Math 239 Spring 2014 Assignment 4 Solutions

1. {6 marks} For each of the following expressions representing sets of binary strings, determine if it is ambiguous or unambiguous. If it is ambiguous, provide an example that can be decomposed in two different ways. If it is unambiguous, determine the generating series (as a simplified rational expression) with respect to the lengths of strings.

   (a) $(\{1\}^*\{001, 0001\}^*)^*$

   Solution. This is ambiguous. In particular, the empty string ε is inside the main $*$, so it can be written as $\varepsilon\varepsilon$ and $\varepsilon\varepsilon\varepsilon\varepsilon$.

   (b) $\{10001, 001, 00110\}^*$

   Solution. This is ambiguous. Consider the string 00110001. This can be decomposed in two different ways: $(001, 10001)$ and $(00110, 001)$.

   (c) $\{00\}^*(\{11\}\{111\}^*\{0\}\{000\}^*)^*\{1, 11, 111\}$

   Solution. This is unambiguous (in the same way that the block decomposition is unambiguous). The generating series is

$$\frac{1}{1-x^2}\left(\frac{1}{1-\frac{x^2}{1-x^3}\frac{x}{1-x^3}}\right)(x+x^2+x^3) = \frac{(1-x^3)^2(x+x^2+x^3)}{(1-x^2)(1-3x^3+x^6)}.$$

2. {9 marks} For each of the following sets of binary strings, determine an unambiguous expression which generates every string in that set (no justification required).

   (a) The set of binary strings where the length of each block of 0's is divisible by 3 and the length of each block of 1's is divisible by 4.

   Solution. $S = \{000, 1111\}^*$.

   (b) The set of binary strings where each block of 1's of odd length is followed by a block of 0's of length at least 3.

   Solution. We use the block decomposition. The middle two-block combination is either odd 1's followed by at least 3 0's, or even 1's followed by any number of 0's. The last part must be either empty or an even block of 1's, for it does not have any block of 0's following it.

$$S = \{0\}^*(\{1\}\{11\}^*\{000\}\{0\}^* \cup \{11\}\{11\}^*\{0\}\{0\}^*)^*\{11\}^*.$$

   (c) The set of binary strings which contain 000111 as a substring.
   (Note: An obvious wrong answer is $\{0,1\}^*(000111)\{0,1\}^*$.)

   Solution. Take all the strings and remove those that do not contain 000111 as a substring. For such a string, any block of 0's of length at least 3 can be followed by a block of 1's of length at most 2, and any block of 0's of length 1 or 2 can be followed by any number of 1's.

$$S = \{0,1\}^* \setminus \{1\}^*(\{000\}\{0\}^*\{1, 11\} \cup \{0, 00\}\{1\}\{1\}^*)^*\{0\}^*.$$

3. {5 marks} On Martin's new game show "It Peis to Play", there is an unlimited supply of gold and silver coins buried inside a box of sand. As a contestant, you dig up one coin at a time, and Martin will offer you \$3 for each gold coin and \$1 for each silver coin. You may stop at any time and keep your winnings. However, if you dig up 3 gold coins in a row or 4 silver coins in a row, the game is over and you lose everything. (For this question, you may represent your answers as coefficients of simplified rational expressions.)

   (a) For some $n \in \mathbb{N}$, how many ways can you win exactly \$$n$ from Martin and walk away with your winnings?

   Solution. We set this up as a binary string problem, where 0 represents a silver coin and 1 represents a gold coin. For a string s, we define the weight function $w(s)$ to be the number of 0's plus 3 times the number of 1's in s.

   In this problem, we need the set of all binary strings with no 4 consecutive 0's and no 3 consecutive 1's. This can be expressed as the following unambiguous expression:

$$\{\varepsilon, 0, 00, 000\}(\{1, 11\}\{0, 00, 000\})^*\{\varepsilon, 1, 11\}.$$

Using the new weight function, we see that the generating series is

$$(1+x+x^2+x^3)\frac{1}{1-(x^3+x^6)(x+x^2+x^3)}(1+x^3+x^6) = \frac{(1+x+x^2+x^3)(1+x^3+x^6)}{1-(x^3+x^6)(x+x^2+x^3)}.$$

The answer is then the coefficient of $[x^n]$ in this generating series.

(b) For some $n \in \mathbb{N}$, suppose you have won \$$n$, but decided to be greedy and then lost everything on the next dig. How many ways can this happen?

Solution. Using the same set up as part (a), we now need the set of all binary strings with no 4 consecutive 0's and no 3 consecutive 1's, which end with either 000 or 11. (This assumes that the next bit will be the same.) An unambiguous expression for this is

$$\{\varepsilon, 0, 00, 000\}(\{1, 11\}\{0, 00, 000\})^*\{11\} \cup \{\varepsilon, 1, 11\}(\{0, 00, 000\}\{1, 11\})^*\{000\}.$$

The generating series for this is

$$\frac{(1+x+x^2+x^3)x^6 + (1+x^3+x^6)x^3}{1-(x^3+x^6)(x+x^2+x^3)}.$$

The answer is then the coefficient of $[x^n]$ in this generating series.

4. {4 marks} Let S be the set of binary strings where a block of 0's cannot be followed by a block of 1's of greater length. For example, 111110010001110 is in S, but 100011110010 is not. Prove that the generating series for S with respect to the lengths of the strings is

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

(Hint: You may want to consider using the set $T = \{01, 0011, 000111, 00001111, \ldots\}$.)

Solution. The configuration that we want to avoid is a block of 0's followed by a block of 1's of greater length, which can be described as $T\{1\}\{1\}^*$. By modifying the block decomposition, we see that

$$S = \{1\}^*(\{0\}\{0\}^*\{1\}\{1\}^* \setminus T\{1\}\{1\}^*)^*\{0\}^*.$$

Now

$$\Phi_T(x) = x^2 + x^4 + x^6 + \cdots = \frac{x^2}{1-x^2}.$$

So,

$$\begin{aligned}
\Phi_S(x) &= \frac{1}{1-x}\frac{1}{1-\left(\frac{x^2}{(1-x)^2} - \frac{x^2}{1-x^2}\frac{x}{1-x}\right)}\frac{1}{1-x} \\
&= \frac{1}{(1-x)^2}\frac{(1-x)^2(1+x)}{(1-x)^2(1+x) - x^2(1+x) + x^3} \\
&= \frac{1+x}{1-x-2x^2+x^3}.
\end{aligned}$$

5. {4 marks} Let T be the set of all binary strings that do not have 0001 as a substring. Determine a recursive definition for T. Briefly explain why your definition is correct and unambiguous. Use this to find the generating series for T with respect to the lengths of the strings.

Solution. We split T into two sets T_0 and T_1 where T_0 are strings in T that do not have any 1's, and T_1 are strings in T that have at least one 1. The strings in T_0 are simply those with only 0's, so $T_0 = \{0\}^*$. For a string in T, break it after the first 1. The first part is either $1, 01$, or 001, and the remaining string is still in T. So we may define T by

$$T = \{0\}^* \cup \{1, 01, 001\}T.$$

This is correct and unambiguous because in the description above, there is only one way to decompose a string in T in this form.

For the generating series, we have

$$\Phi_T(x) = \Phi_{\{0\}^*}(x) + \Phi_{\{1,01,001\}}(x)\Phi_T(x) = \frac{1}{1-x} + (x+x^2+x^3)\Phi_T(x).$$

Solving for T gives us

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