CS3350 **Tuesday October 22, 2013**
Quiz 7: 20 minutes – 25 points
NAME: ....

Important: Please write legibly and justify your answers.

Question 1. [10 points] Consider grammar $G = (\{S\}, \{a, b\}, R, S)$. The set of rules, R, is:

$$S \rightarrow aSb \mid SS \mid \varepsilon$$

Show, by drawing the corresponding parse trees, that G recognizes strings $abab$, $aaabbb$, and $aababab$.

Question 2. [5 points] Let $B = \{1^k y \mid y \in \{0,1\}^* \text{ and } y \text{ contains at least } k \text{ 1s, for } k \geq 1\}$. Show that B is a regular language.

If we design a finite automaton that recognizes B or if we find a regular expression for B, we have proven that it is regular. First, let us think about what B is exactly.

$B = \{1^k y \mid y \in \{0,1\}^* \text{ and } y \text{ contains at least } k \text{ 1s, for } k \geq 1\}$.

What this description of B really means is that if we find ourselves in a situation where we have $w = 1^n.w'$, with $w' \in \{0,1\}$, where $|w'|_1 < n$, then we can borrow $1's$ from the first part of w so that we now can write $w = 1^{|w'|_1}.w_0$ where $w_0 = 1^{n-|w'|_1}.w'$.

In essence, this means that in B, words need to start with a 1, they need to have another 1 somewhere to ensure the balance in the two "parts" of the words and anything else is allowed. We could write it as: $B = 1 \circ 0^* \circ 1 \circ \{0,1\}^*$. (we could stop here).

The corresponding finite automaton is as follows:

Question 3. [10 points] Let $C = \{1^k y \mid y \in \{0,1\}^*$ and y contains at most k 1s, for $k \geq 1\}$. Show that C is not a regular language.

Let us show that this language is not regular. For this as usual, we do not specify p, and for any p, we propose a string s, longer than p, that cannot be pumped.

Let us take $s = 1^p01^p$. We need to look at all possible decompositions of s that are such that:

- $s = xyz$
- $|xy| \leq p$
- $|y| \neq 0$

Since the first part of s is made of $1's$ only, the only possibly decompositions of s are those where the sub-string xy is only made of $1's$.

As a result, we can write s as: $s = 1^k.1^m.1^{p-k-m}.0.1^p$ where: $x = 1^k$, $y = 1^m$, with $k + m \leq p$ and $m > 0$, and $z = 1^{p-k-m}.0.1^p$.

So when we pump "up" the y part of the string, the number of $1's$ in the first part of the string will still be larger than the number of $1's$ in the second part (w in the definition of C).

When we pump down, the new string $x.y^0.z$ is as follows: $x.y^0.z = 1^k.1^{p-k-m}.0.1^p = 1^{p-k}.0.1^p$. Hence, the number of $1's$ in the second part of the word is larger than the number of $1's$ in the first part of it, making the new string $x.y^0.z$ not an element of C.