Today's lecture: Finite Automata

I. Sequential Circuits and Finite Automata
II. Finite Automata as Language Recognizers
III. The Proof from the End of the Sept. 8 Lecture

Schedule:

Today: Finite Automata – Read: *Sipser* 1.1.
Lecture will cover through Example 1.15 (i.e. pages 31–40).

September 13: Regular Languages.
The rest of *Sipser* 1.1 (i.e. pages 40–47).

September 15: Non-Determinism – Read: *Sipser* 1.2.
Lecture will cover through Example 1.35 (i.e. pages 47–52).
Homework 1 goes out (due Sept. 25).

September 18: NFAs
Lecture will cover through Example 1.38 (i.e. pages 53-54).
Homework 0 due.

September 20 and beyond: see Sept. 6 notes.

I. Sequential Circuits and Finite Automata

I may update these notes with details for this section and the next, but that is by no means certain to happen in the near future.

II. Finite Automata as Language Recognizers

III. The Proof from the End of the Sept. 8 Lecture

A. The claim: Let $\Sigma = \{0, 1\}$. Consider the set $S \subseteq \Sigma^*$, such that w is in S iff

$w = \epsilon$; or

There is a string x in S such that $w = 0x1$ or $w = 1x0$; or

There are strings x and y in S such that $w = xy$.

Given a string, $w \in \Sigma^*$, w is in S iff the number of O's in w is equal to the number of 1's.

B. The proof.

Let $nZero(w)$ denote the number of 0's in w, $nOne(w)$ denote the number of 1's in w and $P(w)$ be true iff $nZero(w) = nOne(w)$.

1. $(w \in S) \Rightarrow P(w)$:

Let w be an arbitrary string in S. Our proof is by induction on the structure of w; P is the induction hypothesis for our proof. From the definition of S, we have four cases to consider:

$w = \epsilon$: In this case $nZero(w) = nOne(w) = 0$; therefore, $P(w)$ is satisfied.

$\exists x \in S.\ w = 0x1$: In this case $nZero(w) = nZero(x) + 1$. Likewise, $nOne(w) = nOne(x) + 1$. Because $x \in S$, we can apply the induction hypothesis to conclude $P(x)$, $nZero(x) = nOne(x)$. Thus, $nZero(w) = nOne(w)$ which means that $P(w)$ is satisfied.

$\exists x \in S.\ w = 1x0$: The proof for case is equivalent to the one for $w = 0x1$.

$\exists x, y \in S.\ w = xy$: In this case $nZero(w) = nZero(x) + nZero(y)$, and $nOne(w) = nOne(x) + nOne(y)$. Because x and y are in S, we can apply the induction hypothesis to conclude $nZero(x) = nOne(x)$ and $nZero(y) = nOne(y)$. Thus, $nZero(w) = nOne(w)$ which means that $P(w)$ is satisfied.

This completes the proof for $(w \in S) \Rightarrow P(w)$.

At this point, you might ask: "Why isn't this circular reasoning?" This is because the proof works on the structure of w. The rules for S define derivation *trees*. Trees don't have cycles, and this is what makes it so that the proof isn't circular reasoning.

As an example, let $w = 001101$. Let's define functions for each of the four cases in the definition of S:

$$\begin{aligned} S_0() &= \epsilon \\ S_1(x) &= 0x1 \\ S_2(x) &= 1x0 \\ S_3(x, y) &= xy \end{aligned}$$

Now we have:

- $w = S_1(S_3(S_1(S_0()), S_2(S_0())))$.
- $S_0() = \epsilon$. We showed $P(\epsilon)$ in the first case of the induction proof.
- $S_1(S_0()) = S_1(\epsilon) = 01$.
  Having shown $P(\epsilon)$, this is handled by the second case of the induction proof.
- $S_2(S_0()) = S_2(\epsilon) = 10$.
  Having shown $P(\epsilon)$, this is handled by the third case of the induction proof.
- $S_3(S_1(S_0()), S_2(S_0())) = S_3(01, 10) = 0110$.
  Having shown $P(01)$ and $P(10)$, this is handled by the fourth case of the induction proof.
- $S_1(S_1(S_0()), S_2(S_0())) = S_1(0110) = 001101$.
  Having shown $P(0110)$, this is handled by the second case of the induction proof.

Note that there is another derivation of w by the rules for S. If you want some practice, you can figure out that derivation, and then trace how the induction proof applies to that alternative derivation of w as well.

2. $(w \in S) \Leftarrow P(w)$:

Let $w \in \Sigma^*$ be any string for which $P(w)$ holds. Our proof is by induction on the length of w. We note that the length of $length(w) = nZero(w) + nOne(w) = 2nZero(w)$. Therefore, $length(w)$ must be even. We consider five cases:

$w = \epsilon$: $w \in S$ by the first case in the definition of S.

$w = 0x1$: $nZero(x) = nZero(w) - 1 = nOne(w) - 1 = nOne(x)$. Thus, $P(x)$. Furthermore, $length(x) = length(w) - 2$, in particular $length(x) < length(w)$; so, we can apply the induction hypothesis to conclude $x \in S$. Thus, $w \in S$ by the second case in the defintion of S.

$w = 1x0$: The proof is equivalent to the one given for the previous case.

$w = 0x0$: This is the one where we have to think a little harder. We consider the number of 0's and 1's in x:

$$\begin{array}{lrlll} 1. & nZero(x) & = & nZero(w) - 2, & w = 0x0 \\ 2. & nOne(x) & = & nOne(w), & w = 0x0 \\ 3. & nZero(w) & = & nOne(w), & P(w) \\ 4. & nOne(x) & = & nZero(x) + 2 & 1, 2, 3 \end{array}$$

We conclude that there must be a prefix of x that has one more 1 than 0.

More formally, we can define

$$\begin{aligned} \mathit{prefix}(x, 0) &= \epsilon \\ \mathit{prefix}(c \cdot x, n) &= c \cdot \mathit{prefix}(x, n-1) \end{aligned}$$

In English, $\mathit{prefix}(x, n)$ is the string consisting of the first n symbols of x. By the definition of prefix, $nOne(\mathit{prefix}(x, 0)) - nZero(\mathit{prefix}(x, 0)) = 0$. It was shown above that $nOne(\mathit{prefix}(x, length(x))) - nZero(\mathit{prefix}(x, length(x))) = 2$. Let k be the smallest number such that $nOne(\mathit{prefix}(x, k)) - nZero(\mathit{prefix}(x, k)) > 0$. It is straightforward to show that $nOne(\mathit{prefix}(x, k)) - nZero(\mathit{prefix}(x, k)) = 1$. Now, let $u = \mathit{prefix}(x, k)$. We can define $\mathit{suffix}(x, n)$ to be the last n symbols of x (the definition is similar to that of prefix – can you write it down?). Let $v = \mathit{suffix}(x, length(x) - k)$. Clearly, $x = uv$. Furthermore $P(0u)$, $P(v0)$, and $w = 0uv0$. We can apply the induction hypothesis to conclude $0u \in S$ and $v0 \in S$. Therefore, $w \in S$ by using the fourth case in the definition of S.

$w = 1x1$: The proof is equivalent to that for the previous case.

This completes the proof for $(w \in S) \Leftarrow P(w)$.

Having shown $(w \in S) \Rightarrow P(w)$ and $(w \in S) \Rightarrow P(w)$, we conclude that a string, w is in S the number of 0's in w is equal to the number of 1's.

□