CS 4510: Automata and Complexity

Spring 2015

Extra Home work // Due: Friday, March 27, 2015

1. (10 points) Problem 1.48, page 90 of the text.

   Show that the following language is regular:

   $$L = \{w \in \{0,1\}^* \mid w \text{ contains an equal number of occurrences of the substrings 01 and 10}\}.$$

   Thus, $101 \in L$ because 101 contains a single 01 and a single 10. The string $1010 \notin L$ because 1010 contains two 10s and one 01.

   Solution: L is the set of all strings that start and end with the same symbol.

2. (10 points) Problem 2.28 (b), page 157 of the text.

   Give an unambiguous grammar for the following language:

   $$\{w \in \{0,1\}^* \mid \text{number of } 0's = \text{ and number of } 1's\}.$$

   Solution:

   $$\begin{array}{rcl} S & \rightarrow & 0A1S \mid 1B0S \mid \epsilon \\ A & \rightarrow & 0A1A \mid \epsilon \\ B & \rightarrow & 1B0B \mid \epsilon \end{array}$$

3. (15 points) Problem 2.37, page 158 of the text.

   Prove the following stronger form of the pumping lemma, wherein *both* pieces v and y must be nonempty when the string s is broken up.

   If A is a context-free language, then there is a number k where, if s is any string in A of length at least k, then s may be divided into five pieces, $s = uvxyz$, satisfying the consitions:

   (a) for each $i \geq 0$, $uv^ixy^iz \in A$,

   (b) $v \neq \epsilon$ and $y \neq \epsilon$, and

   (c) $|vxy| \leq k$.

   (Assume that A is generated by a CFG $G = (V, \Sigma, R, S)$ in CNF.)

   Solution: Choose $k = 2^{2|V|+1}$. Let s be any string in A with length at least k. Consider a *smallest* size parse-tree T for the string s. The depth of T must be at least $2|V| + 1$. Choose a path P in T that has length at least $2|V| + 1$. Then, it must be the case that a variable R appears at least three times in P.

   Let R be a variable that occurs at least 3 times among the lowest $2|V| + 1$ modes in P. Let t denote the leaves of the subtree of T rooted at the lowest of these R's. Note that t is not empty. Then, the leaves of the subtree of T rooted at the second of these R's can be written as the concatenation of three strings s_1, t, and s_2 as s_1ts_2. The leaves of the subtree of T rooted at the top most R can be written as the concatenaion of five strings r_1, s_1, t, s_2 and r_2 as $r_1s_1ts_2r_2$. Now, it must be the case that both s_1 and s_2 cannot be empty. Otherwise, we can substitute the tree rooted at the lowest R as the tree for the middle R to get a smaller sized parse-tree. Simlilarly, it must be the case that both r_1 and r_2 cannot be empty.

   Case 1: r_1 and s_1 are empty: Then, $v = t$, $x = s_2$ and $y = r_2$.

   Case 2: r_2 and s_2 are empty: Then, $v = r_1$ (not empty), $x = s_1$ and $y = t$.

   Other cases can be worked out similarly to show that conditions of the pumping lemma are satsified.

4. (15 points) Given a CFG G and a DFA D, describe an algorithm to decide if $L(G) \subseteq L(D)$.

   Hint: We saw in class/text algorithms for the following problems:

   - Constructing a DFA that recognizes the complement of the language of a given DFA.

- Constructing a PDA from a given CFG.
- Constructing a CFG from a given PDA.
- Constructing a PDA that recognizes the intersection of the languages of a given PDA and a given DFA.
- Deciding if the language of a given CFG is empty.

Solution: It is true that $L(G) \subseteq L(D)$ iff $L(\bar{D}) \cap L(G) = \emptyset$.

Here is an algorithm for the given problem:

- Construct the DFA D' that recognizes $L(\bar{D})$ from the DFA D.
- Construct the PDA P from the CFG G.
- Construct the PDA P' that recognizes $L(P) \cap L(D')$.
- Construct the CFG G' from the PDA P'.
- Decide if $L(G) = \emptyset$.

5. Define a *Constrained Two-Headed* (CTH) machine as follows.

   An CTH machine is a DTM with one tape that is unbounded on one side and that has two heads FH and BH.

   - The head FH is a read-only head and the head BH is a write-only head.
   - Both heads can only move right.
   - If a symbol on the tape under the read head FH is read, the head FH moves right by one position.
   - If a symbol is written on the tape under the write head BH, the head BH moves right by one position.
   - In each step, the machine M being in a state and either reading the symbol on the tape under the head FH or not reading the symbol can: (a) change state, and (b) either write a symbol on the tape or not write a symbol.
   - The string on the tape in positions i through j is said to be *current* if FH is on position i and BH is on position $j+1$.

   Let $M = (Q, \Sigma, \Gamma, \delta_M, q_s, q_a, q_r)$ be a 1-tape deterministic Turing machine. Our goal is to simulate M by a RTH machine.

   Let N be the simulating RTH machine. For each symbol $a \in \Gamma$ of M, there are two symbols a, a° in the tape alphabet of N. The symbol a° denotes that the tape head of M is over the symbol a.

   Let $w = w_1 \cdots w_n$ be a bit string that is an input to M. A configuration of M on input w is represented using the current state of N and the current region of N's tape in the form: $u_1 u_2 \cdots u_i{}^\circ \cdots u_k\$$.

   (a) (5 points) How is the input $w_1 w_2 \cdots w_n$ represented in N?

   Solution: $w_1{}^\circ w_2 \cdots w_n$. The head FH is pointing to the symbol $w_1{}^\circ$ and the head BH is pointing to the blank symbol $\sqcup$ after w_n.

   (b) (10 points) Let $\delta_M(p, a) = (q, b, R)$ where $a, b \in \{0, 1\}$. Show how to simulate this move with a RTH machine.

   Solution: Let the current region of N be: $u_1 u_2 \cdots u_i{}^\circ \cdots u_k\$$. The head FH is pointing to the symbol u_1 and the head BH is pointing to the blank symbol $\sqcup$ after the \$ symbol of the current region. The machine N executes the following steps:

```
WHILE the symbol under the FH head is not $ DO
     IF this symbol is not marked THEN write it on the tape
     ELSE write the new symbol dictated by the transition on the tape
          IF the symbol under the FH head is not $ THEN
               write a marked version of the symbol on the tape
          ELSE write a marked blank symbol on the tape
               write a $ on the tape
               Halt

Write $ on the tape
Halt
```

(c) (10 points) Let $\delta_M(p, a) = (q, b, L)$ where $a, b \in \{0, 1\}$. Show how to simulate this move with a RTH machine.

Solution: Let the current region of N be: $u_1 u_2 \cdots u_i{}^\circ \cdots u_k\$$. The head FH is pointing to the symbol u_1 and the head BH is pointing to the blank symbol $\sqcup$ after the \$ symbol of the current region. The machine N executes the following steps:

```
IF the symbol under the head FH is marked THEN
   Write the marked version of the new symbol dictated by the transition
   WHILE the symbol under the FH head is not $ DO
         write this symbol

Write $

IF the symbol under the head FH is unmarked THEN
   store the symbol in the state

WHILE the symbol under the FH head is not $ DO
     IF this symbol is not a marked symbol THEN
           store this symbol in the state
           write the symbol stored in the state
     ELSE
           store the unmarked version of this symbol in the state
           write the marked version of the symbol stored in the state


Write the symbol stored in the state
Write $
Halt
```