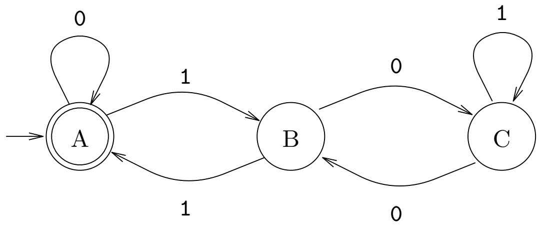
Handout 1: Finite Automata

1.1 Example. Check if a binary number is divisible by 3. While processing the input, we need to remember not the entire processed portion x of the input, but only the remainder of x when divided by 3. This remainder can be updated one input symbol at a time:

$$
\begin{array}{lcl}
x \equiv 0 \bmod 3 & \to & x\mathtt{0} \equiv 0 \bmod 3 \\
x \equiv 0 \bmod 3 & \to & x\mathtt{1} \equiv 1 \bmod 3 \\
x \equiv 1 \bmod 3 & \to & x\mathtt{0} \equiv 2 \bmod 3 \\
x \equiv 1 \bmod 3 & \to & x\mathtt{1} \equiv 0 \bmod 3 \\
x \equiv 2 \bmod 3 & \to & x\mathtt{0} \equiv 1 \bmod 3 \\
x \equiv 2 \bmod 3 & \to & x\mathtt{1} \equiv 2 \bmod 3
\end{array}
$$

We remember "$x \equiv 0 \bmod 3$" in state A, and "$x \equiv 1 \bmod 3$" in state B, and "$x \equiv 2 \bmod 3$" in state C:



1.2 Definition of a finite automaton. A *finite automaton* M consists of

Q ... a finite set of states (the state space),
Σ ... a finite set of input symbols (the input alphabet),
δ: $Q \times \Sigma \to Q$ ... a transition function,
$q_0 \in Q$ ... an initial (or start) state,
$F \subseteq Q$ ... a set of final (or accept) states.

In mathematics, we say that M is the 5-tuple $(Q, \Sigma, \delta, q_0, F)$. In examples, it is often most intuitive to draw the state diagram of M, which is a labeled digraph as shown in Example 1.1. For this example:

$Q = \{\text{A}, \text{B}, \text{C}\}$,
$\Sigma = \{\mathtt{0}, \mathtt{1}\}$,
$\delta(\text{A}, \mathtt{0}) = \text{A}$, $\delta(\text{A}, \mathtt{1}) = \text{B}$, $\delta(\text{B}, \mathtt{0}) = \text{C}$, etc.,
$q_0 = \text{A}$,
$F = \{\text{A}\}$.

1.3 Language of a finite automaton. A *run* of M is a sequence

$$p_0 \overset{a_0}{\to} p_1 \overset{a_1}{\to} \cdots \overset{a_{n-1}}{\to} p_n$$

with $p_0, \ldots, p_n \in Q$ and $a_0, \ldots, a_{n-1} \in \Sigma$ such that

(1) $p_0 = q_0$,
(2) $p_{i+1} = \delta(p_i, a_i)$ for all $i \in [0..n-1]$.

The run *accepts* the input word $a_0 \ldots a_{n-1}$ if

(3) $p_n \in F$.

The *language* of M is

$$L(M) = \{w \in \Sigma^* \mid w \text{ is accepted by some run of } M\}.$$

The language of the automaton from Example 1.1 is the set of binary numbers that are divisible by 3. For instance, the word `011` is accepted by the following run:

$$\text{A} \overset{\texttt{0}}{\to} \text{A} \overset{\texttt{1}}{\to} \text{B} \overset{\texttt{1}}{\to} \text{A}$$

By contrast, the word `010` is not accepted by any run; in particular, the run

$$\text{A} \overset{\texttt{0}}{\to} \text{A} \overset{\texttt{1}}{\to} \text{B} \overset{\texttt{0}}{\to} \text{C}$$

does not end in a final state.