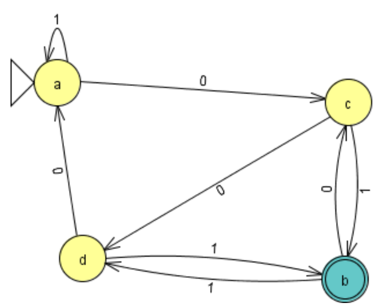
DFSM canonical form

Motivation: First, the reason for canonical form is to determine if two minimal DFAs are equivalent. In general determining whether two graphs are isomorphic is widely believed to be an NP-complete problem (we'll define that concept at the end of the course). But for two minimal DFSMs, there is a much more efficient algorithm.

Firsts, if the two minimal DFSMs M1 and M2 with states $q_1, \ldots, q_n$ and $r_1..r_m$ have different numbers of states, they are not equivalent. If they have the same number of states, they are equivalent iff and there is a way to number the r states in such a way that for every input symbol a, $\delta(r_i, a)$ is the state that corresponds to $\delta(q_i, a)$.

Canonical form is a way of numbering the states of a minimal DFSM in such a way that this will be easy to do. First we order the input symbols. We number the start state as q_0, then to a breadth-first traversal of the graph, always following the transitions in the order of the input symbols. Details of the algorithm are given on page 95 of the textbook.

Example



q_0 is the start state

$\delta(q_0, 0) = c$, so c becomes q_1.

$\delta(q_0, 1`) = q_0$, which we have already seen.

$\delta(q_1, 0) = d$, so d becomes q_2.

$\delta(q_1, 1) = b$, so b becomes q_3.

This is sufficient to number all of the states, but if there were more states, we would look at them in this order: $\delta(q_2, 0)$, $\delta(q_2, 1)$, $\delta(q_2, 0)$, $\delta(q_2, 1)$,