Minimization of DFAs

Note that the equivalence of states is an equivalence relation. We can use this fact to minimize DFAs.

For a given DFA, we run the Table Filling Algorithm, to find all the equivalent states, and hence all the equivalence classes. We call each equivalence class a *block*.

In our last example, the blocks would be:

$$\{E, A\}, \{H, B\}, \{C\}, \{F, D\}, \{G\}$$

The states within each block are equivalent, and the blocks are disjoint.

We now build a minimal DFA with states given by the blocks as follows: $\gamma(S, a) = T$, where $\delta(p, a) \in T$ for $p \in S$.

We must show that γ is well defined; suppose we choose a different $q \in S$. Is it still true that $\delta(q, a) \in T$?

Suppose not, i.e., $\delta(q, a) \in T'$, so $\delta(p, a) = t \in T$, and $\delta(q, a) = t' \in T'$. Since $T \neq T'$, $\{t, t'\}$ is a distinguishable pair. But then so is $\{p, q\}$, which contradicts that they are both in S.

Theorem: We obtain a minimal DFA from the procedure.

Proof: Consider a DFA A on which we run the above procedure to obtain M. Suppose that there exists an N such that $L(N) = L(M) = L(A)$, and N has fewer states than M.

Run the Table Filling Algorithm on M, N together (renaming the states, so they don't have states in common). Since $L(M) = L(N)$ their initial states are indistinguishable. Thus, each state in M is indistinguishable from at least one state in N. But then, two states of M are indistinguishable from the same state of $N \ldots$

Context-Free Grammars and Languages

A CFG is expressed by $G = (V, T, P, S)$ (Variables, Teraminals, Productions, Start variable).

Ex. $P \longrightarrow \varepsilon|0|1|0P0|1P1$.

Ex. $G = (\{E, I\}, T, P, E)$ where $T = \{+, *, (,), a, b, 0, 1\}$ and P is the following set of productions:

$$E \longrightarrow I|E + E|E * E|(E)$$

$$I \longrightarrow a|b|Ia|Ib|I0|I1$$

If $\alpha A \beta \in (V \cup T)^*$, $A \in V$, and $A \longrightarrow \gamma$ is a production, then $\alpha A \beta \Rightarrow \alpha \gamma \beta$. We use $\overset{*}{\Rightarrow}$ to denote 0 or more steps.

$L(G) = \{w \in T^* | S \overset{*}{\Rightarrow} w\}$

$L((\{P\},\{0,1\},\{P \longrightarrow \varepsilon|0|1|0P0|1P1\},P))$ is the set of Palindromes over $\{0,1\}$.

Proof: Suppose w is a palindrome; show by induction on $|w|$ that $P \overset{*}{\Rightarrow} w$.

BS: $|w| \leq 1$, so $w = \varepsilon, 0, 1$, so use $P \longrightarrow \varepsilon, 0, 1$. IS: For $|w| \geq 2$, $w = 0x0, 1x1$, and by IH $P \overset{*}{\Rightarrow} x$.

Suppose that $P \overset{*}{\Rightarrow} w$; show by induction on the number of steps in the derivation that $w = w^R$.

BS: Derivation has 1 step. IS: $P \Rightarrow 0P0 \overset{*}{\Rightarrow} 0x0 = w$ (or with 1 instead of 0).

If $S \overset{*}{\Rightarrow} \alpha$, then $\alpha \in (V \cup T)^*$, so α is a **sentential form**. $L(G)$ is the set of those sentential forms wich are in T^*.