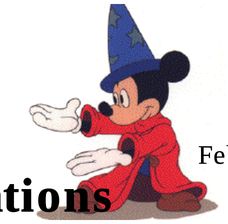
CS103A
Winter 2002

Handout 23
February 22, 2002

Solving Recurrence Relations



Introduction

A wide variety of recurrence problems occur in models. Some of these recurrence relations can be solved using iteration or some other ad hoc technique. However, one very important class of recurrence relations can be explicitly solved in a systematic way. These are the recurrence relations that express the terms of a sequence as a linear combination of previous terms.

Definition: A **linear homogeneous recurrence relation of degree k with constant coefficients** is a recurrence relation of the form:

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

where $c_1, c_2, c_3, \ldots, c_k$ are real numbers, and $c_k \neq 0$.

The recurrence relation in the definition is **linear** since the right-hand side is a sum of multiple of the previous terms of the sequence. The recurrence relation is **homogeneous** since no terms of the recurrence relation fail to involve a previous term of the sequence in some way. The coefficients of the terms of the sequence are all constants, rather than functions that depend on n. The degree is k because a_n is expressed in terms of the previous k terms of the sequence. A consequence of strong induction is that a sequence satisfying the recurrence relation in the definition is uniquely determined by this recurrence relation and the k initial condition:

$$\begin{aligned} a_n &= c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} \\ a_0 &= C_0 \\ a_1 &= C_1 \\ &\ldots\ldots \\ a_{k-1} &= C_{k-1} \end{aligned}$$

Solving the Little Monsters

The basic approach for solving linear homogeneous recurrence relations is to look for solutions of the form $a_n = r^n$, where r is a constant. Note that $a_n = r^n$ is a solution to the recurrence if and only if:

$$r^n = c_1 r^{n-1} + c_2 r^{n-2} + c_3 r^{n-3} + \cdots + c_k r^{n-k}$$

When both sides of the equation are divided by r^{n-k} and the right-hand side is subtracted from the left, we obtain the equivalent equation:

$$r^k - c_1 r^{k-1} - c_2 r^{k-2} - c_3 r^{k-3} - \cdots - c_{k-1} r - c_k = 0$$

Consequently, the sequence with $a_n = r^n$ is a solution **if and only if** r is a solution to this last equation, which is called the **characteristic equation** of the recurrence relation. The solutions of this equation are called **characteristic roots** of the recurrence relation. As we'll see, these characteristic roots can be used to give an explicit formula for all the solutions of the recurrence relation. First, we should develop results that deal with linear homogeneous recurrence relations with constant coefficients of degree two. The results for second order relations can be extended to solve higher-order equations. The mathematics involved in proving everything is really messy, so even though I'll refer to it in lecture, I don't want to place it in here, because you might incorrectly think you're responsible for knowing the proof, and you're not.

Let's turn our undivided attention to linear homogeneous recurrence relations of degree two. First, consider the case where there are two distinct characteristic roots.

A Fact Without Proof: Let c_1 and c_2 be real numbers. Suppose that $r^2 - rc_1 - c_2 = 0$ has two distinct roots r_1 and r_2. Then the sequence $\{a_n\}$ is a solution of the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ if and only if $a_n = b_1 r_1^2 + b_2 r_2^2$ for $n = 0,1,2,3,4,5,\ldots$ where b_1 and b_2 are constants determined by the initial conditions of the recurrence relation.

An Example

Problem: Find an explicit solution to the tiling recurrence relation we developed last time. You may recall that it went a-something like this:

$$T_n = T_{n-1} + T_{n-2}$$
$$T_0 = 1$$
$$T_1 = 1$$

Well, we surmise that the solution will be a linear combination or terms having the form $T_n = r^n$. Following the rule from above, we just plug in $T_n = r^n$ into out recurrence relation and see what constraints are placed on r. Let's listen in:

$$T_n = T_{n-1} + T_{n-2} \qquad r^n = r^{n-1} + r^{n-2}$$
$$\text{guess that } T_n = r^n \qquad r^2 = r + 1$$
$$r^2 - r - 1 = 0$$
$$r_1 = \frac{1+\sqrt{5}}{2}; r_2 = \frac{1-\sqrt{5}}{2}$$

Well, so the guess that $T_n = r^n$ is a good one provided that r be equal to one of the two roots above. In fact, any linear combination of $\left(\frac{1+\sqrt{5}}{2}\right)^n$ and $\left(\frac{1-\sqrt{5}}{2}\right)^n$ will also satisfy the recurrence if you ignore the initial conditions for a moment—that is, $T_n = b_1\left(\frac{1+\sqrt{5}}{2}\right)^n + b_2\left(\frac{1-\sqrt{5}}{2}\right)^n$ for any constant coefficients as factors. It's only when you supply the initial conditions that b_1 and b_2 are required to adopt a specific value. In fact, the initial conditions here dictate that:

$$b_1 + b_2 = 1$$
$$b_1 \frac{1+\sqrt{5}}{2} + b_2 \frac{1-\sqrt{5}}{2} = 1$$

And after solving for b_1 and b_2 do we finally arrive at the unique solution to our recurrence problem:

$$T_n = \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

How very satisfying.

Problem: Find a closed form solution to the bit string problem modeled in our last handout.

Welllllll, the recurrence relation, save the initial conditions, is exactly the same as it that for the tiling problem; that translates to a solution of the same basic form. But the fact that the initial conditions are different hints that the constants multiplying the individual terms:

$$B_n = b_1\left(\frac{1+\sqrt{5}}{2}\right)^n + b_2\left(\frac{1-\sqrt{5}}{2}\right)^n$$

note same form, but possibly different constants

The different initial conditions place different constraint on the values of b_1 and b_2 do we. Now the following system of equations must be solved:

$$b_1 + b_2 = 1$$
$$b_1 \frac{1+\sqrt{5}}{2} + b_2 \frac{1-\sqrt{5}}{2} = 2$$

We end up with something like this as our solution in this case:

$$B_n = \frac{5+3\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{5-3\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

Problem: Solve the recurrence relation given below

$$J_n = J_{n-2}$$
$$J_0 = 3J_1 = 5$$

You may be asking, "Jerry, where in the world is the J_{n-1} term?" Relax—it's in there—it's just that its multiplying coefficient is zero, so I didn't bother writing it in. It's still a homogeneous equation of degree two, so I solve it like any other recurrence of this type.

$$J_n = J_{n-2} \qquad J_0 = 3 \quad J_1 = 5$$

$$\begin{aligned} J_n &= J_{n-2} \\ r^2 &= 1 \\ r^2 - 1 &= 0 \\ r_1 &= 1; r_2 = -1 \\ J_n &= b_1(1)^n + b_2(-1)^n = b_1 + b_2(-1)^n \end{aligned}$$

Funny little twist—dropping the $(1)^n$, since it's transparent to us anyway.

$$\begin{aligned} b_1 + b_2 &= 5 \\ b_1 - b_2 &= \frac{5}{3} \\ b_1 = \frac{10}{3}&, b_2 = \frac{5}{3} \end{aligned} \qquad J_n = \frac{10}{3} + \frac{5}{3}(-1)^n$$

Solving Coupled Recurrence Relations

The first recurrences handout included examples where solutions involved coupled recurrence relations. The circular Tower of Hanoi Problem, the $n \times 3$ tiling problem, the $2 \times 2 \times n$ pillar problem, and the Martian DNA problem were all complex enough to require (or at least benefit from) the invention of a second counting problem and a second recurrence variable. Remember the $2 \times 2 \times n$ pillar recurrence? If not, here it is once more:

$$S_n = \begin{cases} 1 & n = 0 \\ 2 & n = 1 \\ 2S_{n-1} + S_{n-2} + 4T_{n-1} & n \geq 2 \end{cases} \qquad T_n = \begin{cases} 0 & n = 0 \\ 1 & n = 1 \\ S_{n-1} + T_{n-1} & n \geq 2 \end{cases}$$

Because S and T are defined in terms of one another, it's possible to use the second recurrence of the two to eliminate all occurrences or T from the first. It takes a algebra and a little ingenuity, but when we rid of the second variable, we are often left with a single recurrence relation which can be solved like other recurrences we've seen before. The above system of recurrences reduces to a single linear, homogeneous, constant-coefficient equation once we get rid of T. Don't believe me? Read on, Thomas.

Problem: Solve the above system of recurrences for S_n by eliminating T_n and solving the linear equation that results.

I want to completely eliminate all traces of T and arrive at (what will turn out to be) a (cubic) recurrence relation for just S_n, and then solve it. First things first: We want to get rid of the T_{n-1} term from the recurrence relation for S_n, and to do that, I make the neat little observation that the second equality below follows from the first by replacing n by n-1:

$$\begin{aligned} S_n &= 2S_{n-1} + S_{n-2} + 4T_{n-1} \\ S_{n-1} &= 2S_{n-2} + S_{n-3} + 4T_{n-2} \end{aligned}$$

Subtracting the second equation from the first, we get:

$$\begin{aligned} S_n - S_{n-1} &= \left(2S_{n-1} + S_{n-2} + 4T_{n-1}\right) - \left(2S_{n-2} + S_{n-3} + 4T_{n-2}\right) \\ &= 2S_{n-1} - S_{n-2} - S_{n-3} + 4T_{n-1} - 4T_{n-2} \\ S_n &= 3S_{n-1} - S_{n-2} - S_{n-3} + 4T_{n-1} - 4T_{n-2} \\ &= 3S_{n-1} - S_{n-2} - S_{n-3} + 4\left(T_{n-1} - T_{n-2}\right) \end{aligned}$$

Believe it or not, this is progress, because I have something very interesting to say about $T_{n-1} - T_{n-2}$ —it's always equal to S_{n-2}. Just look at the crafty manipulations I work up from the T_n recurrence relation:

$$\begin{aligned} T_n &= S_{n-1} + T_{n-1} \\ T_{n-1} &= S_{n-2} + T_{n-2} \\ T_{n-1} - T_{n-2} &= S_{n-2} + T_{n-2} - T_{n-2} \\ &= S_{n-2} \end{aligned}$$

How crafty! Now I can eradicate any mention of T from the S_n recurrence relation, and I do so like-a this:

$$\begin{aligned} S_n &= 3S_{n-1} - S_{n-2} - S_{n-3} + 4\left(T_{n-1} - T_{n-2}\right) \\ &= 3S_{n-1} - S_{n-2} - S_{n-3} + 4S_{n-2} \\ &= 3S_{n-1} + 3S_{n-2} - S_{n-3} \end{aligned}$$

Therefore, an uncoupled recurrence relation for S_n can be expressed as follows:

$$S_n = \begin{cases} 1 & n = 0 \\ 2 & n = 1 \\ 2S_1 + S_0 + 4T_1 = 9 & n = 2 \\ 3S_{n-1} + 3S_{n-2} - S_{n-3} & n \ge 3 \end{cases}$$

Admittedly, that was a lot of work, but this is pretty exciting, because we're on the verge of taking what is clearly a homogeneous, constant-coefficient equation and coming up with a closed form solution. As always, guess a solution of the form $S_n = a^n$ and substitute to see what values of a make everything work out regardless of the initial conditions. More algebra:

$$\begin{aligned} S_n\big|_{S_n = a^n} &= 3S_{n-1} + 3S_{n-2} - S_{n-3}\big|_{S_n = a^n} \\ a^n &= 3a^{n-1} + 3a^{n-2} - a^{n-3} \\ a^3 &= 3a^2 + 3a - 1 \\ 0 &= a^3 - 3a^2 - 3a + 1 \\ 0 &= (a+1)\left(a^2 - 4a + 1\right) \\ 0 &= (a+1)\left(a - 2 + \sqrt{3}\right)\left(a - 2 - \sqrt{3}\right) \\ a &= -1,\ 2 \pm \sqrt{3} \end{aligned}$$

That means that the general solution to our recurrence (ignoring the initial conditions) is $S_n = x\left(2+\sqrt{3}\right)^n + y\left(2-\sqrt{3}\right)^n + z(-1)^n$, where x, y, and z can be any real numbers. Boundary conditions require that:

$$S_0 = x + y + z = 1$$
$$S_1 = \left(2+\sqrt{3}\right)x + \left(2-\sqrt{3}\right)y - z = 2$$
$$S_2 = \left(7+4\sqrt{3}\right)x + \left(1-4\sqrt{3}\right)y + z = 9$$

This is a system of three linear equations for three unknowns, and it admits exactly one solution. Solving for x, y, and z is a matter of algebra, and after all that algebra is over, we converge on a closed formula of $S_n = \frac{1}{6}\left(2+\sqrt{3}\right)^{n+1} + \frac{1}{6}\left(2-\sqrt{3}\right)^{n+1} + \frac{1}{3}(-1)^n$, which is $\frac{1}{6}\left(2+\sqrt{3}\right)^{n+1}$ rounded to the nearest integer. Neat!

Problem: Revisit the Martian DNA problem, and show that the number of valid Martian DNA strands of length n is given as $a_n + b_n = F_{3n+2}$ (yes, the Fibonacci number!)

Let's bring back the recurrence that defined a and b.

$$a_n = \begin{cases} 0 & n = 0 \\ 2 & n = 1 \\ a_{n-1} + 2b_{n-1} & n \geq 2 \end{cases} \quad b_n = \begin{cases} 1 & n = 0 \\ 3 & n = 1 \\ 2a_{n-1} + 3b_{n-1} & n \geq 2 \end{cases}$$

Eliminate one of the recurrence terms in order to get a closed solution (though this time we'll ultimately need to solve for both variables, because you're interested in their sum.)

$$a_n = a_{n-1} + 2b_{n-1}$$
$$b_n = 2a_{n-1} + 3b_{n-1}$$

Notice that the first one tells us something about $a_n - a_{n-1}$, so let's subtract a shifted version of the second equation from the original to get at something that's all b and no a.

$$b_n = 2a_{n-1} + 3b_{n-1}$$
$$b_{n-1} = 2a_{n-2} + 3b_{n-2}$$
$$b_n - b_{n-1} = 2\left(a_{n-1} - a_{n-2}\right) + 3b_{n-1} - 3b_{n-2}$$
$$b_n = b_{n-1} + 2\left(a_{n-1} - a_{n-2}\right) + 3b_{n-1} - 3b_{n-2}$$

The first equation tells us that $a_{n-1} - a_{n-2} = 2b_{n-2}$. Substitution yields

$$b_n = b_{n-1} + 4b_{n-2} + 3b_{n-1} - 3b_{n-2}$$
$$= 4b_{n-1} + b_{n-2}$$

The characteristic equation here is $r^2 - 4r - 1 = 0$; $b_n = c_1\left(2+\sqrt{5}\right)^n + c_2\left(2-\sqrt{5}\right)^n$. Because $b_0 = 1$ and $b_1 = 3$, we determine c_1, c_2 by solving the following two equations:

$$\begin{aligned} c_1 + c_2 &= 1 \\ \left(2+\sqrt{5}\right)c_1 + \left(2-\sqrt{5}\right)c_2 &= 3 \end{aligned} \qquad c_1 = \frac{5+\sqrt{5}}{10},\ c_2 = \frac{5-\sqrt{5}}{10}.$$

$$b_n = \frac{5+\sqrt{5}}{10}\left(2+\sqrt{5}\right)^n + \frac{5-\sqrt{5}}{10}\left(2-\sqrt{5}\right)^n.$$

You might think you have to do the same thing for a, and in theory you're right in that you need to solve it, but we more or less have. Because

$$b_n = 2a_{n-1} + 3b_{n-1}$$

we actually know that

$$\begin{aligned} b_n &= 2a_{n-1} + 3b_{n-1} \\ 2a_{n-1} &= b_n - 3b_{n-1} \\ a_{n-1} &= \frac{1}{2}b_n - \frac{3}{2}b_{n-1} \\ a_n &= \frac{1}{2}b_{n+1} - \frac{3}{2}b_n \\ a_n + b_n &= \frac{1}{2}b_{n+1} - \frac{3}{2}b_n + b_n \\ &= \frac{1}{2}b_{n+1} - \frac{1}{2}b_n \end{aligned}$$

Recall that we're really just interested in $a_n + b_n$, and since it can be defined just in terms of the b_n, we get:

$$\begin{aligned} a_n + b_n &= \frac{1}{2}b_{n+1} - \frac{3}{2}b_n + b_n \\ &= \frac{1}{2}b_{n+1} - \frac{1}{2}b_n \\ &= \frac{1}{2}\left(\frac{5+\sqrt{5}}{10}\left(2+\sqrt{5}\right)^{n+1} + \frac{5-\sqrt{5}}{10}\left(2-\sqrt{5}\right)^{n+1}\right) - \frac{1}{2}\left(\frac{5+\sqrt{5}}{10}\left(2+\sqrt{5}\right)^{n} + \frac{5-\sqrt{5}}{10}\left(2-\sqrt{5}\right)^{n}\right) \\ &= \frac{1}{2}\frac{5+\sqrt{5}}{10}\left(\left(2+\sqrt{5}\right)-1\right)\left(2+\sqrt{5}\right)^n + \frac{1}{2}\frac{5-\sqrt{5}}{10}\left(\left(2-\sqrt{5}\right)-1\right)\left(2-\sqrt{5}\right)^n \\ &= \frac{1}{2}\frac{5+\sqrt{5}}{10}\left(1+\sqrt{5}\right)\left(2+\sqrt{5}\right)^n + \frac{1}{2}\frac{5-\sqrt{5}}{10}\left(1-\sqrt{5}\right)\left(2-\sqrt{5}\right)^n \\ &= \frac{1}{2}\frac{10+6\sqrt{5}}{10}\left(2+\sqrt{5}\right)^n + \frac{1}{2}\frac{10-6\sqrt{5}}{10}\left(2-\sqrt{5}\right)^n \\ &= \frac{5+3\sqrt{5}}{10}\left(2+\sqrt{5}\right)^n + \frac{5-3\sqrt{5}}{10}\left(2-\sqrt{5}\right)^n \end{aligned}$$

That's typically the closed-form you'd leave it in, but we also want to show that $a_n + b_n = F_{3n+2}$. Here goes:

$$
\begin{aligned}
F_{3n+2} &= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{3n+2} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{3n+2} \\
&= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{2}\left(\frac{1+\sqrt{5}}{2}\right)^{3n} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{2}\left(\frac{1-\sqrt{5}}{2}\right)^{3n} \\
&= \frac{1}{\sqrt{5}}\,\frac{6+2\sqrt{5}}{4}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{3}\right)^{n} - \frac{1}{\sqrt{5}}\,\frac{6-2\sqrt{5}}{4}\left(\left(\frac{1-\sqrt{5}}{2}\right)^{3}\right)^{n} \\
&= \frac{6+2\sqrt{5}}{4\sqrt{5}}\left(2+\sqrt{5}\right)^{n} - \frac{6-2\sqrt{5}}{4\sqrt{5}}\left(2+\sqrt{5}\right)^{n} \\
&= \frac{\sqrt{5}}{\sqrt{5}}\,\frac{6+2\sqrt{5}}{4\sqrt{5}}\left(2+\sqrt{5}\right)^{n} - \frac{\sqrt{5}}{\sqrt{5}}\,\frac{6-2\sqrt{5}}{4\sqrt{5}}\left(2+\sqrt{5}\right)^{n} \\
&= \frac{10+6\sqrt{5}}{20}\left(2+\sqrt{5}\right)^{n} + \frac{10-6\sqrt{5}}{20}\left(2+\sqrt{5}\right)^{n} \\
&= \frac{5+3\sqrt{5}}{10}\left(2+\sqrt{5}\right)^{n} + \frac{5-3\sqrt{5}}{20}\left(2+\sqrt{5}\right)^{n}
\end{aligned}
$$

Therefore, $a_n + b_n = F_{3n+2}$, and we rejoice.

Problem: What about the Circular Tower of Hanoi recurrence? Why can't we solve that one as easily?

Well, you probably already noticed that both the pillar and the DNA recurrences didn't have any inhomogeneous terms anywhere. That's not the case with the Circular Tower of Hanoi recurrence, so while we are certainly invited to eliminate one of the variables from the set of recurrences, there's not much hope for solving it like we did for Martian DNA.

$$
Q_n = \begin{cases} 0; & \text{if } n = 0 \\ 2R_{n-1}+1 & \text{if } n > 0 \end{cases} \qquad R_n = \begin{cases} 0; & \text{if } n = 0 \\ Q_n + Q_{n-1}+1 & \text{if } n > 0 \end{cases}
$$

Clearly, it's a piece of cake to define Q_n in terms of previous Qs, but these 1s just aren't going to go away.

$$
\begin{aligned}
Q_n &= 2R_{n-1}+1 \\
&= 2\left(Q_{n-1}+Q_{n-2}+1\right)+1 \\
&= 2Q_{n-1}+2Q_{n-2}+3
\end{aligned}
$$

so that

$$
Q_n = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \\ 2Q_n + 2Q_{n-1}+3 & \text{if } n > 1 \end{cases}
$$

We're sorta bumming because of that inhomogeneous 3.

There is a trick to solving this one, but it's not something I'm formally going to require you to know, since there's little pedagogical value in memorizing tricks. For those of you with nothing to do on a Friday night, try substituting $Q_n = A_{n-1} - 1$ into the above system (making sure to adjust those base cases as well.) Solve for A_n, then take whatever you got there and subtract 1 from it to get Q_n.

General Approach to Solving all Linear First-Order Recurrences

There is a general technique that can reduce virtually any recurrence of the form

$$a_n T_n = b_n T_{n-1} + c_n$$

to a sum. The idea is to multiply both sides by a summation factor, s_n, to arrive at $s_n a_n T_n = s_n b_n T_{n-1} + s_n c_n$. This factor s_n is chosen to make $s_n b_n = s_{n-1} a_{n-1}$. Then if we write $S_n = s_n a_n T_n$ we have a sum-recurrence for S_n as:

$$S_n = S_{n-1} + s_n c_n$$

Hence $S_n = s_0 a_0 T_0 + \sum_{1\le k\le n} s_k c_k = s_1 b_1 T_0 + \sum_{1\le k\le n} s_k c_k$ and the solution to the original recurrence becomes

$$T_n = \frac{1}{s_n a_n}\left(s_1 b_1 T_0 + \sum_{1\le k\le n} s_k c_k\right).$$

So, you ask: How can we be clever enough to find the perfect s_n? Well, the relation $s_n = s_{n-1} a_{n-1} / b_n$ can be unfolded by repeated substitution for the s_i to tell us that the fraction:

$$s_n = \frac{a_{n-1} a_{n-2} a_{n-3} \ldots a_1}{b_n b_{n-1} b_{n-2} \ldots b_2}$$

(or any convenient constant multiple of this value) will be a suitable summation factor.

Problem: Solve $V_n = \begin{cases} 5 & n = 0 \\ nV_{n-1} + 3\ n! & n \ge 1 \end{cases}$ by finding the appropriate summation factor.

Derivation of the solution just follows protocol. You're specifically told to use summation factors here, so we should do that. Notice here that $a_n = 1$ and $b_n = n$, so that the summation factor should be chosen as:

$$s_n = \frac{1^{n-1}}{n!} = \frac{1}{n!}$$

$V_0 = 5$ for sure, but the recurrence formula $V_n = nV_{n-1} + 3\ n!$ becomes $\frac{1}{n!} V_n = \frac{1}{(n-1)!} V_{n-1} + 3$. Let $T_n = \frac{1}{n!} V_n$ for the time being, just so we can solve an easier recurrence relation. We tackle $T_n = T_{n-1} + 3$, and wouldn't you know it: $T_n = 3n + T_0 = 3n + \frac{1}{0!} V_0 = 3n + 5$. $T_n = \frac{1}{n!} V_n$. We need $V_n = n! T_n$, so therefore we have that $V_n = n!(3n + 5)$. Woo!

An Example

Problem: Solve $C_n = \begin{cases} 0 & n = 0 \\ n + 1 + \frac{2}{n}\sum_{0 \le k \le n-1} C_k & n \ge 1 \end{cases}$ by finding the appropriate summation factor.

Let's write out the recurrences for C_n and C_{n-1} a little more explicitly.

$$C_n = n + 1 + \frac{2}{n}\sum_{0 \le k \le n-1} C_k$$
$$C_{n-1} = n + \frac{2}{n-1}\sum_{0 \le k \le n-2} C_k$$

Subtracting the second one from the first one is a good idea, but only after the multiply through by a factor that'll make each of the summations equal to each other:

$$\begin{aligned} C_n &= n + 1 + \frac{2}{n}\sum_{0 \le k \le n-1} C_k \\ \frac{n-1}{n} C_{n-1} &= \frac{n-1}{n} n + \frac{n-1}{n}\frac{2}{n-1}\sum_{0 \le k \le n-2} C_k \\ &= n - 1 + \frac{2}{n}\sum_{0 \le k \le n-2} C_k \end{aligned}$$

Subtracting the second one from the first, we arrive at:

$$\begin{aligned} C_n - \frac{n-1}{n} C_{n-1} &= \left(n + 1 + \frac{2}{n}\sum_{0 \le k \le n-1} C_k\right) - \left(n - 1 + \frac{2}{n}\sum_{0 \le k \le n-2} C_k\right) \\ &= 2 + \frac{2}{n} C_{n-1} \\ C_n &= \frac{n+1}{n} C_{n-1} + 2 \\ nC_n &= (n+1)C_{n-1} + 2n \end{aligned}$$

If we choose a summation factor of $s_n = \frac{a_{n-1}a_{n-2}a_{n-3}\ldots a_1}{b_n b_{n-1} b_{n-2}\ldots b_2} = \frac{(n-1)(n-2)(n-3)\ldots 1}{(n+1)n(n-1)\ldots 3} = \frac{2}{n(n+1)}$ and multiply through, we arrive at:

$$\begin{aligned} \frac{2}{n(n+1)} nC_n &= \frac{2}{n(n+1)}(n+1)C_{n-1} + \frac{2}{n(n+1)} 2n \\ \frac{2}{(n+1)} C_n &= \frac{2}{n} C_{n-1} + \frac{4}{(n+1)} \\ \frac{C_n}{n+1} &= \frac{C_{n-1}}{n} + \frac{2}{n+1} \end{aligned}$$

If we let $(n+1)D_n = C_n$, then we finally arrive at an equation that looks reasonable: $D_n = D_{n-1} + \frac{2}{n+1}$.

Repeated substitution yields the following:

$$
\begin{aligned}
D_n &= D_{n-1} + \frac{2}{n+1} \\
&= \left(D_{n-2} + \frac{2}{n}\right) + \frac{2}{n+1} \\
&= \left(D_{n-3} + \frac{2}{n-1}\right) + \frac{2}{n} + \frac{2}{n+1} \\
&= \left(D_{n-4} + \frac{2}{n-2}\right) + \frac{2}{n-1} + \frac{2}{n} + \frac{2}{n+1} \\
&\vdots \\
&= D_0 + 2\sum_{1 \le k \le n} \frac{1}{k+1} \\
&= 2\left(H_{n+1} - 1\right) \\
&= 2\left(H_n + \frac{1}{n+1} - 1\right) = 2\left(H_n - \frac{n}{n+1}\right)
\end{aligned}
$$

Recall that $C_n = (n+1)D_n$, so that $C_n = 2(n+1)\left(H_n - \frac{n}{n+1}\right) = 2(n+1)H_n - 2n$.