

CS3133

HW#1

DUE: Tuesday, August 31

1. (14 points) Let L be the language defined by:

- $\varepsilon \in L$.
- If $z \in L$ and z can be written as uvw , then $\{u0v1w, u1v0w\} \subseteq L$.
- A string belongs to L only if it can be derived by a finite number of applications of the above rules.

a (3 points) Prove that L only contains strings of even length.

b (3 points) Describe L . Do not just repeat the definition above, but describe in precise English what the strings in L look like.

c (8 points) Prove your answer in part **b**.

EXTRA CREDIT 1: Give an optimal algorithm to test for membership in L . Prove that your algorithm is optimal.

EXTRA CREDIT 2: If you draw a binary string z from a uniform distribution over the set of all binary strings of length $|z|$, then what is $\Pr\{z \in L\}$? Note that half the binary strings of length 2 belong to L .

2. (10 points) Prove or give a counterexample to each of the following conjectures about languages L_1 and L_2 .

CONJECTURE 1: If $L_1 \subseteq L_2$, then $L_1^* \cup L_2^* \subseteq (L_1 \cup L_2)^*$.

CONJECTURE 2: If $L_1 \subseteq L_2$, then $(L_1 \cup L_2)^* \subseteq L_1^* \cup L_2^*$.

3. (8 points) Prove that each of the following languages over $\{0, 1\}^*$ is regular.

a The set of all strings that do not contain 00. We note that 110101 belongs to the language but 0100011 does not.

b The set of all strings such that the number of 0's plus twice the number of 1's is divisible by 4. So 01110 belongs to the language but 011 does not.

CS3133
Solutions to HW#1

1. **a** Clearly L doesn't contain a string of length 1. Assume that z is a shortest string of odd length in L . The only way for z to belong to L is for there to be a string $uvw \in L$ and either $z=u0v1w$ or $z=u1v0w$. But $|uvw| < |z|$ and uvw is of odd length, contradicting the assumption above. So L can not contain strings of odd length.

b L is the set of all binary strings with the same number of 0's as 1's.

c Claim: Every $z \in L$ has the same number of 0's as 1's.

Proof: If $|z|=0$, then $z = \varepsilon$ and it has the same number of 0's as 1's. We know from part **a** that all strings in L are of even length. Assume that for all some n , all strings in L of length $2n$ have the same number of 0's as 1's. Consider some string $z \in L$ of length $2(n+1)$. In deriving z , the last application of the second rule was applied to a string uvw of length $2n$. By the induction hypothesis, uvw has n 0's and 1's. So z must have $(n+1)$ 0's and 1's.

Claim: Every string over $\{0,1\}^*$ with the same number of 0's as 1's belongs to L .

Proof: If $|z|=0$, then $z = \varepsilon$ and it belongs to L . Assume that for some n , every string with n 0's and 1's belongs to L . String z with $(n+1)$ 0's and 1's can be written as b_1vb_2w , $b_1, b_2 \in \{0,1\}$, $b_1 \neq b_2$, $v, w \in \{0,1\}^*$. That is, b_1 is the first bit of z and b_2 is any occurrence of the other bit, and v is the string separating them. By the induction hypothesis, $vw \in L$. One more application of the second rule yields z , which thus belongs to L .

EXTRA CREDIT 1: The following algorithm tests if $a_1 \dots a_n \in L$.

```

count ← 0
for k ← 1 to n do
    if  $a_k = 0$  then count ← count + 1
    else count ← count - 1
if count = 0 then return  $a_1 \dots a_n \in L$ 
else return  $a_1 \dots a_n \notin L$ 

```

Clearly the time complexity of the algorithm is $\Theta(n)$. To show that it is optimal, we must show that there does not exist an algorithm to test for membership in L with lower time complexity. If there were such an algorithm, then it could not examine every bit of $a_1 \dots a_n$. Choose $a_1 \dots a_n \in L$, and let a_i , $1 \leq i \leq n$, be a bit that wasn't examined. Flip a_i and the new string does not belong to L . But the purported algorithm must do the same tests on the new string and it will get the same answers. Hence it must respond that the new string belongs to L . This contradiction proves that the purported algorithm can not exist.

EXTRA CREDIT 2: If $|z|$ is odd, then $\Pr\{z \in L\} = 0$.

Let $|z|$ be even. We note that there are $2^{|z|}$ binary strings of length $|z|$. Of these strings, there are $\binom{|z|}{|z|/2}$ strings of length $|z|$ in L since there are that many ways to choose $|z|/2$ positions for the 0's (choosing positions for the 0's fixes the positions for the 1's). So

$$\Pr\{z \in L\} = \frac{\binom{|z|}{|z|/2}}{2^{|z|}} .$$

2. CONJECTURE 1 is true. If $z \in L_1^* \cup L_2^*$, then there exist strings $x_1, \dots, x_n \in L_1$ or $x_1, \dots, x_n \in L_2$ such that $z = x_1 \dots x_n$. But in either case, $x_1, \dots, x_n \in L_1 \cup L_2$, implying that $x_1, \dots, x_n \in (L_1 \cup L_2)^*$.

CONJECTURE 2 is true. If $L_1 \subseteq L_2$, then $L_1 \cup L_2 = L_2$, and $(L_1 \cup L_2)^* = L_2^* = L_1^* \cup L_2^*$.

3. **a** The language is accepted by the DFA $(\{s, p, q\}, \{0, 1\}, \delta, s, \{s, p\})$ where

$$\begin{array}{ccc} \delta & 0 & 1 \\ s & p & s \\ p & q & s \\ q & q & q \end{array} .$$

b In the following DFA to accept the language, the "meaning" of state $q_i, 0 \leq i \leq 3$, is that the number of 0's read so far plus twice the number of 1's read so far is $i \pmod{4}$.

$(\{q_0, q_1, q_2, q_3\}, \{0, 1\}, \delta, q_0, \{q_0\})$, where

$$\begin{array}{ccc} \delta & 0 & 1 \\ q_0 & q_1 & q_2 \\ q_1 & q_2 & q_3 \\ q_2 & q_3 & q_0 \\ q_3 & q_0 & q_1 \end{array} .$$