CS3133

HW#2

DUE: Tuesday, September 4

1.(6 points) Prove or give a counterexample to each of the following. If you give a counterexample, justify your answer.

CONJECTURE A: For any languages L_1, L_2,

$$L_1^* = L_2^* \text{ if and only if } L_1 = L_2.$$

CONJECTURE B: If any regular language L_1 over Σ^* is infinite, then there exists a finite language L_2 over Σ^* such that

$$L_1 = \Sigma^* - L_2 = \{w \in \Sigma^* \mid w \notin L_2\}.$$

As an example that may motivate a belief in CONJECTURE B, it is easy to see that $L_1 = \{a\}^*$ over $\Sigma = \{a\}$ is regular. Choosing $L_2 = \varnothing$, it follows that $L_1 = \{a\}^* = \{a\}^* - \varnothing = \Sigma^* - L_2$.

2. (10 points) It's easy to see how we can use a DFA as a spelling checker. However, we want a spelling checker which detects sumple speling ergors. We define the *neighborhood* of any word $x_1 \ldots x_n$ as $\eta(x_1 \ldots x_n) = \{y_1 \ldots y_n \mid \exists k \forall j (j \neq k \rightarrow x_j = y_j)\}$. That is, $\eta(w)$ is the set of all words which are the same length as w and which differ from w in at most one letter. Prove that for any regular language L, the language $\bigcup_{w \in L} \eta(w)$ is also regular. As an example, if $L = \{0, 11\}$ over the alphabet $\{0,1\}$, then $\bigcup_{w \in L} \eta(w) = \{0,1,11,01,10\}$. Note that this could be generalized to more than one erggr.

3. (8 points) We say that string z is a *substring of* string w if there exist strings α, β such that $w = \alpha z \beta$. Prove or give a counterexample to the following.

CONJECTURE: For any regular language L, the language $\{w \mid \exists z \in L (z \text{ is a substring of } w)\}$ is regular.

4. For any language L we define a new language *NoPrefix*(L) to be

$$\{w \in L \mid \neg \exists x (x \neq \varepsilon \rightarrow wx \notin L)\}.$$

So if $L = \{010, 100, 01010000\}$ then $NoPrefix(L) = \{100,\ 01010000\}$.

a (2 points) Describe an infinite regular language L such that $NoPrefix(L)=L$.

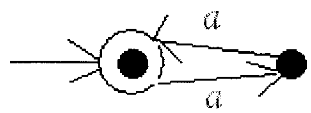
b (3 points) Describe an infinite regular language L such that $NoPrefix(L)=\varnothing$.

c (8 points) Prove that if L is regular then $NoPrefix(L)$ is regular.

CS3133

Solutions to HW#2

1. Both CONJECTUREs are false. As a counterexample to CONJECTURE A, letting $L_1 = \{a\}$ and $L_2 = \{a, aa\}$, then $L_1 \neq L_2$ even though $L_1^* = L_2^* = \{\varepsilon, a, aa, aaa, \ldots\}$. As a counterexample to CONJECTURE B, let $L_1 = \{aa\}^*$. It is easily seen that L_1 is regular:



If $L_1 \subseteq \Sigma^*$ then $a \in \Sigma$. So for any choice of Σ, $\{a^i \mid i \, \text{odd}\} \subseteq \Sigma^* - L_1$. Since $\{a^i \mid i \text{ odd}\}$ is infinite, it does not belong to any finite set.

2. If L is regular, then it is accepted by a DFA $M=(\{q_0,\ldots,q_m\},\Sigma,\delta,q_0,F)$. The idea behind accepting $\bigcup_{w\in L}\eta(w)$ is to have two similar copies, M_1 and M_2, of $(\{q_0,\ldots,q_m\},\Sigma,\delta,q_0,F)$. The computation starts and proceeds in M_1. In reading any input symbol, the computation both traces through M_1, thus accepting each word of L, and it jumps nondeterministically from a state in M_1 to all states in M_2 analogous to a state reachable by a transition in M_1 upon reading a symbol of Σ (this corresponds to the error symbol). So $\bigcup_{w\in L}\eta(w)$ is accepted by the NFA $(\{q_0^1,\ldots,q_m^1,q_0^2,\ldots,q_m^2\},\Sigma,\delta^*,q_0^1,F^*)$. In defining δ^*, we first define for each $q \in \{q_0,\ldots,q_m\}$ the set of states reachable from q over one edge, $\rho(q) = \{\delta(q,a) \mid a \in \Sigma\}$. Then for each $q_i \in \{q_0,\ldots,q_m\}$ and each $a \in \Sigma$, $\delta^*(q_i^1,a) = \{q^1_{\delta(q_i,a)}\} \cup \{q_j^2 \mid q_j \in \rho(q_i)\}$ and $\delta^*(q_i^2,a) = \{q^2_{\delta(q_i,a)}\}$, and $F^* = \bigcup_{q_i\in F}\{q_i^1,q_i^2\}$.

3. The CONJECTURE is true. If L is regular, then there exists a DFA (Q,Σ,δ,q_0,F) to accept it. We now construct a DFA to accept $\{w \mid \exists z \in L(z \text{ is a substring of } w)\}$. The idea behind accepting $\alpha z \beta$ is to guess at α by consuming symbols of Σ nondeterministically while staying in q_0. Once we reach a final state from F after reading z, we just consume the rest of the input, β, while staying in that state. So $\{w \mid \exists z \in L(z \text{ is a substring of } w)\}$ is accepted by the NFA $(Q,\Sigma,\delta^*,q_0,F)$ where, for all $a \in \Sigma$, $\delta^*(q_0,a) = \{q_0\} \cup \{\delta(q_0,a)\}$, $\delta^*(q,a) = \{q\}$ for all $q \in F$, and $\delta^*(q,a) = \{\delta(q,a)\}$ for all $q \in Q - (\{q_0\} \cup F)$.

4. ***a*** 0^*1

b 0^*

c If L is regular, then there is a DFA $M=(Q,\Sigma,\delta,q_0,F)$ to accept L. Construct $M'=(Q,\Sigma,\delta,q_0,F')$ from M by removing from F all states q from which there is a path, of length at least 1, from q to a state in F. We now show that $L(M')=NoPrefix(L(M))$. Since $F'\subseteq F$, it follows that $L(M')\subseteq L(M)$. If a string w is accepted by finishing in a state $q\in F-F'$, then let x be the labels along a path from q to a state of F-F'. So $wx\in L$ and $w\notin NoPrefix(L)$. My apologies for the double negation in $w\notin NoPrefix(L)$ which I could not remove.