Queen's University
School of Computing

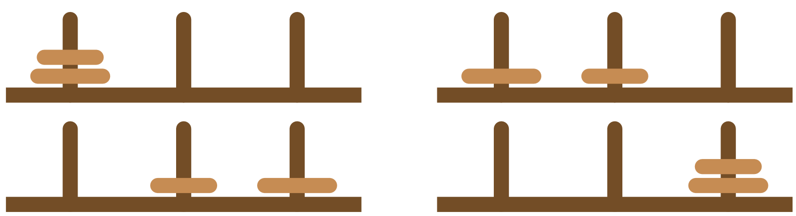
CISC 203: Discrete Mathematics for Computing II
Lecture 4: Recurrence Relations
Winter 2019

1 An Ancient Puzzle

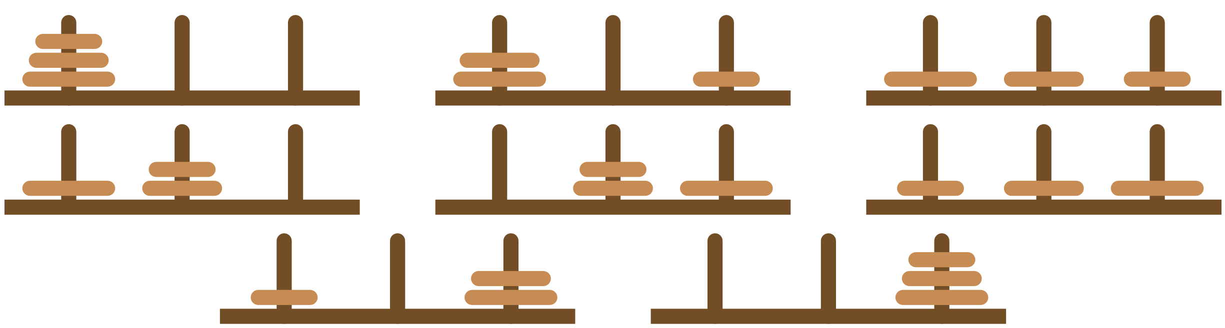
Legend has it that, in the city of Varanasi, India, a group of priests live in the Kashi Vishwanath Temple alongside a priceless artifact. This artifact consists of a base with three pegs holding 64 golden disks, each trimmed with diamonds. In accordance with a Hindu prophecy, the priests take turns moving disks at a rate of one per second in such a way that no larger disk is placed on top of a smaller one. The priests' ultimate goal is to transfer all of the disks from the first peg to the third peg, with all disks ordered by size. The priests have been moving these disks since the beginning of time, and once this act is done, the world will come to an end.

What is the least amount of time it will take for the priests to accomplish their goal? Let's begin by considering a few smaller cases to see if a pattern develops in the number of moves. With only one disk, the priests need to perform only one move: from peg 1 to peg 3 directly.

With two disks, we require a minimum of three moves (with the initial state as the first illustration):



With three disks, we require a minimum of seven moves (with the initial state as the first illustration):



So, there seems to be a pattern developing. If we were to consider a few more cases, we would see that the sequence representing the minimum number of moves is $\{1, 3, 7, 15, 31, 63, 127, \dots\}$, or $2^n - 1$ for n disks.

The priests, having heard this news, are understandably displeased; moving disks at a rate of one per second, it would take the priests $2^{64} - 1 = 18\,446\,744\,073\,709\,551\,615$ seconds to finish. That's almost 585 billion years! The priests want to be sure the sequence we obtained is correct, so they demand that we prove it. In this lecture, we will see a number of methods to prove this result and many others.

2 Defining Recurrences

In an earlier lecture, we touched upon the notion of a **sequence**, but we never concretely defined what sequences were. Let's do that now.

Definition 1 (Sequence). A sequence $A_n = \{a_1, a_2, \ldots, a_n\}$ is a relation from a subset of the set of natural numbers, $\mathbb{N}$, to a set A. Each element a_i of A_n is called a term of the sequence.

Basically, a sequence is just like any other relation, but with different notation. Strictly speaking, our definition of a sequence describes a finite sequence consisting of n terms, but we can also define infinite sequences: two simple examples of infinite sequences are the sets of numbers $\mathbb{N}$ and $\mathbb{Z}$.

Although we define sequences in most cases by listing the terms directly, certain sequences can be defined recursively, where a given term of the sequence is generated from previous terms of the same sequence according to some relation. When a sequence can be defined in such a way, we call the relation that generates terms of the sequence a **recurrence relation** or a **recurrence** for short.

Definition 2 (Recurrence relation). A recurrence relation is a sequence A_n where each term of the sequence is either given as an initial term or produced from one or more previous terms.

Thinking in the other direction, we say that a sequence is a **solution** of a recurrence relation if the terms of the sequence satisfy the recurrence relation. A recurrence relation together with a set of initial terms uniquely defines a sequence, so there exists only one sequence that satisfies a given recurrence relation.

We've already seen a few examples of recurrence relations in earlier lectures. Let's revisit these examples to see how they correspond to our definition.

Example 3. The Fibonacci sequence F is the infinite sequence satisfying the following conditions:

1. $F_1 = 1$;
2. $F_2 = 1$; and
3. $F_n = F_{n-1} + F_{n-2}$ for all $n \in \mathbb{N}$ where $n \geq 3$.

The sequence F_n is generated by the relation $F_n = F_{n-1} + F_{n-2}$ with initial terms F_1 and F_2.

Example 4. The sequence of binomial coefficients B is the infinite sequence satisfying the following conditions:

1. $\binom{n}{0} = 1$ for all $n \in \mathbb{N}$;
2. $\binom{n}{n} = 1$ for all $n \in \mathbb{N}$; and
3. $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ for all $1 \leq k \leq n$.

The sequence $B_{n,k}$ is generated by the relation $B_{n,k} = B_{n-1,k} + B_{n-1,k-1}$ with initial terms $B_{n,0}$ and $B_{n,n}$.

In both of our examples, we obtain new terms of the sequence recursively by adding together two previous terms of the sequence. We are also given two initial terms for each sequence, which we use to obtain the first recursive term of that sequence. However, we can easily define recurrence relations that use more or fewer than two previous terms. In general, if a recurrence relation produces new terms from k previous terms, we say that the **degree** of the recurrence relation is k. We may also refer to such a recurrence relation as a ***k*th-order** recurrence relation. (Thus, the recurrence relations in Examples 3 and 4 are second-order recurrence relations.)

Example 5. We can define common functions using recurrence relations. Consider the exponential function 2^n. Each value of 2^n is given by the first-order recurrence relation $a_n = 2 \cdot a_{n-1}$, and the initial term of the recurrence relation is $a_1 = 1$.

Example 6. How can we define a more interesting function, like $n!$, in terms of a recurrence relation? Simply change the coefficient in the recurrence relation. Each value of $n!$ is given by the first-order recurrence relation $b_n = n \cdot b_{n-1}$, and the initial term of the recurrence relation is $b_1 = 1$.

Observe that, in Examples 5 and 6, the recurrence a_n had a coefficient of 2 for all terms, while the recurrence b_n had a coefficient of n for the nth term. We say that a_n is a recurrence relation with **constant coefficients**; that is, coefficients that do not depend on n. The recurrence relation b_n, however, contains the non-constant coefficient n.

Example 7. Consider the recurrence relation $d_n = d_{n-1} + 1$ with initial term $d_1 = 1$. This first-order, constant-coefficient recurrence relation simply produces the range of numbers from 1 to n.

Example 8. By modifying the Fibonacci recurrence relation to multiply previous terms together, we get a slightly more interesting recurrence relation. Call this recurrence relation $e_n = e_{n-1} \cdot e_{n-2}$, and define the initial terms to be $e_1 = 1$ and $e_2 = 2$. This second-order recurrence relation produces the sequence $\{1, 2, 2, 4, 8, 32, 256, 8192, \dots\}$.

We see even more variations on recurrence relations in Examples 7 and 8. In the recurrence relation d_n, for instance, not all terms in the formula come from previous terms of the sequence d; one of the terms is the constant 1. We say that a recurrence relation where all terms have the same degree is **homogeneous** (and, thus, the recurrence relation d_n is non-homogeneous, since the term 1 is of lower degree than d_{n-1}).

Remark. When we talk about homogeneity, we use the word "degree" to mean "exponent of the term", not "degree of a recurrence relation" as we saw earlier.

In the recurrence relation e_n, we get new terms by multiplying previous terms instead of adding. If a recurrence relation lists previous values as independent terms raised to the first power, then we say it is **linear** (and, thus, the recurrence relation e_n is non-linear, since the terms e_{n-1} and e_{n-2} are not independent).

3 Solving Recurrences

Now that we are familiar with defining recurrence relations, we turn to the big question of this lecture: how do we solve recurrence relations? Here, we use the word "solve" in the sense of finding a closed-form equation that is equivalent to the recursively-defined recurrence relation. There exist a number of approaches we can take when solving a recurrence relation, and these approaches range from the simple to the complex. Our choice of approach ultimately depends on the recurrence relation we are trying to solve.

Before we begin solving, however, let's determine the recurrence relation for the pegs-and-disks problem. Notice that each term of the sequence $H = \{1, 3, 7, 15, 31, 63, 127, \dots\}$ seems to be twice the previous term plus one; for instance, $2(1) + 1 = 3$, $2(3) + 1 = 7$, and so on. This suggests to us that the sequence H can be generated recursively from smaller instances of the same problem. Indeed, this is how we will find the recurrence relation: since we have three pegs, we can model the problem recursively by moving all but the largest disk from the first peg to the second peg, then moving the largest disk to the third peg before placing all of the other disks on the third peg as well.

Lemma 9. *Given n disks, a minimum of $H_n = 2H_{n-1} + 1$ moves are required to transfer all disks from peg 1 to peg 3, where $H_1 = 1$.*

Proof. We begin with all n disks on peg 1. Move the top $n - 1$ disks from peg 1 to peg 2. This step requires H_{n-1} moves. Then, move the nth disk (that is, the largest disk) to peg 3. Finish by moving the top $n - 1$ disks from peg 2 to peg 3. Altogether, this process requires a total of $2H_{n-1} + 1$ moves. □

You might question whether this bound is the true minimum. Indeed, this bound is the best possible; given a finite number of disks, say n, the largest disk must be moved at some point. To move the largest disk, the $n - 1$ smaller disks must be moved out of the way first. Then, to complete the puzzle, the $n - 1$ smaller disks must be moved back on top of the largest disk. Altogether, this process requires at least H_n moves.

Remark. Despite the fact that our earlier story was set in India, the pegs-and-disks problem is frequently referred to as the **tower of Hanoi** problem, named for the Vietnamese city of Hanoi. The tower of Hanoi problem was invented by the French mathematician Édouard Lucas, who was well-known for studying recurrence relations and the Fibonacci sequence.

3.1 Substitution Method

By far the simplest method of solving recurrence relations, the **substitution method** (also called the "guess and check" method) exploits the fact that recurrence relations are recursively defined. Since a recurrence relation gives us initial terms and a recursive term that breaks the problem down into smaller pieces, we could find a solution to the recurrence relation by induction: take the initial terms to be the base case, and take the recursive term to be the inductive case.

With the substitution method, we guess a solution to the recurrence relation and check whether it is correct by induction. Though this method is simple, it is not the most straightforward or intuitive method; there is no general heuristic to help in making a right guess, making a wrong guess is very easy, and a wrong guess means we have to start from scratch. However, with practice, the substitution method can become a quick way to check a potential solution without having to put in as much work as other methods require.

Example 10. By Lemma 9, we know that $H_n = 2H_{n-1} + 1$ and $H_1 = 1$. What is a closed form for H_n?

Let's make a guess... looking at the sequence H, it appears that the nth term is equal to $2^n - 1$. Is our guess correct? Check using a proof by induction. Let $P(n)$ be the statement "$H_n = 2^n - 1$".

Base case: When $n = 1$, we have $2^1 - 1 = 1$. Since $H_1 = 1$, $P(1)$ is true.

Inductive hypothesis: Assume that $P(k)$ is true for some $k \in \mathbb{N}$ where $k \geq 1$. That is, assume that $H_k = 2^k - 1$.

Inductive case: We now show that $P(k+1)$ is true. By the inductive hypothesis, we have $H_k = 2^k - 1$. By our recurrence relation, we have

$$\begin{aligned} H_{k+1} &= 2H_k + 1 \\ &= 2(2^k - 1) + 1 \\ &= 2^{k+1} - 2 + 1 \\ &= 2^{k+1} - 1. \end{aligned}$$

Therefore, $P(k+1)$ is true. By the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$.

Note that, while the substitution method is good for proving easy solutions to recurrence relations, it can be a struggle to prove even a moderately-difficult solution using substitution. In such cases, you might try to show that the recurrence relation is bounded from above by a looser but nicer-looking expression. However, this approach comes with its own set of potential traps. Returning to Example 10, what if we had tried to obtain a loose upper bound of 2^n on our recurrence relation H_n? Our "proof" would have worked right up until the inductive step, where we would have encountered the following problem:

$$\begin{aligned} H_{k+1} &= 2H_k + 1 \\ &\leq 2(2^k) + 1 \\ &\leq 2^{k+1} + 1 \text{ (?)} \\ &\not\leq 2^{k+1} \qquad \text{(??)} \end{aligned}$$

Although we can certainly prove upper bounds for recurrence relations, this example should serve as a fair warning to be careful about the approach you use to do so.

3.2 Iteration Method

As opposed to the substitution method, which uses a proof by induction to find the closed form of a recurrence relation, the **iteration method** uses the recurrence relation itself to find its corresponding closed form. This method works by iteratively replacing occurrences of smaller terms in the recurrence relation with the corresponding equation for that term, then simplifying the expression. Once the initial terms are reached, the entire expression simplifies to the closed-form equation.

Example 11. By Lemma 9, we know that $H_n = 2H_{n-1} + 1$ and $H_1 = 1$. What is a closed form for H_n?

By the iteration method, we have

$$\begin{aligned}
H_n &= 2H_{n-1} + 1 \\
&= 2(\,2H_{n-2} + 1\,) + 1 = 2^2 H_{n-2} + 2 + 1 \\
&= 2^2(\,2H_{n-3} + 1\,) + 2 + 1 = 2^3 H_{n-3} + 2^2 + 2 + 1 \\
&\vdots \\
&= 2^{n-2}(\,2H_1 + 1\,) + 2^{n-3} + \cdots + 2 + 1 \\
&= 2^{n-1} H_1 + 2^{n-2} + \cdots + 2 + 1 \\
&= 2^{n-1} + 2^{n-2} + \cdots + 2 + 1 \\
&= 2^n - 1.
\end{aligned}$$

The second-last line of the previous derivation is the sum of a geometric series: $\sum_{i=0}^{n-1} 2^i = \frac{2^{(n-1)+1}-1}{2-1} = 2^n - 1$.

3.2.1 First-Order Recurrence Relations

First-order recurrence relations are some of the easiest recurrence relations to deal with. Remember that we say a recurrence relation is "first-order" or "degree 1" if it produces new terms from only the previous term. In mathematical terms, a first-order recurrence relation is of the form

$$a_n = ca_{n-1} + x,$$

where the coefficient c and the additive term x are constants.

We can use the iteration method to obtain a general closed form for first-order recurrence relations. Observe that

$$\begin{aligned}
a_n &= ca_{n-1} + x \\
&= c(\,ca_{n-2} + x\,) + x \\
&= c^2(\,ca_{n-3} + x\,) + cx + x \\
&\vdots \\
&= c^{n-1}(\,ca_0 + x\,) + c^{n-2}x + c^{n-3}x + \cdots + cx + x \\
&= c^n a_0 + c^{n-1}x + c^{n-2}x + c^{n-3}x + \cdots + cx + x \\
&= c^n a_0 + (c^{n-1} + c^{n-2} + \cdots + c + 1)x.
\end{aligned}$$

Just like we saw in Example 11, the last line of the previous derivation includes a sum of a geometric series: $\sum_{i=0}^{n-1} c^i = \frac{c^n-1}{c-1}$. Altogether, we have

$$a_n = c^n a_0 + \left(\frac{c^n - 1}{c - 1}\right) x,$$

and, by collecting like terms and rearranging, we get the general closed form.

Theorem 12. *Let* $a_n = ca_{n-1} + x$ *be a recurrence relation where* $c \neq 1$*. Then the sequence* $A_n = \{a_1, a_2, \ldots, a_n\}$ *is a solution of the recurrence relation if and only if*

$$a_n = \left(a_0 + \frac{x}{c-1}\right)c^n - \frac{x}{c-1}$$

for all $n \in \mathbb{N}$*.*

Proof. By the iteration method. □

As Theorem 12 states, we cannot use this closed form if $c = 1$. If we did, then we would be committing the cardinal sin of division by zero. Fortunately, another application of the iteration method to the similar recurrence relation $a_n = a_{n-1} + x$ gives us a result that works when $c = 1$.

Corollary 13. *Let* $a_n = a_{n-1} + x$ *be a recurrence relation. Then the sequence* $A_n = \{a_1, a_2, \ldots, a_n\}$ *is a solution of the recurrence relation if and only if*

$$a_n = a_0 + nx$$

for all $n \in \mathbb{N}$*.*

Proof. By the iteration method. □

3.3 Characteristic Root Method

So far, the methods we have seen for solving recurrence relations have been very general and very brute-force. With the substitution method, we resort to guessing and throwing induction at the problem. With the iteration method, we replace terms and simplify until something falls out of the expression that looks good. Although these methods work for almost any recurrence relation we need to solve, they simply aren't elegant to use. It would be nice to have a more refined method for solving recurrence relations.

Here, we look at the first of two methods for solving a specific type of recurrence relation. The **characteristic root method** is designed to solve linear homogeneous recurrence relations of degree k with constant coefficients. (Remember all of the terminology from earlier?) In mathematical terms, the characteristic root method solves recurrence relations of the form

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

where each of the coefficients $c_1, c_2, \ldots, c_k$ are real numbers and $c_k \neq 0$.

Why do we care about these specific recurrence relations? As it happens, recurrence relations of this form are abundant in all aspects of mathematics and computer science, in part because they take the simple form of a sum of terms each multiplied by some constant. We've already seen a few examples of recurrence relations of this form; consider F_n from Example 3 or a_n from Example 5. Of course, not all recurrence relations are of this form; for instance, the recurrence relation for the pegs-and-disks problem, H_n, is not homogeneous because of its additive constant term. (This means we can't use H_n as our go-to example for this method.) Fortunately, we have seen other methods for solving recurrence relations not of this form.

Where does the name "characteristic root method" come from? Characteristic roots are the values we use to solve the recurrence relation. From centuries of studying recurrence relations, mathematicians know that linear recurrence relations have exponential solutions; that is, solutions of the form $a_n = r^n$ for some constant r. We won't launch into any discussion about how mathematicians know this fact, since it's outside the scope of these notes. However, we will observe that the equation r^n is a solution of the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$ if and only if

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

where we simply substitute all occurrences of a_i in the recurrence relation with r^i for $(n-k) \leq i \leq n$. If we divide both sides of this expression by r^{n-k} to get rid of the highest-order term on the right-hand side, then move all of the terms to one side, we end up with the expression

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

We call such an expression the **characteristic equation** of the recurrence relation a_n, and we call the solutions of this equation the **characteristic roots** of a_n.

All of the preceding observations combined lead to the statement of the characteristic root method.

Theorem 14 (Characteristic root method)**.** *Let $c_1, c_2, \ldots, c_k$ be real numbers where $c_k \neq 0$. Suppose that the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$ has a corresponding characteristic equation*

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

with k distinct roots $r_1, r_2, \ldots, r_k$. Then the sequence $A_n = \{a_1, a_2, \ldots, a_n\}$ is a solution of the recurrence relation if and only if

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n + \cdots + \alpha_k r_k^n$$

for all $n \in \mathbb{N}$, where each of the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_k$ are constant.

Proof. Omitted. □

3.3.1 Second-Order Recurrence Relations with Two Characteristic Roots

At this point, you're likely thinking that Theorem 14 is a lot to take in all at once. You would indeed be correct; the theorem, as we stated it, is the general form of the characteristic root method. We would use this form of the method if we were dealing with kth-order recurrence relations.

Fortunately, the recurrence relations we deal with most frequently are second-order recurrence relations, and we can simplify the statement of the characteristic root method to apply directly to recurrence relations of degree 2.

Corollary 15. *Let c_1 and c_2 be real numbers where $c_2 \neq 0$. Suppose that the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ has a corresponding characteristic equation*

$$r^2 - c_1 r - c_2 = 0$$

with two distinct roots r_1 and r_2. Then the sequence $A_n = \{a_1, a_2, \ldots, a_n\}$ is a solution of the recurrence relation if and only if

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$$

for all $n \in \mathbb{N}$, where α_1 and α_2 are constant.

Proof. Follows from Theorem 14 by setting $k = 2$. □

When we focus only on second-order recurrence relations, the method becomes much more manageable! Let's consider an example of the method in action with our favourite second-order recurrence relation, F_n.

Example 16. Recall that $F_n = F_{n-1} + F_{n-2}$, where $F_1 = F_2 = 1$. What is a closed form for F_n?

From the statement of the recurrence relation for F_n, we see that $c_1 = 1$ and $c_2 = 1$, so the characteristic equation for F_n is

$$r^2 - r - 1 = 0.$$

The two distinct roots of this equation are $r_1 = (1+\sqrt{5})/2$ and $r_2 = (1-\sqrt{5})/2$. Therefore, the recurrence relation for F_n has a solution of the form

$$a_n = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + \alpha_2 \left(\frac{1-\sqrt{5}}{2}\right)^n$$

for all $n \in \mathbb{N}$, where α_1 and α_2 are constant.

To find the values of α_1 and α_2, we can use the initial terms of F_n. We have that

$$F_1 = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right)^1 + \alpha_2 \left(\frac{1-\sqrt{5}}{2}\right)^1 = 1 \text{ and}$$
$$F_2 = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right)^2 + \alpha_2 \left(\frac{1-\sqrt{5}}{2}\right)^2 = 1,$$

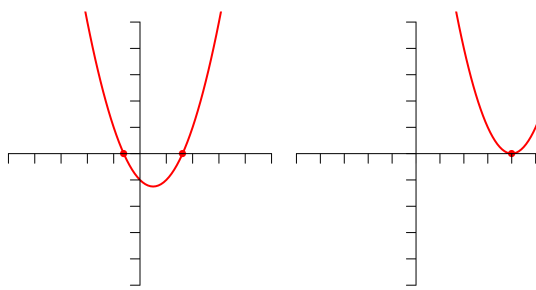
and solving these two expressions gives the values $\alpha_1 = 1/\sqrt{5}$ and $\alpha_2 = -1/\sqrt{5}$.

Therefore, the closed form for F_n is

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

3.3.2 Second-Order Recurrence Relations with One Characteristic Root

Corollary 15 works very well when our characteristic equation has two distinct roots. However, since the characteristic equation is quadratic, we don't have any guarantee that the equation will always have two distinct roots. We may instead have two non-distinct roots, say, when the curve corresponding to the characteristic equation intersects the x-axis at exactly one point. Consider, for example, the following plots:



The plot on the left corresponds to the equation $r^2 - r - 1$, which we saw in Example 16. The plot on the right corresponds to an equation that intersects the x-axis only once; namely, the equation $r^2 - 8r + 16$. In this case, we are still able to use the characteristic root method, but we must make one small modification to its formulation.

Corollary 17. *Let c_1 and c_2 be real numbers where $c_2 \neq 0$. Suppose that the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ has a corresponding characteristic equation*

$$r^2 - c_1 r - c_2 = 0$$

with exactly one root r_1. Then the sequence $A_n = \{a_1, a_2, \ldots, a_n\}$ is a solution of the recurrence relation if and only if

$$a_n = \alpha_1 \, r_1^n + \alpha_2 \, n \, r_1^n$$

for all $n \in \mathbb{N}$, where α_1 and α_2 are constant.

Proof. Omitted. □

The change we made is small, but you can spot it if you look closely at the expression for a_n. To make the characteristic root method work for characteristic equations with one root, we simply multiply the second term (that is, the term including α_2) by a factor of n.

Example 18. Consider the recurrence relation $g_n = 8g_{n-1} - 16g_{n-2}$ with initial terms $g_1 = 1$ and $g_2 = 6$. What is a closed form for g_n?

From the statement of the recurrence relation for g_n, we see that $c_1 = 8$ and $c_2 = -16$, so the characteristic equation for g_n is

$$r^2 - 8r + 16 = 0.$$

The only root of this equation is $r_1 = 4$. Therefore, the recurrence relation for g_n has a solution of the form

$$a_n = \alpha_1\, 4^n + \alpha_2\, n\, 4^n$$

for all $n \in \mathbb{N}$, where α_1 and α_2 are constant.

To find the values of α_1 and α_2, we can use the initial terms of g_n. We have that

$$g_1 = \alpha_1\, 4^1 + \alpha_2\,(1)\,(4^1) = 4\,\alpha_1 + 4\,\alpha_2 = 1, \text{ and}$$
$$g_2 = \alpha_1\, 4^2 + \alpha_2\,(2)\,(4^2) = 16\,\alpha_1 + 32\,\alpha_2 = 6,$$

and solving these two expressions gives the values $\alpha_1 = 1/8$ and $\alpha_2 = 1/8$.

Therefore, the closed form for g_n is

$$g_n = \frac{1}{8}\, 4^n + \frac{1}{8}\, n\, 4^n.$$

3.4 Method of Undetermined Coefficients

Recall that our previous method, the characteristic root method, only worked for recurrence relations of a specific type: the so-called linear homogeneous recurrence relations of degree k with constant coefficients. Although we said earlier that recurrence relations of this type are common, we may occasionally encounter recurrence relations that don't meet all of these specific criteria. In fact, we've already encountered one such recurrence relation: the pegs-and-disks recurrence relation, H_n, which was not homogeneous.

Here, we look at our second method for solving a specific type of recurrence relation. This method is designed to solve linear nonhomogeneous recurrence relations of degree k with constant coefficients. (Note the difference: we are now considering *nonhomogeneous* recurrence relations. Every other property remains the same.) In mathematical terms, recurrence relations of this type are of the form

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

where each of the coefficients $c_1, c_2, \ldots, c_k$ are real numbers with $c_k \neq 0$ and $f(n)$ is a nonzero function dependent on the variable n. In other words, the only difference between the recurrence relations of this section and the recurrence relations of the previous section is the addition of a function $f(n)$ that doesn't incorporate past terms of the recurrence relation.

Before we continue, a little terminology: if we remove the nonhomogeneous term $f(n)$ from the recurrence relation a_n given previously, then we end up with a homogeneous variant of that recurrence relation. We call this recurrence relation (namely, $a_n^{(h)} = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$) the **associated homogeneous recurrence relation**. As a concrete example, consider our familiar recurrence relation $H_n = 2H_{n-1} + 1$. The associated homogeneous recurrence relation for H_n is $H_n^{(h)} = 2H_{n-1}$; we just drop the nonhomogeneous term $f(n) = 1$ from the recurrence relation.

With this notion, we can state the method for solving linear nonhomogeneous recurrence relations of degree k with constant coefficients.

Theorem 19. *Let $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + f(n)$ be a recurrence relation where $c_1, c_2, \ldots, c_k$ are real numbers with $c_k \neq 0$ and $f(n)$ is a nonzero function. If $A^{(p)} = \{a_1^{(p)}, a_2^{(p)}, \ldots, a_n^{(p)}\}$ is a particular solution to a_n, and if $A^{(h)} = \{a_1^{(h)}, a_2^{(h)}, \ldots, a_n^{(h)}\}$ is the homogeneous solution to the associated homogeneous recurrence relation, then $a_n^{(p)} + a_n^{(h)}$ is also a solution to a_n.*

Proof. Omitted. □

Theorem 19 tells us that if we know one particular solution, $A^{(p)}$, to the given recurrence relation as well as the homogeneous solution, $A^{(h)}$, to the associated homogeneous recurrence relation, then all solutions to the given recurrence relation have the form of a termwise sum of the solutions $A^{(p)}$ and $A^{(h)}$.

From the characteristic root method, we know how to find the homogeneous solution, $A^{(h)}$, for an associated homogeneous recurrence relation. But how do we find a "particular solution" to the given recurrence relation? For this, we require the namesake of this section: the **method of undetermined coefficients**.

Theorem 20 (Method of undetermined coefficients)**.** *Let $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + f(n)$ be a recurrence relation where $c_1, c_2, \ldots, c_k$ are real numbers with $c_k \neq 0$ and $f(n)$ is a nonzero function. Suppose that the nonhomogeneous term $f(n)$ is of the form*

$$f(n) = (\beta_r n^r + \beta_{r-1} n^{r-1} + \cdots + \beta_1 n + \beta_0)\, s^n,$$

where $\beta_1, \beta_2, \ldots, \beta_r$ are real numbers and s is a constant.

- *If s is a characteristic root of the associated homogeneous recurrence relation, and s has multiplicity m, then a_n has a particular solution of the form*

$$n^m (\gamma_r n^r + \gamma_{r-1} n^{r-1} + \cdots + \gamma_1 n + \gamma_0)\, s^n.$$

- *If s is a not a characteristic root of the associated homogeneous recurrence relation, then a_n has a particular solution of the form*

$$(\gamma_r n^r + \gamma_{r-1} n^{r-1} + \cdots + \gamma_1 n + \gamma_0)\, s^n.$$

Proof. Omitted. □

With the method of undetermined coefficients, we are able to reduce the problem of finding a particular solution for our recurrence relation to the problem of determining characteristic roots of the associated homogeneous recurrence relation, and we know how to do that via the characteristic root method. Therefore, if you're intimidated by all of the notation in Theorem 20, don't be; you've seen almost everything before!

Note that we broke down Theorem 20 into two cases: one when s is a characteristic root, and one when s is not a characteristic root. In the former case, the particular solution has an extra factor of n^m in order to ensure that the particular solution is not the same as the general solution for the associated homogeneous recurrence relation; the situation here is analogous to the characteristic root method, when the characteristic equation had only one root and we had to introduce an extra factor of n.

Remark. Unfortunately, there is no general method for solving recurrence relations where $f(n)$ is an arbitrary nonzero function. The method of undetermined coefficients only works when $f(n)$ is the product of a polynomial function and an exponential function. If $f(n)$ is not of this form, then we simply need to "guess and check" to obtain a particular solution.

Now, why is this method called the "method of undetermined coefficients"? Recall that, when we solved recurrence relations using the characteristic root method, we ended up with a solution for a_n containing coefficients α_i for $1 \leq i \leq k$. When we use the method of undetermined coefficients, we similarly end up with a particular solution for a_n containing coefficients γ_i for $0 \leq i \leq r$. However, we don't know the values of these coefficients; we only determine the values when we find the overall solution for a_n.

Following the steps laid out previously, let's walk through a few examples of using the method of undetermined coefficients.

Example 21. Let $g'_n = 8g'_{n-1} - 16g'_{n-2} + n \cdot 4^n$. What is a particular solution for this recurrence relation?

Observe that the characteristic equation of the associated homogeneous recurrence relation ${g'_n}^{(h)} = 8g'_{n-1} - 16g'_{n-2}$ is $r^2 - 8r + 16$, and the characteristic root of this equation is 4. Furthermore, this characteristic root has multiplicity 2. Since $s = 4$ is a root of the characteristic equation, by the method of undetermined coefficients, a particular solution for g'_n will be of the form $n^2\,(\gamma_1 n + \gamma_0)\,4^n$.

Example 22. Let $h_n = 4h_{n-1} + 12h_{n-2} + (n^2+1)\cdot 3^n$. What is a particular solution for this recurrence relation?

Observe that the characteristic equation of the associated homogeneous recurrence relation $h_n^{(h)} = 4h_{n-1} + 12h_{n-2}$ is $r^2 - 4r - 12$, and the characteristic roots of this equation are -2 and 6. Since $s = 3$ is not a root of the characteristic equation, by the method of undetermined coefficients, a particular solution for h_n will be of the form $(\gamma_2 n^2 + \gamma_1 n + \gamma_0)\,3^n$.

From here, we conclude by working through a full example of solving a linear nonhomogeneous recurrence relation of degree 2 with constant coefficients.

Example 23. Let $h'_n = 4h'_{n-1} + 12h'_{n-2} + 2^n$. What is a solution for this recurrence relation?

Begin by finding the homogeneous solution for the associated homogeneous recurrence relation ${h'_n}^{(h)} = 4h'_{n-1} + 12h'_{n-2}$. The characteristic equation for ${h'_n}^{(h)}$ is $r^2 - 4r - 12$, and the characteristic roots of this equation are -2 and 6. Therefore, the homogeneous solution is

$${h'_n}^{(h)} = \alpha_1\,(-2)^n + \alpha_2\,6^n,$$

where α_1 and α_2 are constant.

Now, find a particular solution ${h'_n}^{(p)}$. Since $s = 2$ is not a root of the characteristic equation we obtained in the previous step, by the method of undetermined coefficients, a particular solution for h'_n will be of the form

$${h'_n}^{(p)} = \gamma_0\,2^n,$$

where γ_0 is constant.

Substituting the particular solution into the given recurrence relation gives us the following:

$$\begin{aligned}
\gamma_0\,2^n &= 4\,\gamma_0\,2^{n-1} + 12\,\gamma_0\,2^{n-2} + 2^n \\
4\,\gamma_0 &= 8\,\gamma_0 + 12\,\gamma_0 + 4 \quad \text{(via division by } 2^{n-2}\text{)} \\
24\,\gamma_0 &= 4 \\
\gamma_0 &= 4/24.
\end{aligned}$$

Therefore, ${h'_n}^{(p)} = (4/24)\,2^n$ is a particular solution for h'_n, and by Theorem 19, all solutions of h'_n are of the form

$$h'_n = (4/24)\,2^n + \alpha_1\,(-2)^n + \alpha_2\,6^n.$$