CS 611
Advanced Programming Languages

Andrew Myers

Cornell University

Lecture 6: Inductive definitions

6 Sep 00

Administration

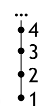
- Homework 1 due September 11

Proofs

- In PL, want to prove various things about inductively defined sets
  - expression termination, equivalence of expressions – abstract syntax
    - **Aexp**, C[ ]
  - equivalence of semantics – legal executions $\langle c, \sigma\rangle \Downarrow \sigma'$, $\langle c, \sigma\rangle \rightarrow^* \langle \mathbf{skip}, \sigma'\rangle$
- What are inductively defined sets, exactly?
- What is the basis for inductive proofs?
  - Winskel: well-founded induction
  - Alternative: induction on proof height

Well-founded induction

- Goal: Prove property P(e) holds for all elements e of a set
- Idea: generalize predecessor relation $\prec$
  - natural numbers: $n \prec n + 1$
  - inductive step: show P(n) & $n \prec n' \Rightarrow$ P(n')



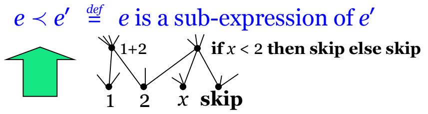
- *Well-founded relation* $\prec$ is any relation with no infinite descending chains
  - must be irreflexive, no cycles

$$\frac{\forall e\,.\,(\forall e' \prec e\,.\,\mathrm{P}(e')) \Rightarrow \mathrm{P}(e)}{\forall e\,.\,\mathrm{P}(e)}$$ *(well-founded induction)*

Structural induction

- Well-founded relation $\prec$:

$e \prec e' \overset{def}{=} e$ is a sub-expression of e'



- To prove $(\forall e' \prec e\,.\,\mathrm{P}(e')) \Rightarrow \mathrm{P}(e)$
  - for expressions e with no predecessors (atoms), prove P(e)
  - for expressions e with ≥1 predecessors e', prove P(e) assuming P(e')

Induction on derivation

- Inductive hypothesis for well-founded structural induction: P(e') for all sub-expressions e'
- Last time, inductive hypothesis slightly stronger: P(e') for all e with shorter AST
  - based on course-of-values induction
  - only alluded to in Winskel Ch. 4
  - caveat: rarely a difference in practice
- How does this work?

Inductive definitions

- Set defined inductively by set of rules (proof system)
- By consistent substitution in agreement with side conditions, rules generate *rule instances* with form $\dfrac{x_1 \ldots x_n}{x}$
- x, x_i are elements of set (no meta-variables)
- *Meaning:* if the elements $x_1 \ldots x_n$ are all members of the set, so is x

Rule operator

- Given a set of elements A assumed to be members of the set being defined, define $R(A)$ to be elements derived by applying all rule instances to A

$$R(A) = \{x \mid \frac{x_1 \ldots x_n}{x} \text{ is a rule inst. \& } \forall_{i \in 1..n}\ x_i \in A\}$$

- $R(\emptyset) = ?$
- $R(R(\emptyset)) = ?$
- $R(A_1 \cup A_2) = ?$

Fixed points

- *Inductively defined set A* is a *fixed point* of rule operator R
- Applying R to A should give us no new elements: $A = R(A)$
- Recall: fixed point of function $f : D \to D$ is $x \in D$ such that $x = f(x)$
- $A = R(A)$ is *equation*, not definition
- Which fixed point of R do we want?

Monotonicity

$$R(A) = \{x \mid \frac{x_1 \ldots x_n}{x} \text{ is a rule inst. \& } \forall_{i \in 1..n}\ x_i \in A\}$$

- If applied to larger set, R yields at least as large a set (monotonic):
  $$A \subseteq B \Rightarrow R(A) \subseteq R(B)$$
- Consider $\emptyset, R(\emptyset), R(R(\emptyset)), R(R(R(\emptyset))), \ldots$
  $= R^0(\emptyset), R^1(\emptyset), R^2(\emptyset), R^3(\emptyset), \ldots$
- By induction: $R^n(\emptyset) \subseteq R^{n+1}(\emptyset)$ for all n
  $R^0(\emptyset) \subseteq R^1(\emptyset) \subseteq R^2(\emptyset) \subseteq R^3(\emptyset), \ldots$

Inductive definition

- The set A defined by the rules is the union of all sets $R^n(\emptyset)$:
  $$A = \bigcup_{n \in \omega} R^n(\emptyset)$$
- A is the *least* fixed point of function R
  - *smallest* set A such that $A = R(A)$
  - finite (but arbitrarily large) number of applications of R
  - elements whose proof trees have finite height

Properties of set

- The least fixed point operator $fix : (D \to D) \to D$
- Why does $\bigcup_{n \in \omega} R^n(\emptyset)$ give $fix\ (R)$?
- First must show $A = R(A)$, *i.e.*
  $$A = \bigcup_{n \in \omega} R^n(\emptyset) = R(\bigcup_{n \in \omega} R^n(\emptyset))$$
- Step 1: $A \supseteq R(A)$
  - if x in $R(A)$, some rule $\dfrac{x_1 \ldots x_m}{x}$ was applied
  - recall $R^n(\emptyset)$ increasing with n
  - must exist n such that $x_1 \ldots x_m$ all in $R^n(\emptyset)$
  - therefore x in $R^{n+1}(\emptyset)$, x in A

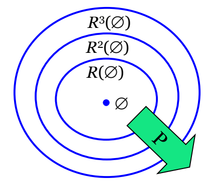
LFP property

$$A = \bigcup_{n\in\omega} R^n(\varnothing)$$

- Step 2: $A \subseteq R(A)$
  - Assume x in A. Then x in $R^n(\varnothing)$ for some n.
  - x in $R(R^{n-1}(\varnothing)) \subseteq R(A)$ (by monotonicity)
  - x in $R(A)$
- Step 3: A is no larger than any fixed pt
  - Suppose $B = R(B)$
  - $R^n(\varnothing) \subseteq B \Rightarrow R^{n+1}(\varnothing) \subseteq R(B) = B$
  - Induction: all $R^n(\varnothing) \subseteq B$, so $A \subseteq B$

Induction on proof height

- Goal: P(e) for all e in set
- Height of derivation of element e is n at which $e \in R^n(\varnothing)$
- Inductive step: prove that P(e) holds for all e in $R^n(\varnothing)$ assuming it holds for all e' in $R^1(\varnothing)...R^{n-1}(\varnothing)$



- *Course-of-values:* $\forall n\,.\,\forall e \in R^n(\varnothing)\,.\,\mathrm{P}(e)$
- *Result:* $\forall e \in \mathit{fix}(R)\,.\,\mathrm{P}(e)$