Notes on The Formal Semantics of Programming Languages

Ray Li

October 22, 2024

Contents

1 Chapter 3: some principles of induction

1.1 Material Notes

In the Book Page 35 to 36, the author defined the *rule instances* and *R-derivation*. I found it confusing. So here I put an example and a non example to demenstrate what the concept is talkign about exactly.

Definitions:

A *set of rule instances* R consists of elements which are pairs (X/y) where X is a finite set and y is an element. Such a pair (X/y) is called a *rule instance* with *premises* X and *conclusion* y.

We are more used to seeing rule instances (X/y) as

$$\frac{}{y} \quad \text{if } X = \emptyset, \quad \text{and as} \quad \frac{x_1, \cdots, x_n}{y} \quad \text{if } X = \{x_1, \cdots, x_n\}.$$

Assume a set of rule instances R. An *R-derivation* of y is either a rule instance $(\emptyset/y)$ or a pair $((d_1, \cdots, d_n)/y)$ where $((x_1, \cdots, x_n)/y)$ is a rule instance and d_1 is an R-derivation of $x_1, \cdots, d_n$ is an R-derivation of x_n. We write $d \Vdash_R y$ to mean d is an R-derivation of y. Thus

$$(\emptyset/y) \Vdash_R y \quad \text{if } (\emptyset/y) \in R,$$
$$((d_1, \cdots, d_n)/y) \Vdash_R y \quad \text{if } ((x_1, \cdots, x_n)/y) \in R \;\&\; d_1 \Vdash_R x_1 \;\&\; \cdots \;\&\; d_n \Vdash_R x_n.$$

We say y is derived from R if there is an R-derivation of y, i.e., $d \Vdash_R y$ for some derivation d. We write $\Vdash_R y$ to mean y is derived from R. When the rules are understood we shall write just $d \Vdash y$ and $\Vdash y$.

Example:

We define the set **Even** as follows:

$$\frac{}{0 \in \text{Even}} \quad \text{and} \quad \frac{n \in \text{Even},\ n \in \mathbb{N}}{n+2 \in \text{Even}}$$

Thus, $\text{Even} = \{(\emptyset/0)\} \cup \{(\{n\}, n+2)\}$ following the syntax given in the book. Alternatively, we can also express this as $\text{Even} = \{\langle\emptyset/0\rangle\} \cup \{\langle\{n\}, n+2\rangle\}$, where we use $\langle\rangle$ to denote the pair. According to the definition, $(\emptyset/0)$ is a rule instance because $(\emptyset/0) = \langle\emptyset, 0\rangle \in \text{Even}$, and similarly, $\langle\{8\}, 10\rangle$ is a rule instance since it belongs to **Even**. However, $\langle\{5\}, 7\rangle$ is not a rule instance.

Now, let me present an example of a derivation. The simplest derivation is that of 0, which is $(\emptyset/0) \Vdash 0$ since $(\emptyset/0) \in \text{Even}$. A slightly more complex derivation is the derivation of 2, which is $(\{(\emptyset/0)\}/2) \Vdash 2$. Let's demonstrate why this is valid: using the symbols from the definition, $X = \{(\emptyset/0)\}$, there is only one derivation $d = (\emptyset/0)$ and one premise $x = 0$. This is valid because $(\{0\}/2) \in \text{Even}$, and we already have $d \Vdash 0$ from the previous example where $(\emptyset/0) \Vdash 0$.

Let's now make the example a bit more complicated.

Consider $R' = \text{Even} \cup \{(\{5\}/7)\} = \text{Even} \cup \{\langle\{5\}, 7\rangle\}$. Clearly, $\langle\{5\}, 7\rangle \in R'$, and thus $\langle\{5\}, 7\rangle$ is a rule instance. **However**, there is no derivation of 7. This is a crucial distinction. Let me explain why:

Assume we have $(X/7) \Vdash 7$. Clearly, $X \neq \emptyset$ since $(\emptyset/0) \Vdash 0$. Then, $5 \in X$ (we will not prove here that X is a singleton), and there must be $(\{x\}/7) \in R'$ with $d \Vdash x$. Thus, $x = 5$, but there is nothing that $\Vdash 5$, hence no derivation of 7 exists.

Let's extend this argument further with a more complicated example while still maintaining the same reasoning:

Let $R' = \text{Even} \cup \{(\{n\}/n+2) \mid n = 1, 3, 5, \ldots\}$. By a similar argument, we can show that $\langle\{1\}, 3\rangle$ and $\langle\{3\}, 5\rangle$ are not derivations.

Assume we have $(X/3)$. Then $X \neq \emptyset$ since $(\emptyset/0) \Vdash 0$. If $(\{k\}/3) \in R'$, then $k = 1$. Now, consider $(\{k_2\}/1) \in R'$, but no such k_2 exists (note that k_2 does not exist does not mean $k_2 = \emptyset$). Therefore, such a pair does not exist.

Understanding this example is neccessary for later understanding Chapter 4 on I_R (the smallest set that is closed under R), in the above example, I_R always equals to $Even$.

1.2 Excercise Answers

Question 1. *(E3.2)*
A string is a sequence of symbols. A string $a_1 a_2 \cdots a_n$ with n positions occupied by symbols is said to have length n. A string can be empty in which case it is said to have length 0. Two strings s and t can be concatenated to form the string st. Use mathematical induction to show there is no string u which satisfies $au = ub$ for two distinct symbols a and b.

Proof. by induction on the length of u
Let $A = \{n \mid \text{len}(u) = n \text{ and } au \neq ub\}$.

1. Base Case:
   - $0 \in A$: By the hypothesis, since $a \neq b$, it is clear that when u is the empty string (i.e., length 0), $au \neq ub$. Thus, $0 \in A$.
   - $1 \in A$: If u has length 1, it is easy to see that $au \neq ub$ because $a \neq b$.
2. Inductive Step: Assume $n \in A$. We want to show that $n+1 \in A$ for $n \geq 1$.

   Proof by contradiction: Suppose $\text{len}(u) = n+1$ and $au = ub$. Since $n+1 \geq 2$, we can write $u = qu'$, where $\text{len}(q) = 1$. Thus, $au = aqu'$ and $ub = qu'b$, leading to the equation $aqu' = qu'b$.

   From this, we deduce that $a = q$. By eliminating the leftmost characters, we are left with $qu' = u'b$. Since $q = a$, the equation becomes $au' = u'b$, where $\text{len}(u') = n$. By the induction hypothesis, $au'p \neq u'pb$, leading to a contradiction.

   Therefore, $n + 1 \in A$.

By mathematical induction, we conclude that there is no string u such that $au = ub$ for two distinct symbols a and b.

□

Question 2. *(E3.6)*
What goes wrong when you try to prove the execution of commands is deterministic by using structural induction on commands?

Proof. The issue arises in verifying the **while** rule. By the **Rules for While-loops**, we aim to show that:

$$\langle \textbf{while } b \textbf{ do } c, \sigma\rangle \to \sigma_1 \quad \text{and} \quad \langle \textbf{while } b \textbf{ do } c, \sigma\rangle \to \sigma_2 \implies \sigma_1 = \sigma_2$$

When $\langle b, \sigma\rangle \to \text{false}$, there is no problem. However, the issue emerges when $\langle b, \sigma\rangle \to \text{true}$.

To argue this, we use **Proposition 2.8** (**while** b **do** $c \sim$ **if** b **then** c; w **else skip**). This allows us to simplify the verification to:

$$\langle \textbf{if } b \textbf{ then } c; w \textbf{ else skip}, \sigma\rangle \to \sigma_1 \quad \textbf{and} \quad \langle \textbf{if } b \textbf{ then } c; w \text{ else skip}, \sigma\rangle \to \sigma_2 \implies \sigma_1 = \sigma_2$$

However, if we attempt to use structural induction, we would need to assume something smaller or a predecessor of w. Yet, in each "previous" case, we still need to argue about w itself (rather than a predecessor or a smaller set). Therefore, it is not possible to prove this by structural induction.

□

[Ray's Note 1: Here there exists two kind of structural induction. Or more specifically one is structural induction and the other one is direvation induction. They are some subtle differences between the two. Not to confuse with the proof for Theorem 3.11 in the book. The structural induction here refers to a way of "syntax deviding", like how you proof arithmetic ones from the definition of

$$a ::= n \mid X \mid a_0 + a_1 \mid a_0 - a_1 \mid a_0 \times a_1.$$

by each case (here the case refers to n, a_0, etc.). Yet the direvation induction is refers the different rules (here the case is one kind of rule.).
This is formally mentioned in Page 33.]

Question 3. *(E 3.8) Let $\prec$ be a well-founded relation on a set B. Prove*

1. *its transitive closure $\prec^+$ is also well-founded,*
2. *its reflexive, transitive closure $\prec^*$ is a partial order.*

Proof. 1. The transitive closure $\prec^+$ is well-founded:

proof by way of contradiction (BWOC). Assume there exists an infinite chain:

$$\cdots \prec^+ a_{n+1} \prec^+ a_n \prec^+ \cdots$$

Now, pick an element a_k, and $\exists b \prec^+ a_k$ such that $b \prec a_k$, call such b a_{k+1} . If no such b exists, we would have $\forall b, a_k \prec b$, which contradicts the assumption of an infinite descending chain. By continuing this process, we can pick an infinite chain of $\prec$, which contradicts the well-foundedness of $\prec$. Hence, $\prec^+$ is well-founded.

2. The reflexive, transitive closure $\prec^*$ is a partial order:

By the definition of a partial order, we need to show that the relation is reflexive, antisymmetric, and transitive. Since $\prec^*$ is already reflexive and transitive by definition, we only need to prove antisymmetry.

BWOC. : Assume there exist $a, b \in B$ such that $a \prec^* b$ and $b \prec^* a$, yet $a \neq b$. If $a \neq b$, then it must be the case that either $a \prec^* b$ or $b \prec^* a$, but not both. This leads to a contradiction, hence $a = b$. Therefore, $\prec^*$ is antisymmetric, and thus a partial order.

□

Question 4. *(E 3.9) For a suitable well-founded relation on strings, use the "no counterexample" approach described above to show there is no string u which satisfies $au = ub$ for two distinct symbols a and b. Compare your proof with another by well-founded induction (and with the proof by mathematical induction asked for in Section 3.1).*

Proof. Define $\prec$ as a relation on the set $U \times U$ (the Cartesian product of the set of strings U) such that for two strings a and b, if $\text{len}(a) < \text{len}(b)$, then $a \prec b$.

Claim: $\prec$ is well-founded on the set of strings: we can identify the minimum element, which is the empty string (having length 0).

Thus, we form the set $A = \{u \mid au = ub, \text{ for all } a \neq b\}$. We aim to show that such a set does not exist.

(BWOC): Assume A exists. Since $A \subseteq U$, there exists a smallest element, call it u, such that $au = ub$. By the argument in *E 3.2*, we know that $\text{len}(u) \neq 0, 1$. Therefore, we can write $u = pu'$, where $\text{len}(p) = 1$.

Thus, $apu' = pu'b$, implying $a = p$. This gives us $apu' = au'b$, and by removing the leftmost characters, we get $pu' = u'b$, where $p = a$. Hence, $au' = u'b$, where $\text{len}(u') < \text{len}(u)$, contradicting the assumption that u is the smallest element in A.

From here, we can argue in two ways:

1. Directly conclude that this leads to a contradiction, thus proving the theorem.
2. Alternatively, we can state that no minimum element exists in A, meaning that $\prec$ is not well-founded on A. Hence, the theorem is proven.

□

2 Inductive Definitions

2.1 Material Notes

2.2 Excercise Answers

Question 5. Exercise 4.2
Proposition 4.1 *With respect to rule instances R*

(i) I_R is R-closed, and

(ii) if Q is an R-closed set, then $I_R \subseteq Q$.

Proof:
(i) It is easy to see that I_R is closed under R. Suppose (X/y) is an instance of a rule in R and that $X \subseteq I_R$. Then, from the definition of I_R, there are derivations of each element of X. If X is nonempty, these derivations can be combined with the rule instance (X/y) to provide a derivation of y, and otherwise, $(\emptyset/y)$ provides a derivation immediately. In either case, we obtain a derivation of y, which must therefore be in I_R too. Hence, I_R is closed under R.

(ii) Suppose that Q is R-closed. We want to show that $I_R \subseteq Q$. Any element of I_R is the conclusion of some derivation. But any derivation is built out of rule instances (X/y). If the premises X are in Q, then so is the conclusion y (in particular, the conclusion of any axiom will be in Q). Hence, we can work our way down any derivation, starting at axioms, to show its conclusion is in Q. More formally, we can do an induction on the proper subderivation relation $\prec$ to show

$$\forall y \in I_R.\ d \Vdash_R y \implies y \in Q$$

for all R-derivations d. Therefore, $I_R \subseteq Q$.

Exercise 4.2 *Do the induction on derivations mentioned in the proof above.*

Proof. We prove this question in two parts:

1. Prove that the relation $\prec$ defined above is well-founded.

2. Prove by induction on all derivations.

1. Claim: The relation $\prec$ defined above is well-founded.

Proof: We prove this by mapping every derivation to a natural number (height) and then utilizing the induction property of natural numbers to complete the proof. We define a function h (height) as $h(d)$ where:

- If d is of the form $(\emptyset/x)$, then $h(d) = 0$.
- If $d = (\{d_1, \ldots, d_n\}/x)$, let $D = \{d_1, \ldots, d_n\}$, then $h(d) = \max\{h(d_1), \ldots, h(d_n)\}+1$. The maximum is well-defined because D is finite by definition.

By this construction, it is easy to show that $h(d_1) < h(d_2)$ if and only if $d_1 \prec d_2$. Since $<$ on natural numbers is well-founded, we conclude that $\prec$ is a well-founded relation. By definition of derivations, we exhusted every derivation (h's domain is all derivations, thus h is well-defined). Thus, it is valid to use induction on $\prec$.

2. Want to Show (WTS): For all derivations d, for all $y \in I_R$, if $d \Vdash_R y$, then $y \in Q$.

We can reformulate this as a property of derivations: Let $P(d)$ denote the property $\forall y \in I_R, d \Vdash_R y \implies y \in Q$. We need to show that for all derivations d_i^1 such that $P(d_i^1)$ holds and $d_i^1 \prec d^2$, we have $P(d^2)$, provided that all rule instances of the form of axioms $(\emptyset/x)$ are in I_R and Q (i.e., $x \in I_R$ and $x \in Q$).

Proof: Let $y \in I_R$ be such that $d^2 \Vdash_R y$. By definition and the construction of the derivation, d^2 has the form (D/y), where $D = \{d_1^1, \ldots, d_n^1\}$. Also, by definition, there exists X such that $(X/y) \in R$ and $X = \{x_1, \ldots, x_n\}$.

By the definition of derivation, we know that $d_i^1 \Vdash_R x_i$. By the induction hypothesis, we conclude that $x_i \in Q$, meaning $X \subseteq Q$. Therefore, $y \in Q$.

Thus, we have $P(d)$ for all derivations d, thereby proving the statement.

□

Question 6. Exercise 4.3 *For rule instances R, show*

$$\bigcap\{Q \mid Q \text{ is } R\text{-closed}\}$$

is R-closed. What is this set?

Proof.

$$\bigcap\{Q \mid Q \text{ is } R\text{-closed}\} = I_R.$$

By showing this, it suffices to answer both parts of the question.

Let $K = \{Q \mid Q \text{ is } R\text{-closed}\}$.

1. $\bigcap K \subseteq I_R$:
   Since $I_R \in K$, we have $\bigcap K \subseteq I_R$.

2. $I_R \subseteq \bigcap K$:
   For all $Q \in K$, we know that $I_R \subseteq Q$ because I_R is the smallest set closed under R (proven before). Thus, $I_R \subseteq \bigcap K$.

Since both directions are proven, we conclude that:

$$\bigcap\{Q \mid Q \text{ is } R\text{-closed}\} = I_R.$$

□

Question 7. Exercise 4.4 *Let the rules consist of* $(\emptyset/0)$ *and* $(\{n\}/(n+1))$ *where* n *is a natural number. What is the set defined by the rules and what is rule induction in this case?*

Proof. Natural Number. Mathematical induction. □

3 Ray's Notes Summary

[Ray's Note 1 (Page 4): Here there exists two kind of structural induction. Or more specifically one is structural induction and the other one is direvation induction. They are some subtle differences between the two. Not to confuse with the proof for Theorem 3.11 in the book. The structural induction here refers to a way of "syntax deviding", like how you proof arithmetic ones from the definition of

$$a ::= n \mid X \mid a_0 + a_1 \mid a_0 - a_1 \mid a_0 \times a_1.$$

by each case (here the case refers to n, a_0, etc.). Yet the direvation induction is refers the different rules (here the case is one kind of rule.).
This is formally mentioned in Page 33.]