CS411 Notes 3 – Induction and Recursion

A. Demers

5 Feb 2001

These notes present inductive techniques for defining sets and subsets, for defining functions over sets, and for proving that a property holds for all elements of a set. They do not deal with simultaneous definitions of multiple sets, a straightforward but tedious extension.

1 Formation Rules (Constructors)

1.1 Defining Sets

Here is some formal justification for the technique used to define the *IMP* syntactic sets. The technique gives a collection *FR* of *formation rules* (or *constructors*). Intuitively, think of formation rules as uninterpreted functions that combine small structures into larger ones. Thus, *FR* can be any (finite or infinite) set. With each $f \in FR$ we associate a nonnegative integer k, called the *arity* of f, that specifies the number of substructures associated with f. We represent an "application" of f as a tuple

$$\langle f, c_1, \ldots, c_k \rangle \qquad \text{where } k = \text{ arity}(f)$$

Intuitively, we would like the set defined by our formation rules to consist of exactly those objects that can be constructed by finitely many applications of the rules. This suggests a "formal" definition like

$$C \;=\; \{\langle f, c_1, \ldots, c_k \rangle | f \in FR, k = \text{arity}(f), c_i \in C\}$$

Such a definition is problematic because of the recursive use of "C" inside the definition of C.

A better approach is to rely on our understanding of definition over the natural numbers. That is, we define a sequence C_i of intermediate sets, indexed by the

natural numbers, such that each set is an approximation to the set C that we are trying to define. We then define C to be the "limit" – that is, the infinite union – of all the C_i. Specifically,

$$C_0 \;=\; \{\langle f\rangle | f \in FR, \text{arity}(f) = 0\}$$

$$C_{i+1} \;=\; C_i \cup \{\langle f, c_1, \ldots, c_k\rangle | f \in FR, \text{arity}(f) = k, c_1, \ldots c_k \in C_i\}$$

Here we are implicitly using the property that any set containing 0 and closed under successor contains all the natural numbers, allowing us to conclude that C_i is defined for every i. We then define C by

$$C \;=\; \bigcup_i C_i$$

This countably infinite union certainly exists and is well-defined.

1.2 Proving Properties by Structural Induction

When a set C is defined by formation rules, we can prove properties of its elements by *structural induction* as follows. Suppose we are given some predicate $P(x)$. To prove

$$(\forall c \in C)\; P(c)$$

it suffices to show, for each $f \in FR$,

$$P(c_1) \wedge \ldots \wedge P(c_k) \implies P(< f, c_1, \ldots, c_k\rangle) \tag{1}$$

where as usual $k = \text{arity}(f)$.

The soundness of this technique follows by ordinary mathematical induction. First, from the proven implication 1 and and the definition of C_i we get

$$(\forall c \in C_i . P(c)) \implies (\forall c \in C_{i+1} . P(c))$$

Then, using ordinary mathematical induction we can establish

$$(\forall i)(\forall c \in C_i . P(c))$$

By definition of C, this is exactly

$$(\forall c \in C) P(c)$$

as desired.

1.3 Defining Functions by Structural Recursion

Given a set C defined by formation rules, we would like exploit the structure of those formation rules to help define a function mapping elements of C into some arbitrary range set D. This turns out to be straightforward. For every $f \in FR$, we define a function

$$[\![f]\!] : D^k \to D$$

where $k =$ arity(f). Then there is a unique function F satisfying

$$F(\langle f, c_1, \ldots, c_k \rangle) = [\![f]\!](F(c_1), \ldots, F(c_k))$$

for all $f \in FR$ and all $c_i \in C$. This fact is most naturally proved by two separate structural inductions: one to show the existence of at least one function F with the desired properties, and another to show that all such functions must agree. These are left as exercises.

The uniqueness of F follows from a property called "unique readability" of constructed objects: for any $c \in C$, there is *at most* one rule $f \in FR$ and one sequence $c_1, \ldots, c_k$ such that $c = \langle f, c_1, \ldots, c_k \rangle$. Totality of F follows because there is always *at least* one such rule and sequence (since it is their existence that witnesses the inclusion $\langle f, c_1, \ldots, c_k \rangle$ in C).

The function $[\![f]\!]$ was called $\mathcal{F}_f$ in lecture, rather disastrously, I thought.

2 Well Founded Induction

2.1 Well Founded Relations

A powerful generalization of the structural techniques presented above is based on the notion of a *well-founded relation*.

A binary relation $\prec$ on a set A is called a *well-founded relation* if either of the following two conditions holds:

1. There are no infinite descending chains; i.e.,

$$\cdots \prec a_i \prec \cdots \prec a_1 \prec a_0 \qquad \text{is not possible} \tag{2}$$

   Note this implies $\prec$ is irreflexive, since if $a \prec a$ we could construct an infinite descending chain consisting entirely of as.

2. Any nonempty $Q \subseteq A$ has a minimal element; i.e.,

$$((Q \subseteq A) \wedge (Q \neq \emptyset)) \implies (\exists m \in Q)(\forall b \prec m.b \notin Q) \tag{3}$$

   Since $\prec$ is not necessarily a total order, a minimal element is not necessary a *least* element, and a set may contain numerous minimal elements.

The text includes a proof that the above two characterizations are equivalent. One direction is immediate: if there is an infinite descending chain, that chain itself is a nonempty subset of A without a minimal element. For the other direction, assuming there exists a nonempty subset $Q \subseteq A$ with no minimal element, the text shows how to construct a countably infinite descending chain.

2.2 Proving Properties by WF Induction

Given a well-founded relation $\prec$ on set A, the principle of *well-founded induction* is: to show a property $P(a)$ holds for every $a \in A$ it is enough to show

$$\forall a \in A.([\forall b \prec a.P(b)] \Rightarrow P(a)) \tag{4}$$

that is, if P is true of all predecessors of a then P is true of a.

This principle easy to prove from characterization 3 above. Suppose it fails; that is, suppose implication 4 holds but $P(a)$ is false for some $a \in A$. Let Q be the set of all elements of A on which P is false. Since Q is assumed nonempty, by 3 it must have a minimal element, which we shall call a'. Since $a' \in Q$, it follows that $P(a')$ is false. But by 3 we know that a' has no predecessor in Q; that is,

$$\forall b \prec a'.(b \notin Q)$$

If $b \notin Q$, it follows that $P(b)$ is true. Thus we have

$$\forall b \prec a'.P(b)$$

By implication 4, which is assumed to hold, this implies $P(a')$ is true, a contradiction.

Well-founded induction is a generalization of the other techniques we have seen up to now. In particular,

1. If we take $\prec$ to be the successor relation, $n \prec n+1$, then the principle of well-founded induction specializes to ordinary mathematical induction on the nonnegative integers.

2. If we take $\prec$ to be the strictly-less relation $<$, then the principle of well-founded induction specializes to course-of-values induction on the nonnegative integers.

3. If we take $\prec$ to be the immediate-substructure relation on a set defined by formation rules,

   $$c_i \ \prec \ \langle f, c_1, \ldots, c_i, \ldots, c_k \rangle$$

   then the principle of well-founded induction specializes to the principle of structural induction.

2.3 Defining Functions by WF recursion

Here is a technique for defining functions analogous to well-formed induction.

Suppose we are given a set A and a well-founded relation $\prec$ on A, and we want to define a function $F : A \to B$ where B is arbitrary. Such a function can be defined by *well-founded recursion* as follows. Define a function $\mathcal{F}$ such that $\mathcal{F}(a, h) \in B$ for all $a \; in A$ and all functions $h : \prec^{-1}(a) \to B$. (Note $\prec^{-1}(a)$ is the set of immediate predecessors of a.) Then there is a unique function $F : A \to B$ such that

$$\forall a \in A. F(a) = \mathcal{F}(a, F| \prec^{-1}(a))$$

This fact is proved in the text (p. 177).

It is worth pointing out one important way in which definition by well-founded recursion differs from the structural recursion we have seen earlier: the absence of unique readability. Recall, for a set defined by formation rules, unique readability guarantees that for any $a \in A$ there is a unique formation rule f and a unique sequence $a_1, \ldots, a_k$ of elements of A such that

$$a \ = \ \langle f, a_1, \ldots, a_k \rangle$$

It follows that $\mathcal{F}$ can be defined quite simply using this deconstruction. Unique readability guarantees that the function

$$\mathcal{F}(a, h) \ = \ [\![f]\!](h(a_1), \ldots, h(a_k))$$

is well-defined for all a and h.

In general it is not trivial to come up with a single-valued function $\{(a, h)$ without relying on unique readability.

For example, let A be finite sets of natural numbers, with the $\prec$ relation that would be induced by building up the sets one element at a time:

$$a' \prec a \iff a' \neq a \ \wedge \ \exists n.(a' \cup \{n\} = a)$$

It is tempting to write a "definition" of the form

$$\mathcal{F}(a' \cup \{n\}, h) \ = \ 3^{h(a')} + 2^n$$

(the details of the function on the right hand side are arbitrary) but such a definition implicitly requires unique readability to determine a' unambiguously. Without it, $\mathcal{F}$ may fail to be well-defined: Let h_0 be the constant function 0, and consider the two derivations

$$\mathcal{F}(\{1,2\}, h_0) \ = \ \mathcal{F}(\{1\} \cup \{2\}) \ = \ 1+4 \ = \ 4$$

$$\mathcal{F}(\{2,1\}, h_0) \ = \ \mathcal{F}(\{2\} \cup \{1\}) \ = \ 1+2 \ = \ 3$$

A little care is required.

3 Rules

3.1 Defining Subsets by Rules

Our structural operational semantics for *IMP* is presented as a collection of inference or proof rules. Each rule contains free metavariables that must be filled in (instantiated) consistently, using actual terms or values, before the rule can be applied. Some rules have additional applicability constraints – for example, "where $n = n_0 + n_1$" in rule L1.add of our long step semantics.

The proofs can be seen as defining a subset of a predefined *universe* set U. For example, the large step semantics for *IMP* formally defines a 3-place relation, a subset of $\mathbf{Com} \times \Sigma \times \Sigma$.

Formally, we deal with the process of instantiating proof rules by the simple expedient of doing all possible instantiations beforehand. An *instantiated rule* consists of a pair $\langle X, y \rangle$, where X is a finite (possibly empty) subset of U and y is an element of U. We begin with a set R of instantiated rules. An *R-derivation* of element y is either a rule instance $\langle \emptyset, y \rangle$ or a pair $\langle \{d_1, \ldots, d_k\}, y \rangle$ where $\langle \{x_1, \ldots, x_k\}, y \rangle$ is a rule instance and, for all i, d_i is an R-derivation of x_i. We write $d \vdash_R y$ when d is an R-derivation of y.

There are actually two sets being defined by R in this process:

- a set of R-derivations, for which the definition is analogous to a definition by formation rules, and
- a set of derivable elements of U (elements y that are the conclusion of some R-derivation), which we denote by I_R. This is discussed below.

Both are useful.

3.2 Induction on Derivations

Properties of operational semantics are sometimes provable by induction on the structure of the program commands themselves. When this approach is inadequate, the best approach to try next is generally induction on derivations.

We already have the necessary tools for this. We define a relation $\prec$ on R-derivations as follows: If d is an R-derivation of the form $\langle\{d_1, \ldots, d_k\}, y\rangle$, then we let $d_i \prec d$ for all i. That is, $d_i \prec d$ iff d_i is an immediate subderivation of d. It is easy to verify that $\prec$ is a well-founded relation. Thus, we can use it to support well-founded induction on R-derivations.

3.3 Rule Induction

Given a set R of instantiated rules, the principle of *Rule Induction* enables us to prove properties of the set I_R of R-derivable elements. By now, the principle should appear completely unsurprising:

Let I_R be defined by rule instances R, and let P be a property. To prove $\forall x \in I_R.P(x)$ it suffices to prove

$$(\forall x \in X.P(x)) \implies P(y) \tag{5}$$

for every rule instance $\langle X, y\rangle$ in R such that $X \subseteq I_R$.

A proof of the correctness Rule Induction appears in the text. The argument is fairly straightforward. It relies on the following notion of an *R-closed* set.

Let Q be a subset of U. Q is R-closed if

$$\forall\langle X, y\rangle \in R \;.\; (X \subseteq Q \Rightarrow y \in Q)$$

In words, if Q contains all the premises of a rule instance, then in must contain its conclusion as well.

It is not too difficult to show that I_R is R-closed, and in fact

$$I_R \;=\; \bigcap\{Q|Q\text{is}R-\text{closed}\}$$

that is, I_R is the *smallest possible* R-closed set.

Next, we can show that a proof of implication 5 is enough to guarantee that the set

$$S \;=\; \{x \in I_R \mid P(x)\}$$

is R-closed. Since I_R is the intersection of all R-closed sets, it follows that $I_R \subseteq S$, as desired.