MATH 239 Spring 2012: Assignment 4

Solutions

1. {16 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 is divisible by 3.

   Solution. $\{000, 111\}^*$.

   (b) The set of binary strings which do not contain 01111 as a substring.

   Solution. $\{1\}^*(\{0\}\{\varepsilon, 1, 11, 111\})^*$.

   (c) The set of binary strings where each block of 1's must be followed by a block of 0's of length at least 3.

   Solution. $\{0\}^*(\{1\}\{1\}^*\{000\}\{0\}^*)^*$. (Note that there is no $\{1\}^*$ at the end, since it is either an empty string, or a block of 1's which must have another block of 0's following it.)

   (d) The set of binary strings which do contain 1111000 as a substring.

   Solution. $\{0,1\}^* \setminus \{0\}^*(\{1, 11, 111\}\{0\}\{0\}^* \cup \{1111\}\{1\}^*\{0, 00\})^*\{1\}^*$. (The idea is that we start with strings that do not contain 1111000, which means that any block of 1's with length 3 or less can be followed by any number of 0's. However, any block of 1's with length 4 or more can be followed by at most 2 0's.)

2. {12 marks} Inside Bertrand's special box, there is an unlimited supply of blue and red balls. You draw one ball at a time, and Bertrand will offer you \$1 for each ball you draw, as long as you do not draw 4 of the same-coloured balls in a row, at which point you lose everything. (For this question, represent your answers as coefficients of rational expressions.)

   (a) How many ways can you win exactly $\$n$ from Bertrand for some $n \geq 1$?

   Solution. We represent drawing a blue ball with a 0, and rawing a red ball with a 1. A sequence of draws can be represented as a binary string. For this question, we are looking for binary strings where each block has length at most 3. By modifying the block decomposition, we get that

   $$S = \{\varepsilon, 0, 00, 000\}(\{1, 11, 111\}\{0, 00, 000\})^*\{\varepsilon, 1, 11, 111\}.$$

   The generating series for S is

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

   The answer is then the coefficient of n in this series.

   (b) How many ways can you win $\$n$, but get greedy and lose everything on the next draw?

   Solution. This is the same as the previous part, except now the last block is either exactly 4 1's or exactly 4 0's. So we are looking for this set of binary strings:

   $$S = \{\varepsilon, 0, 00, 000\}(\{1, 11, 111\}\{0, 00, 000\})^*\{1111\} \cup \{\varepsilon, 1, 11, 111\}(\{0, 00, 000\}\{1, 11, 111\})^*\{0000\}.$$

   The generating series for S is

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

3. {Extra credit: 3 marks} Describe the set of binary strings which is generated by the following expression:

$$(1(0\{1\}^*0)^*1\{0\}^*)^*$$

Solution. This represents the set of binary strings that start with 1 and whose values are multiples of 3, plus the empty string. (We don't expect you to get this unless you have searched online.)

4. {12 marks} Let S be the set of all binary strings where consecutive blocks have different parities. For example, things in S include 00011011111100000110, 1111111, 0011111, ε. Prove that the generating series for S is

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

Solution. We partition S into five parts:

(a) S_0 is the empty string;

(b) S_1 is the set of those strings that begin with an odd block of 0's;

(c) S_2 is the set of those strings that begin with an even block of 0's;

(d) S_3 is the set of those strings that begin with an odd block of 1's;

(e) S_4 is the set of those strings that begin with an even block of 1's.

For S_1, all blocks of 0's have odd length, and all blocks of 1's have even length. So a decomposition for S_1 is

$$S_1 = \{0\}\{00\}^*(\{11\}\{11\}^*\{0\}\{00\}*)^*\{11\}^*.$$

For S_2, all blocks of 0's have even length, and all blocks of 1's have odd lengths. So a decomposition for S_2 is

$$S_2 = \{00\}\{00\}^*(\{1\}\{11\}^*\{00\}\{00\}^*)^*\{\varepsilon \cup \{1\}\{11\}^*\}.$$

Note that we must being with $\{00\}\{00\}^*$ because these strings cannot start with 1's. Also, we need ε in the end because we might end with a block of 0's.

For S_3 and S_4, they are simply S_1 and S_2 where the 0's and 1's switch places. So in particular, the generating series for S_3 and S_4 are the same as the generating series for S_1 and S_2 respectively.

The generating series for S_1 and S_2 are

$$\Phi_{S_1}(x) = \frac{x}{1-x^2}\frac{1}{1-\frac{x^2}{1-x^2}\frac{x}{1-x^2}}\frac{1}{1-x^2} = \frac{x}{1-2x^2-x^3+x^4}$$

$$\Phi_{S_2}(x) = \frac{x^2}{1-x^2}\frac{1}{1-\frac{x}{1-x^2}\frac{x^2}{1-x^2}}\left(1+\frac{x}{1-x^2}\right) = \frac{x^2-x^4+x^3}{1-2x^2-x^3+x^4}.$$

Therefore,

$$\begin{aligned}\Phi_S(x) &= \Phi_{S_0}(x) + 2\Phi_{S_1}(x) + 2\Phi_{S_2}(x)\\ &= 1 + \frac{2x}{1-2x^2-x^3+x^4} + \frac{2(x^2-x^4+x^3)}{1-2x^2-x^3+x^4}\\ &= \frac{1+2x+x^3-x^4}{1-2x^2-x^3+x^4}.\end{aligned}$$

5. {5 marks} For some positive integer m, let $s_1, \ldots, s_k$ be distinct binary strings of length m. Prove that $S = \{s_1, \ldots, s_k\}^*$ is an unambiguous expression.

   Solution. We prove this using strong induction on the length n of a string $s \in S$. When $n = 0$, $s = \varepsilon$, and this can only be generated once from $\{s_1, \ldots, s_k\}^0$. Suppose s has length $n > 0$. Since $\{s_1, \ldots, s_k\}$ are distinct and each string has length m, the first m bits of s is uniquely generated, say it is s_i for some i. Then we can decompose s as $s = s_i t$ where $t \in \{s_1, \ldots, s_k\}^*$. Since t has shorter length than s, by induction hypothesis, there is only one way to generate t from S. Therefore, s can be generated only once from S, and S is unambiguous.

6. {5 marks} Prove that for any choice of positive integers m and n where $m \neq n$, there exist binary strings s and t of lengths m and n respectively where $\{s, t\}^*$ is an ambiguous expression.

   Solution. Let s be the string of m 1's, and let t be the string of n 1's. Then $s^n = t^m$ which is the string of mn 1's, and both $s^n, t^m \in \{s, t\}^*$. Hence this is ambiguous.