Deterministic Finite Automata

An *alphabet* Σ is a finite set. A *string* over an alphabet Σ is a finite sequence of elements of Σ. A string of length 0 is the empty string and denoted ϵ.

Example Let $\Sigma = \{a, b, c\}$. The following are all strings over Σ: aab, $bccabca$, $abcabcabc$, $bbbb$. Note that $ab \neq ba$

We denote the set of all strings over a given alphabet Σ by Σ^*. We define $\emptyset^* = \{\epsilon\}$. For any set other than the empty set, Σ^* is an infinite set.

A *deterministic finite automata* (DFA) is a model with 5 components:

$M = (Q, \Sigma, \delta, s, F)$

$Q =$ a finite set, the *states*.

$\Sigma =$ a finite set, the *alphabet*

$\delta =$ a function $Q \times \Sigma \to Q$, the *transition function*

$s =$ an element of Q, the *start state*

$F =$ a subset of Q, the *final states*

Example Let M be specified by the following data:

$Q = \{0, 1, 2\}$,

$\Sigma = \{a, b\}$,

$\delta(0, a) = 0$, $\delta(0, b) = 1$, $\delta(1, a) = 0$, $\delta(1, b) = 2$, $\delta(2, a) = 2$, $\delta(2, b) = 2$

$s = 0$

$F = 2$

If we input a string in this example with at least two bs, we will end up in state 2, the final state.

We call a string *accepted* by M if we run the DFA on the string and end in a final state. In general, the set of all strings which are accepted by some DFA M is called the *language* accepted by M and written $L(M)$. A set of strings is called *regular* if it is equal to $L(M)$ for some M.

The transition function of a DFA is defined on $Q \times \Sigma$. We would like to extend δ to be defined on strings. We build such a function recursively:

$\hat{\delta} : Q \times \Sigma^* \to Q$

$\hat{\delta}(q, \epsilon) = q$

$\hat{\delta}(q, xa) = \delta(\hat{\delta}(q, x), a)$

With this extended transition function we can define a language accepted by M as: $L(M) = \{x \in \Sigma^* | \delta(s, x) \in F\}$.

Example: The following set: $\{a^n b^n | n \geq 0\}$ is not a regular set. (why?)

Lemma 1. Pumping Lemma *For any regular set R: There exists $k \geq 0$ s.t. for any strings x, y, z with $xyz \in A$ and $|y| \geq k$ there exists u, v, w s.t. $y = uvw$ and for all i, $xuv^i wz \in A$.*

We can form new DFAs from other ones. Here we look at the *product construction*. Suppose we have two DFAs, M_1 and M_2 both with input language Σ. Define M_3 over Σ with statistics:

$Q_3 = Q_1 \times Q_2 = \{(p, q) | p \in Q_1 \text{ and } q \in Q_2\}$

$F_3 = F_1 \times F_2 = \{(p, q) | p \in F_1 \text{ and } q \in F_2\}$

$s_3 = (s_1, s_2)$

$\delta_3((p, q), a) = (\delta_1(p, a), \delta_2(q, a))$