

Tobias Nipkow, Gerwin Klein

Concrete Semantics

with Isabelle/HOL

January 21, 2026



Springer-Verlag

I will not allow books to prove anything.

Jane Austen, *Persuasion*

Preface

This book is two books. Part I is a practical introduction to working with the Isabelle proof assistant. It teaches you how to write functional programs and inductive definitions and how to prove properties about them in Isabelle's structured proof language. Part II is an introduction to the semantics of imperative languages with an emphasis on applications like compilers and program analysers. The distinguishing features are that every bit of mathematics has been formalized in Isabelle and that much of it is executable. Part I focusses on the details of proofs in Isabelle. Part II can be read even without familiarity with Isabelle's proof language: all proofs are described in detail but informally. The Isabelle formalization, including the proofs, is available online: all the material, including accompanying slides, can be downloaded from the book's home page <http://www.concrete-semantics.org>.

Although the subject matter is semantics and applications, the not-so-hidden agenda is to teach the reader two things: the art of precise logical reasoning and the practical use of a proof assistant as a surgical tool for formal proofs about computer science artefacts. In this sense the book represents a formal approach to computer science, not just semantics.

Why?

This book is the marriage of two areas: programming languages and theorem proving. Most programmers feel that they understand the programming language they use and the programs they write. Programming language semantics replaces a warm feeling with precision in the form of mathematical definitions of the meaning of programs. Unfortunately such definitions are often still at the level of informal mathematics. They are mental tools, but their informal nature, their size, and the amount of detail makes them error prone. Since they are typically written in \LaTeX , you do not even know whether they

would type-check, let alone whether proofs about the semantics, e.g., compiler correctness, are free of bugs such as missing cases.

This is where theorem proving systems (or “proof assistants”) come in, and mathematical (im)precision is replaced by logical certainty. A proof assistant is a software system that supports the construction of mathematical theories as formal language texts that are checked for correctness. The beauty is that this includes checking the logical correctness of all proof text. No more ‘proofs’ that look more like LSD trips than coherent chains of logical arguments. Machine-checked (aka “formal”) proofs offer the degree of certainty required for reliable software but impossible to achieve with informal methods.

In research, the marriage of programming languages and proof assistants has led to remarkable success stories like a verified C compiler [53] and a verified operating system kernel [47]. This book introduces students and professionals to the foundations and applications of this marriage.

Concrete?

- The book shows that a semantics is not a collection of abstract symbols on sheets of paper but formal text that can be *checked and executed* by the computer: Isabelle is also a programming environment and most of the definitions in the book are executable and can even be exported as programs in a number of (functional) programming languages. For a computer scientist, this is as concrete as it gets.
- Much of the book deals with concrete applications of semantics: compilers, type systems, program analysers.
- The predominant formalism in the book is operational semantics, the most concrete of the various forms of semantics.
- Foundations are made of concrete.

Exercises!

The idea for this book goes back a long way [65]. But only recently have proof assistants become mature enough for inflicting them on students without causing the students too much pain. Nevertheless proof assistants still require very detailed proofs. Learning this proof style (and all the syntactic details that come with any formal language) requires practice. Therefore the book contains a large number of exercises of varying difficulty. If you want to learn Isabelle, you have to work through (some of) the exercises.

A word of warning before you proceed: theorem proving can be addictive!

Acknowledgements

This book has benefited significantly from feedback by John Backes, Harry Butterworth, Dan Dougherty, Andrew Gacek, Florian Haftmann, Peter Johnson, Yutaka Nagashima, Andrei Popescu, René Thiemann, Andrei Sabelfeld, David Sands, Sean Seefried, Helmut Seidl, Christian Sternagel and Carl Witty. Ronan Nugent provided very valuable editorial scrutiny.

The material in this book has been classroom-tested for a number of years. Sascha Böhme, Johannes Hölzl, Alex Krauss, Peter Lammich and Andrei Popescu worked out many of the exercises in the book.

Alex Krauss suggested the title *Concrete Semantics*.

NICTA, Technische Universität München and the DFG Graduiertenkolleg 1480 PUMA supported the writing of this book very generously.

We are very grateful for all these contributions.

Munich
Sydney
October 2014

TN
GK

Contents

Part I Isabelle

1	Introduction	3
2	Programming and Proving	5
2.1	Basics	5
2.2	Types <i>bool</i> , <i>nat</i> and <i>list</i>	7
2.3	Type and Function Definitions	15
2.4	Induction Heuristics	19
2.5	Simplification	21
3	Case Study: IMP Expressions	27
3.1	Arithmetic Expressions	27
3.2	Boolean Expressions	32
3.3	Stack Machine and Compilation	35
4	Logic and Proof Beyond Equality	37
4.1	Formulas	37
4.2	Sets	38
4.3	Proof Automation	39
4.4	Single Step Proofs	42
4.5	Inductive Definitions	45
5	Isar: A Language for Structured Proofs	53
5.1	Isar by Example	54
5.2	Proof Patterns	56
5.3	Streamlining Proofs	58
5.4	Case Analysis and Induction	61

Part II Semantics

6	Introduction	73
7	IMP: A Simple Imperative Language	75
7.1	IMP Commands	75
7.2	Big-Step Semantics	77
7.3	Small-Step Semantics	85
7.4	Summary and Further Reading	90
8	Compiler	95
8.1	Instructions and Stack Machine	95
8.2	Reasoning About Machine Executions	98
8.3	Compilation	99
8.4	Preservation of Semantics	102
8.5	Summary and Further Reading	112
9	Types	115
9.1	Typed IMP	117
9.2	Security Type Systems	128
9.3	Summary and Further Reading	140
10	Program Analysis	143
10.1	Definite Initialization Analysis	145
10.2	Constant Folding and Propagation	154
10.3	Live Variable Analysis	164
10.4	True Liveness	172
10.5	Summary and Further Reading	178
11	Denotational Semantics	179
11.1	A Relational Denotational Semantics	180
11.2	Summary and Further Reading	188
12	Hoare Logic	191
12.1	Proof via Operational Semantics	191
12.2	Hoare Logic for Partial Correctness	192
12.3	Soundness and Completeness	203
12.4	Verification Condition Generation	208
12.5	Hoare Logic for Total Correctness	212
12.6	Summary and Further Reading	215

13 Abstract Interpretation	219
13.1 Informal Introduction	220
13.2 Annotated Commands	224
13.3 Collecting Semantics	225
13.4 Abstract Values	236
13.5 Generic Abstract Interpreter	241
13.6 Executable Abstract States	253
13.7 Analysis of Boolean Expressions	259
13.8 Interval Analysis	264
13.9 Widening and Narrowing	270
13.10 Summary and Further Reading	279
A Auxiliary Definitions	281
B Symbols	283
C Theories	285
References	287
Index	293

Part I

Isabelle

*It's blatantly clear
You stupid machine, that what
I tell you is true*

Michael Norrish

Introduction

Isabelle is a generic system for implementing logical formalisms, and Isabelle/HOL is the specialization of Isabelle for HOL, which abbreviates Higher-Order Logic. We introduce HOL step by step following the equation

$$\text{HOL} = \text{Functional Programming} + \text{Logic}.$$

We assume that the reader is used to logical and set-theoretic notation and is familiar with the basic concepts of functional programming. Open-minded readers have been known to pick up functional programming through the wealth of examples in [Chapter 2](#) and [Chapter 3](#).

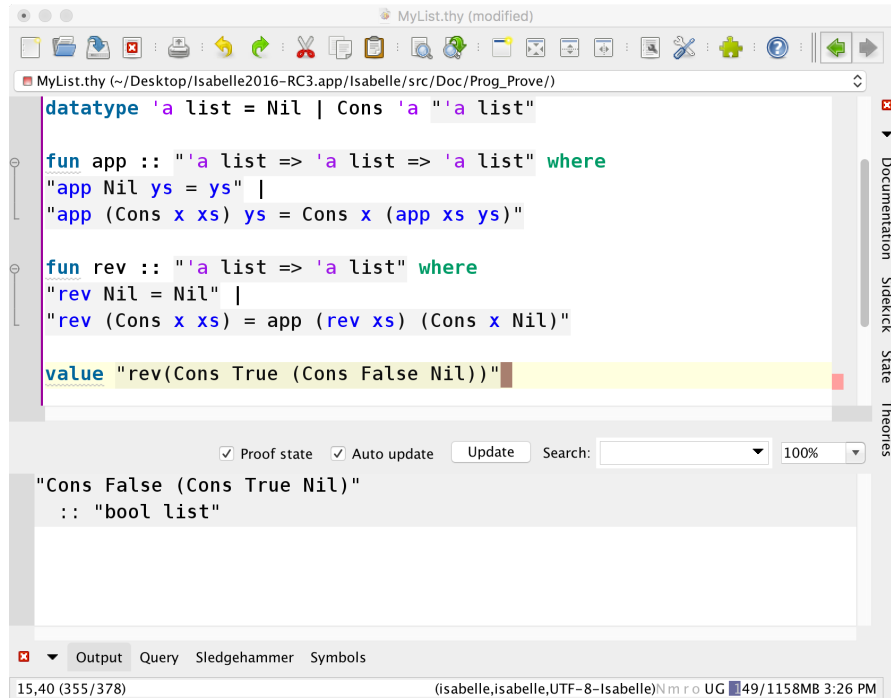
[Chapter 2](#) introduces HOL as a functional programming language and explains how to write simple inductive proofs of mostly equational properties of recursive functions. [Chapter 3](#) contains a small case study: arithmetic and boolean expressions, their evaluation, optimization and compilation. [Chapter 4](#) introduces the rest of HOL: the language of formulas beyond equality, automatic proof tools, single-step proofs, and inductive definitions, an essential specification construct. [Chapter 5](#) introduces Isar, Isabelle’s language for writing structured proofs.

This introduction to the core of Isabelle is intentionally concrete and example-based: we concentrate on examples that illustrate the typical cases without explaining the general case if it can be inferred from the examples. We cover the essentials (from a functional programming point of view) as quickly and compactly as possible. After all, this book is primarily about semantics.

For a comprehensive treatment of all things Isabelle we recommend the *Isabelle/Isar Reference Manual* [92], which comes with the Isabelle distribution. The tutorial by Nipkow, Paulson and Wenzel [68] (in its updated version that comes with the Isabelle distribution) is still recommended for the wealth of examples and material, but its proof style is outdated. In particular it does not cover the structured proof language Isar.

Getting Started with Isabelle

If you have not done so already, download and install Isabelle from <https://isabelle.in.tum.de>. You can start it by clicking on the application icon. This will launch Isabelle’s user interface based on the text editor **jEdit**. Below you see a typical example snapshot of a session. At this point we merely explain the layout of the window, not its contents.



The upper part of the window shows the input typed by the user, i.e., the gradually growing Isabelle text of definitions, theorems, proofs, etc. The interface processes the user input automatically while it is typed, just like modern Java IDEs. Isabelle’s response to the user input is shown in the lower part of the window. You can examine the response to any input phrase by clicking on that phrase or by hovering over underlined text.

! Part I frequently refers to the proof state. You can see the proof state if you press the “State” button. If you want to see the proof state combined with other system output, press “Output” and tick the “Proof state” box.

This should suffice to get started with the jEdit interface. Now you need to learn what to type into it.

Programming and Proving

This chapter introduces HOL as a functional programming language and shows how to prove properties of functional programs by induction.

2.1 Basics

2.1.1 Types, Terms and Formulas

HOL is a typed logic whose type system resembles that of functional programming languages. Thus there are

base types, in particular *bool*, the type of truth values, *nat*, the type of natural numbers (\mathbb{N}), and *int*, the type of mathematical integers (\mathbb{Z}).


type constructors, in particular *list*, the type of lists, and *set*, the type of sets. Type constructors are written postfix, i.e., after their arguments. For example, *nat list* is the type of lists whose elements are natural numbers.

function types, denoted by \Rightarrow .

type variables, denoted by *'a*, *'b*, etc., like in ML.

Note that *'a* \Rightarrow *'b list* means *'a* \Rightarrow (*'b list*), not (*'a* \Rightarrow *'b*) *list*: postfix type constructors have precedence over \Rightarrow .

Terms are formed as in functional programming by applying functions to arguments. If *f* is a function of type $\tau_1 \Rightarrow \tau_2$ and *t* is a term of type τ_1 then *f t* is a term of type τ_2 . We write *t* :: τ to mean that term *t* has type τ .

 There are many predefined infix symbols like + and \leq . The name of the corresponding binary function is (+), not just +. That is, *x* + *y* is nice surface syntax (“syntactic sugar”) for (+) *x* *y*.

HOL also supports some basic constructs from functional programming:

$(if\ b\ then\ t_1\ else\ t_2)$
 $(let\ x = t\ in\ u)$
 $(case\ t\ of\ pat_1 \Rightarrow t_1 \mid \dots \mid pat_n \Rightarrow t_n)$



The above three constructs must always be enclosed in parentheses if they occur inside other constructs.

Terms may also contain λ -abstractions. For example, $\lambda x. x$ is the identity function.

Formulas are terms of type *bool*. There are the basic constants *True* and *False* and the usual logical connectives (in decreasing order of precedence):

$\neg, \wedge, \vee, \longrightarrow$.

Equality is available in the form of the infix function $=$ of type $'a \Rightarrow 'a \Rightarrow bool$. It also works for formulas, where it means “if and only if”.

Quantifiers are written $\forall x. P$ and $\exists x. P$.

Isabelle automatically computes the type of each variable in a term. This is called **type inference**. Despite type inference, it is sometimes necessary to attach an explicit **type constraint** (or **type annotation**) to a variable or term. The syntax is $t :: \tau$ as in $m + (n::nat)$. Type constraints may be needed to disambiguate terms involving overloaded functions such as $+$.

Finally there are the universal quantifier \bigwedge and the implication \Longrightarrow . They are part of the Isabelle framework, not the logic HOL. Logically, they agree with their HOL counterparts \forall and \longrightarrow , but operationally they behave differently. This will become clearer as we go along.



Right-arrows of all kinds always associate to the right. In particular, the formula $A_1 \Longrightarrow A_2 \Longrightarrow A_3$ means $A_1 \Longrightarrow (A_2 \Longrightarrow A_3)$. The (Isabelle-specific¹) notation $\llbracket A_1; \dots; A_n \rrbracket \Longrightarrow A$ is short for the iterated implication $A_1 \Longrightarrow \dots \Longrightarrow A_n \Longrightarrow A$.

Sometimes we also employ inference rule notation:
$$\frac{A_1 \quad \dots \quad A_n}{A}$$

2.1.2 Theories

Roughly speaking, a **theory** is a named collection of types, functions, and theorems, much like a module in a programming language. All Isabelle text needs to go into a theory. The general format of a theory *T* is


```

theory T
imports T1 ... Tn
begin
  definitions, theorems and proofs
end

```

¹ To display implications in this style in Isabelle/jEdit you need to set Plugins > Plugin Options > Isabelle/General > Print Mode to “brackets” and restart.

where $T_1 \dots T_n$ are the names of existing theories that T is based on. The T_i are the direct **parent theories** of T . Everything defined in the parent theories (and their parents, recursively) is automatically visible. Each theory T must reside in a **theory file** named $T.thy$.

 HOL contains a theory *Main*, the union of all the basic predefined theories like arithmetic, lists, sets, etc. Unless you know what you are doing, always include *Main* as a direct or indirect parent of all your theories.

In addition to the theories that come with the Isabelle/HOL distribution (see <https://isabelle.in.tum.de/library/HOL>) there is also the *Archive of Formal Proofs* at <https://isa-afp.org>, a growing collection of Isabelle theories that everybody can contribute to.

2.1.3 Quotation Marks

The textual definition of a theory follows a fixed syntax with keywords like **begin** and **datatype**. Embedded in this syntax are the types and formulas of HOL. To distinguish the two levels, everything HOL-specific (terms and types) must be enclosed in quotation marks: "...". Quotation marks around a single identifier can be dropped. When Isabelle prints a syntax error message, it refers to the HOL syntax as the **inner syntax** and the enclosing theory language as the **outer syntax**.

2.2 Types *bool*, *nat* and *list*

These are the most important predefined types. We go through them one by one. Based on examples we learn how to define (possibly recursive) functions and prove theorems about them by induction and simplification.

2.2.1 Type *bool*

The type of boolean values is a predefined datatype

```
datatype bool = True | False
```

with the two values *True* and *False* and with many predefined functions: \neg , \wedge , \vee , \longrightarrow , etc. Here is how conjunction could be defined by pattern matching:

```
fun conj :: "bool  $\Rightarrow$  bool  $\Rightarrow$  bool" where
  "conj True True = True" |
  "conj _ _ = False"
```

Both the datatype and function definitions roughly follow the syntax of functional programming languages.

2.2.2 Type *nat*

Natural numbers are another predefined datatype:

```
datatype nat = 0 | Suc nat
```

All values of type *nat* are generated by the constructors 0 and *Suc*. Thus the values of type *nat* are 0, *Suc* 0, *Suc* (*Suc* 0), etc. There are many predefined functions: +, *, ≤, etc. Here is how you could define your own addition:

```
fun add :: "nat ⇒ nat ⇒ nat" where
  "add 0 n = n" |
  "add (Suc m) n = Suc(add m n)"
```

And here is a proof of the fact that *add m 0 = m*:

```
lemma add_02: "add m 0 = m"
apply(induction m)
apply(auto)
done
```

The **lemma** command starts the proof and gives the lemma a name, *add_02*. Properties of recursively defined functions need to be established by induction in most cases. Command **apply**(*induction m*) instructs Isabelle to start a proof by induction on *m*. In response, it will show the following proof state²:

1. *add* 0 0 = 0
2. $\bigwedge m. \text{add } m \ 0 = m \implies \text{add } (\text{Suc } m) \ 0 = \text{Suc } m$

The numbered lines are known as *subgoals*. The first subgoal is the base case, the second one the induction step. The prefix $\bigwedge m.$ is Isabelle's way of saying "for an arbitrary but fixed *m*". The \implies separates assumptions from the conclusion. The command **apply**(*auto*) instructs Isabelle to try and prove all subgoals automatically, essentially by simplifying them. Because both subgoals are easy, Isabelle can do it. The base case *add* 0 0 = 0 holds by definition of *add*, and the induction step is almost as simple: *add* (*Suc m*) 0 = *Suc*(*add m* 0) = *Suc m* using first the definition of *add* and then the induction hypothesis. In summary, both subproofs rely on simplification with function definitions and the induction hypothesis. As a result of that final **done**, Isabelle associates the lemma just proved with its name. You can now inspect the lemma with the command

```
thm add_02
```

which displays

```
add ?m 0 = ?m
```

² See page 4 for how to display the proof state.

The free variable m has been replaced by the **unknown** $?m$. There is no logical difference between the two but there is an operational one: unknowns can be instantiated, which is what you want after some lemma has been proved.

Note that there is also a proof method *induct*, which behaves almost like *induction*; the difference is explained in [Chapter 5](#).

! Terminology: We use **lemma**, **theorem** and **rule** interchangeably for propositions that have been proved.

! Numerals (0, 1, 2, ...) and most of the standard arithmetic operations (+, −, *, ≤, <, etc.) are overloaded: they are available not just for natural numbers but for other types as well. For example, given the goal $x + 0 = x$, there is nothing to indicate that you are talking about natural numbers. Hence Isabelle can only infer that x is of some arbitrary type where 0 and + exist. As a consequence, you will be unable to prove the goal. In this particular example, you need to include an explicit type constraint, for example $x+0 = (x::nat)$. If there is enough contextual information this may not be necessary: $Suc\ x = x$ automatically implies $x::nat$ because *Suc* is not overloaded.

An Informal Proof

Above we gave some terse informal explanation of the proof of $add\ m\ 0 = m$. A more detailed informal exposition of the lemma might look like this:

Lemma $add\ m\ 0 = m$

Proof by induction on m .

- Case 0 (the base case): $add\ 0\ 0 = 0$ holds by definition of *add*.
- Case *Suc* m (the induction step): We assume $add\ m\ 0 = m$, the induction hypothesis (IH), and we need to show $add\ (Suc\ m)\ 0 = Suc\ m$. The proof is as follows:

$$\begin{array}{ll} add\ (Suc\ m)\ 0 = Suc\ (add\ m\ 0) & \text{by definition of } add \\ = Suc\ m & \text{by IH} \end{array}$$

Throughout this book, **IH** will stand for “induction hypothesis”.

We have now seen three proofs of $add\ m\ 0 = 0$: the Isabelle one, the terse four lines explaining the base case and the induction step, and just now a model of a traditional inductive proof. The three proofs differ in the level of detail given and the intended reader: the Isabelle proof is for the machine, the informal proofs are for humans. Although this book concentrates on Isabelle proofs, it is important to be able to rephrase those proofs as informal text comprehensible to a reader familiar with traditional mathematical proofs. Later on we will introduce an Isabelle proof language that is closer to traditional informal mathematical language and is often directly readable.

2.2.3 Type *list*

Although lists are already predefined, we define our own copy for demonstration purposes:

```
datatype 'a list = Nil | Cons 'a "'a list"
```

- Type *'a list* is the type of lists over elements of type *'a*. Because *'a* is a type variable, lists are in fact **polymorphic**: the elements of a list can be of arbitrary type (but must all be of the same type).
- Lists have two constructors: *Nil*, the empty list, and *Cons*, which puts an element (of type *'a*) in front of a list (of type *'a list*). Hence all lists are of the form *Nil*, or *Cons x Nil*, or *Cons x (Cons y Nil)*, etc.
- **datatype** requires no quotation marks on the left-hand side, but on the right-hand side each of the argument types of a constructor needs to be enclosed in quotation marks, unless it is just an identifier (e.g., *nat* or *'a*).

We also define two standard functions, *append* and *reverse*:

```
fun app :: "'a list ⇒ 'a list ⇒ 'a list" where
  "app Nil ys = ys" |
  "app (Cons x xs) ys = Cons x (app xs ys)"
```

```
fun rev :: "'a list ⇒ 'a list" where
  "rev Nil = Nil" |
  "rev (Cons x xs) = app (rev xs) (Cons x Nil)"
```

By default, variables *xs*, *ys* and *zs* are of *list* type.

Command **value** evaluates a term. For example,

```
value "rev(Cons True (Cons False Nil))"
```

yields the result *Cons False (Cons True Nil)*. This works symbolically, too:

```
value "rev(Cons a (Cons b Nil))"
```

yields *Cons b (Cons a Nil)*.

Figure 2.1 shows the theory created so far. Because *list*, *Nil*, *Cons*, etc. are already predefined, Isabelle prints qualified (long) names when executing this theory, for example, *MyList.Nil* instead of *Nil*. To suppress the qualified names you can insert the command `declare [[names_short]]`. This is not recommended in general but is convenient for this unusual example.

Structural Induction for Lists

Just as for natural numbers, there is a proof principle of induction for lists. Induction over a list is essentially induction over the length of the list, al-

```

theory MyList
imports Main
begin

datatype 'a list = Nil | Cons 'a "'a list"

fun app :: "'a list => 'a list => 'a list" where
  "app Nil ys = ys" |
  "app (Cons x xs) ys = Cons x (app xs ys)"

fun rev :: "'a list => 'a list" where
  "rev Nil = Nil" |
  "rev (Cons x xs) = app (rev xs) (Cons x Nil)"

value "rev(Cons True (Cons False Nil))"

(* a comment *)

end

```

Fig. 2.1. A theory of lists

though the length remains implicit. To prove that some property P holds for all lists xs , i.e., $P\ xs$, you need to prove

1. the base case $P\ Nil$ and
2. the inductive case $P\ (Cons\ x\ xs)$ under the assumption $P\ xs$, for some arbitrary but fixed x and xs .

This is often called **structural induction** for lists.

2.2.4 The Proof Process

We will now demonstrate the typical proof process, which involves the formulation and proof of auxiliary lemmas. Our goal is to show that reversing a list twice produces the original list.

theorem *rev_rev* [*simp*]: " $rev(rev\ xs) = xs$ "

Commands **theorem** and **lemma** are interchangeable and merely indicate the importance we attach to a proposition. Via the bracketed attribute *simp* we also tell Isabelle to make the eventual theorem a **simplification rule**: future proofs involving simplification will replace occurrences of $rev\ (rev\ xs)$ by xs . The proof is by induction:

apply(*induction xs*)

As explained above, we obtain two subgoals, namely the base case (*Nil*) and the induction step (*Cons*):

1. $rev\ (rev\ Nil) = Nil$
2. $\bigwedge x\ l\ xs.$
 $rev\ (rev\ xs) = xs \implies rev\ (rev\ (Cons\ x\ l\ xs)) = Cons\ x\ l\ xs$

Let us try to solve both goals automatically:

`apply(auto)`

Subgoal 1 is proved, and disappears; the simplified version of subgoal 2 becomes the new subgoal 1:

1. $\bigwedge x\ l\ xs.$
 $rev\ (rev\ xs) = xs \implies$
 $rev\ (app\ (rev\ xs)\ (Cons\ x\ l\ Nil)) = Cons\ x\ l\ xs$

In order to simplify this subgoal further, a lemma suggests itself.

A First Lemma

We insert the following lemma in front of the main theorem:

`lemma rev_app [simp]: "rev(app xs ys) = app (rev ys) (rev xs)"`

There are two variables that we could induct on: *xs* and *ys*. Because *app* is defined by recursion on the first argument, *xs* is the correct one:

`apply(induction xs)`

This time not even the base case is solved automatically:

`apply(auto)`

1. $rev\ ys = app\ (rev\ ys)\ Nil$
- A total of 2 subgoals...

Again, we need to abandon this proof attempt and prove another simple lemma first.

A Second Lemma

We again try the canonical proof procedure:

`lemma app_Nil2 [simp]: "app xs Nil = xs"`

`apply(induction xs)`

`apply(auto)`

`done`

Thankfully, this worked. Now we can continue with our stuck proof attempt of the first lemma:

```
lemma rev_app [simp]: "rev (app xs ys) = app (rev ys) (rev xs)"
apply(induction xs)
apply(auto)
```

We find that this time *auto* solves the base case, but the induction step merely simplifies to

1. $\bigwedge x1\ xs.$

$$\begin{aligned} \text{rev } (\text{app } xs\ ys) &= \text{app } (\text{rev } ys)\ (\text{rev } xs) \implies \\ \text{app } (\text{app } (\text{rev } ys)\ (\text{rev } xs))\ (\text{Cons } x1\ \text{Nil}) &= \\ \text{app } (\text{rev } ys)\ (\text{app } (\text{rev } xs)\ (\text{Cons } x1\ \text{Nil})) \end{aligned}$$

The missing lemma is associativity of *app*, which we insert in front of the failed lemma *rev_app*.

Associativity of *app*

The canonical proof procedure succeeds without further ado:

```
lemma app_assoc [simp]: "app (app xs ys) zs = app xs (app ys zs)"
apply(induction xs)
apply(auto)
done
```

Finally the proofs of *rev_app* and *rev_rev* succeed, too.

Another Informal Proof

Here is the informal proof of associativity of *app* corresponding to the Isabelle proof above.

Lemma $\text{app } (\text{app } xs\ ys)\ zs = \text{app } xs\ (\text{app } ys\ zs)$

Proof by induction on *xs*.

- Case *Nil*: $\text{app } (\text{app } \text{Nil}\ ys)\ zs = \text{app } ys\ zs = \text{app } \text{Nil}\ (\text{app } ys\ zs)$ holds by definition of *app*.
- Case *Cons x xs*: We assume

$$\text{app } (\text{app } xs\ ys)\ zs = \text{app } xs\ (\text{app } ys\ zs) \quad (\text{IH})$$

and we need to show

$$\text{app } (\text{app } (\text{Cons } x\ xs)\ ys)\ zs = \text{app } (\text{Cons } x\ xs)\ (\text{app } ys\ zs).$$

The proof is as follows:

$$\begin{aligned}
& \text{app } (\text{app } (\text{Cons } x \text{ } xs) \text{ } ys) \text{ } zs \\
&= \text{app } (\text{Cons } x \text{ } (\text{app } xs \text{ } ys)) \text{ } zs && \text{by definition of } \text{app} \\
&= \text{Cons } x \text{ } (\text{app } (\text{app } xs \text{ } ys) \text{ } zs) && \text{by definition of } \text{app} \\
&= \text{Cons } x \text{ } (\text{app } xs \text{ } (\text{app } ys \text{ } zs)) && \text{by IH} \\
&= \text{app } (\text{Cons } x \text{ } xs) \text{ } (\text{app } ys \text{ } zs) && \text{by definition of } \text{app}
\end{aligned}$$

Didn't we say earlier that all proofs are by simplification? But in both cases, going from left to right, the last equality step is not a simplification at all! In the base case it is $\text{app } ys \text{ } zs = \text{app } Nil \text{ } (\text{app } ys \text{ } zs)$. It appears almost mysterious because we suddenly complicate the term by appending *Nil* on the left. What is really going on is this: when proving some equality $s = t$, both s and t are simplified until they "meet in the middle". This heuristic for equality proofs works well for a functional programming context like ours. In the base case both $\text{app } (\text{app } Nil \text{ } ys) \text{ } zs$ and $\text{app } Nil \text{ } (\text{app } ys \text{ } zs)$ are simplified to $\text{app } ys \text{ } zs$, the term in the middle.

2.2.5 Predefined Lists

Isabelle's predefined lists are the same as the ones above, but with more syntactic sugar:

- $[]$ is *Nil*,
- $x \# xs$ is *Cons* x xs ,
- $[x_1, \dots, x_n]$ is $x_1 \# \dots \# x_n \# []$, and
- $xs @ ys$ is $\text{app } xs \text{ } ys$.

There is also a large library of predefined functions. The most important ones are the length function $\text{length} :: 'a \text{ list} \Rightarrow \text{nat}$ (with the obvious definition), and the *map* function that applies a function to all elements of a list:

```

fun map :: "('a ⇒ 'b) ⇒ 'a list ⇒ 'b list" where
  "map f Nil = Nil" |
  "map f (Cons x xs) = Cons (f x) (map f xs)"

```

Also useful are the **head** of a list, its first element, and the **tail**, the rest of the list:

```

fun hd :: 'a list ⇒ 'a
  hd (x # xs) = x

fun tl :: 'a list ⇒ 'a list
  tl [] = [] |
  tl (x # xs) = xs

```


Note that since HOL is a logic of total functions, $hd []$ is defined, but we do not know what the result is. That is, $hd []$ is not undefined but underdefined.

From now on lists are always the predefined lists.

Exercises

Exercise 2.1. Use the `value` command to evaluate the following expressions: `"1 + (2::nat)"`, `"1 + (2::int)"`, `"1 - (2::nat)"` and `"1 - (2::int)"`.

Exercise 2.2. Start from the definition of `add` given above. Prove that `add` is associative and commutative. Define a recursive function `double :: nat ⇒ nat` and prove `double m = add m m`.

Exercise 2.3. Define a function `count :: 'a ⇒ 'a list ⇒ nat` that counts the number of occurrences of an element in a list. Prove `count x xs ≤ length xs`.

Exercise 2.4. Define a recursive function `snoc :: 'a list ⇒ 'a ⇒ 'a list` that appends an element to the end of a list. With the help of `snoc` define a recursive function `reverse :: 'a list ⇒ 'a list` that reverses a list. Prove `reverse (reverse xs) = xs`.

Exercise 2.5. Define a recursive function `sum_upto :: nat ⇒ nat` such that `sum_upto n = 0 + ... + n` and prove `sum_upto n = n * (n + 1) div 2`.

2.3 Type and Function Definitions

Type synonyms are abbreviations for existing types, for example

```
type_synonym string = "char list"
```

Type synonyms are expanded after parsing and are not present in internal representation and output. They are mere conveniences for the reader.

2.3.1 Datatypes

The general form of a datatype definition looks like this:

```
datatype ('a1, ..., 'an)t = C1 "τ1,1" ... "τ1,n1"
                        | ...
                        | Ck "τk,1" ... "τk,nk"
```

It introduces the constructors $C_i :: \tau_{i,1} \Rightarrow \dots \Rightarrow \tau_{i,n_i} \Rightarrow ('a_1, \dots, 'a_n)t$ and expresses that any value of this type is built from these constructors in a unique manner. Uniqueness is implied by the following properties of the constructors:

- *Distinctness*: $C_i \dots \neq C_j \dots$ if $i \neq j$
- *Injectivity*: $(C_i x_1 \dots x_{n_i} = C_i y_1 \dots y_{n_i}) = (x_1 = y_1 \wedge \dots \wedge x_{n_i} = y_{n_i})$

The fact that any value of the datatype is built from the constructors implies the **structural induction** rule: to show $P\ x$ for all x of type $('a_1, \dots, 'a_n)t$, one needs to show $P(C_i x_1 \dots x_{n_i})$ (for each i) assuming $P(x_j)$ for all j where $\tau_{i,j} = ('a_1, \dots, 'a_n)t$. Distinctness and injectivity are applied automatically by *auto* and other proof methods. Induction must be applied explicitly.

Like in functional programming languages, datatype values can be taken apart with case expressions, for example

```
(case xs of []  $\Rightarrow$  0 | x # _  $\Rightarrow$  Suc x)
```

Case expressions must be enclosed in parentheses.

As an example of a datatype beyond *nat* and *list*, consider binary trees:

```
datatype 'a tree = Tip | Node "'a tree" 'a "'a tree"
```

with a mirror function:

```
fun mirror :: "'a tree  $\Rightarrow$  'a tree" where
  "mirror Tip = Tip" |
  "mirror (Node l a r) = Node (mirror r) a (mirror l)"
```

The following lemma illustrates induction:

```
lemma "mirror(mirror t) = t"
apply(induction t)
```

yields

1. $\text{mirror}(\text{mirror Tip}) = \text{Tip}$
2. $\bigwedge t1\ x2\ t2.$

$$\llbracket \text{mirror}(\text{mirror } t1) = t1; \text{mirror}(\text{mirror } t2) = t2 \rrbracket$$

$$\implies \text{mirror}(\text{mirror}(\text{Node } t1\ x2\ t2)) = \text{Node } t1\ x2\ t2$$

The induction step contains two induction hypotheses, one for each subtree. An application of *auto* finishes the proof.

A very simple but also very useful datatype is the predefined

```
datatype 'a option = None | Some 'a
```

Its sole purpose is to add a new element *None* to an existing type *'a*. To make sure that *None* is distinct from all the elements of *'a*, you wrap them up in *Some* and call the new type *'a option*. A typical application is a lookup function on a list of key-value pairs, often called an association list:

```
fun lookup :: "('a * 'b) list  $\Rightarrow$  'a  $\Rightarrow$  'b option" where
```

```
"lookup [] x = None" |
"lookup ((a,b) # ps) x = (if a = x then Some b else lookup ps x) "
```

Note that $\tau_1 * \tau_2$ is the type of pairs, also written $\tau_1 \times \tau_2$. Pairs can be taken apart either by pattern matching (as above) or with the projection functions *fst* and *snd*: *fst* $(x, y) = x$ and *snd* $(x, y) = y$. Tuples are simulated by pairs nested to the right: (a, b, c) is short for $(a, (b, c))$ and $\tau_1 \times \tau_2 \times \tau_3$ is short for $\tau_1 \times (\tau_2 \times \tau_3)$.

2.3.2 Definitions

Non-recursive functions can be defined as in the following example:

```
definition sq :: "nat  $\Rightarrow$  nat" where
"sq n = n * n"
```

Such definitions do not allow pattern matching but only $f\ x_1 \dots x_n = t$, where f does not occur in t .

2.3.3 Abbreviations

Abbreviations are similar to definitions:

```
abbreviation sq' :: "nat  $\Rightarrow$  nat" where
"sq' n  $\equiv$  n * n"
```

The key difference is that *sq'* is only syntactic sugar: after parsing, *sq' t* is replaced by $t * t$; before printing, every occurrence of $u * u$ is replaced by *sq' u*. Internally, *sq'* does not exist. This is the advantage of abbreviations over definitions: definitions need to be expanded explicitly (Section 2.5.5) whereas abbreviations are already expanded upon parsing. However, abbreviations should be introduced sparingly: if abused, they can lead to a confusing discrepancy between the internal and external view of a term.

The ASCII representation of \equiv is == or \<equiv>.

2.3.4 Recursive Functions

Recursive functions are defined with **fun** by pattern matching over datatype constructors. The order of equations matters, as in functional programming languages. However, all HOL functions must be total. This simplifies the logic — terms are always defined — but means that recursive functions must terminate. Otherwise one could define a function $f\ n = f\ n + 1$ and conclude $0 = 1$ by subtracting $f\ n$ on both sides.

Isabelle's automatic termination checker requires that the arguments of recursive calls on the right-hand side must be strictly smaller than the arguments on the left-hand side. In the simplest case, this means that one fixed argument position decreases in size with each recursive call. The size is measured as the number of constructors (excluding 0-ary ones, e.g., *Nil*). Lexicographic combinations are also recognized. In more complicated situations, the user may have to prove termination by hand. For details see [49].

Functions defined with **fun** come with their own induction schema that mirrors the recursion schema and is derived from the termination order. For example,

```
fun div2 :: "nat ⇒ nat" where
  "div2 0 = 0" |
  "div2 (Suc 0) = 0" |
  "div2 (Suc (Suc n)) = Suc (div2 n)"
```

does not just define *div2* but also proves a customized induction rule:

$$\frac{P\ 0 \quad P\ (Suc\ 0) \quad \bigwedge n. P\ n \implies P\ (Suc\ (Suc\ n))}{P\ m}$$

This customized induction rule can simplify inductive proofs. For example,

```
lemma "div2 n = n div 2"
apply(induction n rule: div2.induct)
```

(where the infix *div* is the predefined division operation) yields the subgoals

1. *div2* 0 = 0 *div* 2
2. *div2* (Suc 0) = Suc 0 *div* 2
3. $\bigwedge n. \text{div2 } n = n \text{ div } 2 \implies$
 $\text{div2 } (Suc\ (Suc\ n)) = Suc\ (Suc\ n) \text{ div } 2$

An application of *auto* finishes the proof. Had we used ordinary structural induction on *n*, the proof would have needed an additional case analysis in the induction step.

This example leads to the following induction heuristic:

Let f be a recursive function. If the definition of f is more complicated than having one equation for each constructor of some datatype, then properties of f are best proved via $f.induct$.

The general case is often called **computation induction**, because the induction follows the (terminating!) computation. For every defining equation

$$f(e) = \dots f(r_1) \dots f(r_k) \dots$$

where $f(r_i)$, $i=1\dots k$, are all the recursive calls, the induction rule $f.induct$ contains one premise of the form

$$P(r_1) \implies \dots \implies P(r_k) \implies P(e)$$

If $f :: \tau_1 \Rightarrow \dots \Rightarrow \tau_n \Rightarrow \tau$ then $f.induct$ is applied like this:

`apply(induction $x_1 \dots x_n$ rule: $f.induct$)`

where typically there is a call $f\ x_1 \dots x_n$ in the goal. But note that the induction rule does not mention f at all, except in its name, and is applicable independently of f .

Exercises

Exercise 2.6. Starting from the type `'a tree` defined in the text, define a function `contents :: 'a tree \Rightarrow 'a list` that collects all values in a tree in a list, in any order, without removing duplicates. Then define a function `sum_tree :: nat tree \Rightarrow nat` that sums up all values in a tree of natural numbers and prove `sum_tree t = sum_list (contents t)` where `sum_list` is predefined by the equations `sum_list [] = 0` and `sum_list (x # xs) = x + sum_list xs`.

Exercise 2.7. Define the two functions `pre_order` and `post_order` of type `'a tree \Rightarrow 'a list` that traverse a tree and collect all stored values in the respective order in a list. Prove `pre_order (mirror t) = rev (post_order t)`.

Exercise 2.8. Define a function `intersperse :: 'a \Rightarrow 'a list \Rightarrow 'a list` such that `intersperse a [x1, ..., xn] = [x1, a, x2, a, ..., a, xn]`. Now prove that `map f (intersperse a xs) = intersperse (f a) (map f xs)`.

2.4 Induction Heuristics

We have already noted that theorems about recursive functions are proved by induction. In case the function has more than one argument, we have followed the following heuristic in the proofs about the `append` function:

*Perform induction on argument number i
if the function is defined by recursion on argument number i.*

The key heuristic, and the main point of this section, is to *generalize the goal before induction*. The reason is simple: if the goal is too specific, the induction hypothesis is too weak to allow the induction step to go through. Let us illustrate the idea with an example.

Function `rev` has quadratic worst-case running time because it calls `append` for each element of the list and `append` is linear in its first argument.

A linear time version of *rev* requires an extra argument where the result is accumulated gradually, using only #:

```
fun itrev :: "'a list ⇒ 'a list ⇒ 'a list" where
  "itrev [] ys = ys" |
  "itrev (x#xs) ys = itrev xs (x#ys)"
```

The behaviour of *itrev* is simple: it reverses its first argument by stacking its elements onto the second argument, and it returns that second argument when the first one becomes empty. Note that *itrev* is tail-recursive: it can be compiled into a loop; no stack is necessary for executing it.

Naturally, we would like to show that *itrev* reverses its first argument:

```
lemma "itrev xs [] = rev xs"
```

There is no choice as to the induction variable:

```
apply(induction xs)
apply(auto)
```

Unfortunately, this attempt does not prove the induction step:

1. $\bigwedge a \text{ xs. } \text{itrev xs []} = \text{rev xs} \implies \text{itrev xs [a]} = \text{rev xs @ [a]}$

The induction hypothesis is too weak. The fixed argument, [], prevents it from rewriting the conclusion. This example suggests a heuristic:

Generalize goals for induction by replacing constants by variables.

Of course one cannot do this naively: *itrev xs ys = rev xs* is just not true. The correct generalization is

```
lemma "itrev xs ys = rev xs @ ys"
```

If *ys* is replaced by [], the right-hand side simplifies to *rev xs*, as required. In this instance it was easy to guess the right generalization. Other situations can require a good deal of creativity.

Although we now have two variables, only *xs* is suitable for induction, and we repeat our proof attempt. Unfortunately, we are still not there:

1. $\bigwedge a \text{ xs. } \text{itrev xs ys} = \text{rev xs @ ys} \implies \text{itrev xs (a \# ys)} = \text{rev xs @ a \# ys}$

The induction hypothesis is still too weak, but this time it takes no intuition to generalize: the problem is that the *ys* in the induction hypothesis is fixed, but the induction hypothesis needs to be applied with *a # ys* instead of *ys*. Hence we prove the theorem for all *ys* instead of a fixed one. We can instruct induction to perform this generalization for us by adding *arbitrary: ys*.

`apply(induction xs arbitrary: ys)`

The induction hypothesis in the induction step is now universally quantified over ys :

1. $\bigwedge ys. \text{itrev } [] \text{ } ys = \text{rev } [] \text{ } @ \text{ } ys$
2. $\bigwedge a \text{ } xs \text{ } ys.$
 $(\bigwedge ys. \text{itrev } xs \text{ } ys = \text{rev } xs \text{ } @ \text{ } ys) \implies$
 $\text{itrev } (a \text{ } \# \text{ } xs) \text{ } ys = \text{rev } (a \text{ } \# \text{ } xs) \text{ } @ \text{ } ys$

Thus the proof succeeds:

`apply(auto)`
`done`

This leads to another heuristic for generalization:

Generalize induction by generalizing all free variables
 (except the induction variable itself).

Generalization is best performed with *arbitrary*: $y_1 \dots y_k$. This heuristic prevents trivial failures like the one above. However, it should not be applied blindly. It is not always required, and the additional quantifiers can complicate matters in some cases. The variables that need to be quantified are typically those that change in recursive calls.

Exercises

Exercise 2.9. Write a tail-recursive variant of the *add* function on *nat*: *itadd*. Tail-recursive means that in the recursive case, *itadd* needs to call itself directly: $\text{itadd } (\text{Suc } m) \text{ } n = \text{itadd } \dots$. Prove $\text{itadd } m \text{ } n = \text{add } m \text{ } n$.

2.5 Simplification

So far we have talked a lot about simplifying terms without explaining the concept. **Simplification** means

- using equations $l = r$ from left to right (only),
- as long as possible.

To emphasize the directionality, equations that have been given the *simp* attribute are called **simplification rules**. Logically, they are still symmetric, but proofs by simplification use them only in the left-to-right direction. The proof tool that performs simplifications is called the **simplifier**. It is the basis of *auto* and other related proof methods.

The idea of simplification is best explained by an example. Given the simplification rules

$$0 + n = n \quad (1)$$

$$\text{Suc } m + n = \text{Suc } (m + n) \quad (2)$$

$$(\text{Suc } m \leq \text{Suc } n) = (m \leq n) \quad (3)$$

$$(0 \leq m) = \text{True} \quad (4)$$

the formula $0 + \text{Suc } 0 \leq \text{Suc } 0 + x$ is simplified to *True* as follows:

$$(0 + \text{Suc } 0 \leq \text{Suc } 0 + x) \quad \stackrel{(1)}{=}$$

$$(\text{Suc } 0 \leq \text{Suc } 0 + x) \quad \stackrel{(2)}{=}$$

$$(\text{Suc } 0 \leq \text{Suc } (0 + x)) \quad \stackrel{(3)}{=}$$

$$(0 \leq 0 + x) \quad \stackrel{(4)}{=}$$

$$\text{True}$$

Simplification is often also called **rewriting** and simplification rules **rewrite rules**.

2.5.1 Simplification Rules

The attribute *simp* declares theorems to be simplification rules, which the simplifier will use automatically. In addition, **datatype** and **fun** commands implicitly declare some simplification rules: **datatype** the distinctness and injectivity rules, **fun** the defining equations. Definitions are not declared as simplification rules automatically! Nearly any theorem can become a simplification rule. The simplifier will try to transform it into an equation. For example, the theorem $\neg P$ is turned into $P = \text{False}$.

Only equations that really simplify, like $\text{rev } (\text{rev } xs) = xs$ and $xs @ [] = xs$, should be declared as simplification rules. Equations that may be counterproductive as simplification rules should only be used in specific proof steps (see [Section 2.5.4](#) below). Distributivity laws, for example, alter the structure of terms and can produce an exponential blow-up.

2.5.2 Conditional Simplification Rules

Simplification rules can be conditional. Before applying such a rule, the simplifier will first try to prove the preconditions, again by simplification. For example, given the simplification rules

$$p \ 0 = \text{True}$$

$$p \ x \Longrightarrow f \ x = g \ x,$$

the term $f \ 0$ simplifies to $g \ 0$ but $f \ 1$ does not simplify because $p \ 1$ is not provable.

2.5.3 Termination

Simplification can run forever, for example if both $f\ x = g\ x$ and $g\ x = f\ x$ are simplification rules. It is the user's responsibility not to include simplification rules that can lead to nontermination, either on their own or in combination with other simplification rules. The right-hand side of a simplification rule should always be "simpler" than the left-hand side — in some sense. But since termination is undecidable, such a check cannot be automated completely and Isabelle makes little attempt to detect nontermination.

When conditional simplification rules are applied, their preconditions are proved first. Hence all preconditions need to be simpler than the left-hand side of the conclusion. For example

$$n < m \implies (n < \text{Suc } m) = \text{True}$$

is suitable as a simplification rule: both $n < m$ and True are simpler than $n < \text{Suc } m$. But

$$\text{Suc } n < m \implies (n < m) = \text{True}$$

leads to nontermination: when trying to rewrite $n < m$ to True one first has to prove $\text{Suc } n < m$, which can be rewritten to True provided $\text{Suc } (\text{Suc } n) < m$, *ad infinitum*.

2.5.4 The *simp* Proof Method

So far we have only used the proof method *auto*. Method *simp* is the key component of *auto*, but *auto* can do much more. In some cases, *auto* is overeager and modifies the proof state too much. In such cases the more predictable *simp* method should be used. Given a goal

$$1. \llbracket P_1; \dots; P_m \rrbracket \implies C$$

the command

```
apply(simp add: th1 ... thn)
```

simplifies the assumptions P_i and the conclusion C using

- all simplification rules, including the ones coming from **datatype** and **fun**,
- the additional lemmas $th_1 \dots th_n$, and
- the assumptions.

In addition to or instead of *add* there is also *del* for removing simplification rules temporarily. Both are optional. Method *auto* can be modified similarly:

```
apply(auto simp add: ... simp del: ...)
```

Here the modifiers are *simp add* and *simp del* instead of just *add* and *del* because *auto* does not just perform simplification.

Note that *simp* acts only on subgoal 1, *auto* acts on all subgoals. There is also *simp_all*, which applies *simp* to all subgoals.

2.5.5 Rewriting with Definitions

Definitions introduced by the command **definition** can also be used as simplification rules, but by default they are not: the simplifier does not expand them automatically. Definitions are intended for introducing abstract concepts and not merely as abbreviations. Of course, we need to expand the definition initially, but once we have proved enough abstract properties of the new constant, we can forget its original definition. This style makes proofs more robust: if the definition has to be changed, only the proofs of the abstract properties will be affected.

The definition of a function *f* is a theorem named *f_def* and can be added to a call of *simp* like any other theorem:

```
apply(simp add: f_def)
```

In particular, let-expressions can be unfolded by making *Let_def* a simplification rule.

2.5.6 Case Splitting With *simp*

Goals containing if-expressions are automatically split into two cases by *simp* using the rule

$$P \text{ (if } A \text{ then } s \text{ else } t) = ((A \longrightarrow P \ s) \wedge (\neg A \longrightarrow P \ t))$$

For example, *simp* can prove

$$(A \wedge B) = (\text{if } A \text{ then } B \text{ else } \text{False})$$

because both $A \longrightarrow (A \wedge B) = B$ and $\neg A \longrightarrow (A \wedge B) = \text{False}$ simplify to *True*.

We can split case-expressions similarly. For *nat* the rule looks like this:

$$P \text{ (case } e \text{ of } 0 \Rightarrow a \mid \text{Suc } n \Rightarrow b \ n) = \\ ((e = 0 \longrightarrow P \ a) \wedge (\forall n. e = \text{Suc } n \longrightarrow P \ (b \ n)))$$

Case expressions are not split automatically by *simp*, but *simp* can be instructed to do so:

```
apply(simp split: nat.split)
```

splits all case-expressions over natural numbers. For an arbitrary datatype t it is $t.split$ instead of $nat.split$. Method *auto* can be modified in exactly the same way. The modifier *split*: can be followed by multiple names. Splitting if or case-expressions in the assumptions requires *split*: *if_splits* or *split*: *t.splits*.

Exercises

Exercise 2.10. Define a datatype *tree0* of binary tree skeletons which do not store any information, neither in the inner nodes nor in the leaves. Define a function *nodes* :: *tree0* \Rightarrow *nat* that counts the number of all nodes (inner nodes and leaves) in such a tree. Consider the following recursive function:

```
fun explode :: "nat  $\Rightarrow$  tree0  $\Rightarrow$  tree0" where
  "explode 0 t = t" |
  "explode (Suc n) t = explode n (Node t t)"
```

Find an equation expressing the size of a tree after exploding it (*nodes* (*explode* n t)) as a function of *nodes* t and n . Prove your equation. You may use the usual arithmetic operators, including the exponentiation operator “ \wedge ”. For example, $2 \wedge 2 = 4$.

Hint: simplifying with the list of theorems *algebra_simps* takes care of common algebraic properties of the arithmetic operators.

Exercise 2.11. Define arithmetic expressions in one variable over integers (type *int*) as a data type:

```
datatype exp = Var | Const int | Add exp exp | Mult exp exp
```

Define a function *eval* :: *exp* \Rightarrow *int* \Rightarrow *int* such that *eval* e x evaluates e at the value x .

A polynomial can be represented as a list of coefficients, starting with the constant. For example, $[4, 2, -1, 3]$ represents the polynomial $4 + 2x - x^2 + 3x^3$. Define a function *evalp* :: *int list* \Rightarrow *int* \Rightarrow *int* that evaluates a polynomial at the given value. Define a function *coeffs* :: *exp* \Rightarrow *int list* that transforms an expression into a polynomial. This may require auxiliary functions. Prove that *coeffs* preserves the value of the expression: *evalp* (*coeffs* e) $x = eval$ e x . Hint: consider the hint in Exercise 2.10.

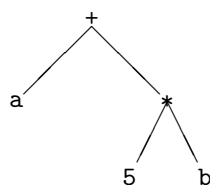
Case Study: IMP Expressions

The methods of the previous chapter suffice to define the arithmetic and boolean expressions of the programming language IMP that is the subject of this book. In this chapter we define their syntax and semantics, write little optimizers for them and show how to compile arithmetic expressions to a simple stack machine. Of course we also prove the correctness of the optimizers and compiler!

3.1 Arithmetic Expressions thy

3.1.1 Syntax

Programming languages have both a concrete and an abstract syntax. **Concrete syntax** means strings. For example, "a + 5 * b" is an arithmetic expression given as a string. The concrete syntax of a language is usually defined by a context-free grammar. The expression "a + 5 * b" can also be viewed as the following tree:



The tree immediately reveals the nested structure of the object and is the right level for analysing and manipulating expressions. Linear strings are more compact than two-dimensional trees, which is why they are used for reading and writing programs. But the first thing a compiler, or rather its parser, will do is to convert the string into a tree for further processing. Now we

are at the level of **abstract syntax** and these trees are **abstract syntax trees**. To regain the advantages of the linear string notation we write our abstract syntax trees as strings with parentheses to indicate the nesting (and with identifiers instead of the symbols $+$ and $*$), for example like this: *Plus a (Times 5 b)*. Now we have arrived at ordinary terms as we have used them all along. More precisely, these terms are over some datatype that defines the abstract syntax of the language. Our little language of arithmetic expressions is defined by the datatype *aexp*:

```
type_synonym vname = string
datatype aexp = N int | V vname | Plus aexp aexp
```

where *int* is the predefined type of integers and *vname* stands for variable name. Isabelle strings require two single quotes on both ends, for example *'abc'*. The intended meaning of the three constructors is as follows: *N* represents numbers, i.e., constants, *V* represents variables, and *Plus* represents addition. The following examples illustrate the intended correspondence:

Concrete	Abstract
5	<i>N 5</i>
x	<i>V 'x'</i>
x + y	<i>Plus (V 'x') (V 'y')</i>
2 + (z + 3)	<i>Plus (N 2) (Plus (V 'z') (N 3))</i>

It is important to understand that so far we have only defined syntax, not semantics! Although the binary operation is called *Plus*, this is merely a suggestive name and does not imply that it behaves like addition. For example, *Plus (N 0) (N 0) \neq N 0*, although you may think of them as semantically equivalent — but syntactically they are not.

Datatype *aexp* is intentionally minimal to let us concentrate on the essentials. Further operators can be added as desired. However, as we shall discuss below, not all operators are as well behaved as addition.

3.1.2 Semantics

The semantics, or meaning of an expression, is its value. But what is the value of $x+1$? The value of an expression with variables depends on the values of its variables. The value of all variables is recorded in the (program) **state**. The state is a function from variable names to values.

```
type_synonym val = int
type_synonym state = vname  $\Rightarrow$  val
```

In our little toy language, the only values are integers.

The value of an arithmetic expression is computed like this:

```

fun aval :: "aexp  $\Rightarrow$  state  $\Rightarrow$  val" where
  "aval (N n) s = n" |
  "aval (V x) s = s x" |
  "aval (Plus a1 a2) s = aval a1 s + aval a2 s"

```

Function *aval* carries around a state and is defined by recursion over the form of the expression. Numbers evaluate to themselves, variables to their value in the state, and addition is evaluated recursively. Here is a simple example:

```
value "aval (Plus (N 3) (V 'x')) (λx. 0)"
```

returns 3. However, we would like to be able to write down more interesting states than $\lambda x. 0$ easily. This is where function update comes in.

To update the state, that is, change the value of some variable name, the generic function update notation $f(a := b)$ is used: the result is the same as f , except that it maps a to b :

$$f(a := b) = (\lambda x. \text{if } x = a \text{ then } b \text{ else } f x)$$

This operator allows us to write down concrete states in a readable fashion. Starting from the state that is 0 everywhere, we can update it to map certain variables to given values. For example, $((\lambda x. 0) ('x' := 7)) ('y' := 3)$ maps $'x'$ to 7, $'y'$ to 3 and all other variable names to 0. Below we employ the following more compact notation

$$\langle 'x' := 7, 'y' := 3 \rangle$$

which works for any number of variables, even for none: $\langle \rangle$ is syntactic sugar for $\lambda x. 0$.

It would be easy to add subtraction and multiplication to *aexp* and extend *aval* accordingly. However, not all operators are as well behaved: division by zero raises an exception and C's ++ changes the state. Neither exceptions nor side effects can be supported by an evaluation function of the simple type $aexp \Rightarrow state \Rightarrow val$; the return type has to be more complicated.

3.1.3 Constant Folding

Program optimization is a recurring theme of this book. We start with an extremely simple example, **constant folding**, i.e., the replacement of constant subexpressions by their value. It is performed routinely by compilers. For example, the expression $Plus (V 'x') (Plus (N 3) (N 1))$ is simplified to $Plus (V 'x') (N 4)$. Function *asimp_const* performs constant folding in a bottom-up manner:

```

fun asimp_const :: "aexp  $\Rightarrow$  aexp" where
  "asimp_const (N n) = N n" |

```

```

"asimp_const (V x) = V x" |
"asimp_const (Plus a1 a2) =
  (case (asimp_const a1, asimp_const a2) of
    (N n1, N n2) => N(n1+n2) |
    (b1,b2) => Plus b1 b2) "

```

Neither N nor V can be simplified further. Given a *Plus*, first the two subexpressions are simplified. If both become numbers, they are added. In all other cases, the results are recombined with *Plus*.

It is easy to show that *asimp_const* is correct. Correctness means that *asimp_const* does not change the semantics, i.e., the value of its argument:

lemma "aval (asimp_const a) s = aval a s"

The proof is by induction on a . The two base cases N and V are trivial. In the *Plus* a_1 a_2 case, the induction hypotheses are $\text{aval } (\text{asimp_const } a_i) s = \text{aval } a_i s$ for $i=1,2$. If $\text{asimp_const } a_i = N n_i$ for $i=1,2$, then

```

aval (asimp_const (Plus a1 a2)) s
= aval (N(n1+n2)) s = n1+n2
= aval (asimp_const a1) s + aval (asimp_const a2) s
= aval (Plus a1 a2) s.

```

Otherwise

```

aval (asimp_const (Plus a1 a2)) s
= aval (Plus (asimp_const a1) (asimp_const a2)) s
= aval (asimp_const a1) s + aval (asimp_const a2) s
= aval (Plus a1 a2) s.

```

This is rather a long proof for such a simple lemma, and boring to boot. In the future we shall refrain from going through such proofs in such excessive detail. We shall simply write "The proof is by induction on a ." We will not even mention that there is a case distinction because that is obvious from what we are trying to prove, which contains the corresponding *case* expression, in the body of *asimp_const*. We can take this attitude because we merely suppress the obvious and because Isabelle has checked these proofs for us already and you can look at them in the files accompanying the book. The triviality of the proof is confirmed by the size of the Isabelle text:

```

apply (induction a)
apply (auto split: aexp.split)
done

```

The *split* modifier is the hint to *auto* to perform a case split whenever it sees a *case* expression over *aexp*. Thus we guide *auto* towards the case distinction we made in our proof above.

Let us extend constant folding: *Plus* (N 0) a and *Plus* a (N 0) should be replaced by a . Instead of extending *asimp_const* we split the optimization

process into two functions: one performs the local optimizations, the other traverses the term. The optimizations can be performed for each *Plus* separately and we define an optimizing version of *Plus*:

```
fun plus :: "aexp ⇒ aexp ⇒ aexp" where
  "plus (N i1) (N i2) = N(i1+i2)" |
  "plus (N i) a = (if i=0 then a else Plus (N i) a)" |
  "plus a (N i) = (if i=0 then a else Plus a (N i))" |
  "plus a1 a2 = Plus a1 a2"
```

It behaves like *Plus* under evaluation:

```
lemma aval_plus: "aval (plus a1 a2) s = aval a1 s + aval a2 s"
```

The proof is by induction on a_1 and a_2 using the computation induction rule for *plus* (*plus.induct*). Now we replace *Plus* by *plus* in a bottom-up manner throughout an expression:

```
fun asimp :: "aexp ⇒ aexp" where
  "asimp (N n) = N n" |
  "asimp (V x) = V x" |
  "asimp (Plus a1 a2) = plus (asimp a1) (asimp a2)"
```

Correctness is expressed exactly as for *asimp_const*:

```
lemma "aval (asimp a) s = aval a s"
```

The proof is by structural induction on a ; the *Plus* case follows with the help of Lemma *aval_plus*.

Exercises

Exercise 3.1. To show that *asimp_const* really folds all subexpressions of the form *Plus* ($N\ i$) ($N\ j$), define a function *optimal* :: *aexp* ⇒ *bool* that checks that its argument does not contain a subexpression of the form *Plus* ($N\ i$) ($N\ j$). Then prove *optimal* (*asimp_const* a).

Exercise 3.2. In this exercise we verify constant folding for *aexp* where we sum up all constants, even if they are not next to each other. For example, *Plus* ($N\ 1$) (*Plus* ($V\ x$) ($N\ 2$)) becomes *Plus* ($V\ x$) ($N\ 3$). This goes beyond *asimp*. Define a function *full_asimp* :: *aexp* ⇒ *aexp* that sums up all constants and prove its correctness: *aval* (*full_asimp* a) s = *aval* a s .

Exercise 3.3. Substitution is the process of replacing a variable by an expression in an expression. Define a substitution function *subst* :: *vname* ⇒ *aexp* ⇒ *aexp* ⇒ *aexp* such that *subst* $x\ a\ e$ is the result of replacing every occurrence of variable x by a in e . For example:

$$\text{subst } 'x'' (N\ 3) (\text{Plus } (V\ 'x'') (V\ 'y'')) = \text{Plus } (N\ 3) (V\ 'y'')$$

Prove the so-called **substitution lemma** that says that we can either substitute first and evaluate afterwards or evaluate with an updated state: $\text{aval } (\text{subst } x\ a\ e)\ s = \text{aval } e\ (s(x := \text{aval } a\ s))$. As a consequence prove $\text{aval } a_1\ s = \text{aval } a_2\ s \implies \text{aval } (\text{subst } x\ a_1\ e)\ s = \text{aval } (\text{subst } x\ a_2\ e)\ s$.

Exercise 3.4. Take a copy of theory *AExp* and modify it as follows. Extend type *aexp* with a binary constructor *Times* that represents multiplication. Modify the definition of the functions *aval* and *asimp* accordingly. You can remove *asimp_const*. Function *asimp* should eliminate 0 and 1 from multiplications as well as evaluate constant subterms. Update all proofs concerned.

Exercise 3.5. Define a datatype *aexp2* of extended arithmetic expressions that has, in addition to the constructors of *aexp*, a constructor for modelling a C-like post-increment operation $x++$, where x must be a variable. Define an evaluation function $\text{aval2} :: \text{aexp2} \Rightarrow \text{state} \Rightarrow \text{val} \times \text{state}$ that returns both the value of the expression and the new state. The latter is required because post-increment changes the state.

Extend *aexp2* and *aval2* with a division operation. Model partiality of division by changing the return type of *aval2* to $(\text{val} \times \text{state})\ \text{option}$. In case of division by 0 let *aval2* return *None*. Division on *int* is the infix *div*.

Exercise 3.6. The following type adds a *LET* construct to arithmetic expressions:

$$\text{datatype } \text{lexp} = \text{Nl } \text{int} \mid \text{Vl } \text{vname} \mid \text{Plusl } \text{lexp } \text{lexp} \mid \text{LET } \text{vname } \text{lexp } \text{lexp}$$

The *LET* constructor introduces a local variable: the value of $\text{LET } x\ e_1\ e_2$ is the value of e_2 in the state where x is bound to the value of e_1 in the original state. Define a function $\text{lval} :: \text{lexp} \Rightarrow \text{state} \Rightarrow \text{int}$ that evaluates *lexp* expressions. Remember $s(x := i)$.

Define a conversion $\text{inline} :: \text{lexp} \Rightarrow \text{aexp}$. The expression $\text{LET } x\ e_1\ e_2$ is inlined by substituting the converted form of e_1 for x in the converted form of e_2 . See Exercise 3.3 for more on substitution. Prove that *inline* is correct w.r.t. evaluation.

3.2 Boolean Expressions thy

In keeping with our minimalist philosophy, our boolean expressions contain only the bare essentials: boolean constants, negation, conjunction and comparison of arithmetic expressions for less-than:

$$\text{datatype } \text{bexp} = \text{Bc } \text{bool} \mid \text{Not } \text{bexp} \mid \text{And } \text{bexp } \text{bexp} \mid \text{Less } \text{aexp } \text{aexp}$$

Note that there are no boolean variables in this language. Other operators like disjunction and equality are easily expressed in terms of the basic ones.

Evaluation of boolean expressions is again by recursion over the abstract syntax. In the *Less* case, we switch to *aval*:

```
fun bval :: "bexp  $\Rightarrow$  state  $\Rightarrow$  bool" where
  "bval (Bc v) s = v" |
  "bval (Not b) s = ( $\neg$  bval b s)" |
  "bval (And b1 b2) s = (bval b1 s  $\wedge$  bval b2 s)" |
  "bval (Less a1 a2) s = (aval a1 s < aval a2 s)"
```

3.2.1 Constant Folding

Constant folding, including the elimination of *True* and *False* in compound expressions, works for *bexp* like it does for *aexp*: define optimizing versions of the constructors

```
fun not :: "bexp  $\Rightarrow$  bexp" where
  "not (Bc True) = Bc False" |
  "not (Bc False) = Bc True" |
  "not b = Not b"

fun "and" :: "bexp  $\Rightarrow$  bexp  $\Rightarrow$  bexp" where
  "and (Bc True) b = b" |
  "and b (Bc True) = b" |
  "and (Bc False) b = Bc False" |
  "and b (Bc False) = Bc False" |
  "and b1 b2 = And b1 b2"

fun less :: "aexp  $\Rightarrow$  aexp  $\Rightarrow$  bexp" where
  "less (N n1) (N n2) = Bc(n1 < n2)" |
  "less a1 a2 = Less a1 a2"
```

and replace the constructors in a bottom-up manner:

```
fun bsimp :: "bexp  $\Rightarrow$  bexp" where
  "bsimp (Bc v) = Bc v" |
  "bsimp (Not b) = not(bsimp b)" |
  "bsimp (And b1 b2) = and (bsimp b1) (bsimp b2)" |
  "bsimp (Less a1 a2) = less (asimp a1) (asimp a2)"
```

Note that in the *Less* case we must switch from *bsimp* to *asimp*.

Exercises

Exercise 3.7. Define functions $Eq, Le :: aexp \Rightarrow aexp \Rightarrow bexp$ and prove $bval (Eq\ a_1\ a_2)\ s = (aval\ a_1\ s = aval\ a_2\ s)$ and $bval (Le\ a_1\ a_2)\ s = (aval\ a_1\ s \leq aval\ a_2\ s)$.

Exercise 3.8. Consider an alternative type of boolean expressions featuring a conditional:

```
datatype ifexp = Bc2 bool | If ifexp ifexp ifexp | Less2 aexp aexp
```

First define an evaluation function $ifval :: ifexp \Rightarrow state \Rightarrow bool$ analogously to $bval$. Then define two functions $b2ifexp :: bexp \Rightarrow ifexp$ and $if2bexp :: ifexp \Rightarrow bexp$ and prove their correctness, i.e., that they preserve the value of an expression.

Exercise 3.9. Define a new type of purely boolean expressions

```
datatype pbool =  
  VAR vname | NEG pbool | AND pbool pbool | OR pbool pbool
```

where variables range over values of type $bool$:

```
fun pbval :: "pbool  $\Rightarrow$  (vname  $\Rightarrow$  bool)  $\Rightarrow$  bool" where  
  "pbval (VAR x) s = s x" |  
  "pbval (NEG b) s = ( $\neg$  pbval b s)" |  
  "pbval (AND b1 b2) s = (pbval b1 s  $\wedge$  pbval b2 s)" |  
  "pbval (OR b1 b2) s = (pbval b1 s  $\vee$  pbval b2 s)"
```

Define a function $is_nnf :: pbool \Rightarrow bool$ that checks whether a boolean expression is in NNF (negation normal form), i.e., if NEG is only applied directly to VAR s. Also define a function $nnf :: pbool \Rightarrow pbool$ that converts a $pbool$ into NNF by pushing NEG inwards as much as possible. Prove that nnf preserves the value ($pbval (nnf\ b)\ s = pbval\ b\ s$) and returns an NNF ($is_nnf (nnf\ b)$).

An expression is in DNF (disjunctive normal form) if it is in NNF and if no OR occurs below an AND . Define a corresponding test $is_dnf :: pbool \Rightarrow bool$. An NNF can be converted into a DNF in a bottom-up manner. The critical case is the conversion of $AND\ b_1\ b_2$. Having converted b_1 and b_2 , apply distributivity of AND over OR . Define a conversion function $dnf_of_nnf :: pbool \Rightarrow pbool$ from NNF to DNF. Prove that your function preserves the value ($pbval (dnf_of_nnf\ b)\ s = pbval\ b\ s$) and converts an NNF into a DNF ($is_nnf\ b \implies is_dnf (dnf_of_nnf\ b)$).

3.3 Stack Machine and Compilation thy

This section describes a simple stack machine and compiler for arithmetic expressions. The stack machine has three instructions:

datatype *instr* = *LOADI val* | *LOAD vname* | *ADD*

The semantics of the three instructions will be the following: *LOADI n* (load immediate) puts *n* on top of the stack, *LOAD x* puts the value of *x* on top of the stack, and *ADD* replaces the two topmost elements of the stack by their sum. A stack is simply a list of values:

type_synonym *stack* = "*val list*"

The top of the stack is its first element, the head of the list.

An instruction is executed in the context of a state and transforms a stack into a new stack:

```
fun exec1 :: "instr  $\Rightarrow$  state  $\Rightarrow$  stack  $\Rightarrow$  stack" where
  "exec1 (LOADI n) _ stk = n # stk" |
  "exec1 (LOAD x) s stk = s(x) # stk" |
  "exec1 ADD _ (j # i # stk) = (i + j) # stk"
```

A list of instructions is executed one by one:

```
fun exec :: "instr list  $\Rightarrow$  state  $\Rightarrow$  stack  $\Rightarrow$  stack" where
  "exec [] _ stk = stk" |
  "exec (i # is) s stk = exec is s (exec1 i s stk)"
```

The simplicity of this definition is due to the absence of jump instructions. Forward jumps could still be accommodated, but backward jumps would cause a serious problem: execution might not terminate.

Compilation of arithmetic expressions is straightforward:

```
fun comp :: "aexp  $\Rightarrow$  instr list" where
  "comp (N n) = [LOADI n]" |
  "comp (V x) = [LOAD x]" |
  "comp (Plus e1 e2) = comp e1 @ comp e2 @ [ADD]"
```

The correctness statement says that executing a compiled expression is the same as putting the value of the expression on the stack:

lemma "*exec (comp a) s stk = aval a s # stk*"

The proof is by induction on *a* and relies on the lemma

exec (is₁ @ is₂) s stk = exec is₂ s (exec is₁ s stk)

which is proved by induction in *is₁*.

Compilation of boolean expressions is covered later and requires conditional jumps.

Exercises

Exercise 3.10. A **stack underflow** occurs when executing an *ADD* instruction on a stack of size less than 2. In our semantics a term *exec1 ADD s stk* where *length stk < 2* is simply some unspecified value, not an error or exception — HOL does not have those concepts. Modify theory *ASM* such that stack underflow is modelled by *None* and normal execution by *Some*, i.e., the execution functions have return type *stack option*. Modify all theorems and proofs accordingly.

Exercise 3.11. This exercise is about a register machine and compiler for *aexp*. The machine instructions are

datatype *instr* = *LDI int reg* | *LD vname reg* | *ADD reg reg*

where type *reg* is a synonym for *nat*. Instruction *LDI i r* loads *i* into register *r*, *LD x r* loads the value of *x* into register *r*, and *ADD r₁ r₂* adds register *r₂* to register *r₁*.

Define the execution of an instruction given a state and a register state (= function from registers to integers); the result is the new register state:

fun *exec1* :: *instr* \Rightarrow *state* \Rightarrow (*reg* \Rightarrow *int*) \Rightarrow *reg* \Rightarrow *int*

Define the execution *exec* of a list of instructions as for the stack machine.

The compiler takes an arithmetic expression *a* and a register *r* and produces a list of instructions whose execution places the value of *a* into *r*. The registers $> r$ should be used in a stack-like fashion for intermediate results, the ones $< r$ should be left alone. Define the compiler and prove it correct: *exec (comp a r) s rs r = aval a s*.

Exercise 3.12. This is a variation on the previous exercise. Let the instruction set be

datatype *instr0* = *LDI0 val* | *LD0 vname* | *MV0 reg* | *ADD0 reg*

All instructions refer implicitly to register 0 as the source (*MV0*) or target (all others). Define a compiler pretty much as explained above except that the compiled code leaves the value of the expression in register 0. Prove that *exec (comp a r) s rs 0 = aval a s*.

Logic and Proof Beyond Equality

4.1 Formulas

The core syntax of formulas (*form* below) provides the standard logical constructs, in decreasing order of precedence:

$$\begin{aligned} \text{form} ::= & (\text{form}) \mid \text{True} \mid \text{False} \mid \text{term} = \text{term} \\ & \mid \neg \text{form} \mid \text{form} \wedge \text{form} \mid \text{form} \vee \text{form} \mid \text{form} \longrightarrow \text{form} \\ & \mid \forall x. \text{form} \mid \exists x. \text{form} \end{aligned}$$

Terms are the ones we have seen all along, built from constants, variables, function application and λ -abstraction, including all the syntactic sugar like infix symbols, *if*, *case*, etc.

! Remember that formulas are simply terms of type *bool*. Hence $=$ also works for formulas. Beware that $=$ has a higher precedence than the other logical operators. Hence $s = t \wedge A$ means $(s = t) \wedge A$, and $A \wedge B = B \wedge A$ means $A \wedge (B = B) \wedge A$. Logical equivalence can also be written with \longleftrightarrow instead of $=$, where \longleftrightarrow has the same low precedence as \longrightarrow . Hence $A \wedge B \longleftrightarrow B \wedge A$ really means $(A \wedge B) \longleftrightarrow (B \wedge A)$.

! Quantifiers need to be enclosed in parentheses if they are nested within other constructs (just like *if*, *case* and *let*).

The most frequent logical symbols and their ASCII representations are shown in Fig. 4.1. The first column shows the symbols, the other columns ASCII representations. The $\langle \dots \rangle$ form is always converted into the symbolic form by the Isabelle interfaces, the treatment of the other ASCII forms depends on the interface. The ASCII forms $/\wedge$ and $\backslash/$ are special in that they are merely keyboard shortcuts for the interface and not logical symbols by themselves.

\forall	<code>\<forall></code>	ALL
\exists	<code>\<exists></code>	EX
λ	<code>\<lambda></code>	%
\longrightarrow	<code>--></code>	
\longleftrightarrow	<code><-></code>	
\wedge	<code>\&</code>	&
\vee	<code>\ </code>	
\neg	<code>\<not></code>	~
\neq	<code>\<noteq></code>	~=

Fig. 4.1. Logical symbols and their ASCII forms

! The implication \implies is part of the Isabelle framework. It structures theorems and proof states, separating assumptions from conclusions. The implication \longrightarrow is part of the logic HOL and can occur inside the formulas that make up the assumptions and conclusion. Theorems should be of the form $\llbracket A_1; \dots; A_n \rrbracket \implies A$, not $A_1 \wedge \dots \wedge A_n \longrightarrow A$. Both are logically equivalent but the first one works better when using the theorem in further proofs.

The ASCII representation of \llbracket and \rrbracket is `[|` and `]`.

4.2 Sets

Sets of elements of type $'a$ have type $'a$ *set*. They can be finite or infinite. Sets come with the usual notation:

- $\{\}, \{e_1, \dots, e_n\}$
- $e \in A, A \subseteq B$
- $A \cup B, A \cap B, A - B, -A$

(where $A - B$ and $-A$ are set difference and complement) and much more. $UNIV$ is the set of all elements of some type. Set comprehension is written $\{x. P\}$ rather than $\{x \mid P\}$.

! In $\{x. P\}$ the x must be a variable. Set comprehension involving a proper term t must be written $\{t \mid x y. P\}$, where $x y$ are those free variables in t that occur in P . This is just a shorthand for $\{v. \exists x y. v = t \wedge P\}$, where v is a new variable. For example, $\{x + y \mid x. x \in A\}$ is short for $\{v. \exists x. v = x + y \wedge x \in A\}$.

Here are the ASCII representations of the mathematical symbols:

\in	<code>\<in></code>	:
\subseteq	<code>\<subsetq></code>	<code><=</code>
\cup	<code>\<union></code>	Un
\cap	<code>\<inter></code>	Int

Sets also allow bounded quantifications $\forall x \in A. P$ and $\exists x \in A. P$.

For the more ambitious, there are also \bigcup and \bigcap :

$$\bigcup A = \{x. \exists B \in A. x \in B\} \quad \bigcap A = \{x. \forall B \in A. x \in B\}$$

The ASCII forms of \bigcup are `\<Union>` and `Union`, those of \bigcap are `\<Inter>` and `Inter`. There are also indexed unions and intersections:

$$\begin{aligned} \left(\bigcup_{x \in A} B\ x\right) &= \{y. \exists x \in A. y \in B\ x\} \\ \left(\bigcap_{x \in A} B\ x\right) &= \{y. \forall x \in A. y \in B\ x\} \end{aligned}$$

The ASCII forms are `UN x:A. B` and `INT x:A. B` where x may occur in B . If A is UNIV you can write `UN x. B` and `INT x. B`.

Some other frequently useful functions on sets are the following:

<code>set :: 'a list \Rightarrow 'a set</code>	converts a list to the set of its elements
<code>finite :: 'a set \Rightarrow bool</code>	is true iff its argument is finite
<code>card :: 'a set \Rightarrow nat</code>	is the cardinality of a finite set
	and is 0 for all infinite sets
<code>f ' A = {y. $\exists x \in A. y = f\ x$}</code>	is the image of a function over a set

See [64] for the wealth of further predefined functions in theory *Main*.

Exercises

Exercise 4.1. Start from the data type of binary trees defined earlier:

```
datatype 'a tree = Tip | Node "'a tree" 'a "'a tree"
```

Define a function `set :: 'a tree \Rightarrow 'a set` that returns the elements in a tree and a function `ord :: int tree \Rightarrow bool` that tests if an *int tree* is ordered.

Define a function `ins` that inserts an element into an ordered *int tree* while maintaining the order of the tree. If the element is already in the tree, the same tree should be returned. Prove correctness of `ins`: `set (ins x t) = {x} \cup set t` and `ord t \Longrightarrow ord (ins i t)`.

4.3 Proof Automation

So far we have only seen *simp* and *auto*: Both perform rewriting, both can also prove linear arithmetic facts (no multiplication), and *auto* is also able to prove simple logical or set-theoretic goals:

```
lemma " $\forall x. \exists y. x = y$ "
by auto
```

```
lemma " $A \subseteq B \cap C \Longrightarrow A \subseteq B \cup C$ "
by auto
```

where

by *proof-method*

is short for

apply *proof-method*
done

The key characteristics of both *simp* and *auto* are

- They show you where they got stuck, giving you an idea how to continue.
- They perform the obvious steps but are highly incomplete.

A proof method is **complete** if it can prove all true formulas. There is no complete proof method for HOL, not even in theory. Hence all our proof methods only differ in how incomplete they are.

A proof method that is still incomplete but tries harder than *auto* is *fastforce*. It either succeeds or fails, it acts on the first subgoal only, and it can be modified like *auto*, e.g., with *simp add*. Here is a typical example of what *fastforce* can do:

lemma "[$\forall xs \in A. \exists ys. xs = ys @ ys; \quad us \in A$]
 $\implies \exists n. \text{length } us = n + n$ "
 by *fastforce*

This lemma is out of reach for *auto* because of the quantifiers. Even *fastforce* fails when the quantifier structure becomes more complicated. In a few cases, its slow version *force* succeeds where *fastforce* fails.

The method of choice for complex logical goals is *blast*. In the following example, *T* and *A* are two binary predicates. It is shown that if *T* is total, *A* is antisymmetric and *T* is a subset of *A*, then *A* is a subset of *T*:

lemma
 "[$\forall x y. T x y \vee T y x;$
 $\forall x y. A x y \wedge A y x \longrightarrow x = y;$
 $\forall x y. T x y \longrightarrow A x y$]
 $\implies \forall x y. A x y \longrightarrow T x y$ "
 by *blast*

We leave it to the reader to figure out why this lemma is true. Method *blast*

- is (in principle) a complete proof procedure for first-order formulas, a fragment of HOL. In practice there is a search bound.
- does no rewriting and knows very little about equality.
- covers logic, sets and relations.
- either succeeds or fails.

Because of its strength in logic and sets and its weakness in equality reasoning, it complements the earlier proof methods.

4.3.1 Sledgehammer

Command `sledgehammer` calls a number of external automatic theorem provers (ATPs) that run for up to 30 seconds searching for a proof. Some of these ATPs are part of the Isabelle installation, others are queried over the internet. If successful, a proof command is generated and can be inserted into your proof. The biggest win of `sledgehammer` is that it will take into account the whole lemma library and you do not need to feed in any lemma explicitly. For example,

```
lemma "[ xs @ ys = ys @ xs; length xs = length ys ] ==> xs = ys"
```

cannot be solved by any of the standard proof methods, but `sledgehammer` finds the following proof:

```
by (metis append_eq_conv_conj)
```

We do not explain how the proof was found but what this command means. For a start, Isabelle does not trust external tools (and in particular not the translations from Isabelle's logic to those tools!) and insists on a proof that it can check. This is what *metis* does. It is given a list of lemmas and tries to find a proof using just those lemmas (and pure logic). In contrast to using *simp* and friends who know a lot of lemmas already, using *metis* manually is tedious because one has to find all the relevant lemmas first. But that is precisely what `sledgehammer` does for us. In this case lemma *append_eq_conv_conj* alone suffices:

$$(xs @ ys = zs) = (xs = take (length xs) zs \wedge ys = drop (length xs) zs)$$

We leave it to the reader to figure out why this lemma suffices to prove the above lemma, even without any knowledge of what the functions *take* and *drop* do. Keep in mind that the variables in the two lemmas are independent of each other, despite the same names, and that you can substitute arbitrary values for the free variables in a lemma.

Just as for the other proof methods we have seen, there is no guarantee that `sledgehammer` will find a proof if it exists. Nor is `sledgehammer` superior to the other proof methods. They are incomparable. Therefore it is recommended to apply *simp* or *auto* before invoking `sledgehammer` on what is left.

4.3.2 Arithmetic

By arithmetic formulas we mean formulas involving variables, numbers, $+$, $-$, $=$, $<$, \leq and the usual logical connectives \neg , \wedge , \vee , \longrightarrow , \longleftrightarrow . Strictly speaking, this is known as **linear arithmetic** because it does not involve multiplication, although multiplication with numbers, e.g., $2*n$, is allowed. Such formulas can be proved by *arith*:

lemma "[(a::nat) ≤ x + b; 2*x < c] ⇒ 2*a + 1 ≤ 2*b + c"
by *arith*

In fact, *auto* and *simp* can prove many linear arithmetic formulas already, like the one above, by calling a weak but fast version of *arith*. Hence it is usually not necessary to invoke *arith* explicitly.

The above example involves natural numbers, but integers (type *int*) and real numbers (type *real*) are supported as well. As are a number of further operators like *min* and *max*. On *nat* and *int*, *arith* can even prove theorems with quantifiers in them, but we will not enlarge on that here.

4.3.3 Trying Them All

If you want to try all of the above automatic proof methods you simply type
try

There is also a lightweight variant *try0* that does not call sledgehammer. If desired, specific simplification and introduction rules can be added:

try0 simp: ... intro: ...

4.4 Single Step Proofs

Although automation is nice, it often fails, at least initially, and you need to find out why. When *fastforce* or *blast* simply fail, you have no clue why. At this point, the stepwise application of proof rules may be necessary. For example, if *blast* fails on $A \wedge B$, you want to attack the two conjuncts A and B separately. This can be achieved by applying *conjunction introduction*

$$\frac{?P \quad ?Q}{?P \wedge ?Q} \text{ conjI}$$

to the proof state. We will now examine the details of this process.

4.4.1 Instantiating Unknowns

We had briefly mentioned earlier that after proving some theorem, Isabelle replaces all free variables x by so called **unknowns** $?x$. We can see this clearly in rule *conjI*. These unknowns can later be instantiated explicitly or implicitly:

- By hand, using *of*. The expression *conjI*[*of* " $a=b$ " "*False*"] instantiates the unknowns in *conjI* from left to right with the two formulas $a=b$ and *False*, yielding the rule

$$\frac{a = b \quad False}{a = b \wedge False}$$

In general, $th[of\ string_1 \dots string_n]$ instantiates the unknowns in the theorem th from left to right with the terms $string_1$ to $string_n$.

- By unification. **Unification** is the process of making two terms syntactically equal by suitable instantiations of unknowns. For example, unifying $?P \wedge ?Q$ with $a = b \wedge False$ instantiates $?P$ with $a = b$ and $?Q$ with $False$.

We need not instantiate all unknowns. If we want to skip a particular one we can write $_$ instead, for example $conjI[of\ _ "False"]$. Unknowns can also be instantiated by name using *where*, for example $conjI[where\ ?P = "a=b"$ and $?Q = "False"]$.

4.4.2 Rule Application

Rule application means applying a rule backwards to a proof state. For example, applying rule *conjI* to a proof state

$$1. \dots \Longrightarrow A \wedge B$$

results in two subgoals, one for each premise of *conjI*:

$$\begin{aligned} 1. \dots &\Longrightarrow A \\ 2. \dots &\Longrightarrow B \end{aligned}$$

In general, the application of a rule $\llbracket A_1; \dots; A_n \rrbracket \Longrightarrow A$ to a subgoal $\dots \Longrightarrow C$ proceeds in two steps:

1. Unify A and C , thus instantiating the unknowns in the rule.
2. Replace the subgoal C with n new subgoals A_1 to A_n .

This is the command to apply rule *xyz*:

apply(rule *xyz*)

This is also called **backchaining** with rule *xyz*.

4.4.3 Introduction Rules

Conjunction introduction (*conjI*) is one example of a whole class of rules known as **introduction rules**. They explain under which premises some logical construct can be introduced. Here are some further useful introduction rules:

$$\frac{?P \Longrightarrow ?Q}{?P \longrightarrow ?Q} \text{ impI} \qquad \frac{\bigwedge x. ?P\ x}{\forall x. ?P\ x} \text{ allI}$$

$$\frac{?P \implies ?Q \quad ?Q \implies ?P}{?P = ?Q} \text{iffI}$$

These rules are part of the logical system of **natural deduction** (e.g., [44]). Although we intentionally de-emphasize the basic rules of logic in favour of automatic proof methods that allow you to take bigger steps, these rules are helpful in locating where and why automation fails. When applied backwards, these rules decompose the goal:

- *conjI* and *iffI* split the goal into two subgoals,
- *impI* moves the left-hand side of a HOL implication into the list of assumptions,
- and *allI* removes a \forall by turning the quantified variable into a fixed local variable of the subgoal.

Isabelle knows about these and a number of other introduction rules. The command

apply rule

automatically selects the appropriate rule for the current subgoal.

You can also turn your own theorems into introduction rules by giving them the *intro* attribute, analogous to the *simp* attribute. In that case *blast*, *fastforce* and (to a limited extent) *auto* will automatically backchain with those theorems. The *intro* attribute should be used with care because it increases the search space and can lead to nontermination. Sometimes it is better to use it only in specific calls of *blast* and friends. For example, *le_trans*, transitivity of \leq on type *nat*, is not an introduction rule by default because of the disastrous effect on the search space, but can be useful in specific situations:

```
lemma "[ (a::nat) ≤ b; b ≤ c; c ≤ d; d ≤ e ] ⟹ a ≤ e"
by(blast intro: le_trans)
```

Of course this is just an example and could be proved by *arith*, too.

4.4.4 Forward Proof

Forward proof means deriving new theorems from old theorems. We have already seen a very simple form of forward proof: the *of* operator for instantiating unknowns in a theorem. The big brother of *of* is *OF* for applying one theorem to others. Given a theorem $A \implies B$ called *r* and a theorem A' called *r'*, the theorem $r[OF\ r']$ is the result of applying *r* to *r'*, where *r* should be viewed as a function taking a theorem *A* and returning *B*. More precisely, *A* and *A'* are unified, thus instantiating the unknowns in *B*, and the result is the instantiated *B*. Of course, unification may also fail.

! Application of rules to other rules operates in the forward direction: from the premises to the conclusion of the rule; application of rules to proof states operates in the backward direction, from the conclusion to the premises.

In general r can be of the form $\llbracket A_1; \dots; A_n \rrbracket \Longrightarrow A$ and there can be multiple argument theorems r_1 to r_m (with $m \leq n$), in which case $r[OF\ r_1 \dots r_m]$ is obtained by unifying and thus proving A_i with r_i , $i = 1 \dots m$. Here is an example, where *refl* is the theorem $?t = ?t$:

```
thm conjI[OF refl[of "a"] refl[of "b"]]
```

yields the theorem $a = a \wedge b = b$. The command **thm** merely displays the result.

Forward reasoning also makes sense in connection with proof states. Therefore *blast*, *fastforce* and *auto* support a modifier *dest* which instructs the proof method to use certain rules in a forward fashion. If r is of the form $A \Longrightarrow B$, the modifier *dest: r* allows proof search to reason forward with r , i.e., to replace an assumption A' , where A' unifies with A , with the correspondingly instantiated B . For example, *Suc_leD* is the theorem $Suc\ m \leq n \Longrightarrow m \leq n$, which works well for forward reasoning:

```
lemma "Suc(Suc(Suc a)) ≤ b ⟹ a ≤ b"
by(blast dest: Suc_leD)
```

In this particular example we could have backchained with *Suc_leD*, too, but because the premise is more complicated than the conclusion this can easily lead to nontermination.

! To ease readability we will drop the question marks in front of unknowns from now on.

4.5 Inductive Definitions

Inductive definitions are the third important definition facility, after datatypes and recursive function. In fact, they are the key construct in the definition of operational semantics in the second part of the book.

4.5.1 An Example: Even Numbers

Here is a simple example of an inductively defined predicate:

- 0 is even
- If n is even, so is $n + 2$.

The operative word “inductive” means that these are the only even numbers. In Isabelle we give the two rules the names *ev0* and *evSS* and write

```
inductive ev :: "nat ⇒ bool" where
  ev0:   "ev 0" |
  evSS:  "ev n ⇒ ev (n + 2)"
```

To get used to inductive definitions, we will first prove a few properties of *ev* informally before we descend to the Isabelle level.

How do we prove that some number is even, e.g., *ev 4*? Simply by combining the defining rules for *ev*:

$$ev\ 0 \implies ev\ (0 + 2) \implies ev((0 + 2) + 2) = ev\ 4$$

Rule Induction

Showing that all even numbers have some property is more complicated. For example, let us prove that the inductive definition of even numbers agrees with the following recursive one:

```
fun evn :: "nat ⇒ bool" where
  "evn 0 = True" |
  "evn (Suc 0) = False" |
  "evn (Suc(Suc n)) = evn n"
```

We prove $ev\ m \implies evn\ m$. That is, we assume *ev m* and by induction on the form of its derivation prove *evn m*. There are two cases corresponding to the two rules for *ev*:

Case *ev0*: *ev m* was derived by rule *ev 0*:

$$\implies m = 0 \implies evn\ m = evn\ 0 = True$$

Case *evSS*: *ev m* was derived by rule $ev\ n \implies ev\ (n + 2)$:

$$\begin{aligned} \implies m &= n + 2 \text{ and by induction hypothesis } evn\ n \\ \implies evn\ m &= evn(n + 2) = evn\ n = True \end{aligned}$$

What we have just seen is a special case of **rule induction**. Rule induction applies to propositions of this form

$$ev\ n \implies P\ n$$

That is, we want to prove a property *P n* for all even *n*. But if we assume *ev n*, then there must be some derivation of this assumption using the two defining rules for *ev*. That is, we must prove

Case *ev0*: *P 0*

Case *evSS*: $\llbracket ev\ n; P\ n \rrbracket \implies P\ (n + 2)$

The corresponding rule is called *ev.induct* and looks like this:

$$\frac{ev\ n \quad P\ 0 \quad \bigwedge n. \llbracket ev\ n; P\ n \rrbracket \implies P\ (n + 2)}{P\ n}$$

The first premise *ev n* enforces that this rule can only be applied in situations where we know that *n* is even.

Note that in the induction step we may not just assume *P n* but also *ev n*, which is simply the premise of rule *evSS*. Here is an example where the local assumption *ev n* comes in handy: we prove $ev\ m \implies ev\ (m - 2)$ by induction on *ev m*. Case *ev0* requires us to prove $ev\ (0 - 2)$, which follows from *ev 0* because $0 - 2 = 0$ on type *nat*. In case *evSS* we have $m = n + 2$ and may assume *ev n*, which implies $ev\ (m - 2)$ because $m - 2 = (n + 2) - 2 = n$. We did not need the induction hypothesis at all for this proof (it is just a case analysis of which rule was used) but having *ev n* at our disposal in case *evSS* was essential. This case analysis of rules is also called “rule inversion” and is discussed in more detail in [Chapter 5](#).

In Isabelle

Let us now recast the above informal proofs in Isabelle. For a start, we use *Suc* terms instead of numerals in rule *evSS*:

$$ev\ n \implies ev\ (Suc\ (Suc\ n))$$

This avoids the difficulty of unifying *n+2* with some numeral, which is not automatic.

The simplest way to prove $ev\ (Suc\ (Suc\ (Suc\ (Suc\ 0))))$ is in a forward direction: *evSS[OF evSS[OF ev0]]* yields the theorem $ev\ (Suc\ (Suc\ (Suc\ (Suc\ 0))))$. Alternatively, you can also prove it as a lemma in backwards fashion. Although this is more verbose, it allows us to demonstrate how each rule application changes the proof state:

lemma "ev(Suc(Suc(Suc(Suc 0))))"

1. $ev\ (Suc\ (Suc\ (Suc\ (Suc\ 0))))$

apply(rule *evSS*)

1. $ev\ (Suc\ (Suc\ 0))$

apply(rule *evSS*)

1. $ev\ 0$

```

apply(rule ev0)
done

```

Rule induction is applied by giving the induction rule explicitly via the *rule*: modifier:

```

lemma "ev m  $\implies$  evn m"
apply(induction rule: ev.induct)
by(simp_all)

```

Both cases are automatic. Note that if there are multiple assumptions of the form *ev t*, method *induction* will induct on the leftmost one.

As a bonus, we also prove the remaining direction of the equivalence of *ev* and *evn*:

```

lemma "evn n  $\implies$  ev n"
apply(induction n rule: evn.induct)

```

This is a proof by computation induction on *n* (see [Section 2.3.4](#)) that sets up three subgoals corresponding to the three equations for *evn*:

1. *evn 0 \implies ev 0*
2. *evn (Suc 0) \implies ev (Suc 0)*
3. $\bigwedge n. \llbracket \text{evn } n \implies \text{ev } n; \text{evn } (\text{Suc } (\text{Suc } n)) \rrbracket \implies \text{ev } (\text{Suc } (\text{Suc } n))$

The first and third subgoals follow with *ev0* and *evSS*, and the second subgoal is trivially true because *evn (Suc 0)* is *False*:

```

by (simp_all add: ev0 evSS)

```

The rules for *ev* make perfect simplification and introduction rules because their premises are always smaller than the conclusion. It makes sense to turn them into simplification and introduction rules permanently, to enhance proof automation. They are named *ev.intros* by Isabelle:

```

declare ev.intros[simp,intro]

```

The rules of an inductive definition are not simplification rules by default because, in contrast to recursive functions, there is no termination requirement for inductive definitions.

Inductive Versus Recursive

We have seen two definitions of the notion of evenness, an inductive and a recursive one. Which one is better? Much of the time, the recursive one is more convenient: it allows us to do rewriting in the middle of terms, and it expresses both the positive information (which numbers are even) and the negative information (which numbers are not even) directly. An inductive definition

only expresses the positive information directly. The negative information, for example, that 1 is not even, has to be proved from it (by induction or rule inversion). On the other hand, rule induction is tailor-made for proving $ev\ n \implies P\ n$ because it only asks you to prove the positive cases. In the proof of $evn\ n \implies P\ n$ by computation induction via *evn.induct*, we are also presented with the trivial negative cases. If you want the convenience of both rewriting and rule induction, you can make two definitions and show their equivalence (as above) or make one definition and prove additional properties from it, for example rule induction from computation induction.

But many concepts do not admit a recursive definition at all because there is no datatype for the recursion (for example, the transitive closure of a relation), or the recursion would not terminate (for example, an interpreter for a programming language). Even if there is a recursive definition, if we are only interested in the positive information, the inductive definition may be much simpler.

4.5.2 The Reflexive Transitive Closure

Evenness is really more conveniently expressed recursively than inductively. As a second and very typical example of an inductive definition we define the reflexive transitive closure. It will also be an important building block for some of the semantics considered in the second part of the book.

The reflexive transitive closure, called *star* below, is a function that maps a binary predicate to another binary predicate: if r is of type $\tau \Rightarrow \tau \Rightarrow bool$ then *star* r is again of type $\tau \Rightarrow \tau \Rightarrow bool$, and *star* $r\ x\ y$ means that x and y are in the relation *star* r . Think r^* when you see *star* r , because *star* r is meant to be the reflexive transitive closure. That is, *star* $r\ x\ y$ is meant to be true if from x we can reach y in finitely many r steps. This concept is naturally defined inductively:

```
inductive star :: "('a  $\Rightarrow$  'a  $\Rightarrow$  bool)  $\Rightarrow$  'a  $\Rightarrow$  'a  $\Rightarrow$  bool" for r where
  refl: "star r x x" |
  step: "r x y  $\implies$  star r y z  $\implies$  star r x z"
```

The base case *refl* is reflexivity: $x = y$. The step case *step* combines an r step (from x to y) and a *star* r step (from y to z) into a *star* r step (from x to z). The “for r ” in the header is merely a hint to Isabelle that r is a fixed parameter of *star*, in contrast to the further parameters of *star*, which change. As a result, Isabelle generates a simpler induction rule.

By definition *star* r is reflexive. It is also transitive, but we need rule induction to prove that:

```
lemma star_trans: "star r x y  $\implies$  star r y z  $\implies$  star r x z"
```

apply(*induction rule: star.induct*)

The induction is over $\text{star } r \ x \ y$ (the first matching assumption) and we try to prove $\text{star } r \ y \ z \implies \text{star } r \ x \ z$, which we abbreviate by $P \ x \ y$. These are our two subgoals:

1. $\bigwedge x. \text{star } r \ x \ z \implies \text{star } r \ x \ z$
2. $\bigwedge u \ x \ y.$
 $\llbracket r \ u \ x; \text{star } r \ x \ y; \text{star } r \ y \ z \implies \text{star } r \ x \ z; \text{star } r \ y \ z \rrbracket$
 $\implies \text{star } r \ u \ z$

The first one is $P \ x \ x$, the result of case *refl*, and it is trivial:

apply(*assumption*)

Let us examine subgoal 2, case *step*. Assumptions $r \ u \ x$ and $\text{star } r \ x \ y$ are the premises of rule *step*. Assumption $\text{star } r \ y \ z \implies \text{star } r \ x \ z$ is $P \ x \ y$, the IH coming from $\text{star } r \ x \ y$. We have to prove $P \ u \ y$, which we do by assuming $\text{star } r \ y \ z$ and proving $\text{star } r \ u \ z$. The proof itself is straightforward: from $\text{star } r \ y \ z$ the IH leads to $\text{star } r \ x \ z$ which, together with $r \ u \ x$, leads to $\text{star } r \ u \ z$ via rule *step*:

apply(*metis step*)

done

4.5.3 The General Case

Inductive definitions have approximately the following general form:

inductive $I :: \text{"}\tau \Rightarrow \text{bool"}$ **where**

followed by a sequence of (possibly named) rules of the form

$$\llbracket I \ a_1; \dots; I \ a_n \rrbracket \implies I \ a$$

separated by $|$. As usual, n can be 0. The corresponding rule induction principle $I.\text{induct}$ applies to propositions of the form

$$I \ x \implies P \ x$$

where P may itself be a chain of implications.



Rule induction is always on the leftmost premise of the goal. Hence $I \ x$ must be the first premise.

Proving $I \ x \implies P \ x$ by rule induction means proving for every rule of I that P is invariant:

$$\llbracket I \ a_1; P \ a_1; \dots; I \ a_n; P \ a_n \rrbracket \implies P \ a$$

The above format for inductive definitions is simplified in a number of respects. I can have any number of arguments and each rule can have additional premises not involving I , so-called **side conditions**. In rule inductions, these side conditions appear as additional assumptions. The **for** clause seen in the definition of the reflexive transitive closure simplifies the induction rule.

Exercises

Exercise 4.2. Formalize the following definition of palindromes

- The empty list and a singleton list are palindromes.
- If xs is a palindrome, so is $a \# xs @ [a]$.

as an inductive predicate $palindrome :: 'a\ list \Rightarrow bool$ and prove that $rev\ xs = xs$ if xs is a palindrome.

Exercise 4.3. We could also have defined $star$ as follows:

inductive $star' :: "('a \Rightarrow 'a \Rightarrow bool) \Rightarrow 'a \Rightarrow 'a \Rightarrow bool"$ **for** r **where**
 $refl'$: $"star'\ r\ x\ x"$ |
 $step'$: $"star'\ r\ x\ y \Longrightarrow r\ y\ z \Longrightarrow star'\ r\ x\ z"$

The single r step is performed after rather than before the $star'$ steps. Prove $star'\ r\ x\ y \Longrightarrow star\ r\ x\ y$ and $star\ r\ x\ y \Longrightarrow star'\ r\ x\ y$. You may need lemmas. Note that rule induction fails if the assumption about the inductive predicate is not the first assumption.

Exercise 4.4. Analogous to $star$, give an inductive definition of the n -fold iteration of a relation r : $iter\ r\ n\ x\ y$ should hold if there are x_0, \dots, x_n such that $x = x_0$, $x_n = y$ and $r\ x_i\ x_{i+1}$ for all $i < n$. Correct and prove the following claim: $star\ r\ x\ y \Longrightarrow iter\ r\ n\ x\ y$.

Exercise 4.5. A context-free grammar can be seen as an inductive definition where each nonterminal A is an inductively defined predicate on lists of terminal symbols: $A(w)$ means that w is in the language generated by A . For example, the production $S \rightarrow aSb$ can be viewed as the implication $S\ w \Longrightarrow S\ (a \# w @ [b])$ where a and b are terminal symbols, i.e., elements of some alphabet. The alphabet can be defined like this: **datatype** $alpha = a \mid b \mid \dots$

Define the two grammars (where ε is the empty word)

$$\begin{array}{lcl} S & \rightarrow & \varepsilon \mid aSb \mid SS \\ T & \rightarrow & \varepsilon \mid TaTb \end{array}$$

as two inductive predicates. If you think of a and b as "(" and ")", the grammars define balanced strings of parentheses. Prove $T\ w \Longrightarrow S\ w$ and $S\ w \Longrightarrow T\ w$ separately and conclude $S\ w = T\ w$.

Exercise 4.6. In [Section 3.1](#) we defined a recursive evaluation function $aval :: aexp \Rightarrow state \Rightarrow val$. Define an inductive evaluation predicate $aval_rel :: aexp \Rightarrow state \Rightarrow val \Rightarrow bool$ and prove that it agrees with the recursive function: $aval_rel\ a\ s\ v \Longrightarrow aval\ a\ s = v$, $aval\ a\ s = v \Longrightarrow aval_rel\ a\ s\ v$ and thus $aval_rel\ a\ s\ v \longleftrightarrow aval\ a\ s = v$.

Exercise 4.7. Consider the stack machine from Chapter 3 and recall the concept of **stack underflow** from [Exercise 3.10](#). Define an inductive predicate $ok :: nat \Rightarrow instr\ list \Rightarrow nat \Rightarrow bool$ such that $ok\ n\ is\ n'$ means that with any initial stack of length n the instructions is can be executed without stack underflow and that the final stack has length n' . Prove that ok correctly computes the final stack size

$$\llbracket ok\ n\ is\ n';\ length\ stk = n \rrbracket \Longrightarrow length\ (exec\ is\ s\ stk) = n'$$

and that instruction sequences generated by $comp$ cannot cause stack underflow: $ok\ n\ (comp\ a)\ ?$ for some suitable value of $?$.

Isar: A Language for Structured Proofs

Apply-scripts are unreadable and hard to maintain. The language of choice for larger proofs is **Isar**. The two key features of Isar are:

- It is structured, not linear.
- It is readable without its being run because you need to state what you are proving at any given point.

Whereas apply-scripts are like assembly language programs, Isar proofs are like structured programs with comments. A typical Isar proof looks like this:

```
proof
  assume "formula0"
  have "formula1" by simp
  ⋮
  have "formulan" by blast
  show "formulan+1" by ...
qed
```

It proves $formula_0 \implies formula_{n+1}$ (provided each proof step succeeds). The intermediate **have** statements are merely stepping stones on the way towards the **show** statement that proves the actual goal. In more detail, this is the Isar core syntax:

```
proof = by method
      | proof [method] step* qed

step = fix variables
      | assume proposition
      | [from fact+] (have | show) proposition proof

proposition = [name:] "formula"

fact = name | ...
```

A proof can either be an atomic **by** with a single proof method which must finish off the statement being proved, for example *auto*, or it can be a **proof-qed** block of multiple steps. Such a block can optionally begin with a proof method that indicates how to start off the proof, e.g., (*induction xs*).

A step either assumes a proposition or states a proposition together with its proof. The optional **from** clause indicates which facts are to be used in the proof. Intermediate propositions are stated with **have**, the overall goal is stated with **show**. A step can also introduce new local variables with **fix**. Logically, **fix** introduces \bigwedge -quantified variables, **assume** introduces the assumption of an implication (\implies) and **have/show** introduce the conclusion.

Propositions are optionally named formulas. These names can be referred to in later **from** clauses. In the simplest case, a fact is such a name. But facts can also be composed with *OF* and *of* as shown in Section 4.4.4 — hence the ... in the above grammar. Note that assumptions, intermediate **have** statements and global lemmas all have the same status and are thus collectively referred to as **facts**.

Fact names can stand for whole lists of facts. For example, if *f* is defined by command **fun**, *f.simps* refers to the whole list of recursion equations defining *f*. Individual facts can be selected by writing *f.simps*(2), whole sublists by writing *f.simps*(2–4).

5.1 Isar by Example

We show a number of proofs of Cantor's theorem that a function from a set to its powerset cannot be surjective, illustrating various features of Isar. The constant *surj* is predefined.

```
lemma "¬ surj (f :: 'a ⇒ 'a set)"
proof
  assume 0: "surj f"
  from 0 have 1: "∀ A. ∃ a. A = f a" by (simp add: surj_def)
  from 1 have 2: "∃ a. {x. x ∉ f x} = f a" by blast
  from 2 show "False" by blast
qed
```

The **proof** command lacks an explicit method by which to perform the proof. In such cases Isabelle tries to use some standard introduction rule, in the above case for \neg :

$$\frac{P \implies False}{\neg P}$$

In order to prove $\neg P$, assume *P* and show *False*. Thus we may assume *surj f*. The proof shows that names of propositions may be (single!) digits —

meaningful names are hard to invent and are often not necessary. Both **have** steps are obvious. The second one introduces the diagonal set $\{x. x \notin f x\}$, the key idea in the proof. If you wonder why 2 directly implies *False*: from 2 it follows that $(a \notin f a) = (a \in f a)$.

5.1.1.1 *this*, then, with, hence, thus, using

Labels should be avoided. They interrupt the flow of the reader who has to scan the context for the point where the label was introduced. Ideally, the proof is a linear flow, where the output of one step becomes the input of the next step, piping the previously proved fact into the next proof, like in a UNIX pipe. In such cases the predefined name *this* can be used to refer to the proposition proved in the previous step. This allows us to eliminate all labels from our proof (we suppress the **lemma** statement):

```
proof
  assume "surj f"
  from this have "∃ a. {x. x ∉ f x} = f a" by(auto simp: surj_def)
  from this show "False" by blast
qed
```

We have also taken the opportunity to compress the two **have** steps into one.

To compact the text further, Isar has a few convenient abbreviations:

```
then  = from this
thus  = then show
hence = then have
```

With the help of these abbreviations the proof becomes

```
proof
  assume "surj f"
  hence "∃ a. {x. x ∉ f x} = f a" by(auto simp: surj_def)
  thus "False" by blast
qed
```

There are two further linguistic variations:

```
(have|show) prop using facts = from facts (have|show) prop
               with facts = from facts this
```

The **using** idiom de-emphasizes the used facts by moving them behind the proposition.

5.1.1.2 Structured Lemma Statements: fixes, assumes, shows

Lemmas can also be stated in a more structured fashion. To demonstrate this feature with Cantor's theorem, we rephrase $\neg \text{surj } f$ a little:

lemma

```
fixes f :: "'a ⇒ 'a set"
assumes s: "surj f"
shows "False"
```

The optional **fixes** part allows you to state the types of variables up front rather than by decorating one of their occurrences in the formula with a type constraint. The key advantage of the structured format is the **assumes** part that allows you to name each assumption; multiple assumptions can be separated by **and**. The **shows** part gives the goal. The actual theorem that will come out of the proof is $\text{surj } f \implies \text{False}$, but during the proof the assumption $\text{surj } f$ is available under the name s like any other fact.

proof —

```
have "∃ a. {x. x ∉ f x} = f a" using s by(auto simp: surj_def)
thus "False" by blast
```

qed

! Note the hyphen after the **proof** command. It is the null method that does nothing to the goal. Leaving it out would be asking Isabelle to try some suitable introduction rule on the goal False — but there is no such rule and **proof** would fail.

In the **have** step the assumption $\text{surj } f$ is now referenced by its name s . The duplication of $\text{surj } f$ in the above proofs (once in the statement of the lemma, once in its proof) has been eliminated.

Stating a lemma with **assumes-shows** implicitly introduces the name *assms* that stands for the list of all assumptions. You can refer to individual assumptions by *assms*(1), *assms*(2), etc., thus obviating the need to name them individually.

5.2 Proof Patterns

We show a number of important basic proof patterns. Many of them arise from the rules of natural deduction that are applied by **proof** by default. The patterns are phrased in terms of **show** but work for **have** and **lemma**, too.

We start with two forms of **case analysis**: starting from a formula P we have the two cases P and $\neg P$, and starting from a fact $P \vee Q$ we have the two cases P and Q :

<pre> show "R" proof cases assume "P" ⋮ show "R" <proof> next assume "¬ P" ⋮ show "R" <proof> qed </pre>	<pre> have "P ∨ Q" <proof> then show "R" proof assume "P" ⋮ show "R" <proof> next assume "Q" ⋮ show "R" <proof> qed </pre>
--	--

How to prove a logical equivalence:

```

show "P ↔ Q"
proof
  assume "P"
  ⋮
  show "Q" <proof>
next
  assume "Q"
  ⋮
  show "P" <proof>
qed

```

Proofs by contradiction (*ccontr* stands for “classical contradiction”):

<pre> show "¬ P" proof assume "P" ⋮ show "False" <proof> qed </pre>	<pre> show "P" proof (rule ccontr) assume "¬ P" ⋮ show "False" <proof> qed </pre>
---	---

How to prove quantified formulas:

<pre> show "∀ x. P(x)" proof fix x ⋮ show "P(x)" <proof> qed </pre>	<pre> show "∃ x. P(x)" proof ⋮ show "P(witness)" <proof> qed </pre>
---	---

In the proof of $\forall x. P(x)$, the step `fix x` introduces a locally fixed variable x into the subproof, the proverbial “arbitrary but fixed value”. Instead of x

we could have chosen any name in the subproof. In the proof of $\exists x. P(x)$, *witness* is some arbitrary term for which we can prove that it satisfies P .

How to reason forward from $\exists x. P(x)$:

```
have " $\exists x. P(x)$ " <proof>
then obtain  $x$  where  $p$ : " $P(x)$ " by blast
```

After the **obtain** step, x (we could have chosen any name) is a fixed local variable, and p is the name of the fact $P(x)$. This pattern works for one or more x . As an example of the **obtain** command, here is the proof of Cantor's theorem in more detail:

```
lemma " $\neg \text{surj}(f :: 'a \Rightarrow 'a \text{ set})$ "
proof
  assume "surj f"
  hence " $\exists a. \{x. x \notin f x\} = f a$ " by(auto simp: surj_def)
  then obtain  $a$  where " $\{x. x \notin f x\} = f a$ " by blast
  hence " $a \notin f a \longleftrightarrow a \in f a$ " by blast
  thus "False" by blast
qed
```

Finally, how to prove set equality and subset relationship:

```
show " $A = B$ "          show " $A \subseteq B$ "
proof                  proof
  show " $A \subseteq B$ " <proof>    fix  $x$ 
next                      assume " $x \in A$ "
  show " $B \subseteq A$ " <proof>    :
qed                        show " $x \in B$ " <proof>
                           qed
```

5.3 Streamlining Proofs

5.3.1 Pattern Matching and Quotations

In the proof patterns shown above, formulas are often duplicated. This can make the text harder to read, write and maintain. Pattern matching is an abbreviation mechanism to avoid such duplication. Writing

```
show formula (is pattern)
```

matches the pattern against the formula, thus instantiating the unknowns in the pattern for later use. As an example, consider the proof pattern for \longleftrightarrow :

```
show "formula1  $\longleftrightarrow$  formula2" (is "?L  $\longleftrightarrow$  ?R")
```

```

proof
  assume "?L"
  ⋮
  show "?R" <proof>
next
  assume "?R"
  ⋮
  show "?L" <proof>
qed

```

Instead of duplicating $formula_i$ in the text, we introduce the two abbreviations $?L$ and $?R$ by pattern matching. Pattern matching works wherever a formula is stated, in particular with **have** and **lemma**.

The unknown $?thesis$ is implicitly matched against any goal stated by **lemma** or **show**. Here is a typical example:

```

lemma "formula"
proof —
  ⋮
  show ?thesis <proof>
qed

```

Unknowns can also be instantiated with **let** commands

```
let ?t = "some-big-term"
```

Later proof steps can refer to $?t$:

```
have "... ?t ..."
```



Names of facts are introduced with *name*: and refer to proved theorems. Unknowns $?X$ refer to terms or formulas.

Although abbreviations shorten the text, the reader needs to remember what they stand for. Similarly for names of facts. Names like 1, 2 and 3 are not helpful and should only be used in short proofs. For longer proofs, descriptive names are better. But look at this example:

```

have x_gr_0: "x > 0"
⋮
from x_gr_0 ...

```

The name is longer than the fact it stands for! Short facts do not need names; one can refer to them easily by quoting them:

```

have "x > 0"
:
from ⟨x > 0⟩ ...

```

The outside quotes in $\langle x > 0 \rangle$ are the standard renderings of the symbols $\backslash\langle\text{open}\rangle$ and $\backslash\langle\text{close}\rangle$. They refer to the fact not by name but “by value”.

5.3.2 moreover

Sometimes one needs a number of facts to enable some deduction. Of course one can name these facts individually, as shown on the right, but one can also combine them with **moreover**, as shown on the left:

have "P ₁ " ⟨proof⟩	have lab ₁ : "P ₁ " ⟨proof⟩
moreover have "P ₂ " ⟨proof⟩	have lab ₂ : "P ₂ " ⟨proof⟩
moreover	:
:	have lab _n : "P _n " ⟨proof⟩
moreover have "P _n " ⟨proof⟩	from lab ₁ lab ₂ ...
ultimately have "P" ⟨proof⟩	have "P" ⟨proof⟩

The **moreover** version is no shorter but expresses the structure a bit more clearly and avoids new names.

5.3.3 Local Lemmas

Sometimes one would like to prove some lemma locally within a proof, a lemma that shares the current context of assumptions but that has its own assumptions and is generalized over its locally fixed variables at the end. This is simply an extension of the basic **have** construct:

```

have B if name: A1 ... Am for x1 ... xn
⟨proof⟩

```

proves $\llbracket A_1; \dots; A_m \rrbracket \implies B$ where all x_i have been replaced by unknowns $?x_i$. As an example we prove a simple fact about divisibility on integers. The definition of *dvd* is $(b \text{ dvd } a) = (\exists k. a = b * k)$.

```

lemma fixes a b :: int assumes "b dvd (a+b)" shows "b dvd a"

```

```

proof -

```

```

  have "∃ k'. a = b*k'" if asm: "a+b = b*k" for k

```

```

  proof

```

```

    show "a = b*(k - 1)" using asm by(simp add: algebra_simps)

```

```

  qed

```

```

  then show ?thesis using assms by(auto simp add: dvd_def)

```

```

qed

```

Exercises

Exercise 5.1. Give a readable, structured proof of the following lemma:

```
lemma assumes T: "∀ x y. T x y ∨ T y x"
  and A: "∀ x y. A x y ∧ A y x ⟶ x = y"
  and TA: "∀ x y. T x y ⟶ A x y" and "A x y"
  shows "T x y"
```

Exercise 5.2. Give a readable, structured proof of the following lemma:

```
lemma "∃ ys zs. xs = ys @ zs ∧
      (length ys = length zs ∨ length ys = length zs + 1)"
```

Hint: There are predefined functions $take :: nat \Rightarrow 'a\ list \Rightarrow 'a\ list$ and $drop :: nat \Rightarrow 'a\ list \Rightarrow 'a\ list$ such that $take\ k\ [x_1, \dots] = [x_1, \dots, x_k]$ and $drop\ k\ [x_1, \dots] = [x_{k+1}, \dots]$. Let sledgehammer find and apply the relevant *take* and *drop* lemmas for you.

5.4 Case Analysis and Induction

5.4.1 Datatype Case Analysis

We have seen case analysis on formulas. Now we want to distinguish which form some term takes: is it 0 or of the form $Suc\ n$, is it \square or of the form $x \# xs$, etc. Here is a typical example proof by case analysis on the form of xs :

```
lemma "length(tl xs) = length xs - 1"
proof (cases xs)
  assume "xs = □"
  thus ?thesis by simp
next
  fix y ys assume "xs = y # ys"
  thus ?thesis by simp
qed
```

Function *tl* ("tail") is defined by $tl\ \square = \square$ and $tl\ (x21.0 \# x22.0) = x22.0$. Note that the result type of *length* is *nat* and $0 - 1 = 0$.

This proof pattern works for any term t whose type is a datatype. The goal has to be proved for each constructor C :

```
fix x1 ... xn assume "t = C x1 ... xn"
```

Each case can be written in a more compact form by means of the *case* command:

```
case (C x1 ... xn)
```

This is equivalent to the explicit **fix-assume** line but also gives the assumption " $t = C x_1 \dots x_n$ " a name: C , like the constructor. Here is the **case** version of the proof above:

```
proof (cases xs)
  case Nil
  thus ?thesis by simp
next
  case (Cons y ys)
  thus ?thesis by simp
qed
```

Remember that *Nil* and *Cons* are the alphanumeric names for \square and $\#$. The names of the assumptions are not used because they are directly piped (via **thus**) into the proof of the claim.

5.4.2 Structural Induction

We illustrate structural induction with an example based on natural numbers: the sum (\sum) of the first n natural numbers ($\{0..n::nat\}$) is equal to $n * (n + 1) \text{ div } 2$. Never mind the details, just focus on the pattern:

```
lemma "sum {0..n::nat} = n*(n+1) div 2"
proof (induction n)
  show "sum {0..0::nat} = 0*(0+1) div 2" by simp
next
  fix n assume "sum {0..n::nat} = n*(n+1) div 2"
  thus "sum {0..Suc n} = Suc n*(Suc n+1) div 2" by simp
qed
```

Except for the rewrite steps, everything is explicitly given. This makes the proof easily readable, but the duplication means it is tedious to write and maintain. Here is how pattern matching can completely avoid any duplication:

```
lemma "sum {0..n::nat} = n*(n+1) div 2" (is "?P n")
proof (induction n)
  show "?P 0" by simp
next
  fix n assume "?P n"
  thus "?P (Suc n)" by simp
qed
```

The first line introduces an abbreviation $?P\ n$ for the goal. Pattern matching $?P\ n$ with the goal instantiates $?P$ to the function $\lambda n. \sum \{0..n\} = n * (n +$

1) *div 2*. Now the proposition to be proved in the base case can be written as $?P\ 0$, the induction hypothesis as $?P\ n$, and the conclusion of the induction step as $?P(Suc\ n)$.

Induction also provides the **case** idiom that abbreviates the **fix-assume** step. The above proof becomes

```
proof (induction n)
  case 0
  show ?case by simp
next
  case (Suc n)
  thus ?case by simp
qed
```

The unknown $?case$ is set in each case to the required claim, i.e., $?P\ 0$ and $?P(Suc\ n)$ in the above proof, without requiring the user to define a $?P$. The general pattern for induction over *nat* is shown on the left-hand side:

```
show "P(n) "
proof (induction n)
  case 0
  ⋮
  show ?case <proof>
next
  case (Suc n)
  ⋮
  show ?case <proof>
qed
```

let ?case = "P(0) "

fix n assume Suc: "P(n) "

let ?case = "P(Suc n) "

On the right side you can see what the **case** command on the left stands for.

In case the goal is an implication, induction does one more thing: the proposition to be proved in each case is not the whole implication but only its conclusion; the premises of the implication are immediately made assumptions of that case. That is, if in the above proof we replace **show** $"P(n) "$ by **show** $"A(n) \implies P(n) "$ then **case 0** stands for

```
assume 0: "A(0) "
let ?case = "P(0) "
```

and **case (Suc n)** stands for

```
fix n
assume Suc: "A(n) \implies P(n) "
           "A(Suc n) "
let ?case = "P(Suc n) "
```

The list of assumptions *Suc* is actually subdivided into *Suc.IH*, the induction hypotheses (here $A(n) \implies P(n)$), and *Suc.premis*, the premises of the goal being proved (here $A(\text{Suc } n)$).

Induction works for any datatype. Proving a goal $\llbracket A_1(x); \dots; A_k(x) \rrbracket \implies P(x)$ by induction on x generates a proof obligation for each constructor C of the datatype. The command `case (C x1 ... xn)` performs the following steps:

1. `fix x1 ... xn`
2. `assume` the induction hypotheses (calling them *C.IH*) and the premises $A_i(C\ x_1 \dots x_n)$ (calling them *C.premis*) and calling the whole list *C*
3. `let ?case = "P(C x1 ... xn)"`

5.4.3 Rule Induction

Recall the inductive and recursive definitions of even numbers in [Section 4.5](#):

```
inductive ev :: "nat ⇒ bool" where
  ev0: "ev 0" |
  evSS: "ev n ⇒ ev (Suc (Suc n))"
```

```
fun evn :: "nat ⇒ bool" where
  "evn 0 = True" |
  "evn (Suc 0) = False" |
  "evn (Suc (Suc n)) = evn n"
```

We recast the proof of $ev\ n \implies evn\ n$ in Isar. The left column shows the actual proof text, the right column shows the implicit effect of the two `case` commands:

<pre>lemma "ev n ⇒ evn n" proof(induction rule: ev.induct) case ev0 show ?case by simp next case evSS thus ?case by simp qed</pre>	<pre>let ?case = "evn 0" fix n assume evSS: "ev n" "evn n" let ?case = "evn (Suc (Suc n))"</pre>
--	--

The proof resembles structural induction, but the induction rule is given explicitly and the names of the cases are the names of the rules in the inductive

definition. Let us examine the two assumptions named *evSS*: *ev n* is the premise of rule *evSS*, which we may assume because we are in the case where that rule was used; *evn n* is the induction hypothesis.

! Because each case command introduces a list of assumptions named like the case name, which is the name of a rule of the inductive definition, those rules now need to be accessed with a qualified name, here *ev.ev0* and *ev.evSS*.

In the case *evSS* of the proof above we have pretended that the system fixes a variable *n*. But unless the user provides the name *n*, the system will just invent its own name that cannot be referred to. In the above proof, we do not need to refer to it, hence we do not give it a specific name. In case one needs to refer to it one writes

```
case (evSS m)
```

like `case (Suc n)` in earlier structural inductions. The name *m* is an arbitrary choice. As a result, case *evSS* is derived from a renamed version of rule *evSS*: *ev m* \implies *ev(Suc(Suc m))*. Here is an example with a (contrived) intermediate step that refers to *m*:

```
lemma "ev n  $\implies$  evn n"
proof(induction rule: ev.induct)
  case ev0 show ?case by simp
next
  case (evSS m)
  have "evn(Suc(Suc m)) = evn m" by simp
  thus ?case using 'evn m' by blast
qed
```

In general, let *I* be a (for simplicity unary) inductively defined predicate and let the rules in the definition of *I* be called *rule*₁, ..., *rule*_{*n*}. A proof by rule induction follows this pattern:

```
show "I x  $\implies$  P x"
proof(induction rule: I.induct)
  case rule1
  :
  show ?case <proof>
next
  :
next
  case rulen
  :
  show ?case <proof>
qed
```

One can provide explicit variable names by writing `case (rulei $x_1 \dots x_k$)`, thus renaming the first k free variables in rule i to $x_1 \dots x_k$, going through rule i from left to right.

5.4.4 Assumption Naming

In any induction, `case name` sets up a list of assumptions also called *name*, which is subdivided into three parts:

name.IH contains the induction hypotheses.

name.hyps contains all the other hypotheses of this case in the induction rule. For rule inductions these are the hypotheses of rule *name*, for structural inductions these are empty.

name.premis contains the (suitably instantiated) premises of the statement being proved, i.e., the A_i when proving $\llbracket A_1; \dots; A_n \rrbracket \implies A$.



Proof method *induct* differs from *induction* only in this naming policy: *induct* does not distinguish *IH* from *hyps* but subsumes *IH* under *hyps*.

More complicated inductive proofs than the ones we have seen so far often need to refer to specific assumptions — just *name* or even *name.premis* and *name.IH* can be too unspecific. This is where the indexing of fact lists comes in handy, e.g., *name.IH*(2) or *name.premis*(1–2).

5.4.5 Rule Inversion

Rule inversion is case analysis of which rule could have been used to derive some fact. The name **rule inversion** emphasizes that we are reasoning backwards: by which rules could some given fact have been proved? For the inductive definition of *ev*, rule inversion can be summarized like this:

$$ev\ n \implies n = 0 \vee (\exists k. n = Suc\ (Suc\ k) \wedge ev\ k)$$

The realisation in Isabelle is a case analysis. A simple example is the proof that $ev\ n \implies ev\ (n - 2)$. We already went through the details informally in [Section 4.5.1](#). This is the Isar proof:

```

assume "ev n"
from this have "ev(n - 2)"
proof cases
  case ev0 thus "ev(n - 2)" by (simp add: ev.ev0)
next
  case (evSS k) thus "ev(n - 2)" by (simp add: ev.evSS)
qed

```

The key point here is that a case analysis over some inductively defined predicate is triggered by piping the given fact (here: **from** *this*) into a proof by *cases*. Let us examine the assumptions available in each case. In case *ev0* we have $n = 0$ and in case *evSS* we have $n = \text{Suc } (\text{Suc } k)$ and *ev* k . In each case the assumptions are available under the name of the case; there is no fine-grained naming schema like there is for induction.

Sometimes some rules could not have been used to derive the given fact because constructors clash. As an extreme example consider rule inversion applied to *ev* (*Suc* 0): neither rule *ev0* nor rule *evSS* can yield *ev* (*Suc* 0) because *Suc* 0 unifies neither with 0 nor with *Suc* (*Suc* n). Impossible cases do not have to be proved. Hence we can prove anything from *ev* (*Suc* 0):

```
assume "ev(Suc 0)" then have P by cases
```

That is, *ev* (*Suc* 0) is simply not provable:

```
lemma "¬ ev(Suc 0)"
```

```
proof
```

```
  assume "ev(Suc 0)" then show False by cases
```

```
qed
```

Normally not all cases will be impossible. As a simple exercise, prove that $\neg \text{ev } (\text{Suc } (\text{Suc } (\text{Suc } 0)))$.

5.4.6 Advanced Rule Induction

So far, rule induction was always applied to goals of the form $I\ x\ y\ z \implies \dots$ where I is some inductively defined predicate and x, y, z are variables. In some rare situations one needs to deal with an assumption where not all arguments r, s, t are variables:

```
lemma "I r s t ⟹ …"
```

Applying the standard form of rule induction in such a situation will lead to strange and typically unprovable goals. We can easily reduce this situation to the standard one by introducing new variables x, y, z and reformulating the goal like this:

```
lemma "I x y z ⟹ x = r ⟹ y = s ⟹ z = t ⟹ …"
```

Standard rule induction will work fine now, provided the free variables in r, s, t are generalized via *arbitrary*.

However, induction can do the above transformation for us, behind the curtains, so we never need to see the expanded version of the lemma. This is what we need to write:

```
lemma "I r s t  $\implies$  ..."
proof(induction "r" "s" "t" arbitrary: ... rule: I.induct)
```

Like for rule inversion, cases that are impossible because of constructor clashes will not show up at all. Here is a concrete example:

```
lemma "ev (Suc m)  $\implies$   $\neg$  ev m"
proof(induction "Suc m" arbitrary: m rule: ev.induct)
  fix n assume IH: " $\bigwedge m. n = \text{Suc } m \implies \neg \text{ev } m$ "
  show " $\neg$  ev (Suc n)"
  proof — contradiction
    assume "ev (Suc n)"
    thus False
  proof cases — rule inversion
    fix k assume "n = Suc k" "ev k"
    thus False using IH by auto
  qed
qed
qed
```

Remarks:

- Instead of the `case` and `?case` magic we have spelled all formulas out. This is merely for greater clarity.
- We only need to deal with one case because the `ev 0` case is impossible.
- The form of the *IH* shows us that internally the lemma was expanded as explained above: $\text{ev } x \implies x = \text{Suc } m \implies \neg \text{ev } m$.
- The goal $\neg \text{ev } (\text{Suc } n)$ may surprise. The expanded version of the lemma would suggest that we have a `fix m assume Suc (Suc n) = Suc m` and need to show $\neg \text{ev } m$. What happened is that Isabelle immediately simplified $\text{Suc } (\text{Suc } n) = \text{Suc } m$ to $\text{Suc } n = m$ and could then eliminate m . Beware of such nice surprises with this advanced form of induction.

! This advanced form of induction does not support the *IH* naming schema explained in [Section 5.4.4](#): the induction hypotheses are instead found under the name *hyps*, as they are for the simpler *induct* method.

Exercises

Exercise 5.3. Give a structured proof by rule inversion:

```
lemma assumes a: "ev (Suc (Suc n))" shows "ev n"
```

Exercise 5.4. Give a structured proof of $\neg \text{ev } (\text{Suc } (\text{Suc } (\text{Suc } 0)))$ by rule inversions. If there are no cases to be proved you can close a proof immediately with `qed`.

Exercise 5.5. Recall predicate *star* from [Section 4.5.2](#) and *iter* from [Exercise 4.4](#). Prove $\text{iter } r \ n \ x \ y \implies \text{star } r \ x \ y$ in a structured style; do not just sledgehammer each case of the required induction.

Exercise 5.6. Define a recursive function $\text{elems} :: 'a \text{ list} \Rightarrow 'a \text{ set}$ and prove $x \in \text{elems } xs \implies \exists ys \ zs. \ xs = ys @ x \# zs \wedge x \notin \text{elems } ys$.

Exercise 5.7. Extend [Exercise 4.5](#) with a function that checks if some *alpha list* is a balanced string of parentheses. More precisely, define a recursive function $\text{balanced} :: \text{nat} \Rightarrow \text{alpha list} \Rightarrow \text{bool}$ such that $\text{balanced } n \ w$ is true iff (informally) $S \ (a^n @ w)$. Formally, prove that $\text{balanced } n \ w = S \ (\text{replicate } n \ a @ w)$ where $\text{replicate} :: \text{nat} \Rightarrow 'a \Rightarrow 'a \text{ list}$ is predefined and $\text{replicate } n \ x$ yields the list $[x, \dots, x]$ of length n .

Part II

Semantics

It is all very well to aim for a more “abstract” and a “cleaner” approach to semantics, but if the plan is to be any good, the operational aspects cannot be completely ignored.

Dana Scott [82]

Introduction

Welcome to the second part of this book. In the first part you have mastered the basics of interactive theorem proving and are now able to navigate the depths of higher-order logic. In this second part, we put these skills to concrete use in the semantics of programming languages.

Why formal semantics? Because there is no alternative when it comes to an unambiguous foundation of what is at the heart of computer science: programs. A formal semantics provides the much needed foundation for their work not just to programmers who want to reason about their programs but also to developers of tools (e.g., compilers, refactoring tools and program analysers) and ultimately to language designers themselves.

This second part of the book is based entirely on a small imperative language called IMP. IMP is our vehicle for showing not just how to formally define the semantics of a programming language but also how to use the semantics to reason about the language, about the behaviour of programs, and about program analyses. Specifically, we examine the following topics: operational semantics, compiler correctness, type systems, program analysis, denotational semantics, Hoare logic, and abstract interpretation.

IMP is a minimal language, with just enough expressive power to be Turing-complete. It does not come with the bells and whistles of a real, mainstream programming language. This is by design: our aim is to show the essence of the techniques, analyses, and transformations that we study in this book, not to get bogged down in detail and sheer volume. This does not mean that formal, machine-checked semantics cannot scale to mainstream languages such as C or Java or more complex features such as object orientation. In fact, this is where proof assistants shine: the ability to automatically check a large amount of error-prone detail relieves the tedium of any sizeable formalization. At the end of each chapter we give pointers to further reading and recent successful applications of the techniques we discuss.

Isabelle

Although the reader will get the most out of this second part of the book if she has studied Isabelle before, it can be read without knowledge of interactive theorem proving. We describe all proofs in detail, but in contrast to the first part of the book, we describe them informally. Nevertheless, everything has been formalized and proved in Isabelle. All these theories can be found in the directory `src/HOL/IMP` of the Isabelle distribution. HTML versions are available online at <http://isabelle.in.tum.de/library/HOL/HOL-IMP>. Many section headings have a link to the corresponding Isabelle theory attached that looks like this: [thy](#). [Appendix C](#) contains a table that shows which sections are based on which theories. When building upon any of those theories, for example when solving an exercise, the **imports** section needs to include `"~/src/HOL/IMP/T"` where T is the name of the required theory.

In this second part of the book we simplify the Isabelle syntax in two minor respects to improve readability:

- We no longer enclose types and terms in quotation marks.
- We no longer separate clauses in function definitions or inductive definitions with `"|"`.

Finally, a note on terminology: We call a proof “automatic” if it requires only a single invocation of a basic Isabelle proof method like *simp*, *auto*, *blast* or *metis*, possibly modified with specific lemmas. Inductions are not automatic, although each case can be.

IMP: A Simple Imperative Language

To talk about semantics, we first need a programming language. This chapter defines one: a minimalistic imperative programming language called **IMP**.

The main aim of this chapter is to introduce the concepts of commands and their abstract syntax, and to use them to illustrate two styles of defining the semantics of a programming language: big-step and small-step operational semantics. Our first larger theorem about IMP will be the equivalence of these two definitions of its semantics. As a smaller concrete example, we will apply our semantics to the concept of program equivalence.

7.1 IMP Commands thy

Before we jump into any formalization or define the abstract syntax of commands, we need to determine which constructs the language IMP should contain. The basic constraints are given by our aim to formalize the semantics of an imperative language and to keep things simple. For an imperative language, we will want the basics such as assignments, sequential composition (semicolon), and conditionals (*IF*). To make it Turing-complete, we will want to include *WHILE* loops. To be able to express other syntactic forms, such as an *IF* without an *ELSE* branch, we also include the *SKIP* command that does nothing. The right-hand side of variable assignments will be the arithmetic expressions already defined in [Chapter 3](#), and the conditions in *IF* and *WHILE* will be the boolean expressions defined in the same chapter. A program is simply a, possibly complex, command in this language.

We have already seen the formalization of expressions and their semantics in [Chapter 3](#). The abstract syntax of commands is:

```

datatype com = SKIP
            | Assign vname aexp
            | Seq com com
            | If bexp com com
            | While bexp com

```

In the definitions, proofs, and examples further along in this book, we will often want to refer to concrete program fragments. To make such fragments more readable, we also introduce concrete infix syntax in Isabelle for the four compound constructors of the *com* datatype. The term *Assign x a* for instance can be written as $x ::= a$, the term *Seq c₁ c₂* as $c_1;; c_2$, the term *If b c₁ c₂* as *IF b THEN c₁ ELSE c₂*, and the while loop *While b c* as *WHILE b DO c*. Sequential composition is denoted by “;;” to distinguish it from the “;” that separates assumptions in the $\llbracket \dots \rrbracket$ notation. Nevertheless we still pronounce “;;” as “semicolon”.

Example 7.1. The following is an example IMP program with two assignments.

$$''x'' ::= Plus (V ''y'') (N 1);; ''y'' ::= N 2$$

We have not defined its meaning yet, but the intention is that it assigns the value of variable *y* incremented by 1 to the variable *x*, and afterwards sets *y* to 2. In a more conventional concrete programming language syntax, we would have written

$$x := y + 1; y := 2$$

We will occasionally use this more compact style for examples in the text, with the obvious translation into the formal form.

! We write concrete variable names as strings enclosed in double quotes, just as in the arithmetic expressions in [Chapter 3](#). Examples are $V ''x''$ or $''x'' ::= exp$. If we write $V x$ instead, *x* is a logical variable for the name of the program variable. That is, in $x ::= exp$, the *x* stands for any concrete name $''x''$, $''y''$, and so on, the same way *exp* stands for any arithmetic expression.

! In our language, semicolon associates to the left. This means $c_1;; c_2;; c_3 = (c_1;; c_2);; c_3$. We will later prove that semantically it does not matter whether semicolon associates to the left or to the right.

The compound commands *IF* and *WHILE* bind stronger than semicolon. That means $WHILE b DO c_1;; c_2 = (WHILE b DO c_1);; c_2$.

While more convenient than writing abstract syntax trees, as we have seen in the example, even the more concrete Isabelle notation above is occasionally somewhat cumbersome to use. This is not a fundamental restriction of the theorem prover or of mechanised semantics. If one were interested in a

more traditional concrete syntax for IMP, or if one were to formalize a larger, more realistic language, one could write separate parsing/printing ML code that integrates with Isabelle and implements the concrete syntax of the language. This is usually only worth the effort when the emphasis is on program verification as opposed to meta-theorems about the programming language.

A larger language may also contain a so-called syntactic de-sugaring phase, where more complex constructs in the language are transformed into simple core concepts. For instance, our IMP language does not have syntax for Java style for-loops, or repeat ... until loops. For our purpose of analysing programming language semantics in general these concepts add nothing new, but for a full language formalization they would be required. De-sugaring would take the for-loop and do ... while syntax and translate it into the standard *WHILE* loops that IMP supports. Therefore definitions and theorems about the core language only need to worry about one type of loop, while still supporting the full richness of a larger language. This significantly reduces proof size and effort for the theorems that we discuss in this book.

7.2 Big-Step Semantics thy

In the previous section we defined the abstract syntax of the IMP language. In this section we show its semantics. More precisely, we will use a **big-step operational semantics** to give meaning to commands.

In an **operational semantics** setting, the aim is to capture the meaning of a program as a relation that describes *how* a program executes. Other styles of semantics may be concerned with assigning mathematical structures as meanings to programs, e.g., in the so-called denotational style in [Chapter 11](#), or they may be interested in capturing the meaning of programs by describing how to reason about them, e.g., in the axiomatic style in [Chapter 12](#).

7.2.1 Definition

In big-step operational semantics, the relation to be defined is between program, initial state, and final state. Intermediate states during the execution of the program are not visible in the relation. Although the inductive rules that define the semantics will tell us how the execution proceeds internally, the relation itself looks as if the whole program was executed in one big step.

We formalize the big-step execution relation in the theorem prover as a ternary predicate *big_step*. The intended meaning of *big_step c s t* is that execution of command *c* starting in state *s* terminates in state *t*. To display such predicates in a more intuitive form, we use Isabelle's syntax mechanism and the more conventional notation $(c, s) \Rightarrow t$ instead of *big_step c s t*.

$\frac{}{(SKIP, s) \Rightarrow s}$	<i>Skip</i>	$\frac{}{(x ::= a, s) \Rightarrow s(x := aval\ a\ s)}$	<i>Assign</i>
$\frac{(c_1, s_1) \Rightarrow s_2 \quad (c_2, s_2) \Rightarrow s_3}{(c_1;; c_2, s_1) \Rightarrow s_3}$	<i>Seq</i>		
$\frac{bval\ b\ s \quad (c_1, s) \Rightarrow t}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \Rightarrow t}$	<i>IfTrue</i>		
$\frac{\neg bval\ b\ s \quad (c_2, s) \Rightarrow t}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \Rightarrow t}$	<i>IfFalse</i>		
$\frac{\neg bval\ b\ s}{(WHILE\ b\ DO\ c, s) \Rightarrow s}$	<i>WhileFalse</i>		
$\frac{bval\ b\ s_1 \quad (c, s_1) \Rightarrow s_2 \quad (WHILE\ b\ DO\ c, s_2) \Rightarrow s_3}{(WHILE\ b\ DO\ c, s_1) \Rightarrow s_3}$	<i>WhileTrue</i>		

Fig. 7.1. The big-step rules of IMP

It remains for us to define which c , s and s' this predicate is made up of. Given the recursive nature of the abstract syntax, it will not come as a surprise that our choice is an inductive definition. Figure 7.1 shows its rules. Predicates such as $(c, s) \Rightarrow t$ that are defined by a set of rules are often also called **judgements**, because the rules decide for which parameters the predicate is true. However, there is nothing special about them, they are merely ordinary inductively defined predicates.

Let us go through the rules and rephrase each of them in natural language:

- If the command is *SKIP*, the initial and final state must be the same.
- If the command is an assignment $x ::= a$ and the initial state is s , then the final state is the same state s where the value of variable x is replaced by the evaluation of the expression a in state s .
- If the command is a sequential composition, rule *Seq* says the combined command $c_1;; c_2$ started in s_1 terminates in s_3 if the the first command started in s_1 terminates in some intermediate state s_2 and c_2 takes this s_2 to s_3 .
- The conditional is the first command that has two rules, depending on the value of its boolean expression in the current state s . If that value is *True*, then the *IfTrue* rule says that the execution terminates in the same state t that the command c_1 terminates in if started in s . The *IfFalse* rule does the same for the command c_2 in the *False* case.
- *WHILE* loops are slightly more interesting. If the condition evaluates to *False*, the whole loop is skipped, which is expressed in rule *WhileFalse*. However, if the condition evaluates to *True* in state s_1 and the body c of the loop takes this state s_1 to some intermediate state s_2 , and if the


```

inductive
  big_step :: com × state ⇒ state ⇒ bool (infix ⇒ 55)
where
  Skip: (SKIP, s) ⇒ s |
  Assign: (x ::= a, s) ⇒ s(x := aval a s) |
  Seq: [| (c1, s1) ⇒ s2; (c2, s2) ⇒ s3 |] ⇒ (c1;; c2, s1) ⇒ s3 |
  IfTrue: [| bval b s; (c1, s) ⇒ t |] ⇒ (IF b THEN c1 ELSE c2, s) ⇒ t |
  IfFalse: [| ¬bval b s; (c2, s) ⇒ t |] ⇒ (IF b THEN c1 ELSE c2, s) ⇒ t |
  WhileFalse: ¬bval b s ⇒ (WHILE b DO c, s) ⇒ s |
  WhileTrue:
    [| bval b s1; (c, s1) ⇒ s2; (WHILE b DO c, s2) ⇒ s3 |]
    ⇒ (WHILE b DO c, s1) ⇒ s3

```

Fig. 7.2. Isabelle definition of the big-step semantics

same *WHILE* loop started in s_2 terminates in s_3 , then the entire loop also terminates in s_3 .

Designing the right set of introduction rules for a language is not necessarily hard. The idea is to have at least one rule per syntactic construct and to add further rules when case distinctions become necessary. For each single rule, one starts with the conclusion, for instance $(c_1;; c_2, s) \Rightarrow s'$, and then constructs the assumptions of the rule by thinking about which conditions have to be true about s , s' , and the parameters of the abstract syntax constructor. In the $c_1;; c_2$ example, the parameters are c_1 and c_2 . If the assumptions collapse to an equation about s' as in the *SKIP* and $x ::= a$ cases, s' can be replaced directly.

Following the rules of Figure 7.1, the corresponding Isabelle definition shown in Figure 7.2 is straightforward, using the command *inductive* (see Section 4.5). The type of *big_step* is $com \times state \Rightarrow state \Rightarrow bool$ rather than the canonical $com \Rightarrow state \Rightarrow state \Rightarrow bool$ merely because that permits us to introduce the concrete syntax $(_,_) \Rightarrow _$ by declaring the transition arrow “ \Rightarrow ” as an infix symbol.

The striking similarity between the rules in the Isabelle definition in Figure 7.2 and the rules in Figure 7.1 is no accident: Figure 7.1 is generated automatically from the Isabelle definition with the help of Isabelle’s \LaTeX pretty-printing facility. In the future we will display inductive definitions only in their pretty-printed form. The interested reader will find the full details in the accompanying Isabelle theories.

7.2.2 Deriving IMP Executions

Figure 7.3 shows a so-called **derivation tree**, i.e., a composition of the rules from Figure 7.1 that visualizes a big-step execution: we are executing the

$$\frac{\frac{}{(''x'' ::= N\ 5, s) \Rightarrow s(''x'' := 5)} \quad \frac{}{(''y'' ::= V\ ''x'', s(''x'' := 5)) \Rightarrow s'}}{(''x'' ::= N\ 5;; ''y'' ::= V\ ''x'', s) \Rightarrow s'}$$

where $s' = s(''x'' := 5, ''y'' := 5)$

Fig. 7.3. Derivation tree for execution of an IMP program

command $''x'' ::= N\ 5;; ''y'' ::= V\ ''x''$, starting it in an arbitrary state s . Our claim is that at the end of this execution, we get the same state s , but with both x and y set to 5. We construct the derivation tree from its root, the bottom of Figure 7.3, starting with the *Seq* rule, which gives us two obligations, one for each assignment. Working on $''x'' ::= N\ 5$ first, we can conclude via the *Assign* rule that it results in the state $s(''x'' := 5)$. We feed this intermediate state into the execution of the second assignment, and again with the assignment rule complete the derivation tree. In general, a derivation tree consists of rule applications at each node and of applications of axioms (rules without premises) at the leaves.

We can conduct the same kind of argument in the theorem prover. The following is the example from Figure 7.3 in Isabelle. Instead of telling the prover what the result state is, we state the lemma with a schematic variable and let Isabelle compute its value as the proof progresses.

```

schematic_goal ex: (''x'' ::= N 5;; ''y'' ::= V ''x'', s) ⇒ ?t
  apply(rule Seq)
  apply(rule Assign)
  apply simp
  apply(rule Assign)
  done

```

After the proof is finished, Isabelle instantiates the lemma statement, and after simplification we get the expected $(''x'' ::= N\ 5;; ''y'' ::= V\ ''x'', s) \Rightarrow s(''x'' := 5, ''y'' := 5)$.

We could use this style of lemma to execute IMP programs symbolically. However, a more convenient way to execute the big-step rules is to use Isabelle's code generator. The following command tells it to generate code for the predicate \Rightarrow and thus make the predicate available in the **values** command, which is similar to **value**, but works on inductive definitions and computes a set of possible results.

```
code_pred big_step .
```

We could now write

```
values {t. (SKIP, λ_. 0) ⇒ t}
```

but this only shows us $\{_ \}$, i.e., that the result is a set containing one element. Functions cannot always easily be printed, but lists can be, so we just ask for the values of a list of variables we are interested in, using the set-comprehension notation introduced in [Section 4.2](#):

`values {map t ['x'', 'y''] | t. ('x'' ::= N 2, λ_. 0) ⇒ t}`

This has the result $\{[2,0]\}$.

This section showed us how to construct program derivations and how to execute small IMP programs according to the big-step semantics. In the next section, we instead deconstruct executions that we know have happened and analyse all possible ways we could have gotten there.

7.2.3 Rule Inversion

What can we conclude from $(SKIP, s) \Rightarrow t$? Clearly $t = s$. This is an example of rule inversion which we had discussed previously in [Section 5.4.5](#). It is a consequence of the fact that an inductively defined predicate is only true if the rules force it to be true, i.e., only if there is some derivation tree for it.

Inversion of the rules for big-step semantics tells us what we can infer from $(c, s) \Rightarrow t$. For the different commands we obtain the following inverted rules:

$$\begin{aligned} (SKIP, s) \Rightarrow t &\implies t = s \\ (x ::= a, s) \Rightarrow t &\implies t = s(x := \text{aval } a \ s) \\ (c_1;; c_2, s_1) \Rightarrow s_3 &\implies \exists s_2. (c_1, s_1) \Rightarrow s_2 \wedge (c_2, s_2) \Rightarrow s_3 \\ (IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \Rightarrow t &\implies \\ bval\ b\ s \wedge (c_1, s) \Rightarrow t \vee \neg bval\ b\ s \wedge (c_2, s) \Rightarrow t \\ (WHILE\ b\ DO\ c, s) \Rightarrow t &\implies \\ \neg bval\ b\ s \wedge t = s \vee \\ bval\ b\ s \wedge (\exists s'. (c, s) \Rightarrow s' \wedge (WHILE\ b\ DO\ c, s') \Rightarrow t) \end{aligned}$$

As an example, we paraphrase the final implication: if $(WHILE\ b\ DO\ c, s) \Rightarrow t$ then either b is false and $t = s$, i.e., rule *WhileFalse* was used, or b is true and there is some intermediate state s' such that $(c, s) \Rightarrow s'$ and $(WHILE\ b\ DO\ c, s') \Rightarrow t$, i.e., rule *WhileTrue* was used.

These inverted rules can be proved automatically by Isabelle from the original rules. Moreover, proof methods like *auto* and *blast* can be instructed to use both the introduction and the inverted rules automatically during proof search. For details see theory *Big_Step*.

One can go one step further and combine the above inverted rules with the original rules to obtain equivalences rather than implications, for example

$$(c_1;; c_2, s_1) \Rightarrow s_3 \iff (\exists s_2. (c_1, s_1) \Rightarrow s_2 \wedge (c_2, s_2) \Rightarrow s_3)$$

Every \Rightarrow in the inverted rules can be turned into \iff because the \Leftarrow direction follows from the original rules.

As an example of the two proof techniques in this and the previous section consider the following lemma. It states that the syntactic associativity of semicolon has no semantic effect. We get the same result, no matter if we group semicolons to the left or to the right.

Lemma 7.2. $(c_1;; c_2;; c_3, s) \Rightarrow s' \iff (c_1;; (c_2;; c_3), s) \Rightarrow s'$

Proof. We show each direction separately. Consider first the execution where the semicolons are grouped to the left: $((c_1;; c_2); c_3, s) \Rightarrow s'$. By rule inversion we can decompose this execution twice and obtain the intermediate states s_1 and s_2 such that $(c_1, s) \Rightarrow s_1$, as well as $(c_2, s_1) \Rightarrow s_2$ and $(c_3, s_2) \Rightarrow s'$. From this, we can construct a derivation for $(c_1;; (c_2;; c_3), s) \Rightarrow s'$ by first concluding $(c_2;; c_3, s_1) \Rightarrow s'$ with the *Seq* rule and then using the *Seq* rule again, this time on c_1 , to arrive at the final result. The other direction is analogous. \square

7.2.4 Equivalence of Commands

In the previous section we have applied rule inversion and introduction rules of the big-step semantics to show equivalence between two particular IMP commands. In this section, we define semantic equivalence as a concept in its own right.

We call two commands c and c' **equivalent** w.r.t. the big-step semantics when *c started in s terminates in t iff c' started in s also terminates in t* . Formally, we define it as an abbreviation:

abbreviation

$equiv_c :: com \Rightarrow com \Rightarrow bool$ (infix \sim 50) **where**
 $c \sim c' \equiv (\forall s\ t. (c, s) \Rightarrow t \iff (c', s) \Rightarrow t)$



Note that the \sim symbol in this definition is not the standard tilde \sim , but the symbol `\sim` instead.

Experimenting with this concept, we see that Isabelle manages to prove many simple equivalences automatically. One example is the unfolding of while loops:

Lemma 7.3.

WHILE b DO $c \sim IF\ b\ THEN\ c;;\ WHILE\ b\ DO\ c\ ELSE\ SKIP$

Another example is a trivial contraction of *IF*:

Lemma 7.4. *IF b THEN c ELSE $c \sim c$*

Of course not all equivalence properties are trivial. For example, the congruence property

Lemma 7.5. $c \sim c' \implies \text{WHILE } b \text{ DO } c \sim \text{WHILE } b \text{ DO } c'$

is a corollary of

Lemma 7.6.

$\llbracket (\text{WHILE } b \text{ DO } c, s) \Rightarrow t; c \sim c' \rrbracket \implies (\text{WHILE } b \text{ DO } c', s) \Rightarrow t$

This lemma needs the third main proof technique for inductive definitions: rule induction. We covered rule induction in [Section 4.5.1](#). Recall that for the big-step semantics, rule induction applies to properties of the form $(c, s) \Rightarrow s' \implies P \ c \ s \ s'$. To prove statements of this kind, we need to consider one case for each introduction rule, and we are allowed to assume P as an induction hypothesis for each occurrence of the inductive relation \Rightarrow in the assumptions of the respective introduction rule. The proof of [Lemma 7.6](#) requires the advanced form of rule induction described in [Section 5.4.6](#) because the big-step premise in the lemma involves not just variables but the proper term $\text{WHILE } b \text{ DO } c$.

This concept of semantic equivalence also has nice algebraic properties. For instance, it forms a so-called equivalence relation.

Definition 7.7. *A relation R is called an equivalence relation iff it is*

reflexive: $R \ x \ x$,
symmetric: $R \ x \ y \implies R \ y \ x$, and
transitive: $\llbracket R \ x \ y; R \ y \ z \rrbracket \implies R \ x \ z$.

Equivalence relations can be used to partition a set into sets of equivalent elements — in this case, commands that are semantically equivalent belong to the same partition. The standard equality $=$ can be seen as the most fine-grained equivalence relation for a given set.

Lemma 7.8. *The semantic equivalence \sim is an equivalence relation. It is reflexive ($c \sim c$), symmetric ($c \sim c' \implies c' \sim c$), and transitive ($\llbracket c \sim c'; c' \sim c'' \rrbracket \implies c \sim c''$).*

Proof. All three properties are proved automatically. □

Our relation \sim is also a so-called **congruence** on the syntax of commands: it respects the structure of commands — if all sub-commands are equivalent, so will be the compound command. This is why we called [Lemma 7.5](#) a congruence property: it establishes that \sim is a congruence relation w.r.t. WHILE . We can easily prove further such rules for semicolon and IF .

We have used the concept of semantic equivalence in this section as a first example of how semantics can be useful: to prove that two programs always have the same behaviour. An important example of where such equivalences are used in practice is the transformation of programs in compiler optimizations, some of which we will show in later chapters of this book.

7.2.5 Execution in IMP is Deterministic

So far, we have proved properties about particular IMP commands and we have introduced the concept of semantic equivalence. We have not yet investigated properties of the language itself. One such property is whether IMP is **deterministic** or not. A language is called deterministic if, for every command and every start state, there is at most one possible final state. Conversely, a language is called **non-deterministic** if it admits multiple final states. Having defined the semantics of the language as a relation, it is not immediately obvious if execution in this language is deterministic or not.

Formally, a language is deterministic if any two executions of the same command from the same initial state will always arrive in the same final state. The following lemma expresses this in Isabelle.

Lemma 7.9 (IMP is deterministic).

$$\llbracket (c, s) \Rightarrow t; (c, s) \Rightarrow t' \rrbracket \implies t' = t$$

Proof. The proof is by induction on the big-step semantics. With our inversion and introduction rules from above, each case is solved automatically by Isabelle. Note that the automation in this proof is not completely obvious. Merely using the proof method `auto` after the induction for instance leads to non-termination, but the backtracking capabilities of `blast` manage to solve each subgoal. Experimenting with different automated methods is encouraged if the standard ones fail. \square

While the above proof is nice for showing off Isabelle’s proof automation, it does not give much insight into why the property is true. [Figure 7.4](#) shows an Isar proof that expands the steps of the only interesting case and omits the boring cases using automation. This is much closer to a blackboard presentation.

So far, we have defined the big-step semantics of IMP, we have explored the proof principles of derivation trees, rule inversion, and rule induction in the context of the big-step semantics, and we have explored semantic equivalence as well as determinism of the language. In the next section we will look at a different way of defining the semantics of IMP.

```

theorem
   $(c,s) \Rightarrow t \implies (c,s) \Rightarrow t' \implies t' = t$ 
proof (induction arbitrary:  $t'$  rule: big_step.induct)
  — the only interesting case, WhileTrue:
  fix  $b\ c\ s\ s_1\ t\ t'$ 
  — The assumptions of the rule:
  assume  $bval\ b\ s$  and  $(c,s) \Rightarrow s_1$  and  $(WHILE\ b\ DO\ c,s_1) \Rightarrow t$ 
  — Ind.Hyp; note the  $\wedge$  because of arbitrary:
  assume  $IHc: \bigwedge t'. (c,s) \Rightarrow t' \implies t' = s_1$ 
  assume  $IHw: \bigwedge t'. (WHILE\ b\ DO\ c,s_1) \Rightarrow t' \implies t' = t$ 
  — Premise of implication:
  assume  $(WHILE\ b\ DO\ c,s) \Rightarrow t'$ 
  with  $\langle bval\ b\ s \rangle$  obtain  $s'_1$  where
     $c: (c,s) \Rightarrow s'_1$  and
     $w: (WHILE\ b\ DO\ c,s'_1) \Rightarrow t'$ 
  by auto
  from  $c\ IHc$  have  $s'_1 = s_1$  by blast
  with  $w\ IHw$  show  $t' = t$  by blast
qed blast+ — prove the rest automatically

```

Fig. 7.4. IMP is deterministic

7.3 Small-Step Semantics thy

The big-step semantics executes a program from an initial to the final state in one big step. Short of inspecting the derivation tree of big-step introduction rules, it does not allow us to explicitly observe intermediate execution states. That is the purpose of a small-step semantics.

Small-step semantics lets us explicitly observe partial executions and make formal statements about them. This enables us, for instance, to talk about the interleaved, concurrent execution of multiple programs. The main idea for representing a partial execution is to introduce the concept of how far execution has progressed in the program. There are many ways of doing this. Traditionally, for a high-level language like IMP, we modify the type of the big-step judgement from $com \times state \Rightarrow state \Rightarrow bool$ to something like $com \times state \Rightarrow com \times state \Rightarrow bool$. The second $com \times state$ component of the judgement is the result state of one small, atomic execution step together with a modified command that represents what still has to be executed. We call a $com \times state$ pair a **configuration** of the program, and use the command *SKIP* to indicate that execution has terminated.

The idea is easiest to understand by looking at the set of rules. They define one atomic execution step. The execution of a command is then a sequence of such steps.

$$\begin{array}{c}
\frac{}{(x ::= a, s) \rightarrow (SKIP, s(x := aval\ a\ s))} \text{Assign} \\
\\
\frac{}{(SKIP;; c_2, s) \rightarrow (c_2, s)} \text{Seq1} \quad \frac{(c_1, s) \rightarrow (c'_1, s')}{(c_1;; c_2, s) \rightarrow (c'_1;; c_2, s')} \text{Seq2} \\
\\
\frac{bval\ b\ s}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow (c_1, s)} \text{IfTrue} \\
\\
\frac{\neg\ bval\ b\ s}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow (c_2, s)} \text{IfFalse} \\
\\
\frac{}{(WHILE\ b\ DO\ c, s) \rightarrow (IF\ b\ THEN\ c;;\ WHILE\ b\ DO\ c\ ELSE\ SKIP, s)} \text{While}
\end{array}$$

Fig. 7.5. The small-step rules of IMP

Going through the rules in Figure 7.5 we see that:

- Variable assignment is an atomic step. As mentioned above, *SKIP* represents the terminated program.
- There are two rules for semicolon: either the first part is fully executed already (signified by *SKIP*), in which case we continue with the second part, or the first part can be executed further, in which case we perform the execution step and replace this first part with its reduced version.
- An *IF* reduces either to the command in the *THEN* branch or the *ELSE* branch, depending on the value of the condition.
- The final rule is for the *WHILE* loop: we define its semantics by merely unrolling the loop once. The subsequent execution steps will take care of testing the condition and possibly executing the body.

Note that we could have used the unrolling definition of *WHILE* in the big-step semantics as well. We were, after all, able to prove it as an equivalence in Section 7.2.4. However, such an unfolding is less natural in the big-step case, whereas in the small-step semantics the whole idea is to transform the command bit by bit to model execution.

Had we wanted to observe partial execution of arithmetic or boolean expressions, we could have introduced a small-step semantics for these as well (see Exercise 7.4) and made the corresponding small-step rules for assignment, *IF*, and *WHILE* non-atomic in the same way as the semicolon rules.

We can now define the execution of a program as the reflexive transitive closure of the *small_step* judgement \rightarrow , using the *star* operator defined in Section 4.5.2:

abbreviation $(\rightarrow^*) :: com \times state \Rightarrow com \times state \Rightarrow bool$ **where**
 $x \rightarrow^* y \equiv star_small_step\ x\ y$

Example 7.10. To look at an example execution of a command in the small-step semantics, we again use the **values** command. This time, we will get multiple elements in the set that it returns — all partial executions of the program. Given the command c with

$$c = ''x'' ::= V\ ''z'';; ''y'' ::= V\ ''x''$$

and an initial state s with

$$s = \langle ''x'' := 3, ''y'' := 7, ''z'' := 5 \rangle$$

we issue the following query to Isabelle

$$\mathbf{values}\ \{(c', map\ t\ [''x'', ''y'', ''z'']) \mid c' \ t.\ (c, s) \rightarrow^* (c', t)\}$$

The result contains four execution steps, starting with the original program in the initial state, proceeding through partial execution of the two assignments, and ending in the final state of the final program *SKIP*:

$$\begin{aligned} &\{ (''x'' ::= V\ ''z'';; ''y'' ::= V\ ''x'', [3, 7, 5]), \\ &\quad (SKIP;; ''y'' ::= V\ ''x'', [5, 7, 5]), \\ &\quad (''y'' ::= V\ ''x'', [5, 7, 5]), \\ &\quad (SKIP, [5, 5, 5]) \} \end{aligned}$$

As a further test of whether our definition of the small-step semantics is useful, we prove that the rules still give us a deterministic language, like the big-step semantics.

Lemma 7.11. $\llbracket cs \rightarrow cs'; cs \rightarrow cs'' \rrbracket \implies cs' = cs''$

Proof. After induction on the first premise (the small-step semantics), the proof is as automatic as for the big-step semantics. \square

! Recall that both sides of the small-step arrow \rightarrow are configurations, that is, pairs of commands and states. If we don't need to refer to the individual components, we refer to the configuration as a whole, such as cs in the lemma above.

We could conduct further tests like this, but since we already have a semantics for IMP, we can use it to show that our new semantics defines precisely the same behaviour. The next section does this.

7.3.1 Equivalence with Big-Step Semantics

Having defined an alternative semantics for the same language, the first interesting question is of course if our definitions are equivalent. This section

shows that this is the case. Both directions are proved separately: for any big-step execution, there is an equivalent small-step execution and vice versa.

We start by showing that any big-step execution can be simulated by a sequence of small steps ending in *SKIP*:

Lemma 7.12. $cs \Rightarrow t \implies cs \rightarrow^* (SKIP, t)$

This is proved in the canonical fashion by rule induction on the big-step judgement. Most cases follow directly. As an example we look at rule *IfTrue*:

$$\frac{bval\ b\ s \quad (c_1, s) \Rightarrow t}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \Rightarrow t}$$

By IH we know that $(c_1, s) \rightarrow^* (SKIP, t)$. This yields the required small-step derivation:

$$\frac{\frac{bval\ b\ s}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow (c_1, s)} \quad (c_1, s) \rightarrow^* (SKIP, t)}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow^* (SKIP, t)}$$

Only rule *Seq* does not go through directly:

$$\frac{(c_1, s_1) \Rightarrow s_2 \quad (c_2, s_2) \Rightarrow s_3}{(c_1;; c_2, s_1) \Rightarrow s_3}$$

The IHs are $(c_1, s_1) \rightarrow^* (SKIP, s_2)$ and $(c_2, s_2) \rightarrow^* (SKIP, s_3)$ but we need a reduction $(c_1;; c_2, s_1) \rightarrow^* \dots$. The following lemma bridges the gap: it lifts a \rightarrow^* derivation into the context of a semicolon:

Lemma 7.13.

$$(c_1, s_1) \rightarrow^* (c, s_2) \implies (c_1;; c_2, s_1) \rightarrow^* (c;; c_2, s_2)$$

Proof. The proof is by induction on the reflexive transitive closure *star*. The base case is trivial and the step is not much harder: If $(c_1, s_1) \rightarrow (c'_1, s'_1)$ and $(c'_1, s'_1) \rightarrow^* (c, s_2)$, we have $(c'_1;; c_2, s'_1) \rightarrow^* (c;; c_2, s_2)$ by IH. Rule *Seq2* and the step rule for *star* do the rest:

$$\frac{\frac{(c_1, s_1) \rightarrow (c'_1, s'_1)}{(c_1;; c_2, s_1) \rightarrow (c'_1;; c_2, s'_1)} \quad (c'_1;; c_2, s'_1) \rightarrow^* (c;; c_2, s_2)}{(c_1;; c_2, s_1) \rightarrow^* (c;; c_2, s_2)}$$

□

Returning to the proof of the *Seq* case, [Lemma 7.13](#) turns the first IH into $(c_1;; c_2, s_1) \rightarrow^* (SKIP;; c_2, s_2)$. From rule *Seq1* and the second IH we have $(SKIP;; c_2, s_2) \rightarrow^* (SKIP, s_3)$:

$$\frac{(SKIP;; c_2, s_2) \rightarrow (c_2, s_2) \quad (c_2, s_2) \rightarrow^* (SKIP, s_3)}{(SKIP;; c_2, s_2) \rightarrow^* (SKIP, s_3)}$$

By transitivity of *star* we finally arrive at $(c_1;; c_2, s_1) \rightarrow^* (SKIP, s_3)$, thus finishing the *Seq* case and the proof of [Lemma 7.12](#).

Let us now consider the other direction:

Lemma 7.14 (Small-step implies big-step).

$$cs \rightarrow^* (SKIP, t) \implies cs \Rightarrow t$$

The proof is again canonical, namely by rule induction on the premise, the reflexive transitive closure. The base case $cs = (SKIP, t)$ is trivial. In the induction step we have $cs \rightarrow cs'$ and $cs' \rightarrow^* (SKIP, t)$ and the IH $cs' \Rightarrow t$. That this implies $cs \Rightarrow t$ is proved as a separate lemma:

Lemma 7.15 (Step case). $\llbracket cs \rightarrow cs'; cs' \Rightarrow t \rrbracket \implies cs \Rightarrow t$

Proof. The proof is automatic after rule induction on the small-step semantics. \square

This concludes the proof of [Lemma 7.14](#). Both directions together ([Lemma 7.12](#) and [Lemma 7.14](#)) let us derive the equivalence we were aiming for in the first place:

Corollary 7.16. $(c, s) \Rightarrow t \iff (c, s) \rightarrow^* (SKIP, t)$

This concludes our proof that the small-step and big-step semantics of IMP are equivalent. Such equivalence proofs are useful whenever there are different formal descriptions of the same artefact. The reason one might want different descriptions of the same thing is that they differ in what they can be used for. For instance, big-step semantics are relatively intuitive to define, while small-step semantics allow us to make more fine-grained formal observations. The next section exploits the fine-grained nature of the small-step semantics to elucidate the big-step semantics.

7.3.2 Final Configurations, Infinite Reductions, and Termination

In contrast to the big-step semantics, in the small-step semantics it is possible to speak about non-terminating executions directly. We can easily distinguish final configurations from those that can make further progress:

definition *final* :: *com* \times *state* \Rightarrow *bool* **where**
final *cs* $\iff (\nexists cs'. cs \rightarrow cs')$

In our semantics, these happen to be exactly the configurations that have *SKIP* as their command.

Lemma 7.17. *final* $(c, s) = (c = SKIP)$

Proof. One direction is easy: clearly, if the command c is *SKIP*, the configuration is final. The other direction is not much harder. It is proved automatically after inducting on c . \square

With this we can show that \Rightarrow yields a final state iff \rightarrow terminates:

Lemma 7.18. $(\exists t. cs \Rightarrow t) \iff (\exists cs'. cs \rightarrow^* cs' \wedge \text{final } cs')$

Proof. Using Lemma 7.17 we can replace *final* with configurations that have *SKIP* as the command. The rest follows from the equivalence of small and big-step semantics. \square

This lemma says that in IMP the absence of a big-step result is equivalent to non-termination. This is not necessarily the case for any language. Another reason for the absence of a big-step result may be a runtime error in the execution of the program that leads to no rule being applicable. In the big-step semantics this is often indistinguishable from non-termination. In the small-step semantics the concept of final configurations neatly distinguishes the two causes.

Since IMP is deterministic, there is no difference between “may” and “must” termination. Consider a language with non-determinism.

In such a language, Lemma 7.18 is still valid and both sides speak about possible (*may*) termination. In fact, the big-step semantics cannot speak about necessary (*must*) termination at all, whereas the small-step semantics can: there must not be an infinite reduction $cs_0 \rightarrow cs_1 \rightarrow \dots$

7.4 Summary and Further Reading

This concludes the chapter on the operational semantics for IMP. In the first part of this chapter, we have defined the abstract syntax of IMP commands, we have defined the semantics of IMP in terms of a big-step operational semantics, and we have experimented with the concepts of semantic equivalence and determinism. In the second part of this chapter, we have defined an alternative form of operational semantics, namely small-step semantics, and we have proved that this alternative form describes the same behaviours as the big-step semantics. The two forms of semantics have different application trade-offs: big-step semantics were easier to define and understand, small-step semantics let us talk explicitly about intermediate states of execution and about termination.

We have looked at three main proof techniques: derivation trees, rule inversion and rule induction. These three techniques form the basic tool set that will accompany us in the following chapters.

Operational semantics in its small-step form goes back to Plotkin [70, 71, 72], who calls it structural operational semantics. Big-step semantics was popularised by Kahn [46] under the name of natural semantics.

There are many programming language constructs that we left out of IMP. Some examples that are relevant for imperative and object-oriented languages are the following.

Syntax. For loops, `do ... while` loops, the `if ... then` command, etc. are just further syntactic forms of the basic commands above. They could either be formalized directly, or they could be transformed into equivalent basic forms by syntactic de-sugaring.

Jumps. The `goto` construct, although considered harmful [27], is relatively easy to formalize. It merely requires the introduction of some notion of program position, be it as an explicit program counter, or a set of labels for jump targets. We will see jumps as part of a machine language formalization in [Chapter 8](#).

Blocks and local variables. Like the other constructs they do not add computational power but are an important tool for programmers to achieve data hiding and encapsulation. The main formalization challenge with local variables is their visibility scope. Nielson and Nielson [62] give a good introduction to this topic.

Procedures. Parameters to procedures introduce issues similar to local variables, but they may have additional complexities depending on which calling conventions the language implements (call by reference, call by value, call by name, etc.). Recursion is not usually a problem to formalize. Procedures also influence the definition of what a program is: instead of a single command, a program now usually becomes a list or collection of procedures. Nielson and Nielson cover this topic as well [62].

Exceptions. Throwing and catching exceptions is usually reasonably easy to integrate into a language formalization. However, exceptions may interact with features like procedures and local variables, because exceptions provide new ways to exit scopes and procedures. The Jinja [48] language formalization is an example of such rich interactions.

Data types, structs, pointers, arrays. Additional types such as fixed size machine words, records, and arrays are easy to include when the corresponding high-level concept is available in the theorem prover, but IEEE floating point values for instance may induce an interesting amount of work [24]. The semantics of pointers and references is a largely orthogonal issue and can be treated at various levels of detail, from raw bytes [87] up to multiple heaps separated by type and field names [13].

Objects, classes, methods. Object-oriented features have been the target of a large body of work in the past decade. Objects and methods lead

to a stronger connection between control structures and memory. Which virtual method will be executed, for instance, depends on the type of the object in memory at runtime. It is by now well understood how to formalize them in a theorem prover. For a study on a Java-like language in Isabelle/HOL, see for instance Jinja [48] or Featherweight Java [45].

All of the features above can be formalized in a theorem prover. Many add interesting complications, but the song remains the same. Schirmer [78], for instance, shows an Isabelle formalization of big-step and small-step semantics in the style of this chapter for the generic imperative language Simpl. The formalization includes procedures, blocks, exceptions, and further advanced concepts. With techniques similar to those described Chapter 12, he develops the language to a point where it can directly be used for large-scale program verification.

Exercises

Exercise 7.1. Define a function $assigned :: com \Rightarrow vname \text{ set}$ that computes the set of variables that are assigned to in a command. Prove that if some variable is not assigned to in a command, then that variable is never modified by the command: $\llbracket (c, s) \Rightarrow t; x \notin assigned\ c \rrbracket \Longrightarrow s\ x = t\ x$.

Exercise 7.2. Define a recursive function $skip :: com \Rightarrow bool$ that determines if a command behaves like *SKIP*. Prove $skip\ c \Longrightarrow c \sim SKIP$.

Exercise 7.3. Define a recursive function $deskip :: com \Rightarrow com$ that eliminates as many *SKIP*s as possible from a command. For example:

$deskip\ (SKIP;;\ WHILE\ b\ DO\ (x ::= a;;\ SKIP)) = WHILE\ b\ DO\ x ::= a$

Prove $deskip\ c \sim c$ by induction on c . Remember Lemma 7.5 for the *WHILE* case.

Exercise 7.4. Define a small-step semantics for the evaluation of arithmetic expressions:

inductive $astep :: aexp \times state \Rightarrow aexp \Rightarrow bool$ ($infix \rightsquigarrow 50$) **where**
 $(V\ x, s) \rightsquigarrow N\ (s\ x) \mid$
 $(Plus\ (N\ i)\ (N\ j), s) \rightsquigarrow N\ (i + j) \mid$

Complete the definition with two rules for *Plus* that model a left-to-right evaluation strategy: reduce the first argument with \rightsquigarrow if possible, reduce the second argument with \rightsquigarrow if the first argument is a number.

Prove that each \rightsquigarrow step preserves the value of the expression: $(a, s) \rightsquigarrow a' \Longrightarrow aval\ a\ s = aval\ a'\ s$. Use the modified induction rule $astep.induct$

[*split_format* (*complete*)]. Do not use the **case** idiom but write down explicitly what you assume and show in each case: **fix** ... **assume** ... **show**

Exercise 7.5. Prove or disprove (by giving a counterexample):

- *IF* *And* $b_1 \ b_2$ *THEN* c_1 *ELSE* $c_2 \sim$
IF b_1 *THEN IF* b_2 *THEN* c_1 *ELSE* c_2 *ELSE* c_2
- *WHILE* *And* $b_1 \ b_2$ *DO* $c \sim$ *WHILE* b_1 *DO WHILE* b_2 *DO* c
- *WHILE* *Or* $b_1 \ b_2$ *DO* $c \sim$ *WHILE* *Or* $b_1 \ b_2$ *DO* $c;;$ *WHILE* b_1 *DO* c

where *Or* $b_1 \ b_2 = \text{Not } (\text{And } (\text{Not } b_1) (\text{Not } b_2))$.

Exercise 7.6. Define a new loop construct *DO* c *WHILE* b (where c is executed once before b is tested) in terms of the existing constructs in *com*: *DO* c *WHILE* $b = \dots$. Define a recursive translation *dewhile* :: *com* \Rightarrow *com* that replaces all *WHILE* b *DO* c by suitable commands that use *DO* c *WHILE* b instead. Prove that your translation preserves the semantics: *dewhile* $c \sim c$.

Exercise 7.7. Let $C :: \text{nat} \Rightarrow \text{com}$ be an infinite sequence of commands and $S :: \text{nat} \Rightarrow \text{state}$ an infinite sequence of states such that $C \ 0 = c;; d$ and $\forall n. (C \ n, S \ n) \rightarrow (C \ (\text{Suc } n), S \ (\text{Suc } n))$. Prove that either all $C \ n$ are of the form $c_n;; d$ and it is always c_n that is reduced, or c_n eventually becomes *SKIP*:

$$\begin{aligned} & \llbracket C \ 0 = c;; d; \forall n. (C \ n, S \ n) \rightarrow (C \ (\text{Suc } n), S \ (\text{Suc } n)) \rrbracket \\ & \implies (\forall n. \exists c_1 \ c_2. \\ & \quad C \ n = c_1;; d \wedge \\ & \quad C \ (\text{Suc } n) = c_2;; d \wedge (c_1, S \ n) \rightarrow (c_2, S \ (\text{Suc } n))) \vee \\ & \quad (\exists k. C \ k = \text{SKIP};; d) \end{aligned}$$

For the following exercises copy theories *Com*, *Big_Step* and *Small_Step* and modify them as required.

Exercise 7.8. Extend IMP with a *REPEAT* c *UNTIL* b command. Adjust the definitions of big-step and small-step semantics, the proof that the big-step semantics is deterministic and the equivalence proof between the two semantics.

Exercise 7.9. Extend IMP with a new command c_1 *OR* c_2 that is a non-deterministic choice: it may execute either c_1 or c_2 . Adjust the definitions of big-step and small-step semantics, prove $(c_1 \ \text{OR} \ c_2) \sim (c_2 \ \text{OR} \ c_1)$ and update the equivalence proof between the two semantics.

Exercise 7.10. Extend IMP with exceptions. Add two constructors *THROW* and *TRY* c_1 *CATCH* c_2 to datatype *com*. Command *THROW* throws

an exception. The only command that can catch an exception is *TRY* c_1 *CATCH* c_2 : if an exception is thrown by c_1 , execution continues with c_2 , otherwise c_2 is ignored. Adjust the definitions of big-step and small-step semantics as follows. The big-step semantics is now of type $com \times state \Rightarrow com \times state$. In a big step $(c, s) \Rightarrow (x, t)$, x can only be *SKIP* (signalling normal termination) or *THROW* (signalling that an exception was thrown but not caught). The small-step semantics is of the same type as before. There are two final configurations now, $(SKIP, t)$ and $(THROW, t)$. Exceptions propagate upwards until an enclosing handler catches them.

Adjust the equivalence proof between the two semantics such that you obtain $cs \Rightarrow (SKIP, t) \iff cs \rightarrow^* (SKIP, t)$ and $cs \Rightarrow (THROW, t) \iff cs \rightarrow^* (THROW, t)$. Also revise the proof of $(\exists cs'. cs \Rightarrow cs') \iff (\exists cs'. cs \rightarrow^* cs' \wedge final\ cs')$.

Compiler

This chapter presents a first application of programming language semantics: proving compiler correctness. To this end, we will define a small machine language based on a simple stack machine. Stack machines are common low-level intermediate languages; the Java Virtual Machine is one example. We then write a compiler from IMP to this language and prove that the compiled program has the same semantics as the source program. The compiler will perform a very simple standard optimization for boolean expressions, but is otherwise non-optimizing.

As in the other chapters, the emphasis here is on showing the structure and main setup of such a proof. Our compiler proof shows the core of the argument, but compared to real compilers we make drastic simplifications: our target language is comparatively high-level, we do not consider optimizations, we ignore the compiler front-end, and our source language does not contain any concepts that are particularly hard to translate into machine language.

8.1 Instructions and Stack Machine thy

We begin by defining the instruction set architecture and semantics of our stack machine. We have already seen a very simple stack machine language in [Section 3.3](#). In this section, we extend this language with memory writes and jump instructions.

Working with proofs on the machine language, we will find it convenient for the program counter to admit negative values, i.e., to be of type *int* instead of the initially more intuitive *nat*. The effect of this choice is that various decomposition lemmas about machine executions have nicer algebraic properties and fewer preconditions than their *nat* counterparts. Such effects are usually discovered during the proof.

As in [Section 3.3](#), our machine language models programs as lists of instructions. Our *int* program counter will need to index into these lists. Isabelle comes with a predefined list index operator *nth*, but it works on *nat*. Instead of constantly converting between *int* and *nat* and dealing with the arising side conditions in proofs, we define our own *int* version of *nth*, i.e., for $i :: \text{int}$:

$$(x \# xs) !! i = (\text{if } i = 0 \text{ then } x \text{ else } xs !! (i - 1))$$

However, we still need the conversion $\text{int} :: \text{nat} \Rightarrow \text{int}$ because the length of a list is of type *nat*. To reduce clutter we introduce the abbreviation

$$\text{size } xs \equiv \text{int}(\text{length } xs)$$

The *!!* operator distributes over *@* in the expected way:

Lemma 8.1. *If $0 \leq i$,*

$$(xs @ ys) !! i = (\text{if } i < \text{size } xs \text{ then } xs !! i \text{ else } ys !! (i - \text{size } xs))$$

We are now ready to define the machine itself. To keep things simple, we directly reuse the concepts of values and variable names from the source language. In a more realistic setting, we would explicitly map variable names to memory locations, instead of using strings as addresses. We skip this step here for clarity, adding it does not pose any fundamental difficulties.

The instructions in our machine are the following. The first three are familiar from [Section 3.3](#):

```
datatype instr =
  LOADI int | LOAD vname | ADD | STORE vname |
  JMP int | JMPLESS int | JMPGE int
```

The instruction *LOADI* loads an immediate value onto the stack, *LOAD* loads the value of a variable, *ADD* adds the two topmost stack values, *STORE* stores the top of stack into memory, *JMP* jumps by a relative value, *JMPLESS* compares the two topmost stack elements and jumps if the second one is less, and finally *JMPGE* compares and jumps if the second one is greater or equal.

These few instructions are enough to compile IMP programs. A real machine would have significantly more arithmetic and comparison operators, different addressing modes that are useful for implementing procedure stacks and pointers, potentially a number of primitive data types that the machine understands, and a number of instructions to deal with hardware features such as the memory management subsystem that we ignore in this formalization.

As in the source language, we proceed by defining the state such programs operate on, followed by the definition of the semantics itself.

Program configurations consist of an *int* program counter, a memory state for which we re-use the type *state* from the source language, and a stack which we model as a list of values:

```
type_synonym stack = val list
type_synonym config = int × state × stack
```

We now define the semantics of machine execution. Similarly to the small-step semantics of the source language, we do so in multiple levels: first, we define what effect a single instruction has on a configuration, then we define how an instruction is selected from the program, and finally we take the reflexive transitive closure to get full machine program executions.

We encode the behaviour of single instructions in the function *ieexec* below. The program counter *i* is usually incremented by 1, except for the jump instructions. Variables are loaded from and stored into the variable state *s* with function application and function update. For the stack *stk*, we use standard list constructs (see [Section 2.2.5](#)) as well as two new abbreviations: *hd2 xs* \equiv *hd (tl xs)* and *tl2 xs* \equiv *tl (tl xs)*.

```
fun ieexec :: instr ⇒ config ⇒ config where
  ieexec (LOADI n) (i, s, stk) = (i + 1, s, n # stk)
  ieexec (LOAD x) (i, s, stk) = (i + 1, s, s x # stk)
  ieexec ADD (i, s, stk) = (i + 1, s, (hd2 stk + hd stk) # tl2 stk)
  ieexec (STORE x) (i, s, stk) = (i + 1, s(x := hd stk), tl stk)
  ieexec (JMP n) (i, s, stk) = (i + 1 + n, s, stk)
  ieexec (JMPLESS n) (i, s, stk) =
    (if hd2 stk < hd stk then i + 1 + n else i + 1, s, tl2 stk)
  ieexec (JMPGE n) (i, s, stk) =
    (if hd stk ≤ hd2 stk then i + 1 + n else i + 1, s, tl2 stk)
```

The next level up, a single execution step selects the instruction the program counter (*pc*) points to and uses *ieexec* to execute it. For execution to be well defined, we additionally check if the *pc* points to a valid location in the list. We call this predicate *exec1* and give it the notation $P \vdash c \rightarrow c'$ for *program P executes from configuration c to configuration c'*.

```
definition exec1 :: instr list ⇒ config ⇒ config ⇒ bool where
  P ⊢ c → c' =
    (∃ i s stk. c = (i, s, stk) ∧ c' = ieexec (P !! i) c ∧ 0 ≤ i < size P)
```

where $x \leq y < z \equiv x \leq y \wedge y < z$ as usual in mathematics.

The last level is the lifting from single step execution to multiple steps using the standard reflexive transitive closure definition that we already used for the small-step semantics of the source language, that is:

abbreviation $P \vdash c \rightarrow^* c' \equiv \text{star} (\text{exec1 } P) c c'$

This concludes our definition of the machine and its semantics. As usual in this book, the definitions are executable. This means, we can try out a simple example. Let $P = [\text{LOAD } ''y'', \text{STORE } ''x''], s ''x'' = 3$, and $s ''y'' = 4$. Then

values $\{(i, \text{map } t [''x'', ''y''], \text{stk}) \mid i \text{ t stk. } P \vdash (0, s, []) \rightarrow^* (i, t, \text{stk})\}$

will produce the following sequence of configurations:

$\{(0, [3, 4], []), (1, [3, 4], [4]), (2, [4, 4], [])\}$

8.2 Reasoning About Machine Executions thy

The compiler proof is more involved than the short proofs we have seen so far. We will need a small number of technical lemmas before we get to the compiler correctness problem itself. Our aim in this section is to execute machine programs symbolically as far as possible using Isabelle's proof tools. We will then use this ability in the compiler proof to assemble multiple smaller machine executions into larger ones.

A first lemma to this end is that execution results are preserved if we append additional code to the left or right of a program. Appending at the right side is easy:

Lemma 8.2. $P \vdash c \rightarrow^* c' \implies P @ P' \vdash c \rightarrow^* c'$

Proof. The proof is by induction on the reflexive transitive closure. For the step case, we observe after unfolding of *exec1* that appending program context on the right does not change the result of indexing into the original instruction list. □

Appending code on the left side requires shifting the program counter.

Lemma 8.3.

$$\begin{aligned} P \vdash (i, s, \text{stk}) &\rightarrow^* (i', s', \text{stk}') \implies \\ P' @ P \vdash (\text{size } P' + i, s, \text{stk}) &\rightarrow^* (\text{size } P' + i', s', \text{stk}') \end{aligned}$$

Proof. The proof is again by induction on the reflexive transitive closure and reduction to *exec1* in the step case. To show the lemma for *exec1*, we unfold its definition and observe [Lemma 8.4](#). □

The execution of a single instruction can be relocated arbitrarily:

Lemma 8.4.

$$\begin{aligned} (n + i', s', \text{stk}') &= \text{iexec } x (n + i, s, \text{stk}) \longleftrightarrow \\ (i', s', \text{stk}') &= \text{iexec } x (i, s, \text{stk}) \end{aligned}$$

Proof. We observe by case distinction on the instruction x that the only component of the result configuration that is influenced by the program counter i is the first one, and in this component only additively. For instance, in the *LOADI* n instruction, we get on both sides $s' = s$ and $stk' = n \# stk$. The *pc* field for input i on the right-hand side is $i' = i + 1$. The *pc* field for $n + i$ on the left-hand side becomes $n + i + 1$, which is $n + i'$ as required. The other cases are analogous. \square

Taking these two lemmas together, we can compose separate machine executions into one larger one.

Lemma 8.5 (Composing machine executions).

$$\begin{aligned} & \llbracket P \vdash (0, s, stk) \rightarrow^* (i', s', stk'); \text{size } P \leq i'; \\ & P' \vdash (i' - \text{size } P, s', stk') \rightarrow^* (i'', s'', stk'') \rrbracket \\ \implies & P @ P' \vdash (0, s, stk) \rightarrow^* (\text{size } P + i'', s'', stk'') \end{aligned}$$

Proof. The proof is by suitably instantiating [Lemma 8.2](#) and [Lemma 8.3](#). \square

8.3 Compilation thy

We are now ready to define the compiler, and will do so in the three usual steps: first for arithmetic expressions, then for boolean expressions, and finally for commands. We have already seen compilation of arithmetic expressions in [Section 3.3](#). We define the same function for our extended machine language:

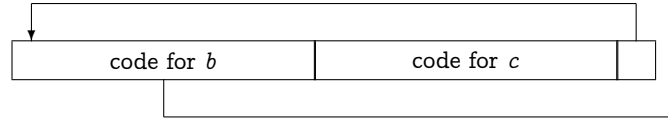
```
fun acomp :: aexp ⇒ instr list where
  acomp (N n)      = [LOADI n]
  acomp (V x)      = [LOAD x]
  acomp (Plus a1 a2) = acomp a1 @ acomp a2 @ [ADD]
```

The correctness statement is not as easy any more as in [Section 3.3](#), because program execution now is a relation, not simply a function. For our extended machine language a function is not suitable because we now have potentially non-terminating executions. This is not a big obstacle: we can still express naturally that the execution of a compiled arithmetic expression will leave the result on top of the stack, and that the program counter will point to the end of the compiled expression.

Lemma 8.6 (Correctness of *acom*).

$$\text{acom } a \vdash (0, s, stk) \rightarrow^* (\text{size } (\text{acom } a), s, \text{aval } a \# stk)$$

Proof. The proof is by induction on the arithmetic expression. \square

Fig. 8.1. Compilation of *WHILE b DO c*

The compilation schema for boolean expressions is best motivated by a preview of the layout of the code generated from *WHILE b DO c* shown in Figure 8.1. Arrows indicate jump instructions. Let *cb* be code generated from *b*. If *b* evaluates to *True*, the execution of *cb* should lead to the end of *cb*, continue with the execution of the code for *c* and jump back to the beginning of *cb*. If *b* evaluates to *False*, the execution of *cb* should jump behind all of the loop code. For example, when executing the compiled code for *WHILE And b₁ b₂ DO c*, after having found that *b₁* evaluates to *False* we can safely jump out of the loop. There can be multiple such jumps: think of *And b₁ (And b₂ b₃)*.

To support this schema, the *bexp* compiler takes two further parameters in addition to *b*: an offset *n* and a flag *f* :: *bool* that determines for which value of *b* the generated code should jump to offset *n*. This enables us to perform a small bit of optimization: boolean constants do not need to execute any code; they either compile to nothing or to a jump to the offset, depending on the value of *f*. The *Not* case simply inverts *f*. The *And* case performs shortcut evaluation as explained above. The *Less* operator uses the *acom* compiler for *a₁* and *a₂* and then selects the appropriate compare and jump instruction according to *f*.

```

fun bcomp :: bexp ⇒ bool ⇒ int ⇒ instr list where
  bcomp (Bc v) f n      = (if v = f then [JMP n] else [])
  bcomp (Not b) f n     = bcomp b (¬ f) n
  bcomp (And b1 b2) f n =
    (let cb2 = bcomp b2 f n;
     m = if f then size cb2 else size cb2 + n;
     cb1 = bcomp b1 False m
    in cb1 @ cb2)
  bcomp (Less a1 a2) f n =
    acomp a1 @ acomp a2 @ (if f then [JMPLESS n] else [JMPGE n])

```

Example 8.7. Boolean constants are optimized away:

```

value bcomp (And (Bc True) (Bc True)) False 3
returns []

```

```

value bcomp (And (Bc False) (Bc True)) True 3
returns [JMP 1, JMP 3]

value bcomp (And (Less (V ''x'') (V ''y'')) (Bc True)) False 3
returns [LOAD ''x'', LOAD ''y'', JMPGE 3]

```

The second example shows that the optimization is not perfect: it may generate dead code.

The correctness statement is the following: the stack and state should remain unchanged and the program counter should indicate if the expression evaluated to *True* or *False*. If $f = \text{False}$ then we end at $\text{size } (bcomp\ b\ f\ n)$ in the *True* case and $\text{size } (bcomp\ b\ f\ n) + n$ in the *False* case. If $f = \text{True}$ it is the other way around. This statement only makes sense for forward jumps, so we require n to be non-negative.

Lemma 8.8 (Correctness of *bcomp*).

Let $pc' = \text{size } (bcomp\ b\ f\ n) + (\text{if } f = \text{bval } b\ s \text{ then } n \text{ else } 0)$. Then $0 \leq n \implies bcomp\ b\ f\ n \vdash (0, s, stk) \rightarrow^* (pc', s, stk)$

Proof. The proof is by induction on b . The constant and *Less* cases are solved automatically. For the *Not* case, we instantiate the induction hypothesis manually to $\neg f$. For *And*, we get two recursive cases. The first needs the induction hypothesis instantiated with $\text{size } (bcomp\ b_2\ f\ n) + (\text{if } f \text{ then } 0 \text{ else } n)$ for n , and with *False* for f , and the second goes through with simply n and f . \square

With both expression compilers in place, we can now proceed to the command compiler *ccomp*. The idea is to compile c into a program that will perform the same state transformation as c and that will always end with the program counter at $\text{size } (ccomp\ c)$. It may push and consume intermediate values on the stack for expressions, but at the end, the stack will be the same as in the beginning. Because of this modular behaviour of the compiled code, the compiler can be defined recursively as follows:

- *SKIP* compiles to the empty list;
- for assignment, we compile the expression and store the result;
- sequential composition appends the corresponding machine programs;
- *IF* is compiled as shown in Figure 8.2;
- *WHILE* is compiled as shown in Figure 8.1.

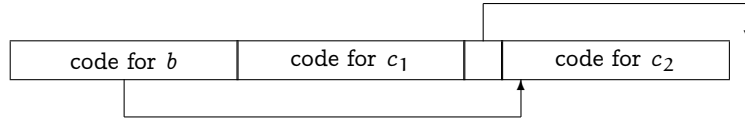
Figure 8.3 shows the formal definition.

Since everything in *ccomp* is executable, we can inspect the results of compiler runs directly in the theorem prover. For example, let p_1 be the command for *IF* $u < 1$ *THEN* $u := u + 1$ *ELSE* $v := u$. Then

```

value ccomp p1

```

Fig. 8.2. Compilation of *IF b THEN c₁ ELSE c₂*

```

fun ccomp :: com ⇒ instr list where
  ccomp SKIP                = []
  ccomp (x ::= a)            = acomp a @ [STORE x]
  ccomp (c1;; c2)          = ccomp c1 @ ccomp c2
  ccomp (IF b THEN c1 ELSE c2) =
    (let cc1 = ccomp c1; cc2 = ccomp c2;
     cb = bcomp b False (size cc1 + 1)
    in cb @ cc1 @ JMP (size cc2) # cc2)
  ccomp (WHILE b DO c)       =
    (let cc = ccomp c; cb = bcomp b False (size cc + 1)
     in cb @ cc @ [JMP (- (size cb + size cc + 1))])

```

Fig. 8.3. Definition of *ccomp*

results in

```

[LOAD "u", LOADI 1, JMPGE 5, LOAD "u", LOADI 1, ADD,
 STORE "u", JMP 2, LOAD "u", STORE "v"]

```

Similarly for loops. Let p_2 be *WHILE u < 1 DO u := u + 1*. Then

```
value ccomp p2
```

results in

```

[LOAD "u", LOADI 1, JMPGE 5, LOAD "u", LOADI 1, ADD,
 STORE "u", JMP (- 8)]

```

8.4 Preservation of Semantics thy

This section shows the correctness proof of our small toy compiler. For IMP, the correctness statement is fairly straightforward: the machine program should have precisely the same behaviour as the source program. It should cause the same state change and nothing more than that. It should terminate if and only if the source program terminates. Since we use the same type for states at the source level and machine level, the first part of the property is easy to express. Similarly, it is easy to say that the stack should not change.

Finally, we can express correct termination by saying that the machine execution started at $pc = 0$ should stop at the end of the machine code with $pc = \text{size } (ccomp\ c)$. In total, we have

$$(c, s) \Rightarrow t \iff ccomp\ c \vdash (0, s, stk) \rightarrow^* (\text{size } (ccomp\ c), t, stk)$$

In other words, the compiled code executes from s to t if and only if the big-step semantics executes the source code from s to t .

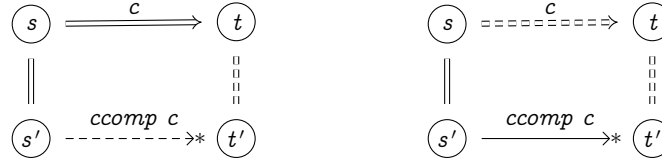


Fig. 8.4. Compiler correctness as two simulations

The two directions of the “ \longleftrightarrow ” are shown diagrammatically in Figure 8.4. The upper level is the big-step execution $(s, c) \Rightarrow t$. The lower level is stack machine execution. The relationship between states and configurations is given by $s' = (0, s, stk)$ and $t' = (\text{size } (ccomp\ c), t, stk)$.

Such diagrams should be read as follows: the solid lines and arrows imply the existence of the dashed ones. Thus the left diagram depicts the “ \longrightarrow ” direction (source code can be simulated by compiled code) and the right diagram depicts the “ \longleftarrow ” direction (compiled code can be simulated by source code). We have already seen the analogous simulation relations between big-step and small-step semantics in Section 7.3.1 and will see more such simulations later in the book.

In the case of the compiler, the crucial direction is the simulation of the compiled code by the source code: it tells us that every final state produced by the compiled code is justified by the source code semantics. Therefore we can trust all results produced by the compiled code, if it terminates. But it is still possible that the compiled code does not terminate although the source code does. That can be ruled out by proving also that the compiled code simulates the source code.

We will now prove the two directions of our compiler correctness statement separately. Simulation of the source code by the compiled code is comparatively easy.

Lemma 8.9 (Correctness of $ccomp$).

If the source program executes from s to t , so will the compiled program.

Formally:

$$(c, s) \Rightarrow t \implies ccomp\ c \vdash (0, s, stk) \rightarrow^* (size\ (ccomp\ c), t, stk)$$

Proof. The proof is by rule induction on the big-step semantics. We go through each of the cases step by step.

SKIP:

The *SKIP* case translates to the empty machine program and trivially satisfies the correctness statement.

$x ::= a:$

In the assignment case, we use the correctness of arithmetic expression compilation (Lemma 8.6) together with the composition lemma (Lemma 8.5) for appending the store instruction. We let the prover symbolically execute this store instruction after the expression evaluation to conclude that the state is updated correctly.

$c_1;; c_2:$

For sequential composition, we get correctness of the two executions for c_1 and c_2 as induction hypothesis, and merely need to lift them into the context of the entire program using our machine execution composition lemmas.

IF b *THEN* c_1 *ELSE* $c_2:$

The two cases for *IF* are solved entirely automatically using correctness of boolean expression compilation (Lemma 8.8), symbolic execution of the jump instruction and the appropriate induction hypothesis for the respective branch.

WHILE b *DO* $c:$

The *False* case of the loop is automatic and follows directly from the correctness lemma for boolean expressions (Lemma 8.8). In the *True* case, we have the two source level executions $(c, s_1) \Rightarrow s_2$ and $(WHILE\ b\ DO\ c, s_2) \Rightarrow s_3$ with corresponding induction hypotheses for the compiled body $ccomp\ c$ started in s_1 and the compiled code of the entire loop started in s_2 ending in s_3 . From this we need to construct an execution of the entire loop from s_1 to s_3 . We again argue by correctness of boolean expression compilation (Lemma 8.8). We know that the *True* case of that compiled code will end in a program counter pointing to the beginning of the compiled body $ccomp\ c$. We use the induction hypothesis and the composition lemma (Lemma 8.5) to execute this code into state s_2 with a pc pointing to the jump instruction that returns us to the beginning of the loop. Executing this jump instruction symbolically, we get to the machine configuration that lets us use the induction hypothesis that executes the rest of the loop into the desired state s_3 .

□

The second direction of the compiler correctness proof is more involved. We have to show that if the machine code executes from state s to t , so does the source code: the source code simulates the compiled code. Formally:

$$ccomp\ c \vdash (0, s, stk) \rightarrow^* (size\ (ccomp\ c), t, stk') \implies (c, s) \Rightarrow t$$

The main reason this direction is harder to show than the other one is the lack of a suitable structural induction principle that we could apply. Since rule induction on the semantics is not applicable, we have only two further induction principles left in the arsenal we have learned so far: structural induction on the command c or induction on the length of the \rightarrow^* execution. Neither is strong enough on its own. The solution is to combine them: we will use an outside structural induction on c which will take care of all cases but *WHILE*, and then a nested, inner induction on the length of the execution for the *WHILE* case.

This idea takes care of the general proof structure. The second problem we encounter is the one of decomposing larger machine executions into smaller ones. Consider the semicolon case. We will have an execution of the form $ccomp\ c_1 @ ccomp\ c_2 \vdash cfg \rightarrow^* cfg'$, and our first induction hypothesis will come with the precondition $ccomp\ c_1 \vdash cfg \rightarrow^* cfg''$. It may seem intuitive that if $cs_1 @ cs_2 \vdash cfg \rightarrow^* cfg'$ then there must be some intermediate executions $cs_1 \vdash cfg \rightarrow^* cfg''$ and $cs_2 \vdash cfg'' \rightarrow^* cfg$, but this is not true in general for arbitrary code sequences cs_1 and cs_2 or configurations cfg and cfg' . For instance, the code may be jumping between cs_1 and cs_2 continuously, and neither execution may make sense in isolation.

However, the code produced from our compiler is particularly well behaved: execution of compiled code will never jump outside that code and will exit at precisely $pc = size\ (ccomp\ c)$. We merely need to formalize this concept and prove that it is adhered to. This requires a few auxiliary notions.

First we define *isuccs*, the possible successor program counters of a given instruction at position n :

definition *isuccs* :: *instr* \Rightarrow *int* \Rightarrow *int set* **where**
isuccs $i\ n = (case\ i\ of$
 JMP $j \Rightarrow \{n + 1 + j\} \mid$
 JMPLESS $j \Rightarrow \{n + 1 + j, n + 1\} \mid$
 JMPGE $j \Rightarrow \{n + 1 + j, n + 1\} \mid$
 $_ \Rightarrow \{n + 1\})$

Then we define *succs* $P\ n$ which yields the successor program counters of an instruction sequence P which itself may be embedded in a larger program at position n . The possible successors program counters of an instruction list are the union of all instruction successors:

definition $\text{succs} :: \text{instr list} \Rightarrow \text{int} \Rightarrow \text{int set}$ where
 $\text{succs } P \ n = \{s. \exists i \geq 0. i < \text{size } P \wedge s \in \text{isuccs } (P !! i) \ (n + i)\}$

Finally, we remove all jump targets internal to P from $\text{succs } P \ 0$ and arrive at the possible exit program counters of P . The notation $\{a..<b\}$ stands for the set of numbers $\geq a$ and $< b$. Similarly, $\{a..b\}$ is the set of numbers $\geq a$ and $\leq b$.

definition $\text{exits} :: \text{instr list} \Rightarrow \text{int set}$ where
 $\text{exits } P = \text{succs } P \ 0 - \{0..<\text{size } P\}$

Unsurprisingly, we will need to reason about the successors and exits of composite instruction sequences.

Lemma 8.10 (Successors over append).

$$\text{succs } (cs @ cs') \ n = \text{succs } cs \ n \cup \text{succs } cs' \ (n + \text{size } cs)$$

Proof. The proof is by induction on the instruction list cs . To solve each case, we derive the equations for Nil and Cons in succs separately:

$$\begin{aligned} \text{succs } [] \ n &= \{\} \\ \text{succs } (x \# xs) \ n &= \text{isuccs } x \ n \cup \text{succs } xs \ (1 + n) \end{aligned}$$

□

We could prove a similar lemma about $\text{exits } (cs @ cs')$, but this lemma would have a more complex right-hand side. For the results below it is easier to reason about succs first and then apply the definition of exits to the result instead of decomposing exits directly.

Before we proceed to reason about decomposing machine executions and compiled code, we note that instead of using the reflexive transitive closure of single-step execution, we can equivalently talk about n steps of execution. This will give us a more flexible induction principle and allow us to talk more precisely about sequences of execution steps. We write $P \vdash c \rightarrow^{\sim n} c'$ to mean that the execution of program P starting in configuration c can reach configuration c' in n steps and define:

$$\begin{aligned} P \vdash c \rightarrow^{\sim 0} c' &= (c' = c) \\ P \vdash c \rightarrow^{\sim (Suc \ n)} c' &= (\exists c'. P \vdash c \rightarrow c' \wedge P \vdash c' \rightarrow^{\sim n} c') \end{aligned}$$

Our old concept of $P \vdash c \rightarrow^* c'$ is equivalent to saying that there exists an n such that $P \vdash c \rightarrow^{\sim n} c'$.

Lemma 8.11. $(P \vdash c \rightarrow^* c') = (\exists n. P \vdash c \rightarrow^{\sim n} c')$

Proof. One direction is by induction on n , the other by induction on the reflexive transitive closure. □

The more flexible induction principle mentioned above is **complete induction** on n : $(\bigwedge n. \forall m < n. P\ m \implies P\ n) \implies P\ n$. That is, to show that a property P holds for any n , it suffices to show that it holds for an arbitrary n under the assumption that P already holds for any $m < n$. In our context, this means we are no longer limited to splitting off single execution steps at a time.

We can now derive lemmas about the possible exits for *acom*, *bcomp*, and *ccomp*. Arithmetic expressions are the easiest, they don't contain any jump instructions, and we only need our append lemma for *succs* (Lemma 8.10).

Lemma 8.12. $exits\ (acom\ a) = \{size\ (acom\ a)\}$

Proof. The proof is by first computing all successors of *acom*, and then deriving the *exits* from that result. For the successors of *acom*, we show

$$succs\ (acom\ a)\ n = \{n + 1..n + size\ (acom\ a)\}$$

by induction on a . The set from $n + 1$ to $n + size\ (acom\ a)$ is not empty, because we can show $1 \leq size\ (acom\ a)$ by induction on a . \square

Compilation for boolean expressions is less well behaved. The main idea is that *bcomp* has two possible *exits*, one for *True*, one for *False*. However, as we have seen in the examples, compiling a boolean expression might lead to the empty list of instructions, which has no successors or exits. More generally, the optimization that *bcomp* performs may statically exclude one or both of the possible exits. Instead of trying to precisely describe each of these cases, we settle for providing an upper bound on the possible successors of *bcomp*. We are also only interested in positive offsets i .

Lemma 8.13. *If $0 \leq i$, then*

$$exits\ (bcomp\ b\ f\ i) \subseteq \{size\ (bcomp\ b\ f\ i), i + size\ (bcomp\ b\ f\ i)\}$$

Proof. Again, we reduce *exits* to *succs*, and first prove

$$\begin{aligned} 0 \leq i &\implies \\ succs\ (bcomp\ b\ f\ i)\ n & \\ \subseteq \{n..n + size\ (bcomp\ b\ f\ i)\} \cup \{n + i + size\ (bcomp\ b\ f\ i)\} & \end{aligned}$$

After induction on b , this proof is mostly automatic. We merely need to instantiate the induction hypothesis for the *And* case manually. \square

Finally, we come to the *exits* of *ccomp*. Since we are building on the lemma for *bcomp*, we can again only give an upper bound: there are either no exits, or the exit is precisely $size\ (ccomp\ c)$. As an example that the former case can occur as a result of *ccomp*, consider the compilation of an endless loop

$$ccomp\ (WHILE\ Bc\ True\ DO\ SKIP) = [JMP\ (-\ 1)]$$

That is, we get one jump instruction that jumps to itself and

$$\text{exits } [JMP \ (-1)] = \{\}$$

Lemma 8.14. $\text{exits } (\text{ccomp } c) \subseteq \{\text{size } (\text{ccomp } c)\}$

Proof. We first reduce the lemma to *succs*:

$$\text{succs } (\text{ccomp } c) \ n \subseteq \{n..n + \text{size } (\text{ccomp } c)\}$$

This proof is by induction on c and the corresponding *succs* results for *acomp* and *bcomp*. \square

We have derived these *exits* results about *acomp*, *bcomp* and *ccomp* because we wanted to show that the machine code produced by these functions is well behaved enough that larger executions can be decomposed into smaller separate parts. The main lemma that describes this decomposition is somewhat technical. It states that, given an n -step execution of machine instructions cs that are embedded in a larger program $(P @ cs @ P')$, we can find a k -step sub-execution of cs in isolation, such that this sub-execution ends at one of the exit program counters of cs , and such that it can be continued to end in the same state as the original execution of the larger program. For this to be true, the initial pc of the original execution must point to somewhere within cs , and the final pc' to somewhere outside cs . Formally, we get the following.

Lemma 8.15 (Decomposition of machine executions).

$$\begin{aligned} & \llbracket P @ cs @ P' \vdash (\text{size } P + pc, stk, s) \rightarrow \sim^n (pc', stk', s'); \\ & \quad pc \in \{0..<\text{size } cs\}; pc' \notin \{\text{size } P..<\text{size } P + \text{size } cs\} \rrbracket \\ \implies & \exists pc'' \ stk'' \ s'' \ k \ m. \\ & \quad cs \vdash (pc, stk, s) \rightarrow \sim^k (pc'', stk'', s'') \wedge \\ & \quad pc'' \in \text{exits } cs \wedge \\ & \quad P @ cs @ P' \vdash (\text{size } P + pc'', stk'', s'') \rightarrow \sim^m (pc', stk', s') \wedge \\ & \quad n = k + m \end{aligned}$$

Proof. The proof is by induction on n . The base case is trivial, and the step case essentially reduces the lemma to a similar property for a single execution step:

$$\begin{aligned} & \llbracket P @ cs @ P' \vdash (\text{size } P + pc, stk, s) \rightarrow (pc', stk', s'); \\ & \quad pc \in \{0..<\text{size } cs\} \rrbracket \\ \implies & cs \vdash (pc, stk, s) \rightarrow (pc' - \text{size } P, stk', s') \end{aligned}$$

This property is proved automatically after unfolding the definition of single-step execution and case distinction on the instruction to be executed.

Considering the step case for $n + 1$ execution steps in our induction again, we note that this step case consists of one single step in the larger context and the n -step rest of the execution, also in the larger context. Additionally we have the induction hypothesis, which gives us our property for executions of length n .

Using the property above, we can reduce the single-step execution to cs . To connect this up with the rest of the execution, we make use of the induction hypothesis in the case where the execution of cs is not at an exit yet, which means the pc is still inside cs , or we observe that the execution has left cs and the pc must therefore be at an exit of cs . In this case k is 1 and $m = n$, and we can just append the rest of the execution from our assumptions. \square

The accompanying Isabelle theories contain a number of more convenient instantiations of this lemma. We omit these here, and directly move on to proving the correctness of $acom$, $bcomp$, $ccomp$.

As always, arithmetic expressions are the least complex case.

Lemma 8.16 (Correctness of $acom$, reverse direction).

$$acom\ a \vdash (0, s, stk) \rightarrow \sim^n (size\ (acom\ a), s', stk') \implies s' = s \wedge stk' = aval\ a\ s \# stk$$

Proof. The proof is by induction on the expression, and most cases are solved automatically, symbolically executing the compilation and resulting machine code sequence. In the *Plus* case, we decompose the execution manually using [Lemma 8.15](#), apply the induction hypothesis for the parts, and combine the results symbolically executing the *ADD* instruction. \square

The next step is the compilation of boolean expressions. Correctness here is mostly about the pc' at the end of the expression evaluation. Stack and state remain unchanged.

Lemma 8.17 (Correctness of $bcomp$, reverse direction).

$$\begin{aligned} & \llbracket bcomp\ b\ f\ j \vdash (0, s, stk) \rightarrow \sim^n (pc', s', stk'); \\ & \quad size\ (bcomp\ b\ f\ j) \leq pc'; 0 \leq j \rrbracket \\ \implies & pc' = size\ (bcomp\ b\ f\ j) + (if\ f = bval\ b\ s\ then\ j\ else\ 0) \wedge \\ & s' = s \wedge stk' = stk \end{aligned}$$

Proof. The proof is by induction on the expression. The *And* case is the only interesting one. We first split the execution into one for the left operand b_1 and one for the right operand b_2 with a suitable instantiation of our splitting lemma above ([Lemma 8.15](#)). We then determine by induction hypothesis that stack and state did not change for b_1 and that the program counter will either exit directly, in which case we are done, or it will point to the instruction sequence of b_2 , in which case we apply the second induction hypothesis to conclude the case and the lemma. \square

We are now ready to tackle the main lemma for *ccomp*.

Lemma 8.18 (Correctness of *ccomp*, reverse direction).

$$\begin{aligned} & ccomp\ c \vdash (0, s, stk) \rightarrow \sim^n (size\ (ccomp\ c), t, stk') \implies \\ & (c, s) \Rightarrow t \wedge stk' = stk \end{aligned}$$

Proof. As mentioned before, the main proof is an induction on *c*, and in the *WHILE* case there is a nested complete induction on the length of the execution. The cases of the structural induction are the following.

SKIP:

This case is easy and automatic.

x ::= a:

The assignment case makes use of the correct compilation of arithmetic expressions (Lemma 8.16) and is otherwise automatic.

c₁;; c₂:

This case comes down to using Lemma 8.15 and combining the induction hypotheses as usual.

IF b THEN c₁ ELSE c₂:

The *IF* case is more interesting. Let *I* stand for the whole *IF* expression. We start by noting that we need to prove

$$\begin{aligned} & ccomp\ I \vdash (0, s, stk) \rightarrow \sim^n (size\ (ccomp\ I), t, stk') \implies \\ & (I, s) \Rightarrow t \wedge stk' = stk \end{aligned}$$

and that we have the same property available for *ccomp c₁* and *ccomp c₂* as induction hypotheses. After splitting off the execution of the boolean expression using Lemma 8.17 and Lemma 8.15, we know

$$\begin{aligned} & ccomp\ c_1 @ JMP\ (size\ (ccomp\ c_2)) \# ccomp\ c_2 \\ & \vdash (if\ bval\ b\ s\ then\ 0\ else\ size\ (ccomp\ c_1) + 1, s, stk) \rightarrow \sim^m \\ & (1 + size\ (ccomp\ c_1) + size\ (ccomp\ c_2), t, stk') \end{aligned}$$

We proceed by case distinction on *bval b s*, and use the corresponding induction hypothesis for *c₁* and *c₂* respectively to conclude the *IF* case.

WHILE b DO c:

In the *WHILE* case, let *w* stand for the loop *WHILE b DO c*, and *cs* for the compiled loop *ccomp w*. Our induction hypothesis is

$$\begin{aligned} & ccomp\ c \vdash (0, s, stk) \rightarrow \sim^n (size\ (ccomp\ c), t, stk') \implies \\ & (c, s) \Rightarrow t \wedge stk' = stk \end{aligned}$$

for any *s*, *stk*, *n*, *t*, and *stk'*. We need to show

$$\begin{aligned} & cs \vdash (0, s, stk) \rightarrow \sim^n (size\ cs, t, stk') \implies \\ & (w, s) \Rightarrow t \wedge stk' = stk \end{aligned}$$

As mentioned, the induction hypothesis above is not strong enough to conclude the goal. Instead, we continue the proof with complete induction on *n*: we still need to show the same goal, for a new arbitrary *n*, but we get the following additional induction hypothesis for an arbitrary *s*.

$$\begin{aligned} \forall m < n. \quad cs \vdash (0, s, stk) \rightarrow^m (size\ cs, t, stk') \longrightarrow \\ (w, s) \Rightarrow t \wedge stk' = stk \end{aligned}$$

We can now start decomposing the machine code of the *WHILE* loop. Recall the definition of compilation for *WHILE*:

$$\begin{aligned} ccomp\ (WHILE\ b\ DO\ c) = \\ (let\ cc = ccomp\ c; \ cb = bcomp\ b\ False\ (size\ cc + 1) \\ in\ cb\ @\ cc\ @\ [JMP\ (-\ (size\ cb + size\ cc + 1))]) \end{aligned}$$

As in the *IF* case, we start with splitting off the execution of the boolean expression followed by a case distinction on its result *bval b s*. The *False* case is easy, it directly jumps to the exit and we can conclude our goal. In the *True* case, we drop into the execution of *ccomp c*. Here, we first need to consider whether *ccomp c* = [] or not, because our decomposition lemma only applies when the program counter points *into* the code of *ccomp c*, which is not possible when that is empty (e.g., compiled from *SKIP*). If it is empty, the only thing left to do in the *WHILE* loop is to jump back to the beginning. After executing that instruction symbolically, we note that we are now in a situation where our induction hypothesis applies: we have executed at least one instruction (the jump), so $m < n$, and we can directly conclude $(w, s) \Rightarrow t \wedge stk' = stk$. In this specific situation, we could also try to prove that the loop will never terminate and that therefore our assumption that machine execution terminates is *False*, but it is easier to just apply the induction hypothesis here.

The other case, where *ccomp c* $\neq []$, lets us apply the decomposition lemma to isolate the execution of *ccomp c* to some intermediate (s'', stk'') with $pc'' = size\ (ccomp\ c)$. With our induction hypothesis about *c* from the outer induction, we conclude $(c, s) \Rightarrow s''$ and $stk'' = stk$.

Symbolically executing the final jump instruction transports us to the beginning of the compiled code again, into a situation where the inner induction hypothesis applies, because we have again executed at least one instruction, so the number of remaining execution steps *m* is less than the original *n*. That is, we can conclude $(w, s'') \Rightarrow t \wedge stk' = stk$. Together with *bval b s* and $(c, s) \Rightarrow s''$, we arrive at $(w, s) \Rightarrow t \wedge stk' = stk$, which is what we needed to show. \square

Combining the above lemma with the first direction, we get our full compiler correctness theorem:

Theorem 8.19 (Compiler correctness).

$$ccomp\ c \vdash (0, s, stk) \rightarrow^* (size\ (ccomp\ c), t, stk) \longleftrightarrow (c, s) \Rightarrow t$$

Proof. Follows directly from [Lemma 8.18](#), [Lemma 8.9](#), and [Lemma 8.11](#). \square

It is worth pointing out that in a deterministic language like IMP, this second direction reduces to preserving termination: if the machine program terminates, so must the source program. If that were the case, we could conclude that, starting in s , the source must terminate in some t' . Using the first direction of the compiler proof and by determinism of the machine language, we could then conclude that $t' = t$ and that therefore the source execution was already the right one. However, showing that machine-level termination implies source-level termination is not much easier than showing the second direction of our compiler proof directly, and so we did not take this path here.

8.5 Summary and Further Reading

This section has shown the correctness of a simple, non-optimizing compiler from IMP to an idealised machine language. The main technical challenge was reasoning about machine code sequences and their composition and decomposition. Apart from this technical hurdle, formally proving the correctness of such a simple compiler is not as difficult as one might initially suspect.

Two related compiler correctness proofs in the literature are the compiler in the Java-like language Jinja [48] in a style that is similar to the one presented here, and, more recently, the fully realistic, optimizing C compiler CompCert by Leroy *et al.* [53], which compiles to PowerPC, x86, and ARM architectures. Both proofs are naturally more complex than the one presented here, both working with the concept of intermediate languages and multiple compilation stages. This is done to simplify the argument and to concentrate on specific issues on each level.

Modelling the machine program as a list of instructions is an abstraction. A real CPU would implement a von Neumann machine, which adds fetching and decoding of instructions to the execution cycle. The main difference is that our model does not admit self-modifying programs, which is not necessary for IMP. It is entirely possible to model low-level machine code in a theorem prover. The CompCert project has done this as part of its compiler correctness statement, but there are also other independent machine language models available in the literature, for instance a very detailed and well-validated model of multiple ARM processor versions from Cambridge [31, 32], and a similarly detailed, but less complete model of the Intel x86 instruction set architecture by Morrisett *et al.* [57].

For our compiler correctness, we presented a proof of semantics preservation in both directions: from source to machine and from machine to source. Jinja only presents one direction; CompCert does both, but uses a different style of argument for the more involved direction from machine to source,

which involves interpreting the semantics co-inductively to include reasoning about non-terminating programs directly.

Exercises

For the following exercises copy and adjust theory *Compiler*. Intrepid readers only should attempt to adjust theory *Compiler2* too.

Exercise 8.1. Modify the definition of *ccomp* such that it generates fewer instructions for commands of the form *IF b THEN c ELSE SKIP*. Adjust the proof of [Lemma 8.9](#) if needed.

Exercise 8.2. Building on [Exercise 7.8](#), extend the compiler *ccomp* and its correctness theorem *ccomp_bigstep* to *REPEAT* loops. Hint: the recursion pattern of the big-step semantics and the compiler for *REPEAT* should match.

Exercise 8.3. Modify the machine language such that instead of variable names to values, the machine state maps addresses (integers) to values. Adjust the compiler and its proof accordingly.

In the simple version of this exercise, assume the existence of a globally bijective function *addr_of* :: *vname* \Rightarrow *int* with *bij addr_of* to adjust the compiler. Use the search facility of the interface to find applicable theorems for bijective functions.

For the more advanced version and a slightly larger project, assume that the function works only on a finite set of variables: those that occur in the program. For the other, unused variables, it should return a suitable default address. In this version, you may want to split the work into two parts: first, update the compiler and machine language, assuming the existence of such a function and the (partial) inverse it provides. Second, separately construct this function from the input program, having extracted the properties needed for it in the first part. In the end, rearrange your theory file to combine both into a final theorem.

Exercise 8.4. This is a slightly more challenging project. Based on [Exercise 8.3](#), and similarly to [Exercise 3.11](#) and [Exercise 3.12](#), define a second machine language that does not possess a built-in stack, but a stack pointer register in addition to the program counter. Operations that previously worked on the stack now work on memory, accessing locations based on the stack pointer.

For instance, let (pc, s, sp) be a configuration of this new machine consisting of program counter, store, and stack pointer. Then the configuration after

an *ADD* instruction is $(pc + 1, s(sp + 1) := s(sp + 1) + s\ sp), sp + 1)$, that is, *ADD* dereferences the memory at $sp + 1$ and sp , adds these two values and stores them at $sp + 1$, updating the values on the stack. It also increases the stack pointer by 1 to pop one value from the stack and leave the result at the top of the stack. This means the stack grows downwards.

Modify the compiler from [Exercise 8.3](#) to work on this new machine language. Reformulate and reprove the easy direction of compiler correctness.

Hint: Let the stack start below 0, growing downwards, and use type *nat* for addressing variable in *LOAD* and *STORE* instructions, so that it is clear by type that these instructions do not interfere with the stack.

Hint: When the new machine pops a value from the stack, this now unused value is left behind in the store. This means, even after executing a purely arithmetic expression, the values in initial and final stores are not all equal. But they are equal above a given address. Define an abbreviation for this concept and use it to express the intermediate correctness statements.

Types

This chapter introduces types into IMP, first a traditional programming language type system, then more sophisticated type systems for information flow analysis.

Why bother with types? Because they prevent mistakes. They are a simple, automatic way to find obvious problems in programs before these programs are ever run.

There are three kinds of types.

The Good Static types that *guarantee* absence of certain runtime faults.

The Bad Static types that have mostly decorative value but do not guarantee anything at runtime.

The Ugly Dynamic types that detect errors only when it can be too late.

Examples of the first kind are Java, ML and Haskell. In Java, for instance, the type system enforces that there will be no memory access errors, which in other languages manifest themselves as segmentation faults. Haskell has an even more powerful type system, in which, for instance, it becomes visible whether a function can perform input/output actions or not.

Famous examples of the bad kind are C and C++. These languages have static type systems, but they can be circumvented easily. The language specification may not even allow these circumventions, but there is no way for compilers to guarantee their absence.

Examples for dynamic types are scripting languages such as Perl and Python. These languages are typed, but typing violations are discovered and reported at runtime only, which leads to runtime messages such as “*TypeError*: ...” in Python for instance.

In all of the above cases, types are useful. Even in Perl and Python, they at least are known at runtime and can be used to conveniently convert values of one type into another and to enable object-oriented features such as dynamic dispatch of method calls. They just don’t provide any compile-time checking.

In C and C++, the compiler can at least report some errors already at compile time and alert the programmer to obvious problems. But only static, sound type systems can enforce the absence of whole classes of runtime errors.

Static type systems can be seen as proof systems, type checking as proof checking, and type inference as proof search. Every time a type checker passes a program, it in effect proves a set of small theorems about this program.

The ideal static type system is permissive enough not to get in the programmer's way, yet strong enough to guarantee Robin Milner's slogan

Well-typed programs cannot go wrong [56].

It is the most influential slogan and one of the most influential papers in programming language theory.

What could go wrong? Some examples of common runtime errors are corruption of data, null pointer exceptions, nontermination, running out of memory, and leaking secrets. There exist type systems for all of these, and more, but in practice only the first is covered in widely used languages such as Java, C#, Haskell, or ML. We will cover this first kind in [Section 9.1](#), and information leakage in [Section 9.2](#).

As mentioned above, the ideal for a language is to be **type safe**. Type safe means that the execution of a well-typed program cannot lead to certain errors. Java and the JVM, for instance, have been *proved* to be type safe. An execution of a Java program may throw legitimate language exceptions such as `NullPointerException` or `OutOfMemory`, but it can never produce data corruption or segmentation faults other than by hardware defects or calls into native code. In the following sections we will show how to prove such theorems for IMP.

Type safety is a feature of a programming language. **Type soundness** means the same thing, but talks about the type system instead. It means that a type system is *sound* or *correct* with respect to the semantics of the language: If the type system says yes, the semantics does not lead to an error. The semantics is the primary definition of behaviour, and therefore the type system must be justified w.r.t. it.

If there is soundness, how about completeness? Remember Rice's theorem:

Nontrivial semantic properties of programs are undecidable.

Hence there is no (decidable) type system that accepts precisely the programs that have a certain semantic property, e.g., termination.

*Automatic analysis of semantic program properties
is necessarily incomplete.*

This applies not only to type systems but to all automatic semantic analyses and is discussed in more detail at the beginning of the next chapter.

9.1 Typed IMP thy

In this section we develop a very basic static type system as a typical application of programming language semantics. The idea is to define the type system formally and to use the semantics for stating and proving its soundness.


The IMP language we have used so far is not well suited for this proof, because it has only one type of value. This is not enough for even a simple type system. To make things at least slightly non-trivial, we invent a new language that computes on real numbers as well as integers.

To define this new language, we go through the complete exercise again, and define new arithmetic and boolean expressions, together with their values and semantics, as well as a new semantics for commands. In the theorem prover we can do this by merely copying the original definitions and tweaking them slightly. Here, we will briefly walk through the new definitions step by step.

We begin with values occurring in the language. Our introduction of a second kind of value means our value type now correspondingly has two alternatives:

```
datatype val = Iv int | Rv real
```

This definition means we tag values with their type at runtime (the constructor tells us which is which). We do this so we can observe when things go wrong, for instance when a program is trying to add an integer to a real. This does not mean that a compiler for this language would also need to carry this information around at runtime. In fact, it is the type system that lets us avoid this overhead! Since it will only admit safe programs, the compiler can optimize and blindly apply the operation for the correct type. It can determine statically what that correct type is.

 Note that the type *real* stands for the mathematical real numbers, not floating point numbers, just as we use mathematical integers in IMP instead of finite machine words. For the purposes of the type system this difference does not matter. For formalizing a real programming language, one should model values more precisely.

Continuing in the formalization of our new type language, variable names and state stay as they are, i.e., variable names are strings and the state is a function from such names to values.

Arithmetic expressions, however, now have two kinds of constants: *int* and *real*:

```
datatype aexp = Ic int | Rc real | V vname | Plus aexp aexp
```

In contrast to vanilla IMP, we can now write arithmetic expressions that make no sense, or in other words have no semantics. For example, the expression

$$\begin{array}{c}
\frac{}{taval (Ic\ i)\ s\ (Iv\ i)} \quad \frac{}{taval (Rc\ r)\ s\ (Rv\ r)} \quad \frac{}{taval (V\ x)\ s\ (s\ x)} \\
\\
\frac{taval\ a_1\ s\ (Iv\ i_1) \quad taval\ a_2\ s\ (Iv\ i_2)}{taval\ (Plus\ a_1\ a_2)\ s\ (Iv\ (i_1 + i_2))} \\
\\
\frac{taval\ a_1\ s\ (Rv\ r_1) \quad taval\ a_2\ s\ (Rv\ r_2)}{taval\ (Plus\ a_1\ a_2)\ s\ (Rv\ (r_1 + r_2))}
\end{array}$$

Fig. 9.1. Inductive definition of $taval :: aexp \Rightarrow state \Rightarrow val \Rightarrow bool$

Plus (Ic 1) (Rc 3) tries to add an integer to a real number. Assuming for a moment that these are fundamentally incompatible types that cannot possibly be added, this expression makes no sense. We would like to express in our semantics that this is not an expression with well-defined behaviour. One alternative would be to continue using a functional style of semantics for expressions. In this style we would now return *val option* with the constructor *None* of the option data type introduced in Section 2.3.1 to denote the undefined cases. It is quite possible to do so, and in later chapters we will demonstrate that variant. However, it implies that we would have to explicitly enumerate all undefined cases.

It is more elegant and concise to only write down the cases that make sense and leave everything else undefined. The operational semantics judgement already lets us do this for commands. We can use the same style for arithmetic expressions. Since we are not interested in intermediate states at this point, we choose the big-step style.

Our new judgement relates an expression and the state it is evaluated in to the value it is evaluated to. We refrain from introducing additional syntax and call this judgement *taval* for *typed arithmetic value* of an expression. In Isabelle, this translates to an inductive definition with type $aexp \Rightarrow state \Rightarrow val \Rightarrow bool$. We show its introduction rules in Figure 9.1. The term *taval a s v* means that arithmetic expression *a* evaluates in state *s* to value *v*.

The definition is straightforward. The first rule *taval (Ic i) s (Iv i)* for instance says that an integer constant *Ic i* always evaluates to the value *Iv i*, no matter what the state is. The interesting cases are the rules that are not there. For instance, there is no rule to add a *real* to an *int*. We only needed to provide rules for the cases that make sense and we have implicitly defined what the error cases are. The following is an example derivation for *taval* where $s''x'' = Iv\ 4$.

$$\frac{\frac{}{taval\ (Ic\ 3)\ s\ (Iv\ 3)} \quad \frac{}{taval\ (V\ ''x'')\ s\ (Iv\ 4)}}{taval\ (Plus\ (Ic\ 3)\ (V\ ''x''))\ s\ (Iv\ 7)}$$

$$\begin{array}{c}
\frac{}{tbval (Bc\ v)\ s\ v} \qquad \frac{tbval\ b\ s\ bv}{tbval (Not\ b)\ s\ (\neg\ bv)} \\
\\
\frac{tbval\ b_1\ s\ bv_1 \quad tbval\ b_2\ s\ bv_2}{tbval (And\ b_1\ b_2)\ s\ (bv_1 \wedge bv_2)} \qquad \frac{taval\ a_1\ s\ (Iv\ i_1) \quad taval\ a_2\ s\ (Iv\ i_2)}{tbval (Less\ a_1\ a_2)\ s\ (i_1 < i_2)} \\
\\
\frac{taval\ a_1\ s\ (Rv\ r_1) \quad taval\ a_2\ s\ (Rv\ r_2)}{tbval (Less\ a_1\ a_2)\ s\ (r_1 < r_2)}
\end{array}$$

Fig. 9.2. Inductive definition of $tbval :: bexp \Rightarrow state \Rightarrow bool \Rightarrow bool$

For $s\ 'x' = Rv\ 3$ on the other hand, there would be no execution of $taval$ that we could derive for the same term.

The syntax for boolean expressions remains unchanged. Their evaluation, however, is different. In order to use the operational semantics for arithmetic expressions that we just defined, we need to employ the same operational style for boolean expressions. Figure 9.2 shows the formal definition. Normal evaluation is straightforward. Then there are the two (missing) error cases $Less\ (Ic\ n)\ (Rc\ r)$ and $Less\ (Rc\ r)\ (Ic\ n)$. Moreover the definition also propagates errors from the evaluation of arithmetic expressions: if there is no evaluation for a_1 or a_2 then there is also no evaluation for $Less\ a_1\ a_2$.

The syntax for commands is again unchanged. We now have a choice: do we define a big-step or a small-step semantics? The answer seems clear: it must be small-step semantics, because only there can we observe when things are going wrong in the middle of an execution. In the small-step case, error states are explicitly visible in intermediate states: if there is an error, the semantics gets stuck in a non-final configuration with no further progress possible. We need executions *to be able to go wrong* if we want a meaningful proof that they do not.

In fact, the big-step semantics could be adjusted as well to perform the same function. By default, in the style we have seen so far, a big-step semantics is not suitable for this, because it conflates non-termination, which is allowed, with runtime errors or undefined execution, which are not. If we mark errors specifically and distinguish them from non-termination in the big-step semantics, we can observe errors just as well as in the small-step case.

So we still have a choice. Small-step semantics are more concise and more traditional for type soundness proofs. Therefore we will choose this one. Later, in Chapter 10, we will show the other alternative.

After all this discussion, the definition of the small-step semantics for typed commands is almost the same as that for the untyped case. As shown in Figure 9.3, it merely refers to the new judgements for arithmetic and boolean expressions, but does not add any new rules on its own.

$$\begin{array}{c}
\frac{taval\ a\ s\ v}{(x ::= a, s) \rightarrow (SKIP, s(x := v))} \\
\\
\frac{}{(SKIP;;\ c, s) \rightarrow (c, s)} \quad \frac{(c_1, s) \rightarrow (c'_1, s')}{(c_1;;\ c_2, s) \rightarrow (c'_1;;\ c_2, s')} \\
\\
\frac{tbval\ b\ s\ True}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow (c_1, s)} \\
\\
\frac{tbval\ b\ s\ False}{(IF\ b\ THEN\ c_1\ ELSE\ c_2, s) \rightarrow (c_2, s)} \\
\\
\frac{}{(WHILE\ b\ DO\ c, s) \rightarrow (IF\ b\ THEN\ c;;\ WHILE\ b\ DO\ c\ ELSE\ SKIP, s)}
\end{array}$$

Fig. 9.3. Inductive definition of $(\rightarrow) :: com \times state \Rightarrow com \times state \Rightarrow bool$

As before, the execution of a command is a sequence of small steps, denoted by star, for example $(c, s) \rightarrow^* (c', s')$.

Example 9.1. For well-behaved programs, our typed executions look as before. For instance, let s satisfy $s\ 'y'' = Iv\ 7$. Then we get the following example execution chain.

$$\begin{aligned}
&('x'' ::= V\ 'y'';;\ 'y'' ::= Plus\ (V\ 'x'')\ (V\ 'y''), s) \rightarrow \\
&(SKIP;;\ 'y'' ::= Plus\ (V\ 'x'')\ (V\ 'y''), s('x'' := Iv\ 7)) \rightarrow \\
&('y'' ::= Plus\ (V\ 'x'')\ (V\ 'y''), s('x'' := Iv\ 7)) \rightarrow \\
&(SKIP, s('x'' := Iv\ 7, 'y'' := Iv\ 14))
\end{aligned}$$

However, programs that contain type errors can get stuck. For example, if in the same state s we take a slightly different program that adds a constant of the wrong type, we get:

$$\begin{aligned}
&('x'' ::= V\ 'y'';;\ 'y'' ::= Plus\ (V\ 'x'')\ (Rc\ 3), s) \rightarrow \\
&(SKIP;;\ 'y'' ::= Plus\ (V\ 'x'')\ (Rc\ 3), s('x'' := Iv\ 7)) \rightarrow \\
&('y'' ::= Plus\ (V\ 'x'')\ (Rc\ 3), s('x'' := Iv\ 7))
\end{aligned}$$

The first assignment succeeds as before, but after that no further execution step is possible because we cannot find an execution for *taval* on the right-hand side of the second assignment.

9.1.1 The Type System

Having defined our new language above, we can now define its type system. The idea of such type systems is to predict statically which values will appear at runtime and to exclude programs in which unsafe values or value combinations might be encountered.

$$\begin{array}{c}
\overline{\Gamma \vdash Ic\ i : Ity} \quad \overline{\Gamma \vdash Rc\ r : Rty} \quad \overline{\Gamma \vdash V\ x : \Gamma\ x} \\
\\
\frac{\Gamma \vdash a_1 : \tau \quad \Gamma \vdash a_2 : \tau}{\Gamma \vdash Plus\ a_1\ a_2 : \tau}
\end{array}$$

Fig. 9.4. Inductive definition of $_ \vdash _ : _ :: tyenv \Rightarrow aexp \Rightarrow ty \Rightarrow bool$

$$\begin{array}{c}
\overline{\Gamma \vdash Bc\ v} \quad \frac{\Gamma \vdash b}{\Gamma \vdash Not\ b} \\
\\
\frac{\Gamma \vdash b_1 \quad \Gamma \vdash b_2}{\Gamma \vdash And\ b_1\ b_2} \quad \frac{\Gamma \vdash a_1 : \tau \quad \Gamma \vdash a_2 : \tau}{\Gamma \vdash Less\ a_1\ a_2}
\end{array}$$

Fig. 9.5. Inductive definition of $(\vdash) :: tyenv \Rightarrow bexp \Rightarrow bool$

The type system we use for this is very rudimentary, it has only two types: *int* and *real*, written as the constructors *Ity* and *Rty*, corresponding to the two kinds of values we have introduced. In Isabelle:

datatype *ty* = *Ity* | *Rty*

The purpose of the type system is to keep track of the type of each variable and to allow only compatible combinations in expressions. For this purpose, we define a so-called typing environment. Where a runtime state maps variable names to values, a static typing environment maps variable names to their static types.

type_synonym *tyenv* = *vname* \Rightarrow *ty*

For example, we could have $\Gamma\ 'x' = Ity$, telling us that variable *x* has type integer and that we should therefore not use it in an expression of type real.

With this, we can give typing rules for arithmetic expressions. The idea is simple: constants have fixed type, variables have the type the typing environment Γ prescribes, and *Plus* can be typed with type τ if both operands have the same type τ . Figure 9.4 shows the definition in Isabelle. We use the notation $\Gamma \vdash a : \tau$ to say that expression *a* has type τ in context Γ .

The typing rules for booleans in Figure 9.5 are even simpler. We do not need a result type, because it will always be *bool*, so the notation is just $\Gamma \vdash b$ for *expression b is well-typed in context* Γ . For the most part, we just need to capture that boolean expressions are well-typed if their subexpressions are well-typed. The interesting case is the connection to arithmetic expressions in *Less*. Here we demand that both operands have the same type τ , i.e., either we compare two ints or two reals, but not an int to a real.

Similarly, commands are well-typed if their subcommands and subexpressions are well-typed. In addition, in an assignment the arithmetic expression

$$\begin{array}{c}
\frac{}{\Gamma \vdash \text{SKIP}} \quad \frac{\Gamma \vdash a : \Gamma x}{\Gamma \vdash x ::= a} \\
\\
\frac{\Gamma \vdash c_1 \quad \Gamma \vdash c_2}{\Gamma \vdash c_1;; c_2} \quad \frac{\Gamma \vdash b \quad \Gamma \vdash c_1 \quad \Gamma \vdash c_2}{\Gamma \vdash \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2} \quad \frac{\Gamma \vdash b \quad \Gamma \vdash c}{\Gamma \vdash \text{WHILE } b \text{ DO } c}
\end{array}$$

Fig. 9.6. Inductive definition of $(\vdash) :: \text{tyenv} \Rightarrow \text{com} \Rightarrow \text{bool}$

must have the same type as the variable it is assigned to. The full set of rules is shown in Figure 9.6. We re-use the syntax $\Gamma \vdash c$ for *command* c is *well-typed in context* Γ .

This concludes the definition of the type system itself. Type systems can be arbitrarily complex. The one here is intentionally simple to show the structure of a type soundness proof without getting side-tracked in interesting type system details.

Note that there is precisely one rule per syntactic construct in our definition of the type system, and the premises of each rule apply the typing judgement only to subterms of the conclusion. We call such rule sets **syntax-directed**. Syntax-directed rules are a good candidate for automatic application and for deriving an algorithm that infers the type simply by applying them backwards, at least if there are no side conditions in their assumptions. Since there is exactly one rule per construct, it is always clear which rule to pick and there is no need for back-tracking. Further, since there is always at most one rule application per syntax node in the term or expression the rules are applied to, this process must terminate. This idea can be extended to allow side conditions in the assumptions of rules, as long as these side conditions are decidable.

Given such a type system, we can now check whether a specific command c is well-typed. To do so, we merely need to construct a derivation tree for the judgment $\Gamma \vdash c$. Such a derivation tree is also called a **type derivation**. Let for instance $\Gamma \vdash x'' = \text{Itty}$ as well as $\Gamma \vdash y'' = \text{Itty}$. Then our previous example program is well-typed, because of the following type derivation.

$$\begin{array}{c}
\frac{\Gamma \vdash y'' = \text{Itty}}{\Gamma \vdash V \ y'' : \text{Itty}} \quad \frac{\Gamma \vdash x'' = \text{Itty}}{\Gamma \vdash V \ x'' : \text{Itty}} \quad \frac{\Gamma \vdash y'' = \text{Itty}}{\Gamma \vdash V \ y'' : \text{Itty}} \\
\\
\frac{\Gamma \vdash V \ x'' ::= V \ y''}{\Gamma \vdash x'' ::= V \ y''} \quad \frac{\Gamma \vdash V \ x'' : \text{Itty} \quad \Gamma \vdash V \ y'' : \text{Itty}}{\Gamma \vdash \text{Plus } (V \ x'') (V \ y'') : \text{Itty}} \\
\\
\frac{\Gamma \vdash x'' ::= V \ y'' \quad \Gamma \vdash y'' ::= \text{Plus } (V \ x'') (V \ y'')}{\Gamma \vdash x'' ::= V \ y'';; y'' ::= \text{Plus } (V \ x'') (V \ y'')}
\end{array}$$

9.1.2 Well-Typed Programs Do Not Get Stuck

In this section we prove that the type system defined above is sound. As mentioned in the introduction to this chapter, Robert Milner coined the phrase *Well-typed programs cannot go wrong*, i.e., well-typed programs will not exhibit any runtime errors such as segmentation faults or undefined execution. In our small-step semantics we have defined precisely what “go wrong” means formally: a program exhibits a runtime error when the semantics gets stuck.

To prove type soundness we have to prove that well-typed programs never get stuck. They either terminate successfully, or they make further progress. Taken literally, the above sentence translates into the following property:

$$\llbracket (c, s) \rightarrow^* (c', s'); \Gamma \vdash c \rrbracket \implies c' = \text{SKIP} \vee (\exists cs''. (c', s') \rightarrow cs'')$$

Given an arbitrary command c , which is well-typed $\Gamma \vdash c$, any execution $(c, s) \rightarrow^* (c', s')$ either has terminated successfully with $c' = \text{SKIP}$, or can make another execution step $\exists cs''. (c', s') \rightarrow cs''$. Clearly, this statement is wrong, though: take c for instance to be a command that computes the sum of two variables: $z := x + y$. This command is well-typed, for example, if the variables are both of type `int`. However, if we start the command in a state that disagrees with this type, e.g., where x contains an `int` but y contains a `real`, the execution gets stuck.

Of course, we want the value of a variable to be of type `int` if the typing says it should be `int`. This means we want not only the program to be well-typed, but the state to be well-typed too.

We so far have the state assigning values to variables and we have the type system statically assigning types to variables in the program. The concept of well-typed states connects these two: we define a judgement that determines if a runtime state is compatible with a typing environment for variables. We call this formal judgement *styping* below, written $\Gamma \vdash s$. We equivalently also say that a state s **conforms** to a typing environment Γ .

With this judgement, our full statement of type soundness is now

$$\llbracket (c, s) \rightarrow^* (c', s'); \Gamma \vdash c; \Gamma \vdash s; c' \neq \text{SKIP} \rrbracket \implies \exists cs''. (c', s') \rightarrow cs''$$

Given a well-typed program, started in a well-typed state, any execution that has not reached `SKIP` yet can make another step.

We will prove this property by induction on the reflexive transitive closure of execution steps, which naturally decomposes this type soundness property into two parts: preservation and progress. **Preservation** means that well-typed states stay well-typed during execution. **Progress** means that in a well-typed state, the program either terminates successfully or can make one more step of execution progress.

In the following, we formalize the soundness proof for typed IMP.

We start the formalization by defining a function from values to types:

```

fun type :: val  $\Rightarrow$  ty where
  type (Iv i) = Ity
  type (Rv r) = Rty

```

Our *styping* judgement for well-typed states is now very simple: for all variables, the type of the runtime value must be exactly the type predicted in the typing environment.

```

definition ( $\vdash$ ) :: tyenv  $\Rightarrow$  state  $\Rightarrow$  bool where
   $\Gamma \vdash s \iff (\forall x. \text{type } (s\ x) = \Gamma\ x)$ 

```

This was easy. In more sophisticated type systems, there may be multiple types that can be assigned to a value and we may need a compatibility or subtype relation between types to define the *styping* judgement.

We now have everything defined to start the soundness proof. The plan is to prove progress and preservation, and to conclude from that that the final type soundness statement that an execution of a well-typed command started in a well-typed state will never get stuck. To prove progress and preservation for commands, we will first need the same properties for arithmetic and boolean expressions.

Preservation for arithmetic expressions means the following: if expression a has type τ under environment Γ , if a evaluates to v in state s , and if s conforms to Γ , then the type of the result v must be τ :

Lemma 9.2 (Preservation for arithmetic expressions).

$\llbracket \Gamma \vdash a : \tau; \text{taeval } a\ s\ v; \Gamma \vdash s \rrbracket \implies \text{type } v = \tau$

Proof. The proof is by rule induction on the type derivation $\Gamma \vdash a : \tau$. If we declare rule inversion on *taeval* to be used automatically and unfold the definition of *styping*, Isabelle proves the rest. \square

The proof of the progress lemma is slightly more verbose. It is almost the only place where something interesting is concluded in the soundness proof — there is the potential of something going wrong: if the operands of a *Plus* were of incompatible type, there would be no value v the expression evaluates to. Of course, the type system excludes precisely this case.

The progress statement is as standard as the preservation statement for arithmetic expressions: given that a has type τ under environment Γ , and given a conforming state s , there must exist a result value v such that a evaluates to v in s .

Lemma 9.3 (Progress for arithmetic expressions).

$\llbracket \Gamma \vdash a : \tau; \Gamma \vdash s \rrbracket \implies \exists v. \text{taeval } a\ s\ v$

Proof. The proof is again by rule induction on the typing derivation. The interesting case is *Plus* $a_1 \ a_2$. The induction hypothesis gives us two values v_1 and v_2 for the subexpressions a_1 and a_2 . If v_1 is an integer, then, by preservation, the type of a_1 must have been *Ity*. The typing rule says that the type of a_2 must be the same. This means, by preservation, the type of v_2 must be *Ity*, which in turn means then v_2 must be an *Iv* value and we can conclude using the *taval* introduction rule for *Plus* that the execution has a result. Isabelle completes this reasoning chain automatically if we carefully provide it with the right facts and rules. The case for reals is analogous, and the other typing cases are solved automatically. \square

For boolean expressions, there is no preservation lemma, because *tbval*, by its Isabelle type, can only return boolean values. The progress statement makes sense, though, and follows the standard progress statement schema.

Lemma 9.4 (Progress for boolean expressions).

$$\llbracket \Gamma \vdash b; \Gamma \vdash s \rrbracket \implies \exists v. \text{tbval } b \ s \ v$$

Proof. As always, the proof is by rule induction on the typing derivation. The interesting case is where something could go wrong, namely where we execute arithmetic expressions in *Less*. The proof is very similar to the one for *Plus*: we obtain the values of the subexpressions; we perform a case distinction on one of them to reason about its type; we infer the other has the same type by typing rules and by preservation on arithmetic expressions; and we conclude that execution can therefore progress. Again, this case is automatic if written carefully; the other cases are trivial. \square

For commands, there are two preservation statements, because the configurations in our small-step semantics have two components: command and state. We first show that the command remains well-typed and then that the state does. Both proofs are by induction on the small-step semantics. They could be proved by induction on the typing derivation as well. Often it is preferable to try induction on the typing derivation first, because the type system typically has fewer cases. On the other hand, depending on the complexity of the language, the more fine-grained information that is available in the operational semantics might make the more numerous cases easier to prove in the other induction alternative. In both cases it pays off to design the structure of the rules in both systems such that they technically fit together nicely, for instance such that they decompose along the same syntactic lines.

Theorem 9.5 (Preservation: commands stay well-typed).

$$\llbracket (c, s) \rightarrow (c', s'); \Gamma \vdash c \rrbracket \implies \Gamma \vdash c'$$

Proof. The preservation of program typing is fully automatic in this simple language. The only mildly interesting case where we are not just transforming

the command into a subcommand is the while loop. Here we just need to apply the typing rules for IF and sequential composition and are done. \square

Theorem 9.6 (Preservation: states stay well-typed).

$$\llbracket (c, s) \rightarrow (c', s'); \Gamma \vdash c; \Gamma \vdash s \rrbracket \implies \Gamma \vdash s'$$

Proof. Most cases are trivial because the state is not modified. In the second $;;$ rule the induction hypothesis applies. In the assignment rule the state is updated with a new value. Type preservation on expressions gives us that the new value has the same type as the expression, and unfolding the *styping* judgement shows that it is unaffected by state updates that are type preserving. In more complex languages, there are likely to be a number of such update cases and the corresponding lemma is a central piece of type soundness proofs. \square

The next step is the progress lemma for commands. Here, we need to take into account that the program might have fully terminated. If it has not, and we have a well-typed program in a well-typed state, we should be able to make at least one step.

Theorem 9.7 (Progress for commands).

$$\llbracket \Gamma \vdash c; \Gamma \vdash s; c \neq SKIP \rrbracket \implies \exists cs'. (c, s) \rightarrow cs'$$

Proof. This time the only induction alternative is on the typing derivation again. The cases with arithmetic and boolean expressions make use of the corresponding progress lemmas to generate the values the small-step rules demand. For IF , we additionally perform a case distinction for picking the corresponding introduction rule. As for the other cases: $SKIP$ is trivial, sequential composition applies the induction hypotheses and makes a case distinction whether c_1 is $SKIP$ or not, and $WHILE$ always trivially makes progress in the small-step semantics, because it is unfolded into an $IF/WHILE$. \square

All that remains is to assemble the pieces into the final type soundness statement: given any execution of a well-typed program started in a well-typed state, we are not stuck; we have either terminated successfully, or the program can perform another step.

Theorem 9.8 (Type soundness).

$$\llbracket (c, s) \rightarrow^* (c', s'); \Gamma \vdash c; \Gamma \vdash s; c' \neq SKIP \rrbracket \implies \exists cs''. (c', s') \rightarrow cs''$$

Proof. The proof lifts the one-step preservation and progress results to a sequence of steps by induction on the reflexive transitive closure. The base case of zero steps is solved by the progress lemma; the step case needs our two preservation lemmas for commands. \square

This concludes the section on typing. We have seen, exemplified by a very simple type system, what a type soundness statement means, how it interacts with the small-step semantics, and how it is proved. While the proof itself will grow in complexity for more interesting languages, the general schema of progress and preservation remains.

For the type soundness theorem to be meaningful, it is important that the failures the type system is supposed to prevent are observable in the semantics, so that their absence can be shown. In a framework like the one above, the definition of the small-step semantics carries the main meaning and strength of the type soundness statement.

Our mantra for type systems:

Type systems have a purpose: the static analysis of programs in order to predict their runtime behaviour. The correctness of this prediction must be provable.

Exercises

Exercise 9.1. Reformulate the inductive predicates $\Gamma \vdash a : \tau$, $\Gamma \vdash b$ and $\Gamma \vdash c$ as three functions $atype :: tyenv \Rightarrow aexp \Rightarrow ty\ option$, $bok :: tyenv \Rightarrow bexp \Rightarrow bool$ and $cok :: tyenv \Rightarrow com \Rightarrow bool$ and prove the three equivalences $(\Gamma \vdash a : \tau) = (atype\ \Gamma\ a = Some\ \tau)$, $(\Gamma \vdash b) = bok\ \Gamma\ b$ and $(\Gamma \vdash c) = cok\ \Gamma\ c$.

Exercise 9.2. Modify the evaluation and typing of $aexp$ by allowing $ints$ to be coerced to $reals$ with the predefined coercion function $real_of_int :: int \Rightarrow real$ where necessary. Now every $aexp$ has a value and a type. Define an evaluation function $aval :: aexp \Rightarrow state \Rightarrow val$ and a typing function $atyp :: tyenv \Rightarrow aexp \Rightarrow ty$ and prove $\Gamma \vdash s \implies atyp\ \Gamma\ a = type\ (aval\ a\ s)$.

For the following two exercises copy theory *Types* and modify it as required.

Exercise 9.3. Add a *REPEAT* loop (see [Exercise 7.8](#)) to the typed version of IMP and update the type soundness proof.

Exercise 9.4. Modify the typed version of IMP as follows. Values are now either integers or booleans. Thus variables can have boolean values too. Merge the two expression types $aexp$ and $bexp$ into one new type exp of expressions that has the constructors of both types (of course without real constants). Combine $taval$ and $tbval$ into one evaluation predicate $eval :: exp \Rightarrow state \Rightarrow val \Rightarrow bool$. Similarly combine the two typing predicates into one: $\Gamma \vdash e : \tau$ where $e :: exp$ and the IMP-type τ can be one of *Ity* or *Bty*. Adjust the small-step semantics and the type soundness proof.

9.2 Security Type Systems

In the previous section we have seen a simple static type system with soundness proof. However, type systems can be used for more than the traditional concepts of integers, reals, etc. In theory, type systems can be arbitrarily complex logical systems used to statically predict properties of programs. In the following, we will look at a type system that aims to enforce a security property: the absence of information flows from private data to public observers. The idea is that we want an easy and automatic way to check if programs protect private data such as passwords, bank details, or medical records.

Ensuring such **information flow control** properties based on a programming language analysis such as a type system is a part of so-called **language-based security**. Another common option for enforcing information flow control is the use of cryptography to ensure the secrecy of private data. Cryptography only admits probabilistic arguments (one could always guess the key), whereas language-based security also allows more absolute statements. As techniques they are not incompatible: both approaches could be mixed to enforce a particular information flow property.

Note that absolute statements in language-based security are always with respect to assumptions on the execution environment. For instance, our proof below will have the implicit assumption that the machine actually behaves as our semantics predicts. There are practical ways in which these assumptions can be broken or circumvented: intentionally introducing hardware-based errors into the computation to deduce private data, direct physical observation of memory contents, deduction of private data by analysis of execution time, and more. These attacks make use of details that are not visible on the abstraction level of the semantic model our proof is based on — they are **covert channels** of information flow.

9.2.1 Security Levels and Expressions thy

We begin developing our security type system by defining security levels. The idea is that each variable will have an associated security level. The type system will then enforce the policy that information may only flow from variables of ‘lower’ security levels to variables of ‘higher’ levels, but never the other way around.

In the literature, levels are often reduced to just two: high and low. We keep things slightly more general by making levels natural numbers. We can then compare security levels by just writing $<$ and we can compute the maximal or minimal security level of two different variables by taking the maximum or minimum respectively. The term $l < l'$ in this system would mean that l is less private or confidential than l' , so level 0 could be equated with ‘public’.

It would be easy to generalize further and just assume a lattice of security levels with $<$, join, and meet operations. We could then also enforce that information does not travel ‘sideways’ between two incomparable security levels. For the sake of simplicity we refrain from doing so and merely use *nat*.

type_synonym *level* = *nat*

For the type system and security proof below it would be sufficient to merely assume the existence of a HOL constant that maps variables to security levels. This would express that we assume each variable to possess a security level and that this level remains the same during execution of the program.

For the sake of showing examples — the general theory does not rely on it! — we arbitrarily choose a specific function for this mapping: a variable of length n has security level n .

The kinds of information flows we would like to avoid are exemplified by the following two:

- explicit: `low := high`
- implicit: `IF high1 < high2 THEN low := 0 ELSE low := 1`

The property we are after is called **noninterference**: a variation in the value of high variables should not interfere with the computation or values of low variables. ‘High should not interfere with low.’

More formally, a program c guarantees noninterference iff for all states s_1 and s_2 : if s_1 and s_2 agree on low variables (but may differ on high variables!), then the states resulting from executing (c, s_1) and (c, s_2) must also agree on low variables.

As opposed to our previous type soundness statement, this definition compares the outcome of two executions of the same program in different, but related initial states. It requires again potentially different, but equally related final states.

With this in mind, we proceed to define the type system that will enforce this property. We begin by computing the security level of arithmetic and boolean expressions. We are interested in flows from higher to lower variables, so we define the security level of an expression as the highest level of any variable that occurs in it. We make use of Isabelle’s overloading and call the security level of an arithmetic or boolean expression *sec* e .

```

fun sec :: aexp  $\Rightarrow$  level where
  sec (N  $n$ )           = 0
  sec (V  $x$ )           = sec  $x$ 
  sec (Plus  $a_1$   $a_2$ ) = max (sec  $a_1$ ) (sec  $a_2$ )

```

```

fun sec :: be  $\Rightarrow$  level where
  sec (Bc v)      = 0
  sec (Not b)      = sec b
  sec (And b1 b2) = max (sec b1) (sec b2)
  sec (Less a1 a2) = max (sec a1) (sec a2)

```

A first lemma indicating that we are moving into the right direction will be that if we change the value of only variables with a higher level than *sec e*, the value of *e* should stay the same.

To express this property, we introduce notation for two states agreeing on the value of all variables below a certain security level. The concept is light-weight enough that a syntactic abbreviation is sufficient; this avoids having to go through the motions of setting up additional proof infrastructure. We will need \leq , but also the strict $<$ later on, so we define both here:

$$\begin{aligned}
 s = s' (\leq l) &\equiv \forall x. \text{sec } x \leq l \longrightarrow s \ x = s' \ x \\
 s = s' (< l) &\equiv \forall x. \text{sec } x < l \longrightarrow s \ x = s' \ x
 \end{aligned}$$

With this, the proof of our first two security properties is simple and automatic: The evaluation of an expression *e* only depends on variables with level below or equal to *sec e*.

Lemma 9.9 (Noninterference for arithmetic expressions).

$$\llbracket s_1 = s_2 (\leq l); \text{sec } e \leq l \rrbracket \implies \text{aval } e \ s_1 = \text{aval } e \ s_2$$

Lemma 9.10 (Noninterference for boolean expressions).

$$\llbracket s_1 = s_2 (\leq l); \text{sec } b \leq l \rrbracket \implies \text{bval } b \ s_1 = \text{bval } b \ s_2$$

9.2.2 Syntax-Directed Typing thy

As usual in IMP, the typing for expressions was simple. We now define a syntax-directed set of security typing rules for commands. This makes the rules directly executable and allows us to run examples. Checking for explicit flows, i.e., assignments from high to low variables, is easy. For implicit flows, the main idea of the type system is to track the security level of variables that decisions are made on, and to make sure that their level is lower than or equal to variables assigned to in that context.

We write $l \vdash c$ to mean that command *c* contains no information flows to variables lower than level *l*, and only safe flows to variables $\geq l$.

Going through the rules of Figure 9.7 in detail, we have defined *SKIP* to be safe at any level. We have defined assignment to be safe if the level of *x* is higher than or equal to both the level of the information source *a* and the level *l*. For semicolon to conform to level *l*, we recursively demand that both parts conform to the same level *l*. As previously shown in the motivating example,

$$\begin{array}{c}
\frac{}{l \vdash \text{SKIP}} \quad \frac{\text{sec } a \leq \text{sec } x \quad l \leq \text{sec } x}{l \vdash x ::= a} \quad \frac{l \vdash c_1 \quad l \vdash c_2}{l \vdash c_1;; c_2} \\
\\
\frac{\max(\text{sec } b) \, l \vdash c_1 \quad \max(\text{sec } b) \, l \vdash c_2}{l \vdash \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2} \quad \frac{\max(\text{sec } b) \, l \vdash c}{l \vdash \text{WHILE } b \text{ DO } c}
\end{array}$$

Fig. 9.7. Definition of $\text{sec_type} :: \text{level} \Rightarrow \text{com} \Rightarrow \text{bool}$

the *IF* command could admit implicit flows. We prevent these by demanding that for *IF* to conform to l , both c_1 and c_2 have to conform to level l or the level of the boolean expression, whichever is higher. We can conveniently express this with the maximum operator \max . The *WHILE* case is similar to an *IF*: the body must have at least the level of b and of the whole command.

Using the \max function makes the type system executable if we tell Isabelle to treat the level and the program as input to the predicate.

Example 9.11. Testing our intuition about what we have just defined, we look at four examples for various security levels.

$0 \vdash \text{IF Less } (V \text{ ''x1''}) (V \text{ ''x''}) \text{ THEN ''x1''} ::= N \ 0 \text{ ELSE SKIP}$

The statement claims that the command is well-typed at security level 0: flows can occur down to even a level 0 variable, but they have to be internally consistent, i.e., flows must still only be from lower to higher levels. According to our arbitrary example definition of security levels that assigns the length of the variable to the level, variable $x1$ has level 2, and variable x has level 1. This means the evaluation of this typing expression will yield *True*: the condition has level 2, and the context is 0, so according to the *IF* rule, both commands must be safe up to level 2, which is the case, because the first assignment sets a level 2 variable, and the second is just *SKIP*.

Does the same work if we assign to a lower-level variable?

$0 \vdash \text{IF Less } (V \text{ ''x1''}) (V \text{ ''x''}) \text{ THEN ''x''} ::= N \ 0 \text{ ELSE SKIP}$

Clearly not. Again, we need to look at the *IF* rule, which still says the assignment must be safe at level 2, i.e., we have to derive $2 \vdash \text{''x''} ::= N \ 0$. But x is of level 1, and the assignment rule demands that we only assign to levels higher than the context. Intuitively, the *IF* decision expression reveals information about a level 2 variable. If we assign to a level 1 variable in one of the branches we leak level 2 information to level 1.

What if we demand a higher security context from our original example?

$2 \vdash \text{IF Less } (V \text{ ''x1''}) (V \text{ ''x''}) \text{ THEN ''x1''} ::= N \ 0 \text{ ELSE SKIP}$

Context of level 2 still works, because our highest level in this command is level 2, and our arguments from the first example still apply.

What if we go one level higher?

$3 \vdash \text{IF Less } (V \text{ ''x1''}) (V \text{ ''x''}) \text{ THEN ''x1''} ::= N \ 0 \text{ ELSE SKIP}$

Now we get *False*, because we need to take the maximum of the context and the boolean expression for evaluating the branches. The intuition is that the context gives the minimum level to which we may reveal information.

As we can already see from these simple examples, the type system is not complete: it will reject some safe programs as unsafe. For instance, if the value of x in the second command was already 0 in the beginning, the context would not have mattered, we only would have overwritten 0 with 0. As we know by now, we should not expect otherwise. The best we can hope for is a safe approximation such that the false alarms are hopefully programs that rarely occur in practice or that can be rewritten easily.

It is the case that the simple type system presented here, going back to Volpano, Irvine, and Smith [89], has been criticised as too restrictive. It excludes too many safe programs. This can be addressed by making the type system more refined, more flexible, and more context-aware. For demonstrating the type system and its soundness proof in this book, however, we will stick to its simplest form.

9.2.3 Soundness

We introduced the correctness statement for this type system as noninterference: two executions of the same program started in related states end up in related states. The relation in our case is that the values of variables below security level l are the same. Formally, this is the following statement:

$$\llbracket (c, s) \Rightarrow s'; (c, t) \Rightarrow t'; 0 \vdash c; s = t (\leq l) \rrbracket \implies s' = t' (\leq l)$$

An important property, which will be useful for this lemma, is the so-called **anti-monotonicity** of the type system: a command that is typeable in l is also typeable in any level smaller than l . Anti-monotonicity is also often called the **subsumption rule**, to say that higher contexts subsume lower ones. Intuitively it is clear that this property should hold: we defined $l \vdash c$ to mean that there are no flows to variables $< l$. If we write $l' \vdash c$ with an $l' \leq l$, then we are only admitting more flows, i.e., we are making a weaker statement.

Lemma 9.12 (Anti-monotonicity). $\llbracket l \vdash c; l' \leq l \rrbracket \implies l' \vdash c$

Proof. The formal proof is by rule induction on the type system. Each of the cases is then solved automatically. \square

The second key lemma in the argument for the soundness of our security type system is **confinement**: an execution of a command that is type correct in context l can only change variables of level l and above, or conversely, all variables below l will remain unchanged. In other words, the effect of c is **confined** to variables of level $\geq l$.

Lemma 9.13 (Confinement). $\llbracket (c, s) \Rightarrow t; l \vdash c \rrbracket \implies s = t (< l)$

Proof. The first instinct may be to try rule induction on the type system again, but the *WHILE* case will only give us an induction hypothesis about the body when we will have to show our goal for the whole loop. Therefore, we choose rule induction on the big-step execution instead. In the *IF* and *WHILE* cases, we make use of anti-monotonicity to instantiate the induction hypothesis. In the *IfTrue* case, for instance, the hypothesis is $l \vdash c_1 \implies s = t (< l)$, but from the type system we only know $\max(\text{sec } b) \vdash c_1$. Since $l \leq \max(\text{sec } b)$, anti-monotonicity allows us to conclude $l \vdash c_1$. \square

With these two lemmas, we can start the main noninterference proof.

Theorem 9.14 (Noninterference).

$$\llbracket (c, s) \Rightarrow s'; (c, t) \Rightarrow t'; 0 \vdash c; s = t (\leq l) \rrbracket \implies s' = t' (\leq l)$$

Proof. The proof is again by induction on the big-step execution. The *SKIP* case is easy and automatic, as it should be.

The assignment case is already somewhat interesting. First, we note that s' is the usual state update $s(x := \text{aval } a \ s)$ in the first big-step execution. We perform rule inversion for the second execution to get the same update for t . We also perform rule inversion on the typing statement to get the relationship between security levels of x and a : $\text{sec } a \leq \text{sec } x$. Now we show that the two updated states s' and t' still agree on all variables below l . For this, it is sufficient to show that the states agree on the new value if $\text{sec } x < l$, and that all other variables y with $\text{sec } y < l$ still agree as before. In the first case, looking at x , we know from above that $\text{sec } a \leq \text{sec } x$. Hence, by transitivity, we have that $\text{sec } a \leq l$. This is enough for our noninterference result on expressions to apply, given that we also know $s = t (\leq l)$ from the premises. This means, we get $\text{aval } a \ s = \text{aval } a \ t$: the new values for x agree as required. The case for all other variables y below l follows directly from $s = t (\leq l)$.

In the semicolon case, we merely need to compose the induction hypotheses. This is solved automatically.

IF has two symmetric cases as usual. We will look only at the *IfTrue* case in more detail. We begin the case by noting via rule inversion that both branches are type correct to level $\text{sec } b$, since the maximum with 0 is the

$$\begin{array}{c}
\frac{}{l \vdash' \text{SKIP}} \qquad \frac{\text{sec } a \leq \text{sec } x \quad l \leq \text{sec } x}{l \vdash' x ::= a} \\
\\
\frac{l \vdash' c_1 \quad l \vdash' c_2}{l \vdash' c_1;; c_2} \qquad \frac{\text{sec } b \leq l \quad l \vdash' c_1 \quad l \vdash' c_2}{l \vdash' \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2} \\
\\
\frac{\text{sec } b \leq l \quad l \vdash' c}{l \vdash' \text{WHILE } b \text{ DO } c} \qquad \frac{l \vdash' c \quad l' \leq l}{l' \vdash' c}
\end{array}$$

Fig. 9.8. Definition of $\text{sec_type}' :: \text{level} \Rightarrow \text{com} \Rightarrow \text{bool}$

identity, i.e., we know $\text{sec } b \vdash c_1$. Then we perform a case distinction: either the level of b is $\leq l$ or it is not. If $\text{sec } b \leq l$, i.e., the *IF* decision is on a more public level than l , then s and t , which agree below l , also agree below $\text{sec } b$. That in turn means by our noninterference lemma for expressions that they evaluate to the same result, so $\text{bval } b \ s = \text{True}$ and $\text{bval } b \ t = \text{True}$. We already noted $\text{sec } b \vdash c_1$ by rule inversion, and with anti-monotonicity, we get the necessary $0 \vdash c_1$ to apply the induction hypothesis and conclude the case. In the other case, if $l < \text{sec } b$, i.e., a condition on a more confidential level than l , then we do not know that both *IF* commands will take the same branch. However, we do know that the whole command is a high-confidentiality computation. We can use the typing rule for *IF* to conclude $\text{sec } b \vdash \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2$ since we know both $\max(\text{sec } b) \ 0 \vdash c_1$ and $\max(\text{sec } b) \ 0 \vdash c_2$. This in turn means we now can apply confinement: everything below $\text{sec } b$ will be preserved — in particular the state of variables up to l . This is true for t to t' as well as s to s' . Together with the initial $s = t$ ($\leq l$) we can conclude $s' = t'$ ($\leq l$). This finishes the *IfTrue* case.

The *IfFalse* and *WhileFalse* cases are analogous. Either the conditions evaluate to the same value and we can apply the induction hypothesis, or the security level is high enough such that we can apply confinement.

Even the *WhileTrue* case is similar. Here, we have to work slightly harder to apply the induction hypotheses, once for the body and once for the rest of the loop, but the confinement side of the argument stays the same. \square

9.2.4 The Standard Typing System

The judgement $l \vdash c$ presented above is nicely intuitive and executable. However, the standard formulation in the literature is slightly different, replacing the maximum computation directly with the anti-monotonicity rule. We introduce the standard system now in Figure 9.8 and show equivalence with our previous formulation.

$$\begin{array}{c}
\frac{}{\vdash \text{SKIP} : l} \quad \frac{\text{sec } a \leq \text{sec } x}{\vdash x ::= a : \text{sec } x} \quad \frac{\vdash c_1 : l_1 \quad \vdash c_2 : l_2}{\vdash c_1;; c_2 : \min l_1 l_2} \\
\\
\frac{\text{sec } b \leq \min l_1 l_2 \quad \vdash c_1 : l_1 \quad \vdash c_2 : l_2}{\vdash \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2 : \min l_1 l_2} \quad \frac{\text{sec } b \leq l \quad \vdash c : l}{\vdash \text{WHILE } b \text{ DO } c : l}
\end{array}$$

Fig. 9.9. Definition of the bottom-up security type system

The equivalence proof goes by rule induction on the respective type system in each direction separately. Isabelle proves each subgoal of the induction automatically.

Lemma 9.15. $l \vdash c \implies l \vdash' c$

Lemma 9.16. $l \vdash' c \implies l \vdash c$

9.2.5 A Bottom-Up Typing System

The type systems presented above are top-down systems: the level l is passed from the context or the user and is checked at assignment commands. We can also give a bottom-up formulation where we compute the greatest l consistent with variable assignments and check this value at *IF* and *WHILE* commands. Instead of *max* computations, we now get *min* computations in Figure 9.9.

We can read the bottom-up statement $\vdash c : l$ as c has a write-effect of l , meaning that no variable below l is written to in c .

Again, we can prove equivalence. The first direction is straightforward and the proof is automatic.

Lemma 9.17. $\vdash c : l \implies l \vdash' c$

The second direction needs more care. The statement $l \vdash' c \implies \vdash c : l$ is not true, Isabelle's nitpick tool quickly finds a counter example:

$$0 \vdash' ''x'' ::= N\ 0, \text{ but } \neg \vdash ''x'' ::= N\ 0 : 0$$

The standard formulation admits anti-monotonicity, the computation of a minimal l does not. If we take this discrepancy into account, we get the following statement that is then again proved automatically by Isabelle.

Lemma 9.18. $l \vdash' c \implies \exists l' \geq l. \vdash c : l'$

9.2.6 What About Termination? thy

In the previous section we proved the following security theorem (Theorem 9.14):

$$\llbracket (c, s) \Rightarrow s'; (c, t) \Rightarrow t'; 0 \vdash c; s = t (\leq l) \rrbracket \Longrightarrow s' = t' (\leq l)$$

We read it as: *If our type system says yes, our data is safe: there will be no information flowing from high to low variables.*

Is this correct? The formal statement is certainly true, we proved it in Isabelle. But: it doesn't quite mean what the sentence above says. It means only precisely what the formula states: given two terminating executions started in states we can't tell apart, we won't be able to tell apart their final states.

What if we don't have two terminating executions? Consider, for example, the following typing statement.

$$0 \vdash \text{WHILE Less } (V \text{ ''}x'') \text{ (N 1) DO SKIP}$$

This is a true statement, the program is type correct at context level 0. Our noninterference theorem holds. Yet, the program still leaks information: the program will terminate if and only if the higher-security variable x (of level 1) is not 0. We can infer information about the contents of a secret variable by observing termination.

This is also called a **covert channel**, that is, an information flow channel that is not part of the security theorem or even the security model, and that the security theorem therefore does not make any claims about.

In our case, termination is observable in the model, but not in the theorem, because it already assumes two terminating executions from the start. An example of a more traditional covert channel is timing. Consider the standard `strcmp` function in C that compares two strings: it goes through the strings from left to right, and terminates with false as soon as two characters are not equal. The more characters are equal in the prefix of the strings, the longer it takes to execute this function. This time can be measured and the timing difference is significant enough to be statistically discernible even over network traffic. Such timing attacks can even be effective against widely deployed cryptographic algorithms, for instance as used in SSL [15].

Covert channels and the strength of security statements are the bane of security proofs. The literature is littered with the bodies of security theorems that have been broken, either because their statement was weak, or their proof was wrong, or because the model made unreasonably strong assumptions, i.e., admitted too many obvious covert channels.

Conducting security proofs in a theorem prover only helps against one of these: wrong proofs. Strong theorem statements and realistic model assumptions, or at least explicit model assumptions, are still our own responsibility.

So what can we do to fix our statement of security? For one, we could prove separately, and manually, that the specific programs we are interested in always terminate. Then the problem disappears. Or we could strengthen the type system and its security statement. The key idea is: *WHILE condi-*

$$\begin{array}{c}
\frac{}{l \vdash \text{SKIP}} \quad \frac{\text{sec } a \leq \text{sec } x \quad l \leq \text{sec } x}{l \vdash x ::= a} \quad \frac{l \vdash c_1 \quad l \vdash c_2}{l \vdash c_1;; c_2} \\
\\
\frac{\text{max } (\text{sec } b) \ l \vdash c_1 \quad \text{max } (\text{sec } b) \ l \vdash c_2}{l \vdash \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2} \quad \frac{\text{sec } b = 0 \quad 0 \vdash c}{0 \vdash \text{WHILE } b \text{ DO } c}
\end{array}$$

Fig. 9.10. Termination-sensitive security type system

tions must not depend on confidential data. If they don't, then termination cannot leak information.

In the following, we formalize and prove this idea.

Formalizing our idea means we replace the *WHILE*-rule with a new one that does not admit anything higher than level 0 in the condition:

$$\frac{\text{sec } b = 0 \quad 0 \vdash c}{0 \vdash \text{WHILE } b \text{ DO } c}$$

This is already it. Figure 9.10 shows the full set of rules, putting the new one into context.

We now need to change our noninterference statement such that it takes termination into account. The interesting case was where one execution terminated and the other didn't. If both executions terminate, our previous statement already applies; if both do not terminate then there is no information leakage, because there is nothing to observe.¹ So, since our statement is symmetric, we now assume *one* terminating execution, a well-typed program of level 0 as before, and two start states that agree up to level l , also as before. We then have to show that the other execution *also* terminates and that the final states still agree up to level l .

$$\llbracket (c, s) \Rightarrow s'; 0 \vdash c; s = t (\leq l) \rrbracket \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t' (\leq l)$$

We build up the proof of this new theorem in the same way as before. The first property is again anti-monotonicity, which still holds.

Lemma 9.19 (Anti-monotonicity).

$$\llbracket l \vdash c; l' \leq l \rrbracket \implies l' \vdash c$$

Proof. The proof is by induction on the typing derivation. Isabelle then solves each of the cases automatically. \square

Our confinement lemma is also still true.

Lemma 9.20 (Confinement).

$$\llbracket (c, s) \Rightarrow t; l \vdash c \rrbracket \implies s = t (< l)$$

¹ Note that if our programs had output, this case might leak information as well.

Proof. The proof is the same as before, first by induction on the big-step execution, then by using anti-monotonicity in the *IF* cases, and automation on the rest. \square

Before we can proceed to noninterference, we need one new fact about the new type system: any program that is type correct, but not at level 0 (only higher), must terminate. Intuitively that is easy to see: *WHILE* loops are the only cause of potential nontermination, and they can now only be typed at level 0. This means, if the program is type correct at some level, but not at level 0, it does not contain *WHILE* loops.

Lemma 9.21 (Termination).

$\llbracket l \vdash c; l \neq 0 \rrbracket \implies \exists t. (c, s) \Rightarrow t$

Proof. The formal proof of this lemma does not directly talk about the occurrence of while loops, but encodes the argument in a contradiction. We start the proof by induction on the typing derivation. The base cases all terminate trivially, and the step cases terminate because all their branches terminate in the induction hypothesis. In the *WHILE* case we have the contradiction: our assumption says that $l \neq 0$, but the induction rule instantiates l with 0, and we get $0 \neq 0$. \square

Equipped with these lemmas, we can finally proceed to our new statement of noninterference.

Theorem 9.22 (Noninterference).

$\llbracket (c, s) \Rightarrow s'; 0 \vdash c; s = t (\leq l) \rrbracket \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t' (\leq l)$

Proof. The proof is similar to that of the termination-insensitive case, it merely has to additionally show termination of the second command. For *SKIP*, assignment, and semicolon this is easy, the first two because they trivially terminate, the third because Isabelle can put the induction hypotheses together automatically.

The *IF* case is slightly more interesting. If the condition does not depend on secret variables, the induction hypothesis is strong enough for us to conclude the goal directly. However, if the condition does depend on secret variables, i.e., $\neg \text{sec } b \leq l$, we make use of confinement again, as we did in our previous proof. However, we first have to show that the second execution terminates, i.e., that a final state exists. This follows from our termination lemma and the fact that if the security level $\text{sec } b$ is greater than l , it cannot be 0. The rest goes through as before.

The *WHILE* case becomes easier than in our previous proof. Since we know from the typing statement that the boolean expression does not contain any high variables, we know that the loops started in s and t will continue

$$\begin{array}{c}
\frac{}{l \vdash' \text{SKIP}} \quad \frac{\text{sec } a \leq \text{sec } x \quad l \leq \text{sec } x}{l \vdash' x ::= a} \quad \frac{l \vdash' c_1 \quad l \vdash' c_2}{l \vdash' c_1;; c_2} \\
\\
\frac{\text{sec } b \leq l \quad l \vdash' c_1 \quad l \vdash' c_2}{l \vdash' \text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2} \quad \frac{\text{sec } b = 0 \quad 0 \vdash' c}{0 \vdash' \text{WHILE } b \text{ DO } c} \\
\\
\frac{l \vdash' c \quad l' \leq l}{l' \vdash' c}
\end{array}$$

Fig. 9.11. Termination-sensitive security type system — standard formulation

to make the same decision whether to terminate or not. That was the whole point of our type system change. In the *WhileFalse* case that is all that is needed; in the *WhileTrue* case, we can make use of this fact to access the induction hypothesis: from the fact that the loop is type correct at level 0, we know by rule inversion that $0 \vdash c$. We also know, by virtue of being in the *WhileTrue* case, that $\text{bval } b \text{ } s, (c, s) \Rightarrow s''$, and $(w, s'') \Rightarrow s'$. We now need to construct a terminating execution of the loop starting in t , ending in some state t' that agrees with s' below l . We start by noting $\text{bval } b \text{ } t$ using noninterference for boolean expressions. Per induction hypothesis we conclude that there is a t'' with $(c, t) \Rightarrow t''$ that agrees with s'' below l . Using the second induction hypothesis, we repeat the process for w , and conclude that there must be such a t' that agrees with s' below l . \square

The predicate $l \vdash c$ is phrased to be executable. The standard formulation, however, is again slightly different, replacing the maximum computation by the anti-monotonicity rule. Figure 9.11 introduces the standard system.

As before, we can show equivalence with our formulation.

Lemma 9.23 (Equivalence to standard formulation).

$$l \vdash c \iff l \vdash' c$$

Proof. As with the equivalence proofs of different security type system formulations in previous sections, this proof goes first by considering each direction of the if-and-only-if separately, and then by induction on the type system in the assumption of that implication. As before, Isabelle then proves each sub case of the respective induction automatically. \square

Exercises

Exercise 9.5. Reformulate the inductive predicate *sec_type* defined in Figure 9.7 as a recursive function $ok :: \text{level} \Rightarrow \text{com} \Rightarrow \text{bool}$ and prove the equivalence of the two formulations.

Try to reformulate the bottom-up system from [Figure 9.9](#) as a function that computes the security level of a command. What difficulty do you face?

Exercise 9.6. Define a bottom-up termination insensitive security type system $\vdash' c : l$ with subsumption rule. Prove equivalence with the bottom-up system in [Figure 9.9](#): $\vdash c : l \implies \vdash' c : l$ and $\vdash' c : l \implies \exists l' \geq l. \vdash c : l'$.

Exercise 9.7. Define a function $\text{erase} :: \text{level} \Rightarrow \text{com} \Rightarrow \text{com}$ that erases those parts of a command that contain variables above some security level. Function $\text{erase } l$ should replace all assignments to variables with security level $\geq l$ by *SKIP*. It should also erase certain *IF*s and *WHILE*s, depending on the security level of the boolean condition. Prove that c and $\text{erase } l c$ behave the same on the variables up to level l :

$$\llbracket (c, s) \Rightarrow s'; (\text{erase } l c, t) \Rightarrow t'; 0 \vdash c; s = t (< l) \rrbracket \implies s' = t' (< l)$$

It is recommended to start with the proof of the very similar looking noninterference [Theorem 9.14](#) and modify that.

In the theorem above we assume that both (c, s) and $(\text{erase } l c, t)$ terminate. How about the following two properties?

$$\begin{aligned} & \llbracket (c, s) \Rightarrow s'; 0 \vdash c; s = t (< l) \rrbracket \\ & \implies \exists t'. (\text{erase } l c, t) \Rightarrow t' \wedge s' = t' (< l) \end{aligned}$$

$$\begin{aligned} & \llbracket (\text{erase } l c, s) \Rightarrow s'; 0 \vdash c; s = t (< l) \rrbracket \\ & \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t' (< l) \end{aligned}$$

Give proofs or counterexamples.

9.3 Summary and Further Reading

In this chapter we have analysed two kinds of type systems: a standard type system that tracks types of values and prevents type errors at runtime, and a security type system that prevents information flow from higher-level to lower-level variables.

Sound, static type systems enjoy widespread application in popular programming languages such as Java, C#, Haskell, and ML, but also on low-level languages such as the Java Virtual Machine and its bytecode verifier [54]. Some of these languages require types to be declared explicitly, as in Java. In other languages, such as Haskell, these declarations can be left out, and types are inferred automatically.

The purpose of type systems is to prevent errors. In essence, a type derivation is a proof, which means type checking performs basic automatic proofs about programs.

The second type system we explored ensured absence of information flow. The field of language-based security is substantial [76]. As mentioned, the type system and the soundness statement in the sections above go back to Volpano, Irvine, and Smith [89], and the termination-sensitive analysis to Volpano and Smith [90]. While language-based security had been investigated before Volpano *et al.*, they were the first to give a security type system with a soundness proof that expresses the enforced security property in terms of the standard semantics of the language. As we have seen, such non-trivial properties are comfortably within the reach of machine-checked interactive proof. Our type system deviated a little from the standard presentation of Volpano *et al.*: we derive anti-monotonicity as a lemma, whereas Volpano, Irvine, and Smith have it as a typing rule. In exchange, they can avoid an explicit *max* calculation. We saw that our syntax-directed form of the rules is equivalent and allowed us to execute examples. Our additional bottom-up type system can be seen as a simplified description of type inference for security types. Volpano and Smith [91] gave an explicit algorithm for type inference and showed that most general types exist if the program is type correct. We also mentioned that our simple security levels based on natural numbers can be generalized to arbitrary lattices. This observation goes back to Denning [26].

A popular alternative to security type systems is dynamic tracking of information flows, or so-called **taint analysis** [81]. It has been long-time folklore in the field that static security analysis of programs must be more precise than dynamic analysis, because dynamic (runtime) analysis can only track one execution of the program at a time, whereas the soundness property of our static type system compares two executions. Many dynamic taint analysis implementations to date do not track implicit flows. Sabelfeld and Russo showed for termination-insensitive noninterference that this is *not* a theoretical restriction, and dynamic monitoring can in fact be more precise than the static type system [77]. However, since their monitor essentially turns implicit flow violations into non-termination, the question is still open for the more restrictive termination-sensitive case. For more sophisticated, so-called *flow-sensitive* type systems, the dynamic and static versions are incomparable: there are some programs where purely dynamic flow-sensitive analysis fails, but the static type system succeeds, and the other way around [75].

The name non-interference was coined by Goguen and Meseguer [35], but the property goes back further to Ellis Cohen who called its inverse *Strong Dependency* [18]. The concept of covert information flow channels already precedes this idea [51]. Non-interference can be applied beyond language-based security, for instance by directly proving the property about a specific system. This is interesting for systems that have inherent security require-

ments and are written in low-level languages such as C or in settings where the security policy cannot directly be attached to variables in a program. Operating systems are an example of this class, where the security policy is configurable at runtime. It is feasible to prove such theorems in Isabelle down to the C code level: the seL4 microkernel is an operating system kernel with such a non-interference theorem in Isabelle [60].

Program Analysis

Program analysis, also known as *static analysis*, describes a whole field of techniques for the static (i.e. compile-time) analysis of programs. Most compilers or programming environments perform more or less ambitious program analyses. The two most common objectives are the following:

Optimization The purpose is to improve the behaviour of the program, usually by reducing its running time or space requirements.

Error detection The purpose is to detect common programming errors that lead to runtime exceptions or other undesirable behaviour.

Program optimization is a special case of program transformation (for example for code refactoring) and consists of two phases: the analysis (to determine if certain required properties hold) and the transformation.

There are a number of different approaches to program analysis that employ different techniques to achieve similar aims. In the previous chapter we used type systems for error detection. In this chapter we employ what is known as *data-flow analysis*. We study three analyses (and associated transformations):

1. Definite initialization analysis determines if all variables have been initialized before they are read. This falls into the category of error detection analyses. There is no transformation.
2. Constant folding is an optimization that tries to replace variables by constants. For example, the second assignment in `x := 1; y := x` can be replaced by the (typically faster) `y := 1`.
3. Live variable analysis determines if a variable is “live” at some point, i.e. if its value can influence the subsequent execution. Assignments to variables that are not live can be eliminated: for example, the first assignment in the sequence `x := 0; x := 1` is redundant.

Throughout this chapter we continue the naive approach to program analysis that ignores boolean conditions. That is, we treat them as nondeterministic: we assume that both values are possible every time the conditions are tested. More precisely, our analyses are correct w.r.t. a (big or small-step) semantics where we have simply dropped the preconditions involving boolean expressions from the rules, resulting in a nondeterministic language.

Limitations

Program analyses, no matter what techniques they employ, are always limited. This is a consequence of Rice's Theorem from computability theory. It roughly tells us that *Nontrivial semantic properties of programs (e.g. termination) are undecidable*. That is, no nontrivial semantic property P has a magic analyser that

- terminates on every input program,
- only says Yes if the input program has property P (correctness), and
- only says No if the input program does not have property P (completeness).

For concreteness, let us consider definite initialization analysis of the following program:

```
FOR ALL positive integers x, y, z, n DO
  IF  $n > 2 \wedge x^n + y^n = z^n$  THEN  $u := u$  ELSE SKIP
```

For convenience we have extended our programming language with a FOR ALL loop and an exponentiation operation: both could be programmed in pure IMP, although it would be painful. The program searches for a counterexample to Fermat's conjecture that no three positive integers x , y , and z can satisfy the equation $x^n + y^n = z^n$ for any integer $n > 2$. It reads the uninitialized variable u (thus violating the definite initialization property) iff such a counterexample exists. It would be asking a bit much from a humble program analyser to determine the truth of a statement that was in the *Guinness Book of World Records* for "most difficult mathematical problems" prior to its 1995 proof by Wiles.

As a consequence, we cannot expect program analysers to terminate, be correct and be complete. Since we do not want to sacrifice termination and correctness, we sacrifice completeness: we allow analysers to say No although the program has the desired semantic property but the analyser was unable to determine that.

10.1 Definite Initialization Analysis

The first program analysis we investigate is called **definite initialization**. The Java Language Specification has the following to say on definite initialization [38, chapter 16, p. 527]:

Each local variable [...] must have a definitely assigned value when any access of its value occurs. [...] A compiler must carry out a specific conservative flow analysis to make sure that, for every access of a local variable [...] *f*, *f* is definitely assigned before the access; otherwise a compile-time error must occur.

Java was the first mainstream language to force programmers to initialize their variables.

In most programming languages, objects allocated on the heap are automatically initialized to zero or a suitable default value, but local variables are not. Uninitialized variables are a common cause of program defects that are very hard to find. For instance, a C program, that uses an uninitialized local integer variable will not necessarily crash on the first access to that integer. Instead, it will read the value that is stored there by accident. On the developer's machine and operating system that value may happen to be 0 and the defect will go unnoticed. On the user's machine, that same memory may contain different values left over from a previous run or from a different application. Worse still, this random value might not directly lead to a crash either, but only cause misbehaviour at a much later point of execution, leading to bug reports that are next to impossible to reproduce for the developer.

Removing the potential for such errors automatically is the purpose of the definite initialization analysis.

Consider the following example with an already initialized *x*. Recall from Section 7.1 that IF and WHILE bind more strongly than semicolon.

```
IF x < 1 THEN y := x ELSE y := x + 1; y := y + 1
IF x < x THEN y := y + 1 ELSE y := x; y := y + 1
```

The first line is clearly fine: in both branches of the IF, *y* gets initialized before it is used in the statement after. The second line is also fine: even though the *True* branch uses *y* where it is potentially uninitialized, we know that the *True* branch can never be taken. However, we only know that because we know that *x < x* will always be *False*.

What about the following example? Assume *x* and *y* are initialized.

```
WHILE x < y DO z := x; z := z + 1
```

Here it depends: if *x < y*, the program is fine (it will never terminate, but at least it does so without using uninitialized variables), but if *x < y* is not the

case, the program is unsafe. So, if our goal is to reject all *potentially* unsafe programs, we have to reject this one.

As mentioned in the introduction, we do not analyse boolean expressions statically to make predictions about program execution. Instead we take both potential outcomes into account. This means, the analysis we are about to develop will only accept the first program, but reject the other two.

Java is more discerning in this case, and will perform the optimization of **constant folding**, which we discuss in [Section 10.2](#), before definite initialization analysis. If during that pass it turns out an expression is always *True* or always *False*, this can be taken into account. This is a nice example of positive interaction between different kinds of optimization and program analysis, where one enhances the precision and predictive power of the other.

As discussed, we cannot hope for completeness of any program analysis, so there will be cases of safe programs that are rejected. For this specific analysis, this is usually the case when the programmer is smarter than the boolean constant folding the compiler performs. As with any restriction in a programming language, some programmers will complain about the shackles of definite initialization analysis, and Java developer forums certainly contain such complaints. Completely eliminating this particularly hard-to-find class of Heisenbugs well justifies the occasional program refactoring, though.

In the following sections, we construct our definite initialization analysis, define a semantics where initialization failure is observable, and then proceed to prove the analysis correct by showing that these failures will not occur.

10.1.1 Definite Initialization [thy](#)

The Java Language Specification quotes a number of rules that definite initialization analysis should implement to achieve the desired result. They have the following form (adjusted for IMP):

Variable x is definitely initialized after *SKIP*
iff x is definitely initialized before *SKIP*.

Similar statements exist for each language construct. Our task is simply to formalize them. Each of these rules talks about variables, or more precisely sets of variables. For instance, to check an assignment statement, we will want to start with a set of variables that is already initialized, we will check that set against the set of variables that is used in the assignment expression, and we will add the assigned variable to the initialized set after the assignment has completed.

So, the first formal tool we need is the set of variables mentioned in an expression. The Isabelle theory *Vars* provides an overloaded function *vars* for this:

$$\begin{array}{c}
\frac{}{D \ A \ SKIP \ A} \quad \frac{vars \ a \subseteq A}{D \ A \ (x ::= a) \ (insert \ x \ A)} \\
\\
\frac{\frac{D \ A_1 \ c_1 \ A_2}{D \ A_1 \ (c_1;; c_2) \ A_3} \quad D \ A_2 \ c_2 \ A_3}{D \ A_1 \ (c_1;; c_2) \ A_3} \\
\\
\frac{vars \ b \subseteq A \quad D \ A \ c_1 \ A_1 \quad D \ A \ c_2 \ A_2}{D \ A \ (IF \ b \ THEN \ c_1 \ ELSE \ c_2) \ (A_1 \cap A_2)} \\
\\
\frac{vars \ b \subseteq A \quad D \ A \ c \ A'}{D \ A \ (WHILE \ b \ DO \ c) \ A}
\end{array}$$

Fig. 10.1. Definite initialization $D :: vname \ set \Rightarrow com \Rightarrow vname \ set \Rightarrow bool$

```

fun vars :: aexp  $\Rightarrow$  vname set where
vars (N n)           = {}
vars (V x)           = {x}
vars (Plus a1 a2) = vars a1  $\cup$  vars a2

fun vars :: bexp  $\Rightarrow$  vname set where
vars (Bc v)          = {}
vars (Not b)          = vars b
vars (And b1 b2)    = vars b1  $\cup$  vars b2
vars (Less a1 a2)  = vars a1  $\cup$  vars a2

```

With this we can define our main definite initialization analysis. The purpose is to check whether each variable in the program is assigned to before it is used. This means we ultimately want a predicate of type $com \Rightarrow bool$, but we have already seen in the examples that we need a slightly more general form for the computation itself. In particular, we carry around a set of variables that we know are definitely initialized at the beginning of a command. The analysis then has to do two things: check whether the command only uses these variables, and produce a new set of variables that we know are initialized afterwards. This leaves us with the following type signature:

$D :: vname \ set \Rightarrow com \Rightarrow vname \ set \Rightarrow bool$

We want the notation $D \ A \ c \ A'$ to mean:

If all variables in A are initialized before c is executed, then no uninitialized variable is accessed during execution, and all variables in A' are initialized afterwards.

Figure 10.1 shows how we can inductively define this analysis with one rule per syntactic construct. We walk through them step by step:

- The *SKIP* rule is obvious, and translates exactly the text rule we have mentioned above.

- Similarly, the assignment rule follows our example above: the predicate $D \ A \ (x ::= a) \ A'$ is *True* if the variables of the expression a are contained in the initial set A , and if A' is precisely the initial A plus the variable x we just assigned to.
- Sequential composition has the by now familiar form: we simply pass through the result A_2 of c_1 to c_2 , and the composition is definitely initialized if both commands are definitely initialized.
- In the *IF* case, we check that the variables of the boolean expression are all initialized, and we check that each of the branches is definitely initialized. We pass back the intersection of the results produced by c_1 and c_2 , because we do not know which branch will be taken at runtime. If we were to analyse boolean expression more precisely, we could introduce further case distinctions into this rule.
- Finally, the *WHILE* case. It also checks that the variables in the boolean expression are all in the initialized set A , and it also checks that the command c is definitely initialized starting in the same set A , but it ignores the result A' of c . Again, this must be so, because we have to be conservative: it is possible that the loop will never be executed at runtime, because b may be already *False* before the first iteration. In this case no additional variables will be initialized, no matter what c does. It may be possible for specific loop structures, such as for-loops, to statically determine that their body will be executed at least once, but no mainstream language currently does that.

We can now decide whether a command is definitely initialized, namely exactly when we can start with the empty set of initialized variables and find a resulting set such that our inductive predicate D is *True*:

$$\mathcal{D} \ c = (\exists A'. \ D \ \{\} \ c \ A')$$

Defining a program analysis such as definite initialization by an inductive predicate makes the connection to type systems clear: in a sense, all program analyses can be phrased as sufficiently complex type systems. Since our rules are syntax-directed, they also directly suggest a recursive execution strategy. In fact, for this analysis it is straightforward to turn the inductive predicate into two recursive functions in Isabelle that compute our set A' if it exists, and check whether all expressions mention only initialized variables. We leave this recursive definition and proof of equivalence as an exercise to the reader and turn our attention to proving correctness of the analysis instead.

10.1.2 Initialization Sensitive Expression Evaluation thy

As in type systems, to phrase what correctness of the definite initialization analysis means, we first have to identify what could possibly go wrong.

Here, this is easy: we should observe an error when the program uses a variable that has not been initialized. That is, we need a new, finer-grained semantics that keeps track of which variables have been initialized and leads to an error if the program accesses any other variable.

To that end, we enrich our set of values with an additional element that we will read as *uninitialized*. As mentioned in [Section 2.3.1](#) in the Isabelle part in the beginning, Isabelle provides the *option* data type, which is useful for precisely such situations:

```
datatype 'a option = None | Some 'a
```

We simply redefine our program state as

```
type_synonym state = vname  $\Rightarrow$  val option
```

and take *None* as the uninitialized value. The *option* data type comes with additional useful notation: $s(x \mapsto y)$ means $s(x := \text{Some } y)$, and $\text{dom } s = \{a. s\ a \neq \text{None}\}$.

Now that we can distinguish initialized from uninitialized values, we can check the evaluation of expressions. We have had a similar example of potentially failing expression evaluation in type systems in [Section 9.1](#). There we opted for an inductive predicate, reasoning that in the functional style where we would return *None* for failure, we would have to consider all failure cases explicitly. This argument also holds here. Nevertheless, for the sake of variety, we will this time show the functional variant with *option*. It is less elegant, but not so horrible as to become unusable. It has the advantage of being functional, and therefore easier to apply automatically in proofs.

```
fun aval :: aexp  $\Rightarrow$  state  $\Rightarrow$  val option where
  aval (N i) s      = Some i
  aval (V x) s      = s x
  aval (Plus a1 a2) s = (case (aval a1 s, aval a2 s) of
    (Some i1, Some i2)  $\Rightarrow$  Some(i1+i2)
  | _  $\Rightarrow$  None)

fun bval :: bexp  $\Rightarrow$  state  $\Rightarrow$  bool option where
  bval (Bc v) s      = Some v
  bval (Not b) s      = (case bval b s of
    None  $\Rightarrow$  None | Some bv  $\Rightarrow$  Some( $\neg$  bv))
  bval (And b1 b2) s = (case (bval b1 s, bval b2 s) of
    (Some bv1, Some bv2)  $\Rightarrow$  Some(bv1  $\wedge$  bv2)
  | _  $\Rightarrow$  None)
  bval (Less a1 a2) s = (case (aval a1 s, aval a2 s) of
    (Some i1, Some i2)  $\Rightarrow$  Some(i1 < i2)
  | _  $\Rightarrow$  None)
```

We can reward ourselves for all these case distinctions with two concise lemmas that confirm that expressions indeed evaluate without failure if they only mention initialized variables.

Lemma 10.1 (Initialized arithmetic expressions).

$\text{vars } a \subseteq \text{dom } s \implies \exists i. \text{aval } a \ s = \text{Some } i$

Lemma 10.2 (Initialized boolean expressions).

$\text{vars } b \subseteq \text{dom } s \implies \exists bv. \text{bval } b \ s = \text{Some } bv$

Both lemmas are proved automatically after structural induction on the expression.

10.1.3 Initialization-Sensitive Small-Step Semantics thy

From here, the development towards the correctness proof is standard: we define a small-step semantics, and we prove progress and preservation as we would for a type system.

$$\begin{array}{c}
 \frac{\text{aval } a \ s = \text{Some } i}{(x ::= a, s) \rightarrow (\text{SKIP}, s(x \mapsto i))} \\
 \\
 \frac{}{(SKIP;; c, s) \rightarrow (c, s)} \quad \frac{(c_1, s) \rightarrow (c'_1, s')}{(c_1;; c_2, s) \rightarrow (c'_1;; c_2, s')} \\
 \\
 \frac{\text{bval } b \ s = \text{Some } \text{True}}{(IF \ b \ THEN \ c_1 \ ELSE \ c_2, s) \rightarrow (c_1, s)} \\
 \\
 \frac{\text{bval } b \ s = \text{Some } \text{False}}{(IF \ b \ THEN \ c_1 \ ELSE \ c_2, s) \rightarrow (c_2, s)} \\
 \\
 \frac{}{(WHILE \ b \ DO \ c, s) \rightarrow (IF \ b \ THEN \ c;; \ WHILE \ b \ DO \ c \ ELSE \ SKIP, s)}
 \end{array}$$

Fig. 10.2. Small-step semantics, initialization-sensitive

In fact, the development is so standard that we only show the small-step semantics in [Figure 10.2](#) and give one hint for the soundness proof. It needs the following lemma.

Lemma 10.3 (D is increasing). $D \ A \ c \ A' \implies A \subseteq A'$

Proof. This lemma holds independently of the small-step semantics. The proof is automatic after structural induction on c . \square

The soundness statement then is as in the type system in [Section 9.1](#).

Theorem 10.4 (*D is sound*).

If $(c, s) \rightarrow^ (c', s')$ and $D \text{ (dom } s) \text{ } c \text{ } A'$ then $(\exists cs''. (c', s') \rightarrow cs'') \vee c' = \text{SKIP}$.*

The proof goes by showing progress and preservation separately and making use of [Lemma 10.3](#). We leave its details as an exercise and present an alternative way of proving soundness of the definite initialization analysis in the next section instead.

10.1.4 Initialization-Sensitive Big-Step Semantics [thy](#)

In the previous section we learned that a formalization in the small-step style and a proof with progress and preservation as we know them from type systems are sufficient to prove correctness of definite initialization. In this section, we investigate how to adjust a big-step semantics such that it can be used for the same purpose of proving the definite initialization analysis correct. We will see that this is equally possible and that big-step semantics can be used for such proofs. This may be attractive for similar kinds of correctness statements, because big-step semantics are often easier to write down. However, we will also see the price we have to pay: a larger number of big-step rules and therefore a larger number of cases in inductive proofs about them.

The plan for adjusting the big-step semantics is simple: we need to be able to observe error states, so we will make errors explicit and propagate them to the result. Formally, we want something of the form

$$com \times state \Rightarrow state \text{ option}$$

where *None* would indicate that an error occurred during execution, in our case that the program accessed an uninitialized variable.

There is a small complication with the type above. Consider for instance this attempt to write the semicolon rule:

$$\frac{(c_1, s_1) \Rightarrow \text{Some } s_2 \quad (c_2, s_2) \Rightarrow s}{(c_1;; c_2, s_1) \Rightarrow s} \quad \frac{(c_1, s_1) \Rightarrow \text{None}}{(c_1;; c_2, s_1) \Rightarrow \text{None}}$$

There is no problem with the soundness of these rules. The left rule is the case where no error occurs, the right rule terminates the execution when an error does occur. The problem is that we will need at least these two cases for any construct that has more than one command. It would be nicer to specify once and for all how error propagates.

We can make the rules more compositional by ensuring that the result type is the same as the start type for an execution, i.e., that we can plug a

result state directly into the start of the next execution without any additional operation or case distinction for unwrapping the *option* type. We achieve this by making the start type *state option* as well.

$$com \times state\ option \Rightarrow state\ option$$

We can now write one rule that defines how error (*None*) propagates:

$$(c, None) \Rightarrow None$$

Consequently, in the rest of the semantics in Figure 10.3 we only have to locally consider the case where we directly produce an error, and the case of normal execution. An example of the latter is the assignment rule, where we update the state as usual if the arithmetic expression evaluates normally:

$$\frac{aval\ a\ s = Some\ i}{(x ::= a, Some\ s) \Rightarrow Some\ (s(x \mapsto i))}$$

An example of the former is the assignment rule, where expression evaluation leads to failure:

$$\frac{aval\ a\ s = None}{(x ::= a, Some\ s) \Rightarrow None}$$

The remaining rules in Figure 10.3 follow the same pattern. They only have to worry about producing errors, not about propagating them.

If we are satisfied that this semantics encodes failure for accessing uninitialized variables, we can proceed to proving correctness of our program analysis *D*.

The statement we want in the end is, paraphrasing Milner, *well-initialized programs cannot go wrong*.

$$\llbracket D\ (dom\ s)\ c\ A'; (c, Some\ s) \Rightarrow s' \rrbracket \Longrightarrow s' \neq None$$

The plan is to use rule induction on the big-step semantics to prove this property directly, without the detour over progress and preservation. Looking at the rules for $D\ A\ c\ A'$, it is clear that we will not be successful with a constant pattern of $dom\ s$ for A , because the rules produce different patterns. This means, both A and A' need to be variables in the statement to produce suitably general induction hypotheses. Replacing $dom\ s$ with a plain variable A in turn means we have to find a suitable side condition such that our statement remains true, and we have to show that this side condition is preserved. A suitable such condition is $A \subseteq dom\ s$, i.e., it is OK if our program analysis succeeds with fewer variables than are currently initialized in the state. After this process of generalizing the statement for induction, we arrive at the following lemma.

$$\begin{array}{c}
\frac{}{(c, \text{None}) \Rightarrow \text{None}} \quad \frac{}{(\text{SKIP}, s) \Rightarrow s} \\
\\
\frac{\text{aval } a \ s = \text{Some } i}{(x ::= a, \text{Some } s) \Rightarrow \text{Some } (s(x \mapsto i))} \quad \frac{\text{aval } a \ s = \text{None}}{(x ::= a, \text{Some } s) \Rightarrow \text{None}} \\
\\
\frac{(c_1, s_1) \Rightarrow s_2 \quad (c_2, s_2) \Rightarrow s_3}{(c_1;; c_2, s_1) \Rightarrow s_3} \\
\\
\frac{\text{bval } b \ s = \text{Some True} \quad (c_1, \text{Some } s) \Rightarrow s'}{(IF \ b \ THEN \ c_1 \ ELSE \ c_2, \text{Some } s) \Rightarrow s'} \\
\\
\frac{\text{bval } b \ s = \text{Some False} \quad (c_2, \text{Some } s) \Rightarrow s'}{(IF \ b \ THEN \ c_1 \ ELSE \ c_2, \text{Some } s) \Rightarrow s'} \\
\\
\frac{\text{bval } b \ s = \text{None}}{(IF \ b \ THEN \ c_1 \ ELSE \ c_2, \text{Some } s) \Rightarrow \text{None}} \\
\\
\frac{\text{bval } b \ s = \text{Some False}}{(WHILE \ b \ DO \ c, \text{Some } s) \Rightarrow \text{Some } s} \\
\\
\frac{\text{bval } b \ s = \text{Some True} \quad (c, \text{Some } s) \Rightarrow s' \quad (WHILE \ b \ DO \ c, s') \Rightarrow s''}{(WHILE \ b \ DO \ c, \text{Some } s) \Rightarrow s''} \\
\\
\frac{\text{bval } b \ s = \text{None}}{(WHILE \ b \ DO \ c, \text{Some } s) \Rightarrow \text{None}}
\end{array}$$

Fig. 10.3. Big-step semantics with error propagation

Lemma 10.5 (Soundness of D).

$\llbracket (c, \text{Some } s) \Rightarrow s'; D \ A \ c \ A'; A \subseteq \text{dom } s \rrbracket$
 $\implies \exists t. s' = \text{Some } t \wedge A' \subseteq \text{dom } t$

Proof. The proof is by rule induction on $(c, \text{Some } s) \Rightarrow s'$; Isabelle solves all sub-cases but *WHILE* automatically. In the *WHILE* case, we apply the induction hypothesis to the body c manually and can then let the automation figure out the rest. Applying the induction hypothesis to c is interesting, because we need to make use of D 's *increasing* property we proved in Lemma 10.3. Recall that the D rule for *WHILE* requires $D \ A \ c \ A'$ for the body c . Per induction hypothesis, we get that the result state t after execution of c has the property $A' \subseteq \text{dom } t$. To apply the induction hypothesis for the rest of the *WHILE* loop, however, we need $A \subseteq \text{dom } t$. From Lemma 10.3 we know that $A \subseteq A'$ and can therefore proceed. \square

After this proof, we can now better compare the small-step and big-step approaches to showing soundness of D : While the small-step semantics is

more concise, the soundness proof is longer, and while the big-step semantics has a larger number of rules, its soundness proof is more direct and shorter. As always, the trade-off depends on the particular application. With machine-checked proofs, it is in general better to err on the side of nicer and easier-to-understand definitions than on the side of shorter proofs.

Exercises

Exercise 10.1. Define the definite initialization analysis as two recursive functions $ivars :: com \Rightarrow vname\ set$ and $ok :: vname\ set \Rightarrow com \Rightarrow bool$ such that $ivars$ computes the set of definitely initialized variables and ok checks that only initialized variable are accessed. Prove $D\ A\ c\ A' \implies A' = A \cup ivars\ c$ and $ok\ A\ c \implies D\ A\ c\ (A \cup ivars\ c)$.

10.2 Constant Folding and Propagation thy

The previous section presented an analysis that prohibits a common programming error, uninitialized variables. This section presents an analysis that enables program optimizations, namely constant folding and propagation.

Constant folding and constant propagation are two very common compiler optimizations. Constant folding means computing the value of constant expressions at compile time and substituting their value for the computation. Constant propagation means determining if a variable has constant value, and propagating that constant value to the use-occurrences of that variable, for instance to perform further constant folding:

```
x := 42 - 5;
y := x * 2
```

In the first line, the compiler would fold the expression $42 - 5$ into its value 37, and in the second line, it would propagate this value into the expression $x * 2$ to replace it with 74 and arrive at

```
x := 37;
y := 74
```

Further liveness analysis could then for instance conclude that x is not live in the program and can therefore be eliminated, which frees up one more register for other local variables and could thereby improve time as well as space performance of the program.

Constant folding can be especially effective when used on boolean expressions, because it allows the compiler to recognize and eliminate further dead code. A common pattern is something like

```
IF debug THEN debug_command ELSE SKIP
```

where `debug` is a global constant that if set to `False` could eliminate debugging code from the program. Other common uses are the explicit construction of constants from their constituents for documentation and clarity.

Despite its common use for debug statements as above, we stay with our general policy in this chapter and will *not* analyse boolean expressions for constant folding. Instead, we leave it as a medium-sized exercise project for the reader to apply the techniques covered in this section.

The semantics of full-scale programming languages can make constant folding tricky (which is why one should prove correctness, of course). For instance, folding of floating point operations may depend on the rounding mode of the machine, which may only be known at runtime. Some languages demand that errors such as arithmetic overflow or division by zero be preserved and raised at runtime; others may allow the compiler to refuse to compile such programs; yet others allow the compiler to silently produce any code it likes in those cases.

A widely known tale of caution for constant folding is that of the Intel Pentium FDIV bug in 1994 which led to a processor recall costing Intel roughly half a billion US dollars. In processors exhibiting the fault, the FDIV instruction would perform an incorrect floating point division for some specific operands (10^{37} combinations would lead to wrong results). Constant folding was not responsible for this bug, but it gets its mention in the test for the presence of the FDIV problem. To make it possible for consumers to figure out if they had a processor exhibiting the defect, a number of small programs were written that performed the division with specific operands known to trigger the bug. Testing for the incorrect result, the program would then print a message whether the bug was present or not.

If a developer compiled this test program naively, the compiler would perform this computation statically and optimize the whole program to a binary that just returned a constant yes or no. This way, every single computer in a whole company could be marked as defective, even though only the developer's CPU actually had the bug.

In all of this, the compiler was operating entirely correctly, and would have acted the same way if it was proved correct. We see that our proofs critically rely on the extra-logical assumption that the hardware behaves as specified. Usually, this assumption underlies everything programmers do. However, trying to distinguish correct from incorrect hardware under the assumption that the hardware is correct is not a good move.

In the following, we are not attempting to detect defective hardware, and can focus on how constant propagation works, how it can be formalized, and how it can be proved correct.

10.2.1 Folding

As usual, we begin with arithmetic expressions. The first optimization is pure constant folding: the aim is to write a function that takes an arithmetic expression and statically evaluates all constant subexpressions within it. In the first part of this book in [Section 3.1.3](#), after our first contact with arithmetic expressions, we already wrote such a function.

At that time we could not simplify variables, i.e., we defined

$$asimp_const (V\ x) = V\ x$$

In this section, however, we are going to mix constant folding with constant propagation, and, if we know the constant value of a variable by propagation, we should use it. To do this, we keep a table or environment that tells us which variables have constant value, and what that value is. This is the same technique we already used in type systems and other static analyses.

$$\text{type_synonym } tab = vname \Rightarrow val\ option$$

We can now formally define our new function *afold* that performs constant folding on arithmetic expressions under the assumption that we already know constant values for some of the variables.

```
fun afold :: aexp  $\Rightarrow$  tab  $\Rightarrow$  aexp where
  afold (N n) _ = N n
  afold (V x) t = (case t x of None  $\Rightarrow$  V x | Some x  $\Rightarrow$  N x)
  afold (Plus e1 e2) t = (case (afold e1 t, afold e2 t) of
    (N n1, N n2)  $\Rightarrow$  N (n1+n2)
    | (e'1, e'2)  $\Rightarrow$  Plus e'1 e'2)
```

For example, the value of *afold* (Plus (V "x'") (N 3)) *t* now depends on the value of *t* at "x'". If *t* "x'" = Some 5, for instance, *afold* will return N 8. If nothing is known about "x'", i.e., *t* "x'" = None, then we get back the original Plus (V "x'") (N 3).

The correctness criterion for this simple optimization is that the result of execution with optimization is the same as without:

$$aval (afold\ a\ t)\ s = aval\ a\ s$$

As with type system soundness and its corresponding type environments, however, we need the additional assumption that the static table *t* conforms with, or in this case *approximates*, the runtime state *s*. The idea is again that the static value needs to agree with the dynamic value if the former exists:

$$\text{definition } approx\ t\ s = (\forall x\ k. t\ x = Some\ k \longrightarrow s\ x = k)$$

With this assumption the statement is provable.

Lemma 10.6 (Correctness of *afold*).

$$\text{approx } t \ s \implies \text{aval } (\text{afold } a \ t) \ s = \text{aval } a \ s$$

Proof. Automatic, after induction on the expression. □

The definitions and the proof reflect that the constant folding part of the folding and propagation optimization is the easy part. For more complex languages, one would have to consider further operators and cases, but nothing fundamental changes in the structure of proof or definition.

As mentioned, in more complex languages, care must be taken in the definition of constant folding to preserve the failure semantics of that language. For some languages it is permissible for the compiler to return a valid result for an invalid program, for others the program must fail in the right way.

10.2.2 Propagation

At this point, we have a function that will fold constants in arithmetic expressions for us. To lift this to commands for full constant propagation, we apply the same technique, defining a new function $\text{fold} :: \text{com} \Rightarrow \text{tab} \Rightarrow \text{com}$. The idea is to take a command and a constant table and produce a new command. The first interesting case in any of these analyses usually is assignment. This is easy here, because we can use *afold*:

$$\text{fold } (x ::= a) \ t = x ::= \text{afold } a \ t$$

What about sequential composition? Given $c_1;; c_2$ and t , we will still need to produce a new sequential composition, and we will obviously want to use *fold* recursively. The question is, which t do we pass to the call $\text{fold } c_2$ for the second command? We need to pick up any new values that might have been assigned in the execution of c_1 . This is basically the analysis part of the optimization, whereas *fold* is the code adjustment.

We define a new function for this job and call it $\text{defs} :: \text{com} \Rightarrow \text{tab} \Rightarrow \text{tab}$ for *definitions*. Given a command and a constant table, it should give us a new constant table that describes the variables with known constant values after the execution of this command.

Figure 10.4 shows the main definition. Auxiliary function *lvars* computes the set of variables on the left-hand side of assignments (see Appendix A). Function *merge* computes the intersection of two tables:

$$\text{merge } t_1 \ t_2 = (\lambda m. \text{if } t_1 \ m = t_2 \ m \text{ then } t_1 \ m \text{ else None})$$

Let's walk through the equations of *defs* one by one.

- For *SKIP* there is nothing to do, as usual.

```

fun defs :: com  $\Rightarrow$  tab  $\Rightarrow$  tab where
defs SKIP t                                = t
defs (x ::= a) t                            = (case afold a t of
                                                N k  $\Rightarrow$  t(x  $\mapsto$  k)
                                                | _  $\Rightarrow$  t(x ::= None))
defs (c1;; c2) t                            = (defs c2  $\circ$  defs c1) t
defs (IF b THEN c1 ELSE c2) t = merge (defs c1 t) (defs c2 t)
defs (WHILE b DO c) t                      = t|(- lvars c)

```

Fig. 10.4. Definition of *defs*

- In the assignment case, we attempt to perform constant folding on the expression. If this is successful, i.e., if we get a constant, we note in the result that the variable has a known value. Otherwise, we note that the variable does not have a known value, even if it might have had one before.
- In the semicolon case, we compose the effects of *c*₁ and *c*₂, in that order.
- In the *IF* case, we can only determine the values of variables with certainty if they have been assigned the same value after both branches; hence our use of the table intersection *merge* defined above.
- The *WHILE* case, as almost always, is interesting. Since we don't know statically whether we will ever execute the loop body, we cannot add any new variable assignments to the table. The situation is even worse, though. We need to *remove* all values from the table that are for variables mentioned on the left-hand side of assignment statements in the loop body, because they may contradict what the initial table has stored. A plain merge as in the *IF* case would not be strong enough, because it would only cover the first iteration. Depending on the behaviour of the body, a different value might be assigned to a variable in the body in a later iteration. Unless we employ a full static analysis on the loop body as in [Chapter 13](#), which constant propagation usually does not, we need to be conservative. The formalization achieves this by first computing the names of all variables on the left-hand side of assignment statements in *c* by means of *lvars*, and by then restricting the table to the complement of that set ($- \textit{lvars } c$). The notation $t|_S$ is defined as follows.

$$t|_S = (\lambda x. \text{if } x \in S \text{ then } t \ x \text{ else } \textit{None})$$

With all these auxiliary definitions in place, our definition of *fold* is now as expected. In the *WHILE* case, we fold the body recursively, but again restrict the set of variables to those not written to in the body.


```

fun fold :: com ⇒ tab ⇒ com where
fold SKIP _ = SKIP
fold (x ::= a) t = x ::= afold a t
fold (c1;; c2) t = fold c1 t;; fold c2 (defs c1 t)
fold (IF b THEN c1 ELSE c2) t = IF b THEN fold c1 t ELSE fold c2 t
fold (WHILE b DO c) t = WHILE b DO fold c (t|(- lvars c))

```

Let's test these definitions with some sample executions. Our first test is the first line in the example program at the beginning of this section. The program was:

```

x := 42 - 5;
y := x * 2

```

Formally, the first line can be encoded as $'x' ::= \text{Plus } (N\ 42) (N\ (-\ 5))$. Running *fold* on this with the empty table λx . *None* gives us $'x' ::= N\ 37$. This is correct. Encoding the second line as a *Plus* in IMP, and running *fold* on it in isolation with the empty table should give us no simplification at all, and this is what we get: $'y' ::= \text{Plus } (V\ 'x') (V\ 'x')$. However, if we provide a table that sets x to some value, say 1, we should get a simplified result: $'y' ::= N\ 2$. Finally, testing propagation over semicolon, we run the whole statement with the empty table and get $'x' ::= N\ 37$; $'y' ::= N\ 74$. This is also as expected.

As always in this book, programming and testing are not enough. We want proof that constant folding and propagation are correct. Because we perform a program transformation, our notion of correctness is semantic equivalence. Eventually, we are aiming for the following statement:

$$\text{fold } c (\lambda x. \text{None}) \sim c$$

Since all our definitions are recursive in the commands, the proof plan is to proceed by induction on the command. Unsurprisingly, we need to generalize the statement from empty tables to arbitrary tables t . Further, we need to add a side condition for this t , namely the same as in our lemma about expressions: t needs to approximate the state s the command runs in. This leads us to the following interlude on equivalence of commands *up to a condition*.

10.2.3 Conditional Equivalence

This section describes a generalization of the equivalence of commands, where commands do not need to agree in their executions for *all* states, but only for those states that satisfy a precondition. In [Section 7.2.4](#), we defined

$$(c \sim c') = (\forall s\ t. (c, s) \Rightarrow t = (c', s) \Rightarrow t)$$

Extending this concept to take a condition P into account is straightforward. We read $P \models c \sim c'$ as c is equivalent to c' under the assumption P .

definition

$$(P \models c \sim c') = (\forall s s'. P s \longrightarrow (c, s) \Rightarrow s' = (c', s) \Rightarrow s')$$

We can do the same for boolean expressions:

definition

$$(P \models b \sim b') = (\forall s. P s \longrightarrow \text{bval } b s = \text{bval } b' s)$$

Clearly, if we instantiate P to the predicate that returns *True* for all states, we get our old concept of unconditional semantic equivalence back.

Lemma 10.7. $((\lambda_. \text{True}) \models c \sim c') = (c \sim c')$

Proof. By unfolding definitions. □

For any fixed predicate, our new definition is an equivalence relation, i.e., it is reflexive, symmetric, and transitive.

Lemma 10.8 (Equivalence Relation).

$$\begin{aligned} P \models c \sim c \\ (P \models c \sim c') &= (P \models c' \sim c) \\ \llbracket P \models c \sim c'; P \models c' \sim c'' \rrbracket &\Longrightarrow P \models c \sim c'' \end{aligned}$$

Proof. Again automatic after unfolding definitions. □

It is easy to prove that, if we already know that two commands are equivalent under a condition P , we are allowed to weaken the statement by strengthening that precondition:

$$\llbracket P \models c \sim c'; \forall s. P' s \longrightarrow P s \rrbracket \Longrightarrow P' \models c \sim c'$$

For the old notion of semantic equivalence we had the concept of congruence rules, where two commands remain equivalent if equivalent sub-commands are substituted for each other. The corresponding rules in the new setting are slightly more interesting. [Figure 10.5](#) gives an overview. The first rule, for sequential composition, has three premises instead of two. The first two are standard, i.e., equivalence of c and c' as well as d and d' . As for the sets of initialized variables in the definite initialization analysis of [Section 10.1](#), we allow the precondition to change. The first premise gets the same P as the conclusion $P \models c;; d \sim c';; d'$, but the second premise can use a new Q . The third premise describes the relationship between P and Q : Q must hold in the states after execution of c , provided P held in the initial state.

The rule for *IF* is simpler; it just demands that the constituent expressions and commands are equivalent under the same condition P . As for the

$$\begin{array}{c}
\frac{P \models c \sim c' \quad Q \models d \sim d' \quad \forall s s'. (c, s) \Rightarrow s' \longrightarrow P s \longrightarrow Q s'}{P \models c;; d \sim c';; d'} \\
\\
\frac{P \models b <\sim> b' \quad P \models c \sim c' \quad P \models d \sim d'}{P \models \text{IF } b \text{ THEN } c \text{ ELSE } d \sim \text{IF } b' \text{ THEN } c' \text{ ELSE } d'} \\
\\
\frac{P \models c \sim c' \quad \forall s s'. (c, s) \Rightarrow s' \longrightarrow P s \longrightarrow \text{bval } b \text{ } s \longrightarrow P s' \quad P \models b <\sim> b'}{P \models \text{WHILE } b \text{ DO } c \sim \text{WHILE } b' \text{ DO } c'}
\end{array}$$

Fig. 10.5. Congruence rules for conditional semantic equivalence

semicolon case, we could provide a stronger rule here that takes into account which branch of the *IF* we are looking at, i.e., adding b or $\neg b$ to the condition P . Since we do not analyse the content of boolean expressions, we will not need the added power and prefer the weaker, but simpler rule.

The *WHILE* rule is similar to the semicolon case, but again in a weaker formulation. We demand that b and b' be equivalent under P , as well as c and c' . We additionally need to make sure that P still holds after the execution of the body if it held before, because the loop might enter another iteration. In other words, we need to prove as a side condition that P is an *invariant* of the loop. Since we only need to know this in the iteration case, we can additionally assume that the boolean condition b evaluates to true.

This concludes our brief interlude into conditional semantic equivalence. As indicated in Section 7.2.4, we leave the proof of the rules in Figure 10.5 as an exercise, as well as the formulation of the strengthened rules that take boolean expressions further into account.

10.2.4 Correctness

So far we have defined constant folding and propagation, and we have developed a tool set for reasoning about conditional equivalence of commands. In this section, we apply this tool set to show correctness of our optimization.

As mentioned before, the eventual aim for our correctness statement is unconditional equivalence between the original and the optimized command:

$$\text{fold } c \text{ } (\lambda x. \text{None}) \sim c$$

To prove this statement by induction, we generalize it by replacing the empty table with an arbitrary table t . The price we pay is that the equivalence is now only true under the condition that the table correctly approximates the state the commands are run from. The statement becomes

$$\text{approx } t \models c \sim \text{fold } c \text{ } t$$

Note that the term *approx t* is partially applied. It is a function that takes a state *s* and returns *True* iff *t* is an approximation of *s* as defined previously in Section 10.2.1. Expanding the definition of equivalence we get the more verbose but perhaps easier to understand form.

$$\forall s s'. \text{approx } t \ s \longrightarrow (c, s) \Rightarrow s' = (\text{fold } c \ t, s) \Rightarrow s'$$

For the proof it is nicer not to unfold the definition equivalence and work with the congruence lemmas of the previous section instead. Now, proceeding to prove this property by induction on *c* it quickly turns out that we will need four key lemmas about the auxiliary functions mentioned in *fold*.

The most direct and intuitive one of these is that our *defs* correctly approximates real execution. Recall that *defs* statically analyses which constant values can be assigned to which variables.

Lemma 10.9 (*defs approximates execution correctly*).

$$\llbracket (c, s) \Rightarrow s'; \text{approx } t \ s \rrbracket \Longrightarrow \text{approx } (\text{defs } c \ t) \ s'$$

Proof. The proof is by rule induction on the big-step execution:

- The *SKIP* base case is trivial.
- The assignment case needs some massaging to succeed. After unfolding of definitions, case distinction on the arithmetic expression and simplification we end up with

$$\forall n. \text{afold } a \ t = N \ n \longrightarrow \text{aval } a \ s = n$$

where we also know our general assumption *approx t s*. This is a reformulated instance of Lemma 10.6.

- Sequential composition is simply an application of the two induction hypotheses.
- The two *IF* cases reduce to this property of *merge* which embodies that it is an intersection:

$$\text{approx } t_1 \ s \vee \text{approx } t_2 \ s \Longrightarrow \text{approx } (\text{merge } t_1 \ t_2) \ s$$

In each of the two *IF* cases we know from the induction hypothesis that the execution of the chosen branch is approximated correctly by *defs*, e.g., *approx (defs c₁ t) s'*. With the above *merge* lemma, we can conclude the case.

- In the *False* case for *WHILE* we observe that we are restricting the existing table *t*, and that approximation is trivially preserved when dropping elements.
- In the *True* case we appeal to another lemma about *defs*. From applying induction hypotheses, we know *approx (defs c t|_(- lvars c)) s'*, but our proof goal for *defs* applied to the while loop is *approx (t|_(- lvars c)) s'*. Lemma 10.10 shows that these are equal.

□

The last case of our proof above rests on one lemma we have not shown yet. It says that our restriction to variables that do not occur on the left-hand sides of assignments is broad enough, i.e., that it appropriately masks any new table entries we would get by running *defs* on the loop body.

Lemma 10.10. $defs\ c\ t|_{(-\ lvars\ c)} = t|_{(-\ lvars\ c)}$

Proof. This proof is by induction on c . Most cases are automatic, merely for sequential composition and *IF* Isabelle needs a bit of hand-holding for applying the induction hypotheses at the right position in the term. In the *IF* case, we also make use of this property of merge:

$$\llbracket t_1|_S = t|_S; t_2|_S = t|_S \rrbracket \implies merge\ t_1\ t_2|_S = t|_S$$

It allows us to merge the two equations we get for the two branches of the *IF* into one. \square

The final lemma we need before we can proceed to the main induction is again a property about the restriction of t to the complement of *lvars*. It is the remaining fact we need for the *WHILE* case of that induction and it says that runtime execution can at most change the values of variables that are mentioned on the left-hand side of assignments.

Lemma 10.11.

$$\llbracket (c, s) \Rightarrow s'; approx\ (t|_{(-\ lvars\ c)})\ s \rrbracket \implies approx\ (t|_{(-\ lvars\ c)})\ s'$$

Proof. This proof is by rule induction on the big-step execution. Its cases are very similar to those of [Lemma 10.10](#). \square

Putting everything together, we can now prove our main lemma.

Lemma 10.12 (Generalized correctness of constant folding).

$$approx\ t \models c \sim fold\ c\ t$$

Proof. As mentioned, the proof is by induction on c . *SKIP* is simple, and assignment reduces to the correctness of *afold*, [Lemma 10.6](#). Sequential composition uses the congruence rule for semicolon and [Lemma 10.9](#). The *IF* case is automatic given the *IF* congruence rule. The *WHILE* case reduces to [Lemma 10.11](#), the *WHILE* congruence rule, and strengthening of the equivalence condition. The strengthening uses the following property

$$\llbracket approx\ t_2\ s; t_1 \subseteq_m t_2 \rrbracket \implies approx\ t_1\ s$$

where $(m_1 \subseteq_m m_2) = (m_1 = m_2\ on\ dom\ m_1)$ and $t|_S \subseteq_m t$. \square

This leads us to the final result.

Theorem 10.13 (Correctness of constant folding).

$$fold\ c\ (\lambda x. None) \sim c$$

Proof. Follows immediately from [Lemma 10.12](#) after observing that $approx\ (\lambda x. None) = (\lambda _ . True)$. \square

Exercises

Exercise 10.2. Extend *afold* with simplifying addition of 0. That is, for any expression e , $e + 0$ and $0 + e$ should be simplified to e , including the case where the 0 is produced by knowledge of the content of variables. Re-prove the results in this section with the extended version.

Exercise 10.3. Strengthen and re-prove the congruence rules for conditional semantic equivalence in [Figure 10.5](#) to take the value of boolean expressions into account in the IF and WHILE cases.

Exercise 10.4. Extend constant folding with analysing boolean expressions and eliminate dead IF branches as well as loops whose body is never executed. Hint: you will need to make use of stronger congruence rules for conditional semantic equivalence.

Exercise 10.5. This exercise builds infrastructure for [Exercise 10.6](#), where we will have to manipulate partial maps from variable names to variable names.

In addition to the function *merge* from theory *Fold*, implement two functions *remove* and *remove_all* that remove one variable name from the range of a map, and a set of variable names from the domain and range of a map.

Prove the following properties:

$$\begin{aligned} \text{ran } (\text{remove } x \ t) &= \text{ran } t - \{x\} \\ \text{ran } (\text{remove_all } S \ t) &\subseteq - S \\ \text{dom } (\text{remove_all } S \ t) &\subseteq - S \\ \text{remove_all } \{x\} (\text{remove } x \ t) &= \text{remove_all } \{x\} \ t \\ \text{remove_all } A (\text{remove_all } B \ t) &= \text{remove_all } (A \cup B) \ t \end{aligned}$$

Reformulate the property $\llbracket t_1 \upharpoonright_S = t \upharpoonright_S; t_2 \upharpoonright_S = t \upharpoonright_S \rrbracket \implies \text{merge } t_1 \ t_2 \upharpoonright_S = t \upharpoonright_S$ from [Lemma 10.10](#) for *remove_all* and prove it.

Exercise 10.6. This is a more challenging exercise. Define and prove correct *copy propagation*. Copy propagation is similar to constant folding, but propagates the right-hand side of assignments if these right-hand sides are just variables. For instance, the program $x := y; z := x + z$ will be transformed into $x := y; z := y + z$. The assignment $x := y$ can then be eliminated in a liveness analysis. Copy propagation is useful for cleaning up after other optimization phases.

10.3 Live Variable Analysis thy

This section presents another important analysis that enables program optimizations, namely the elimination of assignments to a variable whose value is not needed afterwards. Here is a simple example:

```
x := 0; y := 1; x := y
```

The first assignment to x is redundant because x is dead at this point: it is overwritten by the second assignment to x without x having been read in between. In contrast, the assignment to y is not redundant because y is live at that point.

Semantically, variable x is live before command c if the initial value of x before execution of c can influence the final state after execution of c . A weaker but easier to check condition is the following: we call x live before c if there is some potential execution of c where x is read for the first time before it is overwritten. For the moment, all variables are implicitly read at the end of c . A variable is dead if it is not live. The phrase “potential execution” refers to the fact that we do not analyse boolean expressions.

Example 10.14.

- $x := \text{rhs}$
Variable x is dead before this assignment unless rhs contains x .
Variable y is live before this assignment if rhs contains y .
- IF b THEN $x := y$ ELSE SKIP
Variable y is live before this command because execution could potentially enter the THEN branch.
- $x := y; x := 0; y := 1$
Variable y is live before this command (because of $x := y$) although the value of y is semantically irrelevant because the second assignment overwrites the first one. This example shows that the above definition of liveness is strictly weaker than the semantic notion. We will improve on this under the heading of “True Liveness” in [Section 10.4](#).

Let us now formulate liveness analysis as a recursive function. This requires us to generalize the liveness notion w.r.t. a set of variables X that are implicitly read at the end of a command. The reason is that this set changes during the analysis. Therefore it needs to be a parameter of the analysis and we speak of the set of variables live before a command c relative to a set of variables X . It is computed by the function $L\ c\ X$ defined like this:

```
fun L :: com => vname set => vname set where
  L SKIP X                = X
  L (x ::= a) X            = vars a ∪ (X - {x})
  L (c1;; c2) X            = L c1 (L c2 X)
  L (IF b THEN c1 ELSE c2) X = vars b ∪ L c1 X ∪ L c2 X
  L (WHILE b DO c) X       = vars b ∪ X ∪ L c X
```

In a nutshell, $L\ c\ X$ computes the set of variables that are live *before* c given the set of variables X that are live *after* c (hence the order of arguments in $L\ c\ X$).

We discuss the equations for L one by one. The one for $SKIP$ is obvious. The one for $x ::= a$ expresses that before the assignment all variables in a are live (because they are read) and that x is not live (because it is overwritten) unless it also occurs in a . The equation for $c_1;; c_2$ expresses that the computation of live variables proceeds backwards. The equation for IF expresses that the variables of b are read, that b is not analysed and both the $THEN$ and the $ELSE$ branch could be executed, and that a variable is live if it is live on some computation path leading to some point — hence the \cup . The situation for $WHILE$ is similar: execution could skip the loop (hence X is live) or it could execute the loop body once (hence $L \ c \ X$ is live). But what if the loop body is executed multiple times?

In the following discussion we assume this abbreviation:

$$w = WHILE \ b \ DO \ c$$

For a more intuitive understanding of the analysis of loops one should think of w as the control-flow graph in Figure 10.6. A control-flow graph is

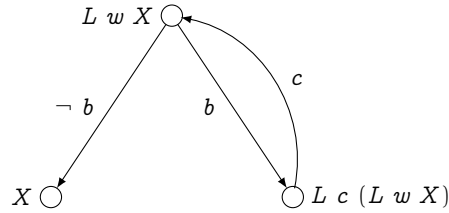


Fig. 10.6. Control-flow graph for $WHILE \ b \ DO \ c$

a graph whose nodes represent program points and whose edges are labelled with boolean expressions or commands. The operational meaning is that execution moves a state from node to node: a state s moves unchanged across an edge labelled with b provided $bval \ b \ s$, and moving across an edge labelled with c transforms s into the new state resulting from the execution of c .

In Figure 10.6 we have additionally annotated each node with the set of variables live at that node. At the exit of the loop, X should be live, at the beginning, $L \ w \ X$ should be live. Let us pretend we had not defined $L \ w \ X$ yet but were looking for constraints that it must satisfy. An edge $Y \xrightarrow{e} Z$ (where e is a boolean expression and Y, Z are liveness annotations) should satisfy $vars \ e \subseteq Y$ and $Z \subseteq Y$ (because the variables in e are read but no variable is written). Thus the graph leads to the following three constraints:

$$\left. \begin{array}{l} \text{vars } b \quad \subseteq L \ w \ X \\ X \quad \subseteq L \ w \ X \\ L \ c \ (L \ w \ X) \subseteq L \ w \ X \end{array} \right\} \quad (10.1)$$

The first two constraints are met by our definition of L , but for the third constraint this is not clear. To facilitate proofs about L we now rephrase its definition as an instance of a general analysis.

10.3.1 Generate and Kill Analyses

This is a class of simple analyses that operate on sets. That is, each analysis in this class is a function $A :: com \Rightarrow \tau \ set \Rightarrow \tau \ set$ (for some type τ) that can be defined as

$$A \ c \ S = gen \ c \cup (S - kill \ c)$$

for suitable auxiliary functions gen and $kill$ of type $com \Rightarrow \tau \ set$ that specify what is to be added and what is to be removed from the input set. Gen/kill analyses satisfy nice algebraic properties and many standard analyses can be expressed in this form, in particular liveness analysis. For liveness, $gen \ c$ are the variables that may be read in c before they are written and $kill \ c$ are the variables that are definitely written in c :

```

fun kill :: com  $\Rightarrow$  vname set where
kill SKIP                                = {}
kill (x ::= a)                          = {x}
kill (c1;; c2)                          = kill c1  $\cup$  kill c2
kill (IF b THEN c1 ELSE c2) = kill c1  $\cap$  kill c2
kill (WHILE b DO c)                  = {}

fun gen :: com  $\Rightarrow$  vname set where
gen SKIP                                = {}
gen (x ::= a)                          = vars a
gen (c1;; c2)                          = gen c1  $\cup$  (gen c2 - kill c1)
gen (IF b THEN c1 ELSE c2) = vars b  $\cup$  gen c1  $\cup$  gen c2
gen (WHILE b DO c)                  = vars b  $\cup$  gen c

```

Note that gen uses $kill$ in the only not-quite-obvious equation $gen \ (c_1;; c_2) = gen \ c_1 \cup (gen \ c_2 - kill \ c_1)$ where $gen \ c_2 - kill \ c_1$ expresses that variables that are read in c_2 but were written in c_1 are not live before $c_1;; c_2$ (unless they are also in $gen \ c_1$).

Lemma 10.15 (Liveness via gen/kill). $L \ c \ X = gen \ c \cup (X - kill \ c)$

The proof is a simple induction on c . As a consequence of this lemma we obtain

Lemma 10.16. $L\ c\ (L\ w\ X) \subseteq L\ w\ X$

Proof. By the previous lemma it follows that $gen\ c \subseteq gen\ w \subseteq L\ w\ X$ and thus that $L\ c\ (L\ w\ X) = gen\ c \cup (L\ w\ X - kill\ c) \subseteq L\ w\ X$. \square

Moreover, we can prove that $L\ w\ X$ is the *least* solution for the constraint system (10.1). This shows that our definition of $L\ w\ X$ is optimal: the fewer live variables, the better; from the perspective of program optimization, the only good variable is a dead variable. To prove that $L\ w\ X$ is the least solution of (10.1), assume that P is a solution of (10.1), i.e., $vars\ b \subseteq P$, $X \subseteq P$ and $L\ c\ P \subseteq P$. Because $L\ c\ P = gen\ c \cup (P - kill\ c)$ we also have $gen\ c \subseteq P$. Thus $L\ w\ X = vars\ b \cup gen\ c \cup X \subseteq P$ by assumptions.

10.3.2 Correctness

So far we have proved that $L\ w\ X$ satisfies the informally derived constraints. Now we prove formally that L is correct w.r.t. the big-step semantics. Roughly speaking we want the following: when executing c , the final value of $x \in X$ only depends on the initial values of variables in $L\ c\ X$. Put differently:

If two initial states of the execution of c agree on $L\ c\ X$
then the corresponding final states agree on X .

To formalize this statement we introduce the abbreviation $f = g\ on\ X$ for two functions $f, g :: 'a \Rightarrow 'b$ being the same on a set $X :: 'a\ set$:

$$f = g\ on\ X \equiv \forall x \in X. f\ x = g\ x$$

With this notation we can concisely express that the value of an expression only depends of the value of the variables in the expression:

Lemma 10.17 (Coincidence).

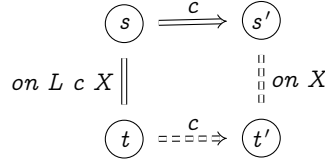
1. $s_1 = s_2\ on\ vars\ a \implies aval\ a\ s_1 = aval\ a\ s_2$
2. $s_1 = s_2\ on\ vars\ b \implies bval\ b\ s_1 = bval\ b\ s_2$

The proofs are by induction on a and b .

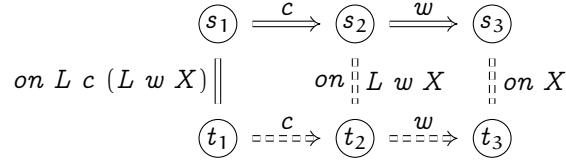
The actual correctness statement for live variable analysis is a simulation property (see Section 8.4). The diagrammatic form of the statement is shown in Figure 10.7 where $(c, s) \Rightarrow t$ is displayed as $s \xRightarrow{c} t$. Theorem 10.18 expresses the diagram as a formula.

Theorem 10.18 (Correctness of L).

$$\llbracket (c, s) \Rightarrow s'; s = t\ on\ L\ c\ X \rrbracket \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t'\ on\ X$$

Fig. 10.7. Correctness of L

Proof. The proof is by rule induction. The only interesting cases are the assignment rule, which is correct by the Coincidence Lemma, and rule *WhileTrue*. For the correctness proof of the latter we assume its hypotheses $\text{bval } b \ s_1$, $(c, s_1) \Rightarrow s_2$ and $(w, s_2) \Rightarrow s_3$. Moreover we assume $s_1 = t_1 \text{ on } L \ w \ X$ and therefore in particular $s_1 = t_1 \text{ on } L \ c \ (L \ w \ X)$ because $L \ c \ (L \ w \ X) \subseteq L \ w \ X$. Thus the induction hypothesis for $(c, s_1) \Rightarrow s_2$ applies and we obtain t_2 such that $(c, t_1) \Rightarrow t_2$ and $s_2 = t_2 \text{ on } L \ w \ X$. The latter enables the application of the induction hypothesis for $(w, s_2) \Rightarrow s_3$, which yields t_3 such that $(w, t_2) \Rightarrow t_3$ and $s_3 = t_3 \text{ on } X$. By means of the Coincidence Lemma, $s_1 = t_1 \text{ on } L \ w \ X$ and $\text{vars } b \subseteq L \ w \ X$ imply $\text{bval } b \ s_1 = \text{bval } b \ t_1$. Therefore rule *WhileTrue* yields $(w, t_1) \Rightarrow t_3$ as required. A graphical view of the skeleton of this argument:



□

Note that the proofs of the loop cases (*WhileFalse* too) do not rely on the definition of L but merely on the constraints (10.1).

10.3.3 Optimization

With the help of the analysis we can program an optimizer *bury* $c \ X$ that eliminates assignments to dead variables from c where X is of course the set of variables live at the end.

```

fun bury :: com  $\Rightarrow$  vname set  $\Rightarrow$  com where
  bury SKIP X      = SKIP
  bury (x ::= a) X = (if x  $\in$  X then x ::= a else SKIP)
  bury (c1;; c2) X = bury c1 (L c2 X);; bury c2 X
  bury (IF b THEN c1 ELSE c2) X =

```

$$\begin{aligned}
& \text{IF } b \text{ THEN } \text{bury } c_1 \text{ } X \text{ ELSE } \text{bury } c_2 \text{ } X \\
& \text{bury } (\text{WHILE } b \text{ DO } c) \text{ } X = \\
& \text{WHILE } b \text{ DO } \text{bury } c \text{ } (L (\text{WHILE } b \text{ DO } c) \text{ } X)
\end{aligned}$$

For simplicity assignments to dead variables are replaced with *SKIP* — eliminating such *SKIPS* in a separate pass was dealt with in [Exercise 7.3](#).

Most of the equations for *bury* are obvious from our understanding of *L*. For the recursive call of *bury* in the *WHILE* case note that the variables live after *c* are *L w X* (just look at [Figure 10.6](#)).

Now we need to understand in what sense this optimization is correct and then prove it. Our correctness criterion is the big-step semantics: the transformed program must be equivalent to the original one: $\text{bury } c \text{ } UNIV \sim c$. Note that *UNIV* (or at least all the variables in *c*) must be considered live at the end because $c' \sim c$ requires that the final states in the execution of *c* and *c'* are identical.

The proof of $\text{bury } c \text{ } UNIV \sim c$ is split into two directions. We start with $(c, s) \Rightarrow s' \implies (\text{bury } c \text{ } UNIV, s) \Rightarrow s'$. For the induction to go through it needs to be generalized to look almost like the correctness statement for *L*. [Figure 10.8](#) is the diagrammatic form of [Lemma 10.19](#) and [Lemma 10.20](#).

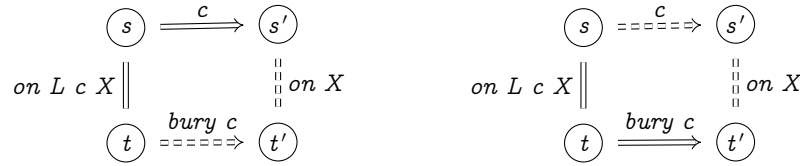


Fig. 10.8. Correctness of *bury* (both directions)

Lemma 10.19 (Correctness of *bury*, part 1).

$$\begin{aligned}
& \llbracket (c, s) \Rightarrow s'; s = t \text{ on } L \text{ } c \text{ } X \rrbracket \\
& \implies \exists t'. (\text{bury } c \text{ } X, t) \Rightarrow t' \wedge s' = t' \text{ on } X
\end{aligned}$$

The proof is very similar to the proof of correctness of *L*. Hence there is no need to go into it.

The other direction $(\text{bury } c \text{ } UNIV, s) \Rightarrow s' \implies (c, s) \Rightarrow s'$ needs to be generalized analogously:

Lemma 10.20 (Correctness of *bury*, part 2).

$$\begin{aligned}
& \llbracket (\text{bury } c \text{ } X, s) \Rightarrow s'; s = t \text{ on } L \text{ } c \text{ } X \rrbracket \\
& \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t' \text{ on } X
\end{aligned}$$

Proof. The proof is also similar to that of correctness of L but induction requires the advanced form explained in Section 5.4.6. As a result, in each case of the induction, there will be an assumption that $bury\ c\ X$ is of a particular form, e.g., $c'_1;; c'_2 = bury\ c\ X$. Now we need to infer that c must be a sequential composition too. The following property expresses this fact:

$$\begin{aligned} (c'_1;; c'_2 = bury\ c\ X) = \\ (\exists c_1\ c_2. c = c_1;; c_2 \wedge c'_2 = bury\ c_2\ X \wedge c'_1 = bury\ c_1\ (L\ c_2\ X)) \end{aligned}$$

Its proof is by cases on c ; the rest is automatic. This property can either be proved as a separate lemma or c'_1 and c'_2 can be obtained locally within the case of the induction that deals with sequential composition. Because $bury$ preserves the structure of the command, all the required lemmas look similar, except for $SKIP$, which can be the result of either a $SKIP$ or an assignment:

$$(SKIP = bury\ c\ X) = (c = SKIP \vee (\exists x\ a. c = x ::= a \wedge x \notin X))$$

We need only the left-to-right implication of these function inversion properties but the formulation as an equivalence permits us to use them as simplification rules. \square

Combining the previous two lemmas we obtain

Corollary 10.21 (Correctness of $bury$). $bury\ c\ UNIV \sim c$

Exercises

Exercise 10.7. Prove the following termination-insensitive version of the correctness of L :

$$\llbracket (c, s) \Rightarrow s'; (c, t) \Rightarrow t'; s = t\ on\ L\ c\ X \rrbracket \implies s' = t'\ on\ X$$

Do not derive it as a corollary to the original correctness theorem but prove it separately. Hint: modify the original proof.

Exercise 10.8. Find a command c such that $bury\ (bury\ c\ \{\})\ \{\} \neq bury\ c\ \{\}$. For an arbitrary command, can you put a limit on the amount of burying needed until everything that is dead is also buried?

Exercise 10.9. Let $lvars\ c/rvars\ c$ be the set of variables that occur on the left-hand/right-hand side of an assignment in c . Let $rvars\ c$ additionally include those variables mentioned in the conditionals of IF and $WHILE$. Both functions are predefined in theory *Vars*. Prove the following two properties of the small-step semantics. Variables that are not assigned to do not change their value:

$$\llbracket (c, s) \rightarrow^* (c', s'); lvars\ c \cap X = \{\} \rrbracket \implies s = s'\ on\ X$$

The reduction behaviour of a command is only influenced by the variables read by the command:

$$\begin{aligned} \llbracket (c, s) \rightarrow^* (c', s'); s = t \text{ on } X; rvars\ c \subseteq X \rrbracket \\ \implies \exists t'. (c, t) \rightarrow^* (c', t') \wedge s' = t' \text{ on } X \end{aligned}$$

Exercise 10.10. An available definitions analysis determines which previous assignments $x := a$ are valid equalities $x = a$ at later program points. For example, after $x := y+1$ the equality $x = y+1$ is available, but after $x := y+1; y := 2$ the equality $x = y+1$ is no longer available. The motivation for the analysis is that if $x = a$ is available before $v := a$ then $v := a$ can be replaced by $v := x$.

Define an available definitions analysis $AD :: (vname \times aexp) \text{ set} \Rightarrow com \Rightarrow (vname \times aexp) \text{ set}$. A call $AD\ A\ c$ should compute the available definitions after the execution of c assuming that the definitions in A are available before the execution of c . This is a gen/kill analysis! Prove correctness of the analysis: if $(c, s) \Rightarrow s'$ and $\forall (x, a) \in A. s\ x = \text{aval}\ a\ s$ then $\forall (x, a) \in AD\ A\ c. s'\ x = \text{aval}\ a\ s'$.

10.4 True Liveness thy

In [Example 10.14](#) we had already seen that our definition of liveness is too simplistic: in $x := y; x := 0$, variable y is read before it can be written, but it is read in an assignment to a dead variable. Therefore we modify the definition of $L\ (x ::= a)\ X$ to consider *vars* a live only if x is live:

$$L\ (x ::= a)\ X = (\text{if } x \in X \text{ then vars } a \cup (X - \{x\}) \text{ else } X)$$

As a result, our old analysis of loops is no longer correct.

Example 10.22. Consider for a moment the following specific w and c

$$\begin{aligned} w &= \text{WHILE Less } (N\ 0)\ (V\ x)\ \text{DO } c \\ c &= x ::= V\ y;; y ::= V\ z \end{aligned}$$

where x, y and z are distinct. Then $L\ w\ \{x\} = \{x, y\}$ but z is live too, semantically: the initial value of z can influence the final value of x . This is the computation of L : $L\ c\ \{x\} = L\ (x ::= V\ y)\ (L\ (y ::= V\ z)\ \{x\}) = L\ (x ::= V\ y)\ \{x\} = \{y\}$ and therefore $L\ w\ \{x\} = \{x\} \cup \{x\} \cup L\ c\ \{x\} = \{x, y\}$. The reason is that $L\ w\ X = \{x, y\}$ is no longer a solution of the last constraint of (10.1): $L\ c\ \{x, y\} = L\ (x ::= V\ y)\ (L\ (y ::= V\ z)\ \{x, y\}) = L\ (x ::= V\ y)\ \{x, z\} = \{y, z\} \not\subseteq \{x, y\}$.

Let us now abstract from this example and reconsider (10.1). We still want $L \ w \ X$ to be a solution of (10.1) because the proof of correctness of L depends on it. An equivalent formulation of (10.1) is

$$\text{vars } b \cup X \cup L \ c \ (L \ w \ X) \subseteq L \ w \ X \quad (10.2)$$

That is, $L \ w \ X$ should be some set Y such that $\text{vars } b \cup X \cup L \ c \ Y \subseteq Y$. For optimality we want the least such Y . We will now study abstractly under what conditions a least such Y exists and how to compute it.

10.4.1 The Knaster-Tarski Fixpoint Theorem on Sets

Definition 10.23. A type $'a$ is a *partial order* if there is binary predicate \leq on $'a$ which is

reflexive: $x \leq x$,
transitive: $\llbracket x \leq y; y \leq z \rrbracket \implies x \leq z$, and
antisymmetric: $\llbracket a \leq b; b \leq a \rrbracket \implies a = b$.

We use “ \leq is a partial order” and “type $'a$ is a partial order” interchangeably.

Definition 10.24. If $'a$ is a partial order and $A :: 'a$ set, then $p \in A$ is a *least element* of A if $p \leq q$ for all $q \in A$.

Least elements are unique: if p_1 and p_2 are least elements of A then $p_1 \leq p_2$ and $p_2 \leq p_1$ and hence $p_1 = p_2$ by antisymmetry.

Definition 10.25. Let τ be a type and $f :: \tau \Rightarrow \tau$. A point $p :: \tau$ is a *fixpoint* of f if $f \ p = p$.

Definition 10.26. Let τ be a type with a partial order \leq and $f :: \tau \Rightarrow \tau$.

- Function f is *monotone* if $x \leq y \implies f \ x \leq f \ y$ for all x and y .
- A point $p :: \tau$ is a *pre-fixpoint* of f if $f \ p \leq p$.

This definition applies in particular to sets, where the partial order is \subseteq . Hence we can rephrase (10.2) as saying that $L \ w \ X$ should be a pre-fixpoint of $\lambda Y. \text{vars } b \cup X \cup L \ c \ Y$, ideally the least one. The Knaster-Tarski fixpoint theorem tells us that all we need is monotonicity.

Theorem 10.27. If $f :: \tau \text{ set} \Rightarrow \tau \text{ set}$ is monotone then $\bigcap \{P. f \ P \subseteq P\}$ is the *least pre-fixpoint* of f .

Proof. Let $M = \{P. f \ P \subseteq P\}$. First we show that $\bigcap M$ is a pre-fixpoint. For all $P \in M$ we have $\bigcap M \subseteq P$ (by definition of \bigcap) and therefore $f \ (\bigcap M) \subseteq f \ P \subseteq P$ (by monotonicity and because $P \in M$). Therefore $f \ (\bigcap M) \subseteq \bigcap M$ (by definition of \bigcap). Moreover, $\bigcap M$ is the least pre-fixpoint: if $P \subseteq P$ then $P \in M$ and thus $\bigcap M \subseteq P$. \square

Lemma 10.28. *Let f be a monotone function on a partial order \leq . Then a least pre-fixpoint of f is also a least fixpoint.*

Proof. Let P be a least pre-fixpoint of f . From $f P \leq P$ and monotonicity we have $f(f P) \leq f P \leq P$ and therefore $P \leq f P$ because P is a least pre-fixpoint. Together with $f P \leq P$ this yields $f P = P$ (by antisymmetry). Moreover, P is the least fixpoint because any fixpoint is a pre-fixpoint and P is the least pre-fixpoint. \square

The **Knaster-Tarski fixpoint theorem** is the combination of both results:

Theorem 10.29. *If $f :: \tau \text{ set} \Rightarrow \tau \text{ set}$ is monotone then*

$$\text{lfp } f = \bigcap \{P. f P \subseteq P\}$$

is the least pre-fixpoint of f , which is also its least fixpoint.

We have separated the two parts because the second part is a generally useful result about partial orders.

In the Isabelle library this theorem is expressed by two lemmas:

lfp_unfold: $\text{mono } f \Longrightarrow \text{lfp } f = f (\text{lfp } f)$

lfp_lowerbound: $f A \leq A \Longrightarrow \text{lfp } f \leq A$

where *mono_def:* $\text{mono } f = (\forall x y. x \leq y \longrightarrow f x \leq f y)$.

Now we boldly define $L w X$ as a least fixpoint:

$$L (WHILE\ b\ DO\ c)\ X = \text{lfp } (\lambda Y. \text{vars } b \cup X \cup L\ c\ Y)$$

This is the full definition of our revised L for true liveness:

```

fun  $L :: com \Rightarrow vname\ set \Rightarrow vname\ set$  where
   $L\ SKIP\ X = X$  |
   $L\ (x ::= a)\ X = (if\ x \in X\ then\ \text{vars } a \cup (X - \{x\})\ else\ X)$  |
   $L\ (c_1;;\ c_2)\ X = L\ c_1\ (L\ c_2\ X)$  |
   $L\ (IF\ b\ THEN\ c_1\ ELSE\ c_2)\ X = \text{vars } b \cup L\ c_1\ X \cup L\ c_2\ X$  |
   $L\ (WHILE\ b\ DO\ c)\ X = \text{lfp } (\lambda Y. \text{vars } b \cup X \cup L\ c\ Y)$ 

```

Only the assignment and *WHILE* cases differ from the version in [Section 10.3](#).

The definition of the *WHILE* case is bold for two reasons: we do not yet know that $\lambda Y. \text{vars } b \cup X \cup L\ c\ Y$ is monotone, i.e., that it makes sense to apply *lfp* to it, and we have no idea how to compute *lfp*.

Lemma 10.30. *$L\ c$ is monotone.*

Proof. The proof is by induction on c . All cases are essentially trivial because \cup is monotone in both arguments and set difference is monotone in the first argument. In the *WHILE* case we additionally need that *lfp* is monotone in the following sense: if $f A \subseteq g A$ for all A , then $\text{lfp } f \subseteq \text{lfp } g$. This is obvious because any pre-fixpoint of g must also be a pre-fixpoint of f . \square

As a corollary we obtain that $\lambda Y. \text{vars } b \cup X \cup L \ c \ Y$ is monotone too. Hence, by the Knaster-Tarski fixpoint theorem, it has a least (pre-)fixpoint. Hence $L \ w \ X$ is defined exactly as required, i.e., it satisfies (10.1).

Lemma 10.31 (Correctness of L).

$$\llbracket (c, s) \Rightarrow s'; s = t \text{ on } L \ c \ X \rrbracket \implies \exists t'. (c, t) \Rightarrow t' \wedge s' = t' \text{ on } X$$

Proof. This is the same correctness lemma as in Section 10.3, proved in the same way, but for a modified L . The proof of the *WHILE* case remains unchanged because it only relied on the pre-fixpoint constraints (10.1) that are still satisfied. The proof of correctness of the new definition of $L \ (x ::= a) \ X$ is just as routine as before. \square

How about an optimizer *bury* as in the previous section? It turns out that we can reuse that *bury* function verbatim, with true liveness instead of liveness. What is more, even the correctness proof carries over verbatim. The reason: the proof of the *WHILE* case relies only on the constraints (10.1), which hold for both L .

10.4.2 Computing the Least Fixpoint

Under certain conditions, the least fixpoint of a function f can be computed by iterating f , starting from the least element, which in the case of sets is the empty set. By iteration we mean f^n defined as follows:

$$\begin{aligned} f^0 \ x &= x \\ f^{n+1} \ x &= f \ (f^n \ x) \end{aligned}$$

In Isabelle, $f^n \ x$ is input as $(f \ ^{\wedge} \ n) \ x$.

Lemma 10.32. *Let $f :: \tau \text{ set} \Rightarrow \tau \text{ set}$ be a monotone function. If the chain $\{\} \subseteq f \ \{\} \subseteq f^2 \ \{\} \subseteq \dots$ stabilizes after k steps, i.e., $f^{k+1} \ \{\} = f^k \ \{\}$, then $\text{lfp } f = f^k \ \{\}$.*

Proof. From $f^{k+1} \ \{\} = f^k \ \{\}$ it follows that $f^k \ \{\}$ is a fixpoint. It is the least fixpoint because it is a subset of any other fixpoint P of f : $f^n \ \{\} \subseteq P$ follows from monotonicity by an easy induction on n . The fact that the $f^n \ \{\}$ form a chain, i.e., $f^n \ \{\} \subseteq f^{n+1} \ \{\}$, follows by a similar induction. \square

This gives us a way to compute the least fixpoint. In general this will not terminate, as the sets can grow larger and larger. However, in our application only a finite set of variables is involved, those in the program. Therefore termination is guaranteed.

Example 10.33. Recall the loop from Example 10.22:

```

w = WHILE Less (N 0) (V x) DO c
c = x ::= V y;; y ::= V z

```

To compute $L\ w\ \{x\}$ we iterate $f = (\lambda Y. \{x\} \cup L\ c\ Y)$. For compactness the notation $X_2\ c_1\ X_1\ c_2\ X_0$ (where the X_i are sets of variables) abbreviates $X_1 = L\ c_2\ X_0$ and $X_2 = L\ c_1\ X_1$. [Figure 10.9](#) shows the computation of $f\ \{\}$ through $f^4\ \{\}$. The final line confirms that $\{x, y, z\}$ is a fixpoint. Of course

$$\begin{aligned}
& \{\} \ x ::= V\ y \ \{\} \quad y ::= V\ z \ \{\} \\
& \quad \implies f\ \{\} = \{x\} \cup \{\} = \{x\} \\
& \{y\} \ x ::= V\ y \ \{x\} \quad y ::= V\ z \ \{x\} \\
& \quad \implies f\ \{x\} = \{x\} \cup \{y\} = \{x, y\} \\
& \{y, z\} \ x ::= V\ y \ \{x, z\} \quad y ::= V\ z \ \{x, y\} \\
& \quad \implies f\ \{x, y\} = \{x\} \cup \{y, z\} = \{x, y, z\} \\
& \{y, z\} \ x ::= V\ y \ \{x, z\} \quad y ::= V\ z \ \{x, y, z\} \\
& \quad \implies f\ \{x, y, z\} = \{x\} \cup \{y, z\} = \{x, y, z\}
\end{aligned}$$

Fig. 10.9. Iteration of f in [Example 10.33](#)

this is obvious because $\{x, y, z\}$ cannot get any bigger: it already contains all the variables of the program.

Let us make the termination argument more precise and derive some concrete bounds. We need to compute the least fixpoint of $f = (\lambda Y. \text{vars } b \cup X \cup L\ c\ Y)$. Informally, the chain of the $f^n\ \{\}$ must stabilize because only a finite set of variables is involved — we assume that X is finite. In the following, let $rvars\ c$ be the set of variables read in c (see [Appendix A](#)). An easy induction on c shows that $L\ c\ X \subseteq rvars\ c \cup X$. Therefore f is bounded by $U = \text{vars } b \cup rvars\ c \cup X$ in the following sense: $Y \subseteq U \implies f\ Y \subseteq U$. Hence $f^k\ \{\}$ is a fixpoint of f for some $k \leq \text{card } U$, the cardinality of U . More precisely, $k \leq \text{card } (rvars\ c) + 1$ because already in the first step $f\ \{\} \supseteq \text{vars } b \cup X$.

It remains to give an executable definition of $L\ w\ X$. Instead of programming the required function iteration ourselves we use a combinator from the HOL-Library theory *While_Combinator*:

```

while :: ('a ⇒ bool) ⇒ ('a ⇒ 'a) ⇒ 'a ⇒ 'a
while b f x = (if b x then while b f (f x) else x)

```

The equation makes *while* executable. Calling *while b f x* leads to a sequence of (tail!) recursive calls *while b f (fⁿ x)*, $n = 0, 1, \dots$, until $\neg b\ (f^k\ x)$ for some k . The equation cannot be a definition because it may not terminate. It is a lemma derived from the actual definition; the latter is a bit tricky and need not concern us here.

Theory *While_Combinator* also contains a lemma that tells us that *while* can implement *lfp* on finite sets provided termination is guaranteed:

Lemma 10.34. *If $f :: 'a \text{ set} \Rightarrow 'a \text{ set}$ is monotone and there is a finite set U such that $A \subseteq U \implies f A \subseteq U$ for all A , then*

$$\text{lfp } f = \text{while } (\lambda A. f A \neq A) f \{\}.$$

This is a consequence of [Lemma 10.32](#) and allows us to prove

$$\begin{aligned} L (WHILE\ b\ DO\ c)\ X = \\ (\text{let } f = \lambda Y. \text{vars } b \cup X \cup L\ c\ Y \text{ in while } (\lambda Y. f Y \neq Y) f \{\}) \end{aligned}$$

for finite X . Finally, L has become executable and for our example

```
value let b = Less (N 0) (V ''x'');
      c = ''x'' ::= V ''y'';; ''y'' ::= V ''z''
      in L (WHILE b DO c) {''x''}
```

Isabelle computes the expected result $\{\text{''x''}, \text{''y''}, \text{''z''}\}$.

To conclude this section we compare true liveness with liveness. The motivation for true liveness was improved precision, and this was achieved: $L (\text{''x''} ::= V \text{''y''}; \text{''x''} ::= N 0) \{\text{''x''}\}$ returns $\{\}$ for true liveness, as opposed to $\{\text{''y''}\}$ for liveness, but at the cost of efficiency: true liveness is no longer a gen/kill analysis that can be performed in a single pass over the program; analysis of loops now needs iteration to compute least fixpoints.

Exercises

Exercise 10.11. Compute $L\ w\ \{\}$ for w as in [Example 10.33](#). The result will be nonempty. Explain why it is not strange that even if we are not interested in the value of any variable after w , some variables may still be live before w . The correctness lemma for L may be helpful.

Exercise 10.12. Find a family of commands c_2, c_3, \dots , such that the computation of $L (WHILE\ b\ DO\ c_n)\ X$ (for suitable X and b) requires n iterations to reach the least fixpoint. Hint: generalize [Example 10.33](#). No need to use Isabelle.

Exercise 10.13. Function *bury* defined in [Section 10.3](#) is not idempotent ([Exercise 10.8](#)). Now define the textually identical function *bury* in the context of true liveness analysis and prove that it is idempotent. Start by proving the following lemma:

$$X \subseteq Y \implies L (\text{bury } c\ Y)\ X = L\ c\ X$$

The proof is straightforward except for the case *While* b c where reasoning about *lfp* is required.

Now idempotence (*bury* (*bury* c X) X = *bury* c X) should be easy.

10.5 Summary and Further Reading

This chapter has explored three different, widely used data-flow analyses and associated program optimizations: definite initialization analysis, constant propagation, and live variable analysis. They can be classified according to two criteria:

Forward/backward

A **forward analysis** propagates information from the beginning to the end of a program.

A **backward analysis** propagates information from the end to the beginning of a program.

May/must

A **may analysis** checks if the given property is true on some path.

A **must analysis** checks if the given property is true on all paths.

According to this schema

- Definite initialization analysis is a forward must analysis: variables must be assigned on all paths before they are used.
- Constant propagation is a forward must analysis: a variable must have the same constant value on all paths.
- Live variable analysis is a backward may analysis: a variable is live if it is used on some path before it is overwritten.

There are also forward may and backward must analyses.

Data-flow analysis arose in the context of compiler construction and is treated in some detail in all decent books on the subject, e.g. [2], but in particular in the book by Muchnik [58]. The book by Nielson, Nielson and Hankin [61] provides a comprehensive and more theoretical account of program analysis.

In [Chapter 13](#) we study “Abstract Interpretation”, a powerful but also complex approach to program analysis that generalizes the algorithms presented in this chapter.

Denotational Semantics thy

So far we have worked exclusively with various operational semantics which are defined by inference rules that tell us how to execute some command. But those rules do not tell us directly what the *meaning* of a command is. This is what **denotational semantics** is about: mapping syntactic objects to their denotation or meaning. In fact, we are already familiar with two examples, namely the evaluation of arithmetic and boolean expressions. The denotation of an arithmetic expression is a function from states to values and $aval :: aexp \Rightarrow (state \Rightarrow val)$ (note the parentheses) is the mapping from syntax to semantics. Similarly, we can think of the meaning of a command as a relation between initial states and final states and can even define

$$Big_step\ c \equiv \{(s, t). (c, s) \Rightarrow t\}$$

If the language is deterministic, this relation is a partial function.

However, Big_step is not a true denotational semantics because all the work happens on the level of the operational semantics. A denotational semantics is characterised as follows:

There is a type *syntax* of syntactic objects, a type *semantics* of denotations and a function $D :: syntax \Rightarrow semantics$ that is defined by primitive recursion. That is, for each syntactic construct C there is a defining equation

$$D\ (C\ x_1\ \dots\ x_n) = \dots\ D\ x_1\ \dots\ D\ x_n\ \dots$$

In words: the meaning of a compound object is defined as a function of the meanings of its subcomponents.

Both $aval$ and $bval$ are denotational definitions, but the big-step semantics is not: the meaning of $WHILE\ b\ DO\ c$ is not defined simply in terms of the meaning of b and the meaning of c : rule *WhileTrue* inductively relies on the meaning of $WHILE\ b\ DO\ c$ in the premise.

One motivation for denotational definitions is that proofs can be conducted by the simple and effective proof principles of equational reasoning and structural induction over the syntax.

11.1 A Relational Denotational Semantics

Although the natural view of the meaning of a deterministic command may be a partial function from initial to final states, it is mathematically simpler to work with relations instead. For this purpose we introduce the identity relation and composition of relations:

$$\begin{aligned} Id &:: ('a \times 'a) \text{ set} \\ Id &= \{p. \exists x. p = (x, x)\} \\ (\circ) &:: ('a \times 'b) \text{ set} \Rightarrow ('b \times 'c) \text{ set} \Rightarrow ('a \times 'c) \text{ set} \\ r \circ s &= \{(x, z). \exists y. (x, y) \in r \wedge (y, z) \in s\} \end{aligned}$$

Note that $r \circ s$ can be read from left to right: first r , then s .

The denotation of a command is a relation between initial and final states:

$$\text{type_synonym } com_den = (state \times state) \text{ set}$$

Function D maps a command to its denotation. The first four equations should be self-explanatory:

$$\begin{aligned} D &:: com \Rightarrow com_den \\ D \text{ SKIP} &= Id \\ D (x ::= a) &= \{(s, t). t = s(x := aval a s)\} \\ D (c_1;; c_2) &= D c_1 \circ D c_2 \\ D (IF b THEN c_1 ELSE c_2) &= \{(s, t). \text{if } bval b s \text{ then } (s, t) \in D c_1 \text{ else } (s, t) \in D c_2\} \end{aligned}$$

Example 11.1. Let $c_1 = ''x'' ::= N 0$ and $c_2 = ''y'' ::= V ''x''$.

$$\begin{aligned} D c_1 &= \{(s_1, s_2). s_2 = s_1(''x'' := 0)\} \\ D c_2 &= \{(s_2, s_3). s_3 = s_2(''y'' := s_2 ''x'')\} \\ D (c_1;; c_2) &= \{(s_1, s_3). s_3 = s_1(''x'' := 0, ''y'' := 0)\} \end{aligned}$$

The definition of $D w$, where $w = \text{WHILE } b \text{ DO } c$, is trickier. Ideally we would like to write the recursion equation

$$D w = \{(s, t). \text{if } bval b s \text{ then } (s, t) \in D c \circ D w \text{ else } s = t\} \quad (*)$$

but because $D w$ depends on $D w$, this is not in denotational style. Worse, it would not be accepted by Isabelle because it does not terminate and may be

inconsistent: remember the example of the illegal ‘definition’ $f\ n = f\ n + 1$ where subtracting $f\ n$ on both sides would lead to $0 = 1$. The fact remains that $D\ w$ should be a solution of $(*)$, but we need to show that a solution exists and to pick a specific one. This is entirely analogous to the problem we faced when analysing true liveness in [Section 10.4](#), and the solution will be the same: the *lfp* operator and the [Knaster-Tarski Fixpoint Theorem](#).

In general, a solution of an equation $t = f\ t$ such as $(*)$ is the same as a fixpoint of f . In the case of $(*)$, the function f is

$$\begin{aligned} W &:: (state \Rightarrow bool) \Rightarrow com_den \Rightarrow (com_den \Rightarrow com_den) \\ W\ db\ dc &= (\lambda dw. \{(s, t). \text{if } db\ s \text{ then } (s, t) \in dc \circ dw \text{ else } s = t\}) \end{aligned}$$

where we have abstracted the terms $bval\ b$, $D\ c$ and $D\ w$ in $(*)$ by the parameters db , dc and dw . Hence $W\ (bval\ b)\ (D\ c)\ (D\ w)$ is the right-hand side of $(*)$. Now we define $D\ w$ as a least fixpoint:

$$D\ (WHILE\ b\ DO\ c) = lfp\ (W\ (bval\ b)\ (D\ c))$$

Why the least? The formal justification will be an equivalence proof between denotational and big-step semantics. The following example provides some intuition why leastness is what we want.

Example 11.2. Let $w = WHILE\ Bc\ True\ DO\ SKIP$. Then

$$\begin{aligned} f &= W\ (bval\ (Bc\ True))\ (D\ SKIP) = W\ (\lambda s. True)\ Id \\ &= (\lambda dw. \{(s, t). (s, t) \in Id \circ dw\}) = (\lambda dw. dw) \end{aligned}$$

Therefore any relation is a fixpoint of f , but our intuition tells us that because w never terminates, its semantics should be the empty relation, which is precisely the least fixpoint of f .

We still have to prove that $W\ db\ dc$ is monotone. Only then can we appeal to the [Knaster-Tarski Fixpoint Theorem](#) to conclude that $lfp\ (W\ db\ dc)$ really is the least fixpoint.

Lemma 11.3. *$W\ db\ dc$ is monotone.*

Proof. For Isabelle, the proof is automatic. The core of the argument rests on the fact that relation composition is monotone in both arguments: if $r \subseteq r'$ and $s \subseteq s'$, then $r \circ s \subseteq r' \circ s'$, which can be seen easily from the definition of \circ . If $dw \subseteq dw'$ then $W\ db\ dc\ dw \subseteq W\ db\ dc\ dw'$ because $dc \circ dw \subseteq dc \circ dw'$. \square

We could already define D without this monotonicity lemma, but we would not be able to deduce anything about $D\ w$ without it. Now we can derive $(*)$. This is a completely mechanical consequence of the way we defined $D\ w$ and W . Given $t = lfp\ f$ for some monotone f , we obtain

$$t = \text{lfp } f = f (\text{lfp } f) = f t$$

where the step $\text{lfp } f = f (\text{lfp } f)$ is the consequence of the [Knaster-Tarski Fixpoint Theorem](#) because f is monotone. Setting $t = D w$ and $f = W (bval b) (D c)$ in $t = f t$ results in $(*)$ by definition of W .

An immediate consequence is

$$\begin{aligned} D (WHILE b DO c) = \\ D (IF b THEN c;; WHILE b DO c ELSE SKIP) \end{aligned}$$

Just expand the definition of D for IF and $;;$ and you obtain $(*)$. This is an example of the simplicity of deriving program equivalences with the help of denotational semantics.

Discussion

Why can't we just define $(*)$ as it is but have to go through the indirection of lfp and prove monotonicity of W ? None of this was required for the operational semantics! The reason for this discrepancy is that inductive definitions require a fixed format to be admissible. For this fixed format, it is not hard to prove that the inductively defined predicate actually exists. Isabelle does this by converting the inductive definition internally into a function on sets, proving its monotonicity and defining the inductive predicate as the least fixpoint of that function. The monotonicity proof is automatic provided we stick to the fixed format. Once you step outside the format, in particular when using negation, the definition will be rejected by Isabelle because least fixpoints may cease to exist and the inductive definition may be plain contradictory:

$$\begin{aligned} P x &\Longrightarrow \neg P x \\ \neg P x &\Longrightarrow P x \end{aligned}$$

The analogous recursive 'definition' is $P x = (\neg P x)$, which is also rejected by Isabelle, because it does not terminate.

To avoid the manual monotonicity proof required for our denotational semantics one could put together a collection of basic functions that are all monotone or preserve monotonicity. One would end up with a little programming language where all functions are monotone and this could be proved automatically. In fact, one could then even automate the translation of recursion equations like $(*)$ into lfp format. Creating such a programming language is at the heart of denotational semantics, but we do not go into it in our brief introduction to the subject.

In summary: Although our treatment of denotational semantics appears more complicated than operational semantics because of the explicit lfp , operational semantics is defined as a least fixpoint too, but this is hidden inside **inductive**. One can hide the lfp in denotational semantics too and allow direct

recursive definitions. This is what Isabelle's `partial_function` command does [50, 92].

11.1.1 Equivalence of Denotational and Big-Step Semantics

We show that the denotational semantics is logically equivalent with our gold standard, the big-step semantics. The equivalence is proved as two separate lemmas. Both proofs are almost automatic because the denotational semantics is relational and thus close to the operational one. Even the treatment of *WHILE* is the same: $D\ w$ is defined explicitly as a least fixpoint and the operational semantics is an inductive definition which is internally defined as a least fixpoint (see the Discussion above).

Lemma 11.4. $(c, s) \Rightarrow t \implies (s, t) \in D\ c$

Proof. By rule induction. All cases are automatic. We just look at *WhileTrue* where we may assume $bval\ b\ s_1$ and the IHs $(s_1, s_2) \in D\ c$ and $(s_2, s_3) \in D\ (WHILE\ b\ DO\ c)$. We have to show $(s_1, s_3) \in D\ (WHILE\ b\ DO\ c)$, which follows immediately from $(*)$. \square

The other direction is expressed by means of the abbreviation *Big_step* introduced at the beginning of this chapter. The reason is purely technical.

Lemma 11.5. $(s, t) \in D\ c \implies (s, t) \in Big_step\ c$

Proof. By induction on c . All cases are proved automatically except $w = WHILE\ b\ DO\ c$, which we look at in detail. Let $B = Big_step\ w$ and $f = W\ (bval\ b)\ (D\ c)$. By definition of W and the big-step \Rightarrow it follows that B is a pre-fixpoint of f , i.e., $f\ B \subseteq B$: given $(s, t) \in f\ B$, either $bval\ b\ s$ and there is some s' such that $(s, s') \in D\ c$ (hence $(c, s) \Rightarrow s'$ by IH) and $(w, s') \Rightarrow t$, or $\neg bval\ b\ s$ and $s = t$; in either case $(w, s) \Rightarrow t$, i.e., $(s, t) \in B$. Because $D\ w$ is the least fixpoint and also the least pre-fixpoint of f (see Knaster-Tarski), $D\ w \subseteq B$ and hence $(s, t) \in D\ w \implies (s, t) \in B$ as claimed. \square

The combination of the previous two lemma yields the equivalence:

Theorem 11.6 (Equivalence of denotational and big-step semantics).

$$(s, t) \in D\ c \iff (c, s) \Rightarrow t$$

As a nice corollary we obtain that the program equivalence \sim defined in Section 7.2.4 is the same as denotational equality: if you replace $(c_i, s) \Rightarrow t$ in the definition of $c_1 \sim c_2$ by $(s, t) \in D\ c_i$ this yields $\forall s\ t. (s, t) \in D\ c_1 \iff (s, t) \in D\ c_2$, which is equivalent with $D\ c_1 = D\ c_2$ because two sets are equal iff they contain the same elements.

Corollary 11.7. $c_1 \sim c_2 \iff D\ c_1 = D\ c_2$

11.1.2 Continuity

Denotational semantics is usually not based on relations but on (partial) functions. The difficulty with functions is that the Knaster-Tarski Fixpoint Theorem does not apply. Abstractly speaking, functions do not form a complete lattice (see [Chapter 13](#)) because the union of two functions (as relations) is not necessarily again a function. Thus one needs some other means of obtaining least fixpoints. It turns out that the functions one is interested in satisfy a stronger property called **continuity** which guarantees least fixpoints also in the space of partial functions. Although we do not need to worry about existence of fixpoints in our setting, the notion of continuity is interesting in its own right because it allows a much more intuitive characterisation of least fixpoints than Knaster-Tarski. This in turn gives rise to a new induction principle for least fixpoints.

Definition 11.8. *A chain is a sequence of sets $S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots$:*

$$\begin{aligned} \text{chain} &:: (\text{nat} \Rightarrow 'a \text{ set}) \Rightarrow \text{bool} \\ \text{chain } S &= (\forall i. S\ i \subseteq S\ (\text{Suc } i)) \end{aligned}$$

Our specific kinds of chains are often called ω -chains.

In the following we make use of the notation $\bigcup_n A\ n$ for the union of all sets $A\ n$ (see [Section 4.2](#)). If $n :: \text{nat}$ this is $A\ 0 \cup A\ 1 \cup \dots$

Definition 11.9. *A function f on sets is called **continuous** if it commutes with \bigcup for all chains:*

$$\begin{aligned} \text{cont} &:: ('a \text{ set} \Rightarrow 'b \text{ set}) \Rightarrow \text{bool} \\ \text{cont } f &= (\forall S. \text{chain } S \longrightarrow f\ (\bigcup_n S\ n) = (\bigcup_n f\ (S\ n))) \end{aligned}$$

That is, $f(S_0 \cup S_1 \cup \dots) = f\ S_0 \cup f\ S_1 \cup \dots$ for all chains S .

To understand why these notions are relevant for us, think in terms of relations between states, or, for simplicity, input and output of some computation. For example, the input-output behaviour of a function *sum* that sums up the first n numbers can be expressed as this infinite relation $\text{Sum} = \{(0,0), (1,1), (2,3), (3,6), (4,10), \dots\}$ on *nat*. We can compute this relation gradually by starting from $\{\}$ and adding more pairs in each step: $\{\} \subseteq \{(0,0)\} \subseteq \{(0,0), (1,1)\} \subseteq \dots$. This is why chains are relevant. Each element S_n in the chain is only a finite approximation of the full semantics of the summation function which is the infinite set $\bigcup_n S\ n$.

To understand the computational meaning of monotonicity, consider a second summation function sum_2 with semantics $\text{Sum}_2 = \{(0, 0), (1, 1), (2, 3)\}$, i.e., sum_2 behaves like *sum* for inputs ≤ 2 and does not terminate otherwise. Let $P[\cdot]$ be a program where we can plug in different subcomponents.

Monotonicity of the semantics of P means that because $Sum_2 \subseteq Sum$, any result that $P[sum_2]$ can deliver, $P[sum]$ can deliver too. This makes computational sense: if $P[sum_2]$ delivers an output, it can only have called sum_2 with arguments ≤ 2 (otherwise the call and thus the whole computation would not have terminated), in which case $P[sum]$ can follow the same computation path in P and deliver the same result.

But why continuity? We give an example where non-computability means non-continuity. Consider the non-computable function $T :: (nat \times nat) \rightarrow bool$ whose output tells us if its input is (the semantics of) an everywhere terminating computation:

$$T\ r = (if\ \forall m. \exists n. (m, n) \in r\ then\ \{True\}\ else\ \{\})$$

This function is monotone but not continuous. For example, let S be the chain of finite approximations of Sum above. Then $T(\bigcup_n S\ n) = \{True\}$ but $(\bigcup_n T(S\ n)) = \{\}$. Going back to the $P[\cdot]$ scenario above, continuity means that a terminating computation of $P[f]$ should only need a finite part of the semantics of f , which means it can only run f for a finite amount of time. Again, this makes computational sense.

Now we study chains and continuity formally. It is obvious (and provable by induction) that in a chain S we have $i \leq j \implies S\ i \subseteq S\ j$. But because \leq is total on nat ($i \leq j \vee j \leq i$), \subseteq must be total on S too:

Lemma 11.10. *chain $S \implies S\ i \subseteq S\ j \vee S\ j \subseteq S\ i$*

Continuity implies monotonicity:

Lemma 11.11. *Every continuous function is monotone.*

Proof. Let f be continuous, assume $A \subseteq B$ and consider the chain $A \subseteq B \subseteq B \subseteq \dots$, i.e., $S = (\lambda i. if\ i = 0\ then\ A\ else\ B)$. Then $f\ B = f(\bigcup_n S\ n) = (\bigcup_n f(S\ n)) = f\ A \cup f\ B$ and hence $f\ A \subseteq f\ B$. \square

Our main theorem about continuous functions is that their least fixpoints can be obtained by iteration starting from $\{\}$.

Theorem 11.12 (Kleene fixpoint theorem).

$$cont\ f \implies lfp\ f = (\bigcup_n f^n\ \{\})$$

Proof. Let $U = (\bigcup_n f^n\ \{\})$. First we show $lfp\ f \subseteq U$, then $U \subseteq lfp\ f$.

Because f is continuous, it is also monotone. Hence $lfp\ f \subseteq U$ follows by [Knaster-Tarski](#) if U is a fixpoint of f . Therefore we prove $f\ U = U$. Observe that by [Lemma 10.32](#) the $f^n\ \{\}$ form a chain.

$$\begin{aligned} f\ U &= \bigcup_n f^{n+1}\ \{\} && \text{by continuity} \\ &= f^0\ \{\} \cup (\bigcup_n f^{n+1}\ \{\}) && \text{because } f^0\ \{\} = \{\} \\ &= U \end{aligned}$$

For the opposite direction $U \subseteq \text{lfp } f$ it suffices to prove $f^n \{\} \subseteq \text{lfp } f$. The proof is by induction on n . The base case is trivial. Assuming $f^n \{\} \subseteq \text{lfp } f$ we have by monotonicity of f and Knaster-Tarski that $f^{n+1} \{\} = f (f^n \{\}) \subseteq f (\text{lfp } f) = \text{lfp } f$. \square

This is a generalization of Lemma 10.32 that allowed us to compute lfp in the true liveness analysis in Section 10.4.2. At the time we could expect the chain of iterates of f to stabilize, now we have to go to the limit.

The Kleene fixpoint theorem is applicable to W :

Lemma 11.13. *Function $W \text{ b } r$ is continuous.*

Proof. Although the Isabelle proof is automatic, we explain the details because they may not be obvious. Let $R :: \text{nat} \Rightarrow \text{com_den}$ be any sequence of state relations — it does not even need to be a chain. We show that $(s, t) \in W \text{ b } r (\bigcup_n R \ n)$ iff $(s, t) \in (\bigcup_n W \text{ b } r (R \ n))$. If $\neg b \ s$ then (s, t) is in both sets iff $s = t$. Now assume $b \ s$. Then

$$\begin{aligned} (s, t) \in W \text{ b } r (\bigcup_n R \ n) &\longleftrightarrow (s, t) \in r \circ (\bigcup_n R \ n) \\ &\longleftrightarrow \exists s'. (s, s') \in r \wedge (s', t) \in (\bigcup_n R \ n) \\ &\longleftrightarrow \exists s' \ n. (s, s') \in r \wedge (s', t) \in R \ n \\ &\longleftrightarrow \exists n. (s, t) \in r \circ R \ n \\ &\longleftrightarrow (s, t) \in (\bigcup_n r \circ R \ n) \\ &\longleftrightarrow (s, t) \in (\bigcup_n W \text{ b } r (R \ n)) \end{aligned}$$

Warning: such \longleftrightarrow chains are an abuse of notation: $A \longleftrightarrow B \longleftrightarrow C$ really means the logically not equivalent $(A \longleftrightarrow B) \wedge (B \longleftrightarrow C)$ which implies $A \longleftrightarrow C$. \square

Example 11.14. In this example we show concretely how iterating W creates a chain of relations that approximate the semantics of some loop and whose union is the full semantics. The loop is

WHILE b *DO* c
 where $b = \text{Not } (Eq \ (V \ 'x') \ (N \ 0))$
 and $c = 'x' ::= Plus \ (V \ 'x') \ (N \ (-1))$.

Function Eq compares its arguments for equality (see Exercise 3.7). Intuitively, the semantics of this loop is the following relation:

$$S = \{(s, t). 0 \leq s \ 'x' \wedge t = s('x' := 0)\}$$

Function f expresses the semantics of one loop iteration:

$f = W \text{ db } dc$
 where $db = bval \ b = (\lambda s. s \ 'x' \neq 0)$
 and $dc = D \ c = \{(s, t). t = s('x' := s \ 'x' - 1)\}$.

This is what happens when we iterate f starting from $\{\}$:

$$\begin{aligned}
f^1 \{\} &= \{(s, t). \text{if } s \text{ ''}x'' \neq 0 \text{ then } (s, t) \in dc \circ \{\} \text{ else } s = t\} \\
&= \{(s, t). s \text{ ''}x'' = 0 \wedge s = t\} \\
f^2 \{\} &= \{(s, t). \text{if } s \text{ ''}x'' \neq 0 \text{ then } (s, t) \in dc \circ f \{\} \text{ else } s = t\} \\
&= \{(s, t). 0 \leq s \text{ ''}x'' \wedge s \text{ ''}x'' < 2 \wedge t = s(\text{''}x'' := 0)\} \\
f^3 \{\} &= \{(s, t). \text{if } s \text{ ''}x'' \neq 0 \text{ then } (s, t) \in dc \circ f^2 \{\} \text{ else } s = t\} \\
&= \{(s, t). 0 \leq s \text{ ''}x'' \wedge s \text{ ''}x'' < 3 \wedge t = s(\text{''}x'' := 0)\}
\end{aligned}$$

Note that the first equality $f^n \{\} = R$ follows by definition of W whereas the second equality $R = R'$ requires a few case distinctions. The advantage of the form R' is that we can see a pattern emerging:

$$f^n \{\} = \{(s, t). 0 \leq s \text{ ''}x'' \wedge s \text{ ''}x'' < \text{int } n \wedge t = s(\text{''}x'' := 0)\}$$

where function int coerces a nat to an int . This formulation shows clearly that $f^n \{\}$ is the semantics of our loop restricted to at most n iterations. The $f^n \{\}$ form a chain and the union is the full semantics: $(\bigcup_n f^n \{\}) = S$. The latter equality follows because the condition $s \text{ ''}x'' < \text{int } n$ in $\bigcup_n f^n \{\}$ is satisfied by all large enough n and hence $\bigcup_n f^n \{\}$ collapses to S .

As an application of the Kleene fixpoint theorem we show that our denotational semantics is deterministic, i.e., the command denotations returned by D are **single-valued**, a predefined notion:

$$\text{single_valued } r = (\forall x \ y \ z. (x, y) \in r \wedge (x, z) \in r \longrightarrow y = z)$$

The only difficult part of the proof is the *WHILE* case. Here we can now argue that if W preserves single-valuedness, then its least fixpoint is single-valued because it is a union of a chain of single-valued relations:

Lemma 11.15. *If $f :: \text{com_den} \Rightarrow \text{com_den}$ is continuous and preserves single-valuedness then $\text{lfp } f$ is single-valued.*

Proof. Because f is continuous we have $\text{lfp } f = (\bigcup_n f^n \{\})$. By [Lemma 10.32](#) (because f is also monotone) the $f^n \{\}$ form a chain. All its elements are single-valued because $\{\}$ is single-valued and f preserves single-valuedness. The union of a chain of single-valued relations is obviously single-valued too. \square

Lemma 11.16. *$\text{single_valued } (D \ c)$*

Proof. A straightforward induction on c . The *WHILE* case follows by the previous lemma because $W \ b \ r$ is continuous and preserves single-valuedness (as can be checked easily, assuming by IH that r is single-valued). \square

11.2 Summary and Further Reading

A denotational semantics is a compositional mapping from syntax to meaning: the meaning of a compound construct is a function of the meanings of its subconstructs. The meaning of iterative or recursive constructs is given as a least fixpoint. In a relational context, the existence of a least fixpoint can be guaranteed by monotonicity via the Knaster-Tarski Fixpoint Theorem. For computational reasons the semantics should be not just monotone but also continuous, in which case the least fixpoint is the union of all finite iterates $f^n \{\}$. Thus we can reason about least fixpoints of continuous functions by induction on n .

Denotational semantics has its roots in the work of Dana Scott and Christopher Strachey [82, 83]. This developed into a rich and mathematically sophisticated theory. We have only presented a simplified set-theoretic version of the foundations. For the real deal the reader is encouraged to consult textbooks [39, 79] and handbook articles [1, 86] dedicated to denotational semantics. Some of the foundations of denotational semantics have been formalized in theorem provers [11, 43, 59].

Exercises

Exercise 11.1. Building on [Exercise 7.8](#), extend the denotational semantics and the equivalence proof with the big-step semantics with a *REPEAT* loop.

Exercise 11.2. Consider [Example 11.14](#) and prove by induction on n that $f^n \{\} = \{(s, t). 0 \leq s \text{ ''}x'' \wedge s \text{ ''}x'' < \text{int } n \wedge t = s(\text{''}x'' := 0)\}$.

Exercise 11.3. Consider [Example 11.14](#) but with the loop condition $b = \text{Less } (N\ 0) \ (V \text{ ''}x'')$. Find a closed expression M (containing n) for $f^n \{\}$ and prove $f^n \{\} = M$.

Exercise 11.4. Define an operator B such that you can express the equation for $D \ (IF\ b\ THEN\ c_1\ ELSE\ c_2)$ in a point-free way. In this context, we call a definition **point free** if it does not mention the state on the left-hand side. For example:

$$D \ (IF\ b\ THEN\ c_1\ ELSE\ c_2) = B\ b\ O\ D\ c_1 \cup B\ (Not\ b)\ O\ D\ c_2$$

A point-wise definition would start

$$D \ (IF\ b\ THEN\ c_1\ ELSE\ c_2)\ s = \dots$$

Similarly, find a point-free equation for $W \ (bval\ b)\ dc$ and use it to write down a point-free version of $D \ (WHILE\ b\ DO\ c)$ (still using *lfp*). Prove that your two equations are equivalent to the old ones.

Exercise 11.5. Let the ‘thin’ part of a relation be its single-valued subset:

$$\text{thin } R = \{(a, b). (a, b) \in R \wedge (\forall c. (a, c) \in R \longrightarrow c = b)\}$$

Prove that if $f :: ('a * 'a) \text{ set} \Rightarrow ('a * 'a) \text{ set}$ is monotone and for all R , $f (\text{thin } R) \subseteq \text{thin } (f R)$, then $\text{single_valued } (\text{lfp } f)$.

Exercise 11.6. Generalize our set-theoretic treatment of continuity and least fixpoints to **chain-complete partial orders** (cpos), i.e., partial orders \leq that have a least element \perp and where every chain $c_0 \leq c_1 \leq \dots$ has a least upper bound $\text{lub } c$ where $c :: \text{nat} \Rightarrow 'a$. A function $f :: 'a \Rightarrow 'b$ between two cpos $'a$ and $'b$ is **continuous** if $f (\text{lub } c) = \text{lub } (f \circ c)$ for all chains c . Prove that if f is monotone and continuous then $\text{lub } (\lambda n. (f \circ \circ n) \perp)$ is the least fixpoint of f .

Exercise 11.7. We define a dependency analysis Dep that maps commands to relations between variables such that $(x, y) \in \text{Dep } c$ means that in the execution of c the initial value of x can influence the final value of y :

```
fun Dep :: com => (vname * vname) set where
  Dep SKIP = Id |
  Dep (x ::= a) = {(u, v). if v = x then u ∈ vars a else u = v} |
  Dep (c1 ;; c2) = Dep c1 O Dep c2 |
  Dep (IF b THEN c1 ELSE c2) = Dep c1 ∪ Dep c2 ∪ vars b × UNIV |
  Dep (WHILE b DO c) = lfp(λR. Id ∪ vars b × UNIV ∪ Dep c O R)
```

where \times is the cross product of two sets. Prove monotonicity of the function lfp is applied to.

For the correctness statement define

```
abbreviation Deps :: com => vname set => vname set where
  Deps c X ≡ (⋃ x ∈ X. {y. (y, x) : Dep c})
```

and prove

lemma $\llbracket (c, s) \Rightarrow s'; (c, t) \Rightarrow t'; s = t \text{ on } \text{Deps } c \ X \rrbracket \Longrightarrow s' = t' \text{ on } X$

Give an example that the following stronger termination-sensitive property

$\llbracket (c, s) \Rightarrow s'; s = t \text{ on } \text{Deps } c \ X \rrbracket \Longrightarrow \exists t'. (c, t) \Rightarrow t' \wedge s' = t' \text{ on } X$

does not hold. Hint: $X = \{\}$.

In the definition of $\text{Dep } (\text{IF } b \text{ THEN } c_1 \text{ ELSE } c_2)$ the variables in b can influence all variables (UNIV). However, if a variable is not assigned to in c_1 and c_2 , it is not influenced by b (ignoring termination). Theory Vars defines a function lvars such that $\text{lvars } c$ is the set of variables on the left-hand side of an assignment in c . Modify the definition of Dep as follows: replace UNIV

by $lvars\ c_1 \cup lvars\ c_2$ (in the case *IF* b *THEN* c_1 *ELSE* c_2) and by $lvars\ c$ (in the case *WHILE* b *DO* c). Adjust the proof of the above correctness statement.

Hoare Logic

So far we have proved properties of IMP, like type soundness, or properties of tools for IMP, like compiler correctness, but almost never properties of individual IMP programs. The Isabelle part of the book has taught us how to prove properties of functional programs, but not of imperative ones.

Hoare logic (due to Tony Hoare), also known as **axiomatic semantics**, is a logic for proving properties of imperative programs. The formulas of Hoare logic are so-called **Hoare triples**

$$\{P\} c \{Q\}$$

which should be read as saying that if formula P is true before the execution of command c then formula Q is true after the execution of c . This is a simple example of a Hoare triple:

$$\{x = y\} y := y+1 \{x < y\}$$

12.1 Proof via Operational Semantics [thy](#)

Before introducing the details of Hoare logic we show that in principle we can prove properties of programs via their operational semantics. Hoare logic can be viewed as the structured essence of such proofs.

As an example, we prove that the program

```
y := 0;
WHILE 0 < x DO (y := y+x; x := x-1)
```

sums up the numbers 1 to x in y . Formally let

```
wsum = WHILE Less (N 0) (V ''x'') DO csum

csum = ''y'' ::= Plus (V ''y'') (V ''x'');;
      ''x'' ::= Plus (V ''x'') (N (-1))
```

The summation property can be expressed as a theorem about the program's big-step semantics:

$$(''y'' ::= N\ 0;;\ wsum,\ s) \Rightarrow t \implies t\ ''y'' = \text{sum}\ (s\ ''x'') \quad (*)$$

where

```
fun sum :: int ⇒ int where
  sum i = (if i ≤ 0 then 0 else sum (i - 1) + i)
```

We prove $(*)$ in two steps. First we show that $wsum$ does the right thing. This will be an induction, and as usual we have to generalize what we want to prove from the special situation where y is 0.

$$(wsum,\ s) \Rightarrow t \implies t\ ''y'' = s\ ''y'' + \text{sum}\ (s\ ''x'') \quad (**)$$

This is proved by an induction on the premise. There are two cases.

If the loop condition is false, i.e., if $s\ ''x'' \leq 0$, then we have to prove $s\ ''y'' = s\ ''y'' + \text{sum}\ (s\ ''x'')$, which follows because $s\ ''x'' < 1$ implies $\text{sum}\ (s\ ''x'') = 0$.

If the loop condition is true, i.e., if $0 < s\ ''x''$, then we may assume $(csum,\ s) \Rightarrow u$ and the IH $t\ ''y'' = u\ ''y'' + \text{sum}\ (u\ ''x'')$ and we have to prove the conclusion of $(**)$. From $(csum,\ s) \Rightarrow u$ it follows by inversion of the rules for $;;$ and $::=$ (Section 7.2.3) that $u = s(''y'' := s\ ''y'' + s\ ''x'', ''x'' := s\ ''x'' - 1)$. Substituting this into the IH yields $t\ ''y'' = s\ ''y'' + s\ ''x'' + \text{sum}\ (s\ ''x'' - 1)$. This is equivalent with the conclusion of $(**)$ because $0 < s\ ''x''$.

Having proved $(**)$, $(*)$ follows easily: From $(''y'' ::= N\ 0;;\ wsum,\ s) \Rightarrow t$ it follows by rule inversion that after the assignment the intermediate state must have been $s(''y'' := 0)$ and therefore $(wsum,\ s(''y'' := 0)) \Rightarrow t$. Now $(**)$ implies $t\ ''y'' = \text{sum}\ (s\ ''x'')$, thus concluding the proof of $(*)$.

Hoare logic can be viewed as the structured essence of such operational proofs. The rules of Hoare logic are (almost) syntax-directed and automate all those aspects of the proof that are concerned with program execution. However, there is no free lunch: you still need to be creative to find generalizations of formulas when proving properties of loops, and proofs about arithmetic formulas are still up to you (and Isabelle).

We will now move on to the actual topic of this chapter, Hoare logic.

12.2 Hoare Logic for Partial Correctness

The formulas of Hoare logic are the Hoare triples $\{P\} c \{Q\}$, where P is called the **precondition** and Q the **postcondition**. We call $\{P\} c \{Q\}$ **valid** if the following implication holds:

If P is true before the execution of c and c terminates
then Q is true afterwards.

Validity is defined in terms of execution, i.e., the operational semantics, our ultimate point of reference. This notion of validity is called **partial correctness** because the postcondition is only required to be true if c terminates. There is also the concept of **total correctness**:

If P is true before the execution of c
then c terminates and Q is true afterwards.

In a nutshell:

Total correctness = partial correctness + termination

Except for the final section of this chapter, we will always work with the partial correctness interpretation, because it is easier.

Pre- and postconditions come from some set of logical formulas that we call **assertions**. There are two approaches to the language of assertions:

Syntactic: Assertions are concrete syntactic objects, like type *expr*.

Semantic: Assertions are predicates on states, i.e., of type *state* \Rightarrow *bool*.

We follow the syntactic approach in the introductory subsection, because it is nicely intuitive and one can sweep some complications under the carpet. But for proofs about Hoare logic, the semantic approach is much simpler and we switch to it in the rest of the chapter.

12.2.1 Syntactic Assertions

Assertions are ordinary logical formulas and include all the boolean expressions of IMP. We are deliberately vague because the exact nature of assertions is not important for understanding how Hoare logic works. Just as we did for the simplified notation for IMP, we write concrete assertions and Hoare triples in typewriter font. Here are some examples of valid Hoare triples:

$\{x = 5\} \ x := x+5 \ \{x = 10\}$

$\{\text{True}\} \ x := 10 \ \{x = 10\}$

$\{x = y\} \ x := x+1 \ \{x \neq y\}$

Note that the precondition *True* is always true; hence the second triple merely says that from whatever state you start, after $x := 10$ the postcondition $x = 10$ is true.

More interesting are the following somewhat extreme examples:

$\{\text{True}\} \ c_1 \ \{\text{True}\}$

$$\{\text{True}\} \ c_2 \ \{\text{False}\}$$

$$\{\text{False}\} \ c_3 \ \{Q\}$$

Which c_i make these triples valid? Think about it before you read on. Remember that we work with partial correctness in this section. Therefore every c_1 works because the postcondition True is always true. In the second triple, c_2 must not terminate, otherwise False would have to be true, which it certainly is not. In the final triple, any c_3 and Q work because the meaning of the triple is that “if False is true . . .”, but False is not true. Note that for the first two triples, the answer is different under a total correctness interpretation.

Proof System

So far we have spoken of Hoare triples being valid. Now we will present a set of inference rules or **proof system** for deriving Hoare triples. This is a new mechanism and here we speak of Hoare triples being **derivable**. Of course being valid and being derivable should have something to do with each other. When we look at the proof rules in a moment they will all feel very natural (well, except for one) precisely because they follow our informal understanding of when a triple is valid. Nevertheless it is essential not to confuse the notions of validity (which employs the operational semantics) and derivability (which employs an independent set of proof rules).

The proof rules for Hoare logic are shown in [Figure 12.1](#). We go through them one by one.

The *SKIP* rule is obvious, but the assignment rule needs some explanation. It uses the substitution notation

$$P[a/x] \equiv P \text{ with } a \text{ substituted for } x.$$

For example, $(x = 5)[5/x]$ is $5 = 5$ and $(x = x)[5+x/x]$ is $5+x = 5+x$. The latter example shows that all occurrences of x in P are simultaneously replaced by a , but that this happens only once: if x occurs in a , those occurrences are not replaced, otherwise the substitution process would go on forever. Here are some instances of the assignment rule:

$$\begin{array}{ll} \{5 = 5\} \ x := 5 & \{x = 5\} \\ \{x+5 = 5\} \ x := x+5 & \{x = 5\} \\ \{2*(x+5) > 20\} \ x := 2*(x+5) & \{x > 20\} \end{array}$$

Simplifying the preconditions that were obtained by blind substitution yields the more readable triples

$$\begin{array}{ll} \{\text{True}\} \ x := 5 & \{x = 5\} \\ \{x = 0\} \ x := x+5 & \{x = 5\} \\ \{x > 5\} \ x := 2*(x+5) & \{x > 20\} \end{array}$$

$$\boxed{
\begin{array}{c}
\overline{\{P\} \text{ SKIP } \{P\}} \\
\\
\overline{\{P[a/x]\} x ::= a \{P\}} \\
\\
\frac{\{P_1\} c_1 \{P_2\} \quad \{P_2\} c_2 \{P_3\}}{\{P_1\} c_1;; c_2 \{P_3\}} \\
\\
\frac{\{P \wedge b\} c_1 \{Q\} \quad \{P \wedge \neg b\} c_2 \{Q\}}{\{P\} \text{ IF } b \text{ THEN } c_1 \text{ ELSE } c_2 \{Q\}} \\
\\
\frac{\{P \wedge b\} c \{P\}}{\{P\} \text{ WHILE } b \text{ DO } c \{P \wedge \neg b\}} \\
\\
\frac{P' \longrightarrow P \quad \{P\} c \{Q\} \quad Q \longrightarrow Q'}{\{P'\} c \{Q'\}}
\end{array}
}$$

Fig. 12.1. Hoare logic for partial correctness (syntactic assertions)

The assignment rule may still puzzle you because it seems to go in the wrong direction by modifying the precondition rather than the postcondition. After all, the operational semantics modifies the post-state, not the pre-state. Correctness of the assignment rule can be explained as follows: if initially $P[a/x]$ is true, then after the assignment x will have the value of a , and hence no substitution is necessary anymore, i.e., P itself is true afterwards. A forward version of this rule exists but is more complicated.

The $;;$ rule strongly resembles its big-step counterpart. Reading it backward it decomposes the proof of $c_1;; c_2$ into two proofs involving c_1 and c_2 and a new intermediate assertion P_2 .

The *IF* rule is pretty obvious: you need to prove that both branches lead from P to Q , where in each proof the appropriate b or $\neg b$ can be conjoined to P . That is, each sub-proof additionally assumes the branch condition.

Now we consider the *WHILE* rule. Its premise says that if P and b are true before the execution of the loop body c , then P is true again afterwards (if the body terminates). Such a P is called an **invariant** of the loop: if you start in a state where P is true then no matter how often the loop body is iterated, as long as b is true before each iteration, P stays true too. This explains the conclusion: if P is true initially, then it must be true at the end because it is invariant. Moreover, if the loop terminates, then $\neg b$ must be

true too. Hence $P \wedge \neg b$ at the end (if we get there). The *WHILE* rule can be viewed as an induction rule where the invariance proof is the step.

The final rule in Figure 12.1 is called the **consequence rule**. It is independent of any particular IMP construct. Its purpose is to adjust the precondition and postcondition. Going from $\{P\} c \{Q\}$ to $\{P'\} c \{Q'\}$ under the given premises permits us to

- strengthen the precondition: $P' \longrightarrow P$
- weaken the postcondition: $Q \longrightarrow Q'$

where A is called stronger than B if $A \longrightarrow B$. For example, from $\{x \geq 0\} c \{x \geq 1\}$ we can prove $\{x = 5\} c \{x \geq 0\}$. The latter is strictly weaker than the former because it tells us less about the behaviour of c . Note that the consequence rule is the only rule where some premises are not Hoare triples but assertions. We do not have a formal proof system for assertions and rely on our informal understanding of their meaning to check, for example, that $x = 5 \longrightarrow x \geq 0$. This informality will be overcome once we consider assertions as predicates on states.

This completes the discussion of the basic proof rules. Although these rules are sufficient for all proofs, i.e., the system is complete (which we show later), the rules for *SKIP*, $::=$ and *WHILE* are inconvenient: they can only be applied backwards if the pre- or postcondition are of a special form. For example, for *SKIP* they need to be identical. Therefore we derive new rules for those constructs that can be applied backwards irrespective of the pre- and postcondition of the given triple. The new rules are shown in Figure 12.2. They are easily derived by combining the old rules with the consequence rule.

$$\boxed{
 \begin{array}{c}
 \frac{P \longrightarrow Q}{\{P\} \text{ SKIP } \{Q\}} \\
 \\
 \frac{P \longrightarrow Q[a/x]}{\{P\} x ::= a \{Q\}} \\
 \\
 \frac{\{P \wedge b\} c \{P\} \quad P \wedge \neg b \longrightarrow Q}{\{P\} \text{ WHILE } b \text{ DO } c \{Q\}}
 \end{array}
 }$$

Fig. 12.2. Derived rules (syntactic assertions)

Here is one of the derivations:

$$\frac{P \longrightarrow Q[a/x] \quad \overline{\{Q[a/x]\} x ::= a \{Q\}} \quad \overline{Q \longrightarrow Q}}{\{P\} x ::= a \{Q\}}$$

Two of the three premises are overlined because they have been proved, namely with the original assignment rule and with the trivial logical fact that anything implies itself.

Examples

We return to the summation program from [Section 12.1](#). This time we prove it correct by means of Hoare logic rather than operational semantics. In Hoare logic, we want to prove the triple

$$\{x = i\} \text{ y := 0; wsum } \{y = \text{sum } i\}$$

We cannot write $y = \text{sum } x$ in the postcondition because x is 0 at that point. Unfortunately the postcondition cannot refer directly to the initial state. Instead, the precondition $x = i$ allows us to refer to the unchanged i and therefore to the initial value of x in the postcondition. This is a general trick for remembering values of variables that are modified.

The central part of the proof is to find and prove the invariant I of the loop. Note that we have three constraints that must be satisfied:

1. It should be an invariant: $\{I \wedge 0 < x\} \text{ csum } \{I\}$
2. It should imply the postcondition: $I \wedge \neg 0 < x \longrightarrow y = \text{sum } i$
3. The invariant should be true initially: $x = i \wedge y = 0 \longrightarrow I$

In fact, this is a general design principle for invariants. As usual, it is a case of generalizing the desired postcondition. During the iteration, $y = \text{sum } i$ is not quite true yet because the first x numbers are still missing from y . Hence we try the following assertion:

$$I = (y + \text{sum } x = \text{sum } i)$$

It is easy to check that the constraints 2 and 3 are true. Moreover, I is indeed invariant as the following proof tree shows:

$$\frac{\frac{I \wedge 0 < x \longrightarrow I[x-1/x] [y+x/y]}{\{I \wedge 0 < x\} \text{ y := y+x } \{I[x-1/x]\}} \quad \frac{}{\{I[x-1/x]\} \text{ x := x-1 } \{I\}}}{\{I \wedge 0 < x\} \text{ csum } \{I\}}$$

Although we have not given the proof rules names, it is easy to see at any point in the proof tree which one is used. In the above tree, the left assignment is proved with the derived rule, the right assignment with the basic rule.

In case you are wondering why $I \wedge 0 < x \longrightarrow I[x-1/x] [y+x/y]$ is true, expand the definition of I and carry out the substitutions and you arrive at the following easy arithmetic truth:

$$y + \text{sum } x = \text{sum } i \wedge 0 < x \longrightarrow y + x + \text{sum}(x-1) = \text{sum } i \quad (12.1)$$

With the loop rule and some more arithmetic we derive

$$\frac{\{I \wedge 0 < x\} \text{ csum } \{I\} \quad \overline{I \wedge \neg 0 < x \longrightarrow y = \text{sum } i}}{\{I\} \text{ wsum } \{y = \text{sum } i\}}$$

Now we only need to connect this result with the initialization to obtain the correctness proof for $y := 0; \text{ wsum}$:

$$\frac{\frac{\overline{x = i \longrightarrow I[0/y]}}{\{x = i\} y := 0 \{I\}} \quad \{I\} \text{ wsum } \{y = \text{sum } i\}}{\{x = i\} y := 0; \text{ wsum } \{y = \text{sum } i\}}$$

The summation program is special in that it always terminates. Hence it does not demonstrate that the proof system can prove anything about nonterminating computations, as in the following example:

$$\frac{\overline{\{True\} \text{ SKIP } \{True\}} \quad \overline{True \wedge \neg True \longrightarrow Q}}{\{True\} \text{ WHILE } True \text{ DO SKIP } \{Q\}}$$

We have proved that if the loop terminates, any assertion Q is true. It sounds like magic but is merely the consequence of nontermination. The proof is straightforward: the invariant is $True$ and Q is trivially implied by the negation of the loop condition.

As a final example consider swapping two variables:

$$\{P\} h := x; x := y; y := h \{Q\}$$

where $P = (x = a \wedge y = b)$ and $Q = (x = b \wedge y = a)$. Drawing the full proof tree for this triple is tedious and unnecessary. A compact form of the tree can be given by annotating the intermediate program points with the correct assertions:

$$\{P\} h := x; \{Q[h/y][y/x]\} x := y; \{Q[h/y]\} y := h \{Q\}$$

Both $Q[h/y]$ and $Q[h/y][y/x]$ are simply the result of the basic assignment rule. All that is left to check is the first assignment with the derived assignment rule, i.e., check $P \longrightarrow Q[h/y][y/x][x/h]$. This is true because $Q[h/y][y/x][x/h] = (y = b \wedge x = a)$.

It should be clear that this proof procedure works for any sequence of assignments, thus reducing the proof to pulling back the postcondition (which is completely mechanical) and checking an implication.

The Method

If we look at the proof rules and the examples it becomes apparent that there is a method in this madness: the backward construction of Hoare logic proofs is partly mechanical. Here are the key points:

- We only need the original rules for `;;` and `IF` together with the derived rules for `SKIP`, `::=` and `WHILE`. This is a syntax-directed proof system and each backward rule application creates new subgoals for the subcommands. Thus the shape of the proof tree exactly mirrors the shape of the command in the Hoare triple we want to prove. The construction of the skeleton of this proof tree is completely automatic.
The consequence rule is built into the derived rules and is not required anymore. This is crucial: the consequence rule destroys syntax-directedness because it can be applied at any point.
- When applying the `;;` rule backwards we need to provide the intermediate assertion P_2 that occurs in the premises but not the conclusion. It turns out that we can compute P_2 by pulling the final assertion P_3 back through c_2 . The variable swapping example illustrates this principle.
- There are two aspects that cannot be fully automated (or program verification would be completely automatic, which is impossible): invariants must be supplied explicitly, and the implications between assertions in the premises of the derived rules must be proved somehow.

In a nutshell, Hoare logic can be reduced to finding invariants and proving assertions. We will carry out this program in full detail in [Section 12.4](#). But first we need to formalize our informal notion of assertions.

12.2.2 Assertions as Functions thy

Our introduction to Hoare logic so far was informal with an emphasis on intuition. Now we formalize assertions as predicates on states:

`type_synonym assn = state \Rightarrow bool`

As an example of the simplicity of this approach we define validity formally:

$$\models \{P\} c \{Q\} \iff (\forall s t. P s \wedge (c, s) \Rightarrow t \longrightarrow Q t)$$

We pronounce $\models \{P\} c \{Q\}$ as “ $\{P\} c \{Q\}$ is valid”.

Hoare logic with functional assertions is defined as an inductive predicate with the syntax $\vdash \{P\} c \{Q\}$ which is read as “ $\{P\} c \{Q\}$ is derivable/provable” (in Hoare logic). The rules of the inductive definition are shown in [Figure 12.3](#); two derived rules are shown in [Figure 12.4](#). These rules are a direct translation

$\frac{}{\vdash \{P\} \text{ SKIP } \{P\}} \text{Skip}$
$\frac{}{\vdash \{\lambda s. P (s[a/x])\} x ::= a \{P\}} \text{Assign}$
$\frac{\vdash \{P\} c_1 \{Q\} \quad \vdash \{Q\} c_2 \{R\}}{\vdash \{P\} c_1;; c_2 \{R\}} \text{Seq}$
$\frac{\vdash \{\lambda s. P s \wedge \text{bval } b \ s\} c_1 \{Q\} \quad \vdash \{\lambda s. P s \wedge \neg \text{bval } b \ s\} c_2 \{Q\}}{\vdash \{P\} \text{ IF } b \text{ THEN } c_1 \text{ ELSE } c_2 \{Q\}} \text{If}$
$\frac{\vdash \{\lambda s. P s \wedge \text{bval } b \ s\} c \{P\}}{\vdash \{P\} \text{ WHILE } b \text{ DO } c \{\lambda s. P s \wedge \neg \text{bval } b \ s\}} \text{While}$
$\frac{\forall s. P' s \longrightarrow P s \quad \vdash \{P\} c \{Q\} \quad \forall s. Q s \longrightarrow Q' s}{\vdash \{P'\} c \{Q'\}} \text{conseq}$

Fig. 12.3. Hoare logic for partial correctness (functional assertions)

of the syntactic ones, taking into account that assertions are predicates on states. Only rule *Assign* requires some explanation. The notation $s[a/x]$ is merely an abbreviation that mimics syntactic substitution into assertions:

$$s[a/x] \equiv s(x := \text{aval } a \ s)$$

What does our earlier $P[a/x]$ have to do with $P(s[a/x])$? We have not formalized the syntax of assertions, but we can explain what is going on at the level of their close relatives, boolean expressions. Assume we have a substitution function $bsubst$ such that $bsubst \ b \ a \ x$ corresponds to $b[a/x]$, i.e., substitutes a for x in b . Then we can prove

Lemma 12.1 (Substitution lemma).

$$\text{bval } (bsubst \ b \ a \ x) \ s = \text{bval } b \ (s[a/x])$$

It expresses that as far as evaluation is concerned, it does not matter if you substitute into the expression or into the state.

12.2.3 Example Proof [thy](#)

We can now perform Hoare logic proofs in Isabelle. For that purpose we go back to the **apply**-style because it allows us to perform such proofs without having to type in the myriad of intermediate assertions. Instead they are

$$\boxed{
\begin{array}{c}
\frac{\forall s. P \ s \longrightarrow Q \ (s[a/x])}{\vdash \{P\} \ x ::= a \ \{Q\}} \text{Assign}' \\
\\
\frac{\vdash \{\lambda s. P \ s \wedge \text{bval } b \ s\} \ c \ \{P\} \quad \forall s. P \ s \wedge \neg \text{bval } b \ s \longrightarrow Q \ s}{\vdash \{P\} \ \text{WHILE } b \ \text{DO } c \ \{Q\}} \text{While}'
\end{array}
}$$

Fig. 12.4. Derived rules (functional assertions)

computed by rule application, except for the invariants. As an example we verify *wsum* (Section 12.1) once more:

lemma $\vdash \{\lambda s. s \ \text{'x''} = i\} \ \text{'y''} ::= N \ 0;; \ wsum \ \{\lambda s. s \ \text{'y''} = \text{sum } i\}$

Rule *Seq* creates two subgoals:

apply(rule Seq)

1. $\vdash \{\lambda s. s \ \text{'x''} = i\} \ \text{'y''} ::= N \ 0 \ \{?Q\}$
2. $\vdash \{?Q\} \ wsum \ \{\lambda s. s \ \text{'y''} = \text{sum } i\}$

As outlined in The Method above, we left the intermediate assertion open and will instantiate it by working on the second subgoal first (**prefer 2**). Since that is a loop, we have to provide it anyway, because it is the invariant:

prefer 2

apply(rule While'[where $P = \lambda s. (s \ \text{'y''} = \text{sum } i - \text{sum } (s \ \text{'x''}))$])

We have rearranged the earlier invariant $y + \text{sum } x = \text{sum } i$ slightly to please the simplifier.

The first subgoal of *While'* is preservation of the invariant:

1. $\vdash \{\lambda s. s \ \text{'y''} = \text{sum } i - \text{sum } (s \ \text{'x''}) \wedge$
 $\text{bval } (\text{Less } (N \ 0) \ (V \ \text{'x''})) \ s\}$
 $\text{csum } \{\lambda s. s \ \text{'y''} = \text{sum } i - \text{sum } (s \ \text{'x''})\}$

A total of 3 subgoals...

Because *csum* stands for a sequential composition we proceed as above:

apply(rule Seq)

prefer 2

1. $\vdash \{?Q6\} \ \text{'x''} ::= \text{Plus } (V \ \text{'x''}) \ (N \ (- \ 1))$
 $\{\lambda s. s \ \text{'y''} = \text{sum } i - \text{sum } (s \ \text{'x''})\}$
2. $\vdash \{\lambda s. s \ \text{'y''} = \text{sum } i - \text{sum } (s \ \text{'x''}) \wedge$
 $\text{bval } (\text{Less } (N \ 0) \ (V \ \text{'x''})) \ s\}$
 $\text{'y''} ::= \text{Plus } (V \ \text{'y''}) \ (V \ \text{'x''}) \ \{?Q6\}$

A total of 4 subgoals...

Now the two assignment rules (basic and derived) do their job.

`apply(rule Assign)`

`apply(rule Assign')`

The resulting subgoal is large and hard to read because of the substitutions; therefore we do not show it. It corresponds to (12.1) and can be proved by *simp* (not shown). We move on to the second premise of *While'*, the proof that at the exit of the loop the required postcondition is true:

$$\begin{aligned} 1. \forall s. s''y'' = \text{sum } i - \text{sum } (s''x'') \wedge \\ \neg \text{bval } (\text{Less } (N\ 0) (V''x''))\ s \longrightarrow \\ s''y'' = \text{sum } i \end{aligned}$$

A total of 2 subgoals...

This is proved by *simp* and all that is left is the initialization.

$$\begin{aligned} 1. \vdash \{ \lambda s. s''x'' = i \} s''y'' ::= N\ 0 \\ \{ \lambda s. s''y'' = \text{sum } i - \text{sum } (s''x'') \} \end{aligned}$$

`apply(rule Assign')`

The resulting subgoal shows a simple example of substitution into the state:

$$\begin{aligned} 1. \forall s. s''x'' = i \longrightarrow \\ (s[N\ 0 / s''y'']) s''y'' = \text{sum } i - \text{sum } ((s[N\ 0 / s''y'']) s''x'') \end{aligned}$$

The proof is again a plain *simp*.

Functional assertions lead to more verbose statements. For the verification of larger programs one would add some Isabelle syntax magic to make functional assertions look more like syntactic ones. We have refrained from that as our emphasis is on explaining Hoare logic rather than verifying concrete programs.

Exercises

Exercise 12.1. Give a concrete counterexample to this naive version of the assignment rule: $\{P\} x ::= a \{P[a/x]\}$.

Exercise 12.2. Define *bsubst* and prove the Substitution Lemma 12.1. This may require a similar definition and proof for *aexp*.

Exercise 12.3. Define a command *cmax* that stores the maximum of the values of the IMP variables *x* and *y* in the IMP variable *z* and prove that $\vdash \{ \lambda s. \text{True} \} \text{cmax } \{ \lambda s. s''z'' = \text{max } (s''x'') (s''y'') \}$ where *max* is the predefined maximum function.

Exercise 12.4. Let $wsum2 = WHILE\ Not(Eq\ (V\ ''x'')\ (N\ 0))\ DO\ csum$ where $bval\ (Eq\ a_1\ a_2)\ s = (aval\ a_1 = aval\ a_2)$ (see [Exercise 3.7](#)). Prove $\vdash \{\lambda s. s\ ''x'' = i \wedge 0 \leq i\}\ ''y'' ::= N\ 0;;\ wsum2\ \{\lambda s. s\ ''y'' = sum\ i\}$.

Exercise 12.5. Prove

$$\begin{aligned} &\vdash \{\lambda s. s\ ''x'' = x \wedge s\ ''y'' = y \wedge 0 \leq x\} \\ &\quad WHILE\ Less\ (N\ 0)\ (V\ ''x'') \\ &\quad DO\ (''x'' ::= Plus\ (V\ ''x'')\ (N\ (-\ 1));; \\ &\quad \quad ''y'' ::= Plus\ (V\ ''y'')\ (N\ (-\ 1))) \\ &\quad \{\lambda s. s\ ''y'' = y - x\} \end{aligned}$$

Exercise 12.6. Define a command $cmult$ that stores the product of x and y in z (assuming $0 \leq y$) and prove $\vdash \{\lambda s. s\ ''x'' = x \wedge s\ ''y'' = y \wedge 0 \leq y\}\ cmult\ \{\lambda t. t\ ''z'' = x * y\}$.

Exercise 12.7. The following command computes an integer approximation r of the square root of $i \geq 0$, i.e., $r^2 \leq i < (r+1)^2$. Prove

$$\begin{aligned} &\vdash \{\lambda s. s\ x = i \wedge 0 \leq i\} \\ &\quad r ::= N\ 0;;\ r2 ::= N\ 1;; \\ &\quad WHILE\ Not\ (Less\ (V\ x)\ (V\ r2)) \\ &\quad DO\ (r ::= Plus\ (V\ r)\ (N\ 1));; \\ &\quad \quad r2 ::= Plus\ (V\ r2)\ (Plus\ (Plus\ (V\ r)\ (V\ r))\ (N\ 1))) \\ &\quad \{\lambda s. s\ x = i \wedge (s\ r)^2 \leq i \wedge i < (s\ r + 1)^2\} \end{aligned}$$

For readability, x , r and $r2$ abbreviate $''x''$, $''r''$ and $''r2''$. Figure out how $r2$ is related to r before formulating the invariant.

Exercise 12.8. Prove $\vdash \{P\}\ c\ \{\lambda s. True\}$.

Exercise 12.9. Design and prove a forward assignment rule of the form $\vdash \{P\}\ x ::= a\ \{?\}$ where $?$ is some suitable postcondition that depends on P , x and a .

12.3 Soundness and Completeness [thy](#)

So far we have motivated the rules of Hoare logic by operational semantics considerations but we have not proved a precise link between the two. We will now prove

Soundness of the logic w.r.t. the operational semantics:

if a triple is derivable, it is also valid.

Completeness of the logic w.r.t. the operational semantics:
if a triple is valid, it is also derivable.

Recall the definition of validity:

$$\models \{P\} c \{Q\} \iff (\forall s t. P s \wedge (c, s) \Rightarrow t \longrightarrow Q t)$$

Soundness is straightforward:

Lemma 12.2 (Soundness of \vdash w.r.t. \models).

$$\vdash \{P\} c \{Q\} \implies \models \{P\} c \{Q\}$$

Proof. By induction on the derivation of $\vdash \{P\} c \{Q\}$: we must show that every rule of the logic preserves validity. This is automatic for all rules except *While*, because these rules resemble their big-step counterparts. Even assignment is easy: To prove $\models \{\lambda s. P (s[a/x])\} x ::= a \{P\}$ we may assume $P (s[a/x])$ and $(x ::= a, s) \Rightarrow t$. Therefore $t = s[a/x]$ and thus $P t$ as required.

The only other rule we consider is *While*. We may assume the IH $\models \{\lambda s. P s \wedge bval b s\} c \{P\}$. First we prove for arbitrary s and t that if $(WHILE b DO c, s) \Rightarrow t$ then $P s \implies P t \wedge \neg bval b t$ by induction on the assumption. There are two cases. If $\neg bval b s$ and $s = t$ then $P s \implies P t \wedge \neg bval b t$ is trivial. If $bval b s$, $(c, s) \Rightarrow s'$ and IH $P s' \implies P t \wedge \neg bval b t$, then we assume $P s$ and prove $P t \wedge \neg bval b t$. Because $P s$, $bval b s$ and $(c, s) \Rightarrow s'$, the outer IH yields $P s'$ and then the inner IH yields $P t \wedge \neg bval b t$, thus finishing the inner induction. Returning to the *While* rule we need to prove $\models \{P\} WHILE b DO c \{\lambda s. P s \wedge \neg bval b s\}$. Assuming $P s$ and $(WHILE b DO c, s) \Rightarrow t$, the lemma we just proved locally yields the required $P t \wedge \neg bval b t$, thus finishing the *While* rule. \square

Soundness was straightforward, as usual: one merely has to plough through the rules one by one. Completeness requires new ideas.

12.3.1 Completeness

Completeness, i.e., $\models \{P\} c \{Q\} \implies \vdash \{P\} c \{Q\}$, is proved with the help of the notion of the **weakest precondition** (in the literature often called the **weakest liberal precondition**):

definition $wp :: com \Rightarrow assn \Rightarrow assn$ where
 $wp c Q = (\lambda s. \forall t. (c, s) \Rightarrow t \longrightarrow Q t)$

Thinking of assertions as sets of states, this says that the weakest precondition of a command c and a postcondition Q is the set of all pre-states such that if c terminates one ends up in Q .

It is easy to see that $wp c Q$ is indeed the weakest precondition:

Fact 12.3. $\models \{P\} c \{Q\} \longleftrightarrow (\forall s. P s \longrightarrow wp\ c\ Q\ s)$

The weakest precondition is a central concept in Hoare logic because it formalizes the idea of pulling a postcondition back through a command to obtain the corresponding precondition. This idea is central to the construction of Hoare logic proofs and we have already alluded to it multiple times.

The following nice recursion equations hold for wp . They can be used to compute the weakest precondition, except for *WHILE*, which would lead to nontermination.

$$\begin{aligned}
 wp\ SKIP\ Q &= Q \\
 wp\ (x ::= a)\ Q &= (\lambda s. Q\ (s[a/x])) \\
 wp\ (c_1;; c_2)\ Q &= wp\ c_1\ (wp\ c_2\ Q) \\
 wp\ (IF\ b\ THEN\ c_1\ ELSE\ c_2)\ Q &= (\lambda s. \text{if } bval\ b\ s \text{ then } wp\ c_1\ Q\ s \text{ else } wp\ c_2\ Q\ s) \\
 wp\ (WHILE\ b\ DO\ c)\ Q &= (\lambda s. \text{if } bval\ b\ s \text{ then } wp\ (c;; WHILE\ b\ DO\ c)\ Q\ s \text{ else } Q\ s)
 \end{aligned}$$

Proof. All equations are easily proved from the definition of wp once you realise that they are equations between functions. Such equations can be proved with the help of **extensionality**, one of the basic rules of HOL:

$$\frac{\bigwedge x. f\ x = g\ x}{f = g} \text{ ext}$$

It expresses that two functions are equal if they are equal for all arguments. For example, we can prove $wp\ SKIP\ Q = Q$ by proving $wp\ SKIP\ Q\ s = Q\ s$ for an arbitrary s . Expanding the definition of wp we have to prove $(\forall t. (SKIP, s) \Rightarrow t \longrightarrow Q\ t) = Q\ s$, which follows by inversion of the big-step rule for *SKIP*. The proof of the other equations is similar. \square

The key property of wp is that it is also a precondition w.r.t. provability:

Lemma 12.4. $\vdash \{wp\ c\ Q\} c \{Q\}$

Proof. By induction on c . We consider only the *WHILE* case, the other cases are automatic (with the help of the wp equations). Let $w = WHILE\ b\ DO\ c$. We show $\vdash \{wp\ w\ Q\} w \{Q\}$ by an application of rule *While'*. Its first premise is $\vdash \{\lambda s. wp\ w\ Q\ s \wedge bval\ b\ s\} c \{wp\ w\ Q\}$. It follows from the IH $\vdash \{wp\ c\ R\} c \{R\}$ (for any R) where we set $R = wp\ w\ Q$, by precondition strengthening: the implication $wp\ w\ Q\ s \wedge bval\ b\ s \longrightarrow wp\ c\ (wp\ w\ Q)\ s$ follows from the wp equations for *WHILE* and $;;$. The second premise we need to prove is $wp\ w\ Q\ s \wedge \neg bval\ b\ s \longrightarrow Q\ s$; it follows from the wp equation for *WHILE*. \square

The completeness theorem is an easy consequence:

Theorem 12.5 (Completeness of \vdash w.r.t. \models).

$$\models \{P\} c \{Q\} \implies \vdash \{P\} c \{Q\}$$

Proof. Because wp is the weakest precondition (Fact 12.3), $\models \{P\} c \{Q\}$ implies $\forall s. P s \longrightarrow wp\ c\ Q\ s$. Therefore we can strengthen the precondition of $\vdash \{wp\ c\ Q\} c \{Q\}$ and infer $\vdash \{P\} c \{Q\}$. \square

Putting soundness and completeness together we obtain that a triple is provable in Hoare logic iff it is provable via the operational semantics:

Corollary 12.6. $\vdash \{P\} c \{Q\} \iff \models \{P\} c \{Q\}$

Thus one can also view Hoare logic as a reformulation of operational semantics aimed at proving rather than executing — or the other way around.

12.3.2 Incompleteness

Having proved completeness we will now explain a related incompleteness result. This section requires some background in recursion theory and logic. It can be skipped on first reading.

Recall from Section 12.2.1 the triple $\{\text{True}\} c \{\text{False}\}$. We argued that this triple is valid iff c terminates for no start state. It is well known from recursion theory that the set of such never terminating programs is not recursively enumerable (r.e.) (see, for example, [42]). Therefore the set of valid Hoare triples is not r.e. either: an enumeration of all valid Hoare triples could easily be filtered to yield an enumeration of all valid $\{\text{True}\} c \{\text{False}\}$ and thus an enumeration of all never terminating programs.

Therefore there is no sound and complete Hoare logic whose provable triples are r.e.. This is strange because we have just proved that our Hoare logic is sound and complete, and its inference rules together with the inference rules for HOL (see, for example, [36]) provide an enumeration mechanism for all provable Hoare triples. What is wrong here? Nothing. Both our soundness and completeness results and the impossibility of having a sound and complete Hoare logic are correct but they refer to subtly different notions of validity of Hoare triples. The notion of validity used in our soundness and completeness proofs is defined in HOL and we have shown that we can prove a triple valid iff we can prove it in the Hoare logic. Thus we have related provability in HOL of two different predicates. On the other hand, the impossibility result is based on an abstract mathematical notion of validity and termination independent of any proof system. This abstract notion of validity is stronger than our HOL-based definition. That is, there are Hoare triples that are valid in this abstract sense but whose validity cannot be proved in HOL. Otherwise the equivalence of \models and \vdash we proved in HOL would contradict the impossibility of having a sound and complete Hoare logic (assuming that HOL is consistent).

What we have just seen is an instance of the incompleteness of HOL, not Hoare logic: there are sentences in HOL that are true but not provable. Gödel [34] showed that this is necessarily the case in any sufficiently strong, consistent and r.e. logic. Cook [20] was the first to analyse this incompleteness of Hoare logic and to show that, because it is due to the incompleteness of the assertion language, one can still prove what he called **relative completeness** of Hoare logic. The details are beyond the scope of this book. Winskel [93] provides a readable account.

Exercises

Exercise 12.10. Prove [Fact 12.3](#).

Exercise 12.11. Replace the assignment command with a new command *Do f* where $f :: \text{state} \Rightarrow \text{state}$ can be an arbitrary state transformer. Update the big-step semantics, Hoare logic and the soundness and completeness proofs.

Exercise 12.12. Consider the following rule schema:

$$\llbracket \vdash \{P\} c \{Q\}; \vdash \{P'\} c \{Q'\} \rrbracket \implies \vdash \{\lambda s. P s \odot P' s\} c \{\lambda s. Q s \odot Q' s\}$$

For each $\odot \in \{\wedge, \vee, \longrightarrow\}$, give a proof or a counterexample.

Exercise 12.13. Based on [Exercise 7.9](#), extend Hoare logic and the soundness and completeness proofs with nondeterministic choice.

Exercise 12.14. Based on [Exercise 7.8](#), extend Hoare logic and the soundness and completeness proofs with a *REPEAT* loop.

Exercise 12.15. The dual of the weakest precondition is the **strongest postcondition** *sp*. Define $sp :: \text{com} \Rightarrow \text{assn} \Rightarrow \text{assn}$ in analogy with *wp* via the big-step semantics. Prove that *sp* really is the strongest postcondition: $\models \{P\} c \{Q\} \iff (\forall s. sp\ c\ P\ s \longrightarrow Q\ s)$. In analogy with the derived equations for *wp* given in the text, give and prove equations for “calculating” *sp* for three constructs: $sp\ (x ::= a)\ P = Q_1$, $sp\ (c_1;; c_2)\ P = Q_2$, and $sp\ (IF\ b\ THEN\ c_1\ ELSE\ c_2)\ P = Q_3$. The Q_i must not involve the semantics and may only call *sp* recursively on the subcommands c_i . Hint: Q_1 requires an existential quantifier.

12.4 Verification Condition Generation thy

This section shows what we have already hinted at: Hoare logic can be automated. That is, we reduce provability in Hoare logic to provability in the assertion language, i.e., HOL in our case. Given a triple $\{P\} c \{Q\}$ that we want to prove, we show how to compute an assertion A from it such that $\vdash \{P\} c \{Q\}$ is provable iff A is provable.

We call A a **verification condition** and the function that computes A a **verification condition generator** or **VCG**. The advantage of working with a VCG is that no knowledge of Hoare logic is required by the person or machine that attempts to prove the generated verification conditions. Most systems for the verification of imperative programs are based on VCGs.

Our VCG works like The Method for Hoare logic we sketched above: it simulates the backward application of Hoare logic rules and gathers up the implications between assertions that arise in the process. Of course there is the problem of loop invariants: where do they come from? We take the easy way out and let the user provide them. In general this is the only feasible solution because we cannot expect the machine to come up with clever invariants in all situations. In [Chapter 13](#) we will present a method for computing invariants in simple situations.

Invariants are supplied to the VCG as annotations of *WHILE* loops. For that purpose we introduce a type *acom* of **annotated commands** with the same syntax as that of type *com*, except that *WHILE* is annotated with an assertion *Inv*:

$$\{Inv\} \text{ WHILE } b \text{ DO } C$$

To distinguish variables of type *com* and *acom*, the latter are capitalised. Function *strip* :: *acom* \Rightarrow *com* removes all annotations from an annotated command, thus turning it into an ordinary command.

Verification condition generation is based on two functions: *pre* is similar to *wp*, *vc* is the actual VCG.

```

fun pre :: acom  $\Rightarrow$  assn  $\Rightarrow$  assn where
  pre SKIP Q      = Q
  pre (x ::= a) Q  = ( $\lambda s.$  Q (s[a/x]))
  pre (C1;; C2) Q = pre C1 (pre C2 Q)
  pre (IF b THEN C1 ELSE C2) Q
    = ( $\lambda s.$  if bval b s then pre C1 Q s else pre C2 Q s)
  pre ({I} WHILE b DO C) Q = I

```

Function *pre* follows the recursion equations for *wp* except in the *WHILE* case where the annotation is returned. If the annotation is an invariant then it must also hold before the loop and thus it makes sense for *pre* to return it.

In contrast to *pre*, *vc* produces a formula that is independent of the state:

```

fun vc :: acom ⇒ assn ⇒ bool where
  vc SKIP Q      = True
  vc (x ::= a) Q  = True
  vc (C1;; C2) Q = (vc C1 (pre C2 Q) ∧ vc C2 Q)
  vc (IF b THEN C1 ELSE C2) Q = (vc C1 Q ∧ vc C2 Q)
  vc ({I} WHILE b DO C) Q =
    ((∀ s. (I s ∧ bval b s ⇒ pre C I s) ∧ (I s ∧ ¬ bval b s ⇒ Q s)) ∧
     vc C I)

```

Function *vc* essentially just goes through the command and produces the following two verification conditions for each $\{I\}$ WHILE *b* DO *C*:

- $\forall s. I s \wedge bval\ b\ s \longrightarrow pre\ C\ I\ s$
It expresses that *I* and *b* together imply the precondition that *I* holds again after *C*, i.e., *I* is an invariant.
- $\forall s. I s \wedge \neg bval\ b\ s \longrightarrow Q\ s$
It expresses that at the end of the loop the postcondition holds.

The recursive invocation *vc C I* merely generates the verification conditions for any loops inside the body *C*.

For the other constructs only trivial verification conditions (*True*) are generated or the results of subcomputations are combined with \wedge .

In examples we revert to syntactic assertions for compactness and readability. We only need to drop the state parameter *s* and perform substitution on assertions instead of states.

Example 12.7. We return to our familiar summation program (ignoring the initialization of *y*) and define the following abbreviations:

```

W  = {I} WHILE 0 < x DO C      C  = y := y+x;; x := x-1
I  = (y + sum x = sum i)      Q  = (y = sum i)

```

The following equations show the computation of *vc W Q*:

```

vc W Q = ((I ∧ 0 < x ⇒ pre C I) ∧ (I ∧ ¬ 0 < x ⇒ Q)
          ∧ vc C I)
pre C I = pre (y := y+x) (pre (x := x-1) I)
          = pre (y := y+x) (I[x-1/x])
          = I[x-1/x][y+x/y] = (y+x + sum (x-1) = sum i)
vc C I = (vc (y := y+x) (pre (x := x-1) I) ∧ vc (x := x-1) I)
          = (True ∧ True)

```

Therefore *vc W Q* boils down to

$$(I \wedge 0 < x \longrightarrow I[x-1/x][y+x/y]) \wedge (I \wedge x \leq 0 \longrightarrow Q)$$

The same conditions arose on pages 197f. in the Hoare logic proof of the triple $\{I\} \text{ wsum } \{y = \text{sum } i\}$ and we satisfied ourselves that they are true.

The example has demonstrated the computation of vc and pre . The generated verification condition turned out to be true, but it remains unclear what that proves. We need to show that our VCG is sound w.r.t. Hoare logic. This will allow us to reduce the problem of proving $\vdash \{P\} c \{Q\}$ to the problem of proving the verification condition. We will obtain the following result:

Corollary 12.8.

$$\llbracket vc \ C \ Q; \forall s. P \ s \longrightarrow pre \ C \ Q \ s \rrbracket \Longrightarrow \vdash \{P\} \text{ strip } C \{Q\}$$

This can be read as a procedure for proving $\vdash \{P\} c \{Q\}$:

1. Annotate c with invariants, yielding C such that $\text{strip } C = c$.
2. Prove the verification condition $vc \ C \ Q$ and that P implies $pre \ C \ Q$.

The actual soundness lemma is a bit more compact than its above corollary which follows from it by precondition strengthening.

Lemma 12.9 (Soundness of pre and vc w.r.t. \vdash).

$$vc \ C \ Q \Longrightarrow \vdash \{pre \ C \ Q\} \text{ strip } C \{Q\}$$

Proof. By induction on c . The *WHILE* case is routine, the other cases are automatic. \square

How about completeness? Can we just reverse this implication? Certainly not: if C is badly annotated, $\vdash \{pre \ C \ Q\} \text{ strip } C \{Q\}$ may be provable but not $vc \ C \ Q$.

Example 12.10. The triple $\{x=1\} \text{ WHILE True DO } x:=0 \{False\}$ is provable with the help of the invariant *True* and precondition strengthening:

$$\frac{\frac{\frac{x=1 \longrightarrow \text{True}}{\text{True} \wedge \text{True} \longrightarrow \text{True}} \quad \frac{\text{True} \wedge \neg \text{True} \longrightarrow \text{False}}{\{True\} \text{ WHILE True DO } x:=0 \{False\}}}{\{True \wedge True\} x:=0 \{True\}} \quad \{True\} \text{ WHILE True DO } x:=0 \{False\}}{\{x=1\} \text{ WHILE True DO } x:=0 \{False\}}$$

But starting from the badly annotated $W = \{x=1\} \text{ WHILE True DO } x := 0$ one of the verification conditions will be that $x=1$ is an invariant, which it is not. Hence $vc \ W \ False$ is not true.

However there always is an annotation that works:

Lemma 12.11 (Completeness of pre and vc w.r.t. \vdash).

$$\vdash \{P\} c \{Q\} \Longrightarrow \exists C. \text{ strip } C = c \wedge vc \ C \ Q \wedge (\forall s. P \ s \longrightarrow pre \ C \ Q \ s)$$

Proof. The proof requires two little monotonicity lemmas:

$$\begin{aligned} \llbracket \forall s. P \ s \longrightarrow P' \ s; \text{pre } C \ P \ s \rrbracket &\Longrightarrow \text{pre } C \ P' \ s \\ \llbracket \forall s. P \ s \longrightarrow P' \ s; \text{vc } C \ P \rrbracket &\Longrightarrow \text{vc } C \ P' \end{aligned}$$

Both are proved by induction on c ; each case is automatic.

In the rest of the proof the formula $\forall s. P \ s$ is abbreviated to P and $\forall s. P \ s \longrightarrow Q \ s$ is abbreviated to $P \rightarrow Q$.

The proof of the completeness lemma is by rule induction on $\vdash \{P\} \ c \ \{Q\}$. We only consider the sequential composition rule in detail:

$$\frac{\vdash \{P_1\} \ c_1 \ \{P_2\} \quad \vdash \{P_2\} \ c_2 \ \{P_3\}}{\vdash \{P_1\} \ c_1;; c_2 \ \{P_3\}}$$

From the IHs we obtain C_1 and C_2 such that $\text{strip } C_1 = c_1$, $\text{vc } C_1 \ P_2$, $P_1 \rightarrow \text{pre } C_1 \ P_2$, $\text{strip } C_2 = c_2$, $\text{vc } C_2 \ P_3$, $P_2 \rightarrow \text{pre } C_2 \ P_3$. We claim that $C' = C_1;; C_2$ is the required annotated command. Clearly $\text{strip } C' = c_1;; c_2$. From $\text{vc } C_1 \ P_2$ and $P_2 \rightarrow \text{pre } C_2 \ P_3$ it follows by monotonicity that $\text{vc } C_1 \ (\text{pre } C_2 \ P_3)$; together with $\text{vc } C_2 \ P_3$ this implies $\text{vc } C' \ P_3$. From $P_2 \rightarrow \text{pre } C_2 \ P_3$ it follows by monotonicity of pre that $\text{pre } C_1 \ P_2 \rightarrow \text{pre } C_1 \ (\text{pre } C_2 \ P_3)$; because $P_1 \rightarrow \text{pre } C_1 \ P_2$ we obtain the required $P_1 \rightarrow \text{pre } C' \ P_3$, thus concluding the case of the sequential composition rule.

The *WHILE* rule is special because we need to synthesise the loop annotation. This is easy: take the invariant P from the *WHILE* rule.

The remaining rules are straightforward. \square

Exercises

Exercise 12.16. Let $\text{asum } i$ be the annotated command $y := 0; W$ where W is defined in [Example 12.7](#). Prove $\vdash \{\lambda s. s \ \text{'x'} = i\} \ \text{strip } (\text{asum } i) \ \{\lambda s. s \ \text{'y'} = \text{sum } i\}$ with the help of [Corollary 12.8](#).

Exercise 12.17. Solve [Exercises 12.4 to 12.7](#) using the VCG: for every Hoare triple $\vdash \{P\} \ c \ \{Q\}$ from one of those exercises define an annotated version C of c and prove $\vdash \{P\} \ \text{strip } C \ \{Q\}$ with the help of [Corollary 12.8](#).

Exercise 12.18. Having two separate functions pre and vc is inefficient. When computing vc one often needs to compute pre too, leading to multiple traversals of many subcommands. Define an optimized function $\text{prevc} :: \text{acom} \Rightarrow \text{assn} \times \text{bool}$ that traverses the command only once and prove that $\text{prevc } C \ Q = (\text{pre } C \ Q, \text{vc } C \ Q)$.

Exercise 12.19. Design a VCG that computes post- rather than preconditions. Start by solving [Exercise 12.9](#). Now modify theory *VCG* as follows. Instead of pre define a function $\text{post} :: \text{acom} \Rightarrow \text{assn} \Rightarrow \text{assn}$ such that

(with the exception of loops) $\text{post } C \ P$ is the strongest postcondition of C w.r.t. the precondition P (see also [Exercise 12.15](#)). Now modify vc such that it uses post instead of pre and prove its soundness and completeness:

$$\begin{aligned} vc \ C \ P &\Longrightarrow \vdash \{P\} \ \text{strip } C \ \{\text{post } C \ P\} \\ \vdash \{P\} \ c \ \{Q\} &\Longrightarrow \exists C. \ \text{strip } C = c \wedge vc \ C \ P \wedge (\forall s. \ \text{post } C \ P \ s \longrightarrow Q \ s) \end{aligned}$$

12.5 Hoare Logic for Total Correctness thy

Recall the informal definition of total correctness of a triple $\{P\} \ c \ \{Q\}$:

If P is true before the execution of c
then c terminates and Q is true afterwards.

Formally, **validity for total correctness** is defined like this:

$$\models_t \{P\} \ c \ \{Q\} \iff (\forall s. \ P \ s \longrightarrow (\exists t. \ (c, s) \Rightarrow t \wedge Q \ t))$$

In this section we always refer to this definition when we speak of validity.

Note that this definition assumes that the language is deterministic. Otherwise $\models_t \{P\} \ c \ \{Q\}$ may hold although only some computations starting from P terminate in Q while others may show any behaviour whatsoever.

Hoare logic for total correctness is defined as for partial correctness, except that we write \vdash_t and that the *WHILE* rule is augmented with a relation $T :: \text{state} \Rightarrow \text{nat} \Rightarrow \text{bool}$ that guarantees termination:

$$\frac{\bigwedge n. \vdash_t \{\lambda s. \ P \ s \wedge \text{bval } b \ s \wedge T \ s \ n\} \ c \ \{\lambda s. \ P \ s \wedge (\exists n' < n. \ T \ s \ n')\}}{\vdash_t \{\lambda s. \ P \ s \wedge (\exists n. \ T \ s \ n)\} \ \text{WHILE } b \ \text{DO } c \ \{\lambda s. \ P \ s \wedge \neg \text{bval } b \ s\}}$$

The purpose of the universally quantified $n :: \text{nat}$ in the premise is to remember the value of T in the precondition to express that it has decreased in the postcondition. The name of the rule is again *While*.

Although this formulation with a relation T has a technical advantage, the following derived rule formulated with a measure function $f :: \text{state} \Rightarrow \text{nat}$ is more intuitive. We call this rule *While_fun*:

$$\frac{\bigwedge n. \vdash_t \{\lambda s. \ P \ s \wedge \text{bval } b \ s \wedge n = f \ s\} \ c \ \{\lambda s. \ P \ s \wedge f \ s < n\}}{\vdash_t \{P\} \ \text{WHILE } b \ \text{DO } c \ \{\lambda s. \ P \ s \wedge \neg \text{bval } b \ s\}}$$

This is like the partial correctness rule except that it also requires a measure function that decreases with each iteration. In case you wonder how to derive the functional version: set $T = (\lambda s \ n. \ f \ s = n)$ in rule *While*.

Example 12.12. We redo the proof of *wsum* from Section 12.2.2. The only difference is that when applying rule *While_fun* (combined with postcondition strengthening as in rule *While'*) we need not only instantiate P as previously but also f : $f = (\lambda s. \text{nat } (s \text{ ''}x''))$ where the predefined function *nat* coerces an *int* into a *nat*, coercing all negative numbers to 0. The resulting invariance subgoal now looks like this:

$$\begin{aligned} \wedge n. \vdash_t \{ & \lambda s. s \text{ ''}y'' = \text{sum } i - \text{sum } (s \text{ ''}x'') \wedge \\ & \text{bval } (\text{Less } (N\ 0) (V \text{ ''}x''))\ s \wedge n = \text{nat } (s \text{ ''}x'') \} \\ & \text{''}y'' ::= \text{Plus } (V \text{ ''}y'') (V \text{ ''}x'');; \\ & \text{''}x'' ::= \text{Plus } (V \text{ ''}x'') (N\ (-\ 1)) \\ & \{ \lambda s. s \text{ ''}y'' = \text{sum } i - \text{sum } (s \text{ ''}x'') \wedge \text{nat } (s \text{ ''}x'') < n \} \end{aligned}$$

The rest of the proof steps are again identical to the partial correctness proof. After pulling back the postcondition the additional conjunct in the goal is $\text{nat } (s \text{ ''}x'' - 1) < n$ which follows automatically from the assumptions $0 < s \text{ ''}x''$ and $n = \text{nat } (s \text{ ''}x'')$.

Lemma 12.13 (Soundness of \vdash_t w.r.t. \models_t).

$$\vdash_t \{P\} c \{Q\} \implies \models_t \{P\} c \{Q\}$$

Proof. By rule induction. All cases are automatic except rule *While*, which we look at in detail. Let $w = \text{WHILE } b \text{ DO } c$. By a (nested) induction on n we show for arbitrary n and s

$$\llbracket P\ s; T\ s\ n \rrbracket \implies \exists t. (w, s) \Rightarrow t \wedge P\ t \wedge \neg \text{bval } b\ t \quad (*)$$

from which $\models_t \{ \lambda s. P\ s \wedge (\exists n. T\ s\ n) \} w \{ \lambda s. P\ s \wedge \neg \text{bval } b\ s \}$, the goal in the *While* case, follows immediately. The inductive proof assumes that $(*)$ holds for all smaller n . For n itself we argue by cases. If $\neg \text{bval } b\ s$ then $(w, s) \Rightarrow t$ is equivalent with $t = s$ and $(*)$ follows. If $\text{bval } b\ s$ then we assume $P\ s$ and $T\ s\ n$. The outer IH $\models_t \{ \lambda s. P\ s \wedge \text{bval } b\ s \wedge T\ s\ n \} c \{ \lambda s'. P\ s' \wedge (\exists n'. T\ s'\ n' \wedge n' < n) \}$ yields s' and n' such that $(c, s) \Rightarrow s', P\ s', T\ s'\ n'$ and $n' < n$. Because $n' < n$ the inner IH yields the required t such that $(w, s') \Rightarrow t, P\ t$ and $\neg \text{bval } b\ t$. With rule *WhileTrue* the conclusion of $(*)$ follows. \square

The completeness proof proceeds like the one for partial correctness. The weakest precondition is now defined in correspondence to \models_t :

$$wp_t\ c\ Q = (\lambda s. \exists t. (c, s) \Rightarrow t \wedge Q\ t)$$

The same recursion equations as for wp can also be proved for wp_t . The crucial lemma is again this one:

Lemma 12.14. $\vdash_t \{wp_t\ c\ Q\} c \{Q\}$

Proof. By induction on c . We focus on the *WHILE* case because the others are automatic, thanks to the wp_t equations. Let $w = \text{WHILE } b \text{ DO } c$. The termination relation T counts the number of iterations of w and is defined inductively by the two rules

$$\frac{\neg \text{bval } b \ s}{T \ s \ 0} \quad \frac{\text{bval } b \ s \quad (c, s) \Rightarrow s' \quad T \ s' \ n}{T \ s \ (n + 1)}$$

Because IMP is deterministic, T is a functional relation:

$$\llbracket T \ s \ n; T \ s \ n' \rrbracket \implies n = n'$$

This is easily proved by induction on the first premise.

Moreover, T is ‘defined’ for any state from which w terminates:

$$(w, s) \Rightarrow t \implies \exists n. T \ s \ n \quad (*)$$

The proof is an easy rule induction on the premise.

Now we come to the actual proof. The IH is $\vdash_t \{wp_t \ c \ R\} \ c \ \{R\}$ for any R , and we need to prove $\vdash_t \{wp_t \ w \ Q\} \ w \ \{Q\}$. In order to apply the *WHILE* rule we use the consequence rule to turn the precondition into $P = (\lambda s. wp_t \ w \ Q \ s \wedge (\exists n. T \ s \ n))$ and the postcondition into $\lambda s. P \ s \wedge \neg \text{bval } b \ s$. Thus we have to prove the following three goals:

$$\begin{aligned} & \forall s. wp_t \ w \ Q \ s \longrightarrow P \ s \\ & \vdash_t \{P\} \ w \ \{\lambda s. P \ s \wedge \neg \text{bval } b \ s\} \\ & \forall s. P \ s \wedge \neg \text{bval } b \ s \longrightarrow Q \ s \end{aligned}$$

The third goal follows because $\neg \text{bval } b \ s$ implies $wp_t \ w \ Q \ s = Q \ s$ and therefore $P \ s$ implies $Q \ s$. The first goal follows directly from $(*)$ by definition of wp_t . Applying the *WHILE* rule backwards to the second goal leaves us with $\vdash_t \{P'\} \ c \ \{R\}$ where $P' = (\lambda s. P \ s \wedge \text{bval } b \ s \wedge T \ s \ n)$ and $R = (\lambda s'. P \ s' \wedge (\exists n' < n. T \ s' \ n'))$. By IH we have $\vdash_t \{wp_t \ c \ R\} \ c \ \{R\}$ and we can obtain $\vdash_t \{P\} \ c \ \{R\}$ by precondition strengthening because $\forall s. P' \ s \longrightarrow wp_t \ c \ R \ s$. To prove the latter we assume $P' \ s$ and show $wp_t \ c \ R \ s$. From $P \ s$ we obtain by definition of wp_t some t such that $(w, s) \Rightarrow t$ and $Q \ t$. Because $\text{bval } b \ s$, rule inversion yields a state s' such that $(c, s) \Rightarrow s'$ and $(w, s') \Rightarrow t$. From $(w, s') \Rightarrow t$ we obtain a number n' such that $T \ s' \ n'$. By definition of T , $T \ s \ (n' + 1)$ follows. Because T is functional we have $n = n' + 1$. Together with $(c, s) \Rightarrow s'$, $(w, s') \Rightarrow t$, $Q \ t$ and $T \ s' \ n'$ this implies $wp_t \ c \ R \ s$ by definition of wp_t . \square

The completeness theorem is an easy consequence:

Theorem 12.15 (Completeness of \vdash_t w.r.t. \models_t).

$$\models_t \{P\} \ c \ \{Q\} \implies \vdash_t \{P\} \ c \ \{Q\}$$

The proof is the same as for [Theorem 12.5](#).

Exercises

Exercise 12.20. Prove total correctness of the commands in Exercises 12.4 to 12.7.

Exercise 12.21. Modify the VCG from Section 12.4 to take termination into account. First modify type *acom* by annotating *WHILE* with a measure function $f :: \text{state} \Rightarrow \text{nat}$ in addition to an invariant:

$$\{I, f\} \text{ WHILE } b \text{ DO } C$$

Functions *strip* and *pre* remain almost unchanged. The only significant change is in the *WHILE* case for *vc*. Finally update the old soundness proof to obtain $vc \ C \ Q \Longrightarrow \vdash_t \{pre \ C \ Q\} \text{ strip } C \ \{Q\}$. You may need the combined soundness and completeness of \vdash_t : $(\vdash_t \{P\} \ c \ \{Q\}) = (\models_t \{P\} \ c \ \{Q\})$.

Exercise 12.22. An alternative version of the *WHILE* rule indexes the invariant by a natural number that must goes down by one with every iteration:

$$\frac{\bigwedge n. \vdash_t \{P \ (Suc \ n)\} \ c \ \{P \ n\} \quad \forall n \ s. P \ (Suc \ n) \ s \longrightarrow bval \ b \ s \quad \forall s. P \ 0 \ s \longrightarrow \neg \ bval \ b \ s}{\vdash_t \{\lambda s. \exists n. P \ n \ s\} \text{ WHILE } b \text{ DO } c \ \{P \ 0\}}$$

Prove soundness and completeness of this alternative set of rules. For the completeness proof it may be helpful to define a recursive function $wpw :: \text{bexp} \Rightarrow \text{com} \Rightarrow \text{nat} \Rightarrow \text{assn} \Rightarrow \text{assn}$ such that $wpw \ b \ c \ n \ Q$ is the weakest precondition such that *WHILE* *b* *DO* *c* terminates after *n* iterations in a state satisfying *Q*.

12.6 Summary and Further Reading

This chapter was dedicated to Hoare logic and the verification of IMP programs. We covered three main topics:

- A Hoare logic for partial correctness and its soundness and completeness w.r.t. the big-step semantics.
- A verification condition generator that reduces the task of verifying a program by means of Hoare logic to the task of annotating all loops in that program with invariants and proving that these annotations are indeed invariants and imply the necessary postconditions.
- A Hoare logic for total correctness and its soundness and completeness.

Hoare logic is a huge subject area and we have only scratched the surface. Therefore we provide further references for the theory and applications of Hoare logic.

12.6.1 Theory

A precursor of Hoare logic is Floyd's method of annotated flowcharts [30]. Hoare [41] transferred Floyd's idea to inference rules for structured programs and Hoare logic (as we now know it) was born. Hence it is sometimes called Floyd-Hoare logic. Soundness and completeness was first proved by Cook [20]. An early overview of the foundations of various Hoare logics is due to Apt [6, 7]. An excellent modern introduction to the many variants of Hoare logic is the book by Apt, de Boer and Olderog [8], which covers procedures, objects and concurrency. Weakest preconditions are due to Dijkstra [28]. He reformulated Hoare logic as a weakest precondition calculus to facilitate the verification of concrete programs.

All of the above references follow the syntactic approach to assertions. The book by Nielson and Nielson [63] is an exception in that its chapters on program verification follow the functional approach. Formalizations of Hoare logic in theorem provers have all followed the functional approach. The first such formalization is due to Gordon [37], who showed the way and proved soundness. Nipkow [65] also proved completeness. Schreiber [80] covered procedures, which was extended with nondeterminism by Nipkow [67, 66]. Nipkow and Prenska [69] formalized a Hoare logic for concurrency.

Formalizing Hoare logics is not a frivolous pastime. Apt observes "various proofs given in the literature are awkward, incomplete or even incorrect" [6]. In fact, Apt himself presents a proof system for total correctness of procedures that was later found to be unsound by America and de Boer [5], who presented a correct version. The formalized system by Schreiber [80] is arguably slicker.

Finally we mention work on verifying programs that manipulate data structures on the heap. An early approach by Burstall [16] was formalized by Bornat [13] and Mehta and Nipkow [55]. A recent and very influential approach is **separation logic** [74]. Its formulas provide special connectives for talking about the layout of data on the heap.

12.6.2 Applications

Our emphasis on foundational issues and toy examples is bound to give the reader the impression that Hoare logic is not practically useful. Luckily, that is not the case.

Among the early practical program verification tools are the KeY [10] and KIV [73] systems. The KeY system can be used to reason about Java programs. One of the recent larger applications of the KIV system is for instance the verification of file-system code in operating systems [29].

Hoare logic-based tools also exist for concurrent programs. VCC [19], for instance, is integrated with Microsoft Visual Studio and can be used to verify

concurrent C. Cohen *et al.* have used VCC to demonstrate the verification of a small operating system hypervisor [4]. Instead of interactive proof, VCC aims for automation and uses the SMT solver Z3 [25] as back-end. The interaction with the tool consists of annotating invariants and function pre- and postconditions in such a way that they are simple enough for the prover to succeed automatically. The automation is stronger than in interactive proof assistants like Isabelle, but it forces specifications to be phrased in first-order logic. The Dafny tool [52] uses the same infrastructure. It does not support concurrency, but is well suited for learning this paradigm of program verification.

Of course, Isabelle also supports practical program verification. As mentioned in [Chapter 7](#), the Simpl [78] verification framework for Isabelle develops Hoare logic from its foundations to a degree that can directly be used for the verification of C programs [88]. While it provides less automation than tools like VCC, the user is rewarded with the full power and flexibility of higher-order logic. Two of the largest formal software verifications to date used Isabelle, both in the area of operating system verification. The Verisoft project [3] looked at constructing a verified software stack from verified hardware up to verified applications, an aspiration of the field that goes back to the 1980s [12]. The seL4 project verified a commercially viable operating system microkernel consisting of roughly 10,000 lines of code [47] in Isabelle. The project made full use of the flexibility of higher-order logic and later extended the verification to include higher-level security properties such as integrity and information flow noninterference [60], as well as verification down to the compiled binary of the kernel [84].

Abstract Interpretation

In [Chapter 10](#) we saw a number of automatic program analyses, and each one was hand crafted. Abstract Interpretation is a generic approach to automatic program analysis. In principle, it covers all of our earlier analyses. In this chapter we ignore the optimizing program transformations that accompany certain analyses.

The specific form of abstract interpretation we consider aims to compute the possible values of all variables at each program point. In order to infer this information the program is interpreted with abstract values that represent sets of concrete values, for example, with intervals instead of integers.

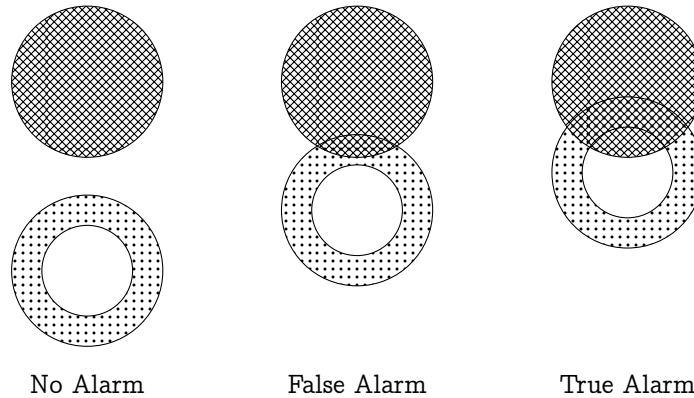
Program analyses are necessarily imprecise because they are automatic. We already discussed this at the beginning of [Chapter 10](#). For analyses that compute if some program point is reachable, which the analyses in this chapter do, this is particularly obvious: the end of the program is reachable iff the program terminates for some input, but termination is undecidable.

Therefore a program analysis should *overapproximate* to be on the safe side: it should compute a superset of the possible values that the variables can take on at each program point.

- If the analysis says that some value cannot arise, this is definitely the case.
- But if the analysis says that some value can arise, this is only potentially the case.

That is, if the analysis says that the variables have no possible value, which means that the program point is unreachable, then it really is unreachable, and the code at that point can safely be removed. But the analysis may claim some points are reachable which in fact are not, thus missing opportunities for code removal. Similarly for constant folding: if the analysis says that x always has value 2 at some point, we can safely replace x by 2 at that point. On the other hand, if the analysis says that x could also have 3 at that point, constant folding is out, although x may in fact only have one possible value.

If we switch from an optimization to a debugging or verification point of view, we think of certain states as erroneous (e.g., where $x = 0$) because they lead to a runtime error (e.g., division by x) or violate some given specification (e.g., $x > 0$). The analysis is meant to find if there is any program point where the reachable states include erroneous states. The following figure shows three ways in which the analysis can behave. The upper circle is the set of erroneous states, the white circle the set of reachable states, and its dotted halo is the superset of the reachable states computed by the analysis, all for one fixed program point.



The term “alarm” describes the situation where the analysis finds an erroneous state, in which case it raises an alarm, typically by flagging the program point in question. In the No Alarm situation, the analysis does not find any erroneous states and everybody is happy. In the True Alarm situation, the analysis finds erroneous states and some reachable states are indeed erroneous. But due to overapproximation, there are also so-called **false alarms**: the analysis finds an erroneous state that cannot arise at this program point. False alarms are the bane of all program analyses. They force the programmer to convince himself that the potential error found is not a real error but merely a weakness of the analysis.

13.1 Informal Introduction

At the centre of our approach to abstract interpretation are **annotated commands**. These are commands interspersed with annotations containing semantic information. This is reminiscent of Hoare logic and we also borrow the $\{ \dots \}$ syntax for annotations. However, annotations may now be placed at all intermediate program points, not just in front of loops as invariants. Loops will be annotated like this:

$$\begin{array}{l} \{I\} \\ \text{WHILE } b \text{ DO } \{P\} \ c \\ \{Q\} \end{array}$$

where I (as in “invariant”) annotates the loop head, P annotates the point before the loop body c , and Q annotates the exit of the whole loop. The annotation points are best visualized by means of the control-flow graph in Figure 13.1. We introduced control-flow graphs in Section 10.3. In fact, Fig-

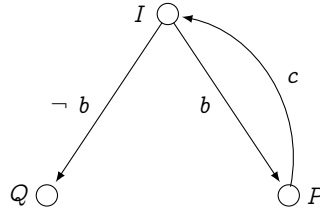


Fig. 13.1. Control-flow graph for $\{I\} \text{ WHILE } b \text{ DO } \{P\} \ c \ \{Q\}$

ure 13.1 is the same as Figure 10.6, except that we now label nodes not with sets of live variables but with the annotations at the corresponding program points. Edges labelled by compound commands stand for whole subgraphs.

13.1.1 Collecting Semantics

Before we come to the actual program analysis we need a semantics to justify it against. This semantics needs to express for each program point the set of states that can arise at this point during an execution. It is known as a **collecting semantics** because it collects together all the states that can arise at each point. Neither the big nor the small-step semantics express this information directly. Here is an example of a program annotated with its collecting semantics:

```

x := 0  {{<x := 0>}} ;
{{<x := 0>, <x := 2>, <x := 4>}}
WHILE x < 3
DO {{<x := 0>, <x := 2>}}
    x := x+2  {{<x := 2>, <x := 4>}}
{{<x := 4>}}
  
```

Annotations are of the form $\{\{...\}\}$ because the object inside the outer annotation braces is a set. Computing the annotations is an iterative process that we explain later.

Annotations can also be infinite sets of states:

```
{ {..., <x := -1>, <x := 0>, <x := 1>, ... } }
WHILE x < 3
DO { { {..., <x := 1>, <x := 2> } }
    x := x+2 { { {..., <x := 3>, <x := 4> } }
{ { <x := 3>, <x := 4>, ... } }
```

And programs can have multiple variables:

```
x := 0; y := 0 { { <x := 0, y := 0> } } ;
{ { <x := 0, y := 0>, <x := 2, y := 1>, <x := 4, y := 2> } }
WHILE x < 3
DO { { <x := 0, y := 0>, <x := 2, y := 1> } }
    x := x+2; y := y+1 { { <x := 2, y := 1>, <x := 4, y := 2> } }
{ { <x := 4, y := 2> } }
```

13.1.2 Abstract Interpretation

Now we move from the semantics to abstract interpretation in two steps. First we replace sets of states with single states that map variables to sets of values:

$(vname \Rightarrow val) \text{ set}$ becomes $vname \Rightarrow val \text{ set}$.

Our first example above now looks much simpler:

```
x := 0 { <x := {0}> } ;
{x := {0, 2, 4}>}
WHILE x < 3
DO { <x := {0, 2}> }
    x := x+2 { <x := {2, 4}> }
{ <x := {4}> }
```

However, this simplification comes at a price: it is an overapproximation that loses relationships between variables. For example, $\{ \langle x := 0, y := 0 \rangle, \langle x := 1, y := 1 \rangle \}$ is overapproximated by $\langle x := \{0, 1\}, y := \{0, 1\} \rangle$. The latter also subsumes $\langle x := 0, y := 1 \rangle$ and $\langle x := 1, y := 0 \rangle$.

In the second step we replace sets of values by “abstract values”. This step is domain specific. For example, we can approximate sets of integers by intervals. For the above example we obtain the following consistent annotation with integer intervals (written $[low, high]$):

```
x := 0 { <x := [0, 0]> } ;
{ <x := [0, 4]> }
WHILE x < 3
DO { <x := [0, 2]> }
```


$$x := x+2 \{<x := [2, 4]>\}$$

$$\{<x := [3, 4]>\}$$

Clearly, we have lost some precision in this step, but the annotations have become finitely representable: we have replaced arbitrary and potentially infinite sets by intervals, which are simply pairs of numbers.

How are the annotations actually computed? We start from an unannotated program and iterate abstract execution of the program (on intervals) until the annotations stabilize. Each execution step is a simultaneous execution of all edges of the control-flow graph. You can also think of it as a synchronous circuit where in each step simultaneously all units of the circuit process their input and make it their new output.

To demonstrate this iteration process we take the example above and give the annotations names:

```

x := 0 {A0} ;
{A1}
WHILE x < 3
DO {A2} x := x+2 {A3}
{A4}
```

In a separate table we can see how the annotations change with each step.

	0	1	2	3	4	5	6	7	8	9
A ₀	⊥	[0, 0]								[0, 0]
A ₁	⊥		[0, 0]			[0, 2]			[0, 4]	[0, 4]
A ₂	⊥			[0, 0]			[0, 2]			[0, 2]
A ₃	⊥				[2, 2]			[2, 4]		[2, 4]
A ₄	⊥									[3, 4]

Instead of the full annotation $<x := ivl>$, the table merely shows ivl . Column 0 shows the initial annotations where \perp represents the empty interval. Unchanged entries are left blank. In steps 1–3, $[0, 0]$ is merely propagated around. In step 4, 2 is added to it, making it $[2, 2]$. The crucial step is from 4 to 5: $[0, 0]$, the invariant A_1 in the making, is combined with the annotation $[2, 2]$ at the end of the loop body, telling us that the value of x at A_1 can in fact be any value from the union $[0, 0]$ and $[2, 2]$. This is overapproximated by the interval $[0, 2]$. The same thing happens again in step 8, where the invariant becomes $[0, 4]$. In step 9, the final annotation A_4 is reached at last: the invariant is now $[0, 4]$ and intersecting it with the negation of $x < 3$ lets $[3, 4]$ reach the end of the loop. The annotations reached in step 9 (which are displayed in full) are stable: performing another step leaves them unchanged.

This is the end of our informal introduction and we become formal again. First we define the type of annotated commands, then the collecting semantics, and finally abstract interpretation. Most of the chapter is dedicated to the stepwise development of a generic abstract interpreter whose precision is gradually increased.

The rest of this chapter builds on [Section 10.4.1](#).

13.2 Annotated Commands thy

The type of commands annotated with values of type $'a$ is called $'a\ acom$. Like com , $'a\ acom$ is defined as a datatype together with concrete syntax. The concrete syntax is described by the following grammar:

$$\begin{aligned}
 'a\ acom &::= SKIP\ \{ 'a \} \\
 &| \textit{string} ::= aexp\ \{ 'a \} \\
 &| 'a\ acom\ ;;\ 'a\ acom \\
 &| IF\ bexp\ THEN\ \{ 'a \}\ 'a\ acom\ ELSE\ \{ 'a \}\ 'a\ acom \\
 &\quad \{ 'a \} \\
 &| \{ 'a \} \\
 &\quad WHILE\ bexp\ DO\ \{ 'a \}\ 'a\ acom \\
 &\quad \{ 'a \}
 \end{aligned}$$

We exclusively use the concrete syntax and do not show the actual datatype.

Variables C , C_1 , C' , etc. will henceforth stand for annotated commands.

The layout of *IF* and *WHILE* is suggestive of the intended meaning of the annotations but has no logical significance. We have already discussed the annotations of *WHILE* but not yet of *IF*:

$$\begin{aligned}
 &IF\ b\ THEN\ \{P_1\}\ C_1\ ELSE\ \{P_2\}\ C_2 \\
 &\{Q\}
 \end{aligned}$$

Annotation P_i refers to the state before the execution of C_i . Annotation $\{Q\}$ is placed on a separate line to emphasize that it refers to the state after the execution of the whole conditional, not just the *ELSE* branch. The corresponding annotated control-flow graph is shown in [Figure 13.2](#).

Our annotated commands are polymorphic because, as we have already seen, we will want to annotate programs with different objects: with sets of states for the collecting semantics and with abstract states, for example, involving intervals, for abstract interpretation.

We now introduce a number of simple and repeatedly used auxiliary functions. Their functionality is straightforward and explained in words. In case of doubt, you can consult their full definitions in [Appendix A](#).

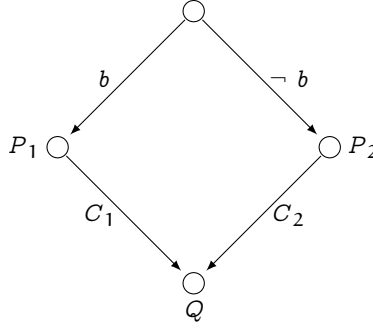


Fig. 13.2. Control-flow graph for $IF\ b\ THEN\ \{P_1\}\ C_1\ ELSE\ \{P_2\}\ C_2\{Q\}$

$strip :: 'a\ acom \Rightarrow com$
strips all annotations from an annotated command.

$annos :: 'a\ acom \Rightarrow 'a\ list$
extracts the list of all annotations from an annotated command, in left-to-right order.

$anno :: 'a\ acom \Rightarrow nat \Rightarrow 'a$
yields the n th annotation of an annotated command, starting at 0:
 $anno\ C\ p = annos\ C\ !\ p$
The standard infix operator $!$ indexes into a list with a natural number.

$post :: 'a\ acom \Rightarrow 'a$
returns the rightmost/last/post annotation of an annotated command:
 $post\ C = last\ (annos\ C)$

$annotate :: (nat \Rightarrow 'a) \Rightarrow com \Rightarrow 'a\ acom$
annotates a command: annotation number p (as counted by $anno$) is set to $f\ p$. The characteristic lemma is
 $p < asize\ c \implies anno\ (annotate\ f\ c)\ p = f\ p$
where $asize$ counts the number of annotation positions in a command.

$map_acom :: ('a \Rightarrow 'b) \Rightarrow 'a\ acom \Rightarrow 'b\ acom$
applies its first argument to every annotation of its second argument.

13.3 Collecting Semantics thy

The aim is to annotate commands with the set of states that can occur at each annotation point. The annotations are generated iteratively by a func-

tion *step* that maps annotated commands to annotated commands, updating the annotations. Each *step* executes all atomic commands (*SKIP*s and assignments) simultaneously and propagates the effects on the annotations forward. You can think of an annotated command as a synchronous circuit where with each clock tick (*step*) the information stored in each node (annotation point) is transformed by the outgoing edges and propagated to the successor nodes. Because we have placed (almost) no annotation points in front of commands, function *step* takes an additional argument, the set of states that are fed into the command. The full type of *step* and its recursive definition is shown in Figure 13.3. We will discuss the different cases one by one.

```

fun step :: state set  $\Rightarrow$  state set acom  $\Rightarrow$  state set acom where
step S (SKIP {Q}) = SKIP {S}
step S (x ::= e {Q}) = x ::= e {{s(x := aval e s) | s. s  $\in$  S}}
step S (C1;; C2) = step S C1;; step (post C1) C2
step S (IF b THEN {P1} C1 ELSE {P2} C2 {Q})
  = IF b THEN {{s  $\in$  S. bval b s}} step P1 C1
    ELSE {{s  $\in$  S.  $\neg$  bval b s}} step P2 C2
    {post C1  $\cup$  post C2}
step S ({I} WHILE b DO {P} C {Q})
  = {S  $\cup$  post C}
    WHILE b
    DO {{s  $\in$  I. bval b s}}
      step P C
    {{s  $\in$  I.  $\neg$  bval b s}}

```

Fig. 13.3. Definition of *step*

In the *SKIP* and the assignment case, the input set *S* is transformed and replaces the old post-annotation *Q*: for *SKIP* the transformation is the identity, $x ::= e$ transforms *S* by updating all of its elements (remember the set comprehension syntax explained in Section 4.2).

In the following let *S27* be the (somewhat arbitrary) state set $\{<x := 2>, <x := 7>\}$. It is merely used to illustrate the behaviour of *step* on the various constructs. Here is an example for assignment:

```

step S27 (x ::= Plus (V x) (N 1) {Q}) =
x ::= Plus (V x) (N 1) {{<x := 3>, <x := 8>}}

```

When applied to $C_1;; C_2$, *step* executes C_1 and C_2 simultaneously: the input to the execution of C_2 is the post-annotation of C_1 , not of *step S C₁*. For example:

$step\ S27\ (x ::= Plus\ (V\ x)\ (N\ 1)\ \{\{<x := 0>\}\});$
 $\quad x ::= Plus\ (V\ x)\ (N\ 2)\ \{Q\} =$
 $x ::= Plus\ (V\ x)\ (N\ 1)\ \{\{<x := 3>, <x := 8>\}\};$
 $x ::= Plus\ (V\ x)\ (N\ 2)\ \{\{<x := 2>\}\}$

On *IF* b *THEN* $\{P_1\}$ C_1 *ELSE* $\{P_2\}$ C_2 $\{_ \}$, *step S* does the following: it pushes filtered versions of S into P_1 and P_2 (this corresponds to the upper two edges in [Figure 13.2](#)), it executes C_1 and C_2 simultaneously (starting with the old annotations in front of them) and it updates the post-annotation with $post\ C_1 \cup post\ C_2$ (this corresponds to the lower two edges in [Figure 13.2](#)). Here is an example:

$step\ S27$
 $(IF\ Less\ (V\ x)\ (N\ 5)$
 $\quad THEN\ \{\{<x := 0>\}\}\ x ::= Plus\ (V\ x)\ (N\ 1)\ \{\{<x := 3>\}\}$
 $\quad ELSE\ \{\{<x := 5>\}\}\ x ::= Plus\ (V\ x)\ (N\ 2)\ \{\{<x := 9>\}\}$
 $\quad \{Q\}) =$
 $IF\ Less\ (V\ x)\ (N\ 5)$
 $\quad THEN\ \{\{<x := 2>\}\}\ x ::= Plus\ (V\ x)\ (N\ 1)\ \{\{<x := 1>\}\}$
 $\quad ELSE\ \{\{<x := 7>\}\}\ x ::= Plus\ (V\ x)\ (N\ 1)\ \{\{<x := 7>\}\}$
 $\quad \{\{<x := 3>, <x := 9>\}\}$

Finally we look at the *WHILE* case. It is similar to the conditional but feeds the post-annotation of the body of the loop back into the head of the loop. Here is an example:

$step\ \{<x := 7>\}\ (\{\{<x := 2>, <x := 5>\}\})$
 $\quad WHILE\ Less\ (V\ x)\ (N\ 5)$
 $\quad DO\ \{\{<x := 1>\}\}$
 $\quad \quad x ::= Plus\ (V\ x)\ (N\ 2)\ \{\{<x := 4>\}\}$
 $\quad \quad \{Q\}) =$
 $\quad \{\{<x := 4>, <x := 7>\}\}$
 $\quad WHILE\ Less\ (V\ x)\ (N\ 5)$
 $\quad DO\ \{\{<x := 2>\}\}$
 $\quad \quad x ::= Plus\ (V\ x)\ (N\ 2)\ \{\{<x := 3>\}\}$
 $\quad \quad \{\{<x := 5>\}\}$

The collecting semantics is now defined by iterating *step S* where S is some fixed set of initial states, typically all possible states, i.e., *UNIV*. The iteration starts with a command that is annotated everywhere with the empty set of states because no state has reached an annotation point yet. We illustrate the process with the example program from [Section 13.1](#):

```

x := 0 {A0} ;
{A1}
WHILE x < 3
DO {A2} x := x+2 {A3}
{A4}

```

In a separate table we can see how the annotations change with each iteration of *step S* (where *S* is irrelevant as long as it is not empty).

	0	1	2	3	4	5	6	7	8	9
A ₀	{}	{0}								{0}
A ₁	{}		{0}			{0,2}			{0,2,4}	{0,2,4}
A ₂	{}			{0}			{0,2}			{0,2}
A ₃	{}				{2}			{2,4}		{2,4}
A ₄	{}									{4}

Instead of the full annotation $\{<x := i_1>, <x:=i_2>, \dots\}$, the table merely shows $\{i_1, i_2, \dots\}$. Unchanged entries are left blank. In steps 1–3, $\{0\}$ is merely propagated around. In step 4, 2 is added to it, making it $\{2\}$. The crucial step is from 4 to 5: $\{0\}$, the invariant A_1 in the making, is combined with the annotation $\{2\}$ at the end of the loop body, yielding $\{0,2\}$. The same thing happens again in step 8, where the invariant becomes $\{0,2,4\}$. In step 9, the final annotation A_4 is reached at last: intersecting the invariant $\{0,2,4\}$ with the negation of $x < 3$ lets $\{4\}$ reach the exit of the loop. The annotations reached in step 9 (which are displayed in full) are stable: performing another step leaves them unchanged.

In contrast to the interval analysis in [Section 13.1](#), the semantics is and has to be exact. As a result, it is not computable in general. The above example is particularly simple in a number of respects that are all interrelated: the initial set S plays no role because x is initialized, all annotations are finite, and we reach a fixpoint after a finite number of steps. Most of the time we will not be so lucky.

13.3.1 Executing *step* in Isabelle

If we only deal with finite sets of states, we can let Isabelle execute *step* for us. Of course we again have to print states explicitly because they are functions. This is encapsulated in the function

```
show_acom :: state set acom  $\Rightarrow$  (vname  $\times$  val)set set acom
```

that turns a state into a set of variable-value pairs. We reconsider the example program above, but now in full Isabelle syntax:

```

definition cex :: com where cex =
  ''x'' ::= N 0;;
  WHILE Less (V ''x'') (N 3)
  DO ''x'' ::= Plus (V ''x'') (N 2)

```

```

definition Cex :: state set acom where Cex = annotate ( $\lambda p.$  {}) cex

```

```

value show_acom (step {<>} Cex)

```

yields

```

''x'' ::= N 0 {{{(''x'', 0)}}};;
{}
WHILE Less (V ''x'') (N 3)
DO {}
  ''x'' ::= Plus (V ''x'') (N 2) {}
{}

```

The triply nested braces are the combination of the outer annotation braces with annotations that are sets of sets of variable-value pairs, the result of converting sets of states into printable form.

You can iterate *step* by means of the function iteration operator $f^{^^n}$ (pretty-printed as f^n). For example, executing four steps

```

value show_acom ((step {<>} ^ 4) Cex)

```

yields

```

''x'' ::= N 0 {{{(''x'', 0)}}};;
{{{(''x'', 0)}}}
WHILE Less (V ''x'') (N 3)
DO {{{(''x'', 0)}}}
  ''x'' ::= Plus (V ''x'') (N 2) {{{(''x'', 2)}}}
{}

```

Iterating *step* {<>} nine times yields the fixpoint shown in the table above:

```

''x'' ::= N 0 {{{(''x'', 0)}}};;
{{{(''x'', 0)}, {''x'', 2)}, {''x'', 4)}}}
WHILE Less (V ''x'') (N 3)
DO {{{(''x'', 0)}, {''x'', 2)}}}
  ''x'' ::= Plus (V ''x'') (N 2) {{{(''x'', 2)}, {''x'', 4)}}}
{{{(''x'', 4)}}}

```

13.3.2 Collecting Semantics as Least Fixpoint

The collecting semantics is a fixpoint of *step* because a fixpoint describes an annotated command where the annotations are consistent with the execution via *step*. This is a fixpoint

```

 $x ::= N\ 0\ \{\{<x := 0>\}\};;$ 
 $\{\{<x := 0>, <x := 2>, <x := 4>\}\}$ 
 $WHILE\ Less\ (V\ x)\ (N\ 3)$ 
 $DO\ \{\{<x := 0>, <x := 2>\}\}$ 
 $\quad x ::= Plus\ (V\ x)\ (N\ 2)\ \{\{<x := 2>, <x := 4>\}\}$ 
 $\{\{<x := 4>\}\}$ 

```

because (in control-flow graph terminology) each node is annotated with the transformed annotation of the predecessor node. For example, the assignment $x ::= Plus\ (V\ x)\ (N\ 2)$ transforms $\{<x := 0>, <x := 2>\}$ into $\{<x := 2>, <x := 4>\}$. In case there are multiple predecessors, the annotation is the union of the transformed annotations of all predecessor nodes: $\{<x := 0>, <x := 2>, <x := 4>\} = \{<x := 0>\} \cup \{<x := 2>, <x := 4>\}$ (there is no transformation).

We can also view a fixpoint as a solution of not just the single equation $step\ S\ C = C$ but of a system of equations, one for each annotation. If $C = \{I\}\ WHILE\ b\ DO\ \{P\}\ C_0\ \{Q\}$ then the equation system is

```

 $I = S \cup post\ C_0$ 
 $P = \{s \in I. bval\ b\ s\}$ 
 $Q = \{s \in I. \neg bval\ b\ s\}$ 

```

together with the equations arising from $C_0 = step\ P\ C_0$. Iterating *step* is one way of solving this equation system. It corresponds to Jacobi iteration in linear algebra, where the new values for the unknowns are computed simultaneously from the old values. In principle, any method for solving an equation system can be used.

In general, *step* can have multiple fixpoints. For example

```

 $\{I\}$ 
 $WHILE\ Bc\ True$ 
 $DO\ \{I\}$ 
 $\quad SKIP\ \{I\}$ 
 $\{\{\}\}$ 

```

is a fixpoint of *step* $\{\}$ for every I . But only $I = \{\}$ makes computational sense: *step* $\{\}$ means that the empty set of states is fed into the execution, and hence no state can ever reach any point in the program. This happens to be the least fixpoint (w.r.t. \subseteq). We will now show that *step* always has a least fixpoint.

At the end of this section we prove that the least fixpoint is consistent with the big-step semantics.

13.3.3 Complete Lattices and Least Fixpoints thy

The Knaster-Tarski fixpoint theorem ([Theorem 10.29](#)) told us that monotone functions on sets have least pre-fixpoints, which are least fixpoints. Unfortunately *step* acts on annotated commands, not sets. Fortunately the theorem easily generalizes from sets to complete lattices, and annotated commands form complete lattices.

Definition 13.1. *A type $'a$ with a partial order \leq is a complete lattice if every set $S :: 'a$ set has a greatest lower bound $\bigwedge S :: 'a$:*

- $\forall s \in S. \bigwedge S \leq s$
- If $\forall s \in S. l' \leq s$ then $l' \leq \bigwedge S$

The greatest lower bound of S is often called the infimum.

Note that in a complete lattice, the ordering determines the infimum uniquely: if l_1 and l_2 are infima of S then $l_1 \leq l_2$ and $l_2 \leq l_1$ and thus $l_1 = l_2$.

The archetypal complete lattice is the powerset lattice: for any type $'a$, its powerset, type $'a$ set, is a complete lattice where \leq is \subseteq and \bigwedge is \bigcap . This works because $\bigcap M$ is the greatest set below (w.r.t. \subseteq) all sets in M .

Lemma 13.2. *In a complete lattice, every set S of elements also has a least upper bound (supremum) $\bigvee S$:*

- $\forall s \in S. s \leq \bigvee S$
- If $\forall s \in S. s \leq u$ then $\bigvee S \leq u$

The least upper bound is the greatest lower bound of all upper bounds: $\bigvee S = \bigwedge \{u. \forall s \in S. s \leq u\}$.

The proof is left as an exercise. Thus complete lattices can be defined via the existence of all infima or all suprema or both.

It follows that a complete lattice does not only have a least element $\bigwedge UNIV$, the infimum of all elements, but also a greatest element $\bigvee UNIV$. They are denoted by \perp and \top (pronounced “bottom” and “top”). [Figure 13.4](#) shows a typical complete lattice and motivates the name “lattice”.

The generalization of the Knaster-Tarski fixpoint theorem ([Theorem 10.29](#)) from sets to complete lattices is straightforward:

Theorem 13.3 (Knaster-Tarski¹). *Every monotone function f on a complete lattice has the least (pre-)fixpoint $\bigwedge \{p. f p \leq p\}$.*

¹ Tarski [85] actually proved that the set of fixpoints of a monotone function on a complete lattice is itself a complete lattice.

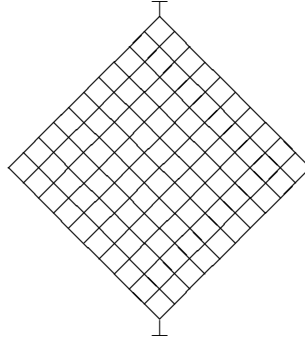


Fig. 13.4. Typical complete lattice

The proof is the same as for the set version but with \leq and \sqcap instead of \subseteq and \cap . This works because $\sqcap M$ is the greatest lower bound of M w.r.t. \subseteq .

13.3.4 Annotated Commands as a Complete Lattice

In order to apply Knaster-Tarski to *step* we need to turn annotated commands into a complete lattice. Given an ordering \leq on $'a$, it can be lifted to $'a\text{ acom}$ in a pointwise manner:

$$\begin{aligned} (C_1 \leq C_2) = & \\ (\text{strip } C_1 = \text{strip } C_2 \wedge & \\ (\forall p < \text{length } (\text{annos } C_1). \text{ anno } C_1\ p \leq \text{ anno } C_2\ p)) & \end{aligned}$$

Two annotated commands C_1 and C_2 are only comparable if they are **structurally equal**, i.e., if $\text{strip } C_1 = \text{strip } C_2$. For example, $x ::= e \{\{a\}\} \leq x ::= e \{\{a, b\}\}$ is *True* and both $x ::= e \{\{a\}\} \leq x ::= e \{\{\}\}$ and $x ::= N\ 0\ \{S\} \leq x ::= N\ 1\ \{S\}$ are *False*.

Lemma 13.4. *If $'a$ is a partial order, so is $'a\text{ acom}$.*

Proof. The required properties carry over pointwise from $'a$ to $'a\text{ acom}$. \square

Because (in Isabelle) \subseteq is just special syntax for \leq on sets, the above definition lifts \subseteq to a partial order \leq on *state set acom*. Unfortunately this order does not turn *state set acom* into a complete lattice: although $\text{SKIP } \{S\}$ and $\text{SKIP } \{T\}$ have the infimum $\text{SKIP } \{S \cap T\}$, the two commands $\text{SKIP } \{S\}$ and $x ::= e \{T\}$ have no lower bound, let alone an infimum. The fact is:

Only structurally equal annotated commands have an infimum.

Below we show that for each $c :: \text{com}$ the set $\{C :: 'a\text{ acom}. \text{strip } C = c\}$ of all annotated commands structurally equal to c is a complete lattice,

assuming that $'a$ is one. Thus we need to generalize complete lattices from types to sets.

Definition 13.5. *Let $'a$ be a partially ordered type. A set $L :: 'a$ set is a complete lattice if every $M \subseteq L$ has a greatest lower bound $\bigcap M \in L$.*

The type-based definition corresponds to the special case $L = UNIV$.

In the following we write $f \in A \rightarrow B$ as a shorthand for $\forall a \in A. f a \in B$. This notation is a set-based analogue of $f :: \tau_1 \Rightarrow \tau_2$.

Theorem 13.6 (Knaster-Tarski). *Let $L :: 'a$ set be a complete lattice and $f \in L \rightarrow L$ a monotone function. On L , f has the least (pre-)fixpoint*

$$lfp f = \bigcap \{p \in L. f p \leq p\}$$

The proof is the same as before, but we have to keep track of the fact that we always stay within L , which is ensured by $f \in L \rightarrow L$ and $M \subseteq L \implies \bigcap M \in L$.

To apply Knaster-Tarski to *state set acom* we need to show that *acom* transforms complete lattices into complete lattices as follows:

Theorem 13.7. *Let $'a$ be a complete lattice and $c :: com$. Then the set $L = \{C :: 'a\ acom. strip C = c\}$ is a complete lattice.*

Proof. The infimum of a set of annotated commands is computed pointwise for each annotation:

$$Inf_acom\ c\ M = annotate\ (\lambda p. \bigcap \{anno\ C\ p \mid C. C \in M\})\ c$$

We have made explicit the dependence on the command c from the statement of the theorem. The infimum properties of *Inf_acom* follow easily from the corresponding properties of \bigcap . The proofs are automatic. \square

13.3.5 Collecting Semantics

Because type *state set* is a complete lattice, [Theorem 13.7](#) implies that $L = \{C :: state\ set\ acom. strip\ C = c\}$ is a complete lattice too, for any c . It is easy to prove ([Lemma 13.26](#)) that *step* $S \in L \rightarrow L$ is monotone: $C_1 \leq C_2 \implies step\ S\ C_1 \leq step\ S\ C_2$. Therefore Knaster-Tarski tells us that *lfp* returns the least fixpoint and we can define the collecting semantics at last:

$CS :: com \Rightarrow state\ set\ acom$ $CS\ c = lfp\ c\ (step\ UNIV)$

The extra argument c of lfp comes from the fact that lfp is defined in [Theorem 13.6](#) in the context of a set L and our concrete L depends on c . The *UNIV* argument of $step$ expresses that execution may start with any initial state; other choices are possible too.

Function CS is not executable because the resulting annotations are usually infinite state sets. But just as for true liveness in [Section 10.4.2](#), we can approximate (and sometimes reach) the lfp by iterating $step$. We already showed how that works in [Section 13.3.1](#).

In contrast to previous operational semantics that were “obviously right”, the collecting semantics is more complicated and less intuitive. In case you still have some nagging doubts that it is defined correctly, the following lemma should help. Remember that $post$ extracts the final annotation ([Appendix A](#)).

Lemma 13.8. $(c, s) \Rightarrow t \implies t \in post (CS\ c)$

Proof. By definition of CS the claim follows directly from $\llbracket (c, s) \Rightarrow t; s \in S \rrbracket \implies t \in post(lfp\ c\ (step\ S))$, which in turn follows easily from two auxiliary propositions:

$$post(lfp\ c\ f) = \bigcap \{post\ C \mid C. strip\ C = c \wedge f\ C \leq C\}$$

$$\llbracket (c, s) \Rightarrow t; strip\ C = c; s \in S; step\ S\ C \leq C \rrbracket \implies t \in post\ C$$

The first proposition is a direct consequence of $\forall C \in M. strip\ C = c \implies post\ (Inf_acom\ c\ M) = \bigcap (post\ 'M)$, which follows easily from the definition of $post$ and Inf_acom . The second proposition is the key. It is proved by rule induction. Most cases are automatic, but *WhileTrue* needs some work. \square

Thus we know that the collecting semantics overapproximates the big-step semantics. Later we show that the abstract interpreter overapproximates the collecting semantics. Therefore we can view the collecting semantics merely as a stepping stone for proving that the abstract interpreter overapproximates the big-step semantics, our standard point of reference.

One can in fact show that both semantics are equivalent. One can also refine the lemma: it only talks about the $post$ annotation but one would like to know that all annotations are correct w.r.t. the big-step semantics. We do not pursue this further but move on to the main topic of this chapter, abstract interpretation.

Exercises

The exercises below are conceptual and should be done on paper, also because many of them require Isabelle material that will only be introduced later.

Exercise 13.1. Show the iterative computation of the collecting semantics of the following program in a table like the one on page [228](#).

```

x := 0; y := 2 {A0} ;
{A1}
WHILE 0 < y
DO {A2} ( x := x+y; y := y - 1 {A3} )
{A4}

```

Note that two annotations have been suppressed to make the task less tedious. You do not need to show steps where only the suppressed annotations change.

Exercise 13.2. Extend type *acom* and function *step* with a construct *OR* for the nondeterministic choice between two commands (see [Exercise 7.9](#)). Hint: think of *OR* as a nondeterministic conditional without a test.

Exercise 13.3. Prove that in a complete lattice $\bigsqcup S = \bigsqcap \{u. \forall s \in S. s \leq u\}$ is the least upper bound of *S*.

Exercise 13.4. Where is the mistake in the following argument? The natural numbers form a complete lattice because any set of natural numbers has an infimum, its least element.

Exercise 13.5. Show that the integers extended with ∞ and $-\infty$ form a complete lattice.

Exercise 13.6. Prove the following slightly generalized form of the Knaster-Tarski pre-fixpoint theorem: If *P* is a set of pre-fixpoints of a monotone function on a complete lattice, then $\bigsqcap P$ is a pre-fixpoint too. In other words, the set of pre-fixpoints of a monotone function on a complete lattice is a complete lattice.

Exercise 13.7. According to [Lemma 10.28](#), least pre-fixpoints of monotone functions are also least fixpoints.

1. Show that leastness matters: find a (small!) partial order with a monotone function that has a pre-fixpoint that is not a fixpoint.
2. Show that the reverse implication does not hold: find a partial order with a monotone function that has a least fixpoint that is not a least pre-fixpoint.

Exercise 13.8. The term *collecting* semantics suggests that the reachable states are collected in the following sense: *step* should not transform $\{S\}$ into $\{S'\}$ but into $\{S \cup S'\}$. Show that this makes no difference. That is, prove that if *f* is a monotone function on sets, then *f* has the same least fixpoint as $\lambda S. S \cup f S$.

13.4 Abstract Values thy

The topic of this section is the abstraction from the state sets in the collecting semantics to some type of abstract values. In [Section 13.1.2](#) we had already discussed a first abstraction of state sets

$$(vname \Rightarrow val) \text{ set} \rightsquigarrow vname \Rightarrow val \text{ set}$$

and that it constitutes a first overapproximation. There are so-called relational analyses that avoid this abstraction, but they are more complicated. In a second step, sets of values are abstracted to **abstract values**:

$$val \text{ set} \rightsquigarrow \text{abstract domain}$$

where the **abstract domain** is some type of abstract values that we can compute on. What exactly that type is depends on the analysis. Interval analysis is one example, parity analysis is another. Parity analysis determines if the value of a variable at some point is always even, is always odd, or can be either. It is based on the following abstract domain

datatype *parity* = *Even* | *Odd* | *Either*

that will serve as our running example.

Abstract values represent sets of concrete values *val* and need to come with an ordering that is an abstraction of the subset ordering. [Figure 13.5](#) shows

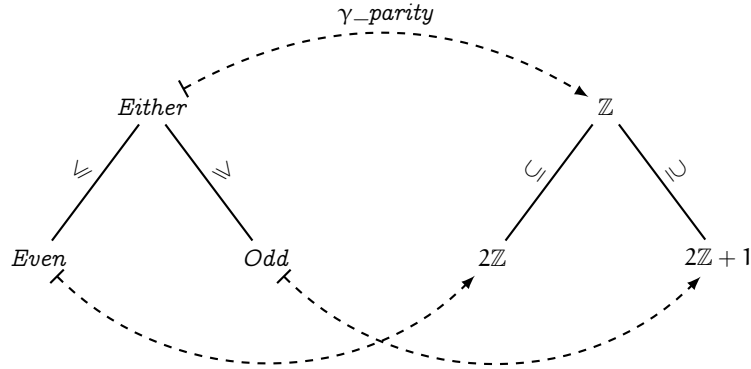


Fig. 13.5. Parity domain and its concretisation

the abstract values on the left and the concrete ones on the right, where $2\mathbb{Z}$ and $2\mathbb{Z} + 1$ are the even and the odd integers. The solid lines represent the orderings on the two types. The dashed arrows represent the concretisation function:

A **concretisation function** maps an abstract value to a set of concrete values.

Concretisation functions will always be named γ , possibly with a suffix. This is the Isabelle definition of the one for *parity*:

```
fun  $\gamma\_parity :: parity \Rightarrow val set$  where
 $\gamma\_parity\ Even = \{i. i \bmod 2 = 0\}$ 
 $\gamma\_parity\ Odd = \{i. i \bmod 2 = 1\}$ 
 $\gamma\_parity\ Either = UNIV$ 
```

The definition of \leq on *parity* is even simpler:

$$(x \leq y) = (y = Either \vee x = y)$$

Of course there should be a relationship between the two:

Bigger abstract values represent bigger sets of concrete values.

That is, concretisation functions should be monotone. This is obviously the case in Figure 13.5. Moreover, \leq should be a partial order. It is easily proved that \leq on *parity* is a partial order.

We want \leq to be a partial order because it abstracts the partial order \subseteq . But we will need more. The collecting semantics uses \cup to join the state sets where two computation paths meet, e.g., after an *IF-THEN-ELSE*. Therefore we require that the abstract domain also provides an operation \sqcup that behaves like \cup . That makes it a semilattice:

Definition 13.9. A type $'a$ is a *semilattice* if it is a partial order and there is a *supremum* operation \sqcup of type $'a \Rightarrow 'a \Rightarrow 'a$ that returns the least upper bound of its arguments:

- Upper bound: $x \leq x \sqcup y$ and $y \leq x \sqcup y$
- Least: $\llbracket x \leq z; y \leq z \rrbracket \implies x \sqcup y \leq z$

The supremum is also called the **join** operation.

A very useful consequence of the semilattice axioms is the following:

Lemma 13.10. $x \sqcup y \leq z \iff x \leq z \wedge y \leq z$

The proof is left as an exercise.

In addition, we will need an abstract value that corresponds to *UNIV*.

Definition 13.11. A partial order has a *top element* \top if $x \leq \top$ for all x .

Thus we will require our abstract domain to be a semilattice with top element, or semilattice with \top for short.

Of course, type *parity* is a semilattice with \top :

$$x \sqcup y = (\text{if } x = y \text{ then } x \text{ else } \textit{Either})$$

$$\top = \textit{Either}$$

The proof that the semilattice and top axioms are satisfied is routine.

13.4.1 Type Classes

We will now sketch how abstract concepts like semilattices are modelled with the help of Isabelle's type classes. A **type class** has a name and is defined by

- a set of required functions, the interface, and
- a set of axioms about those functions.

For example, a partial order requires that there is a function \leq with certain properties. The type classes introduced above are called *order*, *semilattice_sup*, *top*, and *semilattice_sup_top*.

To show that a type τ belongs to some class C we have to

- define all functions in the interface of C on type τ
- and prove that they satisfy the axioms of C .

This process is called **instantiating** C with τ . For example, above we have instantiated *semilattice_sup_top* with *parity*. Informally we say that *parity* is a semilattice with \top .

Note that the function definitions made when instantiating a class are unlike normal function definitions because we define an already existing but overloaded function (e.g., \leq) for a new type (e.g., *parity*).

The instantiation of *semilattice_sup_top* with *parity* was unconditional. There is also a conditional version exemplified by the following proposition:

Lemma 13.12. *If 'a and 'b are partial orders, so is 'a \times 'b.*

Proof. The ordering on 'a \times 'b is defined in terms of the orderings on 'a and 'b in the componentwise manner:

$$(x_1, x_2) \leq (y_1, y_2) \longleftrightarrow x_1 \leq y_1 \wedge x_2 \leq y_2$$

The proof that this yields a partial order is straightforward. \square

It may be necessary to restrict a type variable 'a to be in a particular class C . The notation is 'a :: C . Here are two examples:

```
definition is_true :: 'a::order  $\Rightarrow$  bool where is_true x = (x  $\leq$  x)
lemma is_true (x :: 'a::order)
```

If we drop the class constraints in the above definition and lemma statements, they fail because 'a comes without a class constraint but the given formulas

require one. If we drop the type annotations altogether, both the definition and the lemma statement work fine but in the definition the implicitly generated type variable $'a$ will be of class *ord*, where *ord* is a predefined superclass of *order* that merely requires an operation \leq without any axioms. This means that the lemma will not be provable.

Note that it suffices to constrain one occurrence of a type variable in a given type. The class constraint automatically propagates to the other occurrences of that type variable.

We do not describe the precise syntax for defining and instantiating classes. The semi-formal language used so far is perfectly adequate for our purpose. The interested reader is referred to the Isabelle theories (in particular *Abs_Intl_parity* to see how classes are instantiated) and to the tutorial on type classes [40]. If you are familiar with the programming language Haskell you will recognize that Isabelle provides Haskell-style type classes extended with axioms.

13.4.2 From Abstract Values to Abstract States thy

Let $'a$ be some abstract domain type. So far we had planned to abstract $(vname \Rightarrow val)$ *set* by the abstract state type

$$\text{type_synonym } 'a \text{ st} = vname \Rightarrow 'a$$

but there is a complication: the empty set of states has no meaningful abstraction. The empty state set is important because it arises at unreachable program points, and identifying the latter is of great interest for program optimization. Hence we abstract *state set* by

$$'a \text{ st option}$$

where *None* is the abstraction of $\{\}$. That is, the abstract interpretation will compute with and annotate programs with values of type $'a \text{ st option}$ instead of *state set*. We will now lift the semilattice structure from $'a$ to $'a \text{ st option}$. All the proofs in this subsection are routine and we leave most of them as exercises whose solutions can be found in the Isabelle theories.

Because states are functions we show as a first step that function spaces preserve semilattices:

Lemma 13.13. *If $'a$ is a semilattice with \top , so is $'b \Rightarrow 'a$, for every type $'b$.*

Proof. This class instantiation is already part of theory *Main*. It lifts \leq , \sqcup and \top from $'a$ to $'b \Rightarrow 'a$ in the canonical pointwise manner:

$$\begin{aligned}
f \leq g &= \forall x. f x \leq g x \\
f \sqcup g &= \lambda x. f x \sqcup g x \\
\top &= \lambda x. \top
\end{aligned}$$

It is easily shown that the result is a semilattice with \top if $'a$ is one. For example, $f \leq f \sqcup g$ holds because $f x \leq f x \sqcup g x = (f \sqcup g) x$ for all x . \square

Similarly, but for $'a \text{ option}$ instead of $'b \Rightarrow 'a$:

Lemma 13.14. *If $'a$ is a semilattice with \top , so is $'a \text{ option}$.*

Proof. Here is how \leq , \sqcup and \top are defined on $'a \text{ option}$:

$$\begin{aligned}
\text{Some } x \leq \text{Some } y &\longleftrightarrow x \leq y \\
\text{None} &\leq _ \longleftrightarrow \text{True} \\
\text{Some } _ &\leq \text{None} \longleftrightarrow \text{False} \\
\text{Some } x \sqcup \text{Some } y &= \text{Some } (x \sqcup y) \\
\text{None} \sqcup y &= y \\
x \sqcup \text{None} &= x \\
\top &= \text{Some } \top
\end{aligned}$$

Figure 13.6 shows an example of how the ordering is transformed by going from $'a$ to $'a \text{ option}$. The elements of $'a$ are wrapped up in *Some*

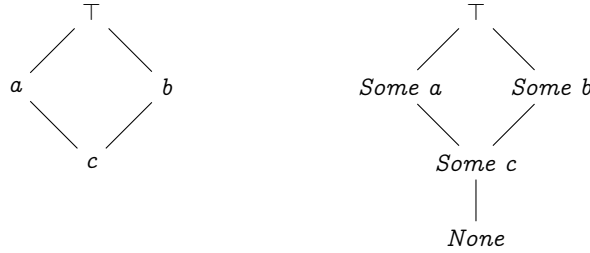


Fig. 13.6. From $'a$ to $'a \text{ option}$

without modifying the ordering between them and *None* is adjoined as the least element.

The semilattice properties of $'a \text{ option}$ are proved by exhaustive case analyses. As an example consider the proof of $x \leq x \sqcup y$. In each of the cases $x = \text{None}$, $y = \text{None}$ and $x = \text{Some } a \wedge y = \text{Some } b$ it holds by definition and because $a \leq a \sqcup b$ on type $'a$. \square

Together with Lemma 13.13 we have:

Corollary 13.15. *If $'a$ is a semilattice with \top , so is $'a$ st option.*

Now we lift the concretisation function γ from $'a$ to $'a$ st option, again separately for \Rightarrow and *option*. We can turn any concretisation function $\gamma :: 'a \Rightarrow 'c$ set into one from $'b \Rightarrow 'a$ to $('b \Rightarrow 'c)$ set:

definition $\gamma_fun :: ('a \Rightarrow 'c$ set) $\Rightarrow ('b \Rightarrow 'a) \Rightarrow ('b \Rightarrow 'c)$ set **where**
 $\gamma_fun \gamma F = \{f. \forall x. f x \in \gamma (F x)\}$

For example, $\gamma_fun \gamma_parity$ concretises $\lambda x::'b. Even$ to the set of all functions of type $'b \Rightarrow int$ whose ranges are even integers.

Lemma 13.16. *If γ is monotone, so is $\gamma_fun \gamma$.*

Lifting of γ to *option* achieves what we initially set out to do, namely to be able to abstract the empty set:

fun $\gamma_option :: ('a \Rightarrow 'c$ set) $\Rightarrow 'a$ option $\Rightarrow 'c$ set **where**
 $\gamma_option \gamma None = \{\}$
 $\gamma_option \gamma (Some a) = \gamma a$

Lemma 13.17. *If γ is monotone, so is $\gamma_option \gamma$.*

Applying monotone functions to all annotations of a command is also monotone:

Lemma 13.18. *If γ is monotone, so is $map_acom \gamma$.*

A useful property of monotone functions is the following:

Lemma 13.19. *If $\varphi :: 'a \Rightarrow 'b$ is a monotone function between two semilattices then $\varphi a_1 \sqcup \varphi a_2 \leq \varphi (a_1 \sqcup a_2)$.*

The proof is trivial: $\varphi a_1 \sqcup \varphi a_2 \leq \varphi (a_1 \sqcup a_2)$ iff $\varphi a_i \leq \varphi (a_1 \sqcup a_2)$ for $i = 1, 2$ (by [Lemma 13.10](#)), and the latter two follow by monotonicity from $a_i \leq a_1 \sqcup a_2$.

Exercises

Exercise 13.9. Prove [Lemma 13.10](#) on paper.

13.5 Generic Abstract Interpreter thy

In this section we define a first abstract interpreter which we refine in later sections. It is generic because it is parameterized with the abstract domain. It works like the collecting semantics but operates on the abstract domain instead of state sets. To bring out this similarity we first abstract the collecting semantics' *step* function.

13.5.1 Abstract *Step* Function thy

The *step* function of the collecting semantics (see Figure 13.3) can be generalized as follows.

- Replace *state set* by an arbitrary type $'a$ which must have a \sqcup operation (no semilattice required yet).
- Parameterize *step* with two functions

$$asem :: vname \Rightarrow aexp \Rightarrow 'a \Rightarrow 'a$$

$$bsem :: bexp \Rightarrow 'a \Rightarrow 'a$$

that factor out the actions of assignment and boolean expression on $'a$.

The result is the function *Step* displayed in Figure 13.7. Note that we have

```

fun Step :: 'a  $\Rightarrow$  'a acom  $\Rightarrow$  'a acom where
  Step a (SKIP { $\_$ }) = SKIP {a}
  Step a (x ::= e { $\_$ }) = x ::= e {asem x e a}
  Step a (C1;; C2) = Step a C1;; Step (post C1) C2
  Step a (IF b THEN {P1} C1 ELSE {P2} C2 {Q})
    = IF b THEN {bsem b a} Step P1 C1 ELSE {bsem (Not b) a} Step P2 C2
      {post C1  $\sqcup$  post C2}
  Step a ({I} WHILE b DO {P} C {Q})
    = {a  $\sqcup$  post C}
      WHILE b
      DO {bsem b I}
        Step P C
      {bsem (Not b) I}

```

Fig. 13.7. Definition of *Step*

suppressed the parameters *asem* and *bsem* of *Step* in the figure for better readability. In reality, they are there and *step* is defined like this:

$$step =$$

$$Step (\lambda x e S. \{s(x := aval e s) \mid s. s \in S\}) (\lambda b S. \{s \in S. bval b s\})$$

(This works because in Isabelle \cup is just nice syntax for \sqcup on sets.) The equations in Figure 13.3 are merely derived from this definition. This way we avoided having to explain the more abstract *Step* early on.

13.5.2 Abstract Interpreter Interface

The abstract interpreter is defined as a parameterized module (a *locale* in Isabelle; see [9] for details). Its parameters are the abstract domain together

with abstractions of all operations on the concrete type $val = int$. The result of the module is an abstract interpreter that operates on the given abstract domain.

In detail, the parameters and their required properties are the following:

Abstract domain A type $'av$ of abstract values.

Must be a semilattice with \top .

Concretisation function $\gamma :: 'av \Rightarrow val\ set$

Must be monotone: $a_1 \leq a_2 \implies \gamma\ a_1 \subseteq \gamma\ a_2$

Must preserve \top : $\gamma\ \top = UNIV$

Abstract arithmetic $num' :: val \Rightarrow 'av$

$plus' :: 'av \Rightarrow 'av \Rightarrow 'av$

Must establish and preserve the $i \in \gamma\ a$ relationship:

$i \in \gamma\ (num'\ i) \quad (num')$

$\llbracket i_1 \in \gamma\ a_1; i_2 \in \gamma\ a_2 \rrbracket \implies i_1 + i_2 \in \gamma\ (plus'\ a_1\ a_2) \quad (plus')$

Remarks:

- Every constructor of *aexp* (except *V*) must have a counterpart on $'av$.
- num' and $plus'$ abstract *N* and *Plus*.
- The requirement $i \in \gamma\ (num'\ i)$ could be replaced by $\gamma\ (num'\ i) = \{i\}$. We have chosen the weaker formulation to emphasize that all operations must establish or preserve $i \in \gamma\ a$.
- Functions whose names end with a prime usually operate on $'av$.
- Abstract values are usually called *a* whereas arithmetic expressions are usually called *e* now.

From γ we define three lifted concretisation functions:

$\gamma_s :: 'av\ st \Rightarrow state\ set$

$\gamma_s = \gamma_fun\ \gamma$

$\gamma_o :: 'av\ st\ option \Rightarrow state\ set$

$\gamma_o = \gamma_option\ \gamma_s$

$\gamma_c :: 'a\ st\ option\ acom \Rightarrow state\ set\ acom$

$\gamma_c = map_acom\ \gamma_o$

All of them are monotone (see the lemmas in [Section 13.4.2](#)).

Now we are ready for the actual abstract interpreter, first for arithmetic expressions, then for commands. Boolean expressions come in a later section.

13.5.3 Abstract Interpretation of Expressions

Abstract interpretation of *aexp* is unsurprising:

```

fun aval' :: aexp ⇒ 'av st ⇒ 'av where
aval' (N i) S      = num' i
aval' (V x) S      = S x
aval' (Plus e1 e2) S = plus' (aval' e1 S) (aval' e2 S)

```

The correctness lemma expresses that *aval'* overapproximates *aval*:

Lemma 13.20 (Correctness of *aval'*).

$s \in \gamma_s S \implies \text{aval } e \ s \in \gamma (\text{aval}' e \ S)$

Proof. By induction on *e* with the help of (*num'*) and (*plus'*). □

Example 13.21. We instantiate the interface to our abstract interpreter module with the *parity* domain described earlier:

```

γ      = γ_parity
num'   = num_parity
plus'  = plus_parity

```

where

```

num_parity i = (if i mod 2 = 0 then Even else Odd)

plus_parity Even Even = Even
plus_parity Odd Odd = Even
plus_parity Even Odd = Odd
plus_parity Odd Even = Odd
plus_parity x Either = Either
plus_parity Either y = Either

```

We had already discussed that *parity* is a semilattice with \top and γ_parity is monotone. Both $\gamma_parity \top = UNIV$ and (*num'*) are trivial, as is (*plus'*) after an exhaustive case analysis on a_1 and a_2 .

Let us call the result *aval'* of this module instantiation *aval_parity*. Then the following term evaluates to *Even*:

```

aval_parity (Plus (V ''x'') (V ''y'')) (λ_. Odd)

```

13.5.4 Abstract Interpretation of Commands

The abstract interpreter for commands is defined in two steps. First we instantiate *Step* to perform one interpretation step, later we iterate this *step'* until a pre-fixpoint is found.

```

step' :: 'av st option ⇒ 'av st option acom ⇒ 'av st option acom
step' = Step asem (λb S. S)

```

where

$$\begin{aligned} \text{asem } x \ e \ S = \\ (\text{case } S \text{ of } \text{None} \Rightarrow \text{None} \mid \text{Some } S \Rightarrow \text{Some } (S(x := \text{aval}' e \ S))) \end{aligned}$$

Remarks:

- From now on the identifier S will (almost) always be of type $'av \ st$ or $'av \ st \ option$.
- Function asem updates the abstract state with the abstract value of the expression. The rest is boiler-plate to handle the *option* type.
- Boolean expressions are not analysed at all. That is, they have no effect on the state. Compare $\lambda b \ S. S$ with $\lambda b \ S. \{s \in S. \text{bval } b \ s\}$ in the collecting semantics. The former constitutes a gross overapproximation to be refined later.

Example 13.22. We continue the previous example where we instantiated the abstract interpretation module with the *parity* domain. Let us call the result step' of this instantiation step_parity and consider the program on the left:

```
x := 3 {None} ;
{None}
WHILE ...
DO {None}
  x := x+2 {None}
{None}
```

Odd			
	Odd		
		Odd	
			Odd
		Odd	

In the table on the right we iterate step_parity , i.e., we see how the annotations in $(\text{step_parity } \top)^k C_0$ change with increasing k (where C_0 is the initial program with *Nones*, and by definition $\top = (\lambda_. \text{Either})$). Each row of the table refers to the program annotation in the same row. For compactness, a parity value p represents the state $\text{Some } (\top(x := p))$. We only show an entry if it differs from the previous step. After four steps, there are no more changes; we have reached the least fixpoint.

Exercise: What happens if the 2 in the program is replaced by a 1?

Correctness of step' means that it overapproximates step :

Corollary 13.23 (Correctness of step').

$$\text{step } (\gamma_o S) (\gamma_c C) \leq \gamma_c (\text{step}' S C)$$

Because step and step' are defined as instances of Step , this is a corollary of a general property of Step :

Lemma 13.24. $\text{Step } f \ g \ (\gamma_o S) (\gamma_c C) \leq \gamma_c (\text{Step } f' \ g' \ S \ C)$ if $f \ x \ e \ (\gamma_o S) \subseteq \gamma_o (f' \ x \ e \ S)$ and $g \ b \ (\gamma_o S) \subseteq \gamma_o (g' \ b \ S)$ for all x, e, b .

Proof. By induction on C . Assignments and boolean expressions are dealt with by the two assumptions. In the *IF* and *WHILE* cases one has to show properties of the form $\gamma_o S_1 \cup \gamma_o S_2 \subseteq \gamma_o (S_1 \sqcup S_2)$; they follow from monotonicity of γ_o by [Lemma 13.19](#). \square

The corollary follows directly from the lemma by unfolding the definitions of *step* and *step'*. The requirement $\{s(x := \text{aval } e \ s) \mid s. s \in \gamma_o S\} \subseteq \gamma_o (\text{asem } x \ e \ S)$ is a trivial consequence of the overapproximation of *aval* by *aval'*.

13.5.5 Abstract Interpreter

The abstract interpreter iterates *step'* with the help of a library function:

```
while_option :: ('a  $\Rightarrow$  bool)  $\Rightarrow$  ('a  $\Rightarrow$  'a)  $\Rightarrow$  'a  $\Rightarrow$  'a option
while_option b f x = (if b x then while_option b f (f x) else Some x)
```

The equation is an executable consequence of the (omitted) definition.

This is a generalization of the *while* combinator used in [Section 10.4.2](#) for the same purpose as now: iterating a function. The difference is that *while_option* returns an optional value to distinguish termination (*Some*) from nontermination (*None*). Of course the execution of *while_option* will simply not terminate if the mathematical result is *None*.

The abstract interpreter *AI* is a search for a pre-fixpoint of *step'* \top :

```
AI :: com  $\Rightarrow$  'av st option acom option
AI c = pfp (step'  $\top$ ) (bot c)
```

```
pfp :: ('a  $\Rightarrow$  'a)  $\Rightarrow$  'a  $\Rightarrow$  'a option
pfp f s = while_option ( $\lambda x. \neg f x \leq x$ ) f s
```

```
bot :: com  $\Rightarrow$  'a option acom
bot c = annotate ( $\lambda p. \text{None}$ ) c
```

Function *pfp* searches for a pre-fixpoint of *f* starting from *s*. In general, this search need not terminate, but if it does, it has obviously found a pre-fixpoint. This follows from the following property of *while_option*:

$$\text{while_option } b \ c \ s = \text{Some } t \implies \neg b \ t$$

Function *AI* starts the search from *bot c* where all annotations have been initialized with *None*. Because *None* is the least annotation, *bot c* is the least annotated program (among those structurally equal to *c*); hence the name.

Now we show partial correctness of *AI*: if abstract interpretation of *c* terminates and returns an annotated program *C*, then the concretisation of the annotations of *C* overapproximate the collecting semantics of *c*.

Lemma 13.25 (Partial correctness of AI w.r.t. CS).

$$AI\ c = Some\ C \implies CS\ c \leq \gamma_c\ C$$

Proof. By definition we may assume $pfp\ (step' \top)\ (bot\ c) = Some\ C$. By the above property of *while_option*, C is a pre-fixpoint: $step' \top C \leq C$. By monotonicity we have $\gamma_c\ (step' \top C) \leq \gamma_c\ C$ and because $step'$ overapproximates $step$ (Corollary 13.23) we obtain $step\ (\gamma_o \top)\ (\gamma_c\ C) \leq \gamma_c\ (step' \top C)$ and thus $\gamma_c\ C$ is a pre-fixpoint of $step\ (\gamma_o \top)$. Because CS is defined as the least pre-fixpoint of $step\ UNIV$ and because $\gamma_o \top = UNIV$ (this is trivial) we obtain $CS\ c \leq \gamma_c\ C$ as required. \square

13.5.6 Monotonicity

Although we know that if *pfp* terminates, by construction it yields a pre-fixpoint, we don't yet know under what conditions it terminates and which pre-fixpoint is found. This is where monotonicity of *Step* comes in:

Lemma 13.26 (Monotonicity of Step). *Let 'a be a semilattice and let $f :: vname \Rightarrow aexp \Rightarrow 'a \Rightarrow 'a$ and $g :: bexp \Rightarrow 'a \Rightarrow 'a$ be two monotone functions: $S_1 \leq S_2 \implies f\ x\ e\ S_1 \leq f\ x\ e\ S_2$ and $S_1 \leq S_2 \implies g\ b\ S_1 \leq g\ b\ S_2$. Then $Step\ f\ g$ is also monotone, in both arguments: $\llbracket C_1 \leq C_2; S_1 \leq S_2 \rrbracket \implies Step\ f\ g\ S_1\ C_1 \leq Step\ f\ g\ S_2\ C_2$.*

Proof. The proof is a straightforward computation induction on *Step*. Additionally it needs the easy lemma $C_1 \leq C_2 \implies post\ C_1 \leq post\ C_2$. \square

As a corollary we obtain the monotonicity of *step* (already used in the collecting semantics to guarantee that *step* has a least fixed point) because *step* is defined as $Step\ f\ g$ for some monotone f and g in Section 13.5.1. Similarly we have $step' = Step\ asem\ (\lambda b\ S.\ S)$. Although $\lambda b\ S.\ S$ is obviously monotone, *asem* is not necessarily monotone. The **monotone framework** is the extension of the interface to our abstract interpreter with a monotonicity assumption for abstract arithmetic:

Abstract arithmetic must be monotone:

$$\llbracket a_1 \leq b_1; a_2 \leq b_2 \rrbracket \implies plus'\ a_1\ a_2 \leq plus'\ b_1\ b_2$$

Monotonicity of *aval'* follows by an easy induction on e :

$$S_1 \leq S_2 \implies aval'\ e\ S_1 \leq aval'\ e\ S_2$$

In the monotone framework, *step'* is therefore monotone too, in both arguments. Therefore $step' \top$ is also monotone.

Monotonicity is not surprising and is only a means to obtain more interesting results, in particular precision and termination of our abstract interpreter. For the rest of this section we assume that we are in the context of the monotone framework.

13.5.7 Precision

So far we have shown correctness (*AI* overapproximates *CS*) but nothing about precision. In the worst case *AI* could annotate every given program with \top everywhere, which would be correct but useless.

We show that *AI* computes not just any pre-fixpoint but the least one. Why is this relevant? Because of the general principle:

Smaller is better because it is more precise.

The key observation is that iteration starting below a pre-fixpoint always stays below that pre-fixpoint:

Lemma 13.27. *Let $'a$ be a partial order and let f be a monotone function. If $x_0 \leq q$ for some pre-fixpoint q of f , then $f^i x_0 \leq q$ for all i .*

The proof is a straightforward induction on i and is left as an exercise.

As a consequence, if x_0 is a least element, then $f^i x_0 \leq q$ for all pre-fixpoints q . In this case $\text{pfp } f x_0 = \text{Some } p$ tells us that p is not just a pre-fixpoint of f , as observed further above, but by this lemma the least pre-fixpoint and hence, by Lemma 10.28, the least fixpoint.

In the end this merely means that the fixpoint computation is as precise as possible, but f itself may still overapproximate too much. In the abstract interpreter f is $\text{step}' \top$. Its precision depends on its constituents, which are the parameters of the module: Type $'av$ is required to be a semilattice, and hence \sqcup is not just any upper bound (which would be enough for correctness) but the least one, which ensures optimal precision. But the assumptions (num') and (plus') for the abstract arithmetic only require correctness and would, for example, allow plus' to always return \top . We do not dwell on the issue of precision any further but note that the canonical approach to abstract interpretation (see Section 13.10.1) covers it.

13.5.8 Termination

We prove termination of *AI* under certain conditions. More precisely, we show that pfp can only iterate step' finitely many times. The key insight is the following:

Lemma 13.28. *Let $'a$ be a partial order, let $f :: 'a \Rightarrow 'a$ be a monotone function, and let $x_0 :: 'a$ be such that $x_0 \leq f x_0$. Then the $f^i x_0$ form an ascending chain: $x_0 \leq f x_0 \leq f^2 x_0 \leq \dots$*

The proof that $f^i x_0 \leq f^{i+1} x_0$ is by induction on i and is left as an exercise. Note that $x_0 \leq f x_0$ holds in particular if x_0 is the least element.

Assume the conditions of the lemma hold for f and x_0 . Because $\text{pfp } f \ x_0$ computes the $f^i x_0$ until a pre-fixpoint is found, we know that the $f^i x_0$ form a strictly increasing chain. Therefore we can prove termination of this loop by exhibiting a measure function m into the natural numbers such that $x < y \implies m \ x > m \ y$. The latter property is called **anti-monotonicity**.

The following lemma summarizes this reasoning. Note that the relativisation to a set L is necessary because in our application the measure function will not be anti-monotone on the whole type.

Lemma 13.29. *Let $'a$ be a partial order, let $L :: 'a$ set, let $f \in L \rightarrow L$ be a monotone function, let $x_0 \in L$ such that $x_0 \leq f \ x_0$, and let $m :: 'a \Rightarrow \text{nat}$ be anti-monotone on L . Then $\exists p. \text{pfp } f \ x_0 = \text{Some } p$, i.e., $\text{pfp } f \ x_0$ terminates.*

In our concrete situation f is $\text{step}' \top$, x_0 is $\text{bot } c$, and we now construct a measure function m_c such that $C_1 < C_2 \implies m_c \ C_1 > m_c \ C_2$ (for certain C_i). The construction starts from a measure function on $'av$ that is lifted to $'av \text{ st option } acom$ in several steps.

We extend the interface to the abstract interpretation module by two more parameters that guarantee termination of the analysis:

Measure function and height: $m :: 'av \Rightarrow \text{nat}$
 $h :: \text{nat}$

Must be anti-monotone and bounded:

$a_1 < a_2 \implies m \ a_1 > m \ a_2$
 $m \ a \leq h$

Under these assumptions the ordering \leq on $'av$ is of **height** at most h : every chain $a_0 < a_1 < \dots < a_n$ has height at most h , i.e., $n \leq h$. That is, \leq on $'av$ is of **finite height**.

Let us first sketch the intuition behind the termination proof. The annotations in an annotated command can be viewed as a big tuple of the abstract values of all variables at all annotation points. For example, if the program has two annotations A_1 and A_2 and three variables x, y, z , then Figure 13.8 depicts some assignment of abstract values a_i to the three variables at the two annotation points. The termination measure is the sum of all $m \ a_i$. Lemma 13.28

A_1			A_2		
x	y	z	x	y	z
a_1	a_2	a_3	a_4	a_5	a_6

Fig. 13.8. Annotations as tuples

showed that in each step of a pre-fixpoint iteration of a monotone function

the ordering strictly increases, i.e., $(a_1, a_2, \dots) < (b_1, b_2, \dots)$, which for tuples means $a_i \leq b_i$ for all i and $a_k < b_k$ for some k . Anti-monotonicity implies both $m\ a_i \geq m\ b_i$ and $m\ a_k > m\ b_k$. Therefore the termination measure strictly decreases. Now for the technical details.

We lift m in three stages to *'av st option acom*:

definition $m_s :: 'av\ st \Rightarrow vname\ set \Rightarrow nat\ where$
 $m_s\ S\ X = (\sum_{x \in X}. m\ (S\ x))$

fun $m_o :: 'av\ st\ option \Rightarrow vname\ set \Rightarrow nat\ where$
 $m_o\ (Some\ S)\ X = m_s\ S\ X$
 $m_o\ None\ X = h * card\ X + 1$

definition $m_c :: 'av\ st\ option\ acom \Rightarrow nat\ where$
 $m_c\ C = (\sum_{a \leftarrow annos\ C}. m_o\ a\ (vars\ C))$

Measure $m_s\ S\ X$ is defined as the sum of all $m\ (S\ x)$ for $x \in X$. The set X is always finite in our context, namely the set of variables in some command.

By definition, $m_o\ (Some\ S)\ X < m_o\ None\ X$ because all our measures should be anti-monotone and *None* is the least value w.r.t. the ordering on *'av st option*.

In words, $m_c\ C$ sums up the measures of all the annotations in C . The definition of m_c uses the notation $\sum_{x \leftarrow xs}. f\ x$ for the summation over the elements of a list, in analogy to $\sum_{x \in X}. f\ x$ for sets. Function *vars* on annotated commands is defined by $vars\ C = vars\ (strip\ C)$ where $vars(c::com)$ is the set of variables in c (see [Appendix A](#)).

It is easy to show that all three measure functions are bounded:

$finite\ X \implies m_s\ S\ X \leq h * card\ X$
 $finite\ X \implies m_o\ opt\ X \leq h * card\ X + 1$
 $m_c\ C \leq length\ (annos\ C) * (h * card\ (vars\ C) + 1)$

The last lemma yields an upper bound for the number of iterations (= calls of *step'*) that *pfp* and therefore *AI* require: $p * ((h + 1) * n + 1)$ where p and n are the number of annotations and the number of variables in the given command. There are at most $p * (h + 1) * n$ many such assignments (see [Figure 13.8](#) for an illustration), and taking into account that an annotation can also be *None* yields the claimed upper bound.

The proof that all three measures are anti-monotone is more complicated. For example, anti-monotonicity of m_s cannot simply be expressed as $finite\ X \implies S_1 < S_2 \implies m_s\ S_2\ X < m_s\ S_1\ X$ because this does not hold: $S_1 < S_2$ could be due solely to an increase in some variable not in X , with all other variables unchanged, in which case $m_s\ X\ S_1 = m_s\ X\ S_2$. We need to take

into account that X are the variables in the command and that therefore the variables not in X do not change. This “does not change” property (relating two states) can more simply be expressed as “is top” (a property of one state): variables not in the program can only ever have the abstract value \top because we iterate $step' \top$. Therefore we define three “is top” predicates relative to some set of variables X :

definition $top_on_s :: 'av\ st \Rightarrow\ vname\ set \Rightarrow\ bool$ where
 $top_on_s\ S\ X = (\forall x \in X. S\ x = \top)$

fun $top_on_o :: 'av\ st\ option \Rightarrow\ vname\ set \Rightarrow\ bool$ where
 $top_on_o\ (Some\ S)\ X = top_on_s\ S\ X$
 $top_on_o\ None\ X = True$

definition $top_on_c :: 'av\ st\ option\ acom \Rightarrow\ bool$ where
 $top_on_c\ C\ X = (\forall a \in set\ (annos\ C). top_on_o\ a\ X)$

With the help of these predicates we can now formulate that all three measure functions are anti-monotone:

$$\begin{aligned} \llbracket finite\ X; S_1 = S_2\ on\ -\ X; S_1 < S_2 \rrbracket &\implies m_s\ S_2\ X < m_s\ S_1\ X \\ \llbracket finite\ X; top_on_o\ o_1\ (-\ X); top_on_o\ o_2\ (-\ X); o_1 < o_2 \rrbracket &\implies m_o\ o_2\ X < m_o\ o_1\ X \\ \llbracket top_on_c\ C_1\ (-\ vars\ C_1); top_on_c\ C_2\ (-\ vars\ C_2); C_1 < C_2 \rrbracket &\implies m_c\ C_2 < m_c\ C_1 \end{aligned}$$

The proofs involve simple reasoning about sums.

Now we can apply [Lemma 13.29](#) to $AI\ c = pfp\ (step' \top)\ (bot\ c)$ to conclude that AI always delivers:

Theorem 13.30. $\exists C. AI\ c = Some\ C$

Proof. In [Lemma 13.29](#), let L be $\{C. top_on_c\ C\ (-\ vars\ C)\}$ and let m be m_c . Above we have stated that m_c is anti-monotone on L . Now we examine that the remaining conditions of the lemma are satisfied. We had shown already that $step' \top$ is monotone. Showing that it maps L to L requires a little lemma

$$top_on_c\ C\ (-\ vars\ C) \implies top_on_c\ (step' \top\ C)\ (-\ vars\ C)$$

together with the even simpler lemma $strip\ (step' S\ C) = strip\ C\ (*)$. Both follow directly from analogous lemmas about $Step$ which are proved by computation induction on $Step$. Condition $bot\ c \leq step' \top\ (bot\ c)$ follows from the obvious $strip\ C = c \implies bot\ c \leq C$ with the help of $(*)$ above. And $bot\ c \in L$ holds because $top_on_c\ (bot\ c)\ X$ is true for all X by definition. \square

Example 13.31. Parity analysis can be shown to terminate because the *parity* semilattice has height 1. Defining the measure function

$$m_parity\ a = (if\ a = \textit{Either}\ then\ 0\ else\ 1)$$

Either is given measure 0 because it is the largest and therefore least informative abstract value. Both anti-monotonicity of m_parity and $m_parity\ a \leq 1$ are proved automatically.

Note that finite height of \leq is actually a bit stronger than necessary to guarantee termination. It is sufficient that there is no infinitely ascending chain $a_0 < a_1 < \dots$. But then there can be ascending chains of any finite height and we cannot bound the number of iterations of the abstract interpreter, which is problematic in practice.

13.5.9 Executability

Above we have shown that for semilattices of finite height, fixpoint iteration of the abstract interpreter terminates. Yet there is a problem: *AI* is not executable! This is why: *pfp* compares programs with annotations of type *'av st option*; but $S_1 \leq S_2$ (where $S_1, S_2 :: 'av\ st$) is defined as $\forall x. S_1\ x \leq S_2\ x$ where x comes from the infinite type *vname*. Therefore \leq on *'av st* is not directly executable.

We learn two things: we need to refine type *st* such that \leq becomes executable (this is the subject to the following section), and we need to be careful about claiming termination. We had merely proved that some term *while_option b f s* is equal to *Some*. This implies termination only if *b*, *f* and *s* are executable, which the given *b* is not. In general, there is no logical notion of executability and termination in HOL and such claims are informal. They could be formalized in principle but this would require a formalization of the code generation process and the target language.

Exercises

Exercise 13.10. Redo [Example 13.22](#) but replace the 2 in the program by 1.

Exercise 13.11. Take the Isabelle theories that define commands, big-step semantics, annotated commands and the collecting semantics and extend them with a nondeterministic choice construct as in [Exercise 13.2](#).

The following exercises require class constraints like *'a :: order* as introduced in [Section 13.4.1](#).

Exercise 13.12. Prove [Lemma 13.27](#) and [Lemma 13.28](#) in a detailed and readable style. Remember that $f^i\ x$ is input as $(f\ \wedge\ i)\ x$.

Exercise 13.13. Let $'a$ be a complete lattice and let $f :: 'a \Rightarrow 'a$ be a monotone function. Prove that if P is a set of pre-fixpoints of f then $\bigcap P$ is also a pre-fixpoint of f .

13.6 Executable Abstract States thy

This section is all about a clever representation of abstract states. We define a new type $'a\ st$ that is a semilattice where \leq , \sqcup and \top are executable (provided $'a$ is a semilattice with executable operations). Moreover it supports two operations that behave like function application and function update:

$$\begin{aligned} fun &:: 'a\ st \Rightarrow vname \Rightarrow 'a \\ update &:: 'a\ st \Rightarrow vname \Rightarrow 'a \Rightarrow 'a\ st \end{aligned}$$

This exercise in data refinement is independent of abstract interpretation and a bit technical in places. Hence it can be skipped on first reading. The key point is that our generic abstract interpreter from the previous section only requires two modifications: $S\ x$ is replaced by $fun\ S\ x$ and $S(x := a)$ is replaced by $update\ S\ x\ a$. The proofs carry over either verbatim or they require only local modifications. We merely need one additional lemma which reduces fun and $update$ to function application and function update: $fun\ (update\ S\ x\ y) = (fun\ S)(x := y)$.

Before we present the new definition of $'a\ st$, we look at two applications, parity analysis and constant propagation.

13.6.1 Parity Analysis thy

We had already instantiated our abstract interpreter with a parity analysis in the previous section. In [Example 13.22](#) we had even shown how the analysis behaves for a simple program. Now that the analyser is executable we apply it to a slightly more interesting example:

```
'x'' ::= N 1;;
WHILE Less (V 'x'') (N 100) DO 'x'' ::= Plus (V 'x'') (N 3)
```

As in [Section 13.3.1](#) we employ a hidden function `show_acom` which displays a function $\{x \mapsto a, y \mapsto b, \dots\}$ as $\{(x,a), (y,b), \dots\}$.

Four iterations of the step function result in this annotated command:

```
'x'' ::= N 1 {Some {('x'', Odd)}};;
{Some {('x'', Odd)}}
WHILE Less (V 'x'') (N 100)
DO {Some {('x'', Odd)}}
   'x'' ::= Plus (V 'x'') (N 3) {Some {('x'', Even)}}
{Some {('x'', Odd)}}
```

Now the *Even* feeds back into the invariant and waters it down to *Either*:

```

''x'' ::= N 1 {Some {''x'', Odd}};;
{Some {''x'', Either}}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', Odd}}
    ''x'' ::= Plus (V ''x'') (N 3) {Some {''x'', Even}}
{Some {''x'', Odd}}

```

A few more steps to propagate *Either* and we reach the least fixpoint:

```

''x'' ::= N 1 {Some {''x'', Odd}};;
{Some {''x'', Either}}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', Either}}
    ''x'' ::= Plus (V ''x'') (N 3) {Some {''x'', Either}}
{Some {''x'', Either}}

```

Variable x can be *Either* everywhere, except after $''x'' ::= N 1$.

13.6.2 Constant Propagation [thy](#)

We have already seen a constant propagation analysis in [Section 10.2](#). That was an *ad hoc* design. Now we obtain constant propagation (although not the folding optimization) as an instance of our abstract interpreter. The advantage of instantiating a generic design and correctness proof is not code reuse but the many hours or even days one saves by obtaining the correctness proof for free.

The abstract domain for constant propagation permits us to analyse that the value of a variable at some point is either some integer i (*Const* i) or could be anything (*Any*):

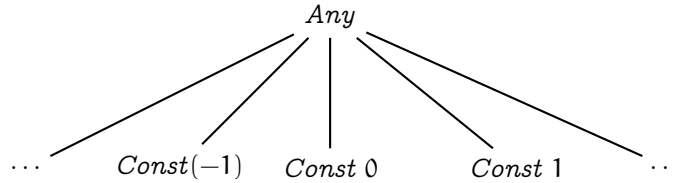
```
datatype const = Const int | Any
```

```
fun  $\gamma_{const} :: const \Rightarrow val\ set\ where$ 
```

```
 $\gamma_{const} (Const\ i) = \{i\}$ 
```

```
 $\gamma_{const}\ Any = UNIV$ 
```

To reflect γ_{const} , the ordering must look like this:



In symbols:

$$(x \leq y) = (y = \text{Any} \vee x = y)$$

$$x \sqcup y = (\text{if } x = y \text{ then } x \text{ else } \text{Any})$$

$$\top = \text{Any}$$

The best we can do w.r.t. addition on *const* is this:

```
fun plus_const :: const ⇒ const ⇒ const where
  plus_const (Const i) (Const j) = Const (i + j)
  plus_const _ _ = Any
```

Termination of the analysis is guaranteed because the ordering is again of height 1 as witnessed by this measure function:

$$m_const\ x = (\text{if } x = \text{Any} \text{ then } 0 \text{ else } 1)$$

Having instantiated the abstract interpreter with the above domain we run some examples. We have now seen enough iterations and merely show the final results.

Some basic arithmetic:

```
"x" ::= N 1 {Some {"x", Const 1}, ("y", Any)};;
"x" ::= Plus (V "x") (V "x") {Some {"x", Const 2}, ("y", Any)};;
"x" ::= Plus (V "x") (V "y") {Some {"x", Any}, ("y", Any)}
```

A conditional where both branches set *x* to the same value:

```
IF Less (N 41) (V "x")
THEN {Some {"x", Any}} "x" ::= N 5 {Some {"x", Const 5}}
ELSE {Some {"x", Any}} "x" ::= N 5 {Some {"x", Const 5}}
{Some {"x", Const 5}}
```

A conditional where both branches set *x* to different values:

```
IF Less (N 41) (V "x")
THEN {Some {"x", Any}} "x" ::= N 5 {Some {"x", Const 5}}
ELSE {Some {"x", Any}} "x" ::= N 6 {Some {"x", Const 6}}
{Some {"x", Any}}
```

Conditions are ignored, which leads to imprecision:

```
"x" ::= N 42 {Some {"x", Const 42}};;
IF Less (N 41) (V "x")
THEN {Some {"x", Const 42}} "x" ::= N 5 {Some {"x", Const 5}}
ELSE {Some {"x", Const 42}} "x" ::= N 6 {Some {"x", Const 6}}
{Some {"x", Any}}
```

This is where the analyser from [Section 10.2](#) beats our abstract interpreter. [Section 13.8](#) will rectify this deficiency.

As an exercise, compute the following result step by step:

```

''x'' ::= N 0 {Some {''x'', Const 0}, (''y'', Any), (''z'', Any)};;
''y'' ::= N 0 {Some {''x'', Const 0}, (''y'', Const 0), (''z'', Any)};;
''z'' ::= N 2
{Some {''x'', Const 0}, (''y'', Const 0), (''z'', Const 2)};;
{Some {''x'', Any}, (''y'', Any), (''z'', Const 2)}
WHILE Less (V ''x'') (N 1)
DO {Some {''x'', Any}, (''y'', Any), (''z'', Const 2)}
  (''x'' ::= V ''y'')
  {Some {''x'', Any}, (''y'', Any), (''z'', Const 2)};;
  ''y'' ::= V ''z''
  {Some {''x'', Any}, (''y'', Const 2), (''z'', Const 2)}}
{Some {''x'', Any}, (''y'', Any), (''z'', Const 2)}

```

13.6.3 Representation Type 'a st_rep thy

After these two applications we come back to the challenge of making \leq on *st* executable. We exploit that states always have a finite domain, namely the variables in the program. Outside that domain, states do not change. Thus an abstract state can be represented by an association list:

type_synonym 'a st_rep = (vname \times 'a) list

Variables of type 'a st_rep will be called *ps*.

Function *fun_rep* turns an association list into a function. Arguments not present in the association list are mapped to the default value \top :

```

fun fun_rep :: ('a::top) st_rep  $\Rightarrow$  vname  $\Rightarrow$  'a where
  fun_rep [] = ( $\lambda x.$   $\top$ )
  fun_rep ((x, a) # ps) = (fun_rep ps)(x := a)

```

Class *top* is predefined. It is the class of all types with a top element \top . Here are two examples of the behaviour of *fun_rep*:

```

fun_rep [('x'', a), ('y'', b)] = ( $\lambda x.$   $\top$ )(''y'' := b)(''x'' := a)
fun_rep [('x'', a), ('x'', b)] = ( $\lambda x.$   $\top$ )(''x'' := b)(''x'' := a)
= ( $\lambda x.$   $\top$ )(''x'' := a).

```

Comparing two association lists is easy now: you only need to compare them on their finite “domains”, i.e., on *map fst*:

$$\begin{aligned}
\text{less_eq_st_rep } ps_1 \text{ } ps_2 = \\
(\forall x \in \text{set } (\text{map fst } ps_1) \cup \text{set } (\text{map fst } ps_2). \\
\text{fun_rep } ps_1 \text{ } x \leq \text{fun_rep } ps_2 \text{ } x)
\end{aligned}$$

Unfortunately this is not a partial order (which is why we gave it the name *less_eq_st_rep* instead of \leq). For example, both $[(''x'', a), (''y'', b)]$ and $[(''y'', b), (''x'', a)]$ represent the same function and the *less_eq_st_rep* relationship between them holds in both directions, but they are distinct association lists.

13.6.4 Quotient Type 'a st

Therefore we define the actual new abstract state type as a **quotient type**. In mathematics, the quotient of a set M by some equivalence relation \sim (see Definition 7.7) is written M/\sim and is the set of all equivalence classes $[m]_\sim \subseteq M, m \in M$, defined by $[m]_\sim = \{m' \mid m \sim m'\}$. In our case the equivalence relation is *eq_st* and it identifies two association lists iff they represent the same function:

$$\text{eq_st } ps_1 \text{ } ps_2 = (\text{fun_rep } ps_1 = \text{fun_rep } ps_2)$$

This is precisely the point of the quotient construction: identify data elements that are distinct but abstractly the same.

In Isabelle the quotient type 'a st is introduced as follows

$$\text{quotient_type 'a st} = ('a::\text{top}) \text{ st_rep} / \text{eq_st}$$

thereby overwriting the old definition of 'a st. The class constraint is required because *eq_st* is defined in terms of *fun_rep*, which requires *top*. Instead of the mathematical equivalence class notation $[ps]_{\text{eq_st}}$ we write *St ps*. We do not go into further details of **quotient_type**: they are not relevant here and can be found elsewhere [92].

Now we can define all the required operations on 'a st, starting with *fun* and *update*:

$$\text{fun } (\text{St } ps) = \text{fun_rep } ps$$

$$\text{update } (\text{St } ps) \text{ } x \text{ } a = \text{St } ((x, a) \# ps)$$

The concretisation function γ_{fun} is replaced by γ_{st} :

$$\gamma_{\text{st}} \gamma \text{ } S = \{f. \forall x. f \text{ } x \in \gamma (\text{fun } S \text{ } x)\}$$

This is how 'a st is turned into a semilattice with top element:

$$\top = St \ \square$$

$$(St \ ps_1 \leq St \ ps_2) = less_eq_st_rep \ ps_1 \ ps_2$$

$$St \ ps_1 \sqcup St \ ps_2 = St(map2_st_rep \ (\sqcup) \ ps_1 \ ps_2)$$

$$\text{fun } map2_st_rep ::$$

$$('a::top \Rightarrow 'a \Rightarrow 'a) \Rightarrow 'a \ st_rep \Rightarrow 'a \ st_rep \Rightarrow 'a \ st_rep$$

$$\text{where}$$

$$map2_st_rep \ f \ \square \ ps_2 = map \ (\lambda(x, y). (x, f \ \top \ y)) \ ps_2$$

$$map2_st_rep \ f \ ((x, y) \# ps_1) \ ps_2 =$$

$$(let \ y_2 = fun_rep \ ps_2 \ x \ in \ (x, f \ y \ y_2) \# map2_st_rep \ f \ ps_1 \ ps_2)$$

The auxiliary function *map2_st_rep* is best understood via this lemma:

$$f \ \top \ \top = \top \implies$$

$$fun_rep \ (map2_st_rep \ f \ ps_1 \ ps_2) =$$

$$(\lambda x. f \ (fun_rep \ ps_1 \ x) \ (fun_rep \ ps_2 \ x))$$

Here and in later uses of *map2_st_rep* the premise $f \ \top \ \top = \top$ is always satisfied.

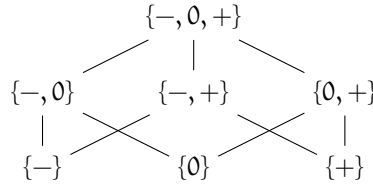
Most of the above functions are defined with the help of some representative *ps* of an equivalence class *St ps*. That is, they are of the form $f(St \ ps) = \dots \ ps \dots$. This is not a definition until one has shown that the right-hand side is independent of the choice of *ps*. Take $fun \ (St \ ps) = fun_rep \ ps$. It must be shown that if $eq_st \ ps_1 \ ps_2$ then $fun_rep \ ps_1 = fun_rep \ ps_2$, which is trivial in this case. Isabelle insists on these proofs but they are routine and we do not go into them here.

Although the previous paragraph is mathematically accurate, the Isabelle details are a bit different because *St* is not a constructor. Hence equations like $fun \ (St \ ps) = fun_rep \ ps$ cannot be definitions but are derived lemmas. The actual definition needs to be made on the *st_rep* level first and is then lifted automatically to the *st* level. For example, *fun_rep* is defined first and *fun* is derived from it almost automatically, except that we have to prove that *fun_rep* is invariant under *eq_st* as explained above.

Exercises

Exercise 13.14. Instantiate the abstract interpreter with an analysis of the sign of variables based on these abstract values: `datatype sign = Neg | Zero | Pos | Any`. Their concretisations are $\{i. i < 0\}$, $\{0\}$, $\{i. 0 < i\}$ and *UNIV*. Formalize your analysis by modifying the parity analysis theory *Abs_Int1_parity*.

Exercise 13.15. The previous sign analysis can be refined as follows. The basic signs are “+”, “−” and “0”, but all combinations are also possible: e.g., $\{0, +\}$ abstracts the set of non-negative integers. This leads to the following semilattice (we do not need $\{\}$):



A possible representation is a datatype with seven constructors, but there is a more elegant solution (including $\{\}$). Formalize your analysis by modifying theory *Abs_Int1_parity* or your solution to the previous exercise.

13.7 Analysis of Boolean Expressions thy

So far we have ignored boolean expressions. A first, quite simple analysis could be the following one: Define abstract evaluation of boolean expressions w.r.t. an abstract state such that the result can be Yes, No, or Maybe. In the latter case, the analysis proceeds as before, but if the result is a definite Yes or No, then we can ignore one of the two branches after the test. For example, during constant propagation we may have found out that x has value 5, in which case the value of $x < 5$ is definitely No.

We want to be more ambitious. Consider the test $x < 5$ and assume that we perform an interval analysis where we have merely determined that x must be in the interval $[0 \dots 10]$. Hence $x < 5$ evaluates to Maybe, which is no help. However, we would like to infer for IF $x < 5$ that in the THEN-branch, x is in the interval $[1 \dots 4]$ and in the ELSE-branch, x is in the interval $[5 \dots 10]$, thus improving the analysis of the two branches.

For an arbitrary boolean expression such an analysis can be hard, especially if multiplication is involved. Nevertheless we will present a generic analysis of boolean expressions. It works well for simple examples although it is far from optimal in the general case.

The basic idea of any non-trivial analysis of boolean expressions is to simulate the collecting semantics *step*. Recall from [Section 13.5.1](#) that *step* is a specialized version of *Step* where $bsem\ b\ S = \{s \in S. \ bval\ b\ s\}$ whereas for *step'* we used $bsem\ b\ S = S$ (see [Section 13.5.4](#)). We would like $bsem\ b\ S$, for $S :: 'av\ st$, to be some $S' \leq S$ such that

No state satisfying b is lost: $\{s \in \gamma_s S. \text{bval } b \ s\} \subseteq \gamma_s S'$

Of course $S' = S$ satisfies this specification, but we can do better.

Computing $S' \leq S$ from S requires a new operation because \sqcup only increases but we need to decrease.

Definition 13.32. A type $'a$ is a *lattice* if it is a semilattice and there is an *infimum* operation $\sqcap :: 'a \Rightarrow 'a \Rightarrow 'a$ that returns the greatest lower bound of its arguments:

- Lower bound: $x \sqcap y \leq x$ and $x \sqcap y \leq y$
- Greatest: $\llbracket x \leq y; x \leq z \rrbracket \implies x \leq y \sqcap z$

A partial order has a *bottom element* \perp if $\perp \leq x$ for all x .

A *bounded lattice* is a lattice with both \top and \perp .

Isabelle provides the corresponding classes *lattice* and *bounded_lattice*. The infimum is sometimes called the *meet* operation.

Fact 13.33. Every complete lattice is a lattice.

This follows because a complete lattice with infimum \sqcap also has a supremum \sqcup (see [Exercise 13.3](#)) and thus we obtain their binary versions as special cases: $x \sqcup y = (\sqcup \{x, y\})$ and $x \sqcap y = (\sqcap \{x, y\})$.

We strengthen the abstract interpretation framework as follows:

- The abstract domain $'av$ must be a lattice
- Infimum must be precise: $\gamma (a_1 \sqcap a_2) = \gamma a_1 \cap \gamma a_2$
- $\gamma \perp = \{\}$

It is left as an easy exercise to show that $\gamma (a_1 \sqcap a_2) \subseteq \gamma a_1 \cap \gamma a_2$ holds in any lattice because γ is monotone (see [Lemma 13.19](#)). Therefore the actual requirement is the other inclusion: $\gamma a_1 \cap \gamma a_2 \subseteq \gamma (a_1 \sqcap a_2)$.

In this context, arithmetic and boolean expressions are analysed. It is a kind of backward analysis. Given an arithmetic expression e , its expected value $a :: 'av$, and some $S :: 'av \text{ st}$, compute some $S' \leq S$ such that

$$\{s \in \gamma_s S. \text{aval } e \ s \in \gamma a\} \subseteq \gamma_s S'$$

Roughly speaking, S' overapproximates the subset of S that makes e evaluate to a .

What if e cannot evaluate to a , i.e., $\{s \in \gamma_s S. \text{aval } e \ s \in \gamma a\} = \{\}$? We need to lift the whole process to $S, S' :: 'av \text{ st option}$; then $S' = \text{None}$ covers this situation. However, there is a subtlety: *None* is not the only abstraction of the empty set of states. For every $S' :: 'av \text{ st}$ where $\text{fun } S' \ x = \perp$ for some x , $\gamma_s S' = \{\}$ by definition of γ_s because $\gamma \perp = \{\}$. Why is this an issue? Let $S' :: 'av \text{ st}$ such that $\gamma_s S' = \{\}$. Both *None* and *Some* S' abstract $\{\}$

but they behave differently in a supremum: $None \sqcup S'' = S''$ but we may have $Some\ S' \sqcup S'' > S''$. Thus $Some\ S'$ can lead to reduced precision, for example when joining the result of the analyses of the THEN and ELSE branches where one of the two branches is unreachable (leading to $\gamma_s\ S' = \{\}$). Hence, for the sake of precision, we adopt the following policy:

Avoid $Some\ S'$ where $fun\ S'\ x = \perp$ for some x ; use $None$ instead.

The backward analysis functions will typically be called inv_f because they invert some function f on the abstract domain in roughly the following sense. Given some expected result r of f and some arguments $args$ of f , $inv_f\ r\ args$ should return reduced arguments $args' \leq args$ such that the reduction from $args$ to $args'$ may eliminate anything that does not lead to the desired result r . For correctness we only need to show that nothing that leads to r is lost by going from $args$ to $args'$. The relation $args' \leq args$ is just for intuition and precision but in the worst case $args' = \top$ would be correct too.

13.7.1 Backward Analysis of Arithmetic Expressions

The analysis is performed recursively over the expression. Therefore we need an inverse function for every constructor of $aexp$ (except V). The abstract arithmetic in the abstract interpreter interface needs to provide two additional executable functions:

- $test_num' :: val \Rightarrow 'av \Rightarrow bool$ such that
 $test_num'\ i\ a = (i \in \gamma\ a)$
- $inv_plus' :: 'av \Rightarrow 'av \Rightarrow 'av \Rightarrow 'av \times 'av$ such that

$$\llbracket inv_plus'\ a\ a_1\ a_2 = (a'_1, a'_2); i_1 \in \gamma\ a_1; i_2 \in \gamma\ a_2; i_1 + i_2 \in \gamma\ a \rrbracket$$

$$\implies i_1 \in \gamma\ a'_1 \wedge i_2 \in \gamma\ a'_2$$

A call $inv_plus'\ a\ a_1\ a_2$ should reduce (a_1, a_2) to some (a'_1, a'_2) such that every pair of integers represented by (a_1, a_2) that adds up to some integer represented by a is still represented by (a'_1, a'_2) . For example, for intervals, if $a = [8..10]$, $a_1 = [4..11]$ and $a_2 = [1..9]$, then the smallest possible result is $(a'_1, a'_2) = ([4..9], [1..6])$; reducing either interval further loses some result in $[8..10]$.

If there were further arithmetic functions, their backward analysis would follow the inv_plus' pattern. Function $test_num'$ is an exception that has been optimized. In standard form it would be called inv_num' and would return $'av$ instead of $bool$.

With the help of these functions we can formulate the backward analysis of arithmetic expressions:

```

fun inv_aval' :: aexp ⇒ 'av st option ⇒ 'av st option where
  inv_aval' (N n) a S = (if test_num' n a then S else None)
  inv_aval' (V x) a S =
    (case S of None ⇒ None
     | Some S ⇒
       let a' = fun S x □ a
         in if a' = ⊥ then None else Some (update S x a'))
  inv_aval' (Plus e1 e2) a S =
    (let (a1, a2) = inv_plus' a (aval'' e1 S) (aval'' e2 S)
     in inv_aval' e1 a1 (inv_aval' e2 a2 S))

fun aval'' :: aexp ⇒ 'av st option ⇒ 'av where
  aval'' e None = ⊥
  aval'' e (Some S) = aval' e S

```

The *let*-expression in the $\text{inv_aval}' (V x)$ equation could be collapsed to $\text{Some } (\text{update } S x (\text{fun } S x \square a))$ except that we would lose precision for the reason discussed above.

Function aval'' is just a lifted version of aval' and its correctness follows easily by induction:

$$s \in \gamma_o S \implies \text{aval } a s \in \gamma (\text{aval}'' a S)$$

This permits us to show correctness of $\text{inv_aval}'$:

Lemma 13.34 (Correctness of $\text{inv_aval}'$).

If $s \in \gamma_o S$ and $\text{aval } e s \in \gamma a$ then $s \in \gamma_o (\text{inv_aval}' e a S)$.

Proof. By induction on e . Case N is trivial. Consider case $V x$. From $s \in \gamma_o S$ we obtain an S' such that $S = \text{Some } S'$ and $s \in \gamma_s S'$, and hence $s x \in \gamma (\text{fun } S' x)$. We need to show $s \in \gamma_o (\text{if } a' = \perp \text{ then None else Some } (\text{update } S' x a'))$ where $a' = \text{fun } S' x \square a$. Note that $\gamma a' = \gamma (\text{fun } S' x \square a) = \gamma (\text{fun } S' x) \cap \gamma a$ because \square must be precise. Assumption $\text{aval } (V x) s \in \gamma a$ implies $s x \in \gamma a$, and thus $s x \in \gamma a'$ (because $s x \in \gamma (\text{fun } S' x)$). Therefore the case $a' = \perp$ cannot arise in this proof because by assumption $\gamma \perp = \{\}$. Now we show the *else* branch is always correct: $s \in \gamma_o (\text{Some } (\text{update } S' x a')) = \gamma_s (\text{update } S' x a')$. This reduces to $s x \in \gamma a'$ (because $s \in \gamma_s S'$), which we have proved above.

The case $\text{Plus } e_1 e_2$ is automatic by means of the assumption about $\text{inv_plus}'$ together with the correctness of aval'' . \square

If the definition of $\text{inv_aval}' (\text{Plus } e_1 e_2)$ still mystifies you: instead of $\text{inv_aval}' e_1 a_1 (\text{inv_aval}' e_2 a_2 S)$ we could have written the more intuitive $\text{inv_aval}' e_1 a_1 S \square \text{inv_aval}' e_2 a_2 S$, except that this would have required us to lift \square to $'av \text{ st option}$, which we avoided by nesting the two $\text{inv_aval}'$ calls.

13.7.2 Backward Analysis of Boolean Expressions

The analysis needs an inverse function for the arithmetic operator “<”, i.e., abstract arithmetic in the abstract interpreter interface needs to provide another executable function:

$$\begin{aligned} \text{inv_less}' :: \text{bool} \Rightarrow 'av \Rightarrow 'av \Rightarrow 'av \times 'av \quad \text{such that} \\ \llbracket \text{inv_less}' (i_1 < i_2) \ a_1 \ a_2 \rrbracket = (a'_1, a'_2); i_1 \in \gamma \ a_1; i_2 \in \gamma \ a_2 \\ \implies i_1 \in \gamma \ a'_1 \wedge i_2 \in \gamma \ a'_2 \end{aligned}$$

The specification follows the *inv_plus'* pattern.

Now we can define the backward analysis of boolean expressions:

$$\begin{aligned} \text{fun inv_bval}' :: \text{bexp} \Rightarrow \text{bool} \Rightarrow 'av \text{ st option} \Rightarrow 'av \text{ st option where} \\ \text{inv_bval}' (Bc \ v) \ \text{res} \ S = (\text{if } v = \text{res} \text{ then } S \text{ else None}) \\ \text{inv_bval}' (\text{Not } b) \ \text{res} \ S = \text{inv_bval}' b \ (\neg \text{res}) \ S \\ \text{inv_bval}' (\text{And } b_1 \ b_2) \ \text{res} \ S = \\ (\text{if } \text{res} \text{ then } \text{inv_bval}' b_1 \ \text{True} \ (\text{inv_bval}' b_2 \ \text{True} \ S) \\ \text{else } \text{inv_bval}' b_1 \ \text{False} \ S \sqcup \text{inv_bval}' b_2 \ \text{False} \ S) \\ \text{inv_bval}' (\text{Less } e_1 \ e_2) \ \text{res} \ S = \\ (\text{let } (a_1, a_2) = \text{inv_less}' \ \text{res} \ (\text{aval}'' e_1 \ S) \ (\text{aval}'' e_2 \ S) \\ \text{in } \text{inv_aval}' e_1 \ a_1 \ (\text{inv_aval}' e_2 \ a_2 \ S)) \end{aligned}$$

The *Bc* and *Not* cases should be clear. The *And* case becomes symmetric when you realise that $\text{inv_bval}' b_1 \ \text{True} \ (\text{inv_bval}' b_2 \ \text{True} \ S)$ is another way of saying $\text{inv_bval}' b_1 \ \text{True} \ S \sqcap \text{inv_bval}' b_2 \ \text{True} \ S$ without needing \sqcap . Case *Less* is analogous to case *Plus* of *inv_aval'*.

Lemma 13.35 (Correctness of *inv_bval'*).

$$s \in \gamma_o \ S \implies s \in \gamma_o \ (\text{inv_bval}' b \ (\text{bval } b \ s) \ S)$$

Proof. By induction on *b* with the help of correctness of *inv_aval'* and *aval''*. □

13.7.3 Abstract Interpretation

As explained in the introduction to this section, we want to replace *bsem b S = S* used in the previous sections with something better. Now we can: *step'* is defined from *Step* with

$$\text{bsem } b \ S = \text{inv_bval}' b \ \text{True} \ S$$

Correctness is straightforward:

Lemma 13.36. $\text{AI } c = \text{Some } C \implies \text{CS } c \leq \gamma_c \ C$

Proof. Just as for [Lemma 13.25](#), with the help of [Corollary 13.23](#), which needs correctness of inv_bval' . \square

This section was free of examples because backward analysis of boolean expressions is most effectively demonstrated on intervals, which require a separate section, the following one.

Exercises

Exercise 13.16. Prove that if $\gamma :: 'a::lattice \Rightarrow 'b::lattice$ is a monotone function, then $\gamma (a_1 \sqcap a_2) \subseteq \gamma a_1 \cap \gamma a_2$. Give an example of two lattices and a monotone γ where $\gamma a_1 \cap \gamma a_2 \subseteq \gamma (a_1 \sqcap a_2)$ does not hold.

Exercise 13.17. Consider a simple sign analysis based on the abstract domain `datatype sign = None | Neg | Pos0 | Any` where $\gamma :: sign \Rightarrow val set$ is defined by $\gamma None = \{\}$, $\gamma Neg = \{i. i < 0\}$, $\gamma Pos0 = \{i. 0 \leq i\}$ and $\gamma Any = UNIV$. Define inverse analyses for “+” and “<”

$$\begin{aligned} inv_plus' &:: sign \Rightarrow sign \Rightarrow sign \Rightarrow sign \times sign \\ inv_less' &:: bool \Rightarrow sign \Rightarrow sign \Rightarrow sign \times sign \end{aligned}$$

and prove the required correctness properties $inv_plus' a a_1 a_2 = (a'_1, a'_2) \implies \dots$ and $inv_less' bv a_1 a_2 = (a'_1, a'_2) \implies \dots$ stated in [Section 13.7.1](#) and [13.7.2](#).

Exercise 13.18. Extend the abstract interpreter from the previous section with the simple forward analysis of boolean expressions sketched in the first paragraph of this section. You need to add a function $less' :: 'av \Rightarrow 'av \Rightarrow bool option$ (where *None*, *Some True* and *Some False* mean Maybe, Yes and No) to locale *Val_semilattice* in theory *Abs_Int0* (together with a suitable assumption), define a function $bval' :: bexp \Rightarrow 'av st \Rightarrow bool option$ in *Abs_Int1*, modify the definition of $step'$ to make use of $bval'$, and update the proof of correctness of *AI*. The remainder of theory *Abs_Int1* can be discarded. Finally adapt constant propagation analysis in theory *Abs_Int1_const*.

13.8 Interval Analysis thy

We had already seen examples of interval analysis in the informal introduction to abstract interpretation in [Section 13.1](#), but only now are we in a position to make formal sense of it.

13.8.1 Intervals

Let us start with the type of intervals itself. Our intervals need to include infinite endpoints, otherwise we could not represent any infinite sets. Therefore we base intervals on **extended integers**, i.e., integers extended with ∞ and $-\infty$. The corresponding polymorphic data type is

```
datatype 'a extended = Fin 'a |  $\infty$  |  $-\infty$ 
```

and *eint* are the extended integers:

```
type_synonym eint = int extended
```

Type *eint* comes with the usual arithmetic operations.

! Constructor *Fin* can be dropped in front of numerals: 5 is as good as *Fin* 5 (except for the type constraint implied by the latter). This applies to input and output of terms. We do not discuss the necessary Isabelle magic.

The naive model of an interval is a pair of extended integers that represent the set of integers between them:

```
type_synonym eint2 = eint  $\times$  eint
```

```
definition  $\gamma\_rep :: eint2 \Rightarrow int\ set$  where  
 $\gamma\_rep\ p = (let\ (l, h) = p\ in\ \{i. l \leq Fin\ i \wedge Fin\ i \leq h\})$ 
```

For example, $\gamma_rep\ (0, 3) = \{0, 1, 2, 3\}$ and $\gamma_rep(0, \infty) = \{0, 1, 2, \dots\}$.

Unfortunately this results in infinitely many empty intervals, namely all pairs (l, h) where $h < l$. Hence we need to perform a quotient construction, just like for *st* in [Section 13.6.4](#). We consider two intervals equivalent if they represent the same set of integers:

```
definition eq_ivl :: eint2  $\Rightarrow$  eint2  $\Rightarrow$  bool where  
eq_ivl  $p_1\ p_2 = (\gamma\_rep\ p_1 = \gamma\_rep\ p_2)$ 
```

```
quotient_type ivl = eint2 / eq_ivl
```

Most of the time we blur the distinction between *ivl* and *eint2* by introducing the abbreviation $[l, h] :: ivl$, where $l, h :: eint$, for the equivalence class of all pairs equivalent to (l, h) w.r.t. *eq_ivl*. Unless $\gamma_rep\ (l, h) = \{\}$, this equivalence class only contains (l, h) . Moreover, let \perp be the empty interval, i.e., $[l, h]$ for any $h < l$. Note that there are two corner cases where $[l, h] = \perp$ does not imply $h < l$: $[\infty, \infty]$ and $[-\infty, -\infty]$.

We will now “define” most functions on intervals by pattern matching on $[l, h]$. Here is a first example:

$\gamma_{ivl} :: ivl \Rightarrow int\ set$
 $\gamma_{ivl} [l, h] = \{i. l \leq Fin\ i \wedge Fin\ i \leq h\}$

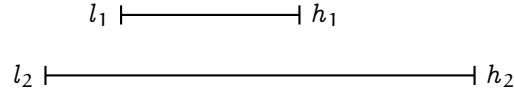
Recall from the closing paragraph of [Section 13.6.4](#) that the actual definitions must be made on the *eint2* level and that the pattern matching equations are derived lemmas.

13.8.2 Intervals as a Bounded Lattice

The following definition turns *ivl* into a partial order:

$([l_1, h_1] \leq [l_2, h_2]) =$
 $(if\ [l_1, h_1] = \perp\ then\ True$
 $\quad else\ if\ [l_2, h_2] = \perp\ then\ False\ else\ l_2 \leq l_1 \wedge h_1 \leq h_2)$

The ordering can be visualized like this:



Making *ivl* a bounded lattice is not hard either:

$[l_1, h_1] \sqcup [l_2, h_2] =$
 $(if\ [l_1, h_1] = \perp\ then\ [l_2, h_2]$
 $\quad else\ if\ [l_2, h_2] = \perp\ then\ [l_1, h_1]\ else\ [min\ l_1\ l_2, max\ h_1\ h_2])$
 $[l_1, h_1] \sqcap [l_2, h_2] = [max\ l_1\ l_2, min\ h_1\ h_2]$
 $\top = [-\infty, \infty]$

The supremum and infimum of two overlapping intervals is shown graphically in [Figure 13.9](#). The loss of precision resulting from the supremum of two non-

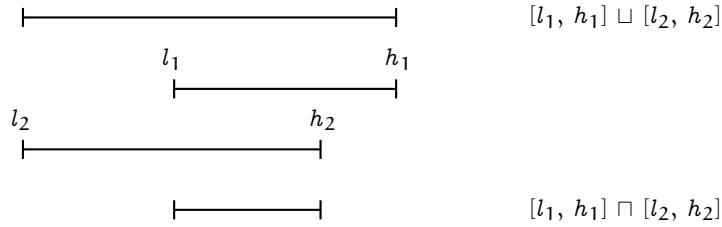


Fig. 13.9. Supremum and infimum of overlapping intervals

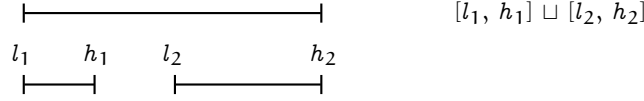


Fig. 13.10. Supremum of non-overlapping intervals

overlapping intervals is shown in Figure 13.10; their empty infimum is not shown.

Intervals do in fact form a complete lattice, but we do not need this fact. It is easy to see that \sqcap is precise:

$$\gamma_{\text{ivl}}(\text{iv}_1 \sqcap \text{iv}_2) = \gamma_{\text{ivl}} \text{iv}_1 \cap \gamma_{\text{ivl}} \text{iv}_2$$

no matter if the two intervals overlap or not.

We will at best sketch proofs to do with intervals. These proofs are conceptually simple: you just need to make a case distinction over all extended integers in the proposition. More importantly, intervals are not germane to abstract interpretation. The interested reader should consult the Isabelle theories for the proof details.

13.8.3 Arithmetic on Intervals

Addition, negation and subtraction are straightforward:

$$\begin{aligned} [l_1, h_1] + [l_2, h_2] &= \\ &(\text{if } [l_1, h_1] = \perp \vee [l_2, h_2] = \perp \text{ then } \perp \text{ else } [l_1 + l_2, h_1 + h_2]) \\ - [l, h] &= [-h, -l] \\ \text{iv}_1 - \text{iv}_2 &= \text{iv}_1 + - \text{iv}_2 \end{aligned}$$

So is the inverse function for addition:

$$\text{inv_plus_ivl } \text{iv } \text{iv}_1 \text{ iv}_2 = (\text{iv}_1 \sqcap (\text{iv} - \text{iv}_2), \text{iv}_2 \sqcap (\text{iv} - \text{iv}_1))$$

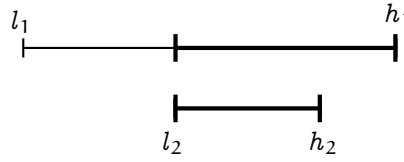
We sketch why this definition meets the requirement that if $i_1 \in \gamma_{\text{ivl}} \text{iv}_1$, $i_2 \in \gamma_{\text{ivl}} \text{iv}_2$ and $i_1 + i_2 \in \gamma_{\text{ivl}} \text{iv}$ then $i_1 \in \gamma_{\text{ivl}}(\text{iv}_1 \sqcap (\text{iv} - \text{iv}_2)) = \gamma_{\text{ivl}} \text{iv}_1 \cap \gamma_{\text{ivl}}(\text{iv} - \text{iv}_2)$ (and similarly for i_2). By assumption $i_1 \in \gamma_{\text{ivl}} \text{iv}_1$. Moreover $i_1 \in \gamma_{\text{ivl}}(\text{iv} - \text{iv}_2) = \{i_1. \exists i \in \gamma_{\text{ivl}} \text{iv}. \exists i_2 \in \gamma_{\text{ivl}} \text{iv}_2. i_1 = i - i_2\}$ also holds, by assumptions $i_2 \in \gamma_{\text{ivl}} \text{iv}_2$ and $i_1 + i_2 \in \gamma_{\text{ivl}} \text{iv}$.

Backward analysis for “<” is more complicated because the empty intervals need to be treated specially and because there are two cases depending on the expected result $\text{res} :: \text{bool}$:

$$\begin{aligned}
& \text{inv_less_ivl } \text{res } iv_1 \text{ } iv_2 = \\
& (\text{if } \text{res} \\
& \quad \text{then } (iv_1 \sqcap (\text{below } iv_2 - [1, 1]), iv_2 \sqcap (\text{above } iv_1 + [1, 1])) \\
& \quad \text{else } (iv_1 \sqcap \text{above } iv_2, iv_2 \sqcap \text{below } iv_1)) \\
& \text{above } [l, h] = (\text{if } [l, h] = \perp \text{ then } \perp \text{ else } [l, \infty]) \\
& \text{below } [l, h] = (\text{if } [l, h] = \perp \text{ then } \perp \text{ else } [-\infty, h])
\end{aligned}$$

The correctness proof follows the pattern of the one for *inv_plus_ivl*.

For a visualization of *inv_less_ivl* we focus on the “ \geq ” case, i.e., $\neg \text{res}$, and consider the following situation, where $iv_1 = [l_1, h_1]$, $iv_2 = [l_2, h_2]$ and the thick lines indicate the result:



All those points that can definitely not lead to the first interval being \geq the second have been eliminated. Of course this is just one of finitely many ways in which the two intervals can be positioned relative to each other.

The “ $<$ ” case in the definition of *inv_less_ivl* is dual to the “ \geq ” case but adjusted by 1 because the ordering is strict.

13.8.4 Abstract Interpretation

Now we instantiate the abstract interpreter interface with the interval domain: $\text{num}' = \lambda i. [\text{Fin } i, \text{Fin } i]$, $\text{plus}' = (+)$, $\text{test_num}' = \text{in_ivl}$ where $\text{in_ivl } i$ $[l, h] = (l \leq \text{Fin } i \wedge \text{Fin } i \leq h)$, $\text{inv_plus}' = \text{inv_plus_ivl}$, $\text{inv_less}' = \text{inv_less_ivl}$. In the preceding subsections we have already discussed that all requirements hold.

Let us now analyse some concrete programs. We mostly show the final results of the analyses only. For a start, we can do better than the naive constant propagation in [Section 13.6.2](#):

```

''x'' ::= N 42 {Some {''x'', [42, 42]}};;
IF Less (N 41) (V ''x'')
THEN {Some {''x'', [42, 42]}} ''x'' ::= N 5 {Some {''x'', [5, 5]}}
ELSE {None} ''x'' ::= N 6 {None}
{Some {''x'', [5, 5]}}

```

This time the analysis detects that the *ELSE* branch is unreachable. Of course the constant propagation in [Section 10.2](#) can deal with this example equally

well. Now we look at examples beyond constant propagation. Here is one that demonstrates the analysis of boolean expressions very well:

```

''y'' ::= N 7 {Some {''x'', [-∞, ∞]}, (''y'', [7, 7])};;
IF Less (V ''x'') (V ''y'')
THEN {Some {''x'', [-∞, 6]}, (''y'', [7, 7])}
    ''y'' ::= Plus (V ''y'') (V ''x'')
    {Some {''x'', [-∞, 6]}, (''y'', [-∞, 13])}
ELSE {Some {''x'', [7, ∞]}, (''y'', [7, 7])}
    ''x'' ::= Plus (V ''x'') (V ''y'')
    {Some {''x'', [14, ∞]}, (''y'', [7, 7])}
{Some {''x'', [-∞, ∞]}, (''y'', [-∞, 13])}

```

Let us now analyse some *WHILE* loops.

```

{Some {''x'', [-∞, ∞])}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', [-∞, 99])}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [-∞, 100])}
{Some {''x'', [100, ∞])}

```

This is very nice; all the intervals are precise and the analysis only takes three steps to stabilize. Unfortunately this is a coincidence. The analysis of almost the same program, but with x initialized, is less well behaved. After five steps we have

```

''x'' ::= N 0 {Some {''x'', [0, 0])};;
{Some {''x'', [0, 1])}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', [0, 0])}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, 1])}
{None}

```

The interval $[0, 1]$ will increase slowly until it reaches the invariant $[0, 100]$:

```

''x'' ::= N 0 {Some {''x'', [0, 0])};;
{Some {''x'', [0, 100])}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', [0, 99])}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, 100])}
{Some {''x'', [100, 100])}

```

(13.1)

The result is as precise as possible, i.e., a least fixpoint, but it takes a while to get there.

In general, the analysis of this loop, for an upper bound of n instead of 100, takes $\Theta(n)$ steps (if $n \geq 0$). This is not nice. We might as well run the

program directly (which this analysis more or less does). Therefore it is not surprising that the analysis of a nonterminating program may not terminate either. The following annotated command is the result of 50 steps:

$$\begin{aligned}
 &''x'' ::= N\ 0\ \{Some\ \{(''x'',\ [0,\ 0])\}\}; \\
 &\{Some\ \{(''x'',\ [0,\ 16])\}\} \\
 &WHILE\ Less\ (N\ (-\ 1))\ (V\ ''x'') \\
 &DO\ \{Some\ \{(''x'',\ [0,\ 15])\}\} \\
 &''x'' ::= Plus\ (V\ ''x'')\ (N\ 1)\ \{Some\ \{(''x'',\ [1,\ 16])\}\} \\
 &\{None\}
 \end{aligned} \tag{13.2}$$

The interval for x keeps increasing indefinitely. The problem is that *ivl* is not of finite height: $[0, 0] < [0, 1] < [0, 2] < \dots$. The Alexandrian solution to this is quite brutal: if the analysis sees the beginning of a possibly infinite ascending chain, it overapproximates wildly and jumps to a much larger point, namely $[0, \infty]$ in this case. An even cruder solution would be to jump directly to \top to avoid nontermination.

Exercises

Exercise 13.19. Construct a terminating program where interval analysis does not terminate. Hint: use a loop with two initialized variables; remember that our analyses cannot keep track of relationships between variables.

13.9 Widening and Narrowing thy

13.9.1 Widening

Widening abstracts the idea sketched at the end of the previous section: if an iteration appears to diverge, jump upwards in the ordering to avoid nontermination or at least accelerate termination, possibly at the cost of precision. This is the purpose of the overloaded **widening operator**:

$$\begin{aligned}
 &(\nabla) :: 'a \Rightarrow 'a \Rightarrow 'a \quad \text{such that} \\
 &x \leq x \nabla y \quad \text{and} \quad y \leq x \nabla y
 \end{aligned}$$

which is equivalent to $x \sqcup y \leq x \nabla y$. This property is only needed for the termination proof later on. We assume that the abstract domain provides a widening operator.

Iteration with widening means replacing $x_{i+1} = f\ x_i$ by $x_{i+1} = x_i \nabla f\ x_i$. That is, we apply widening in each step. Pre-fixpoint iteration with widening is defined for any type $'a$ that provides \leq and ∇ :

definition $iter_widen :: ('a \Rightarrow 'a) \Rightarrow 'a \Rightarrow 'a$ *option* **where**
 $iter_widen\ f = while_option\ (\lambda x. \neg f\ x \leq x)\ (\lambda x. x \nabla f\ x)$

Comparing $iter_widen$ with pfp , we don't iterate f but $\lambda x. x \nabla f\ x$. The condition $\neg f\ x \leq x$ still guarantees that if we terminate we have obtained a pre-fixpoint, but it no longer needs to be the least one because widening may have jumped over it. That is, we trade precision for termination. As a consequence, [Lemma 10.28](#) no longer applies and for the first time we may actually obtain a pre-fixpoint that is not a fixpoint.

This behaviour of widening is visualized in [Figure 13.11](#). The normal iteration of f takes a number of steps to get close to the least fixpoint of f (where f intersects with the diagonal) and it is not clear if it will reach it in a finite number of steps. Iteration with widening, however, boldly jumps over the least fixpoint to some pre-fixpoint. Note that any x where $f\ x$ is below the diagonal is a pre-fixpoint.

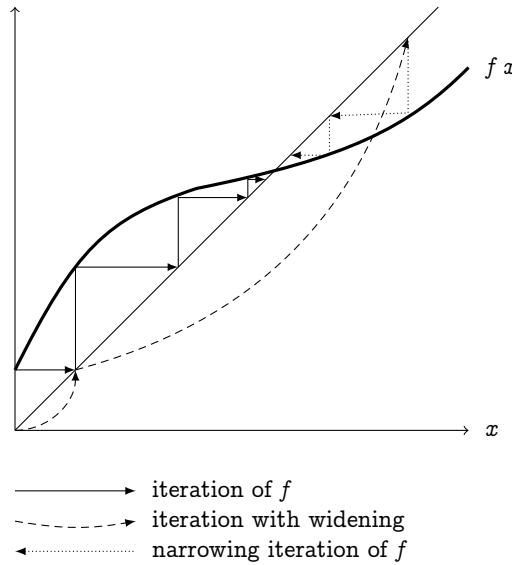


Fig. 13.11. Widening and narrowing

Example 13.37. A simple instance of widening for intervals compares two intervals and jumps to ∞ if the upper bound increases and to $-\infty$ if the lower bound decreases:

$$\begin{aligned}
[l_1, h_1] \nabla [l_2, h_2] &= [l, h] \\
\text{where } l &= (\text{if } l_1 > l_2 \text{ then } -\infty \text{ else } l_1) \\
h &= (\text{if } h_1 < h_2 \text{ then } \infty \text{ else } h_1)
\end{aligned}$$

For example: $[0, 1] \nabla [0, 2] = [0, \infty]$
 $[0, 2] \nabla [0, 1] = [0, 2]$
 $[1, 2] \nabla [0, 5] = [-\infty, \infty]$

The first two examples show that although the symbol ∇ looks symmetric, the operator need not be commutative: its two arguments are the previous and the next value in an iteration and it is important which one is which.

The termination argument for *iter_widen* on intervals goes roughly like this: if we have not reached a pre-fixpoint yet, i.e., $\neg [l_2, h_2] \leq [l_1, h_1]$, then either $l_1 > l_2$ or $h_1 < h_2$ and hence at least one of the two bounds jumps to infinity, which can happen at most twice.

Widening operators can be lifted from *'a* to other types:

- *'a st*: If *st* is the function space (Section 13.5), then
 $S_1 \nabla S_2 = (\lambda x. S_1 \ x \nabla S_2 \ x)$

For the executable version of *st* in Section 13.6 it is

$$St \ ps_1 \nabla St \ ps_2 = St(\text{map2_st_rep } (\nabla) \ ps_1 \ ps_2)$$

- *'a option*:

$$\begin{aligned}
None \nabla x &= x \\
x \nabla None &= x \\
Some \ x \nabla Some \ y &= Some \ (x \nabla y)
\end{aligned}$$

- *'a acom*:

$$C_1 \nabla C_2 = \text{map2_acom } (\nabla) \ C_1 \ C_2$$

where *map2_acom f* combines two (structurally equal) commands into one by combining annotations at the same program point with *f*:

$$\begin{aligned}
\text{map2_acom } f \ C_1 \ C_2 &= \\
&\text{annotate } (\lambda p. f \ (\text{anno } C_1 \ p) \ (\text{anno } C_2 \ p)) \ (\text{strip } C_1)
\end{aligned}$$

Hence we can now perform widening and thus *iter_widen* on our annotated commands of type *'av st option acom*. We demonstrate the effect on two example programs. That is, we iterate $\lambda C. C \nabla \text{step_ivl} \top C$ where *step_ivl* is the specialization of *step'* for intervals that came out of Section 13.8.4 (the name *step_ivl* was not mentioned there). The only difference to Section 13.8.4 is the widening after each step.

Our first example is the nonterminating loop (13.2) from the previous section. After four steps it now looks like this:

```

''x'' ::= N 0 {Some {''x'', [0, 0]}};;
{Some {''x'', [0, 0]}}
WHILE Less (N (- 1)) (V ''x'')
DO {Some {''x'', [0, 0]}}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, 1]}}
{None}

```

Previously, in the next step the annotation at the head of the loop would change from $[0, 0]$ to $[0, 1]$; now this is followed by widening. Because $[0, 0] \nabla [0, 1] = [0, \infty]$ we obtain

```

''x'' ::= N 0 {Some {''x'', [0, 0]}};;
{Some {''x'', [0, \infty]}}
WHILE Less (N (- 1)) (V ''x'')
DO {Some {''x'', [0, 0]}}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, 1]}}
{None}

```

Two more steps and we have reached a pre-fixpoint:

```

''x'' ::= N 0 {Some {''x'', [0, 0]}};;
{Some {''x'', [0, \infty]}}
WHILE Less (N (- 1)) (V ''x'')
DO {Some {''x'', [0, \infty]}}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, \infty]}}
{None}

```

This is very nice because we have actually reached the least fixpoint.

The details of what happened are shown in Table 13.1, where A_0 to A_4 are the five annotations from top to bottom. We start with the situation after four steps. Each iteration step is displayed as a block of two columns: first the

		f	∇	f	∇	f	∇
A_0	$[0, 0]$						$[0, 0]$
A_1	$[0, 0]$	$[0, 1]$	$[0, \infty]$	$[0, 1]$	$[0, \infty]$	$[0, \infty]$	$[0, \infty]$
A_2	$[0, 0]$			$[0, \infty]$	$[0, \infty]$	$[0, \infty]$	$[0, \infty]$
A_3	$[1, 1]$					$[1, \infty]$	$[1, \infty]$
A_4	<i>None</i>						<i>None</i>

Table 13.1. Widening example

result of applying the step function to the previous column, then the result of widening that result with the previous column. The last column is the least pre-fixpoint. Empty entries mean that nothing has changed.

Now we look at the program where previously we needed $\Theta(n)$ steps to reach the least fixpoint. Now we reach a pre-fixpoint very quickly:

```

''x'' ::= N 0 {Some {''x'', [0, 0]}};;
{Some {''x'', [0, ∞]}}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', [0, ∞]}}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, ∞]}}
{Some {''x'', [100, ∞]}}

```

(13.3)

This takes only a constant number of steps, independent of the upper bound of the loop, but the result is worse than previously (13.1): precise upper bounds have been replaced by ∞ .

13.9.2 Narrowing

Narrowing is the attempt to improve the potentially imprecise pre-fixpoint p obtained by widening. It assumes f is monotone. Then we can just iterate f :

$$p \geq f(p) \geq f^2(p) \geq \dots$$

Each $f^i(p)$ is still a pre-fixpoint, and we can stop improving any time, especially if we reach a fixpoint. In Figure 13.11 you can see the effect of two such iterations of f back from an imprecise pre-fixpoint towards the least fixpoint. As an example, start from the imprecise widening analysis (13.3): four applications of $step_widen \top$ take us to the least fixpoint (13.1) that we had computed without widening and narrowing in many more steps.

We may have to stop iterating f before we reach a fixpoint if we want to terminate (quickly) because there may be infinitely descending chains:

$$[0, \infty] > [1, \infty] > [2, \infty] > \dots$$

This is where the overloaded **narrowing operator** comes in:

$$\begin{aligned}
(\triangle) &:: 'a \Rightarrow 'a \Rightarrow 'a \quad \text{such that} \\
y \leq x &\implies y \leq x \triangle y \leq x
\end{aligned}$$

We assume that the abstract domain provides a narrowing operator.

Iteration with narrowing means replacing $x_{i+1} = f x_i$ by $x_{i+1} = x_i \triangle f x_i$. Now assume that x_i is a pre-fixpoint of f . Then $x_i \triangle f x_i$ is again a pre-fixpoint of f : because $f x_i \leq x_i$ and \triangle is a narrowing operator we have $f x_i \leq x_i \triangle f x_i \leq x_i$ and hence $f (x_i \triangle f x_i) \leq f x_i \leq x_i \triangle f x_i$ by monotonicity. That is, iteration of a monotone function with a narrowing operator preserves the pre-fixpoint property.

definition $iter_narrow :: ('a \Rightarrow 'a) \Rightarrow 'a \Rightarrow 'a$ option where
 $iter_narrow f = while_option (\lambda x. x \triangle f x < x) (\lambda x. x \triangle f x)$

In contrast to widening, we are not looking for a pre-fixpoint (if we start from a pre-fixpoint, it remains a pre-fixpoint), but we terminate as soon as we no longer make progress, i.e., no longer strictly decrease. The narrowing operator can enforce termination by making $x \triangle f x$ return x .

Example 13.38. The narrowing operator for intervals only changes one of the interval bounds if it can be improved from infinity to some finite value:

$$\begin{aligned} [l_1, h_1] \triangle [l_2, h_2] &= [l, h] \\ \text{where } l &= (\text{if } l_1 = -\infty \text{ then } l_2 \text{ else } l_1) \\ h &= (\text{if } h_1 = \infty \text{ then } h_2 \text{ else } h_1) \end{aligned}$$

For example: $[0, \infty] \triangle [0, 100] = [0, 100]$
 $[0, 100] \triangle [0, 0] = [0, 100]$
 $[0, 0] \triangle [0, 100] = [0, 0]$

Narrowing operators need not be commutative either.

The termination argument for *iter_narrow* on intervals is very intuitive: finite interval bounds remain unchanged; therefore it takes at most two steps until both bounds have become finite and *iter_narrow* terminates.

Narrowing operators are lifted to *'a st* and *'a acom* verbatim like widening operators, with ∇ replaced by \triangle . On *'a option* narrowing is dual to widening:

$$\begin{aligned} \text{None} \triangle x &= \text{None} \\ x \triangle \text{None} &= \text{None} \\ \text{Some } x \triangle \text{Some } y &= \text{Some } (x \triangle y) \end{aligned}$$

Hence we can now perform narrowing with *iter_narrow* on our annotated commands of type *'av st option acom*. Starting from the imprecise result (13.3) of widening, *iter_narrow (step_ivl \top)* also reaches the least fixpoint (13.1) that iterating *step_ivl \top* had reached without the narrowing operator (see above). In general the narrowing operator is required to ensure termination but it may also lose precision by terminating early.

13.9.3 Abstract Interpretation

The abstract interpreter with widening and narrowing is defined like *AI*

$$AI_wn\ c = pfp_wn\ (step' \top)\ (bot\ c)$$

but instead of the simple iterator *pfp* the composition of widening and narrowing is used:

$$\begin{aligned} pfp_wn\ f\ x &= \\ &(\text{case } iter_widen\ f\ x\ \text{of } \text{None} \Rightarrow \text{None} \mid \text{Some } p \Rightarrow iter_narrow\ f\ p) \end{aligned}$$

In the monotone framework we can show partial correctness of *AI_wn*:

Lemma 13.39. $AI_wn\ c = Some\ C \implies CS\ c \leq \gamma_c\ C$

Proof. The proof is the same as for Lemma 13.25, but now we need that if $pfp_wn\ (step' \top)\ C = Some\ C'$, then C' is a pre-fixpoint. For $iter_widen$ this is guaranteed by the exit condition $f\ x \leq x$. Above we have shown that if f is monotone, then the pre-fixpoint property is an invariant of $iter_narrow\ f$. Because $f = step' \top$ is monotone in the monotone framework, the narrowing phase returns a pre-fixpoint. Thus pfp_wn returns a pre-fixpoint if it terminates (returns *Some*). \square

13.9.4 Termination of Widening

The termination proof for $iter_widen$ is similar to the termination proof in Section 13.5.8. We again require a measure function $m :: 'av \Rightarrow nat$ and a height $h :: nat$ that bounds m : $m\ a \leq h$. The key requirement is that if we have not reached a pre-fixpoint yet, widening strictly decreases m :

$$\neg a_2 \leq a_1 \implies m\ (a_1 \nabla a_2) < m\ a_1 \quad (13.4)$$

Clearly there can only be h many widening steps.

In addition we require anti-monotonicity w.r.t. \leq , not $<$:

$$a_1 \leq a_2 \implies m\ a_1 \geq m\ a_2$$

Note that this does not mean that the ordering is of finite height. Otherwise we could not apply it to intervals further down.

Let us first sketch the intuition behind the termination proof. In Section 13.5.8 we relied on monotonicity to argue that pre-fixpoint iteration gives rise to a strictly increasing chain. Unfortunately, $\lambda x. x \nabla f\ x$ need not be monotone, even if f is. For example, on intervals let f be constantly $[0,1]$. Then $[0,0] \leq [0,1]$ but $[0,0] \nabla f\ [0,0] = [0,\infty] \not\leq [0,1] = [0,1] \nabla f\ [0,1]$. Nevertheless $iter_widen$ gives rise to a strictly decreasing m -chain: if $\neg f\ x \leq x$ then $m\ (x \nabla f\ x) < m\ x$ by (13.4). This is on $'av$. Therefore we need to lift (13.4) to annotated commands. Think in terms of tuples (a_1, a_2, \dots) of abstract values and let the termination measure m' be the sum of all $m\ a_i$, just as in Figure 13.8 and the surrounding explanation. Then $\neg (b_1, b_2, \dots) \leq (a_1, a_2, \dots)$ means that there is a k such that $\neg b_k \leq a_k$. We have $a_i \leq a_i \nabla b_i$ for all i because ∇ is a widening operator. Therefore $m(a_i \nabla b_i) \leq m\ a_i$ for all i (by anti-monotonicity) and $m(a_k \nabla b_k) < m\ a_k$ (by (13.4)). Thus $m'((a_1, a_2, \dots) \nabla (b_1, b_2, \dots)) = m'(a_1 \nabla b_1, a_2 \nabla b_2, \dots) < m'(a_1, a_2, \dots)$. That is, (13.4) holds on tuples too. Now for the technical details.

Recall from Section 13.5.8 that we lifted m to m_s , m_o and m_c . Following the sketch above, (13.4) can be proved for these derived functions too, with a few side conditions:

$$\begin{aligned}
& \llbracket \text{finite } X; \text{fun } S_1 = \text{fun } S_2 \text{ on } -X; \neg S_2 \leq S_1 \rrbracket \\
& \implies m_s (S_1 \nabla S_2) X < m_s S_1 X \\
& \llbracket \text{finite } X; \text{top_on}_o S_1 (-X); \text{top_on}_o S_2 (-X); \neg S_2 \leq S_1 \rrbracket \\
& \implies m_o (S_1 \nabla S_2) X < m_o S_1 X \\
& \llbracket \text{strip } C_1 = \text{strip } C_2; \text{top_on}_c C_1 (-\text{vars } C_1); \\
& \quad \text{top_on}_c C_2 (-\text{vars } C_2); \neg C_2 \leq C_1 \rrbracket \\
& \implies m_c (C_1 \nabla C_2) < m_c C_1
\end{aligned}$$

The last of the above three lemmas implies termination of *iter_widen* $f C$ for $C :: 'av \text{ st option } acom$ as follows. If you set $C_2 = f C_1$ and assume that f preserves the *strip* and *top_on_c* properties then each step of *iter_widen* $f C$ strictly decreases m_c , which must terminate because m_c returns a *nat*.

To apply this result to widening on intervals we merely need to provide the right m and h :

$$\begin{aligned}
m_ivl [l, h] = \\
& (\text{if } [l, h] = \perp \text{ then } 3 \\
& \quad \text{else } (\text{if } l = -\infty \text{ then } 0 \text{ else } 1) + (\text{if } h = \infty \text{ then } 0 \text{ else } 1))
\end{aligned}$$

Strictly speaking m does not need to consider the \perp case because we have made sure it cannot arise in an abstract state, but it is simpler not to rely on this and cover \perp .

Function m_ivl satisfies the required properties: $m_ivl \text{ iv} \leq 3$ and $y \leq x \implies m_ivl x \leq m_ivl y$.

Because *bot* suitably initializes and *step_ivl* preserves the *strip* and *top_on_c* properties, this implies termination of *iter_widen* (*step_ivl* \top):

$$\exists C. \text{iter_widen } (\text{step_ivl } \top) (\text{bot } c) = \text{Some } C$$

13.9.5 Termination of Narrowing

The proof is similar to that for widening. We require another measure function

$$\begin{aligned}
n :: 'av \Rightarrow \text{nat} \quad \text{such that} \\
\llbracket a_2 \leq a_1; a_1 \triangle a_2 < a_1 \rrbracket \implies n (a_1 \triangle a_2) < n a_1
\end{aligned} \tag{13.5}$$

This property guarantees that the measure goes down with every iteration of *iter_narrow* $f a_0$, provided f is monotone and $f a_0 \leq a_0$: let $a_2 = f a_1$; then $a_2 \leq a_1$ is the pre-fixpoint property that iteration with narrowing preserves (see [Section 13.9.2](#)) and $a_1 \triangle a_2 < a_1$ is the loop condition.

Now we need to lift this from *'av* to other types. First we sketch how it works for tuples. Define the termination measure $n'(a_1, a_2, \dots) = n a_1 + n a_2 + \dots$. To show that (13.5) holds for n' too, assume $(b_1, b_2, \dots) \leq (a_1, a_2, \dots)$ and $(a_1, a_2, \dots) \triangle (b_1, b_2, \dots) < (a_1, a_2, \dots)$, i.e., $b_i \leq a_i$ for all i ,

$a_i \triangle b_i \leq a_i$ for all i and $a_k \triangle b_k \leq a_k$ for some k . Thus for all i either $a_i \triangle b_i = a_i$ (and hence $n(a_i \triangle b_i) = n a_i$) or $a_i \triangle b_i < a_i$, which together with $b_i \leq a_i$ yields $n(a_i \triangle b_i) < n a_i$ by (13.5). Thus $n(a_i \triangle b_i) \leq n a_i$ for all i . But $n(a_k \triangle b_k) < n a_k$, also by (13.5), and hence $n'((a_1, a_2, \dots) \triangle (b_1, b_2, \dots)) = n'(a_1 \triangle b_1, a_2 \triangle b_2, \dots) < n'(a_1, a_2, \dots)$. Thus (13.5) holds for n' too. Now for the technical details.

We lift n in three stages to *'av st option acom*:

definition $n_s :: 'av\ st \Rightarrow vname\ set \Rightarrow nat\ where$
 $n_s\ S\ X = (\sum_{x \in X}. n\ (fun\ S\ x))$

fun $n_o :: 'av\ st\ option \Rightarrow vname\ set \Rightarrow nat\ where$
 $n_o\ None\ X = 0$
 $n_o\ (Some\ S)\ X = n_s\ S\ X + 1$

definition $n_c :: 'av\ st\ option\ acom \Rightarrow nat\ where$
 $n_c\ C = (\sum a \leftarrow annos\ C. n_o\ a\ (vars\ C))$

Property (13.5) carries over to *'av st option acom* with suitable side conditions just as (13.4) carried over in the previous subsection:

$$\begin{aligned} & \llbracket strip\ C_1 = strip\ C_2; top_on_c\ C_1\ (-\ vars\ C_1); \\ & \quad top_on_c\ C_2\ (-\ vars\ C_2); C_2 \leq C_1; C_1 \triangle C_2 < C_1 \rrbracket \\ & \implies n_c\ (C_1 \triangle C_2) < n_c\ C_1 \end{aligned}$$

This implies termination of $iter_narrow\ f\ C$ for $C :: 'av\ st\ option\ acom$ provided f is monotone and C is a pre-fixpoint of f (which is preserved by narrowing because f is monotone and \triangle a narrowing operator) because it guarantees that each step of $iter_narrow\ f\ C$ strictly decreases n_c .

To apply this result to narrowing on intervals we merely need to provide the right n :

$$n_ivl\ iv = 3 - m_ivl\ iv$$

It does what it is supposed to do, namely satisfy (a strengthened version of) (13.5): $x \triangle y < x \implies n_ivl\ (x \triangle y) < n_ivl\ x$. Therefore narrowing terminates provided we start with a pre-fixpoint which is \top outside its variables:

$$\begin{aligned} & \llbracket top_on_c\ C\ (-\ vars\ C); step_ivl\ \top\ C \leq C \rrbracket \\ & \implies \exists C'. iter_narrow\ (step_ivl\ \top)\ C = Some\ C' \end{aligned}$$

Because both preconditions are guaranteed by the output of widening we obtain the final termination theorem where AI_wn_ivl is AI_wn for intervals:

Theorem 13.40. $\exists C. AI_wn_ivl\ c = Some\ C$

Exercises

Exercise 13.20. Starting from

```

''x'' ::= N 0 {Some {''x'', [0, 0]}};;
{Some {''x'', [0, 0]}}
WHILE Less (V ''x'') (N 100)
DO {Some {''x'', [0, 0]}}
    ''x'' ::= Plus (V ''x'') (N 1) {Some {''x'', [1, 1]}}
{None}

```

tabulate the three steps that *step_ivl* with widening takes to reach (13.3). Follow the format of Table 13.1.

Exercise 13.21. Starting from (13.3), tabulate both the repeated application of *step_ivl* \top alone and of *iter_narrow* (*step_ivl* \top), the latter following the format of Table 13.1 (with ∇ replaced by \triangle).

13.10 Summary and Further Reading

This concludes the chapter on abstract interpretation. The main points were:

- The reference point is a collecting semantics that describes for each program point which sets of states can arise at that point.
- The abstract interpreter mimics the collecting semantics but operates on abstract values representing (infinite) sets of concrete values. Abstract values should form a lattice: \sqcup abstracts \cup and models the effect of joining two computation paths, \sqcap can reduce abstract values to model backward analysis of boolean expressions.
- The collecting semantics is defined as a least fixpoint of a step function. The abstract interpreter approximates this by iterating its step function. For monotone step functions this iteration terminates if the ordering is of finite height (or at least there is no infinitely ascending chain).
- Widening accelerates the iteration process to guarantee termination even if there are infinitely ascending chains in the ordering, for example on intervals. It trades precision for termination. Narrowing tries to improve the precision lost by widening by further iterations.

Of the analyses in Chapter 10 we have only rephrased constant propagation as abstract interpretation. Definite initialization can be covered if the collecting semantics is extended to distinguish initialized from uninitialized variables. Our simple framework cannot accommodate live variable analysis for two reasons. Firstly, it is a backward analysis, whereas our abstract interpreter is a forward analyser. Secondly, it requires a different semantics:

liveness of a variable at some point is not a property of the possible values of that variable at that point but of the operations executed after that point.

13.10.1 Further Reading

Abstract Interpretation was invented by Patrick and Radhia Cousot [22]. In its standard form, the concrete and abstract level are related by a concretisation function γ together with an adjoint abstraction function α . This allows the verification not just of correctness but also of optimality of an abstract interpreter. In fact, one can even “calculate” the abstract interpreter [21]. The book by Nielson, Nielson and Hankin [61] also has a chapter on abstract interpretation.

There are many variations, generalizations and applications of abstract interpretation. Particularly important are relational analyses which can deal with relations between program variables: they do not abstract $(vname \Rightarrow val) \text{ set}$ to $vname \Rightarrow val \text{ set}$ as we have done. The canonical example is polyhedral analysis that generalizes interval analysis to linear inequalities between variables [23].

Our intentionally simple approach to abstract interpretation can be improved in a number of respects:

- We combine the program and the annotations into one data structure and perform all computations on it. Normally the program points are labelled and the abstract interpreter operates on a separate mapping from labels to annotations.
- We iterate globally over the whole program. Precision is improved by a compositional iteration strategy: the global fixpoint iteration is replaced by a fixpoint iteration for each loop.
- We perform widening and narrowing for all program points. Precision can be improved if we restrict ourselves to one program point per loop.

Cachera and Pichardie [17] have formalized an abstract interpreter in Coq that follows all three improvements. Iteration strategies and widening points have been investigated by Bourdoncle [14] for arbitrary control-flow graphs.

Our abstract interpreter operates on a single data structure, an annotated command. A more modular design transforms the pre-fixpoint requirement $step' \top C \leq C$ into a set of inequalities over the abstract domain and solves those inequalities by some arbitrary method. This separates the programming language aspect from the mathematics of solving particular types of constraints. For example, Gawlitza and Seidl [33] showed how to compute least solutions of interval constraints precisely, without any widening and narrowing.

A

Auxiliary Definitions

This appendix contains auxiliary definitions omitted from the main text.

Variables

```
fun lvars :: com  $\Rightarrow$  vname set where
lvars SKIP                                = {}
lvars (x ::= e)                          = {x}
lvars (c1;; c2)                        = lvars c1  $\cup$  lvars c2
lvars (IF b THEN c1 ELSE c2) = lvars c1  $\cup$  lvars c2
lvars (WHILE b DO c)                = lvars c

fun rvars :: com  $\Rightarrow$  vname set where
rvars SKIP                                = {}
rvars (x ::= e)                          = vars e
rvars (c1;; c2)                        = rvars c1  $\cup$  rvars c2
rvars (IF b THEN c1 ELSE c2) = vars b  $\cup$  rvars c1  $\cup$  rvars c2
rvars (WHILE b DO c)                = vars b  $\cup$  rvars c

definition vars :: com  $\Rightarrow$  vname set where
vars c = lvars c  $\cup$  rvars c
```

Abstract Interpretation

```
fun strip :: 'a acom  $\Rightarrow$  com where
strip (SKIP {P}) = SKIP
strip (x ::= e {P}) = x ::= e
```

```

strip (C1;;C2) = strip C1;; strip C2
strip (IF b THEN {P1} C1 ELSE {P2} C2 {P}) =
  IF b THEN strip C1 ELSE strip C2
strip ({I} WHILE b DO {P} C {Q}) = WHILE b DO strip C

fun annos :: 'a acom ⇒ 'a list where
annos (SKIP {P}) = [P]
annos (x ::= e {P}) = [P]
annos (C1;;C2) = annos C1 @ annos C2
annos (IF b THEN {P1} C1 ELSE {P2} C2 {Q}) =
  P1 # annos C1 @ P2 # annos C2 @ [Q]
annos ({I} WHILE b DO {P} C {Q}) = I # P # annos C @ [Q]

fun asize :: com ⇒ nat where
asize SKIP = 1
asize (x ::= e) = 1
asize (C1;;C2) = asize C1 + asize C2
asize (IF b THEN C1 ELSE C2) = asize C1 + asize C2 + 3
asize (WHILE b DO C) = asize C + 3

definition shift :: (nat ⇒ 'a) ⇒ nat ⇒ nat ⇒ 'a where
shift f n = (λp. f(p+n))

fun annotate :: (nat ⇒ 'a) ⇒ com ⇒ 'a acom where
annotate f SKIP = SKIP {f 0}
annotate f (x ::= e) = x ::= e {f 0}
annotate f (c1;;c2) = annotate f c1;; annotate (shift f (asize c1)) c2
annotate f (IF b THEN c1 ELSE c2) =
  IF b THEN {f 0} annotate (shift f 1) c1
  ELSE {f(asize c1 + 1)} annotate (shift f (asize c1 + 2)) c2
  {f(asize c1 + asize c2 + 2)}
annotate f (WHILE b DO c) =
  {f 0} WHILE b DO {f 1} annotate (shift f 2) c {f(asize c + 2)}

fun map_acom :: ('a ⇒ 'b) ⇒ 'a acom ⇒ 'b acom where
map_acom f (SKIP {P}) = SKIP {f P}
map_acom f (x ::= e {P}) = x ::= e {f P}
map_acom f (C1;;C2) = map_acom f C1;; map_acom f C2
map_acom f (IF b THEN {P1} C1 ELSE {P2} C2 {Q}) =
  IF b THEN {f P1} map_acom f C1 ELSE {f P2} map_acom f C2
  {f Q}
map_acom f ({I} WHILE b DO {P} C {Q}) =
  {f I} WHILE b DO {f P} map_acom f C {f Q}

```

B

Symbols

[[\<lbrakk>
]]	\<rbrakk>
\Rightarrow	\Rightarrow	\<Longrightarrow>
\wedge	!!	\<And>
\equiv	\equiv	\<equiv>
λ	%	\<lambda>
\Rightarrow	\Rightarrow	\<Rightarrow>
\wedge	&	\<and>
\vee		\<or>
\longrightarrow	\longrightarrow	\<longrightarrow>
\rightarrow	\rightarrow	\<rightarrow>
\neg	~	\<not>
\neq	\neq	\<noteq>
\forall	ALL	\<forall>
\exists	EX	\<exists>
\leq	\leq	\<le>
\times	*	\<times>
\in	:	\<in>
\notin	~:	\<notin>
\subseteq	\subseteq	\<subseteq>
\subset	<	\<subset>
\cup	Un	\<union>
\cap	Int	\<inter>
\bigcup	UN, Union	\<Union>
\bigcap	INT, Inter	\<Inter>
\sqcup	sup	\<squnion>
\sqcap	inf	\<sqinter>
\bigsqcup	SUP, Sup	\<Squnion>
\bigsqcap	INF, Inf	\<Sqinter>
\top		\<top>
\perp		\<bottom>

Table B.1. Mathematical symbols, their ASCII equivalents and internal names

C

Theories

The following table shows which sections are based on which theories in the directory `src/HOL/IMP` of the Isabelle distribution.

3.1	AExp	12.1	Hoare_Examples
3.2	BExp	12.2.2	Hoare
3.3	ASM	12.2.3	Hoare_Examples
7.1	Com	12.3	Hoare_Sound_Complete
7.2	Big_Step	12.4	VCG
7.3	Small_Step	12.5	Hoare_Total
8.1	Compiler	13.2	ACom
8.2	Compiler	13.3	Collecting
8.3	Compiler	13.3.3	Complete_Lattice
8.4	Compiler2	13.4	Abs_Int1_parity
9.1	Types	13.4.2	Abs_Int0
9.2.1	Sec_Type_Expr	13.5	Abs_Int0
9.2.2	Sec_Typing	13.5.1	Collecting
9.2.6	Sec_TypingT	13.6	Abs_Int1
10.1.1	Def_Init	13.6.1	Abs_Int1_parity
10.1.2	Def_Init_Exp	13.6.2	Abs_Int1_const
10.1.3	Def_Init_Small	13.6.3	Abs_State
10.1.4	Def_Init_Big	13.7	Abs_Int2
10.2	Fold	13.8	Abs_Int2_ivl
10.3	Live	13.9	Abs_Int3
10.4	Live_True		
11.0	Denotational		

References

1. Samson Abramsky and Achim Jung. Domain theory. In S. Abramsky, D. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3, pages 1–168. Oxford University Press, 1994.
2. Alfred Aho, Monica Lam, Ravi Sethi, and Jeffrey Ullman. *Compilers: Principles, Techniques, & Tools*. Addison-Wesley, 2nd edition, 2007.
3. Eyad Alkassar, Mark Hillebrand, Dirk Leinenbach, Norbert Schirmer, Artem Starostin, and Alexandra Tsyban. Balancing the load — leveraging a semantics stack for systems verification. *Journal of Automated Reasoning: Special Issue on Operating System Verification*, 42, Numbers 2–4:389–454, 2009.
4. Eyad Alkassar, Mark Hillebrand, Wolfgang Paul, and Elena Petrova. Automated verification of a small hypervisor. In Gary Leavens, Peter O’Hearn, and Sriram Rajamani, editors, *Proceedings of Verified Software: Theories, Tools and Experiments 2010*, volume 6217 of *LNCS*, pages 40–54. Springer, 2010.
5. Pierre America and Frank de Boer. Proving total correctness of recursive procedures. *Information and Computation*, 84:129–162, 1990.
6. Krzysztof Apt. Ten Years of Hoare’s Logic: A Survey — Part I. *ACM Trans. Program. Lang. Syst.*, 3(4):431–483, 1981.
7. Krzysztof Apt. Ten Years of Hoare’s Logic: A Survey — Part II: Nondeterminism. *Theoretical Computer Science*, 28:83–109, 1984.
8. Krzysztof Apt, Frank de Boer, and Ernst-Rüdiger Olderog. *Verification of Sequential and Concurrent Programs*. Springer, 3rd edition, 2009.
9. Clemens Ballarin. *Tutorial on Locales and Locale Interpretation*. <http://isabelle.in.tum.de/doc/locales.pdf>.
10. Bernhard Beckert, Reiner Hähnle, and Peter H. Schmitt, editors. *Verification of Object-Oriented Software: The KeY Approach*, volume 4334 of *LNCS*. Springer, 2007.
11. Nick Benton, Andrew Kennedy, and Carsten Varming. Some domain theory and denotational semantics in Coq. In S. Berghofer, T. Nipkow, C. Urban, and M. Wenzel, editors, *Theorem Proving in Higher Order Logics (TPHOLs 2009)*, volume 5674 of *LNCS*, pages 115–130. Springer, 2009.
12. William R. Bevier, Warren A. Hunt Jr., J. Strother Moore, and William D. Young. An approach to systems verification. *J. Autom. Reasoning*, 5(4):411–428, 1989.

13. Richard Bornat. Proving pointer programs in Hoare Logic. In R. Backhouse and J. Oliveira, editors, *Mathematics of Program Construction (MPC 2000)*, volume 1837 of *LNCS*, pages 102–126. Springer, 2000.
14. François Bourdoncle. Efficient chaotic iteration strategies with widenings. In D. Bjørner, Manfred M. Broy, and I. Pottosin, editors, *Formal Methods in Programming and Their Applications*, volume 735 of *LNCS*, pages 128–141. Springer, 1993.
15. David Brumley and Dan Boneh. Remote timing attacks are practical. *Computer Networks*, 48(5):701–716, 2005.
16. Rod Burstall. Some techniques for proving correctness of programs which alter data structures. In B. Meltzer and D. Michie, editors, *Machine Intelligence 7*, pages 23–50. Edinburgh University Press, 1972.
17. David Cachera and David Pichardie. A certified denotational abstract interpreter. In M. Kaufmann and L. Paulson, editors, *Interactive Theorem Proving (ITP 2010)*, volume 6172 of *LNCS*, pages 9–24. Springer, 2010.
18. Ellis Cohen. Information transmission in computational systems. In *Proceedings of the sixth ACM symposium on Operating systems principles (SOSP'77)*, pages 133–139, West Lafayette, Indiana, USA, 1977. ACM.
19. Ernie Cohen, Markus Dahlweid, Mark Hillebrand, Dirk Leinenbach, Michał Moskal, Thomas Santen, Wolfram Schulte, and Stephan Tobies. VCC: A practical system for verifying concurrent C. In Stefan Berghofer, Tobias Nipkow, Christian Urban, and Markus Wenzel, editors, *Theorem Proving in Higher Order Logics (TPHOLs 2009)*, volume 5674 of *LNCS*, pages 23–42, Munich, Germany, 2009. Springer.
20. Stephen Cook. Soundness and completeness of an axiom system for program verification. *SIAM J. on Computing*, 7:70–90, 1978.
21. Patrick Cousot. The calculational design of a generic abstract interpreter. In Broy and Steinbrüggen, editors, *Calculational System Design*. IOS Press, 1999.
22. Patrick Cousot and Radhia Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fix-points. In *Proc. 4th ACM Symp. Principles of Programming Languages*, pages 238–252, 1977.
23. Patrick Cousot and Nicolas Halbwachs. Automatic discovery of linear restraints among variables of a program. In *Proc. 5th ACM Symp. Principles of Programming Languages*, pages 84–97, 1978.
24. Marc Dumas, Laurence Rideau, and Laurent Théry. A generic library for floating-point numbers and its application to exact computing. In R. Boulton and P. Jackson, editors, *Theorem Proving in Higher Order Logics (TPHOLs 2001)*, volume 2152 of *LNCS*, pages 169–184. Springer, 2001.
25. Leonardo Mendonça de Moura and Nikolaj Bjørner. Z3: An efficient SMT solver. In *TACAS*, volume 4963 of *LNCS*, pages 337–340, Budapest, Hungary, March 2008. Springer.
26. Dorothy E. Denning. A lattice model of secure information flow. *Communications of the ACM*, 19(5):236–243, May 1976.
27. Edsger W. Dijkstra. Go to statement considered harmful. *Communications of the ACM*, 11(3):147–148, March 1968.
28. Edsger W. Dijkstra. *A Discipline of Programming*. Prentice-Hall, 1976.

29. Gidon Ernst, Gerhard Schellhorn, Dominik Haneberg, Jörg Pfähler, and Wolfgang Reif. A formal model of a virtual filesystem switch. In *Proc. 7th SSV*, pages 33–45, 2012.
30. Robert Floyd. Assigning meanings to programs. In J. T. Schwartz, editor, *Mathematical Aspects of Computer Science*, volume 19 of *Proceedings of Symposia in Applied Mathematics*, pages 19–32. American Mathematical Society, 1967.
31. Anthony Fox. Formal specification and verification of ARM6. In David Basin and Burkhart Wolff, editors, *Proceedings of the 16th Int. Conference on Theorem Proving in Higher Order Logics (TPHOLs)*, volume 2758 of *LNCS*, pages 25–40, Rome, Italy, September 2003. Springer.
32. Anthony Fox and Magnus Myreen. A trustworthy monadic formalization of the ARMv7 instruction set architecture. In Matt Kaufmann and Lawrence C. Paulson, editors, *1st Int. Conference on Interactive Theorem Proving (ITP)*, volume 6172 of *LNCS*, pages 243–258, Edinburgh, UK, July 2010. Springer.
33. Thomas Gawlitza and Helmut Seidl. Precise fixpoint computation through strategy iteration. In Rocco De Nicola, editor, *Programming Languages and Systems, ESOP 2007*, volume 4421 of *LNCS*, pages 300–315. Springer, 2007.
34. Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38(1):173–198, 1931.
35. Joseph A. Goguen and José Meseguer. Security policies and security models. In *IEEE Symposium on Security and Privacy*, pages 11–20, 1982.
36. Michael J.C. Gordon. HOL: A machine oriented formulation of higher-order logic. Technical Report 68, University of Cambridge, Computer Laboratory, 1985.
37. Michael J.C. Gordon. Mechanizing programming logics in higher order logic. In G. Birtwistle and P.A. Subrahmanyam, editors, *Current Trends in Hardware Verification and Automated Theorem Proving*. Springer, 1989.
38. James Gosling, Bill Joy, Guy Steele, and Gilad Bracha. *Java(TM) Language Specification, 3rd edition*. Addison-Wesley, 2005.
39. Carl Gunter. *Semantics of programming languages: structures and techniques*. MIT Press, 1992.
40. Florian Haftmann. *Haskell-style type classes with Isabelle/Isar*. <http://isabelle.in.tum.de/doc/classes.pdf>.
41. C.A.R. Hoare. An axiomatic basis for computer programming. *Communications of the ACM*, 12:567–580,583, 1969.
42. John Hopcroft, Rajeev Motwani, and Jeffrey Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, 3rd edition, 2006.
43. Brian Huffman. A purely definitional universal domain. In S. Berghofer, T. Nipkow, C. Urban, and M. Wenzel, editors, *Theorem Proving in Higher Order Logics (TPHOLs 2009)*, volume 5674 of *LNCS*, pages 260–275. Springer, 2009.
44. Michael Huth and Mark Ryan. *Logic in Computer Science*. Cambridge University Press, 2004.
45. Atshushi Igarashi, Benjamin Pierce, and Philip Wadler. Featherweight Java: a minimal core calculus for Java and GJ. In *Proceedings of the 14th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications, OOPSLA '99*, pages 132–146. ACM, 1999.

46. Gilles Kahn. Natural semantics. In *STACS 87: Symp. Theoretical Aspects of Computer Science*, volume 247 of *LNCS*, pages 22–39. Springer, 1987.
47. Gerwin Klein, Kevin Elphinstone, Gernot Heiser, June Andronick, David Cock, Philip Derrin, Dhammika Elkaduwe, Kai Engelhardt, Rafal Kolanski, Michael Norrish, Thomas Sewell, Harvey Tuch, and Simon Winwood. seL4: Formal Verification of an OS Kernel. In Jeanna Neefe Matthews and Thomas E. Anderson, editors, *Proc. 22nd ACM Symposium on Operating Systems Principles 2009*, pages 207–220. ACM, 2009.
48. Gerwin Klein and Tobias Nipkow. A machine-checked model for a Java-like language, virtual machine and compiler. *ACM Trans. Program. Lang. Syst.*, 28(4):619–695, 2006.
49. Alexander Krauss. *Defining Recursive Functions in Isabelle/HOL*. <https://isabelle.in.tum.de/doc/functions.pdf>.
50. Alexander Krauss. Recursive definitions of monadic functions. In A. Bove, E. Komendantskaya, and M. Niqui, editors, *Proc. Workshop on Partiality and Recursion in Interactive Theorem Provers*, volume 43 of *EPTCS*, pages 1–13, 2010.
51. Butler W. Lampson. A note on the confinement problem. *Communications of the ACM*, 16(10):613–615, October 1973.
52. K. Rustan M. Leino. Dafny: An automatic program verifier for functional correctness. In *LPAR-16*, volume 6355 of *LNCS*, pages 348–370. Springer, 2010.
53. Xavier Leroy. Formal certification of a compiler back-end or: programming a compiler with a proof assistant. In *Proc. 33rd ACM Symposium on Principles of Programming Languages*, pages 42–54. ACM, 2006.
54. Tim Lindholm, Frank Yellin, Gilad Bracha, and Alex Buckley. *The Java Virtual Machine Specification, Java SE 7 Edition*. Addison-Wesley, February 2013.
55. Farhad Mehta and Tobias Nipkow. Proving pointer programs in higher-order logic. *Information and Computation*, 199:200–227, 2005.
56. Robin Milner. A theory of type polymorphism in programming. *Journal of Computer and System Sciences (JCSS)*, 17(3):348–375, 1978.
57. Greg Morrisett, Gang Tan, Joseph Tassarotti, Jean-Baptiste Tristan, and Edward Gan. Rocksalt: better, faster, stronger SFI for the x86. In *Proceedings of the 33rd ACM SIGPLAN conference on Programming Language Design and Implementation, PLDI '12*, pages 395–404, New York, NY, USA, 2012. ACM.
58. Steven Muchnick. *Advanced Compiler Design and Implementation*. Morgan Kaufmann, 1997.
59. Olaf Müller, Tobias Nipkow, David von Oheimb, and Oskar Slotosch. HOLCF = HOL + LCF. *J. Functional Programming*, 9:191–223, 1999.
60. Toby Murray, Daniel Matichuk, Matthew Brassil, Peter Gammie, Timothy Bourke, Sean Seefried, Corey Lewis, Xin Gao, and Gerwin Klein. seL4: from general purpose to a proof of information flow enforcement. In *IEEE Symposium on Security and Privacy*, pages 415–429, 2013.
61. Flemming Nielson, Hanne Riis Nielson, and Chris Hankin. *Principles of Program Analysis*. Springer, 1999.
62. Hanne Riis Nielson and Flemming Nielson. *Semantics With Applications: A Formal Introduction*. Wiley, 1992.
63. Hanne Riis Nielson and Flemming Nielson. *Semantics with Applications. An Appetizer*. Springer, 2007.

64. Tobias Nipkow. *What's in Main*. <https://isabelle.in.tum.de/doc/main.pdf>.
65. Tobias Nipkow. Winskel is (almost) right: Towards a mechanized semantics textbook. In V. Chandru and V. Vinay, editors, *Foundations of Software Technology and Theoretical Computer Science*, volume 1180 of *LNCS*, pages 180–192. Springer, 1996.
66. Tobias Nipkow. Hoare logics for recursive procedures and unbounded nondeterminism. In J. Bradfield, editor, *Computer Science Logic (CSL 2002)*, volume 2471 of *LNCS*, pages 103–119. Springer, 2002.
67. Tobias Nipkow. Hoare logics in Isabelle/HOL. In H. Schwichtenberg and R. Steinbrüggen, editors, *Proof and System-Reliability*, pages 341–367. Kluwer, 2002.
68. Tobias Nipkow, Lawrence Paulson, and Markus Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, volume 2283 of *Lect. Notes in Comp. Sci.* Springer, 2002.
69. Tobias Nipkow and Leonor Prensa Nieto. Owicki/Gries in Isabelle/HOL. In J.-P. Finance, editor, *Fundamental Approaches to Software Engineering (FASE'99)*, volume 1577 of *LNCS*, pages 188–203. Springer, 1999.
70. G. D. Plotkin. A structural approach to operational semantics. Technical report, University of Aarhus, 1981.
71. Gordon D. Plotkin. The origins of structural operational semantics. *J. Log. Algebr. Program.*, 60-61:3–15, 2004.
72. Gordon D. Plotkin. A structural approach to operational semantics. *J. Log. Algebr. Program.*, 60-61:17–139, 2004.
73. Wolfgang Reif. The KIV system: Systematic construction of verified software. In Deepak Kapur, editor, *11th International Conference on Automated Deduction (CADE)*, volume 607 of *LNCS*, pages 753–757. Springer, June 1992.
74. John C. Reynolds. Separation logic: A logic for shared mutable data structures. In *Proc. 17th IEEE Symposium on Logic in Computer Science (LICS 2002)*, pages 55–74, 2002.
75. Alejandro Russo and Andrei Sabelfeld. Dynamic vs. static flow-sensitive security analysis. In *Proceedings of the 23rd IEEE Computer Security Foundations Symposium (CSF)*, pages 186–199. IEEE Computer Society, 2010.
76. Andrei Sabelfeld and Andrew Myers. Language-based information-flow security. *IEEE Journal on Selected Areas in Communications*, 21(1):5–19, 2003.
77. Andrei Sabelfeld and Alejandro Russo. From dynamic to static and back: Riding the roller coaster of information-flow control research. In A. Pnueli, I. Virbitskaite, and A. Voronkov, editors, *Perspectives of Systems Informatics, 7th International Andrei Ershov Memorial Conference (PSI)*, volume 5947 of *LNCS*, pages 352–365. Springer, 2009.
78. Norbert Schirmer. *Verification of Sequential Imperative Programs in Isabelle/HOL*. PhD thesis, Technische Universität München, 2006.
79. David Schmidt. *Denotational semantics: A methodology for language development*. Allyn and Bacon, 1986.
80. Thomas Schreiber. Auxiliary variables and recursive procedures. In *TAPSOFT'97: Theory and Practice of Software Development*, volume 1214 of *LNCS*, pages 697–711. Springer, 1997.

81. Edward Schwartz, Thanassis Avgerinos, and David Brumley. All you ever wanted to know about dynamic taint analysis and forward symbolic execution (but might have been afraid to ask). In *Proc. IEEE Symposium on Security and Privacy*, pages 317–331. IEEE Computer Society, 2010.
82. Dana Scott. Outline of a mathematical theory of computation. In *Information Sciences and Systems: Proc. 4th Annual Princeton Conference*, pages 169–176. Princeton University Press, 1970.
83. Dana Scott and Christopher Strachey. Toward a mathematical semantics for computer languages. Programming Research Group Technical Monograph PRG-6, Oxford University Computing Lab., 1971.
84. Thomas Sewell, Magnus Myreen, and Gerwin Klein. Translation validation for a verified OS kernel. In *PLDI*, pages 471–481, Seattle, Washington, USA, June 2013. ACM.
85. Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.*, 5:285–309, 1955.
86. Robert Tennent. Denotational semantics. In S. Abramsky, D. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3, pages 169–322. Oxford University Press, 1994.
87. Harvey Tuch. *Formal Memory Models for Verifying C Systems Code*. PhD thesis, School of Computer Science and Engineering, University of NSW, Sydney, Australia, August 2008.
88. Harvey Tuch, Gerwin Klein, and Michael Norrish. Types, bytes, and separation logic. In Martin Hofmann and Matthias Felleisen, editors, *Proceedings of the 34th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 97–108, Nice, France, January 2007. ACM.
89. D. Volpano, C. Irvine, and G. Smith. A sound type system for secure flow analysis. *Journal of computer security*, 4(2/3):167–188, 1996.
90. Dennis Volpano and Geoffrey Smith. Eliminating covert flows with minimum typings. In *Proceedings of the 10th IEEE workshop on Computer Security Foundations, CSFW '97*, pages 156–169. IEEE Computer Society, 1997.
91. Dennis M. Volpano and Geoffrey Smith. A type-based approach to program security. In *Proc. 7th Int. Joint Conference CAAP/FASE on Theory and Practice of Software Development (TAPSOFT '97)*, volume 1214 of *LNCS*, pages 607–621. Springer, 1997.
92. Makarius Wenzel. *The Isabelle/Isar Reference Manual*. <https://isabelle.in.tum.de/doc/isar-ref.pdf>.
93. Glynn Winskel. *The Formal Semantics of Programming Languages*. MIT Press, 1993.

Index

$\llbracket \cdot \rrbracket$	283	$\langle x := i \rangle$	29
\implies	6, 283	$\{\}$	38
\wedge	6, 283	$f \vdash A$	39
\equiv	283	$\{x. P\}$	38
λ	283	$\{t \mid x. P\}$	38
\Rightarrow	283	\dots	60
\wedge	37, 283	$c_1;; c_2$	76
\vee	37, 283	$x ::= a$	76
\longrightarrow	37, 283	$(c, s) \Rightarrow t$	78
\rightarrow	283	$c \sim c'$	82
\neg	37, 283	$(c, s) \rightarrow (c', s')$	86
\neq	283	$(c, s) \rightarrow^* (c', s')$	87
\forall	37, 283	$P \vdash c \rightarrow c'$	97
\exists	37, 283	$P \vdash c \rightarrow^* c'$	98
\leq	173, 283	$xs !! i$	96
\times	283	$P \vdash c \rightarrow \sim^n c'$	106
\in	38, 283	$\Gamma \vdash a : \tau$	121
\notin	283	$\Gamma \vdash b$	121
\subseteq	38, 283	$\Gamma \vdash c$	122
\subset	283	$\Gamma \vdash s$	124
\cup	38, 283	$l \vdash c$	131, 137
\cap	38, 283	$l \vdash' c$	134, 139
\bigcup	38, 283	$s = s' (< l)$	130
\bigcap	38, 283	$s = s' (\leq l)$	130
\sqcup	237, 283	$\vdash c : l$	135
\sqcap	260, 283	$f = g \text{ on } X$	168
\bigsqcup	231, 283	$r \bigcirc s$	180
\bigsqcap	231, 283	$\{P\} c \{Q\}$	191
\top	231, 237, 283	$P[a/x]$	194
\perp	231, 260, 283	$s[a/x]$	200
$f(a := b)$	29	$\models \{P\} c \{Q\}$	199
$\langle \rangle$	29	$\vdash \{P\} c \{Q\}$	200

$\models_t \{P\} c \{Q\}$ 212 $\vdash_t \{P\} c \{Q\}$ 212 γ_{fun} 241 γ_{option} 241 γ 237 γ_c 243 γ_o 243 γ_s 243 $[l, h]$ 265 ∇ 270 \triangle 274

abbreviation 17

abstract domain 236

abstract state 239

abstract syntax 28

abstract syntax tree 28

abstract value 236

acomp 99*ADD* 96*aexp* 28*AI* 246*And* 32*anno* 225*annos* 225, 282*annotate* 225, 282

annotated command 208, 220

anti-monotonicity 132, 249

antisymmetric 173

arbitrary: 20, 68*arith* 41*asem* 242

assertion 193

Assign 76, 78, 200*Assign'* 201*assms* 56*assn* 199

assume 53

assumes 55

assumption 50*auto* 39

available definitions 172

aval 29*aval''* 262*aval'* 244

axiomatic semantics 191

backchaining 43

backward analysis 178

Bc 32*bcomp* 100*bexp* 32

big-step semantics 77

blast 40*bool* 7*bot* 246

bottom element 260

bsem 242*bury* 169*bval* 33

by 53

card 39

case 61–68

case analysis 56, 61–62

case expression 16

case ... of 16*?case* 63–68*cases* 57, 61–68*ccomp* 102

chain 184

chain-complete partial order 189

collecting semantics 221

com 76*com_den* 180

command 75

equivalence 82

comment 11

complete induction 107

computation induction 18

concrete syntax 27

concretisation function 237

config 97

configuration 85

confinement 133

conform 123

congruence 83

Cons 14*conseq* 200

consequence rule 196

constant folding 29, 146

continuity 184

continuous 184, 189

covert channel 128, 136

- cpo 189
- CS 233
- D* 147, 180
- datatype 15
- dead variable 165
- definite initialization 145
- definition 17
- denotational semantics 179
- derivation tree 79
- dest*: 45
- deterministic 84
- equality 6
- equivalence relation 83
- exec1* 97
- exits* 106
- extended integers 265
- extensionality 205
- fact 54
- False* 7
- false alarm 220
- fastforce* 40
- final* 89
- finite* 39
- fix 53, 57
- fixes 55
- fixpoint 173
- for 60
- formula 6
- forward analysis 178
- from 53
- fun* 253
- fun* 17
- generate and kill analysis 167
- greatest lower bound 231
- have 53, 60
- hd* 14
- head 14
- height of ordering 249
- hence 55
- Hoare logic 191
 - completeness 206
 - completeness (total) 214
 - incompleteness 206
- proof system 194
- soundness 204
- soundness (total) 213
- Hoare triple 191
 - derivable 194
 - valid 199
 - valid (total) 212
- .hyps* 66
- Ic* 117
- Id* 180
- ieexec* 97
- IF* 76
- If* 76, 200
- if 60
- IfFalse* 78
- IfTrue* 78
- IH 9
- .IH* 64, 66
- IMP 75
- imports 6
- induction 16, 62–68
- induction heuristics 19
- .induct* 18
- induction ... rule*: 19, 48, 65, 68
- inductive definition 45–51
- infimum 231, 260
- information flow control 128
- inner syntax 7
- instr* 96
- int* 5
- interval analysis 264
- intro* 44
- introduction rule 43
- .intros* 48
- inv_aval'* 262
- inv_bval'* 263
- inv_less'* 263
- inv_plus'* 261
- invariant 195
- is 58
- Isar 53
- isuccs* 105
- Ity* 121
- Iv* 117
- ivl* 265

- jEdit 4
- JMP 96
- JMPGE 96
- JMPLESS 96
- join 237
- judgement 78
- Kleene fixpoint theorem 185
- Knaster-Tarski fixpoint theorem 174, 231, 233
- L* 165, 174
- language-based security 128
- lattice 260
 - bounded 260
 - complete 231, 233
- least element 173
- least upper bound 231
- lemma 9
- lemma 55
- length 14
- Less 32
- let 59
- level 129
- lfp 174
- linear arithmetic 41
- list 10
- live variable 165
- LOAD 96
- LOADI 96
- locale 242
- lwars 281
- Main 7
- map 14
- map_acom 225, 282
- may analysis 178
- meet 260
- metis 41
- mono 174
- monotone 173
- monotone framework 247
- moreover 60
- must analysis 178
- N* 28
- narrowing operator 274
- nat* 8
- natural deduction 44
- Nil* 14
- non-deterministic 84
- None* 16
- noninterference 129
- Not* 32
- num'* 243
- obtain 58
- OF* 44
- of* 42
- operational semantics 77
- option* 16
- outer syntax 7
- parent theory 7
- partial correctness 193
- partial order 173
- pfp* 246
- Plus* 28
- plus'* 243
- point free 188
- polymorphic 10
- post* 225
- postcondition 192
 - strongest 207
- pre* 208
- pre-fixpoint 173
- precondition 192
 - weakest 204
 - weakest liberal 204
- .prems* 64, 66
- preservation 123
- program counter 96
 - exits 106
 - successors 105
- progress 123
- proof 53
- qed 53
- quantifier 6
- quotient type 257
- Rc* 117
- reflexive 83, 173
- rewrite rule 22

- rewriting 22
- Rty* 121
- rule 9
- rule* 43, 44
- rule application 43
- rule induction 46–50, 64–68
- rule inversion 66–68
- Rv* 117
- rvars* 281
-
- sec* 129, 130
- semilattice 237
- separation logic 216
- Seq* 76, 78, 200
- set* (constant) 39
- set* (type) 38
- set comprehension 38
- show* 53
- shows* 55
- side condition 51
- simp* 23
- simplification 21
- simplification rule 21
- simplifier 21
- simulation 103
- single-valued 187
- size* 96
- SKIP* 76
- Skip* 78, 200
- Sledgehammer 41
- small-step semantics 85
- Some* 16
- split*: 25
- .split* 25
- st* 239, 257
- st_rep* 256
- stack* 97
- stack underflow 36, 52
- state 28
- state* 28
- Step* 242
- step* 226, 242
- step'* 244
- STORE* 96
- string* 15
- strip* 208, 225, 281
- structural induction 11, 16, 62–64
- structurally equal 232
- substitution lemma 32, 200
- subsumption rule 132
- Suc* 8
- succs* 105
- supremum 231, 237
- symmetric 83
- syntax-directed 122
-
- tail 14
- taint analysis 141
- taval* 118
- tbval* 119
- term 5
- test_num'* 261
- then* 55
- theorem 9
- theory 6
- theory* 6
- theory file 7
- ?thesis* 59
- this* 55
- thus* 55
- tl* 14
- top element 237
- total correctness 193
- transitive 83, 173
- True* 7
- ty* 121
- tyenv* 121
- type annotation 6
- type class 238
 - instantiation 238
- type constraint 6
- type derivation 122
- type inference 6
- type safe 116
- type soundness 116
-
- ultimately 60
- unification 43
- unknown 9, 42
- update* 253
- using* 55
-
- V* 28
- val* 28

- valid Hoare triple 192
- value 10
- values 80
- vars* 147, 281
- vc* 209
- VCG 208
- verification condition 208
- verification condition generator 208
- vname* 28
- where* 43
- WHILE* 76
- While* 76, 200
- while* 176
- While'* 201
- while_option* 246
- WhileFalse* 78
- WhileTrue* 78
- widening operator 270
- with 55
- wp* 204
- wp_t* 213