MATH 239 Assignment 4

- This assignment is due on Friday, October 12, 2012, at 10am in the drop boxes in St. Jerome's (section 1) or outside MC 4067 (the other two sections).
- You may collaborate with other students in the class, provided that you list your collaborators. However, you MUST write up your solutions individually. Copying from another student (or any other source) constitutes cheating and is strictly forbidden.

1. Let S be the set of binary strings that do not contain the substring 0111.

   (a) Give an unambiguous decomposition for S, and explain why it is unambiguous.

   (b) Find the generating series for S with respect to length. Indicate wherever you use results such as the Product Lemma. Express your answer as a rational function (i.e. in the form $\frac{p(x)}{q(x)}$ where $p(x)$ and $q(x)$ are polynomials).

   Solution:

   (a) Using the 0-decomposition, we have $\{0,1\}^* = \{1\}^*(\{0\}\{1\}^*)^*$, and this expression is unambiguous. To describe S we may modify the 0-decomposition to find

   $$S = \{1\}^*(\{0\}\{\epsilon, 1, 11\})^*.$$

   This describes S since each element of S may start with a string of 1's of arbitrary length, but then all subsequent blocks of 1's (that are therefore preceded by a 0) can have length at most two. Moreover the expression is unambiguous because we have simply restricted the 0-decomposition, which has been proven to be unambiguous.

   (b) Using the definition of generating series we can see that

   $$\Phi_{\{1\}^*}(x) = \frac{1}{1-x},$$

   $$\Phi_{\{0\}}(x) = x,$$

   $$\Phi_{\{\epsilon,1,11\}}(x) = 1 + x + x^2.$$

   Therefore by the *-Lemma and the Product Lemma (which are valid since the decomposition in (a) is unambiguous) we get

   $$\Phi_S(x) = (\frac{1}{1-x})(\frac{1}{1-x(1+x+x^2)}) = \frac{1}{1-x-x(1-x^3)} = \frac{1}{1-2x+x^4}.$$

2. Let S' be the set of binary strings that do not contain the substring 11, in which every block of 0's is of even length.

   (a) Give an unambiguous decomposition for S', and explain why it is unambiguous.

(b) Find the generating series for S' with respect to length. Indicate wherever you use results such as the Product Lemma. Express your answer as a rational function.

Solution:

(a) Using the block decomposition, we have $\{0,1\}^* = \{1\}^*(\{0\}\{0\}^*\{1\}\{1\}^*)^*\{0\}^*$, and this expression is unambiguous. To describe S' we may modify the block decomposition to find

$$S' = \{\epsilon, 1\}(\{00\}\{00\}^*\{1\})^*\{00\}^*.$$

This describes S' since each block of 1's must be of length one, and all blocks of 0's are even. Moreover the expression is unambiguous because we have simply restricted the block decomposition, which has been proven to be unambiguous.

(b) Using the definition of generating series we can see that

$$\Phi_{\{00\}^*}(x) = \frac{1}{1-x^2},$$

$$\Phi_{\{1\}}(x) = x,$$

$$\Phi_{\{00\}}(x) = x^2,$$

$$\Phi_{\{\epsilon,1\}}(x) = 1 + x.$$

Therefore by the *-Lemma and the Product Lemma (which are valid since the decomposition in (a) is unambiguous) we get

$$\Phi_{S'}(x) = (1+x)\left(\frac{1}{1-\frac{x^3}{1-x^2}}\right)\left(\frac{1}{1-x^2}\right) = \frac{1+x}{1-x^2-x^3}.$$

3. Let S'' be the set of binary strings that are empty or end with a 0, in which each block of 1's has odd length, and each block of 0's that is preceded by (at least one) 1 has length exactly two.

(a) Give an unambiguous decomposition for S'', and explain why it is unambiguous.

(b) Find the generating series for S'' with respect to length. Indicate wherever you use results such as the Product Lemma. Express your answer as a rational function.

Solution:

(a) Using the block decomposition, we have $\{0,1\}^* = \{0\}^*(\{1\}\{1\}^*\{0\}\{0\}^*)^*\{1\}^*$, and this expression is unambiguous. To describe S'' we may modify the block decomposition to find

$$S'' = \{0\}^*(\{1\}\{11\}^*\{00\})^*.$$

This describes S'' since each block of 1's must be of odd, all blocks of 0's are of length two (except for the initial block of 0's, if it exists), and no 1's occur at the end. Moreover the expression is unambiguous because we have simply restricted the block decomposition, which has been proven to be unambiguous.

(b) Using the definition of generating series we can see that

$$\Phi_{\{11\}^*}(x) = \frac{1}{1-x^2},$$
$$\Phi_{\{1\}}(x) = x,$$
$$\Phi_{\{00\}}(x) = x^2,$$
$$\Phi_{\{0\}^*}(x) = \frac{1}{1-x}.$$

Therefore by the *-Lemma and the Product Lemma (which are valid since the decomposition in (a) is unambiguous) we get

$$\Phi_{S''}(x) = (\frac{1}{1-x})(\frac{1}{1-\frac{x^3}{1-x^2}}) = \frac{1+x}{1-x^2-x^3}.$$

4. Let S' and S'' be as in Questions 2 and 3. Let $S'(n)$ denote the set of strings in S' of length n, and similarly $S''(n)$ denote the set of strings in S'' of length n. Prove that for every $n \geq 0$ there is a bijection between $S'(n)$ and $S''(n)$. (Hint: you do not necessarily have to find the bijection explicitly.)

Solution:

Recall that there is a bijection between two finite sets A and B if and only if $|A| = |B|$. Here we simply observe that S' and S'' have the same generating function , and therefore

$$|S'(n)| = [x^n]\Phi_{S'}(x) = [x^n]\Phi_{S''}(x) = |S''(n)|.$$

Therefore there is a bijection between $S'(n)$ and $S''(n)$ for each $n \geq 0$.

5. Let S be the set of all binary strings that are either empty or begin with a block of 1's and end with a block of 0's.

(a) Give a recursive decomposition for S, that is an unambiguous expression. Explain why your decomposition is unambiguous.

(b) Use the decomposition in (a) to find the generating series for S with respect to length.

(c) Find the number of strings in S of length n (as an explicit closed-form expression in terms of n).

Solution:

(a) A recursive decomposition for S is given by

$$S = \{\epsilon\} \cup S\{1\}\{1\}^*\{0\}\{0\}^*.$$

To see that this decomposition is correct and unambiguous, let σ be an arbitrary string in S. If σ is the empty string then $\sigma \in \{\epsilon\}$ is the only possibility. If σ is nonempty then let $\sigma = c_1 d_1 \ldots c_k d_k$ be the (uniquely determined) block decomposition of σ, where $k > 0$. Then for each i we must have that c_i is a block of 1's and d_i is a block of 0's. Then there is a unique string in S, namely $\sigma' = c_1 d_1 \ldots c_{k-1} d_{k-1}$, such that σ is obtained by appending a block of 1's and a block of 0's to σ'.

(b) Using the definition of generating series we can see that

$$\Phi_{\{0\}\{0\}^*}(x) = \frac{x}{1-x} = \Phi_{\{1\}\{1\}^*}(x).$$

Therefore by the Sum Lemma and the Product Lemma (which is valid since the decomposition in (a) is unambiguous) we get

$$\Phi_S(x) = 1 + \Phi_S(x)\frac{x^2}{(1-x)^2}.$$

Rearranging this expression gives

$$\Phi_S(x) = \frac{1}{1 - \frac{x^2}{(1-x)^2}} = \frac{1-2x+x^2}{1-2x}.$$

(c) We find the coefficient of x^n in the above generating series.

$$[x^n]\Phi_S(x) = [x^n]\frac{1-2x+x^2}{1-2x} = [x^n](1 + \frac{x^2}{1-2x}).$$

Therefore this number is 1 if $n = 0$, 0 if $n = 1$ and if $n \geq 2$ it is $[x^{n-2}](1-2x) = 2^{n-2}$.

6. For each of the following sets A, either prove that A^* is unambiguous, or give an example to show that A^* is ambiguous. (Hint: one way to prove an expression is unambiguous is by induction on length of strings.)

   (a) $A = \{1, 100, 110011\}$

   (b) $A = \{110, 001, 0001\}$.

   Solution:

   (a) Here A^* is ambiguous, for example $110011 = (1)(100)(1)(1)$.

   (b) Here we show that A^* is an unambiguous expression. To see this, let $\sigma \in A^*$ be an arbitrary string in A^*. We apply induction on the length ℓ of σ. First we can verify the base case $\ell = 0$, since the only way to get the empty string is to use none of the elements of A.

   For the induction hypothesis, assume that $\ell > 0$ and that all strings in A^* of length less than ℓ have a unique expression as an element of A^*.

   Given σ, since $\sigma \in A^*$ there is some way to express $\sigma = a_1 a_2 \cdots a_k$, with each $a_i \in A$. Consider now the first entry in σ. If it is 1, then a_1 must be 110, because this is the only element of A that starts with 1 . Since $a_2 \cdots a_k \in A^*$ has length $< \ell$, the induction hypothesis implies that it has a unique expression as an element of A^*. Therefore the same is true for σ .

   Now suppose the first entry in σ is 0. It follows that, in this case, a_1 is one of 001 or 0001 . In both cases, a_1 is determined by the number of leading 0's in σ (there can be two leading 0's or three leading 0's in σ, if it is two then $a_1 = 001$ and if it is three then $a_1 = 0001$) . Again, $a_2 \cdots a_k$ is unambiguous by induction, so we see that σ has a unique expression as an element of A^*, as required. Therefore by induction A^* is unambiguous.