Math 239 F13 Tutorial 4

October 8, 2013

Problem 1.

For each of the following expressions, decide whether the expression is ambiguous or unambiguous and prove your assertion.

a) $A = \{\epsilon, 00, 01, 10, 11\}^*$

b) $A = \{10, 11, 01110\}^*$

c) $A = \{0, 10, 11, 01\}^*$

d) $A = \{0, 1\}^*\{1100101\}\{0, 1\}^*$

Solution:

a) A is ambiguous because there are multiple ways of writing ϵ: $\epsilon\epsilon, \epsilon\epsilon\epsilon, \ldots$.

b) By induction on the length of strings we'll show that each string in A decomposes unambiguously. Base cases: if a string has length 0, then there is only one way to write it - as ϵ. Also, there are no strings of length 1, hence, there is no ambiguity.

Suppose now that every string in A of length strictly less than k for some $k \geq 2$ is generated unambiguously. Consider any string σ of length k. Since $k \geq 2$, σ must start with an element of $\{10, 11, 01110\}$, say $\sigma = a_1 \cdots a_m$ for some $m \geq 2$. If a_1's first bit is 0, then a_1 *must be* 01110. Otherwise, look at the second bit of a_1. If it is 0, then a_1 *must be* 10. If it is 1, then a_1 *must be* 11. In each of these three cases, we can write $\sigma = a_1\sigma'$ for some σ' in A with length strictly less than k. By the induction hypothesis, it can be expressed uniquely as an element of A. Therefore, by induction, A is unambiguous.

c) The set A is ambiguous since we can write 010 as $0|10$ or $01|0$. (Aside: where exactly would the induction argument from part b) fail if applied to this case?)

d) The set A is ambiguous since, for example, we can write 11001011100101 as $1100101|1100101|\epsilon$ or $\epsilon|1100101|1100101$.

Problem 2.

For the following sets of strings, determine the generating function with respect to length.

a) Set S of strings where each block of ones is followed by a block of exactly two zeroes.

b) Set S of non-empty strings that start and end with the same digit.

c) Set S of strings in which each block of zeros is preceded by a block of ones having twice its length.

d) Set S of strings that do not contain 001 as a substring.

Solution:

a) Decomposing after each block of 0's, we unambiguously express S as $S = \{0\}^*(\{1\}\{1\}^*\{00\})^*$ (notice that there cannot be a suffix $\{1\}^*$ since these 1s would not be followed by two zeroes). By using the sum and product rules for strings, we then obtain the generating function $\Phi_S(x)$ as

$$\Phi_S(x) = \Phi_{\{0\}^*}(x) \cdot \frac{1}{1 - \Phi_{\{1\}\{1\}^*\{00\}}(x)}$$

$$= \frac{1}{1 - \Phi_{\{0\}}(x)} \cdot \frac{1}{1 - (\Phi_{\{1\}}(x) \cdot \frac{1}{1-\Phi_{\{1\}}(x)} \cdot \Phi_{\{00\}}(x))}$$

$$= \frac{1}{1-x} \cdot \frac{1}{1 - \frac{x \cdot x^2}{1-x}} = \frac{1}{1 - x - x^3}.$$

b) The set of strings that start and end with a 0 and are of length at least two can unambiguously be expressed as $\{0\}\{0,1\}^*\{0\}$. Similarly, the set of strings beginning and ending with a 1 and having length at least two can be expressed as $\{1\}\{0,1\}^*\{1\}$. Finally, set S also contains strings 0 and 1. Therefore, $S = \{0\}\{0,1\}^*\{0\} \cup \{1\}\{0,1\}^*\{1\} \cup \{0,1\}$. Since this is an unambiguous decomposition, we compute $\Phi_S(x)$ as

$$\begin{aligned}\Phi_S(x) &= \Phi_{\{0\}}(x)\Phi_{\{0,1\}^*}(x)\Phi_{\{0\}}(x) + \Phi_{\{1\}}(x)\Phi_{\{0,1\}^*}(x)\Phi_{\{1\}}(x) + \Phi_{\{0,1\}}(x)\\ &= x \cdot \frac{1}{1-2x} \cdot x + x \cdot \frac{1}{1-2x} \cdot x + (2x)\\ &= \frac{2x-2x^2}{1-2x}.\end{aligned}$$

c) If we decompose the strings in S after each block of zeroes, then each piece of a given string is in $M := \{110, 111100, 111111000, \ldots\}$ except the last piece that is in $\{1\}^*$ (notice that the first piece cannot be $\{0\}^*$ since these zeroes would not be preceded by ones, as required). Thus,

$$S = M^*\{1\}^*,$$

which is unambiguous. We now compute $\Phi_M(x) = \sum_{i\geq 1} x^{3i} = \frac{x^3}{1-x^3}$. Then,

$$\Phi_S(x) = \Phi_{M^*}(x)\Phi_{\{1\}^*}(x) = \frac{1}{1-\Phi_M(x)} \cdot \frac{1}{1-\Phi_{\{1\}}(x)}$$

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

d) Notice that 'not containing a substring 001' means that each block of zeroes in a given string that is followed by a block of ones must, in fact, consist of only one 0. Now, decompose the strings in S after each block of ones. Then each string in S consists of pieces of the form $\{0\}\{1\}\{1\}^*$, except for the very first piece that may contain only ones $\{1\}^*$ and the very last piece that may consist of arbitrary many zeroes $\{0\}^*$. Thus, $S = \{1\}^*(\{0\}\{1\}\{1\}^*)^*\{0\}^*$ and

$$\begin{aligned}\Phi_S(x) &= \frac{1}{1-\Phi_{\{1\}}(x)} \cdot \frac{1}{1-\Phi_{\{0\}}(x)\Phi_{\{1\}}(x)\Phi_{\{1\}^*}(x)} \cdot \frac{1}{1-\Phi_{\{0\}}(x)}\\ &= \frac{1}{1-x} \cdot \frac{1}{1-x\cdot x \cdot \frac{1}{1-x}} \cdot \frac{1}{1-x}\\ &= \frac{1}{(1-x)(1-x-x^2)}.\end{aligned}$$

Problem 3.

Your friend Ian has an unlimited supply of blue and red balls inside a big shiny box. You play the following game with him: you draw one ball at a time, and Ian offers you \$1 each time you draw a ball from the box. You may stop at any time and keep your winnings. However, if you draw 4 of the same-coloured balls in a row, the game is over and you lose everything. For some $n \in \mathbb{N}$, how many ways can you win exactly \$$n$ from Ian and stop the game?

Solution:

We represent drawing a blue ball with a 0, and drawing 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 decomposing after each block of zeroes, we get that

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

The generating function for S is

$$\begin{aligned}
\Phi_S(x) = \Phi_{\{\epsilon,0,00,000\}}(x) \cdot \frac{1}{1 - \Phi_{\{1,11,111\}}(x)\Phi_{\{0,00,000\}}(x)} \cdot \Phi_{\{\epsilon,1,11,111\}}(x) \\
= (1 + x + x^2 + x^3) \cdot \frac{1}{1 - (x + x^2 + x^3)^2} \cdot (1 + x + x^2 + x^3) \\
= \frac{1 + x + x^2 + x^3}{1 - x - x^2 - x^3}.
\end{aligned}$$

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