Home Work 1: Theory of Computation
October 26, 1998
Bud Mishra
Due On: Oct 28 1998

1 Logic

Problem. 1
It is year 2001! It has been TEOTWAWKI (the end of the world as we knew it). The Y2K bug has brought an end to the human civilization. Fortunately, the computers are not dead! There are two kinds of computers that have survived a Darwinian struggle: Bill's computers (manufactured by William "Bill" Gates of Microsoft) and Gill's computers (manufactured by Guglielmo "Gill" Bates of Megahard). These machines are claimed to give correct answers to any yes/no questions instantly. The output is given by flashing a green or a red light—one for yes, the other for no; unfortunately, it is not known which one is which.

Bill's and Gill's computers have identical appearance, but they differ in one aspect, they use different colors for the yes/no answers. For instance, if Bill's computers use green for yes then Gill's computers use red for yes and vice versa.

Your job is the following:

1. You are shown a machine. You are allowed to ask just one question to find out whether it is a Bill's machine or a Gill's machine. What question would you ask?
2. A machine is called "natural" if it uses green for yes. Your job is to just ask one question to determine which kind of machine is natural. What question would you ask?
3. You are shown a machine. You will get it for free and fifty cows, if you can ask it one yes/no question that it cannot answer. What question would you ask?

Solution. 1 1. "Do Bill's machines answer yes with green?" If Bill's machine uses green for yes, then it will flash green. If Bill's machine uses red for yes, then it will flash green. Bill's machine will always flash green. If Gill's machine uses green for yes, then it will flash red. If Gill's machine uses red for yes, then it will flash red. Gill's machine will always flash red. Green means it's a Bill's machine; red means it's a Gill's machine.

2. "Are you a Bill's machine?" If Bill's machine uses green for yes (and thus natural) then a Bill's machine will flash green (for yes) and a Gill's machine will flash green (for no). If Gill's machine uses green for yes then Bill's machine will flash red (for yes) and Gill's machine will flash red (for no). Green means Bill's machine is natural; red means Gill's machine is natural.

3. "Does the color with which you answer this question stand for no?" Suppose it flashes a green. If green stands for yes, then the answer is wrong as red stands for no. If green stands for no, then the answer is wrong since green in fact stands for no and the answer to the question is yes which is encoded by red. Similarly if it flashes red, there is a contradiction.

2 Finite Automata

Problem. 2

1. Exercise 1.15 (Sipser, pp 86).
2. Exercise 1.25 (Sipser, pp 88).

Solution. 2 Omitted.

3 Pumping Lemma

Problem. 3

1. Consider the language consisting of the set of all strings with sequence of 0's followed by the same number of 1's.

$$\mathcal{L} = \{o^n 1^n | n \geq 0\}.$$

   Show that it's not regular.

2. Using a homomorphism reduction show that the set of palindromes $\mathrm{PAL}(\Sigma)$ over $\Sigma = \{0, 1\}$ is not regular either. (DO NOT USE THE PUMPING LEMMA DIRECTLY.)

$$\mathrm{PAL}(\{0,1\}) = \{\sigma \in \{0,1\}^* | \sigma = \sigma^R\}.$$

3. Consider the set of reverse palindromes $\mathrm{RPAL}(\Sigma)$ over $\Sigma = \{A, T, C, G\}$, defined as follows: The complementation operation is a homomorphism defined by $A^C = T$, $T^C = A$, $C^C = G$, and $G^C = C$. Show that the language $\mathrm{RPAL}(\Sigma)$ is not regular. (Again, DO NOT USE THE PUMPING LEMMA DIRECTLY.)

$$\mathrm{RPAL}(\{A,T,C,G\}) = \{\sigma \in \{A,T,C,G\}^* | \sigma^R = \sigma^C\}.$$

Solution. 3 1. Consider the string $uvw = 0^n 1^n$ for some fixed n. Since $|uv| < n$ and $|v| > 0$, we have $v = 0 \cdots 0$ and $|v| < n$. But then, $uw = 0^{n-|v|}1^n$ (the string uv^iw, with $i = 0$) must be in the same language. An impossibility.

2. Let $\mathcal{L}' = \mathrm{PAL}(\{0, 1\})$. Consider two homomorphisms:

$$\begin{aligned} h_1 &: \{a, b, c\}^* \to \{0, 1\}^* \\ &: a \mapsto 0; b \mapsto 1; c \mapsto 0 \\ h_2 &: \{a, b, c\}^* \to \{0, 1\}^* \\ &: a \mapsto 0; b \mapsto \epsilon; c \mapsto 1 \end{aligned}$$

Now consider the language

$$\begin{aligned} & h_2[h_1^{-1}(\mathcal{L}' \cap 0^*10^*) \cap a^*bc^*] \\ &= h_2[h_1^{-1}(\{0^n 1 0^n | n \geq 0\}) \cap a^*bc^*] \\ &= h_2[\{a^n b c^n | n \geq 0\}] \\ &= \{0^n 1^n | n \geq 0\} = \mathcal{L}] \end{aligned}$$

Since 0^*10^* and a^*bc^* are regular and regular sets are closed under intersection and homomorphisms, we conclude that if $\mathrm{PAL}(\{0, 1\})$ is regular then so is $\{0^n 1^n | n \geq 0\}$.

3. Consider the following surjective homomorphism h

$$\begin{aligned} h &: \{A, T, C, G\}^* \to \{0, 1\}^* \\ &: A \mapsto 0; T \mapsto 0; C \mapsto 1; G \mapsto 1 \end{aligned}$$

Thus $h(X^C) = h(X) = 0$, if $X \in \{A, T\}$ and $h(X^C) = h(X) = 1$, if $X \in \{C, G\}$. Thus for all $\sigma \in \{A, T, C, G\}^*$, $h(\sigma^C) = h(\sigma)$. And for all $\sigma \in \mathrm{RPAL}(\{A, T, C, G\})$, $h(\sigma^C) = h(\sigma) = h(\sigma)^R$. Thus

$$h(\mathrm{RPAL}(\{A, T, C, G\})) = \mathrm{PAL}(\{0, 1\} = \mathcal{L}',$$

which is non-regular.

4 Miscellaneous [HARD]

Problem. 4 Let $\mathcal{L} \subseteq 0^*$. Prove that $\mathcal{L}^*$ is regular.

Solution. 4 Let us assume $0^{k_1}, 0^{k_2}, \ldots, 0^{k_i}, \ldots$ are the members of the language $\mathcal{L}$, where 0^{k_i} stands for a string of 0's of length k_i. Furthermore, we assume that the strings are enumerated in the increasing order of length:

$$0 \leq k_1 < k_2 < \cdots < k_i < \cdots$$

We enumerate the strings of $\mathcal{L}^*$, by breaking them into following sets:

$$\begin{aligned} S_1 &= \{\epsilon, 0^{k_1}, 0^{2k_1}, 0^{3k_1}, \ldots\} \\ S_2 &= \{0^{k_2}, 0^{k_2+k_1}, 0^{k_2+2k_1}, \ldots\} \\ &\qquad \text{iff } 0^{k_2} \notin S_1 \\ &\vdots \\ S_n &= \{0^{k_{n'}}, 0^{k_{n'}+k_1}, 0^{k_{n'}+2k_1}, \ldots\} \\ &\qquad \text{iff } 0^{k_{n'}} \notin \bigcup_{i=1}^{n-1} S_i \\ &\vdots \end{aligned}$$

Note that $\mathcal{L}^*$ can have at most k_1 disjoint infinite subsets, since for a large enough i and $0 \leq j \leq k_1 - 1$, if $0^{j+ik_1} \in \mathcal{L}^*$ then it must belong to one of the S_i's.

Since each S_n can be expressed as a regular expression of the form $0^{k_{n'}}(0^{k_1})^*$ and $\mathcal{L}^*$ can be expressed a finite set union of k_1 regular expressions, $\mathcal{L}^*$ is a regular set.

Another Proof

Assume that $\mathcal{L}$ is an infinite set, since otherwise $\mathcal{L}$ would be regular and so would be $\mathcal{L}^*$ (closed under Kleene star operation).

As before, we assume that the strings 0^{k_1}, 0^{k_2}, ..., 0^{k_i}, ... of $\mathcal{L}$ are enumerated in the increasing order of length:

$$0 \leq k_1 < k_2 < \cdots < k_i < \cdots$$

We may also assume that k_1, k_2, etc. are coprime. It suffices only to prove that $\mathcal{L}^*$ is regular in this case[1].

Let

$$A_L = (k_1, k_2, \ldots, k_i, \ldots).$$

By the *finiteness theorem* for the gcd, we know that the gcd of the infinite set A_L can be obtained as the gcd of a finite subset

$$A_0 = (k_{i_1}, k_{i_2}, \ldots, k_{i_n}) \subset A_L.$$

By the *principle of coprimality*, we know that

$$\exists x_{i_1}, x_{i_2}, \ldots, x_{i_n} \,.\, x_{i_1}k_{i_1} + x_{i_2}k_{i_2} + \cdots + x_{i_n}k_{i_n} = 1. \qquad (1)$$

We claim that all the strings of length greater than or equal to

$$r = k_{min}(|x_{i'_1}|k_{i'_1} + |x_{i'_2}|k_{i'_2} + \cdots + |x_{i'_m}|k_{i'_m})$$

[1] If k_i's are not coprime, then we can define a homomorphism $h : 0^d \mapsto a$, where $d = \gcd k_1, k_2, \ldots$, such that $\mathcal{L}' = h(\mathcal{L}) \subseteq a^*$ and has the desired properties. Since the regular sets are closed under homomorphisms, the proof remains valid independent of the value of d.

is in $\mathcal{L}^*$. In the above expression,

$$\begin{aligned} k_{min} &= \min(k_{i_1}, k_{i_2}, \ldots, k_{i_n}) \\ & \quad x_{i'_1}, x_{i'_2}, \ldots, x_{i'_m} \text{ are all negative coefficients in eq 1 .} \end{aligned}$$

Assume that $k_{min} = k_{i_1}$, without loss of generality. Clearly, 0^s, $s \geq r$, is in $\mathcal{L}^*$. Write s as

$$\begin{aligned} s &= r + j(k_{min}) + i, \quad j \geq 0, 0 \leq i < k_{min} \\ &= k_{min}(j + |x_{i'_1}|k_{i'_1} + |x_{i'_2}|k_{i'_2} + \cdots + |x_{i'_m}|k_{i'_m}) \\ &\quad + i(x_{i_1}k_{i_1} + x_{i_2}k_{i_2} + \cdots + x_{i_n}k_{i_n}) \\ &= y_{i_1}k_{i_1} + y_{i_2}k_{i_2} + \cdots + y_{i_n}k_{i_n}. \end{aligned}$$

In this expression, either $y_{i_1} = ix_{i_1} + j$, if the x coefficient in eq 1 is positive, and $y_{i_1} = (i - k_{min})x_{i_1} + j$ if the x coefficient is negative. Similarly, for $p > 1$, $y_{i_p} = ix_{i_p}$, if the x coefficient in 1 is positive, and $y_{i_p} = (i - k_{min})x_{i_p}$ if the x coefficient is negative. In all cases, y's are positive, and

$$0^s = (0^{k_{i_1}})^{y_{i_1}}(0^{k_{i_2}})^{y_{i_2}} \cdots (0^{k_{i_n}})^{y_{i_n}} \in \mathcal{L}^*.$$

Since r is finite, we see that

$$\mathcal{L}^* \equiv R_1 + R_2,$$

where $R_2 = (0^r)0^*$ and R_1 is a finite set of strings of length less than r. Thus $\mathcal{L}^*$ is regular.