INDUCTIVELY GENERATED DATA

4.501-2023 1

Generative processes

 Virtually every infinite set considered in programming is *generated by a process*.

The fundamental example is the set \mathbb{N} of natural numbers:

▶ Initial object ("base"): The number 0 is in \mathbb{N} .

Generative processes

 Virtually every infinite set considered in programming is *generated by a process*.

The fundamental example is the set \mathbb{N} of natural numbers:

► Initial object ("base"): The number 0 is in N.

Generative step: If $n \in N$ then "next" of n, s n, is $\in \mathbb{N}$

Generative processes

 Virtually every infinite set considered in programming is *generated by a process*.

The fundamental example is the set \mathbb{N} of natural numbers:

▶ Initial object ("base"): The number 0 is in \mathbb{N} .

Generative step: If $n \in N$ then "next" of n, s n, is $\in \mathbb{N}$

• Implicit assumptions:

The meanings of 0 and "next" are known and given.

Generating $\{0,1\}^*$

- ▶ Base. The *empty string* is in {0,1}*.
- ► Generative step.

If $w \in \{0,1\}^*$ then 0w and 1w are $\in \{0,1\}^*$

Generating {0,1}*

- ▶ Base. The *empty string* is in {0,1}*.
- ► Generative step.

If $w \in \{0,1\}^*$ then 0w and 1w are $\in \{0,1\}^*$

• Implicit assumptions:

The meanings of the empty string and of juxtaposition are known.

Generating {0,1}*

- ▶ Base. The *empty string* is in {0,1}*.
- ► Generative step.

If $w \in \{0,1\}^*$ then 0w and 1w are $\in \{0,1\}^*$

• Implicit assumptions:

The meanings of the empty string and of juxtaposition are known.

 Note: We generate strings here "from the head"; This conforms with the general use of constructors, and relflected in the functions *head* and *tail*.

Generating {0,1}*

- **Base.** The *empty string* is in $\{0, 1\}^*$.
- Generative step.

If $w \in \{0,1\}^*$ then 0w and 1w are $\in \{0,1\}^*$

• Implicit assumptions:

The meanings of the empty string and of juxtaposition are known.

- Note: We generate strings here "from the head"; This conforms with the general use of constructors, and relflected in the functions *head* and *tail*.
- But numerals are in fact generated from the tail:

 $[7654321]_{10} = 1 + 10 \cdot [765432]_{10}$

Format of generative definitions

• Two parts in a generative dfn of set S:

► Base:

Particular known objects are in S.

Format of generative definitions

- Two parts in a generative dfn of set *S*:
 - ► Base:

Particular known objects are in S.

Generative steps:

If certain objects are in S

then so are vertain objects obtained from those.

Another example: Binary trees

 Binary tree means here a finite, ordered, unlabeled binary tree

Base: The singleton tree • is in **BT**. **Generative step:**

If t_0 , t_1 are binary trees then so is

Implicit assumptions:

We know what a singleton tree and juncture of trees mean.



• Generate the set E of even natural numbers.

- Generate the set E of even natural numbers.
 - ► Base: 0
 - Generative step: If $n \in E$ then $n-2 \in E$.

• Two ways to define boolean terms (closed ones: no variables):

- Two ways to define boolean terms (closed ones: no variables):
- IBT: Infix boolean terms:
 - ▶ 0 and 1 are in IBT
 - ▶ If $t, t' \in IBT$ then $(t) \land (t') \in IBT$ and $(t) \lor (t') \in IBT$

- Two ways to define boolean terms (closed ones: no variables):
- IBT: Infix boolean terms:
 - ▶ 0 and 1 are in IBT
 - If $t, t' \in IBT$ then $(t) \land (t') \in IBT$ and $(t) \lor (t') \in IBT$
- **PBT**: Prefix boolean terms:
 - ▶ 0 and 1 are in PBT
 - If $t, t' \in \mathbf{PBT}$ then $\wedge t t' \in \mathbf{PBT}$ and $\vee t t' \in \mathbf{PBT}$

- Two ways to define boolean terms (closed ones: no variables):
- IBT: Infix boolean terms:
 - ▶ 0 and 1 are in IBT
 - If $t, t' \in IBT$ then $(t) \land (t') \in IBT$ and $(t) \lor (t') \in IBT$
- **PBT**: Prefix boolean terms:
 - ▶ 0 and 1 are in PBT
 - If $t, t' \in \mathbf{PBT}$ then $\wedge t t' \in \mathbf{PBT}$ and $\vee t t' \in \mathbf{PBT}$
- Main difference between IBT and PBT:

No parentheses in PBT !

Lists of natural numbers

- Generate $\mathbb{L}(\mathbb{N})$ the **lists of natural numbers**.
- Fix a textual coding of \mathbb{N} , say binary numerals.

Lists of natural numbers

- Generate $\mathbb{L}(\mathbb{N})$ the **lists of natural numbers**.
- Fix a textual coding of \mathbb{N} , say binary numerals.
 - \blacktriangleright \Box is a list of naturals.
 - If ℓ is a list and k a numeral then $k: \ell$ is a list.

Lists of natural numbers

- Generate $\mathbb{L}(\mathbb{N})$ the **lists of natural numbers**.
- Fix a textual coding of \mathbb{N} , say binary numerals.
 - \blacktriangleright \Box is a list of naturals.
 - If ℓ is a list and k a numeral then $k: \ell$ is a list.
- Examples: $1:\Box$, $0:101:10011:10:\Box$.

REASONING ABOUT INDUCTIVE DATA

4.501-2023 9

• \mathbb{N} is infinite. $\{0, 1\}^*$ is infinite.

But our minds and our proofs are finite.

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about $\mathbb N$?

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about №?
- Trying many cases is never sufficient.
 Example: "For all n at least one of 2ⁿ + 1 and 2ⁿ 1 is prime."
 2, 3, 5, 7, 17, 31 Hooray!

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about N ?
- Trying many cases is never sufficient.
 Example: "For all n at least one of 2ⁿ + 1 and 2ⁿ 1 is prime."
 2,3,5,7,17,31 Hooray!

Oops: Both 63 and 65 are composite.

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about N ?
- Trying many cases is never sufficient.
 Example: "For all n at least one of 2ⁿ + 1 and 2ⁿ 1 is prime." 2,3,5,7,17,31 Hooray!
 Oops: Both 63 and 65 are composite.
- Another try (Fermat): All numbers $2^{2^n} + 1$ are prime 3, 5, 17, 257, 65537. Yahoo!

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about №?
- Trying many cases is never sufficient.
 Example: "For all n at least one of 2ⁿ + 1 and 2ⁿ 1 is prime." 2,3,5,7,17,31 Hooray!
 Oops: Both 63 and 65 are composite.
- Another try (Fermat): All numbers $2^{2^n} + 1$ are prime 3, 5, 17, 257, 65537. Yahoo!

Oops (Euler): Next one is 4,294,967,297, which is divisible by 641.

- N is infinite. {0, 1}* is infinite.
 But our minds and our proofs are finite.
- So how can we prove anything about №?
- Trying many cases is never sufficient.
 Example: "For all n at least one of 2ⁿ + 1 and 2ⁿ 1 is prime." 2,3,5,7,17,31 Hooray!
 Oops: Both 63 and 65 are composite.
- Another try (Fermat): All numbers 2^{2ⁿ} + 1 are prime 3, 5, 17, 257, 65537. Yahoo!
 Oops (Euler): Next one is 4,294,967,297, which is divisible by 641.
- So how can we hope to prove that all natural numbers are such-and-such ?

Finitely generated infinities!

- The secret is that inductive data is generated by *finite rules.*
- Therefore we have a finite tool for proving that all generated objects satisfy certain properties.

• Suppose we generate ${\mathbb N}$ using a green pen.

- Suppose we generate \mathbb{N} using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

- Suppose we generate ${\mathbb N}$ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0

- Suppose we generate ${\mathbb N}$ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0 1

- Suppose we generate ${\mathbb N}$ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0 1 2

- Suppose we generate ${\mathbb N}$ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.
 - 0 1 2 3

- Suppose we generate ${\mathbb N}$ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0 1 2 3 4

Following the process

- Suppose we generate ℕ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0 1 2 3 4

• They all come out green:

As we generate N we make sure that we start with green, and that each step maintains green-ness.

Following the process

- Suppose we generate ℕ using a green pen.
 - ▶ 0 is a green natural.
 - If x is a green natural, then so is its successor.

0 1 2 3 4

• They all come out green:

As we generate N we make sure that we start with green, and that each step maintains green-ness.

 Green-ness is here the process' *invariant:* True at the outset, and preserved by the steps.

Suppose P(x) is a property of natural numbers x.
 P(x) abbreviates here "x has the property P"

- Suppose P(x) is a property of natural numbers x.
 P(x) abbreviates here "x has the property P"
- Assume:
 - ► **Base. P**(0) and
 - ▶ Step. For all $n \in \mathbb{N}$, P(n) implies P(n+1).

- Suppose P(x) is a property of natural numbers x.
 P(x) abbreviates here "x has the property P"
- Assume:
 - **Base.** P(0) and
 - ▶ Step. For all $n \in \mathbb{N}$, P(n) implies P(n+1).
- Conclude: P(x) for all $x \in \mathbb{N}$.

- Suppose P(x) is a property of natural numbers x. P(x) abbreviates here "x has the property P"
- Assume:
 - **Base.** P(0) and
 - ▶ Step. For all $n \in \mathbb{N}$, P(n) implies P(n+1).
- Conclude: P(x) for all $x \in \mathbb{N}$.
- As natural numbers are being generated, they all come out satisfying *P*.

- Suppose P(x) is a property of natural numbers x.
 P(x) abbreviates here "x has the property P"
- Assume:
 - **Base.** P(0) and
 - ▶ Step. For all $n \in \mathbb{N}$, P(n) implies P(n+1).
- Conclude: P(x) for all $x \in \mathbb{N}$.
- As natural numbers are being generated, they all come out satisfying *P*.
- A property of natural numbers that holds for zero and is invariant under successor is true of every natural number.

- Suppose P(x) is a property of natural numbers x.
 P(x) abbreviates here "x has the property P"
- Assume:
 - **Base.** P(0) and
 - ▶ Step. For all $n \in \mathbb{N}$, P(n) implies P(n+1).
- Conclude: P(x) for all $x \in \mathbb{N}$.
- As natural numbers are being generated, they all come out satisfying *P*.
- A property of natural numbers that holds for zero and is invariant under successor is true of every natural number.
- The premise of the STEP is often called the "induction assumption" or the Induction Hypothesis (IH).

• Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$. What is the property?

- Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- P(x) is $2^x < 2^{x+1}$

- Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- If we know that
 - $2^x < 2^{x+1}$ is true for x = 0; and
 - ► $2^x < 2^{x+1}$ for x = nimplies that $2^x < 2^{x+1}$ for x = n+1
- then $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.

- Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- If we know that
 - $2^x < 2^{x+1}$ is true for x = 0; and
 - ► $2^x < 2^{x+1}$ for x = nimplies that $2^x < 2^{x+1}$ for x = n+1
- then $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- · But we do have
 - ► Base: $2^0 = 1 < 2 = 2^{0+1}$

- Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- · If we know that
 - $2^x < 2^{x+1}$ is true for x = 0; and
 - ► $2^x < 2^{x+1}$ for x = nimplies that $2^x < 2^{x+1}$ for x = n+1
- then $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- But we do have
 - Base: $2^0 = 1 < 2 = 2^{0+1}$
 - Step: If $2^n < 2^{n+1}$ (P(x) for x = n) then

 $2^{n+1} = 2^n + 2^n < 2^{n+1} + 2^{n+1} = 2^{n+2}$ next(P(x) for x = n+1)

- Show that $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- · If we know that
 - $2^x < 2^{x+1}$ is true for x = 0; and
 - ► $2^x < 2^{x+1}$ for x = nimplies that $2^x < 2^{x+1}$ for x = n+1
- then $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.
- But we do have
 - Base: $2^0 = 1 < 2 = 2^{0+1}$
 - ► Step: If $2^n < 2^{n+1}$ (P(x) for x = n) then $2^{n+1} = 2^n + 2^n < 2^{n+1} + 2^{n+1} = 2^{n+2}$ next(P(x) for x = n+1)
- By Induction, $2^x < 2^{x+1}$ for all $x \in \mathbb{N}$.

Try this...

- Prove by induction on \mathbb{N} that $x \leq 2^x$ for all $x \in \mathbb{N}$. We are given that exponentiation is an increasing function.
- By Induction $x \leq 2^x$ for all $x \in \mathbb{N}$.

Try this...

- Prove by induction on \mathbb{N} that $x \leq 2^x$ for all $x \in \mathbb{N}$. We are given that exponentiation is an increasing function.
 - Base: For x = 0 we have $x^2 = 0 < 1 = 2^x$.
 - ▶ Step: Assume $n \leq 2^n$. Then

$$n+1 \leqslant 2^{n}+1 \quad (IH)$$

$$= 2^{n}+2^{0}$$

$$\leqslant 2^{n}+2^{n} \quad (exponentiation is increasing)$$

$$= 2^{n+1}$$

• By Induction $x \leq 2^x$ for all $x \in \mathbb{N}$.

- P(x): $x^3 + 2x$ is divisible by 3.
- By Induction:

Example: Divisibility

- P(x): $x^3 + 2x$ is divisible by 3.
- By Induction:
 - Base. For x = 0 $x^3 + 2x = 0^3 + 2 \cdot 0 = 0$ which is divisible by 3.

- P(x): $x^3 + 2x$ is divisible by 3.
- By Induction:
 - **Base.** For x = 0

 $x^{3} + 2x = 0^{3} + 2 \cdot 0 = 0$ which is divisible by 3.

► Step. Assume
$$P(n)$$
 (IH). Then for $x = n+1$
 $x^3 + 2x = (n+1)^3 + (2n+2)$
 $= (n^3 + 3n^2 + 3n + 1) + (2n+2)$
 $= (n^3 + 2n) + 3(n^2 + n + 1)$

 $x^3 + 2x$ is the sum of numbers divisible by 3, and is therefore divisible by 3.

- P(x): $x^3 + 2x$ is divisible by 3.
- By Induction:
 - ▶ Base. For x = 0 x³ + 2x = 0³ + 2 ⋅ 0 = 0 which is divisible by 3.
 ▶ Step. Assume P(n) (IH). Then for x = n+1 x³ + 2x = (n + 1)³ + (2n + 2)

$$= (n^{3} + 3n^{2} + 3n + 1) + (2n + 2)$$
$$= (n^{3} + 3n^{2} + 3n + 1) + (2n + 2)$$
$$= (n^{3} + 2n) + 3(n^{2} + n + 1)$$

 $x^3 + 2x$ is the sum of numbers divisible by 3, and is therefore divisible by 3.

• By Induction $x^3 + 2x$ is divisible by 3, for all $x \in \mathbb{N}$.

- P(x): $x^3 + 2x$ is divisible by 3.
- By Induction:
 - ► Base. For x = 0 x³ + 2x = 0³ + 2 ⋅ 0 = 0 which is divisible by 3.
 ► Step. Assume P(n) (IH). Then for x = n+1 x³ + 2x = (n + 1)³ + (2n + 2)

$$= (n^{3} + 3n^{2} + 3n + 1) + (2n + 2)$$

$$= (n^{3} + 3n^{2} + 3n + 1) + (2n + 2)$$

$$= (n^{3} + 2n) + 3(n^{2} + n + 1)$$

 $x^3 + 2x$ is the sum of numbers divisible by 3, and is therefore divisible by 3.

• By Induction $x^3 + 2x$ is divisible by 3, for all $x \in \mathbb{N}$.

(*) $0+1+2+\cdots+x = x(x+1)/2$

• By Induction :

- (*) $0+1+2+\cdots+x = x(x+1)/2$
- By Induction :
 - ▶ **Base.** (*) is true for x = 0: $0 = 0 \cdot (0 + 1)/2$.

(*)
$$0+1+2+\cdots+x = x(x+1)/2$$

• By Induction :

• Base. (*) is true for x = 0: $0 = 0 \cdot (0 + 1)/2$. • Step. Assume (*) for x = n. Then for x = n+1 $0 + 1 + \dots + x = 0 + 1 + \dots + n + (n+1)$ $= \frac{n(n+1)}{2} + (n+1)$ (IH) $= (n + 1)(\frac{1}{2}n + 1)$ $= \frac{1}{2}(n+1)(n+2)$ $= \frac{1}{2}x(x + 1)$

That is, (\star) for x = n+1.

(*)
$$0+1+2+\cdots+x = x(x+1)/2$$

• By Induction :

• Base. (*) is true for x = 0: $0 = 0 \cdot (0 + 1)/2$. • Step. Assume (*) for x = n. Then for x = n+1 $0+1+\dots+x = 0+1+\dots+n+(n+1)$ $= \frac{n(n+1)}{2} + (n+1)$ (IH) $= (n+1)(\frac{1}{2}n+1)$ $= \frac{1}{2}(n+1)(n+2)$ $= \frac{1}{2}x(x+1)$

That is, (\star) for x = n+1.

• Conclude: (*) holds for every $x \in \mathbb{N}$.

- A property of natural numbers may refer to non-numeric data!
 - (*) Every set with x elements has 2^x subsets
- By Induction.

- A property of natural numbers may refer to non-numeric data!
 - (*) Every set with x elements has 2^x subsets
- By Induction.
 - ▶ Base. x = 0. The only set with 0 elements is Ø, which has just $2^0 = 1$ subset, namely Ø itself.

• A property of natural numbers may refer to non-numeric data!

(*) Every set with x elements has 2^x subsets

- By Induction.
 - ▶ Base. x = 0. The only set with 0 elements is Ø, which has just $2^0 = 1$ subset, namely Ø itself.
 - Step. Assume P(n) (IH).

For x = n+1 let S be a set with n+1 elements.

Choose $a \in S$ (*S* can't be empty!) and let $S^- =_{df} S - \{a\}$.

- A property of natural numbers may refer to non-numeric data!
 - (*) Every set with x elements has 2^x subsets
- By Induction.
 - ▶ **Base.** x = 0. The only set with 0 elements is Ø, which has just $2^0 = 1$ subset, namely Ø itself.
 - Step. Assume P(n) (IH).

For x = n+1 let S be a set with n+1 elements. Choose $a \in S$ (S can't be empty!) and let $S^- =_{df} S - \{a\}$. By IH S^- has 2^n subsets A_1, \ldots, A_{2^n} . Subsets of $S : A_1, \ldots, A_{2^n}, A_1 \cup \{a\}, \ldots, A_{2^n} \cup \{a\}$ which are all different. So S has $2^n + 2^n = 2^{n+1}$ subsets.

- A property of natural numbers may refer to non-numeric data!
 - (*) Every set with x elements has 2^x subsets
- By Induction.
 - ▶ Base. x = 0. The only set with 0 elements is Ø, which has just $2^0 = 1$ subset, namely Ø itself.
 - ▶ Step. Assume P(n) (IH).

For x = n+1 let S be a set with n+1 elements.

Choose $a \in S$ (*S* can't be empty!) and let $S^- =_{df} S - \{a\}$. By IH S^- has 2^n subsets A_1, \ldots, A_{2^n} .

Subsets of *S*: $A_1, ..., A_{2^n}, A_1 \cup \{a\}, ..., A_{2^n} \cup \{a\}$

which are all different. So *S* has $2^n + 2^n = 2^{n+1}$ subsets.

• By Induction (*) for all $x \in \mathbb{N}$.

- Show $x^2 > x$ for all x > 1.
- We wish to start induction from 2.

- Show $x^2 > x$ for all x > 1.
- We wish to start induction from 2.

But that's the same as Induction for the property

 $(x+2)^2 > (x+2)$!

- Show $x^2 > x$ for all x > 1.
- We wish to start induction from 2.

But that's the same as Induction for the property

 $(x+2)^2 > (x+2)$!

• We refer to this as Shifted Induction:

► Base.
$$2^2 = 4 > 2$$
► Step. $n^2 > n$ implies
 $(n+1)^2 = n^2 + 2n + 1$
> $n+2n+1$ (IH)
> $n+1$ since $n > 0$)

- Show $x^2 > x$ for all x > 1.
- We wish to start induction from 2.

But that's the same as Induction for the property

 $(x+2)^2 > (x+2)$!

• We refer to this as Shifted Induction:

► Base.
$$2^2 = 4 > 2$$
► Step. $n^2 > n$ implies
 $(n+1)^2 = n^2 + 2n + 1$
> $n+2n+1$ (IH)
> $n+1$ since $n > 0$)

• Conclusion: $x^2 > x$ for all integers x > 1.

Shifted Induction

- The template for such reasoning is **Shifted Induction**
- Given a property P(x) of natural numbers, and $b \in \mathbb{N}$,
- Assume: Shifted Base. *P* true of *b*; and

▶ Shifted Step. For all $n \ge b$, P(n) implies P(n+1)

- Conclude: P(x) for all $x \ge b$.
- Induction is a special case, with b = 0.

Another example

- $3^n > 5 \cdot 2^n$ for all $n \ge 4$.
- By Shifted Induction with initial value 4.

Another example

- $3^n > 5 \cdot 2^n$ for all $n \ge 4$.
- By Shifted Induction with initial value 4.
 - <u>Basis</u>. $3^4 = 81 > 80 = 5 \cdot 2^4$

Another example

- $3^n > 5 \cdot 2^n$ for all $n \ge 4$.
- By Shifted Induction with initial value 4.
 - <u>Basis</u>. $3^4 = 81 > 80 = 5 \cdot 2^4$
 - Step. If $3^n > 5 \cdot 2^n$ then

$$3^{n+1} = 3 \cdot 3^n$$

> $3 \cdot (5 \cdot 2^n)$ (IH)
> $2 \cdot 5 \cdot 2^n$
= $5 \cdot 2^{n+1}$

Inductive reasoning in general

- The principle of inductive reasoning applies to any inductively generated set *S*, not just N.
- If P(x) makes sense for x ∈ S,
 is true for every base element of S
 and remains true under the generative steps for S,
 then P(x) is true for all x ∈ S.

Inductive reasoning in general

- The principle of inductive reasoning applies to any inductively generated set *S*, not just N.
- If P(x) makes sense for x ∈ S,
 is true for every base element of S
 and remains true under the generative steps for S,
 then P(x) is true for all x ∈ S.
- The underlying reason is the same as for N: as the elements of *S* are generated, the property *P* invariantly holds.

- P(x) may mention also non-numeric data.
- Example: Take "Every set with x elements has 2^x subsets"

- P(x) may mention also non-numeric data.
- Example: Take "Every set with x elements has 2^x subsets"
- We prove by Induction on \mathbb{N} that P(x) for all $x \in \mathbb{N}$.

- P(x) may mention also non-numeric data.
- Example: Take "Every set with x elements has 2^x subsets"
- We prove by Induction on N that P(x) for all $x \in \mathbb{N}$.
 - Base. P(0): "Every set with 0 elements has $2^0 = 1$ subsets". Indeed \emptyset has one subset.

- P(x) may mention also non-numeric data.
- Example: Take "Every set with x elements has 2^x subsets"
- We prove by Induction on N that P(x) for all $x \in \mathbb{N}$.
 - Base. P(0): "Every set with 0 elements has $2^0 = 1$ subsets". Indeed \emptyset has one subset.
 - Step. Assume P(n).

To prove P(n+1) let S be a set with n+1 elements. Choose $a \in S$ and let $S^- = S - \{a\}$. By IH S^- has 2^n subsets A_1, \ldots, A_{2^n} . The subsets of S are $A_1, \ldots, A_{2^n}, A_1 \cup \{a\}, \ldots, A_{2^n} \cup \{a\}$, which are all different.

So S has $2^n + 2^n = 2^{n+1}$ subsets.

- P(x) may mention also non-numeric data.
- Example: Take "Every set with x elements has 2^x subsets"
- We prove by Induction on N that P(x) for all $x \in \mathbb{N}$.
 - Base. P(0): "Every set with 0 elements has $2^0 = 1$ subsets". Indeed \emptyset has one subset.
 - Step. Assume P(n).

To prove P(n+1) let S be a set with n+1 elements. Choose $a \in S$ and let $S^- = S - \{a\}$. By IH S^- has 2^n subsets A_1, \ldots, A_{2^n} . The subsets of S are $A_1, \ldots, A_{2^n}, A_1 \cup \{a\}, \ldots, A_{2^n} \cup \{a\}$, which are all different.

So *S* has $2^n + 2^n = 2^{n+1}$ subsets.

– By Induction on \mathbb{N} P(x) holds for all $x \in \mathbb{N}$.

Induction on strings

• Let P(x) be a property of Σ -strings.

Induction on strings

- Let P(x) be a property of Σ -strings.
- Assume:
 - ▶ Base. $P(\varepsilon)$
 - ▶ Steps. For each $\sigma \in \Sigma$ and $w \in \Sigma^*$ P(w) implies $P(\sigma w)$

Induction on strings

- Let P(x) be a property of Σ -strings.
- Assume:
 - ▶ Base. $P(\varepsilon)$
 - ▶ Steps. For each $\sigma \in \Sigma$ and $w \in \Sigma^*$ P(w) implies $P(\sigma w)$
- **Conclude:** P(w) for all $w \in \Sigma^*$.

 For w ∈ {0,1}* let ∽(w) ("swap w") be w with 0 and 1 interchanged: ∽001 = 110. We show (*) ∽(∽(w)) = w

- For w ∈ {0,1}* let ∽(w) ("swap w") be w with 0 and 1 interchanged: ∽001 = 110. We show (*) ∽(∽(w)) = w
- The proof is by induction on $\{0,1\}^*$.

- For w ∈ {0,1}* let ∽(w) ("swap w") be w with 0 and 1 interchanged: ∽001 = 110.
 We show (*) ∽(∽(w)) = w
- The proof is by induction on $\{0,1\}^*$.
 - ▶ Basis. $\backsim(\backsim(\varepsilon)) = \backsim(\varepsilon) = \varepsilon$

- For $w \in \{0, 1\}^*$ let $\backsim(w)$ ("swap w") be w with 0 and 1 interchanged: $\backsim001 = 110$. We show $(\star) \backsim(\backsim(w)) = w$
- The proof is by induction on $\{0,1\}^*$.
 - ► Basis. $\backsim(\backsim(\varepsilon)) = \backsim(\varepsilon) = \varepsilon$ ► Step for 0. If $\backsim(\backsim(x)) = x$ then $\backsim(\backsim(0x)) = \backsim(1\backsim(x))$ $= 0 \backsim(\backsim(x))$ = 0x (IH)

Step for 1 is similar.

- For w ∈ {0,1}* let ∽(w) ("swap w") be w with 0 and 1 interchanged: ∽001 = 110. We show (★) ∽(∽(w)) = w
- The proof is by induction on $\{0, 1\}^*$.
 - ► Basis. $\neg(\neg(\varepsilon)) = \neg(\varepsilon) = \varepsilon$ ► Step for 0. If $\neg(\neg(x)) = x$ then $\neg(\neg(0x)) = \neg(1\neg(x))$ $= 0 \neg(\neg(x))$ = 0x (IH)

Step for 1 is similar.

• By induction on $\{0, 1\}^*$ (*) for all $w \in \{0, 1\}^*$.

• Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).

- Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).
- Problem: This is a property of a pair of strings!

- Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).
- Solution: Read it as a property of one *x*:

 (\star)

 $|x \cdot u| = |x| + |u|$ for all $u \in \Sigma^*$

- Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).
- Solution: Read it as a property of one *x*:

 $|x \cdot u| = |x| + |u|$ for all $u \in \Sigma^*$

► Basis: $x = \varepsilon$. $|\varepsilon \cdot u| = |u|$ since $\varepsilon \cdot u = u$ $|\varepsilon| + |u| = 0 + |u| = |u|$ (\star)

- Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).
- Solution: Read it as a property of one *x*:

 (\star)

 $|x \cdot u| = |x| + |u|$ for all $u \in \Sigma^*$

► Basis: $x = \varepsilon$. $|\varepsilon \cdot u| = |u|$ since $\varepsilon \cdot u = u$ $|\varepsilon| + |u| = 0 + |u| = |u|$ ► Step: Assume (*) for x = w.

For $x = \sigma w$ we have for all $u \in \Sigma^*$ $|\sigma w \cdot u| = |\sigma(w \cdot u)|$ $= 1 + |w \cdot u|$ = 1 + |w| + |u| (IH) $= (|\sigma w|) + |u|$

- Prove $|x \cdot u| = |x| + |u|$ ($x, u \in \Sigma^*$).
- Solution: Read it as a property of one *x*:

 $|x \cdot u| = |x| + |u|$ for all $u \in \Sigma^*$

► Basis: $x = \varepsilon$. $|\varepsilon \cdot u| = |u|$ since $\varepsilon \cdot u = u$ $|\varepsilon| + |u| = 0 + |u| = |u|$ ► Step: Assume (*) for x = w.

For $x = \sigma w$ we have for all $u \in \Sigma^*$ $|\sigma w \cdot u| = |\sigma w| + |u|$

• By induction on Σ^* conclude (\star) for all $x \in \Sigma^*$.

 (\star)

Unambiguous PBT's

- A PBT *t* is **unambiguous** if it is exactly one of:
 - ▶ 0 or 1
 - $\blacktriangleright \wedge t_0 t_1$ or $\lor t_0 t_1$ for some unique terms t_0 and t_1 .

- A PBT *t* is **unambiguous** if it is exactly one of:
 - ▶ 0 or 1
 - $\blacktriangleright \wedge t_0 t_1$ or $\lor t_0 t_1$ for some unique terms t_0 and t_1 .
- That is: t can't be both $\wedge t_0 t_1$ and $\wedge t'_0 t'_1$ unless $t'_0 = t_0$ and $t'_1 = t_1$.

- A PBT *t* is **unambiguous** if it is exactly one of:
 - ▶ 0 or 1
 - $\blacktriangleright \wedge t_0 t_1$ or $\lor t_0 t_1$ for some unique terms t_0 and t_1 .
- That is: t can't be both $\wedge t_0 t_1$ and $\wedge t'_0 t'_1$ unless $t'_0 = t_0$ and $t'_1 = t_1$.
- Theorem: Every PBT is unambiguous

- A PBT *t* is **unambiguous** if it is exactly one of:
 - ▶ 0 or 1
 - $\blacktriangleright \wedge t_0 t_1$ or $\lor t_0 t_1$ for some unique terms t_0 and t_1 .
- That is: t can't be both $\wedge t_0 t_1$ and $\wedge t'_0 t'_1$ unless $t'_0 = t_0$ and $t'_1 = t_1$.
- Theorem: Every PBT is unambiguous
- How to prove this?

- A PBT *t* is **unambiguous** if it is exactly one of:
 - ▶ 0 or 1
 - $\blacktriangleright \wedge t_0 t_1$ or $\lor t_0 t_1$ for some unique terms t_0 and t_1 .
- That is: t can't be both $\wedge t_0 t_1$ and $\wedge t'_0 t'_1$ unless $t'_0 = t_0$ and $t'_1 = t_1$.
- Theorem: Every PBT is unambiguous
- How to prove this?
- Induction on terms does not work:
 - If $t = \wedge t_0 t_1$ what can we possibly conclude from assuming that
 - t_0 and t_1 are unambiguous?

• Maybe induction on strings over $\Sigma = \{0, 1, \Lambda, V\}$

- Maybe induction on strings over $\Sigma = \{0, 1, \Lambda, V\}$
- But we can't conclude that $\wedge t_1 t_2$ is unambiguous from the assumption that t_1, t_2 are unambiguous!

- Maybe induction on strings over $\Sigma = \{0, 1, \Lambda, V\}$
- But we can't conclude that $\wedge t_1 t_2$ is unambiguous from the assumption that t_1, t_2 are unambiguous!
- And the string t_1t_2 is not a term!

- Maybe induction on strings over $\Sigma = \{0, 1, \Lambda, V\}$
- But we can't conclude that $\wedge t_1 t_2$ is unambiguous from the assumption that t_1, t_2 are unambiguous!
- And the string t_1t_2 is not a term!
- Solution: A broader notion of "non-ambiguity"!
 w ∈ Σ* is *unambiguous* if it can'e be read as a *concatenation* of terms in more than one way:

- Maybe induction on strings over $\Sigma = \{0, 1, \Lambda, V\}$
- But we can't conclude that $\wedge t_1 t_2$ is unambiguous from the assumption that t_1, t_2 are unambiguous!
- And the string t_1t_2 is not a term!
- Solution: A broader notion of "non-ambiguity"!
 w ∈ Σ* is *unambiguous* if it can'e be read as a *concatenation* of terms in more than one way:
- If $w = t_1 \cdots t_k = t'_1 \cdots t'_m$

then m = k and $t_i = t'_i$ for $i \in [1..k]$.

A generalized non-ambiguity theorem

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .

A generalized non-ambiguity theorem

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.

A generalized non-ambiguity theorem

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.
 - Step: Assume that w is unambiguous and

 $(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

$$(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$$

(*) $\sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

- Case $\sigma = 0$. ($\sigma = 1$ is similar.)

0 is the only term starting with 0. So $t_1 = t'_1 = 0$ and $w = t_2 \cdots t_k = t'_2 \cdots t'_m$.

 $(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

- Case $\sigma = 0$. ($\sigma = 1$ is similar.) 0 is the only term starting with 0. So $t_1 = t'_1 = 0$ and $w = t_2 \cdots t_k = t'_2 \cdots t'_m$. By IH k-1 = m-1 and $t_i = t'_i$ for i = 2..k. So k = m and $t_i = q_i$ for i = 1..k.

 $(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

- Case $\sigma = 0$. ($\sigma = 1$ is similar.) 0 is the only term starting with 0. So $t_1 = t'_1 = 0$ and $w = t_2 \cdots t_k = t'_2 \cdots t'_m$. By IH k-1 = m-1 and $t_i = t'_i$ for i = 2..k. So k = m and $t_i = q_i$ for i = 1..k.

- Case
$$\sigma = \wedge$$
. ($\sigma = \lor$ is similar.)
 $t_1 = \wedge q r$ and $t'_1 = \wedge q' r'$ for some q, r, q', r'

$$(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$$

- Case $\sigma = 0$. ($\sigma = 1$ is similar.) 0 is the only term starting with 0. So $t_1 = t'_1 = 0$ and $w = t_2 \cdots t_k = t'_2 \cdots t'_m$. By IH k-1 = m-1 and $t_i = t'_i$ for i = 2..k. So k = m and $t_i = q_i$ for i = 1..k.

- Case
$$\sigma = \wedge$$
. ($\sigma = \lor$ is similar.)
 $t_1 = \wedge q r$ and $t'_1 = \wedge q' r'$ for some q, r, q', r' .

So $w = q r t_2 \cdots t_k = q' r' t'_2 \cdots t'_m$. By IH k+1 = m+1, q = q', r = r' and $t_i = q_i$ for i = 2..k.

$$(\star) \quad \sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$$

- Case
$$\sigma = 0$$
. ($\sigma = 1$ is similar.)
0 is the only term starting with 0.
So $t_1 = t'_1 = 0$ and $w = t_2 \cdots t_k = t'_2 \cdots t'_m$.
By IH $k-1 = m-1$ and $t_i = t'_i$ for $i = 2..k$.
So $k = m$ and $t_i = q_i$ for $i = 1..k$.

- Case
$$\sigma = \wedge$$
. ($\sigma = \vee$ is similar.)
 $t_1 = \wedge q r$ and $t'_1 = \wedge q' r'$ for some q, r, q', r' .

So $w = q r t_2 \cdots t_k = q' r' t'_2 \cdots t'_m$. By IH k+1 = m+1, q = q', r = r' and $t_i = q_i$ for i = 2..k.

So k = m and $t_i = t'_i$ for i = 1..k.

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.
 - Step: Assume that w is unambiguous and

 $\sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.
 - Step: Assume that w is unambiguous and $\sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$

We concluded that k = m and $t_i = t'_i$ for i = 1..k.

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.
 - Step: Assume that w is unambiguous and $\sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$
- By induction on Σ* we conclude that every w ∈ Σ* is unambiguous.

- Non-ambiguity Theorem. Every $w \in \Sigma^*$ is unambiguous.
- The proof is by induction on Σ^* .
 - **Basis.** If w is ε then it is unambiguous vacuously.
 - Step: Assume that w is unambiguous and $\sigma w = t_1 \cdots t_k = t'_1 \cdots t'_m$
- By induction on Σ* we conclude that every w ∈ Σ* is unambiguous.
- In particular, every PBT t is a concatenation of 1 string, and therefore must be unambiguous as a term.

Induction over binary trees

 Recall that the set of binary trees is generated from a base tree
 by juncture:

if t_0, t_1 are binary trees then so is

• Let P(x) be a property that makes sense for any binary tree t.

 $t_0 = t_1$

- If we can show that
 - ▶ **Base:** $P(\bullet)$; and
 - Step: If both $P(t_0)$ and $P(t_1)$

then P(t) for the juncture t above of t_0 and t_1

then P(t) is true for all binary trees t.

• Can a binary tree have an even number of nodes?

• Every binary tree has an odd number of nodes.

- Every binary tree has an odd number of nodes.
- Let P(t) be the property

"t has an odd number of nodes"

- Every binary tree has an odd number of nodes.
- Let P(t) be the property

"t has an odd number of nodes"

Induction on trees:

• **Basis:** $P(\bullet)$ (since 1 is odd)

- Every binary tree has an odd number of nodes.
- Let P(t) be the property

"t has an odd number of nodes"

Induction on trees:

- **Basis:** $P(\bullet)$ (since 1 is odd)
- Step: Suppose t_0, t_1 are trees of odd sizes n_0 and n_1 .

Let *t* be obtained from t_0 and t_1 . then the size of *t* is $n_0 + n_1 + 1$ which is again odd.

- Every binary tree has an odd number of nodes.
- Let P(t) be the property

"t has an odd number of nodes"

Induction on trees:

- **Basis:** $P(\bullet)$ (since 1 is odd)
- Step: Suppose t_0, t_1 are trees of odd sizes n_0 and n_1 .

Let *t* be obtained from t_0 and t_1 . then the size of *t* is $n_0 + n_1 + 1$ which is again odd.

By induction on binary tree we conclude that

P(t) for all binary trees t.

The mother of all inductions

- The Induction principle applies to all inductively generated sets.
- But induction on \mathbb{N} is the one usually invoked. *Why*?

The mother of all inductions

- The Induction principle applies to all inductively generated sets.
- But induction on N is the one usually invoked. Why?
- Take generated set \mathbb{G} , P(x) a property of $x \in \mathbb{G}$.
- Obtain induction over $x \in \mathbb{G}$ for property P(x)as induction over $n \in \mathbb{N}$ for the property:

P(n) is true for all $x \in \mathbb{G}$ generated in $\leq n$ steps

The mother of all inductions

- The Induction principle applies to all inductively generated sets.
- But induction on N is the one usually invoked. Why?
- Take generated set \mathbb{G} , P(x) a property of $x \in \mathbb{G}$.
- Obtain induction over $x \in \mathbb{G}$ for property P(x)as induction over $n \in \mathbb{N}$ for the property:

P(n) is true for all $x \in \mathbb{G}$ generated in $\leq n$ steps

Note that N is the *simplest* infinite generated set:
 one initial object, *one* generative rule, involving *one* premise!