is finite, we may assume $P = \{m\}$ without loss of generality.

Proof of Schützenberger's theorem

Claim : $\varphi^{-1}(m)$ is star-free, for all $m \in M$

proof : We use induction on $r(m) := | M \setminus MmM |$

Base Case :

if $r(m) = 0$ then $M = MmM$. Therefore $\exists\ u, v \in M$ such that $umv = 1$. Applying simplification lemma, $(um)1(v) = 1$ and $(u)1(mv) = 1 \implies u = v = 1$ and thus $m = 1$.
Now, let $B := \{a \in A : \varphi(a) = 1\}$,
then $u \in B^* \implies u \in \varphi^{-1}(1)$. Also if, $u \in \varphi^{-1}(1)$ then by simplification lemma, $\varphi(b) = 1$ for each letter b of u.
Therefore $\varphi^{-1}(m) = B^*$, which is star-free.

Proof of Schützenberger's theorem

Induction hypothesis :

Assume $r(m) > 0$ and $\varphi^{-1}(s)$ is star-free if $r(s) < r(m)$.

Induction step :

Claim :

$$\varphi^{-1}(m) = (UA^* \cap A^*V) \setminus (A^*CA^* \cup A^*WA^*) \qquad (1)$$

$U := \sum_{(n,a)\in E} \varphi^{-1}(n)a$; $V := \sum_{(a,n)\in F} a\varphi^{-1}(n)$
$C := \{a \in A : m \notin M\varphi(a)M\}$; $W := \sum_{(a,n,b)\in G} a\varphi^{-1}(n)b$

$$E := \{(n,a) \in M \times A : n\varphi(a)\mathcal{R}m \wedge n \notin mM\}$$

$$F := \{(a,n) \in A \times M : \varphi(a)n\mathcal{L}m \wedge n \notin Mm\}$$

$G := \{(a,n,b) \in A \times M \times A :$
$m \in (M\varphi(a)nM \cap Mn\varphi(b)M) \setminus M\varphi(a)n\varphi(b)M\}$

Proof of Schützenberger's theorem

Let $L := (UA^* \cap A^*V) \setminus (A^*CA^* \cup A^*WA^*)$

proof :

Let $u \in \varphi^{-1}(m)$ and let p be the shortest prefix of u such that $\varphi(p)\mathcal{R}m$.
Then $p \neq \epsilon$, otherwise $m\mathcal{R}1$, whence $m = 1$ by simplification lemma.

Put $p = ra$, with $r \in A^*$ and $a \in A$ and $n = \varphi(r)$.
By construction, $(n, a) \in E$ since

a) $n\varphi(a) = \varphi(r)\varphi(a) = \varphi(p)\ \mathcal{R}m$.
b) since $m \leqslant_R \varphi(p) = n\varphi(a) \leqslant_R n$, $n \notin mM$ otherwise $n\mathcal{R}m$.

Proof of Schützenberger's theorem

It follows that $p \in \varphi^{-1}(n)a$ and $u \in UA^*$. A symmetric argument shows $u \in A^*V$.

If $u \in A^*CA^*$, $\exists a \in C$ such that $m = \varphi(u) \in M\varphi(a)M \Rightarrow\Leftarrow a \in C$.

Similarly, if $u \in A^*WA^*$, $\exists\ (a, n, b) \in G$ such that $m \in M\varphi(a)n\varphi(b)M \Rightarrow\Leftarrow (a, n, b) \in G$.

Therefore $u \in L$.

Proof of Schützenberger's theorem

Conversely, assume $u \in L$ and $s := \varphi(u)$.
Since, $u \in UA^*$ we have $u \in \varphi^{-1}(n)aA^*$, for some $(n, a) \in E$, and hence $s = \varphi(u) \in n\varphi(a)M$.
Now, since $(n, a) \in E$, $n\varphi(a)M = mM$ and thus $s \in mM$.
A dual argument shows $u \in VA^*$ implies $s \in mM$.

By Lemma 3. to prove that $s = m$ and hence $u \in \varphi^{-1}(m)$, it suffices to prove that $s \notin J_m$ i.e $m \in MsM$

Proof of Schützenberger's theorem

On the contrary, consider a factor f of u of minimal length such that $m \notin M\varphi(f)M$. Then $f \neq \epsilon$.

If $f \in A$ then $f \in C$ and $u \in A^*CA^*$, which is impossible.

Set $f = agb$ where $a, b \in A$. Set $n = \varphi(g)$.
Since f is of minimal length, we have $m \in M\varphi(a)nM$ and $m \in Mn\varphi(b)M$.

Consequently, $(a, n, b) \in G$, and $f \in W$, which is impossible.

Equation (1) is thus established.

Proof of Schützenberger's theorem

A^*CA^* is star-free.

Let $(n, a) \in E$. Since $n\varphi(a)M = mM$, we have $MmM \subseteq MnM$, and hence $r(n) \leqslant r(m)$.

Moreover, as $m \leqslant_R n$, by Theorem 1., $MmM = MnM \implies n\mathcal{R}m \Rightarrow\Leftarrow n \notin mM$.

Therefore $r(n) < r(m)$ and thus U is star-free by *Induction hypothesis*.

A similar argument works for V.

Proof of Schützenberger's theorem

Finally, let $(a, n, b) \in G$.
One has $r(n) \leqslant r(m)$ since, $m \in MnM$.

Suppose that $MmM = MnM$. Then, $n \in MmM$
also $m \in M\varphi(a)nM$ and $m \in Mn\varphi(b)M$,

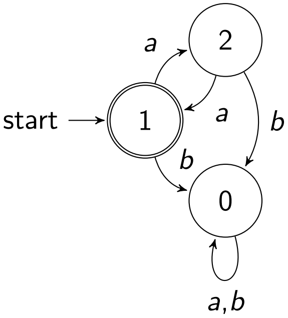
it follows $n \in M\varphi(a)nM$ and $n \in Mn\varphi(b)M$, whence
$n\mathcal{L}\varphi(a)n$ and $n\mathcal{R}n\varphi(b)$.

By Green's lemma,
$n\varphi(b)\mathcal{L}\varphi(a)n\varphi(b)$ and hence $m\mathcal{J}\varphi(a)n\varphi(b) \Rightarrow\Leftarrow$
$(a, n, b) \in G$.

Therefore $r(n) < r(m)$ and hence W is star-free by
Induction hypothesis. $\square$

Examples

Let $A = \{a, b\}$ and $L = (aa)^*$. Then L is accepted by the minimal DFA, $\mathcal{A}$



Examples

The syntactic monoid of L consists of three permutations $I = (0\ 1\ 2)$, $\alpha = (2\ 1\ 0)$ and $\beta = (0\ 0\ 0)$, defined by the relations $\alpha^2 = I$, $\beta \circ \alpha = \alpha \circ \beta = \beta$, where I is the identity.

This monoid is not aperiodic since $\forall\ n \in \mathbb{N}$, $\alpha^n \neq \alpha^{n+1}$, and so L is not star-free.

References

- M.-P. Schützenberger, On finite monoids having only trivial sub-groups,Information and Control(1965), 190–194
- M. Perles, M. Rabin and E. Shamir, The theory of definite automata,IEEE Trans. Electron. Comput.(1963), 233–243
- Straubing, Howard (1994). Finite automata, formal logic, and circuit complexity. Progress in Theoretical Computer Science. Basel: Birkhäuser. p. 79.
- Kamp, Johan Antony Willem (1968). Tense Logic and the Theory of Linear Order. University of California at Los Angeles (UCLA).