Introduction to Techniques for Counting

‘A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detachedly, which could be embarrassing, we put them all in a bag, and then we have only one object to carry, the bag.’

—George Polya, *Mathematics and plausible reasoning* (1954)

‘A generating function is a clothesline on which we hang up a sequence of numbers for display.’

—Herbert Wilf, *generatingfunctionology* (1994)

Some days I go to the coffee shop, and other days not. If I never go to the coffee shop two days in a row, how many choices of schedule are possible for visiting the coffee shop over a period of n consecutive days? Implicitly, ‘schedule’ here refers to a specification of which days I went to the coffee shop, and which days not; so each schedule is uniquely represented by a sequence of 0’s and 1’s indicating the days I did not, or did, go to the shop, respectively. We translate our original problem into an equivalent one involving bitstrings, as follows.

A *bit* (abbreviation for ‘*binary digit*’) is defined to be a symbol ‘0’ or ‘1’. (Note that we consider bits a symbols, to be interpreted literally rather than numerically (so you should view them as simply letters, not numbers). By a *bitstring* (or *binary string*), we mean a finite sequence of bits. For each $n \geqslant 0$, there are exactly 2^n bitstrings of *length* n, where the length of a bitstring is defined to be the number of bits in the string. We include the case $n = 0$, which yields the *null string* ‘’ of length zero. Usually we omit the quotation marks, abbreviating ‘01101’ as 01101 for example; but for the null string, clearly such abbreviation won’t work. Here we list explicitly all bitstrings of length $\leqslant 3$:

n	bitstrings of length n
0	‘’
1	0, 1
2	00, 01, 10, 11
3	000, 001, 010, 011, 100, 101, 110, 111

Our problem is to determine the number of bitstrings of length n having no two consecutive 1’s. Let us call such a bitstring a 11-free bitstring, and denote by a_n the number of 11-free bitstrings of length n. The following table shows that the first few terms in the sequence a_n look remarkably like Fibonacci numbers:

n	11-free bitstrings of length n	a_n = number of such bitstrings	b_n = no. ending in '1'	c_n = no. *not* ending in '1'
0	''	1	0	1
1	0, 1	2	1	1
2	00, 01, 10	3	1	2
3	000, 001, 010, 100, 101	5	2	3
4	0000, 0001, 0010, 0100, 1000, 0101, 1001, 1010	8	3	5

For future reference, I have also indicated how many such bitstrings end (or do not end) in '1': there are b_n and c_n such bitstrings respectively, so that $a_n = b_n + c_n$. Based on this limited evidence, we conjecture (i.e. guess) that $a_{n-1} = b_{n+1} = c_n$ is the n-th Fibonacci number (see Chapter 5 of the textbook). The Fibonacci sequence 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ... is defined recursively by

$$F_n = \begin{cases} 1, & \text{if } n = 0 \text{ or } 1; \\ F_{n-1} + F_{n-2}, & \text{if } n \geqslant 2. \end{cases}$$

In order to establish the connection, we prove

Theorem. *For all $n \geqslant 0$, we have $a_{n+2} = a_{n+1} + a_n$.*

Proof. Every 11-free bitstring of length $n+1$ *not* ending in '1' (call it w) has the form $w = w'0$ where w' is an arbitrary 11-free bitstring of length n; this says that

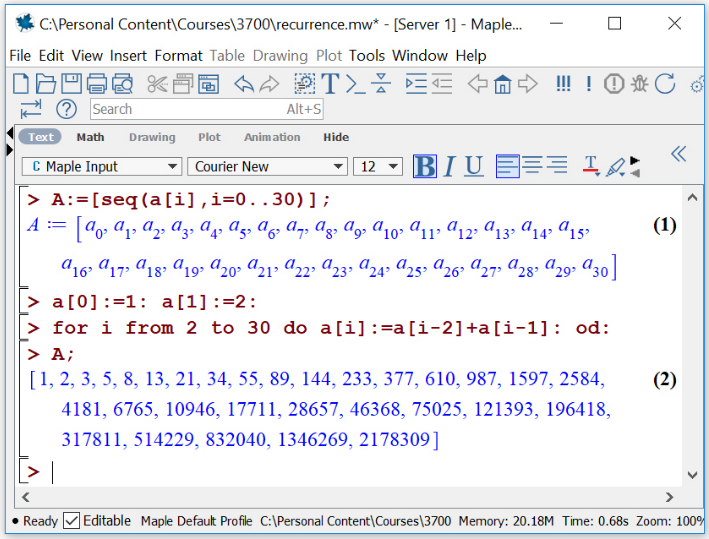
$$c_{n+1} = a_n = b_n + c_n.$$

Every 11-free bitstring w of length $n+1$ *ending in* '1' has the form $w = w'1$ where w' is necessarily a 11-free bitstring of length n *not* ending in '1'; this says that

$$b_{n+1} = c_n.$$

Substituting $b_n = c_{n-1}$ into the previous formula gives $c_{n+1} = c_n + c_{n-1}$ for $n \geqslant 1$; and since $c_0 = c_1 = 1$, we must have $c_n = F_n$, the nth Fibonacci number. This also yields $a_{n+2} = a_{n+1} + a_n$. $\square$

We demonstrate how easily Maple® generates the first few terms of the sequence using the *recurrence relation* $a_n = a_{n-1} + a_{n-2}$:



Our recursive formula for a_n clearly provides a way to compute many terms of the sequence. However, for large values of n, this method requires iterating the recursion many times, which may prove to be impractical. In some situations it may therefore be preferable to have a *closed formula* for a_n, i.e. one which yields a_n directly without having to compute all the preceding terms of the sequence. Although closed formulas are not available for every counting problem, we will obtain such a formula in this case, thereby obtaining further insight into the sequence a_n (including its asymptotic growth rate, which turns out to be exponential). Our primary tool for this purpose is the *generating function* of the sequence $a_0, a_1, a_2, \ldots$, defined (as in Chapter 5 of the textbook) by

$$\begin{aligned} F(x) &= \sum_{n\geqslant 0} a_n x^n = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots \\ &= x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + 13x^6 + \cdots. \end{aligned}$$

Note that a single expression $F(x)$ carries all the same information as the entire infinite sequence a_n. It is important to note that this expression is purely symbolic. Here x is a symbol, not a number; and $F(x)$ is also purely symbolic. In particular, calling $F(x)$ a function is a misnomer. We never evaluate $F(a)$ for any number a; and so convergence of the power series is never an issue. For this reason, none of the analytic properties of power series learned in Calculus II are needed here; all that we require is the formal (and quite naive) algebraic manipulation of power series.

One of the most important power series is the *geometric series*

$$\frac{1}{1-u} = \sum_{n=0}^{\infty} u^n = 1 + u + u^2 + u^3 + u^4 + \cdots .$$

The identity is verified by cross-multiplying:

$$\begin{aligned}(1-u)(1+u+u^2+u^3+u^4+\cdots) &= 1-u+u-u^2+u^2-u^3+u^3-\cdots \\ &= 1.\end{aligned}$$

Here, as always, u and the series expansion for $\frac{1}{1-u}$, are purely symbolic. (In Calculus II you would have been taught that the series is only valid when the series converges, i.e. for $|u| < 1$; but in our context it is inappropriate to require such a caveat since u is not a number, and in particular $|u|$ has no meaning.)

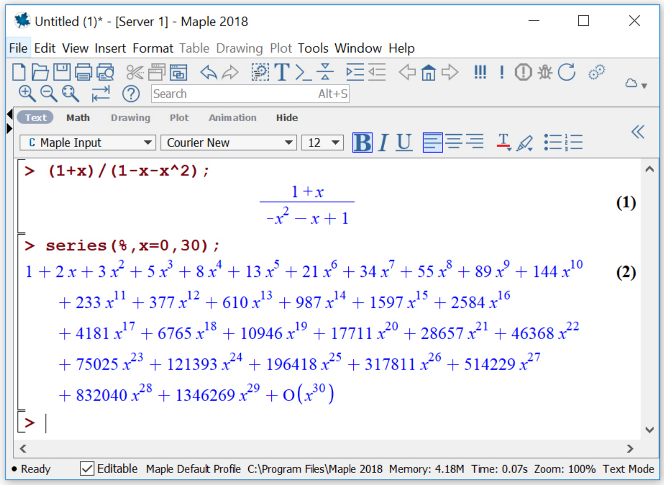
To obtain a closed-form expression for $F(x)$, use the identity $a_{n+2} = a_{n+1} + a_n$, valid for all $n \geqslant 0$, thus:

$$\begin{aligned}F(x) = \sum_{n=0}^{\infty} a_n x^n &= 1 + 2x + \sum_{n=2}^{\infty} a_n x^n \\ &= 1 + 2x + \sum_{n=0}^{\infty} (a_{n+1}+a_n) x^{n+2} \\ &= 1 + 2x + x\sum_{n=0}^{\infty} a_{n+1} x^{n+1} + x^2 \sum_{n=0}^{\infty} a_n x^n \\ &= 1 + 2x + x\sum_{n=1}^{\infty} a_n x^n + x^2 \sum_{n=0}^{\infty} a_n x^n \\ &= 1 + 2x + x(F(x)-1) + x^2 F(x).\end{aligned}$$

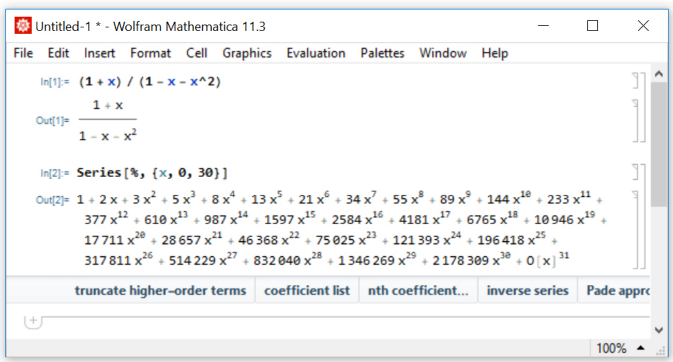
Thus $(1-x-x^2)F(x) = 1+x$, i.e.

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

This is a *rational function* of x, i.e. a quotient of two polynomials. From it we may extract each term a_n of the original sequence as the coefficient of x^n. We demonstrate how readily this is possible using Maple®:



or using Mathematica®:



Let us describe several ways to extract the coefficients a_n from this expression, thereby answering the counting problem for 11-free bitstrings of a given length. Which of these approaches is most appropriate depends the tools available (usually computer or hand computation) and the question we are trying to answer (such as, determining a_5? or a_{100}? or a_n for general n? Do we require only a recursive formula for a_n, or a closed formula? Do we require a simple asymptotic formula for a_n expressing roughly its rate of growth?) The following methods are possible:

(i) Start generating terms using the recursive formula for a_n.

(ii) Compute the n-th derivative $F^{(n)}(x)$ for the first few values of $n = 0, 1, 2, 3, \ldots$ and evaluate at 0 to obtain $a_n = \frac{F^{(n)}}{n!}$ as in Calculus II.

(iii) Expand the first few terms using a geometric series expansion for the denominator of $F(x)$.

(iii) Obtain a partial fraction decomposition for $F(x)$. Assuming the denominator factors with distinct roots, each term in the partial fraction decomposition expands as a geometric series, yielding an exact formula for a_n; also an asymptotic formula for a_n. If the denominator has repeated roots, essentially the same idea works if we use a more general binomial expansion in place of the geometric series expansion for each repeated root.

We have already described method (i). Method (ii) is mentioned for completeness, and because it will be familiar from Calculus II; however it suffers from the technical difficulty of computing higher order derivatives. To demonstrate method (iii), the first few terms in our sequence are easily generated by using a series expansion

$$\begin{aligned}F(x) &= (1+x)\big[1 + (x+x^2) + (x+x^2)^2 + (x+x^2)^3 + (x+x^2)^4 + \cdots\big]\\ &= (1+x)\big[1 + (x+x^2) + (x^2+2x^3+x^4) + (x^3+3x^4+3x^5+x^6)\\ &\qquad\qquad\qquad\qquad\qquad\qquad + (x^4+4x^5+6x^6+4x^7+x^8) + \cdots\big]\\ &= (1+x)\big[1 + x + 2x^2 + 3x^3 + 5x^4 + \cdots\big]\\ &= 1 + 2x + 3x^3 + 5x^4 + 8x^5 + \cdots.\end{aligned}$$

This gives a few terms without any trouble, but no exact general formula for a_n. Finally, we demonstrate method (iv). In order to decompose $F(x)$ into two terms with linear (rather than quadratic) denominators, we first factor the denominator as

$$1 - x - x^2 = (1-\alpha x)(1-\beta x)$$

where α and β are the reciprocal roots of the quadratic polynomial (i.e. $\frac{1}{\alpha}$ and $\frac{1}{\beta}$ are the actual roots of $1-x-x^2$). In particular for $\frac{1}{\alpha}$ to be a root, we require $1 - \frac{1}{\alpha} - \frac{1}{\alpha^2} = 0$ and so $\alpha^2 - \alpha - 1 = 0$, and similarly $\beta^2 - \beta - 1 = 0$. Thus

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2}\,.$$

It doesn't matter which root has the '+' sign and which has the '−' sign; we might as well take

$$\alpha = \frac{1+\sqrt{5}}{2} = 1.618\ldots; \qquad \beta = \frac{1-\sqrt{5}}{2} = -0.618\ldots.$$

For future reference, observe that $\alpha - \beta = \sqrt{5}$. The decimal approximations for α and β are not strictly needed here, but they are shown to satisfy our curiosity. Note that α is the famous irrational number known as the *golden ratio*. Next, as promised, we split $F(x)$ into two terms as

$$F(x) = \frac{1+x}{1-x-x^2} = \frac{1+x}{(1-\alpha x)(1-\beta x)} = \frac{A}{1-\alpha x} + \frac{B}{1-\beta x}$$

where A and B are constants. This decomposition, known as the *partial fraction decomposition of* $F(x)$, is often introduced in Calculus II as a technique for integrating rational functions. The fact that such constants A, B exist is a result in linear algebra; and while we do not require any knowledge of calculus, we do require you to know some linear algebra. In particular the constants A and B are found by solving two linear equations in two unknowns. Start by multiplying both sides of our formula for $F(x)$ by $1 - x - x^2 = (1 - \alpha x)(1 - \beta x)$ and cancelling factors where possible to obtain

$$1 + x = (1 - x - x^2)F(x) = (1 - \alpha x)(1 - \beta x)F(x) = (1 - \beta x)A + (1 - \alpha x)B.$$

This is an identity of polynomials. Evaluating at $x = \frac{1}{\alpha}$ (so that the last term vanishes) yields

$$\alpha = 1 + \tfrac{1}{\alpha} = \big(1 - \tfrac{\beta}{\alpha}\big)A = \tfrac{\alpha-\beta}{\alpha}A = \tfrac{\sqrt{5}}{\alpha}A\,,$$

so that $A = \frac{\alpha^2}{\sqrt{5}}$; and evaluating similarly at $x = \frac{1}{\beta}$ yields

$$\beta = 1 + \tfrac{1}{\beta} = \big(1 - \tfrac{\alpha}{\beta}\big)B = \tfrac{\beta-\alpha}{\beta}B = -\tfrac{\sqrt{5}}{\beta}B$$

so that $B = -\frac{\beta^2}{\sqrt{5}}$. This gives the partial fraction decomposition

$$F(x) = \frac{1+x}{1-x-x^2} = \frac{1}{\sqrt{5}}\Big(\frac{\alpha^2}{1-\alpha x} - \frac{\beta^2}{1-\beta x}\Big)\,.$$

Using our geometric series expansion, we obtain

$$\begin{aligned} F(x) &= \tfrac{\alpha^2}{\sqrt{5}}\Big(1 + \alpha x + \alpha^2x^2 + \alpha^3x^3 + \cdots\Big) - \tfrac{\beta^2}{\sqrt{5}}\Big(1 + \beta x + \beta^2x^2 + \beta^3x^3 + \cdots\Big) \\ &= \sum_{n\geqslant 0} \tfrac{\alpha^{n+2}-\beta^{n+2}}{\sqrt{5}}x^n. \end{aligned}$$

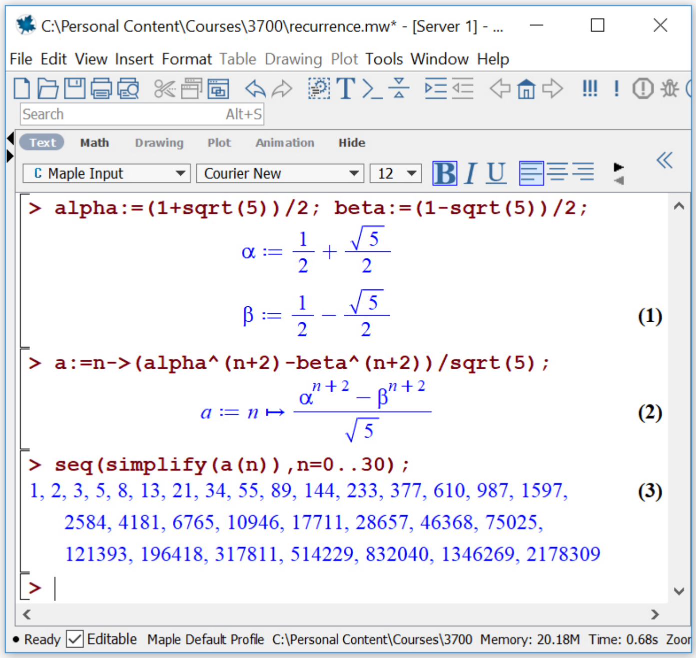
Our closed form expression for the n-th Fibonacci number is obtained by simply reading off the coefficient of x^n:

$$a_n = \frac{\alpha^{n+2} - \beta^{n+2}}{\sqrt{5}}\,.$$

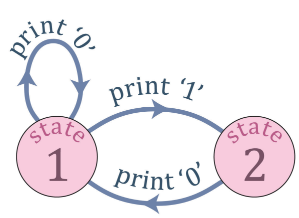
Note that for large values of n, the value of β^n tends to zero since $|\beta| < 1$. A consequence of this is that

$$a_n \approx \frac{\alpha^{n+2}}{\sqrt{5}};$$

in fact we may obtain a_n from $\frac{\alpha^{n+2}}{\sqrt{5}}$ by simply rounding off to the nearest integer. A consequence of this formula is the fact that a_n grows at an exponential rate, with the golden ratio α as the base of the exponential function. We demonstrate the use of the closed formula for a_n using Maple®:



Finally, we demonstrate an alternative derivation of the generating function $F(x)$ which does not begin with a proof of the recurrence formula as we have done. By this alternative method, the generating function (and implicitly, the recurrence formula, which depends only on the denominator of $F(x)$) pop out automatically! We conceive of a machine with two states (labelled states 1 and 2), as shown:



The machine begins in state 1; and at each time step, the state transitions by following one of the arrows, also printing out a bit as indicated by the chosen arrow. After n time steps, the machine will have printed an arbitrary 11-free bitstring of length n. So the number of 11-free bitstrings of length n is the number of possible computational paths our machine can take in n steps. If we view the diagram as a graph on 2 vertices, then a_n is the number of walks of length n starting at vertex 1. This is a directed graph with a loop. The adjacency matrix $A = \left[\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right]$ has as its (i,j)-entry the number of directed edges from vertex i to vertex j. The number of walks of length n from vertex i to vertex j is the (i,j)-entry of A^n.

Since the characteristic polynomial of A is $x^2 - x - 1 = (x-\alpha)(x-\beta)$, its eigenvalues are α and β. The corresponding eigenvectors are the columns of $M = \left[\begin{smallmatrix}\alpha & \beta\\ 1 & 1\end{smallmatrix}\right]$, i.e. $AM = MD$ where $D = \left[\begin{smallmatrix}\alpha & 0\\ 0 & \beta\end{smallmatrix}\right]$. We solve to obtain $A = MDM^{-1}$ where $M^{-1} = \frac{1}{\sqrt{5}}\left[\begin{smallmatrix}1 & -\beta\\ -1 & \alpha\end{smallmatrix}\right]$; so

$$\begin{aligned} A^n = (MDM^{-1})^n = MD^nM^{-1} &= \frac{1}{\sqrt{5}}\begin{bmatrix}\alpha & \beta\\ 1 & 1\end{bmatrix}\begin{bmatrix}\alpha^n & 0\\ 0 & \beta^n\end{bmatrix}\begin{bmatrix}1 & -\beta\\ -1 & \alpha\end{bmatrix}\\ &= \frac{1}{\sqrt{5}}\begin{bmatrix}\alpha^{n+1}-\beta^{n+1} & \alpha\beta^{n+1}-\alpha^{n+1}\beta\\ \alpha^n-\beta^n & \alpha\beta^n-\alpha^n\beta\end{bmatrix}. \end{aligned}$$

To obtain the number of walks of length n starting at vertex 1 (and ending at either vertex 1 or 2), we must add the $(1,1)$- and $(1,2)$-entries of A^n. Thus

$$\begin{aligned} a_n &= \tfrac{1}{\sqrt{5}}\left[(\alpha^{n+1}-\beta^{n+1}) + (\alpha\beta^{n+1}-\alpha^{n+1}\beta)\right]\\ &= \tfrac{1}{\sqrt{5}}\left[\alpha^{n+1}(1-\beta) - \beta^{n+1}(1-\alpha)\right]\\ &= \tfrac{1}{\sqrt{5}}\left(\alpha^{n+2} - \beta^{n+2}\right) \end{aligned}$$

since $\alpha+\beta = 1$. This of course agrees with the closed formula we had before. In order to obtain the generating function directly (and without factoring its denominator), note that the number of walks of length n from vertex i to vertex j is the (i,j)-entry of

$$(I - Ax)^{-1} = I + Ax + A^2x^2 + A^3x^3 + \cdots.$$

However,

$$(I - Ax)^{-1} = \begin{bmatrix}1-x & -x\\ -x & 1\end{bmatrix}^{-1} = \frac{1}{1-x-x^2}\begin{bmatrix}1 & x\\ x & 1-x\end{bmatrix},$$

so adding the $(1,1)$- and $(1,2)$-entries gives

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

as the generating function for a_n.