of $\bar{\mathscr{A}}$ equals $\mathscr{H}(\mathscr{A})$. If $A \in \mathbf{I}(\mathscr{A}^c)$ and $h(A) > 1$ then $\mathbf{H}(\mathscr{A}) = \mathbf{H}(B)$ where $A = B \cup \{z\}$ and $B \in \mathbf{I}(\mathscr{A})$. If $h(A) = 1$ then $\mathbf{H}(A) = \mathbf{H}(B) + 1$.

4.6 Let τ be an admissible partition on the transformation semigroup $\mathscr{A} = (Q, S)$ and $(f, g): \mathscr{A} \to \mathscr{A}/\langle\tau\rangle$ the natural epimorphism. Show that $\mathbf{I}(\mathscr{A}/\langle\tau\rangle) = f(\mathbf{I}(\mathscr{A}))$ (as sets).

4.7 Let $h: I(\mathscr{A}) \to \mathbf{Z}$ be a height function with $h(Q) = n$, and $\pi^n > \pi^{n-1} > \ldots > 1$ the derived sequence. Suppose that π^i is an orthogonal partition for some $1 < i < n$ and let $\pi^i \cap \tau = 1$ with τ an admissible partition. Let $(f, g): \mathscr{A} \to \mathscr{A}/\langle\tau\rangle$ be the natural epimorphism. Show that $f(\pi^i) \geq f(\pi^{i-1}) \geq \ldots \geq f(1)$ is a sequence of admissible subset systems in $\mathscr{A}/\langle\tau\rangle$.

4.8 With the notation of 4.7 define a function $k: \mathbf{I}(\mathscr{A}/\langle\tau\rangle) \to \mathbf{Z}$ by $k(B) = \inf\{h(A) \mid A \in \mathbf{I}(\mathscr{A}),\ f(A) = B,\ B \in \mathbf{I}(\mathscr{A}/\langle\tau\rangle)\}$. Prove that $k(B) = \inf\{j \mid B \in f(\pi^j)\}$.

5

Recognizers

We have seen how Mealy machines can be used to model the connections between inputs and outputs of complex systems, and how to decompose the underlying state machines that are central to this procedure. There is another area in which state machines play a major role. In the development of computer systems it is important to distinguish between certain *sequences* of inputs. The computer must be able to recognize those instructions that are compatible with its system and these instructions will take the form of input words from an input alphabet.

This chapter is concerned with the mathematical theory of recognizers; these are state machines that are able to *discriminate* between two disjoint sets of input words. The foundations for this theory, initially developed by S. C. Kleene in 1956, had an important influence on the construction of compilers for computers. It is also of independent mathematical interest and is closely related to the study of languages and psycholinguistics.

5.1 Automata or recognizers

Let $\mathscr{M} = (Q, \Sigma, F)$ be a state machine (as usual Q and Σ are finite and F is a partial function, $F: Q \times \Sigma \to Q$). Let $i \in Q$ be a fixed state called the *initial state* and suppose that $T \subseteq Q$ is a set of states called the *set of terminal states.*

The collection $\mathfrak{M} = (\mathscr{M}, i, T)$ is called an *automaton* or a *recognizer.* We will use the second term since automaton is often used as a generic noun to describe all types of machine. The recognizer is able to distinguish between certain types of word from the monoid Σ^*. For example, let $\alpha \in \Sigma^*$, then $iF_\alpha \in Q$ or $iF_\alpha = \varnothing$ and we say that $\mathfrak{M}$ *recognizes* α if and only if $iF_\alpha \in T$. The set Σ^* is partitioned into two disjoint subsets, the set of words recognized by $\mathfrak{M}$ and the set of words not recognized by $\mathfrak{M}$. The set of words of Σ^* recognized by $\mathfrak{M}$ is called the *behaviour* of $\mathfrak{M}$ and is denoted by $|\mathfrak{M}|$. Thus $|\mathfrak{M}| = \{\alpha \in \Sigma^* \mid iF_\alpha \in T\}$.

One major aim is the characterization of the subsets of Σ^* that can arise as the behaviour of a recognizer. We shall see that some subsets of Σ^* can never be the behaviour of a recognizer. Another fact that will soon become apparent is that different recognizers can have the same behaviour.

We need some straightforward notation for describing subsets of Σ^*. Let $A \subseteq \Sigma^*$, $B \subseteq \Sigma^*$, with $A \neq \varnothing$, $B \neq \varnothing$, we define

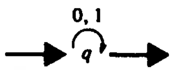
$$A \cdot B = \{\alpha \in \Sigma^* \mid \alpha = ab;\ a \in A,\ b \in B\},$$

$$A^+ = \{\alpha \in \Sigma^* \mid \alpha = a_1 \cdot \ldots \cdot a_n;\ a_i \in A,\ 1 \leq i \leq n,\ n > 0\},$$

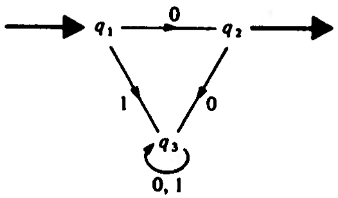
$$A^* = A^+ \cup \{\Lambda\}.$$

Examples 5.1

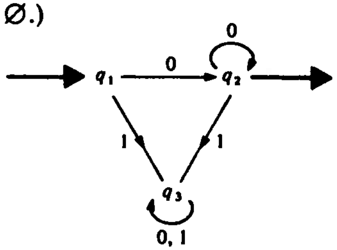
These examples will be described by using directed graphs to describe the recognizer with the initial state indicated by a bold arrow, unlabelled, and pointing towards the state. The terminal states are shown with a bold arrow, unlabelled, pointing away from the state. Let $\Sigma = \{0, 1\}$ in all of these examples.

(i) 

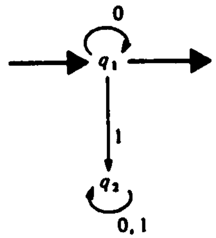
Here $Q = T = \{q\}$. The initial state is also q. Any word from Σ^* will be recognized by this machine and so $|\mathfrak{M}| = \Sigma^*$.

(ii) 

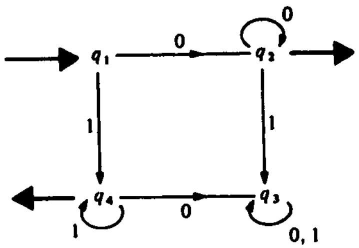
Then the behaviour, $|\mathfrak{M}|$, is $\{0\}$. (Notice that 01 is not recognized since $q_1F_{01} = \varnothing$.)

(iii) 

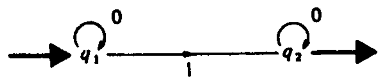
This recognizer has behaviour $\{0\}^+$ (which can be written as $\{0\}^* \cdot \{0\}$ or $\{0\} \cdot \{0\}^*$ or $\{0\}^* \backslash \{\Lambda\}$).

(iv) 

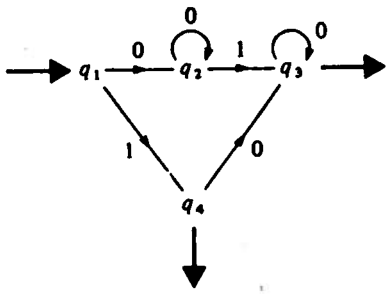
This has behaviour $\{0\}^*$.

(v) 

The behaviour of this machine is $\{0^+\} \cup \{1\}^+$.

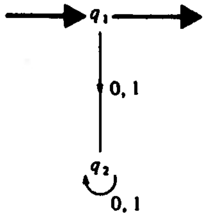
(vi) 

The behaviour is $\{0\}^* \cdot \{1\} \cdot \{0\}^*$, that is all words in Σ^* containing precisely one occurrence of 1.

(vii) 

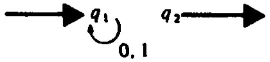
This has behaviour $\{0\}^+ \cdot \{1\} \cdot \{0\}^* \cup \{1\} \cup \{1\} \cdot \{0\}^+$ which is the set of all words of Σ^* containing precisely one occurrence of 1, that is the behaviour is equal to $\{0\}^* \cdot \{1\} \cdot \{0\}^*$ as in (vi).

(viii)



This recognizer has behaviour $\{\Lambda\}$.

(ix)



The behaviour of this recognizer is $\varnothing$.

We will meet other examples as we proceed further.

The concept of the completion of a state machine has been studied in chapter 2. We can immediately deduce the following result:

Theorem 5.1.1

Let $\mathcal{M} = (Q, \Sigma, F)$ be a state machine, $i \in Q$ and $T \subseteq Q$. Consider the completion $\mathcal{M}^c$ of $\mathcal{M}$, and let $\mathfrak{M}^c = (\mathcal{M}^c, i, T)$, then

$$|\mathfrak{M}^c| = |\mathfrak{M}|.$$

Proof We assume first that $\mathcal{M}$ is not complete and let $\mathcal{M}^c = (Q \cup \{x\}, \Sigma, F')$ where $x \notin Q$, $xF'_\sigma = x$, $qF'_\sigma = x$ if $qF_\sigma = \varnothing$ and $qF'_\sigma = qF_\sigma$ if $qF_\sigma \neq \varnothing$. We now assume that $|\mathfrak{M}| \neq \varnothing$. Let $\alpha \in |\mathfrak{M}|$, then $iF_\alpha \in T$. Since $iF_\alpha \neq \varnothing$ we may deduce that $iF'_\alpha = iF_\alpha \in T$ and so $\alpha \in |\mathfrak{M}^c|$. Now let $\alpha \in |\mathfrak{M}^c|$, then $iF'_\alpha \in T$. If $iF'_\alpha \neq iF_\alpha$ then $iF'_\beta = x$ for some $\beta \in \Sigma^*$ such that $\alpha = \beta\gamma$, $\gamma \in \Sigma^*$. But then $iF'_\alpha = xF'_\gamma = x \notin T$ and so $\alpha \notin |\mathfrak{M}^c|$. Hence $|\mathfrak{M}^c| = |\mathfrak{M}|$.

If $|\mathfrak{M}| = \varnothing$ then $iF_\alpha \notin T$ for any $\alpha \in \Sigma^*$ and so $iF'_\alpha \notin T$ for any $\alpha \in \Sigma^*$. □

The recognizer $\mathfrak{M}^c$ will be called the *completion* of $\mathfrak{M}$. It is clear from theorem 5.1.1 that we will lose very little if we concentrate our studies on complete recognizers.

Let Σ be a finite set, a subset A of Σ^* will be called *recognizable* if there exists a recognizer $\mathfrak{M}$ such that A is the behaviour of $\mathfrak{M}$, that is if $A = |\mathfrak{M}|$ for some recognizer $\mathfrak{M}$. We say that *$\mathfrak{M}$ recognizes A.*

Given a recognizable subset A it is usually possible to find many distinct recognizers that recognize A and one of our tasks in the next

section is to construct a standard complete recognizer that recognizes A and is also the 'most efficient' recognizer with this property. We will now explain the term 'most efficient'.

Intuitively a recognizer $\mathfrak{M}$ recognizing the set $A \subseteq \Sigma^*$ would be considered efficient if there were no 'wasted states'. For example, suppose that $\mathfrak{M} = (\mathcal{M}, i, T)$ where $\mathcal{M} = (Q, \Sigma, F)$ and consider the set of states

$$R = \{iF_\alpha \mid \alpha \in \Sigma^*\}.$$

R then consists of all those states of Q that can be reached from the initial state i. These are the only states that can influence the behaviour $|\mathfrak{M}| = A$ and consequently if $R \subsetneq Q$ there will be some states in Q that will never feature in our discussions about A. If $\mathfrak{M}$ has the property that $R = Q$ we will call $\mathfrak{M}$ *accessible.* Given a recognizer $\mathfrak{M} = (\mathcal{M}, i, T)$ we can remove the states in the set $Q \backslash R$, and obtain an accessible recognizer which clearly has the same behaviour as $\mathfrak{M}$. This is called the *accessible part*, $\mathfrak{M}^a$, of $\mathfrak{M}$. Thus

$$\mathfrak{M}^a = (\mathcal{M}^a, i, T)$$

where $\mathcal{M}^a = (R, \Sigma, F^a)$, $R = \{iF_\alpha \mid \alpha \in \Sigma^*\}$, and $qF^a_\alpha = qF_\alpha$ for $q \in R$, $\alpha \in \Sigma^*$. Note that $|\mathfrak{M}^a| = |\mathfrak{M}|$. It is clear that if $\mathfrak{M}$ is complete then $\mathfrak{M}^a$ is also complete.

Another way in which states may be redundant is if there are states in the recognizer that never lead to a terminal state. Thus if $q \in Q$ and $qF_\alpha \notin T$ for all $\alpha \in \Sigma^*$ then q can never lie on a successful 'route' from the initial state i to a final state in T. Consider the set S, of all states that can lead to a terminal state, so that

$$S = \{q \mid qF_\alpha \in T \quad \text{for some } \alpha \in \Sigma^*\}.$$

If $S = Q$ we call $\mathfrak{M}$ *coaccessible.* The *coaccessible* part of a recognizer $\mathfrak{M} = (\mathcal{M}, i, T)$ is defined to be $\mathfrak{M}^b = (\mathcal{M}^b, i, T)$ where

$$\mathcal{M}^b = (S, \Sigma, F^b),$$
$$S = \{q \mid qF_\alpha \in T \text{ for some } \alpha \in \Sigma^*\}$$

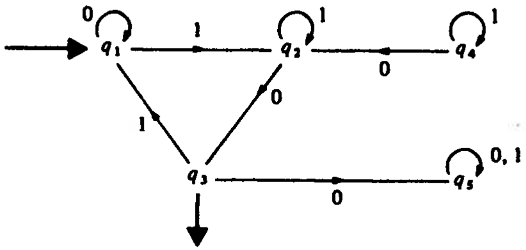
and

$$qF^b_\alpha = qF_\alpha \text{ for } q \in S,\ \alpha \in \Sigma^*.$$

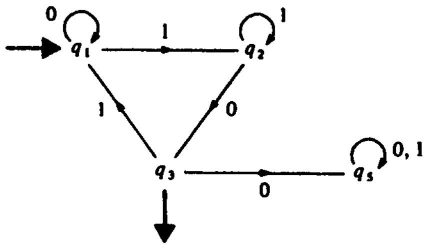
Clearly $|\mathfrak{M}^b| = |\mathfrak{M}|$. A recognizer $\mathfrak{M} = (\mathcal{M}, i, T)$ is called *trim* if it is both accessible and coaccessible.

Example 5.2

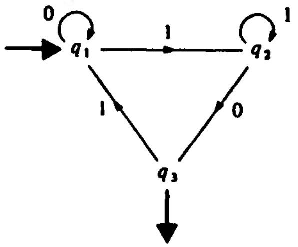
Let $\Sigma = \{0, 1\}$, $Q = \{q_1, q_2, q_3, q_4, q_5\}$.



This defines a complete recognizer $\mathfrak{M} = (\mathcal{M}, q_1, \{q_3\})$. Then $\mathfrak{M}^a$ is given by



$(\mathfrak{M}^a)^b$ is given by



and this is a trim recognizer which satisfies $|(\mathfrak{M}^a)^b| = |\mathfrak{M}|$, but is no longer complete.

We could equally well have constructed $(\mathfrak{M}^b)^a$ and this would have produced the same machine.

Our final task for this section is the introduction of some useful notation.

Let $\mathcal{M} = (Q, \Sigma, F)$ be a complete state machine and suppose that $q \in Q$, $\alpha \in \Sigma^*$; define $q * \alpha = qF_\alpha$, and then for $A \subseteq \Sigma^*$, $S \subseteq Q$ we have:

$$q * A = \{q * \alpha \mid \alpha \in A\}$$
$$S * \alpha = \{q * \alpha \mid q \in S\}$$
$$S * A = \{q * \alpha \mid q \in S, \alpha \in A\}.$$

$$q * \alpha^{-1} = \{p \in Q \mid q = p * \alpha\}$$
$$q * A^{-1} = \{p \in Q \mid q = p * \alpha \text{ for some } \alpha \in A\}$$
$$S * \alpha^{-1} = \{p \in Q \mid p * \alpha \in S\}$$
$$S * A^{-1} = \{p \in Q \mid p * \alpha \in S \text{ for some } \alpha \in A\}.$$

If $A \subseteq \Sigma^*$ and $B \subseteq \Sigma^*$, $a \in A$ and $b \in B$ then define

$$a \cdot b^{-1} = \{\alpha \in \Sigma^* \mid \alpha b = a\}$$
$$a^{-1} \cdot b = \{\alpha \in \Sigma^* \mid a\alpha = b\}$$
$$a^{-1} \cdot B = \{\alpha \in \Sigma^* \mid a\alpha \in B\}$$
$$a \cdot B^{-1} = \{\alpha \in \Sigma^* \mid \alpha b = a \text{ for some } b \in B\}$$
$$A \cdot b^{-1} = \{\alpha \in \Sigma^* \mid \alpha b \in A\}$$
$$A^{-1} \cdot b = \{\alpha \in \Sigma^* \mid a\alpha = b \text{ for some } a \in A\}$$
$$A \cdot B^{-1} = \{\alpha \in \Sigma^* \mid \alpha b \in A \text{ for some } b \in B\}$$
$$A^{-1} \cdot B = \{\alpha \in \Sigma^* \mid a\alpha \in B \text{ for some } a \in A\}.$$

With $\mathcal{M} = (Q, \Sigma, F)$ and $p, q \in Q$ we put $q^{-1} \circ p = \{\alpha \in \Sigma^* \mid p = q * \alpha\}$, that is the set of words that 'send q to p'.

If $R, S \subseteq Q$ then we let

$$q^{-1} \circ R = \{\alpha \in \Sigma^* \mid q * \alpha \in R\}$$
$$S^{-1} \circ R = \{\alpha \in \Sigma^* \mid q * \alpha \in R \text{ for some } q \in S\}.$$

Some elementary results can now be stated, their proof will be left as exercises. Some useful identities are to be found in exercise 5.8.

Proposition 5.1.2

Let $\mathcal{M} = (Q, \Sigma, F)$ be a state machine, $A, B, C \subseteq \Sigma^*$ and $S \subseteq Q$.

(i) $(S * A) * B = S * (A \cdot B)$

(ii) $(S * A^{-1}) * B^{-1} = S * (B \cdot A)^{-1}$

Proposition 5.1.3

Let $\mathcal{M} = (Q, \Sigma, F)$ and $\mathfrak{M} = (\mathcal{M}, i, T)$ and $A = |\mathfrak{M}|$ then

$A = i^{-1} \circ T$.

If $q = i * \alpha$, $\alpha \in \Sigma^*$, then

$q^{-1} \circ T = \alpha^{-1} A$.

Proof Recall that $i^{-1} \circ T = \{\alpha \in \Sigma^* \mid i * \alpha \in T\}$

$= \{\alpha \in \Sigma^* \mid iF_\alpha \in T\}$

$= A.$

Now let $q = i * \alpha$, then

$$\begin{aligned} q^{-1} \circ T &= \{\beta \in \Sigma^+ \mid q * \beta \in T\} \\ &= \{\beta \in \Sigma^* \mid (i * \alpha) * \beta \in T\} \\ &= \{\beta \in \Sigma^* \mid iF_{\alpha\beta} \in T\} \\ &= \{\beta \in \Sigma^* \mid \alpha\beta \in A\} \\ &= \alpha^{-1}A. \end{aligned}$$ □

Proposition 5.1.4

Let $\mathscr{M} = (Q, \Sigma, F)$ and $\mathfrak{M} = (\mathscr{M}, i, T)$, then $\mathfrak{M}$ is accessible if and only if $Q = i * (\Sigma^*)$ and $\mathfrak{M}$ is coaccessible if and only if $Q = T * (\Sigma^*)^{-1}$.

Proof This follows from the definitions since

$$R = \{iF_\alpha \mid \alpha \in \Sigma^*\} = \{i * \alpha \mid \alpha \in \Sigma^*\} = i * (\Sigma^*)$$

and

$$\begin{aligned} S &= \{q \mid qF_\alpha \in T \text{ for some } \alpha \in \Sigma^*\} \\ &= \{q \mid q * \alpha \in T \text{ for some } \alpha \in \Sigma^*\} \\ &= T * (\Sigma^*)^{-1}. \end{aligned}$$ □

5.2 Minimal recognizers

Let Σ be a finite set and $A \subseteq \Sigma^*$. If A is recognizable then there exists a recognizer

$$\mathfrak{M} = (\mathscr{M}, i, T) \quad \text{where } \mathscr{M} = (Q, \Sigma, F) \quad \text{and} \quad A = |\mathfrak{M}|.$$

We shall now construct a recognizer with behaviour equal to A directly.

Let us consider all subsets of Σ^* of the form

$$\alpha^{-1} \cdot A = \{\beta \in \Sigma^* \mid \alpha\beta \in A\},$$

where $\alpha \in \Sigma^*$. Put Q_A to be the *set* of all such subsets, noting that this may include the empty set, $\varnothing$.

Thus $Q_A = \{\alpha^{-1} \cdot A \mid \alpha \in \Sigma^*\}$ and clearly $A \in Q_A$ since $A = \Lambda^{-1} \cdot A$. The state function $F^A : Q_A \times \Sigma \to Q_A$ is defined by

$$\left.\begin{aligned} (\alpha^{-1} \cdot A)F^A_\sigma &= (\alpha\sigma)^{-1} \cdot A \\ \varnothing F^A &= \varnothing \end{aligned}\right\} \text{for } \sigma \in \Sigma.$$

Put $i_A = A$ and define $T_A = \{\alpha^{-1} \cdot A \in Q_A \mid \alpha \in A\}$. (Note that $\alpha^{-1} \cdot A \in T_A \Leftrightarrow \Lambda \in \alpha^{-1} \cdot A$.) This defines a state machine

$$\mathscr{M}_A = (Q_A, \Sigma, F^A)$$

and a recognizer

$$\mathfrak{M}_A = (\mathscr{M}_A, i_A, T_A)$$

once we have established that $F^A : Q_A \times \Sigma \to Q_A$ is a well-defined mapping and Q_A is a finite set.

Note that if $A = \varnothing$ then $Q_A = \{\varnothing\}$, $i_A = \varnothing$ and $T_A = \varnothing$ (that is there are no final states).

Theorem 5.2.1

If $A \subseteq \Sigma^*$ is recognizable then $\mathfrak{M}_A$ is a recognizer with the property that $|\mathfrak{M}_A| = A$.

Proof Let $\alpha^{-1} \cdot A$, $\gamma^{-1} \cdot A \in Q_A$ with $\alpha^{-1} \cdot A = \gamma^{-1} \cdot A \neq \varnothing$. If $\sigma \in \Sigma$ then

$$\begin{aligned} (\alpha^{-1} \cdot A)F^A_\sigma &= (\alpha\sigma)^{-1} \cdot A \\ &= \sigma^{-1} \cdot (\alpha^{-1}A) \text{ by exercise 5.8.} \\ &= \sigma^{-1} \cdot (\gamma^{-1} \cdot A) \\ &= (\gamma\sigma)^{-1} \cdot A \\ &= (\gamma^{-1} \cdot A)F^A_\sigma \end{aligned}$$

and so F^A is a well-defined function.

Next we show that Q_A is finite. Let $\alpha^{-1} \cdot A \in Q_A$ and put $q = i\alpha$ where $\mathfrak{M} = (\mathscr{M}, i, T)$ is a recognizer that recognizes A. Now

$$\begin{aligned} \alpha^{-1} \cdot A &= \{\beta \in \Sigma^* \mid \alpha\beta \in A\} \\ &= \{\beta \in \Sigma^* \mid i\alpha\beta \in T\} \\ &= \{\beta \in \Sigma^* \mid q\beta \in T\} \\ &= q^{-1} \circ T. \end{aligned}$$

Since Q is finite there can only be a finite number of sets of this form and so Q_A is finite.

If $a \in A$ then $a^{-1}A \in T_A$ and so $AF^A_a \in T_A$ which means that $a \in |\mathfrak{M}_A|$. Now let $x \in |\mathfrak{M}_A|$, then $AF^A_x \in T_A$ which gives $x^{-1} \cdot A \in T_A$. Suppose that $x^{-1} \cdot A = a^{-1} \cdot A$ where $a \in A$, then $a\Lambda = a \in A$ and so $\Lambda \in a^{-1} \cdot A = x^{-1} \cdot A$.

Therefore $x\Lambda \in A$ and so $x \in A$, proving that $|\mathfrak{M}_A| \subseteq A$. Consequently

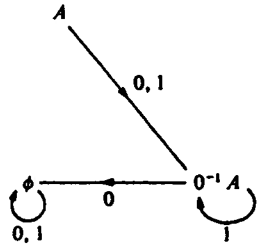
$$|\mathfrak{M}_A| = A.$$ □

Note that $\mathfrak{M}_A$ is complete and accessible, but will not be coaccessible if $\varnothing \in Q_A$.

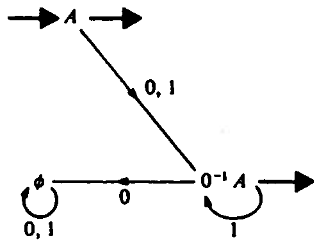
Examples 5.3

Let $\Sigma = \{0, 1\}$

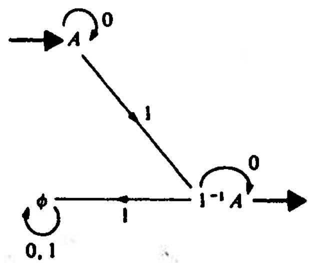
(i) $A = \{0\} \cdot \{1\}^* \cup \{1\}^*$, $0^{-1}A = \{1\}^* = 1^{-1}A$, $(0^2)^{-1}A = \varnothing$, etc. and $Q_A = \{A, 0^{-1}A, \varnothing\}$ with state function:



$T_A = \{A, 0^{-1}A\}$ and so the complete recognizer $\mathfrak{M}_A$ is:



(ii) $A = \{0\}^* \cdot \{1\} \cdot \{0\}^*$, $1^{-1}A = \{0\}^*$, $Q_A = \{A, \{0\}^*, \varnothing\}$, $T_A = \{1^{-1}A\}$.



(the completion of example 5.1 (vi))

The recognizer $\mathfrak{M}_A$ has the following minimality property.

Theorem 5.2.2

Let $A \subseteq \Sigma^*$ be recognizable and suppose that $\mathfrak{M} = (\mathcal{M}, i, T)$, where $\mathcal{M} = (Q, \Sigma, F)$, is a complete accessible recognizer with behaviour A. There exists a function $f: Q \to Q_A$ such that:

(i) $f(i) = i_A$,

(ii) $f^{-1}(T_A) = T$,

(iii) $(f(q))F_\sigma^A = f(qF_\sigma)$ for all $q \in Q$, $\sigma \in \Sigma$,

(iv) f is surjective.

Proof Define $f: Q \to Q_A$ by

$$f(q) = q^{-1} \circ T = \{\alpha \in \Sigma^* \mid q * \alpha \in T\}.$$

We must first show that $f(q) \in Q_A$. Since Q is accessible there exists $\beta \in \Sigma^*$ such that $q = iF_\beta = i * \beta$. Then

$$\begin{aligned}\beta^{-1} \cdot A &= \{\gamma \in \Sigma^* \mid \beta\gamma \in A\}\\ &= \{\gamma \in \Sigma^* \mid i * (\beta\gamma) \in T\}\\ &= \{\gamma \in \Sigma^* \mid (i * \beta) * \gamma \in T\}\\ &= \{\gamma \in \Sigma^* \mid q * \gamma \in T\}\\ &= q^{-1} \circ T.\end{aligned}$$

Thus f is a function. Then

(i) $f(i) = i^{-1} \circ T = \{\alpha \in \Sigma^* \mid i * \alpha \in T\} = A = i_A$.

(ii) Let $f(q) \in T_A$, then $q^{-1} \circ T = a^{-1}A$ for some $a \in A$ and so $\Lambda \in q^{-1} \circ T$, that is $q * \Lambda \in T$ and so $q \in T$. Hence $f^{-1}(T_A) \subseteq T$. Now for $t \in T$ we have $f(t) = t^{-1} \circ T$ and since $\Lambda \in t^{-1} \circ T$ we see that $t^{-1} \circ T \in T_A$. Thus $f(T) \subseteq T_A$ and so $f^{-1}(T_A) = T$.

(iii)
$$\begin{aligned}(f(q))F_\sigma^A &= (q^{-1}T)F_\sigma^A\\ &= (\beta^{-1}A)F_\sigma^A \quad \text{if } q = i * \beta\\ &= (\beta\sigma)^{-1}A\\ &= \{\gamma \in \Sigma^* \mid \beta\sigma\gamma \in A\}\\ &= \{\gamma \in \Sigma^* \mid i\beta\sigma\gamma \in T\}\\ &= \{\gamma \in \Sigma^* \mid q\sigma\gamma \in T\}\\ &= (q\sigma)^{-1} \circ T = f(qF_\sigma) \quad \text{for } q \in Q, \sigma \in \Sigma.\end{aligned}$$

(iv) Let $s^{-1}A \in Q_A$ where $s \in \Sigma^*$, then put $p = is \in Q$ and note that $p^{-1} \circ T = s^{-1}A$ and so $s^{-1}A = f(p)$.

Therefore f is surjective. □

We can now regard the recognizer $\mathfrak{M}_A$ as being the *minimal complete recognizer* of the recognizable subset A, where the term 'minimal' refers to the properties described in theorem 5.2.2, in particular (iv) implies that $|Q_A| \leq |Q|$. If we try to construct the recognizer $\mathfrak{M}_A$ in the case where A is not recognizable we will find that the set of states Q_A is no longer finite and so $\mathfrak{M}_A$ will not then be a recognizer according to our definition. This is examined in the next theorem.

Theorem 5.2.3

Let $A \subseteq \Sigma^*$, then A is recognizable if and only if the collection $\{\beta^{-1}A \mid \beta \in \Sigma^*\}$ is finite.

Proof If A is recognizable then the proof of 5.2.1 establishes that Q_A is finite and so $\{\beta^{-1}A \mid \beta \in \Sigma^*\}$ is finite. Clearly if $\{\beta^{-1}A \mid \beta \in \Sigma^*\}$ is finite then we may construct the recognizer $\mathfrak{M}_A$ which will then establish the fact that A is recognizable. □

5.3 Recognizable sets

The examples of recognizable sets that we have already seen will now be augmented by developing general techniques for constructing more recognizable sets from given recognizable sets.

Notice first that the following are examples of recognizable sets where Σ is a given finite set and $\sigma \in \Sigma$.

$$\{\sigma\}, \{\Lambda\}, \varnothing, \Sigma^*.$$

Now suppose that A, B are recognizable subsets of Σ^*. We will show that $A \cup B$, $A \cdot B$, A^*, $\Sigma^* \backslash A$, $A \cap B$, A^+ are also recognizable. The basic technique is the same in all cases, namely that we construct a recognizer with the desired property using recognizers of A and B. The machines so formed may not be minimal in the sense of 5.2.2. but that is irrelevant here.

Theorem 5.3.1

Let $A, B \subseteq \Sigma^*$. If A and B are recognizable then $A \cup B$ is also recognizable.

Proof Let $\mathfrak{M} = (\mathcal{M}, i, T)$, $\mathfrak{M}' = (\mathcal{M}', i', T')$ be recognizers with $\mathcal{M} = (Q, \Sigma, F)$, $\mathcal{M}' = (Q', \Sigma, F')$ and such that $|\mathfrak{M}| = A$, $|\mathfrak{M}'| = B$. Consider $\mathcal{M} \vee \mathcal{M}' = (Q \times Q', \Sigma, \bar{F})$ where $(q, q')\bar{F}_\sigma = (qF_\sigma, q'F'_\sigma)$ for $\sigma \in \Sigma$, $q \in Q$, $q' \in Q'$.

Let $\mathfrak{M} \vee \mathfrak{M}' = (\mathcal{M} \vee \mathcal{M}', (i, i'), (T \times Q') \cup (Q \times T'))$. We show that $|\mathfrak{M} \vee \mathfrak{M}'| = A \cup B$. Let $\gamma \in |\mathfrak{M} \vee \mathfrak{M}'|$, then $(i, i')\bar{F}_\gamma \in (T \times Q') \cup (Q \times T')$ so $(iF_\gamma, i'F'_\gamma) \in (T \times Q') \cup (Q \times T')$ and either $iF_\gamma \in T$ or $i'F'_\gamma \in T'$, that is either $\gamma \in A$ or $\gamma \in B$, and so $\gamma \in A \cup B$.

Now let $\gamma \in A \cup B$, then either $qF_\gamma \in T$ or $q'F'_\gamma \in T'$. If $qF_\gamma \in T$ then $(q, q')\bar{F}_\gamma = (qF_\gamma, q'F'_\gamma) \in T \times Q'$ and if $q'F'_\gamma \in T'$ then $(q, q')\bar{F}_\gamma = (qF_\gamma, q'F'_\gamma) \in Q \times T'$ and in either case $\gamma \in |\mathfrak{M} \vee \mathfrak{M}'|$. □

Theorem 5.3.2

Let $A, B \subseteq \Sigma^*$. If A and B are recognizable sets then $A \cdot B$ is also recognizable.

Proof Let $\mathfrak{M} = (\mathcal{M}, i, T)$, $\mathfrak{M}' = (\mathcal{M}', i', T')$ be recognizers with $\mathcal{M} = (Q, \Sigma, F)$, $\mathcal{M}' = (Q', \Sigma, F')$ and such that $|\mathfrak{M}| = A$, $|\mathfrak{M}'| = B$.

Consider $\mathcal{M} \triangle \mathcal{M}' = (Q \times \mathcal{P}(Q'), \Sigma, F^\triangle)$ where

$$(q, P)F_\sigma = \begin{cases} (qf_\sigma, PF_\sigma) & \text{if } qF_\sigma \notin T \\ (qF_\sigma, PF_\sigma \cup \{i'\}) & \text{if } qF_\sigma \in T \end{cases}$$

for $q \in Q$, $P \in \mathcal{P}(Q')$, $\sigma \in \Sigma$ (here $PF_\sigma = \{pF_\sigma \mid p \in P\}$).

Now put $T^\triangle = \{(q, P) \mid q \in Q, P \in \mathcal{P}(Q'), P \cap T' \neq \varnothing\}$ and examine the recognizer.

$$\mathfrak{M} \triangle \mathfrak{M}' = (\mathcal{M} \triangle \mathcal{M}', (i, \varnothing), T^\triangle).$$

Let $\alpha \in A \cdot B$, then $\alpha = a \cdot b$ for some $a \in A$, $b \in B$ and $iF_a \in T$. Now

$$\begin{aligned} (i, \varnothing)F^\triangle_\alpha &= (i, \varnothing)F^\triangle_{ab} \\ &= (iF_a, i')F^\triangle_b \\ &= (iF_{ab}, P) \quad \text{where } i'F_b \in P, P \in \mathcal{P}(Q') \\ &\in T^\triangle. \end{aligned}$$

Hence $A \cdot B \subseteq |\mathfrak{M} \triangle \mathfrak{M}'|$.

Now let $\beta \in |\mathfrak{M} \triangle \mathfrak{M}'|$, then $(i, \varnothing)F^\triangle_\beta \in T^\triangle$, so that $(iF_\beta, P) \in T^\triangle$ for some $P \in \mathcal{P}(Q)$. Clearly $P \neq \varnothing$ and so there exists $\gamma, \delta \in \Sigma^*$ such that $\beta = \gamma \cdot \delta$ and $iF_\gamma \in T$. We call γ an initial segment of β and note that $\gamma \in A$. Let $C = \{\gamma \in \Sigma^* \mid \beta = \gamma \cdot \delta$ for some $\delta \in \Sigma^*$ and $\gamma \in A\} = \beta \cdot (\Sigma^*)^{-1}$, then we have seen that C is not empty. Each initial segment γ in C defines an 'end segment' δ such that $\beta = \gamma \cdot \delta$. Let $R = \{i'F'_\delta \mid \delta \in \Sigma^*$ and $\beta = \gamma \cdot \delta$ for some $\gamma \in C\}$, then $R \cap T' \neq \varnothing$ otherwise β would not be recognized. Let $q' \in R \cap T'$ be such that $q' = i'F'_{\delta_0}$ and suppose that $\beta = \gamma_0 \cdot \delta_0$ with $\gamma_0 \in C$. Then $\gamma_0 \in A$ and $\delta_0 \in B$, hence $\beta \in A \cdot B$ as required. □

Theorem 5.3.3

Let $A \subseteq \Sigma^*$. If A is recognizable then so is A^*.

Proof Let $\mathfrak{M} = (\mathcal{M}, i, T)$ where $\mathcal{M} = (Q, \Sigma, F)$ and $|\mathfrak{M}| = A$. Define $\mathcal{M}^* = (\mathcal{P}(Q), \Sigma, F^*)$ where

$$PF^*_\sigma = \begin{cases} PF_\sigma & \text{if } PF_\sigma \cap T = \varnothing \\ PF_\sigma \cup \{i\} & \text{otherwise} \end{cases}$$

for $P \in \mathcal{P}(Q)$, $\sigma \in \Sigma$.

Let $T^* = \{P \in \mathcal{P}(Q) \mid P \cap T \neq \varnothing\}$ and put $\mathfrak{M}^* = (\mathcal{M}^*, \{i\}, T^*)$. If $\alpha \in A^*$ then $\alpha = a_1 \ldots a_n$ for some $n \in \mathcal{N}$ and $a_i \in A$, $1 \le i \le n$. Now

$$\begin{aligned}\{i\}F^*_\alpha &= \{i\}F^*_{a_1\ldots a_n}\\ &= (\{iF_{a_1}\}\cup\{i\})F^*_{a_2\ldots a_n}\\ &= (\{iF_{a_1a_2}, iF_{a_2}\}\cup\{i\})F^*_{a_3\ldots a_n}\\ &\ \vdots\\ &= \{iF_\alpha, iF_{a_2\ldots a_n}, \ldots, iF_{a_n}, i\}\in T^*\end{aligned}$$

since $iF_{a_n} \in T$. Therefore $A^* \subseteq |\mathfrak{M}^*|$. The inequality $|\mathfrak{M}^*| \subseteq A$ will be left as an exercise. □

Theorem 5.3.4

Let $A \subseteq \Sigma^*$ be a recognizable set, then $\Sigma^*\backslash A$ is also recognizable.

Proof If $\mathfrak{M} = (\mathcal{M}, i, T)$ where $\mathcal{M} = (Q, \Sigma, F)$ is such that $|\mathfrak{M}| = A$ then $\bar{\mathfrak{M}} = (\mathcal{M}, i, Q\backslash T)$ is such that $|\bar{\mathfrak{M}}| = \Sigma^*\backslash A$. □

Theorem 5.3.5

If A and B are recognizable subsets of Σ^* then so is $A \cap B$.

Proof See exercises. □

So far we have established that there are a considerable number of recognizable sets but we have yet to meet a subset of Σ^* that is not recognizable. We will, shortly, develop techniques for testing the recognizability of certain subsets of Σ^*, but in the meantime we will briefly examine a subset which is not recognizable.

Example 5.4

Let $\Sigma = \{0, 1\}$ and put $A = \{0^n1^n \mid n \in N\}$. Suppose that $\mathfrak{M} = (\mathcal{M}, i, T)$ is such that $A = |\mathfrak{M}|$. If $\mathcal{M} = (Q, \Sigma, F)$, as usual, then $iF_{0^n1^n} \in T$ for each $n \in N$. Let $q_n = iF_{0^n}$ and suppose that $q_n = q_m$ where $m \in \mathcal{N}$, then $iF_{0^m1^n} = q_nF_{1^n} = iF_{0^n1^n} \in T$ and so $0^m1^n \in A$. Therefore $0^m1^n = 0^n1^n$ which implies $m = n$. Consequently the set of states $q_1, q_2, \ldots, q_n, \ldots$ is infinite and $\mathfrak{M}$ cannot then be a recognizer. Hence we have a contradiction to A being recognizable.

Theorem 5.3.6

Let Σ, Γ be finite non-empty sets and $f : \Sigma^* \to \Gamma^*$ a function satisfying the condition $f^{-1}(\Lambda_\Gamma) = \Lambda_\Sigma$ where Λ_Γ and Λ_Σ are the empty

words in Γ^* and Σ^* respectively. If $A \subseteq \Sigma^*$ is recognizable then so is $f(A)$.

Proof This is left as an exercise. □

5.4 The syntactic monoid

Suppose that $\mathcal{M} = (Q, \Sigma, F)$ is a state machine and consider the relation $\sim_\mathcal{M}$ defined on Σ^* by

$$\alpha \sim_\mathcal{M} \beta \Leftrightarrow F_\alpha = F_\beta$$

where $\alpha, \beta \in \Sigma^*$. We can immediately deduce the following proposition (cf. section 2.2).

Proposition 5.4.1

If $\mathcal{M} = (Q, \Sigma, F)$ is a state machine then $\sim_\mathcal{M}$ is a congruence on Σ^*.

Proof Clearly $\alpha \sim_\mathcal{M} \beta \Leftrightarrow x\alpha y \sim_\mathcal{M} x\beta y$ for all $x, y \in \Sigma^*$. □

If $\mathfrak{M} = (\mathcal{M}, i, T)$ is a recognizer with $\mathcal{M} = (Q, \Sigma, F)$ we note that if $\alpha \sim \beta$ then for $x, y \in \Sigma^*$ *either* $x\alpha y$ and $x\beta y$ both belong to $|\mathfrak{M}|$ *or* $x\alpha y$ and $x\beta y$ both do not belong to $|\mathfrak{M}|$, thus

$$\alpha \sim_\mathcal{M} \beta \Leftrightarrow [x\alpha y \in |\mathfrak{M}| \Leftrightarrow x\beta y \in |\mathfrak{M}|, \text{ for all } x, y \in \Sigma^*].$$

It is now possible to define a relation on Σ^* based on any given subset $A \subseteq \Sigma^*$; we put

$$\alpha \approx_A \beta \Leftrightarrow [x\alpha y \in A \Leftrightarrow x\beta y \in A \text{ for all } x, y \in \Sigma^*].$$

Then we have seen that $\alpha \sim_\mathcal{M} \beta \Leftrightarrow \alpha \approx_{|\mathfrak{M}|} \beta$. Can we obtain a closer connection between these two relations? In general $\mathcal{M}$ may have too many equivalent functions for the relations to be identical. We can, however, replace $\mathcal{M}$ by a more efficient machine, namely the minimal complete recognizer of $|\mathfrak{M}|$.

Theorem 5.4.2

Let $A \subseteq \Sigma^*$ be a recognizable subset of Σ^* with minimal complete recognizer $\mathfrak{M}_A = (\mathcal{M}_A, i_A, T_A)$. Then for $\alpha, \beta \in \Sigma^*$ we have

$$\alpha \sim_{\mathcal{M}_A} \beta \Leftrightarrow \alpha \approx_A \beta.$$

Proof Since $A = |\mathfrak{M}_A|$ we already have $\alpha \sim_{\mathcal{M}_A} \beta \Rightarrow \alpha \approx_A \beta$. Let $\alpha \approx_A \beta$, then $x\alpha y \in A \Leftrightarrow x\beta y \in A$ for all $x, y \in \Sigma^*$. Hence $y \in (x\alpha)^{-1} \cdot A \Leftrightarrow$

$y \in (x\beta)^{-1} \cdot A$ for all $y \in \Sigma^*$. Thus $(x\alpha)^{-1} \cdot A = (x\beta)^{-1} \cdot A$ and so

$$(x^{-1} \cdot A)F_\alpha^A = (x^{-1} \cdot A)F_\beta^A \quad \text{for all } x \in \Sigma^*$$

which means that $F_\alpha^A = F_\beta^A$ and so

$$\alpha \sim_{\mathcal{M}_A} \beta. \qquad \square$$

For a given recognizable set $A \subseteq \Sigma^*$ the congruence $\approx_A$ is called the *Myhill congruence of A*. If this congruence is factored out of the monoid Σ^* we obtain the *syntactic monoid* of A, this is given by $\Sigma^*/\approx_A = \Sigma^*/\sim_{\mathcal{M}_A} \cong \mathrm{M}(\mathcal{M}_A)$, the monoid of the minimal complete state machine $\mathcal{M}_A$. See chapter 2.

Since the congruence $\approx_A$ can be defined with respect to any subset $A \subseteq \Sigma^*$ it is of interest to see what happens when A is not recognizable. This is explored in the next result.

Theorem 5.4.3

Let $A \subseteq \Sigma^*$. The following statements are equivalent:

(i) A is recognizable.

(ii) $\Sigma^*/\approx_A$ is finite.

(iii) A is the union of congruence classes of a congruence on Σ^* of finite index.

Proof (i)$\Rightarrow$(ii). A is recognizable implies that a minimal complete recognizer $\mathfrak{M}_A$ exists and $\mathrm{M}(\mathcal{M}_A)$ is finite so that $\Sigma^*/\approx_A$ is also finite.

(ii)$\Rightarrow$(iii). If $\Sigma^*/\approx_A$ is finite then the congruence $\approx_A$ on Σ^* is of finite index. Let $[\alpha]$ denote the congruence class containing α where $\alpha \in \Sigma^*$. Now put

$$\begin{aligned} B &= \cup\{[\alpha] \mid i_A F_\alpha^A \in T_A\} \\ &= \cup\{[\alpha] \mid \alpha^{-1} \cdot A = a^{-1} \cdot A \text{ for some } a \in A\}. \end{aligned}$$

Clearly each $\alpha \in A$ since $i_A F_\alpha^A \in T_A$ and so $B \subseteq A$. Now let $a \in A$, then $i_A F_a^A \in T_A$ and so $[a] \subseteq B$. Hence $B = A$.

(iii)$\Rightarrow$(i). Suppose that $\sim$ is a congruence of finite index on Σ^* and let $A = \cup\{[\alpha_i] \mid i = 1, \ldots, n\}$ where $\alpha_i \in \Sigma^*$ and $[\alpha_i]$ is the $\sim$-congruence class containing α_i.

Let $\mathcal{M} = (Q, \Sigma, F)$ where $Q = \Sigma^*/\sim$, $F: Q \times \Sigma \to Q$ is defined by

$$[\alpha]F_\sigma = [\alpha\sigma] \quad \text{for } [\alpha] \in \Sigma^*/\sim, \sigma \in \Sigma.$$

Put $i = [\Lambda]$ and $T = \{[\alpha] \mid \alpha \in A\}$. $\mathfrak{M} = (\mathcal{M}, i, T)$ is a recognizer since Q is finite. Let $a \in A$, then

$$iF_a = [\Lambda] \cdot F_a = [\Lambda a] = [a] \in T,$$

hence $A \subseteq |\mathfrak{M}|$. If $b \in |\mathfrak{M}|$ then $iF_b \in T$ so $[\Lambda]F_b = [b] \in T$ and $b \in A$. Hence $|\mathfrak{M}| = A$ and A is thus recognizable. $\square$

The criterion (ii) can often be used to establish that a particular set is not recognizable since it means that the Myhill congruence is then of infinite index.

Example 5.5

Let $\Sigma = \{0, 1\}$ and put

$$A = \{0^n 1 0^n \mid n \in \mathcal{N}\}.$$

Consider the Myhill congruence $\approx_A$ defined by A. The infinite sequence of elements $0, 0^2, \ldots, 0^m, \ldots$ must all belong to different congruence classes for if $0^p \approx_A 0^q$ then

$$x0^p y \in A \Leftrightarrow x0^q y \in A \quad \text{for all } x, y \in \Sigma^*$$

and in particular, if we assume that $p > q$, then put $x = \Lambda$, $y = 10^q$ we get

$$0^{p-q}0^q 10^q \in A \Leftrightarrow 0^q 10^q \in A,$$

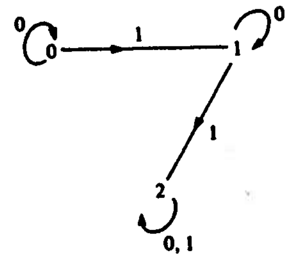
that is

$$0^{p-q}0^q 10^q \in A,$$

which is false. Thus $p = q$ and so $\approx_A$ is not of finite index and A cannot be recognizable.

Example 5.6

Consider the recognizer of example 5.3(ii) where $\Sigma = \{0, 1\}$ and $A = \{0\}^* \cdot \{1\} \cdot \{0\}^*$. The state machine is isomorphic to



We will calculate the monoid of this machine, it is generated by $\{\Lambda, 1, 1^2\}$

with the table

	Λ	1	1^2
Λ	Λ	1	1^2
1	1	1^2	1^2
1^2	1^2	1^2	1^2

This monoid is the syntactic monoid of A. Notice that 1 is not $\approx_A$-related to 1^2 since

$$\Lambda 1 \Lambda \in A \text{ but } \Lambda 1^2 \Lambda \notin A,$$

similarly 0 is not $\approx_A$-related to 1 since $\Lambda 0 \Lambda \notin A$. There are three distinct $\approx_A$-classes and $A = [1]$.

5.5 Rational decompositions of recognizable sets

In this section we examine one of two methods of decomposing a recognizable set. This first method is the classical approach of Kleene and gives a constructive characterization of a recognizable set.

It will be recalled that any singleton word from Σ^* is recognizable, as indeed is the empty set of words. Furthermore if A and B are recognizable subsets of Σ^* then so are $A \cup B$, $A \cdot B$ and A^*. Consequently we can start with a finite collection of singleton sets of words, apply the operations of union, 'dot' product and the star operation to them a finite number of times and obtain more recognizable subsets. The question Kleene answered was whether any recognizable subsets exist that cannot be produced in this way.

Let Σ^* be the free monoid on the non-empty set Σ and consider the set $\mathscr{P}(\Sigma^*)$ consisting of all sets of words in Σ^*. We can define three operations on $\mathscr{P}(\Sigma^*)$, namely

$$A \cup B$$
$$A \cdot B$$
$$A^*.$$

They are called the *rational operations* on $\mathscr{P}(\Sigma^*)$ where $A, B \in \mathscr{P}(\Sigma^*)$. Now let $\mathscr{H} \subseteq \mathscr{P}(\Sigma^*)$, we say that $\mathscr{H}$ is *closed under* the rational operations if given $A, B \in \mathscr{H}$ then $A \cup B \in \mathscr{H}$, $A \cdot B \in \mathscr{H}$ and $A^* \in \mathscr{H}$.

We now define a subset $\text{Rat}(\Sigma) \subseteq \mathscr{P}(\Sigma^*)$ as follows. $\text{Rat}(\Sigma)$ is the smallest subset of $\mathscr{P}(\Sigma^*)$ that contains the singleton subsets and $\varnothing$, and is closed under the rational operations.

Suppose that a set $A \in \mathscr{P}(\Sigma^*)$ is either $\varnothing$ or $\{x\}$ (where $x \in \Sigma^*$) or is formed from sets of this type by a finite number of rational operations, then clearly $A \in \text{Rat}(\Sigma)$. We will call such sets *regular* sets of words. The collection of all regular words is written $\text{Reg}(\Sigma)$ and clearly

$$\text{Reg}(\Sigma) \subseteq \text{Rat}(\Sigma).$$

Notice, however, that the set $\text{Reg}(\Sigma)$ is itself closed under the rational operations, it contains the singleton subsets and the empty set, consequently it equals $\text{Rat}(\Sigma)$ which was supposed to be the smallest such set. Thus

$$\text{Reg}(\Sigma) = \text{Rat}(\Sigma)$$

and $\text{Rat}(\Sigma)$ is contained in the set of all recognizable subsets of Σ^* by 5.3.1, 5.3.2, 5.3.3 noting that $\varnothing$, $\{\Lambda\}$, $\{x\}$ ($x \in \Sigma^*$) are all recognizable. (The first two can be found in examples 5.1 and the last one is exercise 5.1.)

Proposition 5.5.1

Let Σ and Γ be non-empty finite sets and suppose that $f: \Sigma \to \Gamma$ is a mapping. Define $f^*: \Sigma^* \to \Gamma^*$ by

$$f^*(\sigma_1 \ldots \sigma_n) = f(\sigma_1) \ldots f(\sigma_n), \quad \sigma_1 \ldots \sigma_n \in \Sigma^+$$
$$f^*(\Lambda) = \Lambda.$$

If A is a regular set of Σ^* then $f^*(A)$ is a regular set of Γ^*.

Proof If A is a singleton then $f^*(A)$ is also a singleton, similarly if A is $\varnothing$, then $f^*(A)$ is $\varnothing$. Suppose that B and C are regular sets in Σ^*, $f^*(B)$ and $f^*(C)$ are regular sets in Γ^*. Then

$$f^*(B \cup C) = f^*(B) \cup f(C) \quad \text{is regular in } \Gamma^*,$$
$$f^*(B \cdot C) = f^*(B) \cdot f^*(C) \quad \text{is regular in } \Gamma^*,$$
$$f^*(B^*) = (f^*(B))^* \quad \text{is regular in } \Gamma^*.$$

An inductive proof based on the number of regular operations in the decomposition of A will establish that $f^*(A)$ is regular in Σ^*. □

The proof of the fact that recognizable sets are regular is best examined with the help of some more abstract terminology, otherwise the details can become rather daunting.

Let S be any finite non-empty set and suppose that $\mathscr{R}$ is a relation on S. We will write

$$a\mathscr{R}a' \quad \text{to mean} \quad (a, a') \in \mathscr{R} \text{ or } a \text{ is related to } a' \text{ under } \mathscr{R}.$$

Now suppose that $\alpha = s_1 \dots s_n \in S^+$ we call α an $\mathcal{R}$-word if

$$s_i \mathcal{R} s_{i+1} \quad \text{for all } i = 1, \dots, n-1.$$

The empty word Λ will also be called an $\mathcal{R}$-word. Given two $\mathcal{R}$-words $\alpha = s_1 \dots s_n$ and $\alpha' = s'_1 \dots s'_n$ we can form further $\mathcal{R}$-words, namely

$$\alpha \cdot \alpha' = s_1 \dots s_n \cdot s'_1 \dots s'_n \quad \text{if } s_n \mathcal{R} s'_1.$$

Given two sets X, Y of $\mathcal{R}$-words then we define the sets

$$X \cup Y$$
$$X \cdot Y = \{x \cdot y \mid x \in X, y \in Y \text{ and } x \cdot y \text{ is an } \mathcal{R}\text{-word}\}$$
$$X^* = \{x_1 \cdot x_2 \dots x_m \mid x_i \in X \text{ and } x_1 \cdot x_2 \dots x_m \text{ is an } \mathcal{R}\text{-word}\}.$$

These are all sets of $\mathcal{R}$-words in S^*.

Given $s_1, s_n \in S$ we define $\mathcal{R}(s_1, s_n)$ to be the set of all $\mathcal{R}$-words in S^* of the form $s_1 \dots s_n$.

Theorem 5.5.2

Let S be a non-empty finite set, $\mathcal{R}$ a binary relation on S and $s_1, s_n \in S$, then the set $\mathcal{R}(s_1, s_n)$ is a regular set of words of S^*.

Proof We proceed by induction on the size of the finite set S. Let $|S| = k$. Consider the case $k = 1$. Suppose that $S = \{s\}$, then we have two possibilities, either $s \mathcal{R} s$ or s is not related to s under $\mathcal{R}$. In the former case the set $\mathcal{R}(s, s) = \{\Lambda, s, s \cdot s, s \cdot s \cdot s, \dots\} = \{s\}^*$, in the latter case $\mathcal{R}(s, s) = \{\Lambda\}$. In both cases $\mathcal{R}(s, s)$ is regular.

Not let $k > 1$ and assume that the result is true for all finite sets S of order less than m. Consider a set S of order m and put $S' = S \backslash \{s_1\}$. Let α be an $\mathcal{R}$-word belonging to $\mathcal{R}(s_1, s_n)$. Then $\alpha = s_1 \cdot \alpha' \cdot s_n$ for some $\mathcal{R}$-word $\alpha' \in S^*$. We can write α in the following form. Either

$$\alpha = s_1^{n_1} \cdot \beta_1 \cdot s_1^{n_2} \cdot \beta_2 \dots s_1^{n_r} \cdot \beta_r \cdot s_n$$

or

$$\alpha = s_1^{n_1} \cdot \beta_1 \cdot s_1^{n_2} \cdot \beta_2 \dots s_1^{n_r} \cdot \beta_r \cdot s_1^{n_{r+1}} \cdot s_n$$

where the $\mathcal{R}$-words $\beta_1, \dots, \beta_r$ do not contain the symbol s_1, and are not the empty word. It is clear that $\beta_1, \dots, \beta_r$ are $\mathcal{R}'$-words in $(S')^*$ if we consider the restriction $\mathcal{R}'$ of the relation $\mathcal{R}$ to the set S'.

Now let $\beta_1 = \gamma_{11} \dots \gamma_{1t_1}$ where $\gamma_{11}, \dots, \gamma_{1t_1} \in S'$ then $\beta_1 \in \mathcal{R}'(\gamma_{11}, \gamma_{1t_1})$ which is a regular set in $(S')^*$. If $\beta_1 \in S'$ then $\beta_1 \in \{\beta_1\}$ which is also regular in $(S')^*$. Similarly for $\beta_2, \dots, \beta_r$. Hence

$$\alpha \in \{s_1\}^* \cdot A_1 \cdot \{s_1\}^* \cdot A_2 \dots \{s_1\}^* \cdot A_r \cdot \{s_n\}$$

or

$$\alpha \in \{s_1\}^* \cdot A_1 \cdot \{s_1\}^* \cdot A_2 \dots \{s_1\}^* \cdot A_r \cdot \{s_1\}^* \cdot \{s_n\}$$

where the $A_1, \dots, A_r$ are all regular sets. If

$$B = \left[\bigcup_{\gamma, \gamma' \in S'} \mathcal{R}'(\gamma, \gamma')\right] \cup \left[\bigcup_{\gamma \in S'} \{\gamma\}\right]$$

then B is also a regular set in $(S')^*$, $A_i \subseteq B$ and so either

$$\alpha \in \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B \dots \{s_1\}^* \cdot B \cdot \{s_n\}$$

or

$$\alpha \in \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B \dots \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot \{s_n\}.$$

It is also clear that if

$$\alpha_1 \in \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B \dots \{s_1\}^* \cdot B \cdot \{s_n\}$$

or

$$\alpha_1 \in \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B \dots \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot \{s_n\}$$

then

$$\alpha_1 \in \mathcal{R}(s_1, s_n)$$

and hence

$$\begin{aligned}\mathcal{R}(s_1, s_n) = {} & [\{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B \dots \{s_1\}^* \cdot B \cdot \{s_n\}] \\ & \cup [\{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot B_1 \dots \\ & \dots \{s_1\}^* \cdot B \cdot \{s_1\}^* \cdot \{s_n\}]\end{aligned}$$

which is a regular set in S^*; and this completes the inductive proof. □

Theorem 5.5.3

If A is a recognizable set of Σ^* then A is regular.

Proof Let $\mathfrak{M} = (\mathcal{M}, i, T)$, $\mathcal{M} = (Q, \Sigma, F)$ be such that $A = |\mathfrak{M}|$. Put $S = Q \times \Sigma$ and consider the set of words S^*. Define the relation $\mathcal{R}$ on S by

$$(q, \sigma) \mathcal{R} (q', \sigma') \Leftrightarrow q' = qF_\sigma \quad \text{for } q, q' \in Q, \sigma, \sigma' \in \Sigma.$$

Now let $\alpha \in A$ then $\alpha \in \Sigma^*$ and $iF_\alpha \in T$. If $\alpha = \sigma_1 \dots \sigma_n$ then the sequence of states $i, iF_{\sigma_1}, iF_{\sigma_1\sigma_2}, \dots, iF_{\sigma_1 \dots \sigma_n}$ defines an $\mathcal{R}$-word in S^* namely:

$$(i, \sigma_1) \cdot (iF_{\sigma_1}, \sigma_2) \dots (iF_{\sigma_1 \dots \sigma_{n-1}}, \sigma_n)$$

which belongs to $\mathcal{R}((i, \sigma_1), (iF_{\sigma_1 \dots \sigma_{n-1}}, \sigma_n))$ which is a regular set of S^*. Let

$$A' = \bigcup \{\mathcal{R}((i, \sigma), (q, \sigma')) \mid \sigma, \sigma' \in \Sigma, q \in Q, qF_{\sigma'} \in T\},$$

then $\alpha \in A'$. Conversely let $\beta \in A'$, then $\beta \in \mathcal{R}((i, \sigma), (q, \sigma'))$ for some $\sigma, \sigma' \in \Sigma$, $q \in Q$, and where $qF_{\sigma'} \in T$. If $\beta = (i, \sigma)(q_1, \sigma_1) \dots (q_n, \sigma_n)(q, \sigma')$ then $\sigma\sigma_1 \dots \sigma_n\sigma' \in A$. Define a function $f: S^* \to \Sigma^*$ by

$$f((q_1, \sigma_1) \dots (q_n, \sigma_n)) = \sigma_1 \dots \sigma_n,$$
$$f(\Lambda) = \Lambda,$$

then $f(A') = A$. Furthermore A' is regular by construction and theorem 5.5.2, and by using proposition 5.5.1 we see that A is also regular. □

We will now reformulate theorem 5.5.3 along with our results from section 5.3 to obtain:

Theorem 5.5.4
(Kleene) Let Σ be a finite non-empty set. The class of recognizable sets of Σ^* equals the class $\mathrm{Reg}(\Sigma)$ of all regular sets of Σ^*.

This result then tells us that the only recognizable sets are those sets constructed from the singletons and $\varnothing$ using the rational operations.

5.6 Prefix decompositions of recognizable sets

The other decomposition of recognizable sets is based on an analysis of the type of sets that are recognized by recognizers with single final states.

Let $\mathfrak{M} = (\mathcal{M}, i, T)$ be a recognizer such that T is a singleton; we call $\mathfrak{M}$ a *direct recognizer.*

A recognizable set $A \subseteq \Sigma^*$ is called *unitary* if the minimal complete recognizer $\mathfrak{M}_A$ is direct.

Theorem 5.6.1
Let $A \subseteq \Sigma^*$ be a recognizable set. Then A is unitary if and only if $A \neq \varnothing$ and $\alpha^{-1} \cdot A = \beta^{-1} \cdot A$ for all $\alpha, \beta \in A$.

Proof Let $\mathfrak{M}_A = (\mathcal{M}_A, i_A, T_A)$ be the minimal complete recognizer for A and suppose that $T_A = \{t_A\}$. Let $\alpha, \beta \in A$, then

$$i_A F^A_\alpha = i_A F^A_\beta = t_A$$

and so

$$\alpha^{-1} \cdot A = \beta^{-1} \cdot A.$$

Clearly $A \neq \varnothing$ as $\mathfrak{M}_A$ is accessible. Now let $\alpha^{-1} \cdot A = \beta^{-1} \cdot A$ for all $\alpha, \beta \in A$, then

$$T_A = \{\gamma^{-1} \cdot A \in Q_A \mid \gamma \in A\} = \{\alpha^{-1} \cdot A\}$$

and so T_A is a singleton. □

Theorem 5.6.2
Let $A \subseteq \Sigma^*$ be recognizable with $A \neq \varnothing$, then $A = \bigcup_{j=1}^{r} A_j$ where the A_j are unitary and $A_j \cap A_k = \varnothing$ if $j \neq k$.

Proof Let $\mathfrak{M}_A = (\mathcal{M}_A, i_A, T_A)$ be the minimal complete recognizer and suppose that $T_A = \{t_1, \dots, t_r\}$.

Now each $t_j \in T_A$ is of the form $\alpha_j^{-1} \cdot A$ for some $\alpha_j \in A$. Let $A_j = \{\beta \in A \mid \beta^{-1} \cdot A = \alpha_j^{-1} \cdot A\}$ for $j = 1, \dots, r$. Then A_j is the behaviour of the recognizer $\mathfrak{M}_j = (\mathcal{M}_A, i_A, \{t_j\})$ and so A_j is recognizable; furthermore if $\beta \in A_j$ then

$$i_A F^A_\beta = t_j$$

and so

$$\beta^{-1} \cdot A_j = \alpha_j^{-1} \cdot A$$

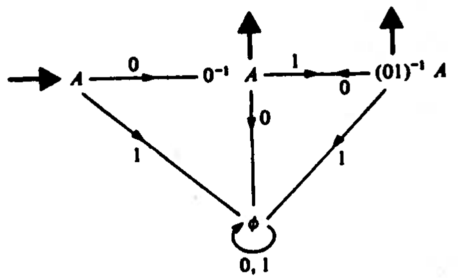
and this holds for all $\beta \in A_j$, and thus A_j is unitary. Since $a \in A$ if and only if $i_A F^A_a = t_j$ for some $j \in \{1, \dots, r\}$ we have $a \in \bigcup_{j=1}^{r} A_j$ and thus $A = \bigcup_{j=1}^{r} A_j$. If $\gamma \in A_j \cap A_k$ with $j \neq k$ then

$$i_A F^A_\gamma = t_j \quad \text{and} \quad i_A F^A_\gamma = t_k$$

which is clearly false. Thus $A_j \cap A_k = \varnothing$. □

We call the sets A_j $(j = 1, \dots, r)$ the *unitary components* of A.

Example 5.7
Let $\Sigma = \{0, 1\}$ and $A = \{0\} \cdot \{10\}^* \cup \{01\}^+$, then A is recognizable and the minimal complete recognizer is given by



Let $A_1 = \{0\} \cdot \{10\}^*$, $A_2 = \{01\}^+$, then $A = A_1 \cup A_2$ where A_1 and A_2 are both unitary sets. Now put $B_1=\{0\}$, $B_2=\{010\}\cdot\{10\}^*$, $B_3=\{01\}^+$; these are all unitary sets and $A=B_1\cup B_2\cup B_3$. Thus we see that the unitary decomposition may not be unique.

Let $A\subseteq\Sigma^*$ and suppose that $\alpha^{-1}\cdot A=\{\Lambda\}$ for all $\alpha\in A$. We call A a *prefix*. A prefix A then has the property that if $\alpha\in A$ the word α cannot be the start of another word from A, that is $\alpha x\notin A$ for all $x\in\Sigma^*$ except $x=\Lambda$. This concept is of considerable interest in coding theory. Here words are encoded by various methods so that transmission of messages across noisy channels can be achieved with as little distortion of the message as possible.

Example 5.8

Let $\Gamma=\{a,b,c,d,e\}$, $\Sigma=\{0,1\}$. We will encode a message, that is a word in Γ^*, into a word in Σ^* by specifying a function $f:\Gamma\to\Sigma^*$. Let $f(a)=1$, $f(b)=01$, $f(c)=001$, $f(d)=0001$, $f(e)=00001$, then the message

cbcdea

is encoded to

001010010001000011.

Now consider another coding function $f':\Gamma\to\Sigma^*$ given by $f'(a)=1$, $f'(b)=10$, $f'(c)=100$, $f'(d)=1000$, $f'(e)=10000$; then the message

cbcdea

is encoded to

100101001000100001.

The receiver will attempt to decode this as it is received and after three symbols all it can decide is that the first decoded symbol is not a or b. In our earlier example, after the first three symbols, the receiver knows that the first decoded symbol is c. We describe the function f as defining a code that can be *immediately decoded.* The function f' defines a code that cannot be immediately decoded. The algebraic difference between the two functions is characterized by the fact that

$$f(\Gamma)=\{0^n1\mid 0\le n\le 4\}$$

is a prefix whereas

$$f'(\Gamma)=\{10^n\mid 0\le n\le 4\}$$

is not a prefix.

It is immediate from theorem 5.6.1 that a prefix is unitary. Furthermore if A is recognizable and a prefix we can characterize the type of recognizer that recognizes A.

Theorem 5.6.3

Let $A\subseteq\Sigma^*$ be a recognizable set. Then A is a prefix if and only if the minimal complete recognizer $\mathfrak{M}_A$ is direct and $T_A*\Sigma=\varnothing$.

Proof If A is a prefix then $\alpha^{-1}A=\{\Lambda\}$ for all $\alpha\in A$ and so $T_A=\{\alpha^{-1}\cdot A\mid\alpha\in A\}=\{\{\Lambda\}\}$. Furthermore for $\sigma\in\Sigma$, $(\alpha^{-1}\cdot A)F_\sigma^A=(\alpha\sigma)^{-1}\cdot A$ and if $\beta\in(\alpha\sigma)^{-1}\cdot A$ then $\alpha\sigma\beta\in A$ which implies that $\sigma\beta\in\alpha^{-1}\cdot A=\{\Lambda\}$. Thus $(\alpha\sigma)^{-1}\cdot A=\varnothing$.

Conversely if A is recognizable and $\mathfrak{M}_A$ has the stated properties let $\alpha\in A$, then $i_AF_\alpha^A=\alpha^{-1}\cdot A\in T_A$ and so $\alpha^{-1}\cdot A=t_A$ where $T_A=\{t_A\}$. Suppose that $\beta\in\alpha^{-1}\cdot A$ with $\beta\ne\Lambda$, then $\alpha\beta\in A$. Let $\beta=\sigma\gamma$ where $\gamma\in\Sigma^*$, then $\alpha\sigma\gamma\in A$ and so $\gamma\in(\alpha\sigma)^{-1}\cdot A=t_AF_\sigma^A=\varnothing$ which is a contradiction. Hence $\alpha^{-1}\cdot A=\{\Lambda\}$ and A is a prefix. □

We have seen that a set A is a prefix if there are no words of the form $\alpha=\beta\gamma$ where both α and β belong to A. It is easy to construct the prefix part of any subset of Σ^* by removing all such words.

Let $A\subseteq\Sigma^*$ be recognizable, define the *prefix part of A* to be $A_p=A\backslash A\cdot\Sigma^+$. It is immediate that a recognizable subset A will be a prefix if and only if $A=A_p$. (Exercise 5.9 is concerned with the task of verifying that A_P is recognizable.)

We have seen that for a given recognizable subset A the prefix part A_P has some special properties. What can be said of the remainder of A? We first examine an example.

Example 5.9

Let $\Sigma=\{0,1\}$ and $A=\{0\}^*\{1\}\cdot\{0\}^*$, then

$$A_P=A\backslash A\cdot\Sigma^+=A\backslash\{0^n10^m\mid m>0\}=\{0\}^*\cdot\{1\}.$$

Notice that A is a unitary set and any element $\alpha\in A$ can be written in the form $\beta\cdot\gamma$ where $\beta\in A_P$ and $\gamma\in\{0\}^*$. Thus

$$A=A_P\cdot\{0\}^*.$$

We will now investigate the properties of $\{0\}^*$. Notice firstly that $\{0\}^*$ is a monoid, and secondly $\gamma^{-1}\cdot\{0\}^*=\{0\}^*$ for all $\gamma\in\{0\}^*$, that is $\{0\}^*$ is a unitary subset (it is clearly recognizable). We call $\{0\}^*$ a *unitary monoid.*

Our basic aim is the decomposition of a recognizable set into subsets of the form $A \cdot M$ where A is a prefix and M is a unitary monoid.

Let $B \subseteq \Sigma^*$; we call B a *unitary monoid* if

(i) B is a unitary subset of Σ^* (B is thus recognizable);

(ii) B is a submonoid of Σ^*.

Since B is a submonoid of Σ^* we see that $\Lambda \in B$ and thus $\gamma^{-1} \cdot B = B$ for all $\gamma \in B$.

Theorem 5.6.4

Given any unitary subset A of Σ^* the set

$$A_M = A^{-1} \cdot A = \{\gamma \in \Sigma^* \mid \alpha\gamma \in A \text{ for some } \alpha \in A\}$$

is a unitary monoid and $A = A_P \cdot A_M$.

Proof Since A is unitary the minimal complete recognizer $\mathfrak{M}_A = (\mathcal{M}_A, i_A, T_A)$ is direct. Let $T_A = \{t_A\}$ and consider the recognizer $\bar{\mathfrak{M}}_A = (\mathcal{M}_A, t_A, T_A)$. Now

$$\begin{aligned} \beta \in |\bar{\mathfrak{M}}_A| &\Leftrightarrow t_A F^A_\beta = t_A \\ &\Leftrightarrow i_A F^A_{\alpha\beta} = t_A \quad \text{for any } \alpha \in A \\ &\Leftrightarrow \alpha\beta \in A \quad \text{for any } \alpha \in A \\ &\Leftrightarrow \beta \in \alpha^{-1} \cdot A \quad \text{for any } \alpha \in A \\ &\Leftrightarrow \beta \in A^{-1} \cdot A. \end{aligned}$$

Thus $A^{-1} \cdot A$ is recognizable and unitary. Furthermore, if $\beta, \beta' \in A^{-1} \cdot A$ then clearly $t_A F^A_{\beta\beta'} = t_A$ and so $\beta\beta' \in A^{-1} \cdot A$ and $A^{-1} \cdot A$ is a unitary monoid. Now let $\alpha \in A$, then $i_A F^A_\alpha = t_A$. The sequence of states defined by α may contain t_A several times. If we put $\alpha = \gamma\beta$ when $\gamma \in A$ and $\beta \in A^{-1} \cdot A$ such that the path from i_A to t_A labelled by γ contains only one occurrence of t_A, namely the last one, then $\gamma \in A_P = A \backslash A\Sigma^+$. Thus $A \subseteq A_P \cdot A_M$ and the reverse inclusion is obvious. □

Theorem 5.6.5

Let A be a recognizable subset of Σ^*. Then

$$A = B_1C_1 \cup B_2C_2 \cup \ldots \cup B_rC_r$$

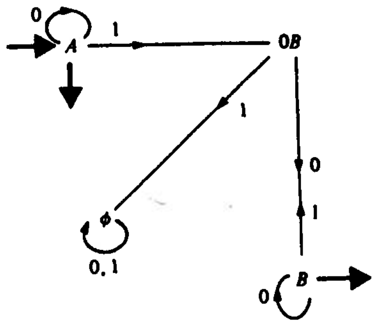
where B_iC_i are unitary subsets, B_i are prefixes and C_i are unitary monoids for $i = 1, 2, \ldots, r$.

Proof Using theorem 5.6.2 we have $A = A_1 \cup A_2 \cup \ldots \cup A_r$ where each A_i is a unitary subset. Now let $B_i = (A_i)_P = A_i \backslash A_i\Sigma^+$, $C_i = (A_i)_M = (A_i)^{-1} \cdot A_i$ for $i = 1, \ldots, r$, then $A_i = B_iC_i$ and B_i is a prefix and C_i is a unitary monoid. □

This decomposition is called the *unitary-prefix decomposition*. We finish our discussion with some examples.

Example 5.10

Let $\Sigma = \{0, 1\}$ and $A = \{0\}^* \cdot \{\{10\} \cup \{0\}\}^* \cdot \{0\}^*$. This is recognizable by construction. Let $B = \{\{10\} \cup \{0\}\}^* \cdot \{0\}^*$. Now $0^{-1} \cdot A = A$, $1^{-1} \cdot A = 0B$, $(01)^{-1} \cdot A = 1^{-1} \cdot A = 0B$, $(10)^{-1} \cdot A = B$, $(11)^{-1} \cdot A = \varnothing$, $(100)^{-1} \cdot A = B$, $(101)^{-1} \cdot A = 0B$. The minimal complete recognizer is given by:

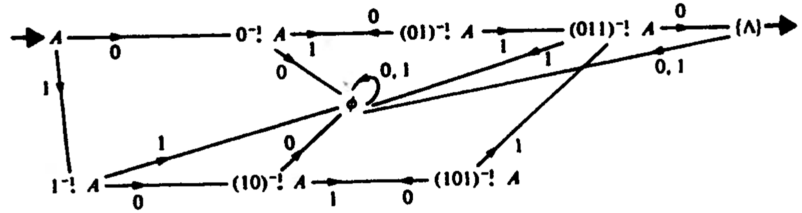


with $t_A = \{A, B\}$. The unitary decomposition is $A = \{0\}^* \cup \{0\}^*\{10\}B$. Now $\{0\}^* = \{\Lambda\} \cdot \{0\}^*$ where $\{\Lambda\}$ is a prefix and $\{0\}^*$ is a unitary monoid. Also $\{0\}^*\{10\}$ is a prefix and B is a unitary monoid.

Example 5.11

Let $\Sigma = \{0, 1\}$ and $A = (\{01\}^+ \cdot \{10\}) \cup (\{10\}^+ \cdot \{110\})$. Then $0^{-1} \cdot A = \{1\} \cdot \{01\}^* \cdot \{10\}$, $1^{-1} \cdot A = \{0\} \cdot \{10\}^* \cdot \{110\}$, $(01)^{-1} \cdot A = \{01\}^* \cdot \{10\}$, $(10)^{-1} \cdot A = \{10\}^* \cdot \{110\}$, $(011)^{-1} \cdot A = \{0\}$, $\{101\}^{-1} \cdot A = 1^{-1} \cdot A \cup \{10\}$, $(010)^{-1} \cdot A = 0^{-1} \cdot A$, $(0110)^{-1} \cdot A = \{\Lambda\} = (10110)^{-1} \cdot A$ etc.

The minimal complete recognizer is:



From the diagram we note that $A = A_P$ is a prefix and so $A = A \cdot \{\Lambda\}$ is the unitary decomposition.

Example 5.12

Let $\Sigma = \{0, 1\}$ and suppose that A is the set of words of Σ^* containing an equal number of 0s and 1s. Let A_1 be the set of words of A containing an odd number of 0s and A_2 the set of words of A containing an even number of 0s. Then $A = A_1 \cup A_2$. Let us try to construct the minimal complete recognizer for A.

Let

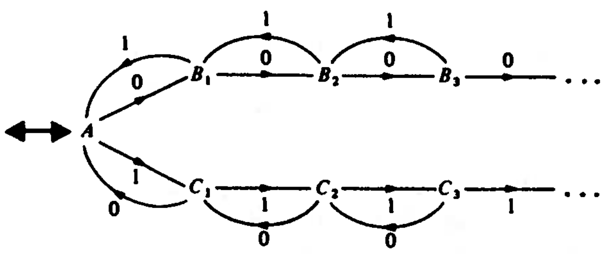
$$B_j = \{\alpha \in \Sigma^* \text{ such that } \alpha \text{ has } j \text{ more 1s than 0s}\}$$

$$C_j = \{\alpha \in \Sigma^* \text{ such that } \alpha \text{ has } j \text{ more 0s than 1s}\}$$

Then $0^{-1} \cdot A = B_1, (01)^{-1} \cdot A = A, (00)^{-1} \cdot A = B_2, \ldots$

$1^{-1} \cdot A = C_1, (10)^{-1} \cdot A = A, (11)^{-1} \cdot A = C_2, \ldots$

The 'machine' will have a graph of the following form:



and it is clear that the set of states will have to be infinite. In fact it can easily be shown that A is not recognizable. However it is possible to devise a decomposition for A that is similar, in some respects, to the unitary–prefix decomposition. Notice that A_2 is a monoid and satisfies the condition $\alpha^{-1} \cdot A_2 = \beta^{-1} \cdot A_2 = A_2$ for any $\alpha, \beta \in A_2$. However A_2 is not a unitary monoid since it is not recognizable. Similarly $A_1 = \{01, 10\} \cdot A_2$ and $\{01, 10\}$ is a prefix. Thus

$$A = \{01, 10\} \cdot A_2 \cup \{\Lambda\} \cdot A_2$$

where $\{01, 10\}$ and $\{\Lambda\}$ are prefixes and A_2 is a monoid satisfying the condition $\alpha^{-1} \cdot A_2 = \beta^{-1} \cdot A_2$ for all $\alpha, \beta \in A_2$.

5.7 The pumping lemma and the size of a recognizable set

We examine a useful technique for testing the recognizability of subsets of Σ^*. This then leads us to a method for deciding if a recognizable subset is finite. Following this we investigate the size of an infinite recognizable set.

Lemma 5.7.1

(Pumping lemma) Let $A \subseteq \Sigma^*$ be recognizable and suppose that $n = |Q_A|$, the number of states in the minimal complete recognizer of A. If $\alpha \in A$ and the length of α is greater than or equal to n then

$$\alpha = \beta\gamma\delta$$

such that

(i) $\gamma \neq \Lambda$,
(ii) $\{\beta\} \cdot \{\gamma\}^* \cdot \{\delta\} \subseteq A$.

Proof Suppose that $\alpha \in A$, then $\alpha^{-1}A \in T_A$. The sequence of states $i_A = q_0, q_1, \ldots, q_r = \alpha^{-1} \cdot A$ defined by the word α is of length $n+1$. There must therefore be repetitions so that $q_j = q_k$ with $j \neq k$. Consider the word $\gamma \in \Sigma^*$ obtained by passing along the path defined by α between q_j and q_k. Then clearly a word $\beta \in \Sigma^*$ and a word $\delta \in \Sigma^*$ exist such that

$$i_A F^A_\beta = q_j, \; q_j F^A_\gamma = q_j, \; q_j F^A_\delta = \alpha^{-1} \cdot A = i_A F^A_\alpha.$$

Then $\alpha = \beta\gamma\delta$, $\gamma \neq \Lambda$ and any word of the form $\beta\gamma^m\delta$ is recognized. $\square$

Example 5.13

Consider the recognizer in example 5.10. Here $n = 4$ and if $\alpha = 00010100$ we see that $\beta = 000$, $\gamma = 1010$, $\delta = 0$ gives a suitable decomposition $\alpha = \beta\gamma\delta$. Others exist, for example $\beta' = 0$, $\gamma' = 00$, $\delta' = 10100$. Notice that $\{\beta\} \cdot \{\gamma\}^* \cdot \{\delta\} \neq \{\beta'\} \cdot \{\gamma'\}^* \cdot \{\delta'\}$.

We see then that the existence in the recognizable set of a word of length at least n will guarantee that the set is infinite. If $A \subseteq \Sigma^*$ is a finite recognizable set and $n = |Q_A|$ then no words of length n can exist in A.

If $A \subseteq \Sigma^*$ let us define $A^{(n)}$ to be the set of all words of A that are of length n for $n = 0, 1, \ldots$. Then

$$A = \bigcup_{n=0}^{\infty} A^{(n)}.$$

For a finite set A we will have $A^{(n)} = A^{(n+1)} = \ldots = \varnothing$ for some value of n. For an infinite set each $A^{(n)}$ is finite, in fact

$$|A^{(n)}| \leq k^n \quad \text{where } k = |\Sigma|.$$

Our next task is to find some information about the size of the sets $A^{(n)}$ when A is a recognizable subset of Σ^*.

First let $\mathcal{M}=(Q, \Sigma, F)$ be a complete finite state machine and let $|Q| = m$. The machine $\mathcal{M}$ can be described by a set of $m \times m$ matrices that effectively define the action of F.

First we let $Q = \{q_1, \ldots, q_m\}$ and then for each $\sigma \in \Sigma$ define the matrix

$$f_\sigma = (f_{ij}^\sigma) \quad \text{where } f_{ij}^\sigma = \begin{cases} 1 & \text{if } q_iF = q_j \\ 0 & \text{otherwise} \end{cases}$$

for $i, j \in \{1, \ldots, m\}$.

Each row of the matrix f_σ will consist of one 1 and $(m-1)$ 0s. Each state q_j will be represented by a $1 \times m$ row vector e_j of the form $(0\ldots 010 \ldots 0)$ with a 1 in the j-th position. So that $q_jF_\sigma = q_k$ will be replaced by the matrix equation $e_j \cdot f_\sigma = e_k$.

Given $\alpha = \sigma_1 \ldots \sigma_n \in \Sigma^*$ we define $f_\alpha = f_{\sigma_1} \ldots f_{\sigma_n}$ and notice that

$$q_jF_\alpha = q_k \Leftrightarrow e_j \cdot f_\alpha = e_k.$$

Finally we put $f_\Lambda = I_m$, the $m \times m$ identity matrix. If $\mathfrak{M} = (\mathcal{M}, i, T)$ is a recognizer, let $i = q_1$ and define

$$\mathscr{E} = \{e_j \mid q_j \in T\},$$

then for each $\alpha \in |\mathfrak{M}|$ we have $e_1 \cdot f_\alpha \in \mathscr{E}$ and clearly

$$|\mathfrak{M}| = \{\alpha \in \Sigma^* \mid e_1 \cdot f_\alpha \in \mathscr{E}\}.$$

Let $\mathscr{F} = \sum_{\sigma \in \Sigma} f_\sigma$, which is again an $m \times m$ matrix (it belongs to the set of all $m \times m$ matrices over the integers); we call $\mathscr{F}$ the *matrix* of $\mathcal{M}$. For any subset $R \subseteq Q$ we define

$$\mathscr{E}(R) = \{e_j \mid q_j \in R\}$$

and consider

$$\mathfrak{E}(R) = \sum_{e_j \in \mathscr{E}(R)} = e_j^T.$$

(e_j^T is the transpose of e_j and thus $\mathfrak{E}(R)$ is a column vector.)

Theorem 5.7.2

Let $\mathfrak{M} = (\mathcal{M}, i, T)$ be a recognizer with matrix $\mathscr{F}$. Let R be a set of states of $\mathcal{M}$ and $k \geq 0$, then the number of words of Σ^* of length k which send the initial state i to a state in R is given by

$$e_1 \cdot (\mathscr{F})^k \cdot \mathfrak{E}(R).$$

Proof Let Σ^k be the set of words of Σ^* of length k. Then $\mathscr{F}^k = \sum_{\alpha \in \Sigma^k} f_\alpha$ and $e_1 \cdot (\mathscr{F})^k = \sum_{\alpha \in \Sigma^k} e_1 \cdot f_\alpha = (a_{11}, \ldots, a_{1m})$ where m is the number of states in $\mathcal{M}$ and a_{ij} is the number of words of length k

that send state i to state q_j. The number of words of length k that send state i to a state in R is then given by

$$a_{11}b_1 + \ldots + a_{1m}b_m \quad \text{where } b_j = 1 \text{ if } q_j \in R \text{ and } b_j = 0 \text{ if } q_j \notin R.$$

This is just $(a_{11}, \ldots, a_{1m}) \cdot \mathfrak{E}(R)$. □

Corollary 5.7.3

The number of words in $|\mathfrak{M}|$ of length k is given by

$$e_1 \cdot (\mathscr{F})^k \cdot \mathfrak{E}(T).$$

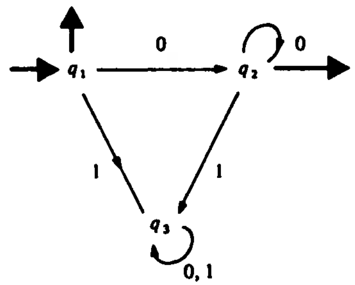
The total number of words in $|\mathfrak{M}|$ is given by

$$e_1 \cdot (\mathscr{F}^0 + \mathscr{F}^1 + \mathscr{F}^2 + \ldots + \mathscr{F}^k + \ldots) \cdot \mathfrak{E}(T) = e_1 \cdot (I - \mathscr{F})^{-1} \cdot \mathfrak{E}(T).$$

In the case of an infinite set $|\mathfrak{M}|$ the matrix $I - \mathscr{F}$ will be singular and so great care must be taken with this notation.

Example 5.14

Let $\Sigma = \{0, 1\}$ and consider the state machine $\mathcal{M}$ given by



Let $i = q_1$, $T = \{q_1, q_2\}$. Then if $\mathfrak{M} = (\mathcal{M}, i, T)$, $|\mathfrak{M}| = \{0\}^*$.

$$f_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad f_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathscr{F} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

$$e_1 = (1, 0, 0), \qquad \mathfrak{E}(T) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

and the number of words of $|\mathfrak{M}|$ of length 2 is given by

$$e_1 \cdot \mathscr{F}^2 \cdot \mathfrak{E}(T) = (1, 0, 0) \cdot \begin{pmatrix} 0 & 1 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1.$$