# The Theory of Computation

Bernard M.E. Moret

Addison-Wesley-Longman, 1998

## *Solution Manual*
## *Chapter 3, Additional Problem 3.1*

The key observation (a very simple one) is that, if $x$ is a suitable string—let's call it a *distinguishing* string, then, for every prefix $y$ of $x$, we must have $\delta(p, y) \neq \delta(q, y)$. So, until the end of the string is reached, the two sequences of moves go through pairs of different states; moreover, both states in each pair are accepting or both are rejecting—otherwise we would have found a shorter distinguishing string. There are $(n - |F|)(n - |F| - 1)$ pairs of different rejecting states and $|F|(|F| - 1)$ pairs of different accepting states, or $f(n, F) = n^2 - n + 2|F|(|F| - n)$ pairs in all. The shortest distinguishing string never goes through the same pair twice—if it did, that loop could simply be cut out. So the shortest distinguishing string takes the machine through at most $f(n, F)$ transitions and thus has at most $f(n, F)$ symbols. This is not very long: even when $|F| = 1$, the value is at most $(n - 1)(n - 2)$.

On the other hand, it is unlikely that this bound is tight, because we have approached the two copies of the machine as if they were completely independent; since they are, in fact, copies of each other, certain transitions may not be possible in the paired machine. For instance, if we are in combined state $(q_i, q_j)$ and, under input $a$, have a transition to combined state $(q_j, q_k)$, then, from the latter state, under input $a$, the transition must be to a combined state of the form $(q_k, \ )$—not just to any combined state.

Can we then do better? That is, can we derive a tighter bound through a different approach? Suppose we want to divide all of the $n$ states of the DFA into groups, such that states within any group are indistinguishable from each other. We can do this constructively as follows: begin with two groups, one being $F$, the other $Q - F$—any accepting state is immediately distinguishable from any rejecting state. At any step of the process, we will have a certain number of groups, call them $G_1, G_2, \ldots$ . Now look at the states reachable in one transition on symbol $a$ from any of the states within a group $G_i$; if all of these states are in the same group, then the states in group $G_i$ remain indistinguishable under additional symbol $a$; if that is the case for every input symbol $a$, then the group of states in $G_i$ remains indistinguishable under any input sequence longer by one symbol and thus $G_i$ remains unchanged to the next step. On the other hand, if there is some input symbol $a$ such that the successors of some states in $G_i$ are in some group $G_j$ and the successors of some other states in $G_i$ are in some other group $G_k$ $(j \neq k)$, then the states in the first group and the states in the second group can be distinguished and thus group $G_i$ gets split into two or more groups for the next step. If all groups remain unchanged from one step to the next, the final answer has been found: no more distinguishing of any type can take place. Thus the process stops as soon as two consecutive steps return the same groups; viewed the other way, the process continues as long as at least one group is split, thereby increasing the number of groups. But the number of groups cannot exceed $n$, the number of states of the machine—all states are then distinguishable—and there were two groups initially, so that the number of steps is at most $n - 2$.

We claim that this process creates distinguishing sequences automatically, so that no distinguishing sequence need be longer than $n - 2$ symbols. If we have two distinguishable states $p$ and $q$, we find in the process the first step, say $i$, at which the two are no longer in the same group; then a sequence of $i - 1$ symbols will serve as a distinguishing sequence and can be reconstructed by following the sequence of groups splits that led to the last group in which $p$ and $q$ remain together, followed by the symbol that led to the split which sundered $p$ from $q$.