COMS 331: Theory of Computing, Spring 2023
Homework Assignment 4

Due at 10:00PM, Wednesday, February 22, on Gradescope.

Problem 23. Call a string $x \in \{0,1\}^*$ *prefix-balanced* if, for every prefix w of x,

$$|\#(0,w) - \#(1,w)| \leq 1.$$

Design a DFA M such that

$$L(M) = \{x \in \{0,1\}^* \mid x \text{ is prefix-balanced and } \#(0,x) = \#(1,x)\}.$$

Problem 24. Design a DFA M that decides the language

$$A = \{x \in \{0,1\}^* \mid \#(1,x) \text{ is odd and } (bnum(x) \text{ is a multiple of 2 or a multiple of 3})\}.$$

Define a *finite-state compressor (FSC)* to be a 4-tuple

$$C = (Q, \delta, \nu, s),$$

where

- Q is a finite set of states;
- $\delta : Q \times \{0,1\} \longrightarrow Q$ is the transition function;
- $\nu : Q \times \{0,1\} \longrightarrow \{0,1\}^*$ is the output function;
- $s \in Q$ is the start state.

Define the extended transition function $\hat{\delta} : Q \times \{0,1\}^* \longrightarrow Q$ as for DFAs. Define the output from state $q \in Q$ on input string $w \in \{0,1\}^*$ to be the string $\hat{\nu}(q,w)$ defined by the recursion

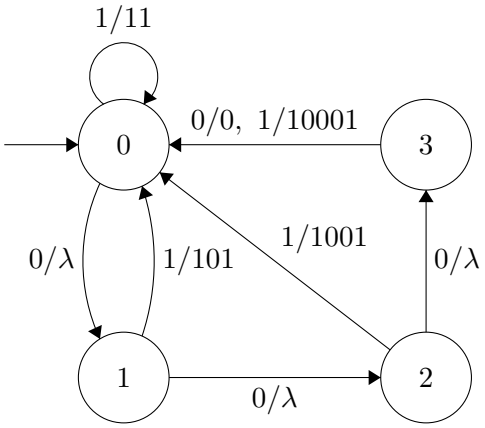
$$\begin{aligned} \nu(q,\lambda) &= \lambda \\ \hat{\nu}(q,wb) &= \hat{\nu}(q,w)\nu(\hat{\delta}(q,w),b) \end{aligned}$$

for all $q \in Q$, $w \in \{0,1\}^*$, and $b \in \{0,1\}$.

Finally, define the output of the FSC C on input $w \in \{0,1\}^*$ to be the string

$$C(w) = \hat{\nu}(s, w).$$

Example



The diagram above denotes the FSC $C = (Q, \delta, \nu, s)$, where

- $Q = \{0, 1, 2, 3\}$;
- $\delta(q, 0) = (q + 1) mod 4$ for all $q \in Q$;
- $\delta(q, 1) = 0$ for all $q \in Q$;
- $\nu(q, 0) = \begin{cases} \lambda, & \text{if } q \leq 2 \\ 0, & \text{if } q = 3 \end{cases}$
- $\nu(q, 1) = 10^q 1$ for all $q \in Q$.

Note that each transition arrow is labeled by b/y, where b is the input bit producing the transition and y is the output string produced by the transition.

Intuitively, an FSC *compresses* an input string w if $|C(w)|$ is significantly less than $|w|$.

Problem 25. For the specific example FSC C above, give formulas for $|C(0^n)|$ and $|C(1^n)|$ in terms of n. Intuitively, what kinds of strings does C compress?

Problem 26. For the specific example FSC C above, assume that $C(w) = 00101$.

(a) What are the possible values of $\hat{\delta}(s, w)$?

(b) For each of the states in (a), what is the value of w?

An FSC $C = (Q, \delta, \nu, s)$ is *information lossless (IL)* if the function

$$g : \{0,1\}^* \longrightarrow \{0,1\}^* \times Q$$

defined by

$$g(w) = (C(w), \hat{\delta}(s, w))$$

is one-to-one.

Problem 27.

(a) Explain why our example FSC is IL.

(b) Give an example of an FSC that is *not* IL.

Problem 28. Give an example of an IL FSC C such that $|C(w)| < |w|$ holds for every string $w \in \{0,1\}^+$.

Problem 29. Prove that, for every IL FSC $C = (Q, \delta, \nu, s)$ and every $n \in \mathbb{N}$, there is an input string $w \in \{0,1\}^n$ such that $|C(w)| \geq n - \log_2 |Q|$.