COMP 330 - Fall 2016 - Assignment 3

Due 7:00 pm Oct 28, 2016

General rules: In solving these questions you may consult books but you may not consult with each other. There are in total 110 points, but your grade will be considered out of 100. You should either drop your solutions in the assignment drop-off box located in the Trottier Building on the 2nd floor, or submit it through mycourses. Late homeworks can be submitted until 48 hours after the deadline. Late submissions must be submited through mycourses as the drop-off box will not be checked after the deadline. There will be a penalty of -25% for one-day delays, and -40% for two-day delays on late homeworks.

1. (20 points) Prove that $\{0^m 1^n \mid m, n \geq 1 \text{ and } m \text{ is divisible by } n\}$ is not context-free.

   Answer: Suppose it is, and let p be the pumping constant. Consider the string $w = 0^{q^2} 1^q$ for $m = 100p^2$ (We are picking a large enough value for m), which belongs to the language. By pumping lemma there exists a decomposition $w = uvxyz$ such that $|uy| > 0$, $k = |vxy| \leq p$, and moreover $uv^i xy^i z$ is in the language for all i.

   - If vxy consists of all 1's, then $uv^2xy^2z = 0^{m^2}1^{m+k}$. Since $k \leq p < m$ we have $m + k \nmid m^2$, a contradiction!
   - If vxy consists of all 0's, then $uv^2xy^2z = 0^{2m+k}1^m$. Since $k \leq p < m$, we have $m \nmid m^2 + k$, a contradiction!
   - If any of v or y contains both 0's and 1's, then uv^2xy^2z will have 0's that come after some 1's and thus it will not be in the language, a contradiction!
   - The only remaining case is $v = 0^a$ and $u = 1^b$ for some integers $a, b \geq 0$ where at least one of a or b is not zero. In this case $uv^ixy^iz = 0^{m^2+(i-1)b}1^{m+(i-1)a}$. Now to get a contradiction we need to find an i such that $m^2 + ai$ is not divisible by $m + bi$. Here actually again taking $i = 2$ will work. Note that

     $$(m + b) \times (m - b) = m^2 - b^2 < m^2 + a < m^2 - b^2 + m + b = (m + b) \times (m - b + 1).$$

     The lower-bound is obvious as both $a, b \leq p$ and at east one of them is non-zero. For the upper-bound we chose m so that $m - b^2 + b > p$ for every $b \leq p$, and so $m^2 - b^2 + m^2 + b > m + p \geq m^2 + a$. Consequently, $m^2 + a$ cannot be divisible by $m + b$.

2. (20 points) Prove that the intersection of a context-free language C and a *regular language* R is context-free.

   Answer: To simplify the presentation we assume that both languages are over the same alphabet Σ (otherwise we can take the union of the two alphabets). Let $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$ be a push-down automaton that accepts C, and $D = (Q', \Sigma, \delta', q_0', F')$ be a DFA that accepts R. We will construct a push-down automaton $(Q'', \Sigma, \delta'', q_0'', F'')$ that accepts $C \cap R$. The state set of this PDA is the Cartesian product $Q \times Q'$. An state $[q, q']$ is an accept state if both q and q' are accepts states in P and D, respectively. The start state is $[q_0, q_0']$. It remains to describe the transition function δ''. We add an arrow $a, \alpha \to \beta$ from state $[q_1, q_1']$ to $[q_2, q_2']$ if $a, \alpha \to \beta$ is an arrow from q_1 to q_2 in the original PDA P, and moreover either there is an arrow with label a from q_1' to q_2' in the original DFA or $a = \varepsilon$ and $q_1' = q_2'$.

Note that this new PDA simulates both the PDA P and the DFA D simultaneously, and accepts a string if both of them accept it. Thus it's language is the intersection of the two languages.

3. (20 points) Prove that $A = \{wtw^R \,|w, t \in \{0,1\}^* \text{ and } |w| = |t|\}$ is not context-free.

Answer: Assume otherwise, and let p be the pumping constant. Let $m > 100p$ and consider the string $0^m10^m110^m$, which obviously belongs to A as it can be divided as $(0^m1)(0^m1)(10^m)$. By pumping lemma, there is a partition $0^m10^m110^m = uvxyz$ such that $|vxy| \leq p$, $|vy| > 0$ and uv^ixy^iz belongs to A for every i. By considering various cases, we show that this cannot be the case.

- **$|u| < m$:**
In this case let $i = 2$, and note that since $|uv^2xy^2z| > 3(m+1)$, the first $m+2$ letters must be the reverse of the last $m+2$ letters. By $|uv^2xy^2z|$ starts with $0^{m'}10$ where $m' \geq m$ and ends with 110^m, and thus this is not possible.
- **$|u| = m$ and $|v| = 1$:**
In this case let $i = 6$, we get $|uv^2xy^2z| \geq 3(m+1) + 6 = 3(m+3)$ and thus the first $m+3$ letters must be the reverse of the last $m+4$ letters. However the last $m+4$ letters are 0110^m while the first $m+3$ letters are 0^m111.
- **$|u| = m$ and either $|v| > 1$ or $|v| = 0$:**
In this case let $i = 2$, we get $|uv^2xy^2z| > 3(m+1)$ and thus the first $m+2$ letters must be the reverse of the last $m+2$ letters. However the first $m+2$ letters are 0^m10 and the last $m+2$ letters are 110^m.
- **$|z| \leq m$ and y contains at least one 1:** Set $i = 7$, and note that $|uv^7xy^7z| \geq 3m + 3 + 6|y| = 3(m + 2|y| + 1)$, and thus the first $m + 2|y|$ letters must be the reverse of the last. However, the first $m + 2|y|$ letters contain only one 1 (we use the fact that $m > 100p$ and so the the first part is not altered), while the last $m + 2|y|$ letters contain two copies of y and thus at least two 1's.[1]
- **$|z| \leq m$ and y contains no 1's (it can be empty):**
Set $i = 2$, and note that $|uv^2xy^2z| > 3m + 3$, and thus the first $m+2$ letters must be the reverse of the last. However, it is easy to see that the last $m+2$ letters cannot be 010^m.
- **$|z| \geq m+1$ and $|u| \geq m+1$:**
This is the only remaining case. Take $i = 2$ so that $|uv^2xy^2z| > 3m + 3$, and note that uv^2xy^2z starts with 0^m10 and ends with 110^m.

[1] We have to be careful in this case. We can have for example $yz = (10^b)(0^{m-b})$, and thus $y^2z = 10^b10^m$ and the last part of that can be the reverse of 0^m10^b from the beginning of the string.

4. **The solution to the following question will not be posted.**

   (a) (40 points) Draw the state diagram of a Turing Machine that decides the strings of the form "$u < v$" where u and v are two *positive* integers in binary and the string is valid as an inequality. Here the alphabet is $\{0, 1, <\}$. (For example $10 < 100$ is in the language but $11 < 10$ is not; Also note that the leftmost digit of the binary representation of every positive integer is 1.) Explain why your Turing Machine works.

   (b) (10 points) Run your Turing Machine on $100 < 11$. You have to list the sequence of configurations.