CSE 20 Discussion: Week 7

Structural, mathematical, and strong induction

Part 1: Key concepts

Recursive definitions

- A recursively defined set is specified by a **basis step**, which lists finitely many elements, and a **recursive step**, which gives rules for building new elements from existing ones.
  - S: basis $\texttt{A}, \texttt{C}, \texttt{U}, \texttt{G}$; if $s \in S$ and $b \in B$, then $sb \in S$.
  - $\mathbb{N}$: basis 0; if $n \in \mathbb{N}$, then $n + 1 \in \mathbb{N}$.
  - L: basis $[]$; if $l \in L$ and $n \in \mathbb{N}$, then $(n, l) \in L$.
- A function on a recursively defined set may be defined in the same two steps: its value at each basis element, and its value at a new element in terms of its value at the elements used to build it. For example,

$$length(\ []\ ) = 0, \qquad length(\ (n, l)\ ) = 1 + length(l).$$

- Every element is obtained from the basis elements by finitely many applications of the recursive step. Hence a property that holds for the basis elements and is preserved by the recursive step holds for every element.

Structural induction

Proof by Structural Induction: To prove a universal quantification over a recursively defined set:

> **Basis Step**: Show the statement holds for elements specified in the basis step of the definition.
>
> **Recursive Step**: Show that if the statement is true for each of the elements used to construct new elements in the recursive step of the definition, the result holds for these new elements.

To prove $\forall x \in X \ P(x)$ for a recursively defined set X:

- **Basis step**: prove $P(e)$ for each element e in the basis step of X.
- **Recursive step**: for arbitrary elements $u, \ldots$ used by a rule, assume $P(u)$ (the **induction hypothesis**, IH) and prove P for the new element.
- In the recursive step, apply the recursive definition of each function to express the new element's value in terms of u; then apply the IH. If a definition has cases, the proof has the same cases.

Mathematical induction

Proof by Mathematical Induction: To prove a universal quantification over the set of all integers greater than or equal to some base integer b,

Basis Step: Show the property holds for b.

Recursive Step: Consider an arbitrary integer n greater than or equal to b, assume (as the **induction hypothesis**) that the property holds for n, and use this and other facts to prove that the property holds for $n + 1$.

- Mathematical induction is structural induction on the recursively defined set $\mathbb{N}$, or on the set of integers greater than or equal to b.
- The induction hypothesis is the property for n only.

Strong induction

Proof by Strong Induction: To prove that a universal quantification over the set of all integers greater than or equal to some base integer b holds, pick a fixed nonnegative integer j and then:

Basis Step: Show the statement holds for b, $b+1$, ..., $b+j$.

Recursive Step: Consider an arbitrary integer n greater than or equal to $b+j$, assume (as the **strong induction hypothesis**) that the property holds for **each of** b, $b+1$, ..., n, and use this and other facts to prove that the property holds for $n+1$.

- Strong induction is used when the property for $n+1$ depends on the property for values other than n.
- Each value to which the IH is applied must lie between b and n; this must be verified in the proof.

Theorems, lemmas, and invariants

Theorem: Statement that can be shown to be true, usually an important one.

Lemma: A less important theorem that is useful in proving a theorem.

Corollary: A theorem that can be proved directly after another one has been proved, without needing a lot of extra work.

Invariant: A theorem that describes a property that is true about an algorithm or system no matter what inputs are used.

- A lemma such as $\forall (x, y) \in Pos$ ($x + y$ is an even integer) can be written as $\forall (x, y)$ $\big((x, y) \in Pos \rightarrow x + y$ is an even integer$\big)$. If the conclusion is false for a particular (x, y), then $(x, y) \notin Pos$.
- The lemma does not imply that a point with an even coordinate sum is in Pos; that inference would use the converse.

Part 2: Examples

Example: structural induction

Claim: $\forall s \in S \ (\ rnalen(s) > 0\ )$

Basis step Each basis element of S is a base $b \in B$. By the basis step of $rnalen$, $rnalen(b) = 1 > 0$.

Recursive step WTS
$\forall u \in S\ \forall b \in B\ \big(\ rnalen(u) > 0 \rightarrow rnalen(ub) > 0\ \big)$.
Let $u \in S$ and $b \in B$ be arbitrary, and assume (IH) $rnalen(u) > 0$. Then
$rnalen(ub) \overset{\text{rec. def.}}{=} 1 + rnalen(u) \overset{\text{IH}}{>} 1 + 0 = 1 > 0$. $\square$

Key observations

- The recursive step is a universal conditional whose hypothesis is the IH for the simpler element u.
- The recursive definition of $rnalen$ relates the value at ub to the value at u, where the IH applies.

Example: mathematical induction

$sumPow : \mathbb{N} \to \mathbb{N}$ is defined by $sumPow(0) = 1$ and, for $x \in \mathbb{N}$, $sumPow(x+1) = sumPow(x) + 2^{x+1}$.

Claim: $\forall n \in \mathbb{N} \left(\; sumPow(n) = 2^{n+1} - 1 \; \right)$

Basis step $sumPow(0) = 1$ by the basis step of $sumPow$, and $2^{0+1} - 1 = 1$.

Recursive step Let $n \in \mathbb{N}$ be arbitrary, and assume (IH) $sumPow(n) = 2^{n+1} - 1$. WTS $sumPow(n+1) = 2^{(n+1)+1} - 1$.

$LHS = sumPow(n) + 2^{n+1} \overset{\text{IH}}{=} (2^{n+1} - 1) + 2^{n+1}$,
$RHS = 2 \cdot 2^{n+1} - 1 = (2^{n+1} - 1) + 2^{n+1}$. Hence $LHS = RHS$. $\square$

Key observations

- The recursive step uses the recursive definition of $sumPow$ to reduce $sumPow(n+1)$ to $sumPow(n)$.
- Simplifying each side separately avoids assuming the equality to be proved.

Example: strong induction

Theorem: Every positive integer is a sum of (one or more) distinct powers of 2.

Basis step $(b = 1,\ j = 0)\ \ 1 = 1 \cdot 2^0$.

Recursive step Let $n \geq 1$ be an arbitrary integer, and assume (IH) the property for each of $1, \ldots, n$. WTS the property for $n + 1$.
Apply the IH to $(n + 1)$ **div** 2. This is valid: it is an integer; since $n + 1 \geq 2$, it is at least 2 **div** $2 = 1$; since $n + 1 \leq 2n$, it is at most $2n$ **div** $2 = n$.
The expansion of $n + 1$ is the expansion of $(n + 1)$ **div** 2 followed by the coefficient $(n + 1)$ **mod** 2, because
$n + 1 = 2\big((n + 1)$ **div** $2\big) + (n + 1)$ **mod** 2. $\square$

Key observations

- The IH is applied to $(n + 1)$ **div** 2, not to n; mathematical induction would not suffice.
- The proof verifies that $(n + 1)$ **div** 2 lies in the range $1, \ldots, n$ covered by the IH.

Example: induction for an existential statement

Claim: $\forall n \in \mathbb{N}\ \exists l \in L\ \left(\ length(l) = n\ \right)$

Basis step WTS $\exists l \in L\ (length(l) = 0)$. The witness $l = []$ works, since $length(\ []\ ) = 0$.

Recursive step Let $n \in \mathbb{N}$ be arbitrary, and assume (IH) that some $l_w \in L$ satisfies $length(l_w) = n$. WTS $\exists l \in L\ (length(l) = n + 1)$.
Consider $l = (0, l_w)$, which is in L by the recursive step, since $0 \in \mathbb{N}$ and $l_w \in L$. Then
$length(\ (0, l_w)\ ) \overset{\text{rec. def.}}{=} 1 + length(l_w) \overset{\text{IH}}{=} 1 + n.$ $\square$

Key observations

- For an existential property, the IH provides a witness for n, and the recursive step constructs a witness for $n + 1$ from it.
- The constructed witness must be shown to belong to the domain L.

Part 3: Practice

Q1. Structural induction

Claim: For every $s \in S$,

$$rnalen(s) = basecount(\ (s, \texttt{A})\ ) + basecount(\ (s, \texttt{C})\ ) + basecount(\ (s, \texttt{U})\ ) + basecount(\ (s, \texttt{G})\ ).$$

We prove the claim by structural induction on S.

(a) State what the basis step must show, and verify it for the basis element `A`.

(b) State the induction hypothesis and what must be shown in the recursive step.

▸ Definitions

Q1. Structural induction

(c) Complete the recursive step. (The definition of $basecount$ has two cases; determine how they apply to the four terms.)

▸ Definitions

Q2. Strong induction

Theorem: Every positive integer greater than 1 is a product of (one or more) primes.

Proof outline (strong induction with $b = 2$, $j = 0$). Let $n \geq 2$ be an arbitrary integer, and assume (IH) the property for each of $2, \ldots, n$. Case 1: $n + 1$ is prime. Case 2: $n + 1$ is composite, so $n + 1 = xy$ for some integers x, y with $2 \leq x \leq n$ and $2 \leq y \leq n$; the IH applies to x and y.

(a) Explain why Cases 1 and 2 cover every possibility for $n + 1$.

(b) In Case 2, justify that $n + 1 = xy$ for integers x, y with $2 \leq x \leq n$ and $2 \leq y \leq n$.

▸ Definitions

Q2. Strong induction

(c) Why would mathematical induction, whose hypothesis is the property for n only, not suffice for this proof?

▸ Definitions

Q3. Recursive data

Recall the set L of linked lists of natural numbers and the functions $length$, $append$, and $prepend$.

(a) Prove or disprove:
$\forall l \in L\ \forall m \in \mathbb{N}\ \Big(\ append(\ (l, m)\ ) = prepend(\ (l, m)\ )\ \Big)$

(b) Consider the claim
$\forall l \in L\ \forall m \in \mathbb{N}\ \Big(\ length(\ append(\ (l, m)\ )\ ) = 1 + length(l)\ \Big)$. On which variable should the induction be performed? Justify your answer.

▸ Definitions

Q3. Recursive data

(c) Prove the claim by structural induction on L.

▸ Definitions

Appendix: Definitions from the lecture notes

Definitions for Q1

Proof by Structural Induction: To prove a universal quantification over a recursively defined set:

Basis Step: Show the statement holds for elements specified in the basis step of the definition.

Recursive Step: Show that if the statement is true for each of the elements used to construct new elements in the recursive step of the definition, the result holds for these new elements.

◀ Back to Q1

Definitions for Q1

RNA strands: The set of RNA strands S is defined (recursively) by:
Basis Step: $\mathtt{A} \in S, \mathtt{C} \in S, \mathtt{U} \in S, \mathtt{G} \in S$
Recursive Step: If $s \in S$ and $b \in B$, then $sb \in S$
where sb is string concatenation.

rnalen: The function *rnalen* that computes the length of RNA strands in S is defined recursively by:

		$\mathit{rnalen} : S$	$\to \mathbb{Z}^+$
Basis Step:	If $b \in B$ then	$\mathit{rnalen}(b)$	$= 1$
Recursive Step:	If $s \in S$ and $b \in B$, then	$\mathit{rnalen}(sb)$	$= 1 + \mathit{rnalen}(s)$

◂ Back to Q1

Definitions for Q1

basecount: The function *basecount* that computes the number of a given base b appearing in a RNA strand s is defined recursively by:

$$\textit{basecount} : S \times B \to \mathbb{N}$$

Basis Step: If $b_1 \in B, b_2 \in B$

$$\textit{basecount}(\ (b_1, b_2)\ ) = \begin{cases} 1 & \text{when } b_1 = b_2 \\ 0 & \text{when } b_1 \neq b_2 \end{cases}$$

Recursive Step: If $s \in S, b_1 \in B, b_2 \in B$

$$\textit{basecount}(\ (sb_1, b_2)\ ) = \begin{cases} 1 + \textit{basecount}(\ (s, b_2)\ ) & \text{when } b_1 = b_2 \\ \textit{basecount}(\ (s, b_2)\ ) & \text{when } b_1 \neq b_2 \end{cases}$$

◀ Back to Q1

Definitions for Q2

Proof by Strong Induction: To prove that a universal quantification over the set of all integers greater than or equal to some base integer b holds, pick a fixed nonnegative integer j and then:

Basis Step: Show the statement holds for b, $b+1$, ..., $b+j$.

Recursive Step: Consider an arbitrary integer n greater than or equal to $b+j$, assume (as the **strong induction hypothesis**) that the property holds for **each of** b, $b+1$, ..., n, and use this and other facts to prove that the property holds for $n+1$.

◀ Back to Q2

Definitions for Q2

Definition: An integer p greater than 1 is called **prime** means the only positive factors of p are 1 and p. A positive integer that is greater than 1 and is not prime is called composite.

Facts about numbers:

1. Addition and multiplication of real numbers are each commutative and associative.
2. The product of two positive numbers is positive, of two negative numbers is positive, and of a positive and a negative number is negative.
3. The sum of two integers, the product of two integers, and the difference between two integers are each integers.
4. For every integer x there is no integer strictly between x and $x + 1$,
5. When x, y are positive integers, $xy \geq x$ and $xy \geq y$.

◂ Back to Q2

Definitions for Q3

Definition: The set of linked lists of natural numbers L is defined recursively by

$$\begin{array}{ll} \text{Basis Step:} & [] \in L \\ \text{Recursive Step:} & \text{If } l \in L \text{ and } n \in \mathbb{N}, \text{ then } (n, l) \in L \end{array}$$

Definition: The length of a linked list of natural numbers L, $length : L \to \mathbb{N}$ is defined by

$$\begin{array}{lllll} \text{Basis Step:} & & length([]) & = 0 \\ \text{Recursive Step:} & \text{If } l \in L \text{ and } n \in \mathbb{N}, \text{ then} & length((n, l)) & = 1 + length(l) \end{array}$$

◀ Back to Q3

Definitions for Q3

Definition: The function $append : L \times \mathbb{N} \to L$ that adds an element at the end of a linked list is defined by

Basis Step: If $m \in \mathbb{N}$ then $append(\ ([], m)\ ) = (m, [])$

Recursive Step: If $l \in L$ and $n \in \mathbb{N}$ and $m \in \mathbb{N}$, then
$append(\ (\ (n, l), m\ )\ ) = (n, append(\ (l, m)\ )\ )$

Definition: The function $prepend : L \times \mathbb{N} \to L$ that adds an element at the front of a linked list is defined by

$$prepend(\ (l, n)\ ) = (n, l)$$

◀ Back to Q3

Definitions for Q3

Proof by Structural Induction: To prove a universal quantification over a recursively defined set:

Basis Step: Show the statement holds for elements specified in the basis step of the definition.

Recursive Step: Show that if the statement is true for each of the elements used to construct new elements in the recursive step of the definition, the result holds for these new elements.

◀ Back to Q3