



# CSE373: Data Structures and Algorithms

## Lecture 2: Proof by Induction

Catie Baker

Spring 2015

# *Background on Induction*

- Type of mathematical proof
- Typically used to establish a given statement for all natural numbers (e.g. integers  $> 0$ )
- Proof is a sequence of deductive steps
  1. Show the statement is true for the first number.
  2. Show that if the statement is true for any one number, this implies the statement is true for the next number.
  3. If so, we can infer that the statement is true for all numbers.

# *Think about climbing a ladder*



1. Show you can get to the first rung (base case)
2. Show you can get between rungs (inductive step)
3. Now you can climb forever.

# *Why you should care*

- Induction turns out to be a useful technique
  - AVL trees
  - Heaps
  - Graph algorithms
  - Can also prove things like  $3^n > n^3$  for  $n \geq 4$
- Exposure to rigorous thinking

# *Example problem*

- Find the sum of the integers from 1 to  $n$
- $1 + 2 + 3 + 4 + \dots + (n-1) + n$

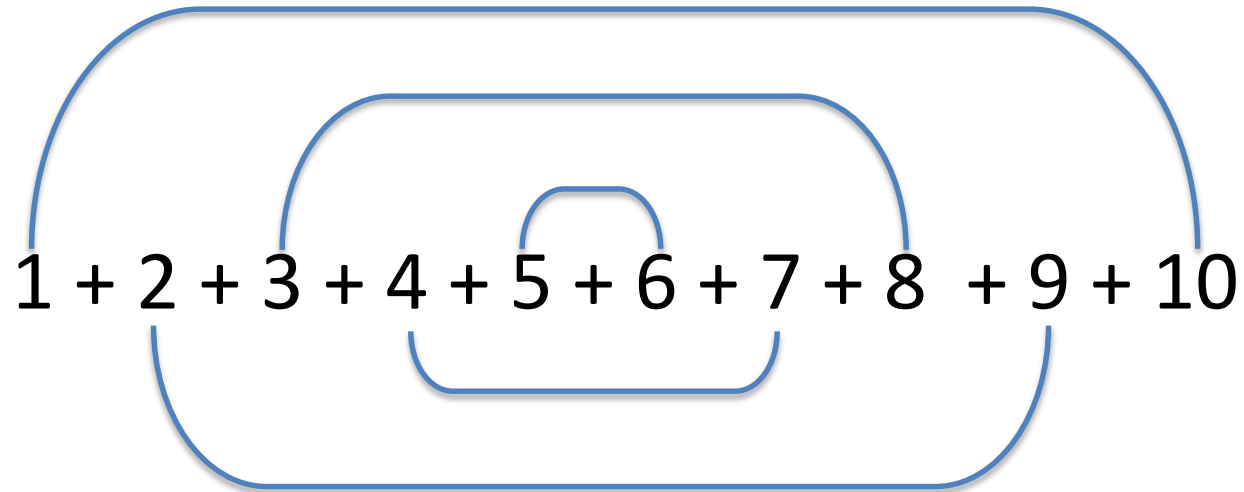
$$\sum_{i=1}^n i$$

- For any  $n \geq 1$
- Could use brute force, but would be slow
- There's probably a clever **shortcut**

# *Finding the formula*

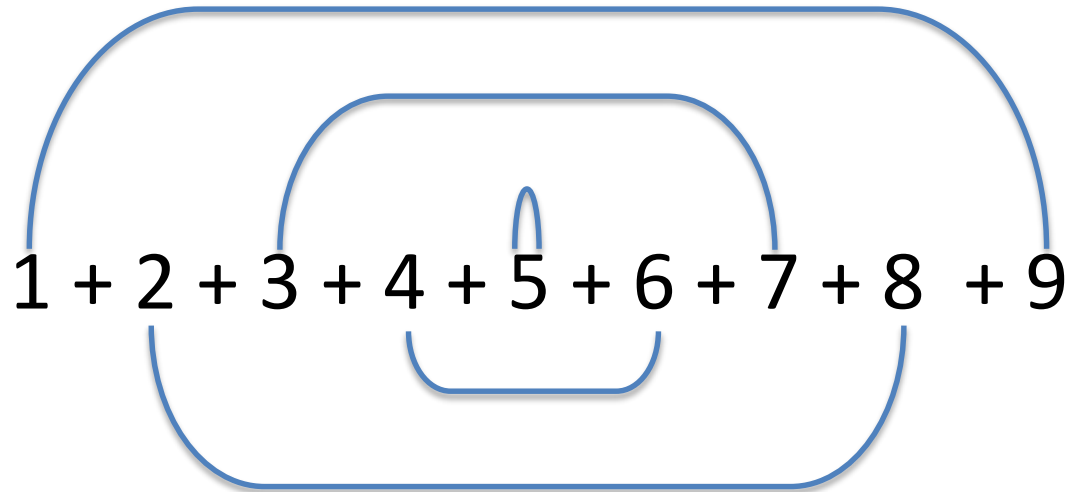
- Shortcut will be some **formula** involving  $n$
- Compare examples and look for patterns
  - Not something I will ask you to do!
- Start with  $n = 10$ :  
 $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10$ 
  - Large enough to be a pain to add up
  - Worthwhile to find shortcut

# *Finding the formula*



$$= 5 \times 11$$

# *Finding the formula*



$$= 4 \times 10 + 5$$



# *Finding the formula*

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8

$$= 4 \times 9$$

# *Finding the formula*

The diagram shows the arithmetic series 1 + 2 + 3 + 4 + 5 + 6 + 7. Three blue brackets are drawn to illustrate a pairing strategy. The first bracket groups the first four terms (1, 2, 3, 4) from below. The second bracket groups the last three terms (4, 5, 6) from above. The third, larger bracket groups all seven terms (1, 2, 3, 4, 5, 6, 7) from below.

$$= 3 \times 8 + 4$$

## *Finding the formula*

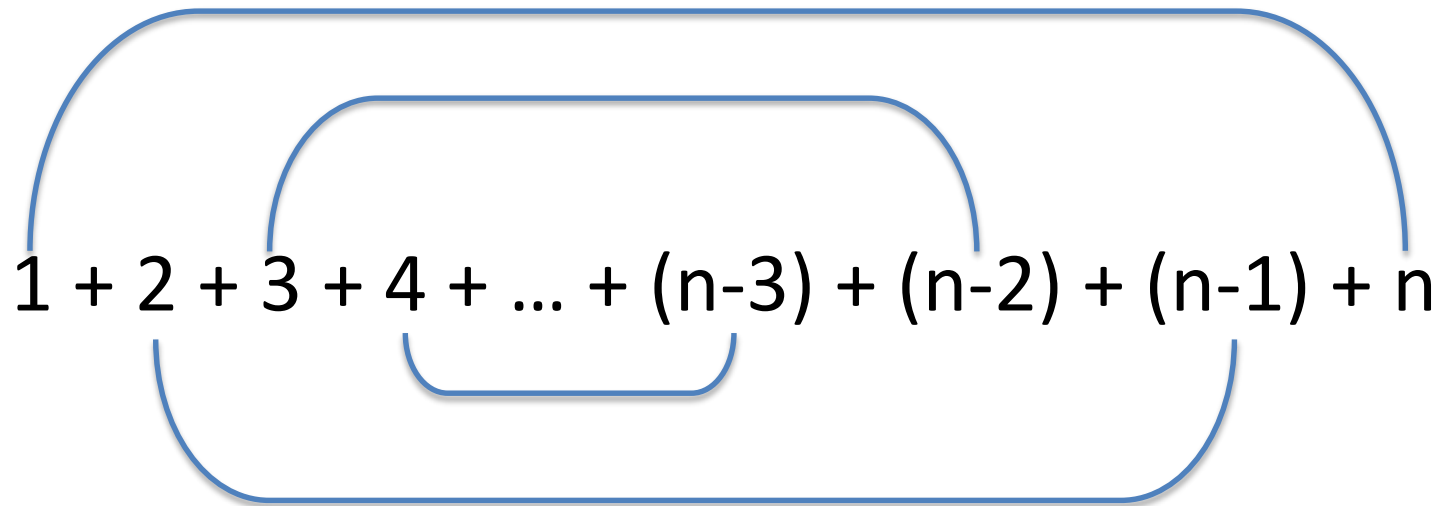
n=7	$3 \times 8 + 4$
n=8	$4 \times 9$
n=9	$4 \times 10 + 5$
n=10	$5 \times 11$

## *Finding the formula*

n=7	$3 \times 8 + 4$	n is odd
n=8	$4 \times 9$	n is even
n=9	$4 \times 10 + 5$	n is odd
n=10	$5 \times 11$	n is even

# *Finding the formula*

When  $n$  is even



The diagram illustrates the summation of an arithmetic series for an even number of terms  $n$ . The series is written as  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Three blue brackets are drawn above the terms to show pairings: the first and last terms (1 and  $n$ ), the second and second-to-last terms (2 and  $n-1$ ), and the third and third-to-last terms (3 and  $n-2$ ). This visualizes that each pair of terms sums to  $n+1$ , and there are  $n/2$  such pairs.

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

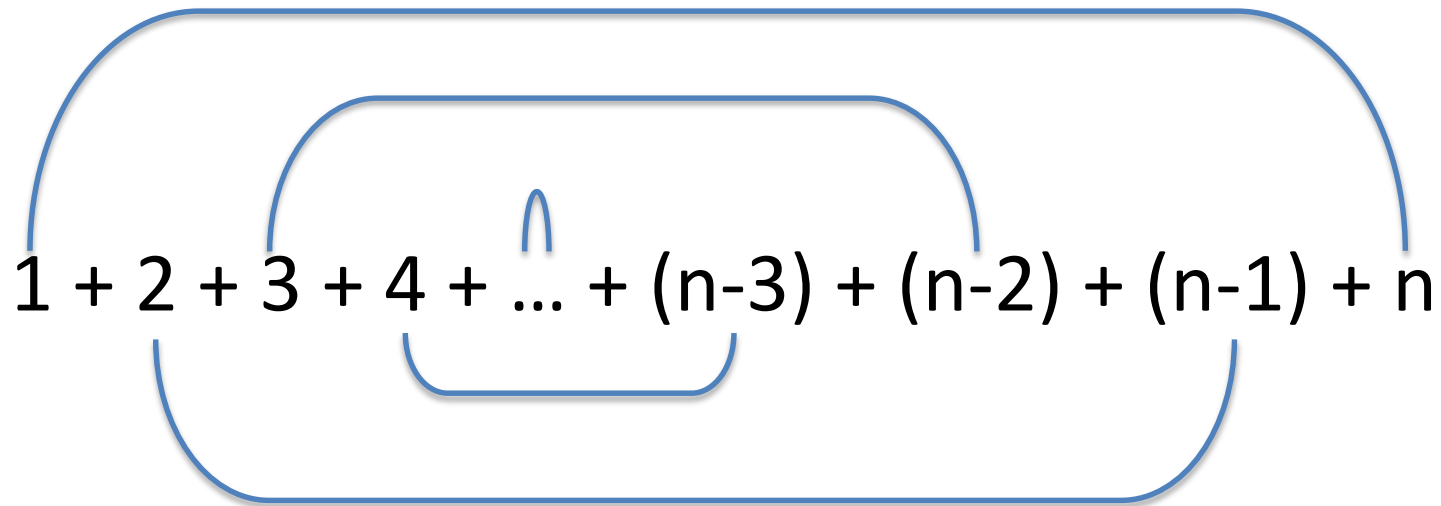
$$= (n/2) \times (n+1)$$

# *Finding the formula*

$3 \times 8 + 4$	
$4 \times 9$	$n(n+1)/2$
$4 \times 10 + 5$	
$5 \times 11$	$n(n+1)/2$

# *Finding the formula*

When n is odd



The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Three blue brackets are drawn to group terms: a large bracket at the top groups the entire series; a smaller bracket below it groups the terms from 3 to  $(n-2)$ ; and a third bracket below that groups the terms from 2 to  $(n-1)$ .

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= ((n-1)/2) \times (n+1) + (n+1)/2$$

# *Finding the formula*

When  $n$  is odd

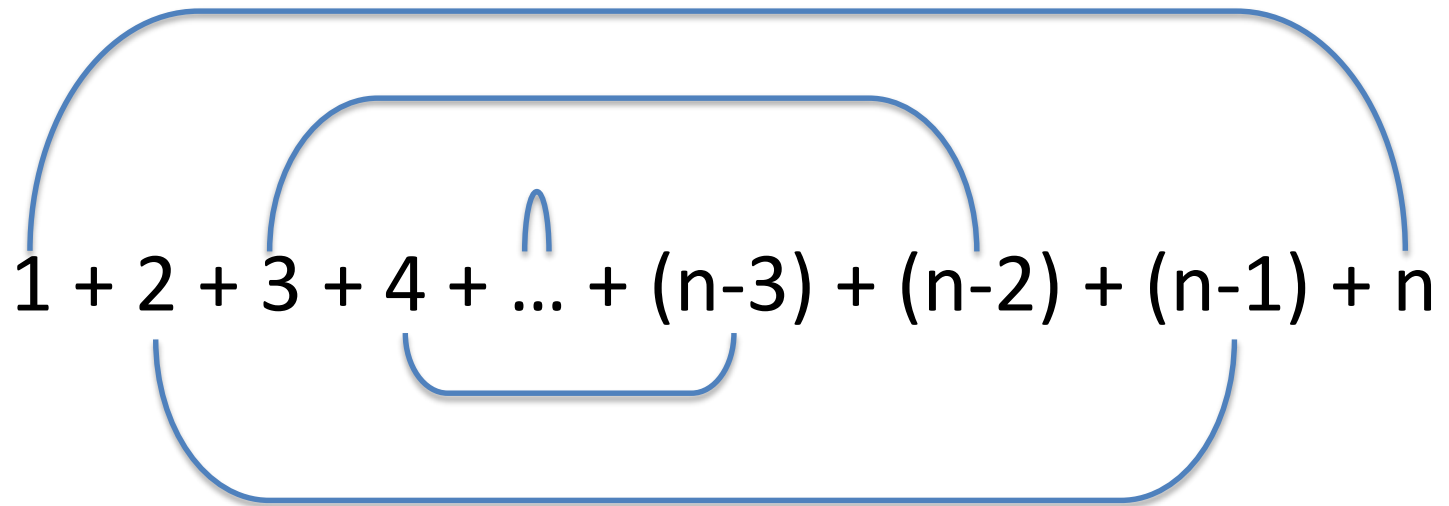
The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Three blue brackets are used to group terms: a large bracket at the top groups the entire series; a smaller bracket below it groups the terms from 3 to  $(n-2)$ ; and a third bracket below that groups the terms from 2 to  $(n-1)$ . The ellipsis  $\dots$  is positioned between 4 and  $(n-3)$ .

$$= ((n-1)/2) \times (n+1) + (n+1)/2$$



# *Finding the formula*

When  $n$  is odd



The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Blue brackets are drawn to group terms: a large bracket groups the entire series; a smaller bracket groups the terms from 3 to  $(n-2)$ ; and a tiny bracket is placed above the ellipsis. Additionally, a bracket is drawn below the terms from 2 to  $(n-1)$ .

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= ((n-1) \times (n+1) + (n+1)) / 2$$

# *Finding the formula*

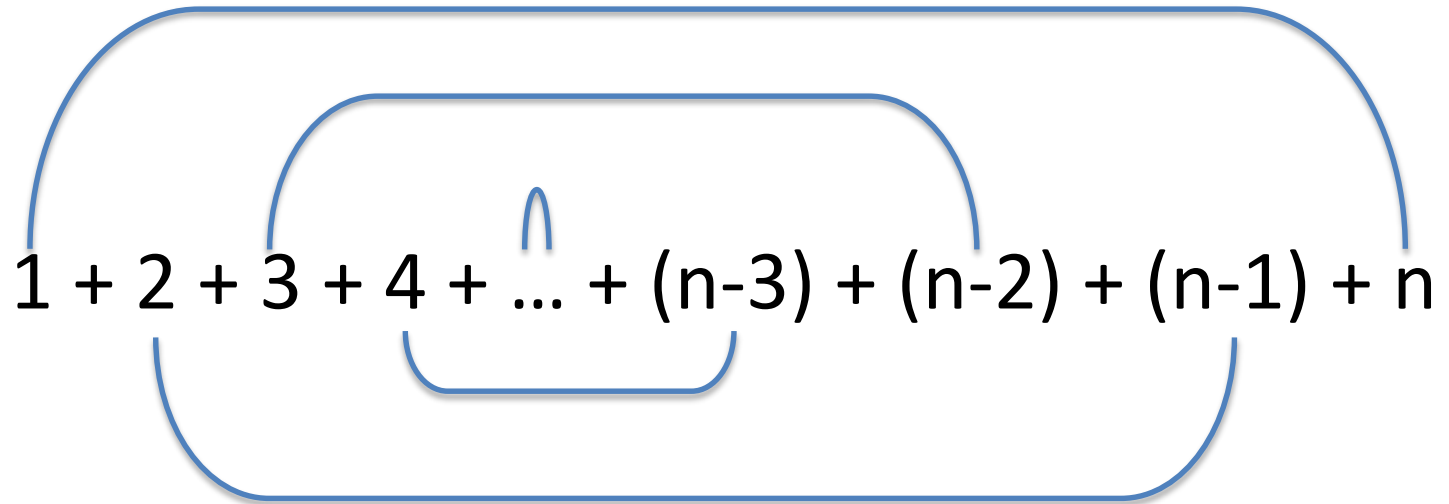
When  $n$  is odd

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= ((n-1) \times (n+1) + (n+1)) / 2$$

# *Finding the formula*

When  $n$  is odd



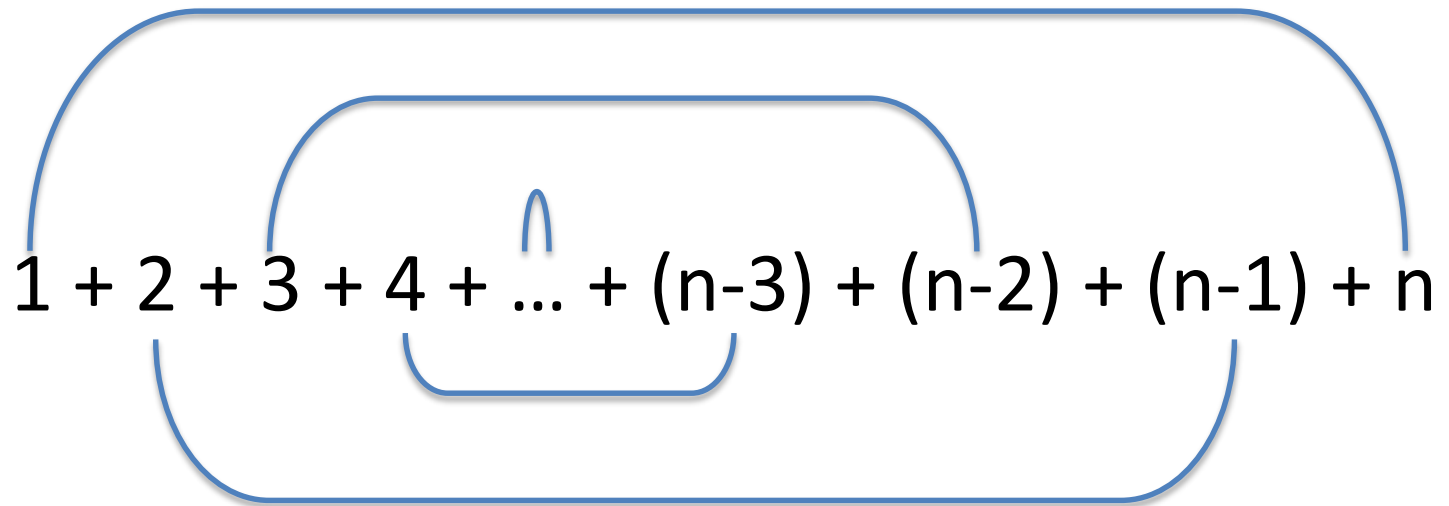
The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Blue brackets are drawn above and below the series to pair terms. The top bracket connects 1 and  $n$ , 2 and  $(n-1)$ , 3 and  $(n-2)$ , and 4 and  $(n-3)$ . The bottom bracket connects  $(n-3)$  and 4,  $(n-2)$  and 3,  $(n-1)$  and 2, and  $n$  and 1. A small blue arch is drawn above the ellipsis  $\dots$ .

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= ((n-1 + 1) \times (n+1)) / 2$$

# *Finding the formula*

When n is odd



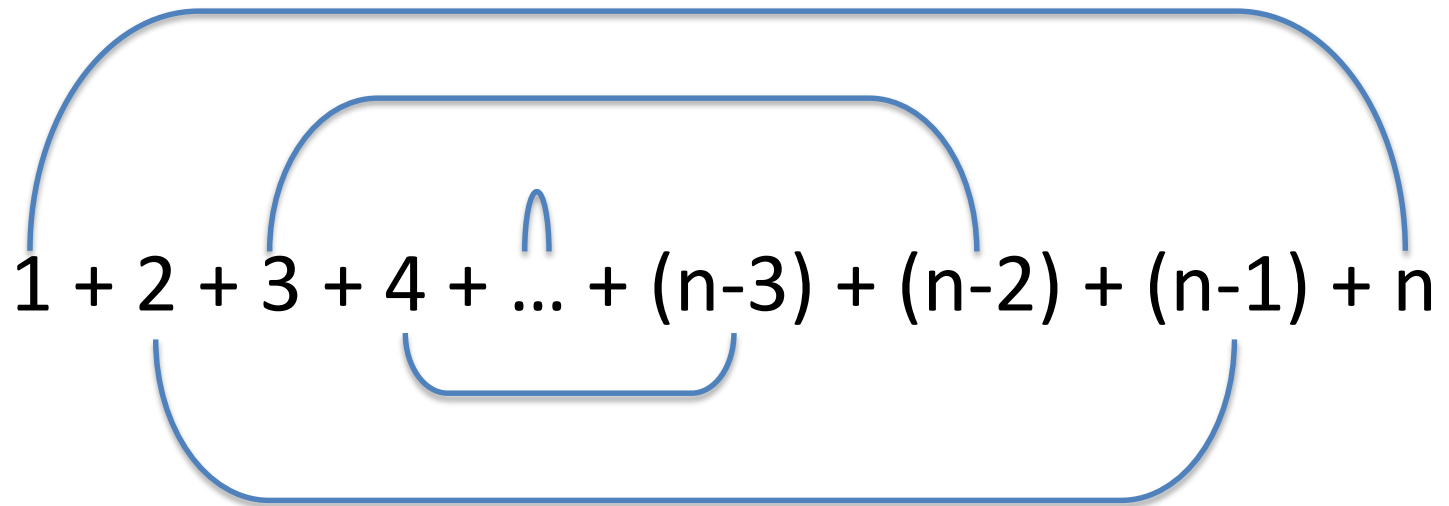
The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Blue brackets are drawn above and below the series to pair terms. The outermost bracket pairs 1 and n. The next bracket pairs 2 and n-1. A smaller bracket pairs 3 and n-2. The innermost bracket pairs 4 and n-3. This illustrates that each pair sums to n+1, and there are (n+1)/2 such pairs.

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= ((n-1 + 1) \times (n+1)) / 2$$

# *Finding the formula*

When  $n$  is odd



The diagram shows the arithmetic series  $1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$ . Blue brackets are drawn above and below the series to pair terms from both ends. The outermost bracket pairs 1 and  $n$ . The next bracket pairs 2 and  $n-1$ . A smaller bracket pairs 3 and  $n-2$ . The innermost bracket pairs 4 and  $n-3$ . This illustrates that each pair sums to  $n+1$ , and there are  $n/2$  such pairs.

$$1 + 2 + 3 + 4 + \dots + (n-3) + (n-2) + (n-1) + n$$

$$= (n(n+1))/2$$

# *Finding the formula*

$3 \times 8 + 4$	$n(n+1)/2$
$4 \times 9$	$n(n+1)/2$
$4 \times 10 + 5$	$n(n+1)/2$
$5 \times 11$	$n(n+1)/2$

# *Are we done?*

- The pattern seems pretty clear
  - Is there any reason to think it changes?
- But we want something for **any**  $n \geq 1$
- A mathematical approach is **skeptical**

$$\frac{n(n + 1)}{2}$$

# *Are we done?*

- The pattern seems pretty clear
  - Is there any reason to think it changes?
- But we want something for *any*  $n \geq 1$
- A mathematical approach is *skeptical*
- All we know is  $n(n+1)/2$  works for 7 to 10
- We must *prove* the formula works in all cases
  - A *rigorous* proof



# *Proof by Induction*

- Prove the formula works for all cases.
- Induction proofs have four components:
  1. The thing you want to prove, e.g., *sum of integers from 1 to  $n = n(n+1)/2$*
  2. The base case (usually "let  $n = 1$ "),
  3. The assumption step ("assume true for  $n = k$ ")
  4. The induction step ("now let  $n = k + 1$ ").

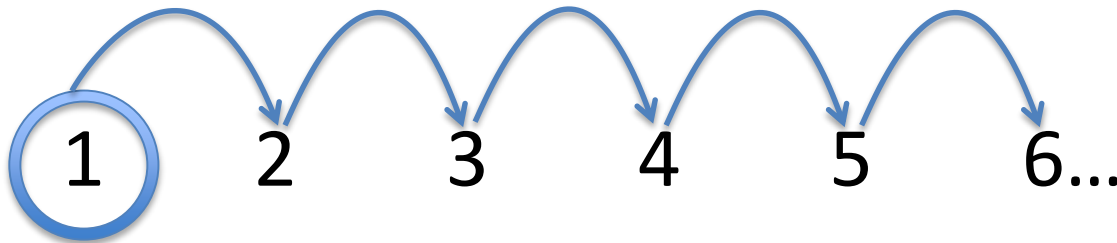
$n$  and  $k$  are just *variables*!

# *Proof by induction*

- $P(n)$  = sum of integers from 1 to  $n$
- We need to do
  - Base case *prove for  $P(1)$*
  - Assumption *assume for  $P(k)$*
  - Induction step *show for  $P(k+1)$*
- $n$  and  $k$  are just *variables!*

# *Proof by induction*

- $P(n)$  = sum of integers from 1 to  $n$
- We need to do
  - Base case *prove for  $P(1)$*
  - Assumption *assume for  $P(k)$*
  - Induction step *show for  $P(k+1)$*



# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Base case
  - $P(1) = 1$
  - $1(1+1)/2 = 1(2)/2 = 1(1) = 1$



# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$   
 $= 1 + 2 + \dots + k + (k+1)$

# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$

# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2$

# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2 = (k(k+1) + 2(k+1))/2$



# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2 = (k(k+1) + 2(k+1))/2$
  - $= (k+1)(k+2)/2$

# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2 = (k(k+1) + 2(k+1))/2$
  - $= (k+1)(k+2)/2$

# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2 = (k(k+1) + 2(k+1))/2$
  - $= (k+1)(k+2)/2 = (k+1)((k+1)+1)/2$



# *Proof by induction*

- What we are trying to prove:  $P(n) = n(n+1)/2$
- Assume true for  $k$ :  $P(k) = k(k+1)/2$
- Induction step:
  - Now consider  $P(k+1)$
  - $= 1 + 2 + \dots + k + (k+1)$
  - $= k(k+1)/2 + (k+1)$
  - $= k(k+1)/2 + 2(k+1)/2 = (k(k+1) + 2(k+1))/2$
  - $= (k+1)(k+2)/2 = (k+1)((k+1)+1)/2$



# *We're done!*

- $P(n)$  = sum of integers from 1 to  $n$
- We have shown
  - Base case *proved for  $P(1)$*
  - Assumption *assumed for  $P(k)$*
  - Induction step *proved for  $P(k+1)$*

Success: we have proved that  $P(n)$  is true for any  $n \geq 1$  😊

## *Another one to try*

- What is the sum of the first  $n$  powers of 2?
- $2^0 + 2^1 + 2^2 + \dots + 2^{n-1}$
- $k = 1: 2^0 = 1$
- $k = 2: 2^0 + 2^1 = 1 + 2 = 3$
- $k = 3: 2^0 + 2^1 + 2^2 = 1 + 2 + 4 = 7$
- $k = 4: 2^0 + 2^1 + 2^2 + 2^3 = 1 + 2 + 4 + 8 = 15$
- For general  $n$ , the sum is  $2^n - 1$

# *How to prove it*

$P(n)$  = “the sum of the first  $n$  powers of 2 (starting at 0) is  $2^n - 1$ ”

Theorem:  $P(n)$  holds for all  $n \geq 1$

Proof: By induction on  $n$

- **Base case:**  $n=1$ . Sum of first 1 power of 2 is  $2^0$ , which equals  $1 = 2^1 - 1$ .
- **Inductive case:**
  - Assume the sum of the first  $k$  powers of 2 is  $2^k - 1$
  - Show the sum of the first  $(k+1)$  powers of 2 is  $2^{k+1} - 1$

# *How to prove it*

- The sum of the first  $k+1$  powers of 2 is

$$2^0 + 2^1 + 2^2 + \dots + 2^{(k-1)} + 2^k$$

sum of the first  $k$  powers of 2

by inductive hypothesis

$$= 2^k - 1 + 2^k$$

$$= 2(2^k) - 1 = 2^{k+1} - 1$$



# *Conclusion*

- Mathematical induction is a technique for proving something is true for all integers starting from a small one, usually 0 or 1.
- A proof consists of three parts:
  1. Prove it for the base case.
  2. Assume it for some integer  $k$ .
  3. With that assumption, show it holds for  $k+1$
- It can be used for complexity and correctness analyses.

# *End of Inductive Proofs!*



# *Powers of 2*

- A bit is 0 or 1 (just two different “letters” or “symbols”)
- A sequence of  $n$  bits can represent  $2^n$  distinct things
  - For example, the numbers 0 through  $2^n-1$
- $2^{10}$  is 1024 (“about a thousand”, kilo in CSE speak)
- $2^{20}$  is “about a million”, mega in CSE speak
- $2^{30}$  is “about a billion”, giga in CSE speak

Java: an **int** is 32 bits and signed, so “max int” is “about 2 billion”

a **long** is 64 bits and signed, so “max long” is  $2^{63}-1$

# *Therefore...*

Could give a unique id to...

- Every person in the U.S. with 29 bits
- Every person in the world with 33 bits
- Every person to have ever lived with 38 bits (estimate)
- Every atom in the universe with 250-300 bits

So if a password is 128 bits long and randomly generated, do you think you could guess it?