Regular Expressions

Sample regular expressions over the alphabet \{a, b\}:

- \(a + ab\): this expression represents the language \(\{a, ab\}\).
- \(aa^* + \epsilon\): this expression represents the language containing the empty string only, and their union is the language of strings containing any number of \(a\)s (including none). A more concise regular expression for the language is \(a^*\).
- \(a(a^*(bb^*a + \epsilon))^* + \epsilon\): this is the language of strings in which the number of \(ab\)s is equal to the number of \(ba\)s. To see this, note that the three regular expressions joined by + at the top level correspond respectively to strings beginning with \(a\), strings beginning with \(b\), and the empty string. You can think of the * operator as roughly corresponding to a loop. Consider the expression \(a(a^*(bb^*a + \epsilon))^*\): this is the language of all strings beginning with \(a\), and continuing with the following repeating pattern: any number of \(a\)s, optionally followed by at least one \(b\) and one \(a\), and the empty string in the language.

The Pumping Lemma

Use the pumping lemma to prove that the language \(L = \{a^n b^n | n > 0\}\) is not regular.

Suppose \(L\) is regular. Then, by the pumping lemma, there is some \(p > 0\) such that every string \(s \in L\) with \(|s| \geq p\) can be written as \(s = xyz\), and \(|xy| \leq p\), \(|y| \geq 1\), and \(xy^iz \in L\) for all natural \(i\).

Consider the string \(s = a^p b^p\). \(s \in L\), and \(|s| = 2p \geq p\). Then, write \(s = xyz\), where \(|xy| \leq p\) and \(|y| \geq 1\). Consider the string \(xy\): since it is of length at most \(p\), it is composed of some number of \(a\)s only. Therefore, \(y = a^q\) for some \(q \geq 1\). The string \(xy^2z\) is in \(L\) by the pumping lemma, yet \(xy^2z\) contains more \(a\)s than \(b\)s. This is a contradiction, so \(L\) is not regular.

Generating Functions

Suppose you wake up with a terrible headache. You have three tablets of Tylenol, two tablets of Advil, and six tablets of Motrin. Tablets of Motrin must be taken in pairs to be effective. Create a generating function for the number of different ways you can take \(n\) tablets of medicine.

Directly from the problem, the polynomial is \((1+x+x^2+x^3)(1+x+x^2)(1+x^2+x^4+x^6)\). The first multiplicand represents the possible choices of Tylenol tablets, the second of Advil tablets, and the last of Motrin tablets. The result of multiplying this out is \(1 + 2x + 4x^2 + 5x^3 + 6x^4 + 6x^5 + 6x^6 + 6x^7 + 5x^8 + 4x^9 + 2x^{10} + x^{11}\).

If you wanted to use this to find out how many ways you could pick one tablet, for example, you would look at the coefficient of the term \(2x^4\). There are two choices: you can either pick one Tylenol tablet or one Advil tablet, but you cannot pick a Motrin tablet because they come in pairs.
**Automaton**

Write a regular expression for the language accepted by this automaton. The alphabet is \( \{a, b\} \).

The automaton accepts the language of strings that include the sequence \( abbab \). A regular expression for this language is \((a + b)^* abbab (a + b)^*\).

**Generating Functions Again**

Create a generating function for the number of ways three gold bars can be divided among two pirates. You don’t need to expand it. Suppose you did, however. What term would you be most interested in looking at, and why? What do the other terms tell you?

The generating function here is \((1 + x + x^2 + x^3)^2 = 1 + 2x + 3x^2 + 4x^3 + 3x^4 + 2x^5 + x^6\). A term \(ax^b\) can be interpreted as “there are \(a\) ways to divide \(b\) gold bars among two pirates, such that no pirate gets more than three bars of gold.” To see this, consider all the cross-terms contributing to each term: for example, \(3x^4 = xx^3 + x^2x^2 + x^3x\). The exponents give the possible ways to divide the four bars of gold such that neither pirate gets more than three bars.

The specific term we are interested in is the one where \(b = 3\), as there are three bars of gold.