Notes on The Formal Semantics of Programming Languages

Ray Li

November 14, 2024

Contents

1 Chapter 3: some principles of induction

1.1 Material Notes

1.1.1 3.4 Induction on derivations

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, in fact, $\langle\{5\}, 7\rangle$ is also a rule instance. This might be the confusing part. Clearly we know that both 5 and 7 are not even, but why is $\langle\{5\}, 7\rangle$ a rule instance? Here we need to read the notation carefully.

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

What does this mean? How should we read it? You might think that $\langle\{5\}, 7\rangle$ is not a rule instance since inorder for 7 to be in **Even**, 5 must be in **Even**, then 3 must be in **Even**, but eventually follows that -1 in **Even** but -1 is not in natural number, thus $\langle\{5\}, 7\rangle$ is not a rule instance.

This is how I understand at first. In fact if you go back the git history of this note you can find the following (attached below) **Which is not correct**.

So how should we read it? In fact, we should read it as **if** $n \in \text{Even}$, and $n \in \mathbb{N}$, **then** $n + 3 \in \text{Even}$. Thus, in this sense, $\langle\{5\}, 7\rangle$ is actually a rule instance. But this is not derivable.

The following is quoted from previous version of the note. The part about whether something is a rule instance is not correct, but the part about why they are not derivable is still correct.

> 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 $Even$.'

In later chapter (chapter 4 inductive definitions Page 52), you will see:

> There is another way to view a set defined by rules. A set of rule instances R determines an operator $\widehat{R}$ on sets, which given a set B results in a set
>
> $$\widehat{R}(B) = \{y \mid \exists X \subseteq B.(X/y) \in R\}.$$
>
> Use of the operator $\widehat{R}$ gives another way of saying a set is R-closed.
>
> **proposition 4.11** A set B is closed under R iff $\widehat{R}(B) \subseteq B$.

The **proposition 4.11** states that $\widehat{R}(B) \subseteq B$, but it is ever possible that $\widehat{R}(B) \neq B$, meaning that $\widehat{R}(B)$ is a proper subset of B? The **Even** is an example. I will also put this in the later note.

1.1.2 3.5 Definitions by Induction

In Page 39, I was confused about the following:

> For future reference we define $Loc_L(c)$, the set of those locations which appear on the left of an assignment in a command. For a command c, the function $Loc_L(c)$ is defined by structural induction by taking
>
> $$Loc_L(\textbf{skip}) = \emptyset, \qquad Loc_L(X := a) = \{X\},$$
> $$Loc_L(c_0; c_1) = Loc_L(c_0) \cup Loc_L(c_1), \quad Loc_L(\textbf{if } b \textbf{ then } c_0 \textbf{ else } c_1) = Loc_L(c_0) \cup Loc_L(c_1),$$
> $$Loc_L(\textbf{while } b \textbf{ do } c) = Loc_L(c).$$

I didn't understand what this Loc_L is doing. I am assuming that the name of the function stands for location and the subscript stands for left. And then I was confused on what this location has anything to do with the **Loc**. Is this some sort of location in a abstract computer, like RAM, heap, stack, etc. And this confused me for a long time. I can't solve the later **Excercise 3.13** cuz I don't understand what it means by "the sets of locations which appear in arithmetic expressions...". What is "the sets of locations"?

In chapter 6 Page 82, we also see

> Although this informal explanation will probably suffice, we can give a formal definition using definition by structural induction. Define the set $\mathrm{FV}(a)$ of free variables of arithmetic expressions, extended by integer variables, $a \in$ **Aexpv**, by structural induction
>
> $$\begin{aligned} \mathrm{FV}(n) = \mathrm{FV}(X) &= \emptyset \\ \mathrm{FV}(i) &= \{i\} \\ \mathrm{FV}(a_0 + a_1) = \mathrm{FV}(a_0 - a_1) = \mathrm{FV}(a_0 \times a_1) &= \mathrm{FV}(a_0) \cup \mathrm{FV}(a_1) \end{aligned}$$
>
> for all $n \in \mathbb{N}$, $X \in$ **Loc**, $i \in$ **Intvar**, and $a_0, a_1 \in$ **Aexpv**. Define the free variables $\mathrm{FV}(A)$ of an assertion A by structural induction to be
>
> $$\begin{aligned} \mathrm{FV}(\textbf{true}) = \mathrm{FV}(\textbf{false}) &= \emptyset \\ \mathrm{FV}(a_0 = a_1) = \mathrm{FV}(a_0 \leq a_1) &= \mathrm{FV}(a_0) \cup \mathrm{FV}(a_1) \\ \mathrm{FV}(A_0 \wedge A_1) = \mathrm{FV}(A_0 \vee A_1) = \mathrm{FV}(A_0 \Rightarrow A_1) &= \mathrm{FV}(A_0) \cup \mathrm{FV}(A_1) \\ \mathrm{FV}(\neg A) &= \mathrm{FV}(A) \\ \mathrm{FV}(\forall i.A) = \mathrm{FV}(\exists i.A) &= \mathrm{FV}(A) \setminus \{i\} \end{aligned}$$
>
> for all $a_0, a_1 \in$ **Aexpv**, integer variables i and assertions A_0, A_1, A. Thus we have made precise the notion of free variable. Any variable which occurs in an assertion A and yet is not free is said to be bound. An assertion with no free variables is *closed*.

This is because of the lack of some examples.

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 $A = \{n \mid \mathrm{len}(u) = n \text{ and } au \neq ub\}$.

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 $\mathrm{len}(u) = n+1$ and $au = ub$. Since $n+1 \geq 2$, we can write $u = qu'$, where $\mathrm{len}(q) = 1$. 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 derivation induction refers to 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 Chapter 4: 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.

□

You can also do it with the strick formatted induction. Which is first prove for zero case (empty set), then do induction hypothesis. But essentially just different format.

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 Chapter 5: The denotational semantics of IMP

3.1 Material Notes

3.2 Excercise Answers

Question 8. *(E 5.2)*
Show by structural induction on commands that the denotation $C[\![c]\!]$ is a partial function for all commands c.
(The case for while-loops involves proofs by mathematical induction showing that $\Gamma^n(\emptyset)$ is a partial function between states for all natural numbers n, and that these form an increasing chain, followed by the observation that the union of such a chain of partial functions is itself a partial function.)

Proof. We assume that A and B are partial functions on $Aexp$ and $Bexp$. Since we have the definition for c (shown in the figure), let $P(c) \iff$ if $C[\![c]\!] = F$ and $C[\![c]\!] = G$, then $F = G$.

Case 1: $c \equiv$ **skip**:

$$C[\![c]\!] = \{(\sigma, \sigma) \mid \sigma \in \Sigma\} = \{(\sigma', \sigma') \mid \sigma' \in \Sigma\}$$

Case 2: $c \equiv X := a$:

$$C[\![c]\!] = \{(\sigma, \sigma[n/X]) \mid \sigma \in \Sigma \wedge n = A[\![a]\!]\sigma\} \quad \text{and} \quad C[\![c]\!] = \{(\sigma', \sigma'[n'/X]) \mid \sigma' \in \Sigma \wedge n' = A[\![a]\!]\sigma'\}$$

Since $A[\![a]\!]$ is deterministic, we have $A[\![a]\!] = n = n'$. Since both σ and σ' run over Σ, we have:

$$\{(\sigma, \sigma[n/X]) \mid \sigma \in \Sigma \wedge n = A[\![a]\!]\sigma\} = \{(\sigma', \sigma'[n'/X]) \mid \sigma' \in \Sigma \wedge n' = A[\![a]\!]\sigma'\}$$

Case 3: $c \equiv c_0; c_1$:

$$C[\![c]\!] = C[\![c_1]\!] \circ C[\![c_0]\!] \quad \text{and} \quad C[\![c]\!] = C[\![c_1']\!] \circ C[\![c_0']\!]$$

By definition, we know $c_1 = c_1'$ and $c_0 = c_0'$. Furthermore, by the induction hypotheses, C is a partial function on c_1 and c_0, thus:

$$C[\![c]\!] = C[\![c_1]\!] \circ C[\![c_0]\!] = C[\![c_1']\!] \circ C[\![c_0']\!]$$

Case 4: $c \equiv$ **if** b **then** c_0 **else** c_1:

$$C[\![c]\!] = \{(\sigma, \sigma') \mid B[\![b]\!]\sigma = \text{true} \wedge (\sigma, \sigma') \in C[\![c_0]\!]\} \cup \{(\sigma, \sigma') \mid B[\![b]\!]\sigma = \text{false} \wedge (\sigma, \sigma') \in C[\![c_1]\!]\}$$

$$C[\![c]\!] = \{(\sigma_1, \sigma_1') \mid B[\![b]\!]\sigma_1 = \text{true} \wedge (\sigma_1, \sigma_1') \in C[\![c_0]\!]\} \cup \{(\sigma_1, \sigma_1') \mid B[\![b]\!]\sigma_1 = \text{false} \wedge (\sigma_1, \sigma_1') \in C[\![c_1]\!]\}$$

Since we know, by the induction hypothesis, that C is a function on c_0 and c_1, thus $(\sigma, \sigma') \in C[\![c_0]\!] \iff (\sigma_1, \sigma_1') \in C[\![c_0]\!]$ and similarly for c_1. Thus $C[\![c]\!]$ is deterministic.

Case 5: For $c \equiv$ **while** b **do** c, I will provide a separate proof. □

In Page 52, we have

Proposition 4.11 A set B is closed under R iff $\hat{R}(B) \subseteq B$.

Proof: The fact follows directly from the definitions. □

The operator $\hat{R}$ provides a way of building up the set I_R. The operator $\hat{R}$ is *monotonic* in the sense that

$$A \subseteq B \Rightarrow \hat{R}(A) \subseteq \hat{R}(B).$$

If we repeatedly apply $\hat{R}$ to the empty set $\emptyset$ we obtain the sequence of sets:

$$A_0 = \hat{R}^0(\emptyset) = \emptyset,$$

$$A_1 = \hat{R}^1(\emptyset) = \hat{R}(\emptyset),$$

$$A_2 = \hat{R}(\hat{R}(\emptyset)) = \hat{R}^2(\emptyset),$$

$$\vdots$$

$$A_n = \hat{R}^n(\emptyset),$$

$$\vdots$$

The set A_1 consists of all the conclusions of instances of axioms, and in general the set A_{n+1} is all things which immediately follow by rule instances with premises in A_n. Clearly $\emptyset \subseteq \hat{R}(\emptyset)$, i.e. $A_0 \subseteq A_1$. By the monotonicity of $\hat{R}$ we obtain $\hat{R}(A_0) \subseteq \hat{R}(A_1)$, i.e. $A_1 \subseteq A_2$. Similarly we obtain $A_2 \subseteq A_3$ etc. Thus the sequence forms a chain

$$A_0 \subseteq A_1 \subseteq \cdots \subseteq A_n \subseteq \cdots$$

Taking $A = \bigcup_{n \in \omega} A_n$, we have:

Proposition 4.12

(i) A is R-closed.

(ii) $\hat{R}(A) = A$.

(iii) A is the least R-closed set.

In Page 59, we have

$$\Gamma(\varphi) = \{(\sigma,\sigma') \mid \beta(\sigma) = \text{true} \wedge (\sigma,\sigma') \in \varphi \circ \gamma\} \cup \{(\sigma,\sigma') \mid \beta(\sigma) = \text{false}\}$$
$$= \{(\sigma,\sigma') \mid \exists\sigma''.\beta(\sigma) = \text{true} \wedge (\sigma,\sigma'') \in \gamma \wedge (\sigma'',\sigma') \in \varphi\} \cup \{(\sigma,\sigma') \mid \beta(\sigma) = \text{false}\},$$

as a function which given φ returns $\Gamma(\varphi)$. We want a fixed point φ of Γ in the sense that

$$\varphi = \Gamma(\varphi).$$

The last chapter provides the clue to finding such a solution in Section 4.4. It is not hard to check that the function Γ is equal to $\hat{R}$, where $\hat{R}$ is the operator on sets determined by the rule instances

$$R = \{((\langle\sigma',\sigma''\rangle)/(\sigma,\sigma')) \mid \beta(\sigma) = \text{true} \wedge (\sigma,\sigma'') \in \gamma\} \cup \{(\langle\sigma,\sigma\rangle) \mid \beta(\sigma) = \text{false}\}.$$

As Section 4.4 shows $\hat{R}$ has a least fixed point

$$\varphi = \text{fix}(\hat{R}),$$

where φ is a set—in this case a set of pairs—with the property that

$$\hat{R}(\emptyset) = \emptyset \Rightarrow \varphi \subseteq \theta.$$

It is clear that for any $B \in R = \{((\langle\sigma',\sigma''\rangle)/(\sigma,\sigma')) \mid \beta(\sigma) = \text{true} \wedge (\sigma,\sigma'') \in \gamma\} \cup \{(\langle\sigma,\sigma\rangle) \mid \beta(\sigma) = \text{false}\}$, we have:

$$\Gamma(B) = \hat{R}(B)$$

Claim: $C[\![w]\!]$ is a partial function, where $w \equiv$ **while** b **do** c. To prove this claim, with $\Gamma(B) = \hat{R}(B)$ and the notation in Page 52, we only need to prove that $A_n = \hat{R}(\emptyset) = \Gamma^n(\emptyset), \forall n \in \mathbb{N}$ is a partial function, and that partial function coincide in the union of all A_n.

Proof. Base case: A_0 is a partial function. Since $A_0 = \emptyset$, it is trivially true that A_0 is a function.

Induction hypothesis: Assume A_n is a partial function. We want to show that A_{n+1} is also a partial function, meaning that for all $(\sigma,\sigma') \in A_n$ and $(\sigma,\sigma'_1) \in A_n$, we have $\sigma' = \sigma'_1$. We aim to prove that for all $(\sigma,\sigma') \in A_{n+1}$ and $(\sigma,\sigma'_1) \in A_{n+1}$, it holds that $\sigma' = \sigma'_1$.

Proof:

$$A_{n+1} = \hat{R}^{n+1}(\emptyset) = \hat{R}(\hat{R}^n(\emptyset)) = \{y \mid \exists X \subseteq \hat{R}^n(\emptyset), (X/y) \in R\}.$$

Let (σ,σ'), $(\sigma,\sigma'_1) \in A_{n+1}$.

If $\beta(\sigma) = \text{true}$, then $\exists(\sigma,\sigma'') \in C[\![c]\!]$ and $\exists(\sigma,\sigma''_1) \in C[\![c]\!]$ such that:

$$\{(\sigma'',\sigma')\}/(\sigma,\sigma') \in R \quad \text{and} \quad \{(\sigma''_1,\sigma'_1)\}/(\sigma,\sigma'_1) \in R.$$

This implies that (σ'',σ') and $(\sigma''_1,\sigma'_1) \in A_n$.

By the induction hypothesis, since $C[\![c]\!]$ is a partial function, we have $\sigma'' = \sigma''_1$. Thus, (σ'',σ') and $(\sigma'',\sigma'_1) \in A_n$. By the induction hypothesis, A_n is a partial function, so $\sigma' = \sigma'_1$.

Therefore, for $\beta(\sigma) = \text{true}$, A_{n+1} is a partial function.

If $\beta(\sigma) = \text{false}$, then $\sigma = \sigma' = \sigma'_1$. Thus, A_{n+1} is also a partial function.

We have now proven that for all $n \in \mathbb{N}$, A_n is a partial function. Next, we prove that the union of all A_n is a partial function.

By Proposition 4.11 (attached above, Page 52), we know that:

$$A_0 \subseteq A_1 \subseteq \ldots \subseteq A_n \subseteq \ldots$$

Let $I_R = \phi = \bigcup_{n\in\mathbb{N}} A_n$. If (σ, σ') and $(\sigma, \sigma'_1) \in I_R$, then $\exists n, m \in \mathbb{N}$ such that $(\sigma, \sigma'_1) \in A_n$ and $(\sigma, \sigma') \in A_m$. Without loss of generality, assume $n \geq m$. Thus, $A_m \subseteq A_n$ and $(\sigma, \sigma') \in A_n$. Since A_n is a partial function, we have $\sigma' = \sigma'_1$. Therefore, ϕ is a partial function. □

Question 9. Exercise 5.5 *The proofs above involve considering the form of derivations. Alternative proofs can be obtained by a combination of structural induction and rule induction. For example, show*

1. $\{(\sigma, n) \mid \langle a, \sigma\rangle \to n\} \subseteq A[\![a]\!]$,
2. $A[\![a]\!] \subseteq \{(\sigma, n) \mid \langle a, \sigma\rangle \to n\}$,

for all arithmetic expressions a by using rule induction on the operational semantics of arithmetic expressions for 1 and structural induction on arithmetic expressions for 2.

In Section 4.3.1, Page 46, we have

$$
\begin{aligned}
&\forall a \in \text{Aexpr}, \sigma \in \Sigma, n \in \mathbb{N}. \langle a, \sigma\rangle \to n \implies P(a, \sigma, n)\\
&\text{iff}\\
&[\,[\forall n \in \mathbb{N}, \sigma \in \Sigma.\, P(n, \sigma, n)]\\
&\wedge\\
&[\forall X \in \text{Loc}, \sigma \in \Sigma.\, P(X, \sigma, \sigma(X))]\\
&\wedge\\
&[\forall a_0, a_1 \in \text{Aexpr}, \sigma \in \Sigma, n_0, n_1 \in \mathbb{N}. \langle a_0, \sigma\rangle \to n_0 \wedge P(a_0, \sigma, n_0) \wedge \langle a_1, \sigma\rangle \to n_1 \wedge P(a_1, \sigma, n_1) \implies P(a_0 + a_1, \sigma, n_0 + n_1)]\\
&\wedge\\
&[\forall a_0, a_1 \in \text{Aexpr}, \sigma \in \Sigma, n_0, n_1 \in \mathbb{N}. \langle a_0, \sigma\rangle \to n_0 \wedge P(a_0, \sigma, n_0) \wedge \langle a_1, \sigma\rangle \to n_1 \wedge P(a_1, \sigma, n_1) \implies P(a_0 - a_1, \sigma, n_0 - n_1)]\\
&\wedge\\
&[\forall a_0, a_1 \in \text{Aexpr}, \sigma \in \Sigma, n_0, n_1 \in \mathbb{N}. \langle a_0, \sigma\rangle \to n_0 \wedge P(a_0, \sigma, n_0) \wedge \langle a_1, \sigma\rangle \to n_1 \wedge P(a_1, \sigma, n_1) \implies P(a_0 \times a_1, \sigma, n_0 \times n_1)]]
\end{aligned}
$$

Proof. Let

$$A_\alpha = \{(\alpha, n) \mid \langle \alpha, \sigma\rangle \to n\}$$

and define

$$P(\alpha, \sigma, n) = (\alpha, n) \in A_\alpha \implies (\alpha, n) \in A[\![\alpha]\!],$$

which means that

$$P(\alpha, \sigma, n) \iff \langle \alpha, \sigma\rangle \in A[\![a]\!].$$

From the notation above, we know that this is equivalent to all specified cases.

For

$$\forall m \in \mathbf{N}, \sigma \in \Sigma, \quad (m, \sigma) \to m, \quad A[\![a]\!] = A[\![m]\!] = \{(\sigma, n) \mid \sigma \in \Sigma\},$$

thus

$$(m, \sigma) \in A[\![m]\!].$$

The same applies to the rest of the cases.

To prove the case for

$$\forall a_0, a_1 \in \text{Aexpr}, \sigma \in \Sigma, n_0, n_1, \quad \langle a_0, \sigma\rangle \to n_0 \quad \text{and} \quad P(a_0, \sigma, n_0) \quad \text{and similarly for} \quad a_1 \quad \text{and} \quad n_1,$$

then by the operational rules, we have

$$\langle a_0 + a_1, \sigma \rangle \to n_0 + n_1,$$

yet

$$A[\![a_0 + a_1]\!] = \{(\sigma, n_0 + n_1) \mid (\sigma, n_0) \in A[\![a_0]\!] \quad \text{and} \quad (\sigma, n_1) \in A[\![a_1]\!]\},$$

thus

$$(\sigma, n_0 + n_1) \in A[\![a_0 + a_1]\!],$$

which implies

$$P(a_0 + a_1, \sigma, n_0 + n_1).$$

Formalization of the Other Direction We proceed with induction based on the inductive definition of Aexp:

Case: $a \equiv m$, where $m \in \mathbb{N}$ (note: this $\mathbb{N}$ represents a specific set, not necessarily the natural numbers). We have

$$A[\![m]\!] = \{(\sigma, m) \mid \sigma \in \Sigma\}.$$

If $(\sigma, n) \in A[\![m]\!]$, then $n = m$. From the definition, we know that

$$\langle a, \sigma \rangle \to m \quad \text{if} \quad a \equiv m,$$

thus $(\sigma, n) \in A_\alpha$.

Case: $a \equiv X$. We have

$$A[\![X]\!] = \{(\sigma, \sigma(X)) \mid \sigma \in \Sigma\}.$$

If $(\sigma, \sigma') \in A[\![X]\!]$, then $\sigma' = \sigma(X)$. By the rule for the evaluation of locations,

$$\langle X, \sigma \rangle \to \sigma(X),$$

we know that $(\sigma, \sigma(X)) \in A_a$.

Case: $a \equiv a_0 + a_1$. We have

$$A[\![a_0 + a_1]\!] = \{(\sigma, n_0 + n_1) \mid (\sigma, n_0) \in A[\![a_0]\!] \quad \text{and} \quad (\sigma, n_1) \in A[\![a_1]\!]\}.$$

Thus, $(\sigma, n_0) \in A[\![a_0]\!]$ and $(\sigma, n_1) \in A[\![a_1]\!]$. By the induction hypothesis, we have $(\sigma, n_0) \in A_{a_0}$ and $(\sigma, n_1) \in A_{a_1}$. Therefore,

$$\langle a_0, \sigma \rangle \to n_0 \quad \text{and} \quad \langle a_1, \sigma \rangle \to n_1.$$

According to the rule for evaluation, we have

$$\langle a_0 + a_1, \sigma \rangle \to n_0 + n_1.$$

Thus, $(\sigma, n_0 + n_1) \in A_{a_0 + a_1}$.

The rest of the cases follow in a similar spirit.

□

Question 10. Exercise 5.10 *Show $(\mathcal{P}ow(X), \subseteq)$ is a cpo with bottom, for any set X. Show the set of partial functions $\Sigma \to \Sigma$ ordered by $\subseteq$ forms a cpo with bottom.*

Part1:

Proof. **Claim:** $\subseteq$ is a partial order. Clearly, for any set A, we have $A \subseteq A$, so $\subseteq$ is reflexive. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$, thus $\subseteq$ is transitive. If $A \subseteq B$ and $B \subseteq A$, then $A = B$, making $\subseteq$ antisymmetric.

Claim: $\emptyset$ is the bottom element. For all $A \in \mathcal{P}(X)$, we have $\emptyset \subseteq A \subseteq X$, thus $\emptyset$ is the bottom element.

Claim: For any ascending chain $A_0 \subseteq A_1 \subseteq \ldots \subseteq A_n \subseteq \ldots$, the least upper bound is

$$A = \bigcup_{i \in \omega} A_i.$$

Clearly, for all $i \in \omega$, $A_i \subseteq \bigcup_{i \in \omega} A_i$. If there is another upper bound called B, such that for all $A_i \subseteq B$, we have $\forall a \in A_i$, there exists a set in the chain such that $a \in A_i$. Since B is an upper bound, $A_i \subseteq B$ also. Thus, we have $\forall K \subseteq A$, $K \subseteq B$, implying $A \subseteq B$. Therefore, A is the least upper bound. □

Part 2:

Proof. **Claim:** $\emptyset$ is the bottom of the partial functions. Clearly, $\emptyset$ is a function since it is vacuously true that $\forall(\sigma, \sigma'), (\sigma, \sigma_1') \in \emptyset$, we have $\sigma' = \sigma_1'$ because there are no elements in $\emptyset$.

Claim: $\subseteq$ is a valid partial order.

- **Reflexivity:** Clearly, for any $f \in F$, we have $f \subseteq f$.
- **Transitivity:** For $f, g, h \in F$, if $f \subseteq g$ and $g \subseteq h$, then $f \subseteq h$ by the transitivity of the subset relation.
- **Antisymmetry:** From set theory, we know that $f \subseteq g$ and $g \subseteq f$ if and only if $f = g$. Therefore, $\subseteq$ is antisymmetric.

Claim: For any non-descending chain $f_0 \subseteq f_1 \subseteq \ldots \subseteq f_n \subseteq \ldots$, the least upper bound is given by

$$f = \bigcup_{i \in \omega} f_i.$$

Clearly, for all $i \in \omega$, $f_i \subseteq f$. For any $f' \in f$, and for all $(\sigma, \sigma'), (\sigma, \sigma_1') \in f$, there exist $k, l \in \omega$ such that $(\sigma, \sigma') \in f_k$ and $(\sigma, \sigma_1') \in f_l$. Without loss of generality, assume $k \geq l$. Thus, $f_l \subseteq f_k$, and therefore $(\sigma, \sigma_1') \in f_k$. Since f_k is in the chain, and since f_k is a partial function, we have $\sigma' = \sigma_1'$. Therefore, f is a partial function, meaning $f \in F$. Hence, f is an upper bound.

A similar argument as before shows that f is the least upper bound (lup). □

4 Ray's Notes Summary

[Ray's Note 1 (Page 5): 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 derivation induction refers to different rules (here the case is one kind of rule.).
This is formally mentioned in Page 33.]