

## Binary Search

**Theorem:** Binary search takes at most  $\lfloor \log_2(n) \rfloor + 1$  loop iterations on a list of  $n$  items.

**Proof:** Let  $P(n)$  be the statement that if  $L - F = n \geq 0$ , then we go through the loop at most  $\lfloor \log_2(L + 1 - F) \rfloor + 1$  times.

*Basis:* If  $L - F = 0$ , then we go through the loop at most once (0 times if the  $w = w_i$  is actually on the list), and  $\log_2(1) + 1 = 1$ .

*Inductive step:* Assume  $P(0), \dots, P(n)$ . If  $L - F = n + 1$ , then either  $w = w_{\lfloor (F+L)/2 \rfloor}$  (in which case we quit), or (a)  $w < w_{\lfloor (F+L)/2 \rfloor}$  or (b)  $w > w_{\lfloor (F+L)/2 \rfloor}$ . Let  $L', F'$  be values of  $L$  and  $F$  on the next iteration.

In case (a),  $L' = \lfloor (F + L)/2 \rfloor - 1$ ,  $F' = F$ , so

$$L' + 1 - F' = \lfloor (F + L)/2 \rfloor - F = \lfloor (L - F)/2 \rfloor$$

In case (b)  $F' = \lfloor (F + L)/2 \rfloor + 1$ ,  $L' = L$ , so

$$L' + 1 - F' = L - \lfloor (F + L)/2 \rfloor = \lceil (L - F)/2 \rceil$$

Either way, by strong induction, it takes at most

$$1 + \lfloor \log_2(\lceil (L - F)/2 \rceil) \rfloor + 1$$

times through the loop. (One more than it takes starting at  $(L', F')$ ).

Two facts about the floor function:

- $\lfloor x/2 \rfloor \leq \frac{x}{2} + \frac{1}{2}$  if  $x$  is an integer
- $1 + \lfloor x \rfloor = \lfloor 1 + x \rfloor$  for all  $x \in \mathbb{R}$

Therefore:

$$\begin{aligned} & 1 + \lfloor \log_2(\lceil (L - F)/2 \rceil) \rfloor + 1 \\ & \leq 1 + \lfloor \log_2((L + 1 - F)/2) \rfloor + 1 \\ & = \lfloor 1 + \log_2(L + 1 - F)/2 \rfloor + 1 \\ & = \lfloor \log_2(2) + \log_2(L + 1 - F)/2 \rfloor + 1 \\ & = \lfloor \log_2(L + 1 - F) \rfloor + 1 \end{aligned}$$

This is what we wanted to prove!

1

2

## Bubble Sort

Suppose we wanted to sort  $n$  items. Here's one way to do it:

**Input**  $n$  [number of items to be sorted]  
 $w_1, \dots, w_n$  [items]

**Algorithm BubbleSort**  
**for**  $i = 1$  **to**  $n - 1$   
    **for**  $j = 1$  **to**  $n - i$   
        **if**  $w_j > w_{j+1}$  **then**  $\text{switch}(w_j, w_{j+1})$  **endif**  
    **endfor**  
**endfor**

Why is this right:

- Intuitively, because highest elements "bubble up" to the top

How many comparisons?

- Best case, worst case, average case all the same:  
 $\circ (n - 1) + (n - 2) + \dots + 1 = n(n - 1)/2$

3

## Proving Bubble Sort Correct

We want to show that the algorithm is correct by induction. What's the statement of the induction?

$P(k)$  is the statement that after  $k$  iterations of the outer loop,  $w_{n-k+1}, \dots, w_n$  are the  $k$  highest items, sorted in the right order.

*Basis:* How do we prove  $P(1)$ ? By a nested induction!

This time, take  $Q(l)$  to be the statement that after  $l$  iterations of the inner loop,  $w_{l+1}$  is higher than  $\{w_1, \dots, w_l\}$ .

*Basis:*  $Q(1)$  holds because after the first iteration of the inner loop,  $w_2 > w_1$  (thanks to the switch statement).

*Inductive step:* After  $l$  iterations, the algorithm guarantees that  $w_{l+1} > w_l$ . Using the induction hypothesis,  $w_{l+1}$  is also higher than  $\{w_1, \dots, w_{l-1}\}$ .

$Q(n - 1)$  implies  $P(1)$ , so we're done with the base case of the main induction.

[**Note:** For a really careful proof, we need better notation (for value of  $w_l$  before and after the switch).]

4

*Inductive step (for main induction):* Assume  $P(k)$ .  
 By the subinduction, after  $n - k - 1$  iterations of the inner loop,  $w_{n-k}$  is alphabetically after  $\{w_1, \dots, w_{n-(k+1)}\}$ .  
 Combined with  $P(k)$ , this tells us  $w_{n-k}, \dots, w_n$  are the  $k + 1$  highest elements. This proves  $P(k + 1)$ .

## How to Guess What to Prove

Sometimes formulating  $P(n)$  is straightforward; sometimes it's not. This is what to do:

- Compute the result in some specific cases
- Conjecture a generalization based on these cases
- Prove the correctness of your conjecture (by induction)

## Example

Suppose  $a_1 = 1$  and  $a_n = a_{\lceil n/2 \rceil} + a_{\lfloor n/2 \rfloor}$  for  $n > 1$ .  
 Find an explicit formula for  $a_n$ .

Try to see the pattern:

- $a_1 = 1$
- $a_2 = a_1 + a_1 = 1 + 1 = 2$
- $a_3 = a_2 + a_1 = 2 + 1 = 3$
- $a_4 = a_2 + a_2 = 2 + 2 = 4$

Suppose we modify the example. Now  $a_1 = 3$  and  $a_n = a_{\lceil n/2 \rceil} + a_{\lfloor n/2 \rfloor}$  for  $n > 1$ . What's the pattern?

- $a_1 = 3$
- $a_2 = a_1 + a_1 = 3 + 3 = 6$
- $a_3 = a_2 + a_1 = 6 + 3 = 9$
- $a_4 = a_2 + a_2 = 6 + 6 = 12$

$$a_n = 3n!$$

**Theorem:** If  $a_1 = k$  and  $a_n = a_{\lceil n/2 \rceil} + a_{\lfloor n/2 \rfloor}$  for  $n > 1$ , then  $a_n = kn$  for  $n \geq 1$ .

**Proof:** By strong induction. Let  $P(n)$  be the statement that  $a_n = kn$ .

*Basis:*  $P(1)$  says that  $a_1 = k$ , which is true by hypothesis.

*Inductive step:* Assume  $P(1), \dots, P(n)$ ; prove  $P(n + 1)$ .

$$\begin{aligned} a_{n+1} &= a_{\lceil (n+1)/2 \rceil} + a_{\lfloor (n+1)/2 \rfloor} \\ &= k\lceil (n+1)/2 \rceil + k\lfloor (n+1)/2 \rfloor \text{ [Induction hypothesis]} \\ &= k(\lceil (n+1)/2 \rceil + \lfloor (n+1)/2 \rfloor) \\ &= k(n+1) \end{aligned}$$

We used the fact that  $\lceil n/2 \rceil + \lfloor n/2 \rfloor = n$  for all  $n$  (in particular, for  $n + 1$ ). To see this, consider two cases:  $n$  is odd and  $n$  is even.

- if  $n$  is even,  $\lceil n/2 \rceil + \lfloor n/2 \rfloor = n/2 + n/2 = n$
- if  $n$  is odd, suppose  $n = 2k + 1$ 
  - $\lceil n/2 \rceil + \lfloor n/2 \rfloor = (k + 1) + k = 2k + 1 = n$

This proof has a (small) gap:

- We should check that  $\lceil (n + 1)/2 \rceil \leq n$

In general, there is no rule for guessing the right inductive hypothesis. However, if you have a sequence of numbers

$$r_1, r_2, r_3, \dots$$

and want to guess a general expression, here are some guidelines for trying to find the *type* of the expression (exponential, polynomial):

- Compute  $\lim_{n \rightarrow \infty} r_{n+1}/r_n$ 
  - if it looks like  $\lim_{n \rightarrow \infty} r_{n+1}/r_n = b \notin \{0, 1\}$ , then  $r_n$  probably has the form  $Ab^n + \dots$ .
  - You can compute  $A$  by computing  $\lim_{n \rightarrow \infty} r_n/b^n$
  - Try to compute the form of  $\dots$  by considering the sequence  $r_n - Ab^n$ ; that is,

$$r_1 - Ab, r_2 - Ab^2, r_3 - Ab^3, \dots$$

- $\lim_{n \rightarrow \infty} r_{n+1}/r_n = 1$ , then  $r_n$  is most likely a polynomial.
- $\lim_{n \rightarrow \infty} r_{n+1}/r_n = 0$ , then  $r_n$  may have the form  $A/b^{f(n)}$ , where  $f(n)/n \rightarrow \infty$ 
  - $f(n)$  could be  $n \log n$  or  $n^2$ , for example

Once you have guessed the form of  $r_n$ , prove that your guess is right by induction.

9

## One more example

Find a formula for

$$\frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \frac{1}{7 \cdot 10} + \dots + \frac{1}{(3n-2)(3n+1)}$$

Some values:

- $r_1 = 1/4$
- $r_2 = 1/4 + 1/28 = 8/28 = 2/7$
- $r_3 = 1/4 + 1/28 + 1/70 = (70 + 10 + 4)/280 = 84/280 = 3/10$

Conjecture:  $r_n = n/(3n+1)$ . Let this be  $P(n)$ .

*Basis:*  $P(1)$  says that  $r_1 = 1/4$ .

*Inductive step:*

$$\begin{aligned} r_{n+1} &= r_n + \frac{1}{(3n+1)(3n+4)} \\ &= \frac{n}{3n+1} + \frac{1}{(3n+1)(3n+4)} \\ &= \frac{n(3n+4)+1}{(3n+1)(3n+4)} \\ &= \frac{3n^2+4n+1}{(3n+1)(3n+4)} \\ &= \frac{(n+1)(3n+1)}{(3n+1)(3n+4)} \\ &= \frac{n+1}{3n+4} \end{aligned}$$

11

## More examples

Come up with a simple formula for the sequence

$$1, 5, 13, 41, 121, 365, 1093, 3281, 9841, 29525$$

Compute limit of  $r_{n+1}/r_n$ :

$$5/1 = 5, \quad 13/5 \approx 2.6, \quad 41/13 \approx 3.2, \quad 121/41 \approx 2.95, \dots, 29525/9841 \approx 3.000$$

Guess: limit is 3 ( $\Rightarrow r_n = A3^n + \dots$ )

Compute limit of  $r_n/3^n$ :

$$1/3 \approx .33, \quad 5/9 \approx .56, \quad 13/27 \approx .5, \quad 41/81 \approx .5, \dots, 29525/3^{10} \approx .5000$$

Guess: limit is  $1/2$  ( $\Rightarrow r_n = \frac{1}{2}3^n + \dots$ ) +

Compute  $r_n - 3^n/2$ :

$$(1 - 3/2), (5 - 9/2), (13 - 27/2), (41 - 81/2), \dots = -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \dots$$

Guess: general term is  $3^n/2 + (-1)^n/2$

Verify (by induction ...)

10

## Faulty Inductions

Part of why I want you to write out your assumptions carefully is so that you don't get led into some standard errors.

**Theorem:** All women are blondes.

**Proof by induction:** Let  $P(n)$  be the statement: For any set of  $n$  women, if at least one of them is a blonde, then all of them are.

*Basis:* Clearly OK.

*Inductive step:* Assume  $P(n)$ . Let's prove  $P(n+1)$ .

Given a set  $W$  of  $n+1$  women, one of which is blonde. Let  $A$  and  $B$  be two subsets of  $W$ , each of which contains the known blonde, whose union is  $W$ .

By the induction hypothesis, each of  $A$  and  $B$  consists of all blondes. Thus, so does  $W$ . This proves  $P(n) \Rightarrow P(n+1)$ .

12

Take  $W$  to be the set of women in the world, and let  $n = |W|$ . Since there is clearly at least one blonde in the world, it follows that all women are blonde!

Where's the bug?

**Theorem:** Every integer  $> 1$  has a unique prime factorization.

[The result is true, but the following proof is not:]

**Proof:** By strong induction. Let  $P(n)$  be the statement that  $n$  has a unique factorization, for  $n > 1$ .

*Basis:*  $P(2)$  is clearly true.

*Induction step:* Assume  $P(2), \dots, P(n)$ . We prove  $P(n+1)$ . If  $n+1$  is prime, we are done. If not, it factors somehow. Suppose  $n+1 = rs$ ,  $r, s > 1$ . By the induction hypothesis,  $r$  has a unique factorization  $\prod_i p_i$  and  $s$  has a unique prime factorization  $\prod_j q_j$ . Thus,  $\prod_i p_i \prod_j q_j$  is a prime factorization of  $n+1$ , and since none of the factors of either piece can be changed, it must be unique.

What's the flaw??

13

Problem: Suppose  $n+1 = 36$ . That is, you've proved that every number up to 36 has a unique factorization. Now you need to prove it for 36.

36 isn't prime, but  $36 = 3 \times 12$ . By the induction hypothesis, 12 has a unique prime factorization, say  $p_1 p_2 p_3$ . Thus,  $36 = 3 p_1 p_2 p_3$ .

However, 36 is also  $4 \times 9$ . By the induction hypothesis,  $4 = q_1 q_2$  and  $9 = r_1 r_2$ . Thus,  $36 = q_1 q_2 r_1 r_2$ .

How do you know that  $3 p_1 p_2 p_3 = q_1 q_2 r_1 r_2$ .

(They do, but it doesn't follow from the induction hypothesis.)

This is a *breakdown error*. If you're trying to show something is unique, and you break it down (as we broke down  $n+1$  into  $r$  and  $s$ ) you have to argue that nothing changes if we break it down a different way. What if  $n+1 = tu$ ?

- The actual proof of this result is quite subtle

15

14

**Theorem:** The sum of the internal angles of a regular  $n$ -gon is  $180(n-2)$  for  $n \geq 3$ .

**Proof:** By induction. Let  $P(n)$  be the statement of the theorem. For  $n = 3$ , the result was shown in high school. Assume  $P(n)$ ; let's prove  $P(n+1)$ . Given a regular  $(n+1)$ -gon, we can lop off one of the corners:

By induction, the sum of the internal angles of the  $n$ -gon is  $180(n-2)$  degrees; the sum of the internal angles of the triangle is 180 degrees. Thus, the internal angles of the original  $(n+1)$ -gon is  $180(n-1)$ . What's wrong??

- When you lop off a corner, you don't get a *regular*  $n$ -gon.

The fix: **Strengthen the induction hypothesis.**

- Let  $P(n)$  say that the sum of the internal angles of *any*  $n$ -gon is  $180(n-2)$ .

16

Consider 0-1 sequences in which 1's may not appear consecutively, except in the rightmost two positions.

- 010110 is not allowed, but 010011 is

Prove that there are  $2^n$  allowed sequences of length  $n$  for  $n \geq 1$

Why can't this be right?

**“Proof”** Let  $P(n)$  be the statement of the theorem.

*Basis:* There are 2 sequences of length 1—0 and 1—and they're both allowed.

*Inductive step:* Assume  $P(n)$ . Let's prove  $P(n+1)$ . Take any allowed sequence  $x$  of length  $n$ . We get a sequence of length  $n+1$  by appending either a 0 or 1 at the end. In either case, it's allowed.

- If  $x$  ends with a 1, it's OK, because  $x1$  is allowed to end with 2 1's.

Thus,  $s_{n+1} = 2s_n = 2 \cdot 2^n = 2^{n+1}$ .

Where's the flaw?

- What if  $x$  already ends with 2 1's?

Correct expression involves separating out sequences which end in 0 and 1 (it's done in Chapter 5, but I'm not sure we'll get to it)

17

## Example

**Theorem**  $n$  is odd iff  $n^2$  is odd, for  $n \in \mathbb{N}^+$ .

**Proof:** We have to show

1.  $n$  odd  $\Rightarrow n^2$  odd
2.  $n^2$  odd  $\Rightarrow n$  odd

For (1), if  $n$  is odd, it is of the form  $2k+1$ . Hence,

$$n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

Thus,  $n^2$  is odd.

For (2), we proceed by contradiction. Suppose  $n^2$  is odd and  $n$  is even. Then  $n = 2k$  for some  $k$ , and  $n^2 = 4k^2$ . Thus,  $n^2$  is even. This is a contradiction. Thus,  $n$  must be odd.

19

## Methods of Proof

Typically you're trying to prove a statement like “Given  $X$ , prove (or show that)  $Y$ ”. This means you have to prove

$$X \Rightarrow Y$$

In the proof, you're allowed to assume  $X$ , and then show that  $Y$  is true, using  $X$ .

- A special case: if there is no  $X$ , you just have to prove  $Y$  or *true*  $\Rightarrow Y$ .

Alternatively, you can do a *proof by contradiction*: Assume that  $Y$  is false, and show that  $X$  is false.

- This amounts to proving

$$\neg Y \Rightarrow \neg X$$

18

## A Proof By Contradiction

**Theorem:**  $\sqrt{2}$  is irrational.

**Proof:** By contradiction. Suppose  $\sqrt{2}$  is rational. Then  $\sqrt{2} = a/b$  for some  $a, b \in \mathbb{N}^+$ . We can assume that  $a/b$  is in lowest terms.

- Therefore,  $a$  and  $b$  can't both be even.

Squaring both sides, we get

$$2 = a^2/b^2$$

Thus,  $a^2 = 2b^2$ , so  $a^2$  is even. This means that  $a$  must be even.

Suppose  $a = 2c$ . Then  $a^2 = 4c^2$ .

Thus,  $4c^2 = 2b^2$ , so  $b^2 = 2c^2$ . This means that  $b^2$  is even, and hence so is  $b$ .

Contradiction!

Thus,  $\sqrt{2}$  must be irrational.

20

## A Bad Proof

Prove  $\log(x/y) = \log(x) - \log(y)$

**Proof:**

$$\log(x/y) = \log(x) - \log(y)$$

$$\log(x) + \log(1/y) = \log(x) - \log(y)$$

$$\log(x) + \log(y^{-1}) = \log(x) - \log(y)$$

$$\log(x) - \log(y) = \log(x) - \log(y)$$

What's wrong?

- You need to connect the statements (using  $\Leftrightarrow$ , for example)