CS323 Assignment 1 - Answers

(For reference only. DO NOT distribute.)

1 Required Exercises (100 points)

Exercise 1: 7

Exercise 2:

1. $n+1$
2. $n-1$
3. 1
4. 1
5. 1 if $m < n$; 0 if $m = n$
6. $n \times (n+1)/2 + 1$ or $C_{n+1}^{2} + 1$
7. $\Sigma_{i=0}^{n} C_n^i$

Exercise 3:

1. Strings of a's and b's.
2. Strings of a's and b's that the third character from the last is a.
3. Strings of a's and b's that only contain three b's.

Exercise 4:

1. $digit \rightarrow [0-9]$
   $nonZeroDigit \rightarrow [1-9]$
   $countryCode \rightarrow 86$
   $countryCode \rightarrow 755$
   $phoneNumber \rightarrow countryCode - areaCode - nonZeroDigit\ digit^7$
2. a(a|b)*b [5 points]
3. $consonant \rightarrow [bcdfghjklmnpqrstvwxyz]$
   $d \rightarrow consonant^*\ a\ (consonant|a)^*\ e\ (consonant|e)^*\ i\ (consonant|i)^*\ o\ (consonant|o)^*\ u\ (consonant|u)^*$

2 Optional Exercises (10 bonus points)

L_1 is equivalent to L_2 and we can prove this by contradition (there may be other ways to prove this). Since L_1 and L_2 are essentially two sets, to prove that they are equivalent is basically to prove that (1) $L_1 \subseteq L_2$ and (2) $L_2 \subseteq L_1$.

Part 1: We first prove that $L_1 \subseteq L_2$.

Suppose that $L_1 \subseteq L_2$ does not hold, then we can find a string s such that $s \in L_1$ but $s \notin L_2$. Without loss of generality, let's assume that $|s| = n$ and we use (i, j) to denote a substring of s, where i indicates the index of the starting character and j indicates the index of the ending character. Here, the index starts from 1, $1 \leq i, j \leq n$, and $i \leq j$.

If s is not in L_2, then its substring $(1, n-1)$ is also not in L_2, because if $(1, n-1)$ is in L_2, then obviously s is also in L_2 according to the language definition. Next, we can show that substrings $(1, n-2)$, $\ldots$, $(1, 1)$ are not in L_2. This is impossible since $(1, 1)$ is either the string "a" or "b", which must be in L_2.

So $L_1 \subseteq L_2$ must hold.

Part 2: We then prove that $L_2 \subseteq L_1$.

Suppose that $L_2 \subseteq L_1$ does not hold, then we can find a string s such tat $s \in L_2$ but $s \notin L_1$. Without loss of generality, let's assume that $|s| = n$ and we use (i, j) to denote a substring of s, where i indicates the index of the starting character and j indicates the index of the ending character. Here, the index starts from 1, $1 \leq i, j \leq n$, and $i \leq j$.

The language $L(a^*b^*)$ is equivalent to the language $L(a|b|a^*b^*)$, i.e., adding two strings "a" and "b" does not change the set. Then $L((a^*b^*)^*)$ is equivalent to $L((a|b|a^*b^*)^*)$.

If s is not in L_1, then its substring $(1, n-1)$ is not in L_1, because if $(1, n-1)$ is in L_1, then obviously s is also in L_1 according to the language definition. Next, we can show that substring $(1, n-2)$, $\ldots$, $(1, 1)$ are not in L_1. This is impossible since $(1, 1)$ is either the string "a" or "b", which must be in L_1.

So $L_2 \subseteq L_1$ must hold.