Toward Reliable Modular Programs

Thesis by
K. Rustan M. Leino

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1995
(Submitted 5 January 1995)
Abstract

Software is being applied in an ever-increasing number of areas. Computer programs and systems are becoming more complex and consisting of more delicately interconnected components. Errors surfacing in programs are still a conspicuous and costly problem. It’s about time we employ some techniques that guide us toward higher reliability of practical programs. The goal of this thesis is just that.

This thesis presents a theory for verifying programs based on Dijkstra’s weakest-precondition calculus. A variety of program paradigms used in practice, such as exceptions, procedures, object orientation, and modularity, are dealt with.

The thesis sheds new light on the theory behind programs with exceptions. It develops an elegant algebra, and shows it to be the foundation on which the semantics of exceptions rests. It develops a trace semantics for programs with exceptions, from which the weakest-precondition semantics is derived. It also proves a theorem on programming methodology relating to exceptions, and applies this theorem in the novel derivation of a simple program.

The thesis presents a simple model for object-oriented data types, in which concerns have been separated, resulting in the simplicity of the model.

To deal with large programs, this thesis takes a practical look at modularity and abstraction. It reveals a problem that arises in writing specifications for modular programs where previous techniques fail. The thesis introduces a new specification construct that solves that problem, and gives a formal proof of soundness for modular verification using that construct. The model is a generalization of Hoare’s classical data refinement. However, there are more problems to be solved. The thesis reports on some of these problems and suggests some future directions toward more reliable modular programs.
Preface

I’m Rustan Leino. I’ll be your host through this thesis. During the last three and a half years, I’ve been advised, in different capacities, by three people. Jan van de Snepscheut led me through my Master’s thesis [49]. His taste and skill for elegant solutions to practical problems put me on the path that has led to the present thesis. I hope we will be able to keep up his enthusiasm. Mani Chandy with his slogan of “Bringing theory to the marketplace” guided me from there. His apt ability to authoritatively sieve the important from the unimportant, resulting in a focus on the road and not on the curbside distractions, was instrumental in making the journey expeditious. At Digital’s Systems Research Center (DEC SRC), Greg Nelson, a champ in applying his (and others’) theory in practice, directed me to problems yearning for my solutions. His ample support and fruitful collaboration have had a great impact on the destination of this thesis. I’m indebted to each of these individuals for his inspiration and guidance.

As a Microsoftee —before becoming a graduate student and starting a research career—, I learned the craft of programming in the large, and became more aware of the challenges involved in producing correct software and the discipline that requires. I also learned the importance of the run-time check as a device for more quickly detecting errors in a program. This thesis applies to verification in general, but has been motivated by the hope of proving statically that no run-time check will fail during run-time, an area known as extended static checking.

In addition to the people mentioned above, numerous other people at Caltech and DEC SRC have been of great help during discussions and in reviewing my thesis. Of these colleagues, I mention Dave Detlefs, Robert Harley, Allan Heydon, Peter Hofstee, Rajit Manohar, Berna Massingill, Adam Rifkin, Paul Sivilotti, John Thornley, and, of course, the members of my thesis committee: Mani Chandy, Alain Martin, Greg Nelson, Beverly Sanders, and Rick Wilson. I am grateful to them all for their stimulus and loyal support. I am particularly grateful to Rajit and Paul for their various contributions to proofs, Berna for her careful proofreading of my thesis, and Dave for his putting these formulas to work in the SRC Extended Static Checker. I’d also like to express much appreciation to DEC SRC for their financial support.

These pages have been prepared in \LaTeX, using many useful macros by Marcel van der Goot and Rajit Manohar. The fancy font used in various headers is a public domain POSTSCRIPT font called Civitype, developed by S.G. Moyer. All figures were made from POSTSCRIPT programs I wrote. I have written the thesis in a personal
style so that I may get to know my readers better.

Last, but not least, I want to thank Indi for her never-ending loving support. You’ve been of tremendous help in every way during this time.

Tjoflöjt, K.R.M.L.

January 1995
Pasadena, CA, U.S.A.

This thesis is available as Technical Report Caltech-CS-TR-95-03.
Contents

0 Introduction

0.0 Motivation .................................................. 0
0.1 Contents ..................................................... 2
0.2 Preliminaries ................................................ 4

I Control Structures ............................................ 10

Imperative programming languages .......................... 12
Outline ........................................................... 12

1 Semantics of programs with exceptions .................. 14
1.0 Weakest precondition ...................................... 14
1.1 Assignment ................................................ 16
1.2 Unit statements and compositions ....................... 16
1.3 Block ....................................................... 17
1.4 Partial commands .......................................... 18
1.5 Choice compositions ..................................... 20
1.6 Iteration .................................................. 21
1.7 Specification statement .................................. 22
1.8 Refinement ................................................ 24

2 Functions of two arguments and their compositions .... 26
2.0 Function compositions ................................... 26
2.1 Left and right composition ............................... 27
2.2 Double composition ...................................... 28
2.3 Ceiling and floor ......................................... 30
2.4 Transposition ............................................. 32
2.5 The connection with programs ........................... 33
2.6 Concluding remarks ..................................... 34

3 Trace semantics for exceptions ............................. 36
3.0 Introduction ................................................. 36
3.1 Trace sets .................................................. 37
3.2 Program constructs as trace sets ....................... 38
3.3 Weakest preconditions of trace sets ........................................ 43
3.4 Calculating the weakest preconditions ................................. 45

4 A theorem on programming methodology ........................................ 48
4.0 Hoare triples .............................................................. 48
4.1 Free occurrences of raise .............................................. 49
4.2 Usage of exceptions ..................................................... 51

5 Constructing a program with exceptions ................................. 52
5.0 A program derivation .................................................... 52
5.1 Discussion ................................................................. 53

6 Modeling common programming languages .................................. 56
6.0 Procedures ................................................................. 56
6.1 Alternative statements ................................................. 59
6.2 Statements that “go wrong” ............................................ 60
6.3 Expressions ................................................................. 61

II Data Structures ............................................................... 64
Data structures ................................................................. 66
Outline ................................................................. 66

7 Data types ........................................................................ 68
7.0 Types ....................................................................... 68
7.1 Types in common programming languages .............................. 70
7.2 Declaring new types .................................................... 73
7.3 Maps and specifications ................................................ 75

8 Objects ........................................................................ 78
8.0 Subtypes .................................................................. 78
8.1 Data fields ................................................................. 79
8.2 Methods ................................................................. 79
8.3 Method implementations .............................................. 79
8.4 Object simplicity ........................................................ 80
8.5 Language implementations of objects ................................ 81
8.6 Objects in common programming languages ...................... 81

9 Abstraction .................................................................. 84
9.0 Abstract variables ......................................................... 84
9.1 Abstract variables and refinement .................................... 86
9.2 Abstract data fields .................................................... 89
### III Modularity

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modules and modular verification</td>
<td>94</td>
</tr>
<tr>
<td>Outline</td>
<td>94</td>
</tr>
</tbody>
</table>

### 10 Specifications in modular programs

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0 Motivation</td>
<td>96</td>
</tr>
<tr>
<td>10.1 Problem</td>
<td>97</td>
</tr>
<tr>
<td>10.2 Solution</td>
<td>101</td>
</tr>
<tr>
<td>10.3 Enforcing the requirements</td>
<td>104</td>
</tr>
<tr>
<td>10.4 Soundness of modular verification</td>
<td>105</td>
</tr>
<tr>
<td>10.5 A generalization of classical data refinement</td>
<td>106</td>
</tr>
<tr>
<td>10.6 Other specification languages</td>
<td>108</td>
</tr>
</tbody>
</table>

### 11 Generating verification conditions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.0 A notation for modular programs</td>
<td>110</td>
</tr>
<tr>
<td>11.1 Definitions</td>
<td>112</td>
</tr>
<tr>
<td>11.2 Proving refinements</td>
<td>122</td>
</tr>
<tr>
<td>11.3 The importance of residues</td>
<td>124</td>
</tr>
</tbody>
</table>

### 12 Soundness of modular verification

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.0 Requirements</td>
<td>128</td>
</tr>
<tr>
<td>12.1 Soundness</td>
<td>128</td>
</tr>
<tr>
<td>12.2 Proof outline</td>
<td>129</td>
</tr>
<tr>
<td>12.3 All side effects are benevolent</td>
<td>132</td>
</tr>
<tr>
<td>12.4 X</td>
<td>135</td>
</tr>
<tr>
<td>12.5 Epilogue</td>
<td>141</td>
</tr>
</tbody>
</table>

### 13 depends in perspective

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.0 Specification of a consumer</td>
<td>144</td>
</tr>
<tr>
<td>13.1 Shortcomings of depends</td>
<td>146</td>
</tr>
<tr>
<td>13.2 Private values</td>
<td>148</td>
</tr>
</tbody>
</table>

### IV Epilogue

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 Summary</td>
<td>154</td>
</tr>
</tbody>
</table>
Chapter 0

Introduction

In this chapter, I present an introduction to, and the motivation for, this thesis. I also provide an outline of the contents of the thesis, show the dependencies between the chapters, and discuss some preliminaries such as notation.

0.0 Motivation

Today, computer programs are being written for an ever-growing number of purposes. Unfortunately, not all of this software is correct. Programmers introduce errors into programs for a variety of reasons. The errors may be the result of, e.g., typos, logical mistakes, incorrect assumptions, vague or changing specifications, or lack of specifications. Whatever the cause, the effect of software errors can be very costly, e.g., a malfunctioning computerized radiation therapy machine has claimed the lives of humans [52], a broken telephone switch has resulted in loss of service [44], problems with an automated baggage-handling system have delayed the opening of an airport [27], an erroneous word processor can cause the loss of important information, and an error-prone program can degrade the reputation of a software company. Even the errors that are found prior to the shipping of or use of a software product are costly, primarily in terms of man-hours spent finding and correcting errors, and in terms of delayed time to market, resulting in loss of market share and revenues.

To reduce the number of errors in a program, or to increase one’s confidence in a program, one can test the program on a given test suite. If the program is observed to behave correctly for these test cases, the program is shipped to the customer. One then hopes there will be other cases that customers try for which the program also behaves correctly.

Another way to reduce the number of errors in and increase one’s confidence in a program is to write precise specifications and mathematically verify that the program meets those specifications.

Remark 0.0. Notice that I said “reduce the number of errors” and “increase one’s confidence”, as opposed to “eliminate all errors”, because the specifications, too, can be written incorrectly, or the program could
be used incorrectly by the customer. However, specifications are generally
concerned with fewer details than programs, and may thus be easier to
get right. Moreover, the probability of making the same error in both
specification and program is much lower than just making an error in the
program text itself.

Mathematically proving a program correct shows the program will work not just for
one test suite, but for all permitted uses of the program. This technique also has
the advantage over testing in that it can be applied prior to the completion of an
implementation; thus, errors can be found earlier, which reduces costs.

To sustain mathematical proofs of program correctness, programs must be given
a mathematical meaning, a semantics. An example thereof is Dijkstra’s weakest-
precondition semantics [17], which has achieved considerable success in modeling pro-
grams mathematically, because of its high level of expressivity (using predicates over
the state space) and its simplicity.

Weakest preconditions have been around for almost two decades, and other tech-
niques for reasoning about programs (e.g., Hoare triples [36]) for over two decades.
So why is it that not every programmer uses these techniques every day? Part of the
answer is that programmers actually do, implicitly—the study of semantics has had
a profound influence on the design of programming languages that programmers use,
and programmers may have firm their understanding of programs through learning
about semantics as undergraduates.

A major reason these techniques have not penetrated the everyday life of pro-
grammers more visibly is the challenges that the proving of large programs poses.
For example, a large program gives rise to large formulas to be proven. Proving all
of these by hand would be a virtually impossible task, especially for large programs
that undergo change—one small change in the program may necessitate reproofing
the entire program. Instead, we may consider receiving assistance from automatic
theorem provers (see, e.g., [69, 74]). This is an area that still needs more work.

Another task that presents challenges is the mathematical modeling of programs.
Although their design is influenced by semantics, common programming languages
provide some features whose mathematical meanings are difficult to capture concisely,
e.g., arbitrary pointers.

Yet another area that presents challenges is the writing of precise specifications!
If we don’t know how to write specifications, how can we expect to be able to verify
that a program behaves correctly?

A step toward the (full) verification of programs is an area called extended static
checking. The idea is to prove programs correct, but only with respect to certain pro-
erties. In particular, a program is given enough specifications to prove the absence of
checked run-time errors, like nil-dereferencing, array index out-of-bounds errors,
and failing assertions. This is likely to lighten the burden in the areas of (automatic)
theorem proving and specification writing. For example, rather than needing a means
to express and reason about “permutation of” in the full specification and verification
of a sorting routine, the specification would only need to be strong enough to show
that the implementation never indexes the given array outside its bounds. Despite the fact that the verification of the latter does not guarantee the array to be sorted upon termination of the sorting routine, such a verification removes the possibility of a program execution resulting in a checked run-time error and greatly increases our confidence in the program.

**Remark 0.1.** Drawing from my personal experience with developing programs at Microsoft, and also since then (e.g., the implementation in [49]), I have found that almost all errors ever detected were detected as results of failing run-time checks. Proving the absence of checked run-time errors would thus have shown the absence of most errors detected in these programs.

The technique also offers the benefit of detecting errors earlier. As a bonus, run-time checks can be removed from the executable code, making the executable smaller and faster [26] (see also [78]).

Extended static checking gets its name from the idea that it is to be incorporated into a compiler, much like static type checkers in today’s compilers. This would indeed bring the direct benefit of the science of program correctness to programmers, an event one may expect to do marvels for the correctness of their programs.

### 0.1 Contents

This thesis concerns the specification and verification of modular programs that feature exceptions and objects. "Modular" means that programs are divided into pieces that can be compiled separately. Moreover, through the application of modular verification, the modules of such programs can be verified separately. This is important, because it is modularity and modular verification that allow us to tackle large programs. An exception is a form of a structured jump and is sometimes convenient in programming. A small example thereof is given in Chapter 5, but exceptions are more frequently used in large programs. Objects are a means of organizing data in a program and facilitate the sharing of code. Object-oriented programming is gaining popularity. Its utility is most visible when programming in the large.

The main contributions of this thesis are several aspects of reasoning about programs with exceptions, a separation of concerns in the meaning of objects, and the invention of a new specification construct for the specification and verification of modular programs. Of these three main contributions, the last mentioned, which in some sense is an extension of classical data refinement [38], is bound to have the greatest impact on making the specification and verification of large programs feasible in practice.

Although they work independently of each other, these three contributions, collectively, are steps toward making modular programs more reliable.
0.1.0 Outline

Compiled from sources like [17, 70, 67, 51], I kick off Part I, Control Structures, by presenting a mathematical semantics, based on weakest preconditions, of sequential imperative programming constructs for languages with exceptions (Chapter 1). Then, for the duration of a few chapters, I devote my attention to the control flow of programs with exceptions (so-called exceptional programs, no pun intended). I show that the weakest preconditions of the basic exceptional statements and compositions have their foundation in a beautiful algebra over functions of two arguments (Chapter 2). I show an operational semantics of programs with exceptions and its connection with the weakest precondition semantics (Chapter 3). I also prove a theorem that suggests a use for exceptions (Chapter 4), and from it show a heuristic for program construction applied to the derivation of a simple program (Chapter 5). With an eye to “real” programming languages and their common usage, I present procedures, and show how run-time errors, partial expressions, short-circuit boolean operators, and expressions with side effects are modeled mathematically (Chapter 6).

In Part II, Data Structures, I turn my attention to the data structures of programs and show how common data structures are modeled mathematically (Chapter 7). The most interesting of these concerns objects (Chapter 8), for which I propose a simple mathematical treatment that separates the concerns involved. As a warm-up to Part III, I also present data abstraction in its classical form [38] (Chapter 9).

Part III, Modularity, concerns modular specification and verification of programs. I demonstrate our lack of understanding of writing specifications in modular programs (with an emphasis on object-oriented programs), and contribute a new specification construct, depends, as a necessary aid in writing such specifications (Chapter 10). I define a precise interpretation of modular programs and their specifications (Chapter 11), and prove that this interpretation lends itself to sound modular verification of programs (Chapter 12). Finally, I show some remaining problems in the area of writing specifications of modular programs (Chapter 13).

Chapter 14 summarizes the thesis and offers some concluding remarks.

Figure 0.0 shows the dependencies between the chapters. I recommend reading the chapters in any order suggested by the partial order in that figure.
0.1.1 Originality and Collaboration

Roughly speaking, it is Chapters 2–5, 8, and 10–12 that contain the new work presented in this thesis. The other chapters present pertinent material from the literature, composed and presented in such a way as to set the stage for the chapters containing new material.

Most of Chapters 1 and 6 is a composition of well-known work in semantics (cf. [17, 51, 70, 67, 2, 40, 29, 33]). Chapters 2–4 contain joint work with Jan L.A. van de Snepscheut, and are published in [51]. Although the program in Chapter 5 occurs in [51], the derivation of this program and the heuristic used in that process appear first in [50]. Chapter 7 is based on [40, 37, 17, 58], and the first part of Chapter 8 is from [15]. The portion of Chapter 8 that concerns object simplicity, however, resulted from work I did with Greg Nelson at Digital’s Systems Research Center. The ideas in Chapter 9 are mostly from [38]. Chapters 10–12 stem from joint work with Greg Nelson. Chapter 13 presents my assessment of the consequences of the work in Chapters 10–12, and some future directions for that work.

0.2 Preliminaries

In this section, I explain the notation I use in this thesis. I take most of the notation and proof format from [21]. I assume familiarity with the predicate calculus (see, e.g., [21, 31, 19]).

Functions

Function application is written with an infix dot. For example, a function \( f \) applied to a value \( a \) is denoted \( f.a \). Function application is left associative. Thus, \( \text{wp.5.Q} \) means \( \text{wp.5}.Q \).

Substitution

For a list of identifiers \( x \) and an (equally long) list of expressions \( y \), I use \( [x := y] \) as the postfix, left-associative substitution function. The expression

\[
Q[x := y]
\]

denotes \( Q \) in which all free occurrences of (each identifier in) \( x \) are replaced by (the corresponding expression in) \( y \).

Operators and binding powers

Function application (\( . \)) and substitution (\( [:=] \)) have the highest binding power. Then comes negation (\( \neg \)), followed by the arithmetic and set operators whose relative binding powers are the usual ones. Next are relations like \( = \) (when used in predicates and arithmetic expressions), \( \leq \), and \( < \).
The remaining boolean connectives are given lower binding power. Among them, \( \land \) and \( \lor \) bind the strongest, then \( \Rightarrow \) and \( \Leftarrow \), and last \( \equiv \).

The operators of statement compositions bind weaker than the boolean connectives. Of these, \( ; \) and \( < \) have the strongest binding power, then \( \rightarrow \), and last \( \sqcap \) and \( \emptyset \). Refinement \( (\sqsubseteq) \) is given lower binding power than all of these.

In the equality among programs or among predicates, \( = \) binds weaker than any other operator.

**Predicates**

A *predicate* is a boolean function over some state space. Conjunction \( (\land, \text{“and”}) \) and disjunction \( (\lor, \text{“or”}) \) are examples of operators on predicates. The universal quantification, \( \forall \) (conjunction), of an expression \( T \) over the values of a list of variables \( x \) constrained by a predicate \( R \), is written

\[
\{ \forall x \mid R \Rightarrow T \}
\]

The expression can be read as "for all \( x \) such that \( R \) holds, \( T \) holds" or "the conjunction over \( x \) such that \( R \) of \( T \)." For example,

\[
\{ \forall \theta \mid 0 \leq \theta \leq \pi \Rightarrow 0 \leq \sin \theta \}
\]

expresses that for all values of \( \theta \) satisfying \( 0 \leq \theta \leq \pi \), the sine of \( \theta \) is at least 0.

In the general expression, \( x \) is called the *dummy*, \( R \) the *range* of the dummy, and \( T \) the *term* of the quantified expression.

Existential quantification, \( \exists \) (disjunction), and union quantification, \( \cup \), are written similarly.

If the range is *true* or if it is understood from the context, then it is often omitted for brevity. The quantified expression is then written

\[
\{ \forall x \Rightarrow T \}
\]

and similarly for the other quantifiers. Rules for manipulating quantifiers can be found in, e.g., [21, 31].

A function from predicates to predicates is called a *predicate transformer*.

**Everywhere and lifting**

For any predicate \( P \), \([P]\) (pronounced "\( P \) everywhere") denotes that \( P \) holds in every state. The brackets \( [\cdot] \) are called *everywhere* brackets. Thus, for a predicate on the state space whose variables are denoted by \( z \), \([P]\) is a shorthand for

\[
\{ \forall z \Rightarrow P.z \}
\]

In general, \([P]\) is preferred to \( \{ \forall z \Rightarrow P.z \} \), because it allows expressions to omit irrelevant details. Similarly, I often write

\[
(P \Rightarrow Q).z
\]
instead of

$$P.z \Rightarrow Q.z$$

and similarly for operators other than \( \Rightarrow \). This process is called *lifting*. A special case is \( P \equiv Q \), which denotes the predicate that is *true* in exactly those states where \( P \) and \( Q \) yield the same value. Consequently, \( [P \equiv Q] \) means that \( P \) and \( Q \) are equal as predicates, a fact that is frequently expressed as \( P = Q \).

A predicate whose value, as a function, equals the predicate *true* or the predicate *false* is called a *boolean scalar*. For example, \([P]\) is a boolean scalar.

**Junctivity, distributivity, and monotonicity**

A function \( f \) is *conjunctive* if it distributes over conjunction, that is,

$$f.\{ \forall X \mid X \in B \triangleright X \} = \{ \forall X \mid X \in B \triangleright f.X \} \quad (0.0)$$

A description of the bags (multisets) \( B \) of predicates for which this property holds is usually used as a prefix of "conjunctive". For example, \( f \) is said to be *universally conjunctive* if \( (0.0) \) holds for any \( B \), *positively conjunctive* if \( (0.0) \) holds for nonempty bags \( B \) ("positively" refers to the positive cardinality of \( B \)), and so on.

Similarly, a function is *disjunctive* if it distributes over disjunction.

As junctivity is simply distributivity, "universally distributive over conjunction" or "universally conjunction-distributive" are synonymous to "universally conjunctive". This notion then also extends to, for example, a function being positively *union-distributive*, meaning that it distributes over any nonempty bag of sets.

A function is positively, finitely, and linearly conjunctive (disjunctive), *i.e.*, it distributes over any nonempty finite conjunction (disjunction, respectively) of predicates totally ordered by \([ \Rightarrow \]) exactly when it is *monotonic* [21]. Thus, conjunctivity (or disjunctivity) implies monotonicity. Monotonicity of a predicate transformer \( f \) is often written

$$\{ \forall X, Y \triangleright [X \Rightarrow Y] \Rightarrow [f.X \Rightarrow f.Y] \} \quad .$$

In proof hints (see below), I often abbreviate "distribution of \( \bigvee \) over \( \forall \)" by " \( \bigvee \) over \( \forall \)", and similarly for other operators.

**Sets**

I use standard mathematical notation for sets and pairs. Operator \( \setminus \) denotes asymmetric set difference, defined for any sets \( A \) and \( B \) as

$$\{ \forall x \triangleright x \in A \setminus B \equiv x \in A \land x \notin B \} \quad .$$

Deviating from standard mathematical notation, I write a quantified set constructor in a notation similar to that of other quantifiers. For example,

$$\{ n \mid 0 \leq n < N \triangleright n^2 \}$$
is the set of the first $N$ squares.

The definition of the set constructor can be written

$$\langle \forall y \triangleright y \in \{ x \mid R \triangleright T \} \equiv \langle \exists x \mid R \triangleright y = T \rangle \rangle .$$

Alternatively, the definition can be written

$$\{ x \mid R \triangleright T \} = \langle \cup x \mid R \triangleright \{T\} \rangle ,$$

where the latter is a quantification over set union. Quantified set constructors can be used, for example, in expressing properties like

$$A \setminus B = \{ x \mid x \in A \land x \notin B \triangleright x \} .$$

As for other quantified expressions, if the range is true or is understood, I omit the range and simply write

$$\{ x \triangleright T \} .$$

**Proof format**

I use an explicit proof format proposed by W.H.J. Feijen. It gives a hint for each step. Let me give an example. For any positive integer $x$, let $p.x$ denote the (unique) bag of prime factors whose product equals $x$. Then, the calculation, for all $x$,

\[ \text{even.}(x^2) \]

\[ = \{ \text{even.y} \equiv 2 \in p.y, \text{with } y := x^2 \} \]

\[ 2 \in p.(x^2) \]

\[ = \{ p.(y \cdot z) = p.y \cup p.z, \text{with } y, z := x, x \} \]

\[ 2 \in p.x \cup p.x \]

\[ = \{ y \in X \cup Y \equiv y \in X \lor y \in Y, \text{with } y, X, Y := 2, p.x, p.x \} \]

\[ 2 \in p.x \lor 2 \in p.x \]

\[ = \{ \lor \text{ is idempotent} \} \]

\[ 2 \in p.x \]

\[ = \{ \text{even.y} \equiv 2 \in p.y, \text{with } y := x \} \]

\[ \text{even.x} \]

demonstrates

$$\langle \forall x \triangleright \text{even.}(x^2) \equiv \text{even.x} \rangle .$$

For predicates, a calculation like

$$A$$

$$= \{ \text{hint why } [A \equiv B] \}$$

$$B$$

$$\Rightarrow \{ \text{hint why } [B \Rightarrow C] \}$$

$$C$$

$$= \{ \text{hint why } [C \equiv D] \}$$

$$D$$

shows that $[A \Rightarrow D]$. 
Programming languages

I often make references to common programming languages. I refer more often to Modula-3 [71] than to other languages. However, principles and techniques generally extend to other languages like Ada [1], Modula-2 [81], Pascal [42], C [45], or C++ [23] as well.

Nomenclature

I do not make any distinction between the terms statement, command, and guarded command. Each refers to a component of a program. I use these terms interchangeably.

A client of an interface is another module or interface that makes use of the first interface.

Verification process refers to the process of verifying a program. The process may be conducted by a human or by a machine.
Part I

Control Structures
Imperative programming languages

My goal is to reason about the correctness of programs. I focus our attention on sequential imperative languages, where I have in mind some language like (the sequential subset of) Modula-3, Ada, Modula-2, Pascal, or maybe even a disciplined subset of C or C++. These languages provide a variety of program constructs. Since many of these constructs are but variations of, or shorthands for, other constructs, we find that we can write most of the constructs in Dijkstra's guarded command language [17], with some extensions [51, 70, 67].

The advantage of using guarded commands as the programming notation is that we have a simple, precise, and concise mathematical meaning, or semantics, for such programs. The idea is to map each program to a predicate transformer [17]. A predicate is a boolean function on the state space of a program, and a predicate transformer is a function from predicates to predicates. Predicate transformers, as do predicates, draw their mathematical properties from complete boolean lattices [82, 7, 76]. The particular mappings from programs to predicate transformers that I use are called weakest precondition and weakest liberal precondition [17].

Outline

The structure of Part I is as follows. Chapter 1 introduces the program constructs and their weakest-precondition semantics.

Chapters 2 through 5 are concerned with exceptions—a form of structured jumps. Motivated by the weakest preconditions of exceptional statements, I show a nice algebra over functions of two arguments in Chapter 2. In Chapter 3, I justify the particular weakest preconditions given to the basic exceptional statements and compositions. I do that by showing a more concrete semantics (viz., a trace semantics) for exceptions, from which I derive the weakest preconditions. In Chapter 4, I deal with the use of exceptions, and develop a theorem that suggests a method for using exceptions in program development. An example application of that theorem is presented in Chapter 5, where I show a novel derivation of a simple program.

Chapters 2 through 5 are quite independent of each other, and also of the subsequent material in this thesis. Thus, the reader can study those chapters according to interest. To guide in that selection, the following table associates these chapters with interest areas.

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Algebra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 3</td>
<td>Semantics</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Programming methodology</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Program derivation</td>
</tr>
</tbody>
</table>

The material presented in Chapters 2 through 5 appears in modified form in [51] and [50].

12
I end this Part with Chapter 6, which concerns the use of the semantics from Chapter 1 to model popular programming constructs in common programming languages.
Semantics of programs with exceptions

In this chapter, I define the control structures of the programming notation (guarded commands) I use in this thesis. I define each statement in terms of its mathematical interpretation, viz., its weakest precondition and weakest liberal precondition. In Chapter 6, I deal with the relation between these statements and those found in common programming languages, whenever this relation is not immediately apparent. I conclude the present chapter by defining the notion of refinement.

1.0 Weakest precondition

For any statement $S$ and predicate $Q$ on the final state of $S$, Dijkstra [17] defines $wp.S.Q$ to be a predicate on the initial state of $S$:

$\textit{wp}.S.Q$ is true of exactly those initial states from which execution of $S$ is guaranteed to terminate and to terminate in a state satisfying $Q$.

I consider program statements that have two ways of terminating, normally and exceptionally. Therefore, the weakest precondition of a statement maps a pair of predicates on the final state to a predicate on the initial state. I use the notation $wp.S.(P,Q)$ and present the following interpretation [13, 60, 51].

$\textit{wp}.S.(P,Q)$ is true of exactly those initial states from which execution of $S$ is guaranteed to terminate and to either terminate exceptionally in a state satisfying $P$ or normally in a state satisfying $Q$.

So, for example, if, for some statement $S$,

$\textit{wp}.S.(\textit{false, true}) = \textit{true}$,

then $S$ always terminates normally, never exceptionally, because no state satisfies false.
I am restricting my attention to programs with one exceptional outcome. A generalization to an arbitrary number of outcomes is straightforward (see the aforementioned references, [61], Section 2.6, or Section 6.2).

1.0.0 Termination or Lack Thereof

The fact that the weakest precondition captures that programs do terminate is referred to as total correctness. However, when verifying nontrivial programs, we are often willing to settle for less or to prove termination separately. For that purpose, we can consider another attribute of a program statement: its weakest liberal precondition (wlp) [17]. Like wp, wlp maps a program to a predicate transformer, but wlp only guarantees that the postcondition will be reached if the program terminates. This is called partial correctness.

I introduce wlp for exceptional program S and postcondition pair (P, Q) as follows.

\[
\text{wlp.S}(P, Q) \text{ is true of exactly those initial states from which execution of } S \text{ is guaranteed to terminate exceptionally in a state satisfying } P \text{ or normally in a state satisfying } Q \text{ or to not terminate at all.}
\]

Stated differently, \( \text{wp.S}(P, Q) \) guarantees that \( S \) terminates exceptionally in \( P \) or normally in \( Q \), whereas \( \text{wlp.S}(P, Q) \) only guarantees that \( S \) will not terminate exceptionally in \( \neg P \) nor normally in \( \neg Q \).

Note that \( \text{wp.S} \) and \( \text{wlp.S} \) differ only if \( S \) might not terminate. For brevity, I introduce only \( \text{wp} \) in this chapter. Except for the iterative statement, the equation defining the wlp of each statement I introduce is the same as the equation defining the wp of the statement but with every occurrence of wp replaced by wlp.

1.0.1 Monotonicity

In the next several sections, I introduce the program statements and compositions of a simple programming notation. Two kinds of monotonicity are of importance. First is the monotonicity of wp.S, for each statement S. wp.S is monotonic (with respect to \( [\Rightarrow] \) —“implication everywhere”) if for all predicates \( P, P', Q, Q' \), we have

\[
[P \Rightarrow P'] \land [Q \Rightarrow Q'] \Rightarrow [\text{wp.S}(P, Q) \Rightarrow \text{wp.S}(P', Q')]
\]

The importance of this monotonicity will be clear in Section 1.6.

There is an ordering on commands called the refinement ordering, explained in Section 1.8. The other important kind of monotonicity is that every statement composition is monotonic with respect to this refinement ordering in its constituent statements. The importance of this monotonicity is explained in Section 1.8.

Every statement composition I introduce satisfies the second kind of monotonicity, and every simple statement I introduce satisfies the first kind of monotonicity. Consequently, every command that can be constructed from my simple statements and statement compositions satisfies the first kind of monotonicity.
1.1 Assignment

The state of a program consists of a number of independent coordinates called variables. The assignment statement updates the values of these variables.

I now define the assignment statement. As previously advertised, I do so by giving its weakest precondition. For any list \( v \) of program variables and an (equally long) list \( E \) of expressions, I define \( v := E \) by

\[
wp.(v := E).(P, Q) = Q[v := E] \quad .
\]

The right-hand side of this formula is the predicate \( Q \) with every free occurrence of \( v \) replaced by \( E \) (see Section 0.2). The operational interpretation of \( v := E \) is that the list of expressions \( E \) is computed, after which variables \( v \) are updated with the respective computed values of \( E \). For example,

\[
x, y := y, x
\]

has the effect of swapping the values of variables \( x \) and \( y \).

I assume that \( E \) is total, meaning that \( E \) is defined in every state in which the command is ever executed. I discuss partial expressions in Section 6.3.

Here, and throughout this thesis, I assume the evaluation of expressions to have no effect on the program state. Programs written in common programming languages often contain expressions with such so-called side effects, a topic I treat in Section 6.3.

We calculate,

\[
wp.(x := E).(false, true)
= \{ x := \}
true[x := E]
= \{ substitution \}
true
\]

This calculation lets us conclude that the assignment statement always terminates normally.

1.2 Unit statements and compositions

I now define two statements that do not modify the program state, \( skip \) and \( raise \). The former always terminates normally, the latter always exceptionally.

\[
wp.skip.(P, Q) = Q \quad \quad (1.1)
wp.raise.(P, Q) = P \quad \quad (1.2)
\]

Sequential (normal) composition of two statements \( S \) and \( T \), written \( S;T \) (and pronounced "\( S \) semi \( T \)") is defined as

\[
wp.(S;T).(P, Q) = wp.S.(P, wp.T.(P, Q)) \quad \quad (1.3)
\]

16
In words, $wp(S; T). (P, Q)$ is true of those initial states from which either $S$ terminates exceptionally in $P$ (then $T$ is ignored), or $S$ terminates normally in a state from which $T$ either terminates exceptionally in $P$ or terminates normally in $Q$.

I also define exceptional composition, also known as the exception handler. It is written $S \triangleleft T$ (and pronounced “$S$ try $T$”—that’s “try”, almost as in “tri-angle”). The idea is that $T$ “handles” any exception raised by $S$.

\[
wp(S \triangleleft T). (P, Q) = wp.S.(wp.T.(P, Q), Q) \tag{1.4}
\]

Hence, $wp(S \triangleleft T). (P, Q)$ is true of those initial states from which either $S$ terminates normally in $Q$ (then the handler $T$ is ignored), or $S$ terminates exceptionally in a state from which the handler $T$ terminates exceptionally in $P$ or normally in $Q$.

**Remark 1.0.** If it were not clear from the above English descriptions of $;$ and $\triangleleft$, it is certainly clear from formulas (1.3) and (1.4) that there is some duality between the two program compositions. Reviewing (1.1) and (1.2), we also detect a duality. Indeed, by identifying a program with its weakest precondition, we find that functions \textit{skip} and \textit{raise} project to one of the two components of a pair. I will write these functions as $R$ and $L$, respectively. Furthermore, we can write $;$ as \textit{right composition} and $\triangleleft$ as \textit{left composition}, denoted $\circ$ and $\circ$, respectively, over functions of some type $D \times D \rightarrow D$. That is, for any domain $D$, functions $f, g : D \times D \rightarrow D$, and elements $p, q \in D$, we have

\[
(f \circ g). (p, q) = f.(g.(p, q), q), \text{ and } (f \circ g). (p, q) = f.(p, g.(p, q)).
\]

Propelled by this discovery, I explore the phenomenon in Chapter 2.

Other convenient statements that can be defined in terms of \textit{skip}, \textit{raise}, $;$, and $\triangleleft$ are presented in Section 2.5.

### 1.3 Block

The \textit{block} statement, written

\[
[ v \bullet S ]
\]

where $v$ is a list of identifiers, introduces local variables $v$ for use in $S$, the \textit{body} of the block statement.

\[
wp([ v \bullet S ]). (P, Q) = (\forall v \triangleright wp.S.(P, Q)) \tag{1.5}
\]

In words, the block statement guarantees $(P, Q)$ upon termination exactly when $wp.S.(P, Q)$ holds initially, \textit{for any value of} $v$. Thus, $S$ must not depend on $v$ having any particular initial value.
1.4 Partial commands

The weakest precondition of a command $S$ is said to be strict just when

$$wp.S.(false,false) = false$$

In [17], this property is referred to as the (Law of the) Excluded Miracle, because statements that lack this property do not, in general, lend themselves to a practical implementation—$wp.S.(false,false)$ characterizes those initial states from which $S$ is guaranteed to terminate in a state satisfying false!

A statement $S$ whose weakest precondition is not strict is called partial [60, 70] (or miraculous or feasible [67]), because one may think of it as being executable only from the initial states satisfying

$$\neg wp.S.(false,false)$$

(See also Remark 1.1 below.) A command that is not partial is said to be total.

Despite the fact that they do not always admit a realistic implementation, partial commands are important and useful when handled with care [70, 67, 66]. I give some examples below.

1.4.0 The guard statement

A primitive partial command is the guard statement, $g \rightarrow S$, where $g$ is a predicate and $S$ is a command. Operator $\rightarrow$ binds stronger than $;$ and $<$, but weaker than $\land$ logic connectives. The non-exceptional definition of the guard statement is

$$wp.(g \rightarrow S).Q = \neg g \lor wp.S.Q$$

I extend this definition in the obvious way.

$$wp.(g \rightarrow S).(P,Q) = \neg g \lor wp.S.(P,Q)$$

Note that (1.6) can also be written

$$wp.(g \rightarrow S).(P,Q) = g \Rightarrow wp.S.(P,Q)$$

which has appeal because $\rightarrow$ and $\Rightarrow$ look similar.

An operational interpretation of the guard statement is that $g \rightarrow S$ is like $S$, except that in addition to the states from which $S$ cannot be started, $g \rightarrow S$ cannot be started in states where $g$ does not hold. An alternative operational interpretation is that $g \rightarrow S$ “invokes a miracle” when $g$ does not hold and invokes $S$ otherwise.

Remark 1.1. In the first of these interpretations, one can let the execution of a partial command from a state in which it cannot begin cause the entire program to backtrack. This is what the text processing language LIM does [10]. Nevertheless, I will continue to use the terminology “invoke a miracle.”
For a total command \( S \), \( g \) is called the \textit{guard} of the command \( g \rightarrow S \). This concept can be generalized: The \textit{guard} of any (total or partial) command \( S \), denoted \( \text{guard}.S \), is defined by
\[
\text{guard}.S = \neg \text{wp}.S.(\text{false},\text{false})
\]
This is, the guard of a statement characterizes those initial states from which execution of \( S \) can start.

Note that
\[
\text{guard}.S \rightarrow S = S
\]
In fact, for any \( g \) such that \([\text{guard}.S \Rightarrow g]\),
\[
g \rightarrow S = S
\]
This shows, for example, that \( g \rightarrow \) is idempotent, \( i.e., \)
\[
g \rightarrow (g \rightarrow S) = g \rightarrow S
\]
It also shows that \( \text{true} \) is a left identity of \( \rightarrow \).
\[
\text{true} \rightarrow S = S
\]

1.4.1 \textbf{Examples}

I give a few examples. I make use of the \textit{assert} statement, which I discuss in Section 6.2.0. For now, all we need to know is that \textit{assert} \( g \) terminates normally just when \( g \) holds initially.
\[
\text{wp}.(\text{assert} \ g).(P,Q) = g \land Q
\]
If \( S \) is total, then \( g \rightarrow S \) invokes a miracle precisely when \( g \) does not hold initially. Thus, despite its use of \( \rightarrow \),
\[
\text{assert} \ g ; g \rightarrow S
\]
is total, because the statement \( g \rightarrow S \) is only reached if \( g \) holds. In general, for any \( S \),
\[
\text{assert} \ \text{guard}.S ; S
\]
is total.
Using a guard statement as the body of a block proves convenient. For example, for a total command \( S \), executing
\[
\llbracket x \cdot x^2 = g \rightarrow S \rrbracket
\]
has the effect of executing \( S \) from a state where \( x = 3 \lor x = -3 \) holds. The reader is invited to prove
\[
\text{wp}.(\llbracket x \cdot x^2 = g \rightarrow \text{assert} \ x = 3 \lor x = -3 \rrbracket).(\text{false},\text{true}) = \text{true}
\]
Remark 1.2. Juno-2 [35] is a language that achieves expressive power by taking advantage of partial commands in this way.

Other convenient uses of blocks with partial commands are described in [70]. Partial commands come in handy also when developing programs through refinements, as is shown, for example, in [67, 66].

1.5 Choice compositions

Next, I introduce two choice compositions, □ ("box") and □ ("else") [70], each with lower binding power than →.

\[ wp.(S \circ T).(P, Q) = wp.S.(P, Q) \land wp.T.(P, Q) \]
\[ S \circ T = guard.S \rightarrow S \circ \neg guard.S \rightarrow T \]

We have that □ is associative and so is □ (the fact that □ is follows directly from its definition; the fact that □ is makes a nice exercise for the reader). Beware that, contrary to what the appearance of its symbol suggests, □ is not symmetric.

The execution of S □ T consists of the execution of exactly one of S and T, and so does the execution of S □ T. The difference is that while S □ T guarantees nothing about which of S and T is chosen for execution, execution of S □ T reduces to execution of S whenever S □ T is started in a state where execution of S can begin (i.e., does not invoke a miracle) and reduces to T otherwise.

An example of a choice composition is

\[ g0 \rightarrow S0 \circ g1 \rightarrow S1 \circ g2 \rightarrow S2 \]  \hspace{1cm} (1.7)

If S0, S1, S2 are total commands, execution of (1.7) can result in the execution of S0 if g0 holds (initially), S1 if g1 holds, and S2 if g2 holds. If g0 \lor g1 \lor g2 holds initially, exactly one of S0, S1, S2 is executed; if \neg g0 \land \neg g1 \land \neg g2 holds initially, then (1.7) invokes a miracle.

The statement

\[ g0 \rightarrow S0 \circ g1 \rightarrow S1 \circ g2 \rightarrow S2 \]  \hspace{1cm} (1.8)

is like (1.7), except that S0 is chosen just when g0 holds, S1 just when \neg g0 \land g1 holds, and S2 just when \neg g0 \land \neg g1 \land g2 holds. The statement

\[ g0 \rightarrow S0 \circ g1 \rightarrow S1 \circ g2 \rightarrow S2 \circ \text{skip} \]  \hspace{1cm} (1.9)

is similar to (1.8), except that it terminates without changing the state (it "skips") where (1.8) invokes a miracle. (Note that, since true is a left identity element of →, the last \text{skip} can also be written as true → \text{skip}.)

For similarity with common notation, I permit myself to surround total commands with if \text{fi} brackets. So that I don’t need to attach any meaning to these brackets, I will refrain from the use of if S \text{fi} if S is partial. So, I may write (1.9) equivalently as

\[ \text{if } g0 \rightarrow S0 \circ g1 \rightarrow S1 \circ g2 \rightarrow S2 \circ \text{true} → \text{skip } \text{fi} \]  \hspace{1cm} (1.10)
1.6 Iteration

Iteration of a statement $S$ is written $\text{do } g \mapsto S \text{ od}$.  

Remark 1.3. The symbol $\mapsto$ is not to be confused with the guard operator $\rightarrow$. The construct $\text{do } g \mapsto S \text{ od} $ is a construct parameterized by a predicate $g$ and a statement $S$.

Let this iterative statement (or loop) be denoted by $DO$. Operationally, the execution of $DO$ consists of repeated executions of $S$ for as long as $g$ holds. We think of $DO$ as satisfying

$$ DO \ = \ g \rightarrow S; \ DO \ \lnot g \rightarrow \text{skip} \ . $$

Inspired by this, I define $\wp.DO.(P,Q)$ as the least fixed point of the equation

$$ X \ = \ (g \Rightarrow \wp.S.X) \land (\lnot g \Rightarrow Q) \ , \qquad (1.11) $$

solved for $X$. Because $\wp.S$ is monotonic (cf. Section 1.0.1), the right-hand side of (1.11) is a monotonic function of $X$; hence, due to the Knaster-Tarski Theorem (see, e.g., [76]), (1.11) does indeed have a least fixed point, so $\wp.DO.(P,Q)$ is well-defined. It is for this reason that the first kind of monotonicity mentioned in Section 1.0.1 is important.

Remark 1.4. The fact that the loop satisfies the second kind of monotonicity mentioned in that section follows from the fact that the right-hand side of (1.11) is monotonic in $S$ (see, e.g., (130) in [76]). Note that this property holds even for partial statements $S$.

Similarly, $\wpp.DO.(P,Q)$ is defined as the greatest fixed point of

$$ X \ = \ (g \Rightarrow \wpp.S.X) \land (\lnot g \Rightarrow Q) \ , $$

solved for $X$ (cf. [21]).

The fact that the definition of $DO$ involves a fixed point may appear worrisome—if proving something about a loop would require computing a particular fixed point, the practical application of the semantics of the loop would be hampered. However, the Invariance Theorem [21] shows a sufficient condition for proving that a loop establishes a particular postcondition. It states that

$$ [\text{Pre} \Rightarrow \wpp.(\text{do } g \mapsto S \text{ od}).(P,Q)] \quad (1.12) $$

follows from

$$ (\exists J \Rightarrow [\text{Pre} \Rightarrow J] \land [J \land g \Rightarrow \wpp.S.(P,J)] \land [J \land \lnot g \Rightarrow Q] ) \quad (1.13) $$

and a proof that the loop eventually terminates. The $J$ in (1.13) is called the loop invariant. Equation (1.13) states that there exists an invariant $J$ that satisfies three conditions. First, the invariant holds prior to the execution of the loop. Second,
provided the invariant holds initially, an execution of the loop body terminates exceptionally in a state satisfying \( P \) (upon which the loop terminates) or normally in a state satisfying, once again, the invariant. Third, the invariant conjoined with the negation of the guard imply \( Q \), the desired normal postcondition of the loop.

The existential quantification in (1.13) may look intimidating because \( f \) ranges over all predicates. However, the programmer who writes the loop has a good idea of what an invariant of the loop may be. Having the programmer supply that invariant \( f \) simplifies (1.13) to the satisfiability of each of the conjuncts in the term of its quantification.

**Remark 1.5.** Instead of having the programmer provide the invariant, methods of “widening” and “narrowing” can be used in an attempt to synthesize a proper invariant (see [12, 8]).

Termination can be handled in a similar way by letting the programmer supply a *bound function* [17]. If termination is not of concern, \( wlp \) can be used. The \( wlp \) equation corresponding to (1.12) follows from the \( wlp \) equation corresponding to (1.13) alone, without any further proof of termination.

The loop that never terminates, in its simplest form written as \( \text{do } \text{true } \mapsto \text{skip } \text{od} \), is often referred to as \( \text{abort} \) [17].

### 1.7 Specification statement

As a final statement, I introduce the *specification statement* [67]. Its non-exceptional form takes the shape

\[
  w : [\text{Pre}, \text{Post}],
\]

where \( w \) (called the *frame*) is an (unordered) list of variables, and \( \text{Pre} \) (the *precondition*) and \( \text{Post} \) (the *postcondition*) are predicates. If started in a state satisfying \( \text{Pre} \), the statement will terminate in a state satisfying \( \text{Post} \), having modified the values only of variables \( w \). The weakest precondition of this statement is thus given by

\[
  wp.(w : [\text{Pre}, \text{Post}]).Q = \text{Pre} \land (\forall w \mid \text{Post } \triangleright Q).
\]

(1.14)

That is, for \( w : [\text{Pre}, \text{Post}] \) to establish \( Q \), \( \text{Pre} \) must hold initially. In addition, the following must hold initially: For any values of variables \( w \) that satisfy \( \text{Post} \), \( Q \) holds. (I defer until Section 6.2 describing what the operational interpretation of the statement is when \( \text{Pre} \) does not hold initially.)

We often want to specify the final values of \( w \) in terms of their initial values. We can then think of “saving” the initial values of some variables \( v \) (admittedly usually a subset of \( w \)), as in

\[
  \left[ v_0 \bullet v_0 := v ; w : [\text{Pre}, \text{Post}] \right].
\]

(1.15)
Then, \textit{Post} can refer to \(v_0\). I abbreviate (1.15) simply as

\[ w : [\text{Pre, Post}] \]

where \textit{Post} may subscript the name of any variable with 0 to refer to the initial value of that variable. I assume that actual names of program variables never end with a subscripted 0. I refer to a variable subscripted with 0 as an \textit{initial-value} variable. (The terms \textit{entry}, \textit{logical}, and \textit{mathematical} variables are also in use.) \textit{Post} can be thought of as a \textit{two-state} predicate, since it is a predicate on both the initial and final states.

Note that \(w\) and \textit{Pre} do not mention any initial-value variables. Also note that it doesn’t change the semantics of the specification statement if more variables are saved in (1.15) than are needed in \textit{Post}; hence, when referring to the initial-value variables in \textit{Post}, we may safely consider any superset of those initial-value variables that actually occur in \textit{Post}.

Let \(v_0\) be (any superset of) the initial-value variables in \textit{Post}. The following calculation arrives at the definition of the specification statement that permits initial-value variables.

\[
wp.(w : [\text{Pre, Post}]).Q
= \begin{cases} 
\text{ shorthand for (1.15) } \\
wp.([v_0 \bullet v_0 := v ; w : [\text{Pre, Post}] ]).Q \\
\{ (1.5): \text{ wp of } [\bullet ] \} \\
(\forall v_0 \triangleright wp.(v_0 := v ; w : [\text{Pre, Post}]).Q) \\
\{ (1.3,1.0): \text{ wp of ; and := } \} \\
(\forall v_0 \triangleright (wp.(w : [\text{Pre, Post}]).Q)[v_0 := v]) \\
\{ v_0 \text{ does not occur free in term of quantification } \} \\
(wp.(w : [\text{Pre, Post}]).Q)[v_0 := v] \\
\{ (1.14): \text{ wp of specification statement without initial-value variables } \} \\
(\text{Pre } \land (\forall w \triangleright (Q))[v_0 := v]) \\
\{ v_0 \text{ does not occur free in Pre } \} \\
\text{Pre } \land (\forall w \triangleright Q)[v_0 := v]
\end{cases}
\]

(1.16)

\textbf{Remark 1.6.} The observation about \(v_0 := v\) is one I thought to be folklore. Later, I traced it back to having been a discovery [18, pp. 217–219]. Maybe this attests how the general understanding of semantics has grown during the last two decades.

In the realm of exceptional programs, I extend the postcondition in the specification statement to be a pair of predicates.

\[
wp.(w : [\text{Pre, (cPost, nPost)}]).(P, Q) = \begin{cases} 
\text{Pre } \land (\forall w \triangleright (cPost \Rightarrow P) \land (nP \Rightarrow Q))[v_0 := v] \\
\end{cases}
\]

(1.17)

Note that the specification statement can be a partial command, \textit{e.g.,} \(w : [\text{true, (false, false)}]\).
1.8 Refinement

The specification statement offers a high-level notation that conveniently expresses what the command does. However, not only does it have the possibility of being partial, the specification statement cannot easily be compiled automatically into more primitive statements, because it does not give any clues as to how the command is to arrive at the specified final state. The statement thus lends itself well to writing a specification, the implementation of which requires guidance from the programmer. We want to be able to prove that the implementation meets the specification. This leads to the concept of program refinement, first proposed by Dijkstra [16] and Wirth [80], and first given a mathematical foundation by Back [2].

Informally, a statement \( S \) is refined by a statement \( T \), denoted \( S \sqsubseteq T \), just when \( T \) meets any specification that \( S \) does. Formally, the non-exceptional form of \( \sqsubseteq \) is defined by

\[
(S \sqsubseteq T) = (\forall R \triangleright [wp.S.R \Rightarrow wp.T.R]),
\]  

(1.18)

where \( R \) ranges over all predicates on the final state. In the complete boolean lattice of predicate transformers, elements are ordered by \( \sqsubseteq \), which is the lifting of \( [ \Rightarrow ] \) in the complete boolean lattice of predicates [82]. If \( wlp \), not \( wp \), is of interest, then (1.18) is written with \( wp \) replaced by \( wlp \).

Examples of refinements are

\[
x : [true, x_0 < x] \sqsubseteq x := x + 3
\]

and

\[
S \sqcap T \sqsubseteq S
\]

In both of these examples, the left-hand sides allow a greater degree of nondeterminism than the respective right-hand sides.

So, given \( S \sqsubseteq T \), we can always replace the program \( S \) by the program \( T \). \( T \) is therefore sometimes said to be “better than” \( S \) [39]. Note, however, that some commands are “too good”. For example, \( wp.(false \rightarrow skip).R = true \), so \( false \rightarrow skip \) refines (or “is better than”) any command. However, \( false \rightarrow skip \) is a partial command that usually cannot realistically be implemented (the exception is again backtracking, see Remark 1.1). This shows that when refining a program, there is a risk of ending up with a non-implementable program [39, 68].

We may wonder if \( S \sqsubseteq T \) allows us to replace \( S \) by \( T \) in any context. That is, if \( S \) is a subcomponent of a larger program \( P \), does replacing \( S \) by \( T \) in \( P \) result in a refinement of \( P \)? Because this is important in program development, I require that this property hold for all programs under consideration. The fact that \( S \) is a subcomponent of \( P \) is captured by writing \( P \) as a function of \( S \), say \( f.S \). Then, the requirement can be written down as

\[
(S \sqsubseteq T) \Rightarrow (f.S \sqsubseteq f.T)
\]
a formula we recognize as expressing the monotonicity of $f$. It is primarily for this reason that I consider only those statement compositions that are monotonic in their constituent statements (see Section 1.0.1).

For exceptional programs $S$ and $T$, I define refinement by

$$ (S \sqsubseteq T) = \left\{ \forall P, Q \triangleright [\text{wp}.S(P, Q) \Rightarrow \text{wp}.T(P, Q)] \right\}, \quad (1.19) $$

and similar for the $\text{wp}$ counterpart.
Chapter 2

Functions of two arguments and their compositions

To show that program constructs in a language with exceptions are not appreciably more difficult to reason about than those in a language without exceptions, I show that the weakest preconditions of these constructs make up a nice algebra over functions of two arguments.

I first introduce the algebra, and then, as alluded to in Remark 1.0 in Section 1.2, make the connection with program statements.

2.0 Function compositions

Consider functions \( f \) and \( g \) of type \( D \to D \), for any domain \( D \). We are accustomed to composing these functions, that is, applying one after the other. We use \( \circ \) to denote function composition, and recognize its familiar definition, for any element \( x \) of \( D \),

\[
(f \circ g).x = f.(g.x)
\]

Now, consider functions \( f \) and \( g \) of type \( D \times D \to D \). These cannot be composed in the same way as the previous functions, because the expression

\[
f.(g.(x,y))
\]

where \((x,y)\) is a pair of type \( D \times D \), doesn’t make any sense since the types don’t match up: the expression \( g.(x,y) \) has type \( D \), whereas the domain of \( f \) is \( D \times D \).

From this, we conclude that composing functions of two arguments requires an operator different from \( \circ \). In fact, such functions can be composed in several different ways.

In the rest of this chapter, I use \( f \), \( g \), and \( h \) to denote any functions of type \( D \times D \to D \) for any domain \( D \), and \( p \) and \( q \) to denote any elements in \( D \). As before, an ordered pair with components \( p \) and \( q \) is written \((p,q)\).
2.1 Left and right composition

Of the different ways functions from pairs to elements can be composed, we first consider left and right composition, written \( \circ \) and \( \cdot \), respectively. I define these as follows.

\[
(f \circ g).(p, q) = f.(g.(p, q), q) \quad \text{(2.0)}
\]

\[
(f \cdot g).(p, q) = f.(p, g.(p, q)) \quad \text{(2.1)}
\]

**Theorem**

\( \circ \) is associative.

\( \cdot \) is associative.

**Proof.**

\[
(f \circ (g \circ h).(p, q))
\]

\[
= \quad \{ \text{(2.0): def. of } \circ \} \]

\[
f.(g \circ h).(p, q), q)
\]

\[
= \quad \{ \text{(2.0): def. of } \circ \} \]

\[
f.(h.(p, q), q)
\]

\[
= \quad \{ \text{(2.0): def. of } \circ \} \]

\[
((f \circ g) \circ h).(p, q)
\]

I omit the proof of the other case as it is similar to the present case (and will do so in many more proofs).

Two functions of special interest are \( L \) and \( R \), defined as follows.

\[
L.(p, q) = p \quad \text{(2.4)}
\]

\[
R.(p, q) = q \quad \text{(2.5)}
\]

**Theorem**

\( L \) is the identity of \( \circ \).

\( R \) is the identity of \( \cdot \).

**Proof.**

\[
(L \circ f).(p, q)
\]

\[
= \quad \{ \text{(2.0): def. of } \circ \} \]

\[
L.(f.(p, q), q)
\]

\[
= \quad \{ \text{(2.4): def. of } L \} \]

\[
f.(p, q)
\]

\[
= \quad \{ \text{(2.4): def. of } L \} \]

\[
f.(L.(p, q), q)
\]

\[
= \quad \{ \text{(2.0): def. of } \circ \} \]

\[
(f \circ L).(p, q)
\]

\[\square\]
Theorem

$L$ is a left zero of $\circ$.

$R$ is a left zero of $\circ$.

Proof.

\[(L \circ g). (p,q)\]

= \[
\begin{cases}
(2.1): \text{def. of } \circ 
\end{cases}
\]

\[L.(p,g.(p,q))\]

= \[
\begin{cases}
(2.4): \text{def. of } L 
\end{cases}
\]

\[p\]

= \[
\begin{cases}
(2.4): \text{def. of } L 
\end{cases}
\]

\[L.(p,q)\] □

2.2 Double composition

As a different way to compose functions of two arguments, I define double composition, written $\langle \circ \rangle$.

\[(f \langle \circ \rangle g). (p,q) = f.(g.(p,q),g.(p,q))\] (2.10)

We have the following correspondences between single (left and right) compositions and double composition.

Theorem

\[f \langle \circ \rangle g = (f \langle \circ R \rangle) \circ g\] (2.11)

\[f \langle \circ \rangle g = (f \langle \circ L \rangle) \circ g\] (2.12)

Proof.

\[(f \langle \circ R \rangle \circ g).(p,q)\]

= \[
\begin{cases}
(2.1): \text{def. of } \circ 
\end{cases}
\]

\[(f \circ R).(p,g.(p,q))\]

= \[
\begin{cases}
(2.0): \text{def. of } \circ 
\end{cases}
\]

\[f.(R.(p,g.(p,q)),g.(p,q))\]

= \[
\begin{cases}
(2.5): \text{def. of } R 
\end{cases}
\]

\[f.(g.(p,q),g.(p,q))\]

= \[
\begin{cases}
(2.10): \text{def. of } \langle \circ \rangle 
\end{cases}
\]

\[(f \circ g).(p,q)\] □

Remark 2.0. I stated this property as an equality between functions; however, its proof applies those functions to an arbitrary pair $(p,q)$. I strive toward calculations at the level of functions, since they tend to be more concise and easier to read. As it turns out, having shown the above relation between single and double compositions, I am now able to carry out the calculations at the level of functions. This phenomenon is commonly referred to as lifting (Section 0.2).
I continue with some theorems regarding the associativity and distributivity of the composition operators.

**Theorem**

\[
(f \circ g) \circ h = f \circ (g \circ h) \quad (2.13)
\]

\[
(f \circ g) \circ h = f \circ (g \circ h) \quad (2.14)
\]

**Proof.**

\[
(f \circ g) \circ h = \quad \{ (2.12): \text{double/single trade, and (2.2): } \circ \text{ is associative } \} \\
(f \circ L) \circ g \circ h = \quad \{ (2.12): \text{double/single trade, and (2.2): } \circ \text{ is associative } \} \\
f(\circ) (g \circ h) \quad \square
\]

Now for a useful theorem whose proof is rather curious—maybe the most interesting proof in this chapter.

**Theorem**

\( \circ \) is associative. \quad (2.15)

**Proof.**

\[
f(\circ) (g \circ h) = \quad \{ (2.11): \text{double/single trade } \} \\
f(\circ) ((g \circ R) \circ h) = \quad \{ (2.14): \text{mutual associativity of } \circ \} \\
(f \circ \circ) (g \circ R) \circ h = \quad \{ (2.13): \text{mutual associativity of } \circ \} \\
((f \circ g) \circ \circ R) \circ h = \quad \{ (2.11): \text{double/single trade } \} \\
(f \circ g) \circ \circ h \quad \square
\]

**Theorem**

\( L \) and \( R \) are left identities of \( \circ \). \quad (2.16)

**Proof.**

\[
L(\circ) g = \quad \{ (2.11): \text{double/single trade } \} \\
(L(\circ) R) \circ g = \quad \{ (2.6): L \text{ is identity of } \circ \} \\
R \circ g = \quad \{ (2.7): R \text{ is identity of } \circ \} 
\]
\[
\begin{align*}
  g &= \{ \text{(2.6): } \ L \text{ is identity of } \circ \} \\
  L \circ g &= \{ \text{(2.7): } \ R \text{ is identity of } \circ \} \\
  (R \circ L) \circ g &= \{ \text{(2.12): double/single trade } \} \\
  R \circ g &= \text{ } \square
\end{align*}
\]

A consequence of this theorem, since \(L\) and \(R\) differ, is that \(\circ\) lacks a right identity. However, \(\circ\) with \(L\) or \(R\) as a second argument is still interesting, as is shown by the next theorem.

**Theorem**

\[
\begin{align*}
  f \circ L &= f \circ L \\
  f \circ R &= f \circ R
\end{align*}
\]  \tag{2.17} \tag{2.18}

**Proof.**

\[
\begin{align*}
  f \circ L &= \{ \text{(2.12): double/single trade } \} \\
  (f \circ L) \circ L &= \{ \text{(2.6): } \ L \text{ is identity of } \circ \} \\
  f \circ L &= \text{ } \square
\end{align*}
\]

I find these instances where \(\circ\) equals \(\circ\) or \(\circ\) curious—in fact, so curious that I will devote the next section to it.

### 2.3 Ceiling and floor

Intrigued by (2.17) and (2.18), I introduce some special notation, \(\lceil \ \rceil\) and \(\lfloor \ \rfloor\), defined as follows.

\[
\begin{align*}
  [f] &= f \circ L \\
  [f] &= f \circ R
\end{align*}
\]  \tag{2.19} \tag{2.20}

This leads us to the following theorems.

**Theorem**

\[
\begin{align*}
  [L] &= L = [R] \\
  [L] &= R = [R]
\end{align*}
\]  \tag{2.21} \tag{2.22}
Proof.

\[
\begin{align*}
[L] &= \{ \text{(2.19): def. of } [\ ] \} \\
L(\diamond) L &= \{ \text{(2.16): } L \text{ is left identity of } \langle \diamond \rangle \} \\
L &= \{ \text{(2.16): } R \text{ is left identity of } \langle \diamond \rangle \} \\
R(\diamond) L &= \{ \text{(2.19): def. of } [\ ] \} \\
[R] &\quad \square
\end{align*}
\]

How we think about an operator influences the notation we choose. This is important, because the notation we choose in turn inspires how we think about the operator! For example, we know well not to make the mistake of writing + and \cdot for \lor and \land, respectively; doing that effectively hides the fact that \lor distributes over \land, a property not enjoyed by the arithmetic operators + and \cdot \cite{25}. The following theorem justifies the use of ceiling and floor. These ceiling and floor operators can also be shown to be monotonic (with respect to any ordering over functions of two arguments), just like the ones operating on real numbers.

**Theorem**

\[
[\ ] \text{ and } [\ ] \text{ are idempotent.} \quad (2.23)
\]

Proof.

\[
\begin{align*}
[[f]] &= \{ \text{(2.19): def. of } [\ ], \text{ twice, and (2.15): } \langle \diamond \rangle \text{ is associative } \} \\
f(\diamond) L(\diamond) L &= \{ \text{(2.16): } L \text{ is left identity of } \langle \diamond \rangle \} \\
f(\diamond) L &= \{ \text{(2.19): def. of } [\ ] \} \\
[f] &\quad \square
\end{align*}
\]

**Theorem**

\[
\begin{align*}
f(\diamond) g &= [f] \langle \diamond \rangle g \\
f(\diamond) g &= [f] \langle \diamond \rangle g \quad (2.24)
\end{align*}
\]

Proof.

\[
\begin{align*}
f(\diamond) g &= \{ \text{(2.12): double/single trade } \} \\
(f(\diamond) L) \langle \diamond \rangle g &= \{ \text{(2.17): } \langle \diamond \rangle L \text{ and } \langle \diamond \rangle L \}
\end{align*}
\]

31
\[
(f \circ L) \circ g
= \begin{cases}
(2.19): \text{def. of } [ ] & \\
[f] \circ g
\end{cases}
\]
\[
\square
\]

**Theorem**

\[
[f \circ g] = [f] \circ [g]
\]
\[
[f \circ g] = [f] \circ [g]
\]  
(2.26)

Proof. Using the associativity of \( \circ \) in every step, we calculate,

\[
[f \circ g]
= \begin{cases}
(2.19): \text{def. of } [ ] & \\
[f] \circ [g]
\end{cases}
\]
\[
f \circ g \circ L
= \begin{cases}
(2.16): \text{L is left identity of } \circ & \\
f \circ L \circ [g] \circ L
\end{cases}
\]
\[
[f \circ g] = \begin{cases}
(2.19): \text{def. of } [ ] \text{, twice } & \\
[f] \circ [g]
\end{cases}
\]
\[
\square
\]

**Theorem**

\[
[f \circ g] = [f] \circ [g]
\]
\[
[f \circ g] = [f] \circ [g]
\]

Proof.

\[
[f \circ g]
= \begin{cases}
(2.26): \text{\( \circ \) over } [ ] & \\
[f] \circ [g]
\end{cases}
\]
\[
[f] \circ [g]
= \begin{cases}
(2.24) & \\
[[f]] \circ [g]
\end{cases}
\]
\[
[f] \circ [g]
= \begin{cases}
(2.23): \text{idempotence of } [ ] & \\
[f] \circ [g]
\end{cases}
\]
\[
\square
\]

2.4 Transposition

I introduce operator \( \sim \) with higher binding power than composition and function application, defined as follows.

\[
\sim f.(p, q) = f.(q, p)
\]  
(2.28)

Clearly, \( \sim \) is an *involution*, that is, \( \sim \sim \) is the identity function \((\sim \sim f = f)\).

**Remark 2.1.** Having introduced a new operator, defined at the level of pairs, the proof of the next theorem is done at the point level.
Theorem

\[ \sim(f \circ g) = \sim f \circ \sim g \]
\[ \sim(f \circ g) = \sim f \circ \sim g \]
\[ \sim(f \circ g) = \sim f \circ \sim g \]

Proof. For left composition, we have,

\[ \sim(f \circ g).p.q \]
\[ = \{ \text{(2.28): def. of } \sim \} \]
\[ (f \circ g).q.p \]
\[ = \{ \text{(2.0): def. of } \circ \} \]
\[ f.(g.(q,p),p) \]
\[ = \{ \text{(2.28): def. of } \sim \} \]
\[ f.(\sim g.(p,q),p) \]
\[ = \{ \text{(2.28): def. of } \sim \} \]
\[ f.(p,\sim g.(p,q)) \]
\[ = \{ \text{(2.1): def. of } \circ \} \]
\[ (\sim f \circ \sim g).(p,q) \]

and similar for right composition. For double composition, we have,

\[ \sim(f \circ g).p.q \]
\[ = \{ \text{(2.28): def. of } \sim \} \]
\[ (f \circ g).q.p \]
\[ = \{ \text{(2.10): def. of } \circ \} \]
\[ f.(g.(q,p),g.(q,p)) \]
\[ = \{ \text{(2.28): def. of } \sim \text{, twice } \} \]
\[ f.(\sim g.(p,q),\sim g.(p,q)) \]
\[ = \{ \text{(2.28): def. of } \sim \} \]
\[ f.(\sim g.(p,q),\sim g.(p,q)) \]
\[ = \{ \text{(2.10): def. of } \circ \} \]
\[ (\sim f \circ \sim g).(p,q) \]

This theorem shows the duality between \( \circ \) and \( \circ \).

2.5 The connection with programs

Section 1.2 motivated this chapter. By identifying a program with its weakest precondition, the connection between programs and functions of two arguments is summarized as follows.

\[ \text{skip} \quad = \quad R \]
\[ \text{raise} \quad = \quad L \]
\[ S;T \quad = \quad S \circ T \quad \text{or more succinctly: } ; \quad = \quad \circ \]
\[ S \triangleleft T \quad = \quad S \circ T \quad \text{or more succinctly: } \triangleleft \quad = \quad \circ \]

33
I proceed to explain the connections between the other operators in the algebra and programs. Using (2.25) and (2.24), we have

\[
S \langle \circ \rangle T = (S \triangleleft \text{skip}); T \\
S \langle \circ \rangle T = (S; \text{raise}) \triangleleft T
\]

From the definitions of \([\ ]\) and \(\lbrack \ ]\), and from (2.18) and (2.17), we have

\[
[S] = S \triangleleft \text{skip} \\
[S] = S; \text{raise}
\]

In words, the execution of \([S]\) is like that of \(S\), except that when \(S\) terminates at all, \([S]\) terminates normally. Similarly, the execution of \([S]\) is like that of \(S\), except that when \(S\) terminates at all, \([S]\) terminates exceptionally. The execution of \(S \langle \circ \rangle T\) consists of the execution of \(S\) followed by, provided \(S\) terminates at all, the execution of \(T\).

Transposition \(\sim S\) is the statement that terminates just when \(S\) does, and upon termination complements the outcome. We can implement \(\sim S\) as

\[
[ b \bullet ((S; b := \text{true}) \triangleleft b := \text{false}) ; \text{if } b \rightarrow \text{raise} \quad \lnot b \rightarrow \text{skip} \quad \text{fi }]
\]

As a program construct, \(\sim\) seems to be of limited use, but maybe that's just our lack of imagination. In the algebra, however, it allows us to prove the duality between \(\langle \circ\rangle\) and \(\circ\).

Modula-3 is an example of a programming language with exceptions. In addition to the \(\triangleleft\) construct, it has a so-called try finally statement. Execution of

\[
\text{TRY } S \text{ FINALLY } T \text{ END}
\]

consists of the execution of \(S\) followed by the execution of \(T\). If the execution of \(S\) terminates exceptionally, then execution of \(T\) is followed by reraising the exception. This construct can be captured by

\[
(S \triangleleft (T; \text{raise})); T
\]

Finally, I remark on the relation between the theory presented herein and existing programming languages. We find that usual programming languages introduce an asymmetry between left and right composition. For example, statements begin their execution in a normal state, \(\triangleleft\) is often much longer to type than \(;\), and \(\triangleleft\) may not be as efficient as \(;\) (see, e.g., [71]). However, the properties presented in this chapter suggest a more symmetric treatment of \(;\) and \(\triangleleft\).

2.6 Concluding remarks

We have seen that the algebra over functions of two arguments and their compositions—the algebra that underlies weakest preconditions for programs with exceptions—is simple and elegant. Consequently, there is good hope of getting a computer to
perform calculations within this algebra, as is needed to do automatic verification of programs with exceptions.

A generalization of the present algebra to functions of type $D^n \to D$ for any $n$ is found in [61]—in fact, the generalization does not even require that $n$ be finite. [61] also presents a way, via partitioned predicates (see Remark 3.3), to handle exceptions without requiring $\wp.S$ be a function from a pair of predicates to a predicate, but instead allowing it to remain a function from one predicate to another.

In previous studies of exceptions (and mathematics in general), this algebra has gone unnoticed. One reason for this has been suboptimal choices of notation for exceptional weakest preconditions. Where I write $\wp.S.(P,Q)$, [60] writes $\wp(S,Q,P)$ and [13] uses the two functions $\wp(S,e,P)$ and $\wp(S,;,Q)$. My notation has the advantage over the others of permitting the separation of the function $\wp.S$ from its argument $(P,Q)$. Only then does the opportunity to discover the algebra over functions of type $D \times D \to D$ arise.

**Remark 2.2.** Dijkstra came to the same conclusion regarding his non-exceptional $\wp$, as is witnessed by contrasting the notation in [17] with that in [21].

The notation in [13] is readily extensible to more than one exception. However, to specify that a program does not raise any exceptions, one needs information about all declared exceptions, so the notation is not as extensible as it may first appear.

One of two other approaches to allowing more than one exception is to add a special variable which indicates which exception is raised. The variable would be updated immediate preceding a `raise` statement. The other possibility is to extend the pair to an arbitrary tuple. This is done in [4] and [61].

Application of the present algebra is not limited to the semantics of programs with exceptions—it can be used for any functions of type $D \times D \to D$. Another appearance of such functions in computer science is [22]. Also, [61] considers some other application areas, including relational databases and embedded systems.
Trace semantics for exceptions

In Chapter 1, I introduced the weakest preconditions for exceptional program constructs. More precisely, I let the weakest preconditions (and weakest liberal preconditions) define these programs constructs. In Section 2.5, I described constructs like \(<\), \([\ ]\), and try finally operationally. We may thus ask, “Does our operational interpretation of these constructs correspond to their mathematical weakest precondition definitions?” Not only is this issue relevant when using the constructs to write programs, but it becomes unavoidable when we try to implement the program constructs.

So, the question is, “Are these constructs the ones we think they are?”. Focusing on assignment, the unit statements, and normal and exceptional composition, I lead us in this chapter to come to grips with the answer to that question.

3.0 Introduction

One approach to convincing ourselves that we have indeed defined the intended constructs is taken by Manasse and Nelson in [60], where exceptional programs are translated into simpler constructs. The implementation of these simpler constructs is more familiar to us, and is discussed in [21, Ch. 10]. In the present chapter, I take an approach different from that in [60]—in some sense, I approach the problem from the other end. To understand my approach, let’s start by contrasting different semantics of a program.

3.0.0 Concreteness of a semantics

One reason the weakest-precondition semantics is so useful is that it provides a high-level view of programs—only initial and final states play a rôle.

In contrast, the advantage of a more concrete semantics is that it gets us closer to the implementation of the construct under consideration. On the other hand, a more concrete semantics suffers from being unwieldy to work with when proving programs.
3.0.1 Outline of Chapter

As the more concrete semantics, I choose a trace semantics (cf. [75]). In the rest of this chapter, I take programs that can raise and handle exceptions along the path of Lukkan [59], which describes first an operational semantics in terms of traces and then derives a weakest-precondition semantics from it.

Section 3.1 describes my trace semantics model. In Section 3.2, I start off with a clean slate and define the basic exceptional statements and compositions by their trace semantics. Section 3.3 defines the meaning of weakest preconditions ($wp$) in the trace semantics setting. Section 3.4 calculates $wp$ for the statements defined in Section 3.2. We can then compare the $wp$ in this chapter with the $wp$ from Chapter 1, and will find the two equal. Thus, the definitions in Chapter 1 do match our operational interpretation of the commands.

3.1 Trace sets

For a program without exceptions, the program state space is the Cartesian product of the program variables. For example, the state space of a program with two program variables, $a$ and $b$ say, can be thought of as a two-dimensional space, as depicted in Figure 3.0.

To model exceptional programs, I augment the state space with a binary "outcome" coordinate, $oc$, depicted for the two-variable example in Figure 3.1. Coordinate $oc$ partitions the resulting state space into normal ($oc = \bot$) and exceptional ($oc = \top$).
states. For state \( x \), I write \( \text{nor}.x \) to indicate that the outcome is normal, and \( \text{exc}.x \) to indicate that the outcome is exceptional. I write \( X \) for the set of all states, including those with an exceptional outcome.

The semantics of a program is defined via traces. In this chapter, a trace is a nonempty sequence of states that starts in a normal state; no actions are recorded in the traces. A trace set is a (possibly infinite) set of (possibly infinite) traces. For program \( S \), I identify \( S \) with the set of all traces that can be the result of executing \( S \). Figure 3.2 depicts a sample trace.

Catenation, which binds stronger than function application, will be denoted by juxtaposition. Variables \( s \) and \( t \) range over (possibly empty) sequences of states, and \( x \) and \( y \) over states. I define \( \text{fin}.s \) to hold just when the length of \( s \) is finite, and \( \text{inf}.s \) to hold otherwise. For nonempty sequence \( s \), I let \( \text{first}.s \) denote the first state in \( s \), and, if \( s \) is finite, \( \text{last}.s \) the last state in \( s \). For state \( x \), I write \( [x] \) for \( x[\text{oc} := \bot] \) and \( [x] \) for \( x[\text{oc} := \top] \), that is, \( x \) in which the value of coordinate \( \text{oc} \) has been replaced by \( \bot \) and \( \top \), respectively. Stated differently, \( [x] \) and \( [x] \) are the two states that are the projections of \( x \) onto the normal and exceptional planes, respectively.

Remark 3.0. The given trace semantics models programs with one exception. To model \( n \) distinct exceptions directly in the trace semantics, one can change \( \text{oc} \) from being a binary-valued coordinate to a coordinate that can assume \( 1+n \) values: the normal value plus one for each exception.

3.2 Program constructs as trace sets

In this section, I define each program construct as a trace set.

3.2.0 Primitive statements

By way of introduction, I define the trace semantics of the assignment statement. Let \( v \) be a regular program variable and \( E \) be a total expression (see Section 1.1). Then, in the absence of exceptions, one would write

\[
v := E = \{ x \mapsto x[\!\!\!\!\!\!\!\!\!v := E.x]!\!\!\!\!\!\!\!\!\} \quad ,
\]
in which every trace has length two: it consists of initial state \( x \) and final state \( x[v := E.x] \), that is, state \( x \) in which the value of coordinate \( v \) has been replaced by the value of expression \( E \) evaluated in state \( x \). The set \( v := E \) contains such a trace for every state \( x \in X \). In the presence of exceptions, I restrict \( x \) to be a normal state and write

\[
v := E = \{ x \mid nor.x \triangleright x \ (x[v := E.x]) \} \quad .
\]

Statement \( \text{skip} \) is defined as

\[
\text{skip} = \{ x \mid nor.x \triangleright x \}
\]

(3.1)
in which the latter occurrence of \( x \) denotes a trace of length one. I could have chosen

\[
\text{skip} = \{ x \mid nor.x \triangleright xx \}
\]
to get traces of length two, but, for reasons discussed below, I prefer (3.1).

I write the raising of an exception as the statement \( \text{raise} \), and define its trace semantics as

\[
\text{raise} = \{ x \mid nor.x \triangleright x[x] \}
\]

(3.2)
that is, the set of all traces of length two starting with a normal state and ending with an exceptional state; the two states are equal in all other coordinates. Figure 3.3 shows a sample trace from set \( \text{raise} \). Alternatively, I might have written

\[
\text{raise} = oc := T
\]

except that \( oc \) is not a regular program variable; it is a variable that I have introduced for describing the trace semantics only.

### 3.2.1 Normal Composition

The definition of sequential composition is changed to accommodate exceptional outcomes. If there were no exceptions, one could define

\[
S;T = \{ s,x,t \mid sx \in S \land xt \in T \land fin.s \triangleright sx \} \cup \\
\{ s \mid s \in S \land inf.s \triangleright s \}
\]

which distinguishes between those traces in which execution of \( S \) does or does not terminate. In words, \( S;T \) contain the set of traces that start with a finite trace from \( S \ (sx) \) and continue with a trace from \( T \ (xt) \), where the last state in the trace from \( S \ (x) \) is the same as the first state of \( T \ (x) \). \( S;T \) also contains the infinite traces in \( S \).

In the presence of exceptions, I refine the definition to

\[
S;T = \{ s,x,t \mid sx \in S \land xt \in T \land fin.s \land nor.x \triangleright sx \} \cup \\
\{ s \mid s \in S \land (inf.s \lor exc.(last.s)) \triangleright s \}
\]

(3.3)
What distinguishes this definition from the previous one is that the connecting state \((x)\) is restricted to be a normal state. The infinite traces of \(S\) and the finite traces of \(S\) that end with an exceptional state are not joined with any trace from \(T\). This definition captures the fact that execution of \(S;T\) reduces to execution of \(S\) in the case where that execution terminates exceptionally.

We have

**Theorem**

\[
; \text{ is associative.} \tag{3.4}
\]

I omit the proof of this theorem.

**Theorem**

\(skip\) is the left identity of \(;\). \tag{3.5}

Proof.

\[
skip;T
= \{ (3.3): \text{def. of } ; \}
\{ s, x, t \mid sx \in skip \land xt \in T \land fin.s \land nor.x \triangleright sx \} \cup
\{ s \mid s \in skip \land (inf.s \lor exc.(last.s)) \triangleright s \}
= \{ (3.1): \text{def. of } skip \}
\{ x, t \mid xt \in T \land nor.x \triangleright xt \}
= \{ \text{each trace starts in a normal state } \}
\]

\(T\)

\[
\]

**Theorem**

\(skip\) is the right identity of \(;\). \tag{3.6}

Proof.

\[
S;skip
= \{ (3.3): \text{def. of } ; \}
\{ s, x, t \mid sx \in S \land xt \in skip \land fin.s \land nor.x \triangleright sx \} \cup
\{ s \mid s \in S \land (inf.s \lor exc.(last.s)) \triangleright s \}
= \{ (3.1): \text{def. of } skip \}
\{ s, x \mid sx \in S \land fin.s \land nor.x \triangleright sx \} \cup
\{ s \mid s \in S \land (inf.s \lor exc.(last.s)) \triangleright s \}
= \{ \text{fin and inf are each other's complements, and ditto for nor and exc } \}
\{ s \mid s \in S \triangleright s \}
= S
\]

\[
\]

40
For the above two theorems to hold, it is essential that skip does not duplicate the state in a trace when joined by a semicolon with another statement. This is why the trace set of skip contains traces of length one instead of traces of length two.

In the next theorem, I consider ; applied to arbitrary trace sets.

**Theorem**

; is positively ∪-distributive in both arguments. \( (3.7) \)

**Remark 3.1.** In its left argument, ; is even universally ∪-distributive, as the proof shows, but that property is not needed in the present discussion.

**Proof.** With \( S \) ranging over any bag of trace sets, we calculate,

\[
\langle \cup S \triangleright S \rangle ; T \\
= \{ \text{ (3.3): def. of } ; \} \\
\{ s, x, t | s x \in (\cup S \triangleright S) \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \cup \\
\{ s | s \in (\cup S \triangleright S) \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ interchange unions } \} \\
\{ S, s, x, t | s x \in S \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \cup \\
\{ S, s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ nesting } \} \\
\langle \cup S \triangleright \{ s, x, t | s x \in S \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \rangle \cup \\
\langle \cup S \triangleright \{ s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \rangle \\
= \{ \text{ combine terms } \} \\
\langle \cup S \triangleright \{ s, x, t | s x \in S \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \rangle \cup \\
\{ s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ (3.3): def. of } ; \} \\
\langle \cup S \triangleright S ; T \rangle.
\]

For the other argument, and with \( T \) ranging over any nonempty bag of trace sets, we calculate,

\[
S ; \langle \cup T \triangleright T \rangle \\
= \{ \text{ (3.3): def. of } ; \} \\
\{ s, x, t | s x \in S \land x t \in (\cup T \triangleright T) \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \cup \\
\{ s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ interchange union } \} \\
\{ T, s, x, t | s x \in S \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \cup \\
\{ s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ nesting } \} \\
\langle \cup T \triangleright \{ s, x, t | s x \in S \land x t \in T \land \text{ fin.s } \land \text{ nor.s } \triangleright s x t \} \rangle \cup \\
\{ s | s \in S \land (\text{ inf.s } \lor \text{ exc.(last.s)}) \triangleright s \} \\
= \{ \text{ range of } T \text{ is nonempty } \}
\]
\[ \langle \cup T \rangle \{ s, x, t \mid s x \in S \land x t \in T \land \text{fin} \cdot s \land \text{nor} \cdot s \triangleright s x t \} \cup \{ s \mid s \in S \land (\text{inf} \cdot s \lor \text{exc} \cdot (\text{last} \cdot s)) \triangleright s \} \]

= \{ \text{(3.3): def. of } ; \} \langle \cup T \triangleright S; T \rangle \quad . \]

3.2.2 Exceptional composition

Next, I define the trace semantics of the exception handler.

\[ S \triangleleft T = \{ s, x, t \mid s x \in S \land [x] t \in T \land \text{fin} \cdot s \land \text{exc} \cdot x \triangleright s x [x] t \} \cup \{ s \mid s \in S \land (\text{inf} \cdot s \lor \text{nor} \cdot (\text{last} \cdot s)) \triangleright s \} \quad (3.8) \]

This set is similar to \( S; T \), but has some important differences. Finite traces of \( S \) where the last state is exceptional (\( s x \)) are joined with traces from \( T \) that begin with the normal projection (\([x]\)) of the last state of the trace from \( S \). Moreover, unlike normal composition, no state is dropped here, so both the last state in the trace from \( S \) (\( x \)) and the first state in the trace from \( T \) (\([x]\)) appear in the traces in \( S \triangleleft T \).

Also, infinite traces of \( S \) and finite traces of \( S \) that end with a normal state are included in \( S \triangleleft T \), but are not joined with any trace from \( T \).

Remark 3.2. It would be nice to have

\[ \text{raise} \text{ is the left and right identity of } \triangleleft \quad , \]

but neither part of this property holds, because the traces of \( \text{raise} \) have length two, and therefore add an extra state to the traces of the exception handler. Furthermore, the exception handler, too, repeats the connecting state, \( x \), with \( x := \bot \) in the second occurrence.

3.2.3 Other statements

I omit discussion of the remaining statements; aside from differences noted in this section, their definitions are as in [59].

Compared to [59], the definitions of \( \text{abort} \) and of the if-statement need not be changed because the initial state is always normal. The definition of the do-statement need not be changed because it is defined in terms of the if-statement, sequential composition, and \( \text{skip} \). The latter two have already been redefined to cater for exceptional states, and the only properties used in the context of the do-statement are that sequential composition is associative (which it still is—see (3.4)), that \( \text{skip} \) is its identity element (which it still is—see (3.5) and (3.6)), and that sequential composition is positively \( \cup \)-distributive in both arguments (which it still is—see (3.7)). For the purpose of defining the do-statement, [59] introduces a partial command \( g? \), which can be seen as the command \( g \rightarrow \text{skip} \). However, this command is not recognized as a partial command in [59].
3.3 Weakest preconditions of trace sets

I define function \( wp.S.(P, Q) \) to be the weakest condition on the initial state such that: execution of program \( S \) terminates, every exceptional terminating state satisfies \( P \), and every normal terminating state satisfies \( Q \). Since the \( ac \) coordinate is not part of the program but of the trace semantics only, I require that \( P \), \( Q \), and \( wp.S.(P, Q) \) be independent of the \( ac \) coordinate, that is, \( Q_x = Q_x \). I do so by restricting \( P \) and \( Q \) to predicates in which \( ac \) does not occur, and by designing \( wp \) carefully (see (3.11) below). As a result, \( wp.S.(false, Q) \) in this chapter coincides with Dijkstra's \( wp.S.Q \) [17].

In the sequel, I often need to distinguish between conditions on the exceptional and on the normal states. I write a pair of conditions to capture this distinction.

\[
(P, Q)_x = (exc.x \Rightarrow P_x) \land (nor.x \Rightarrow Q_x) \tag{3.9}
\]

**Remark 3.3.** In [61], the generalization of such a pair to any tuple is called a *partition predicate*.

This construction is universally conjunctive, that is,

\[
\langle \forall i \triangleright (P_i, Q_i) \rangle = \langle \langle \forall i \triangleright P_i \rangle, \langle \forall i \triangleright Q_i \rangle \rangle \tag{3.10}
\]

as is shown by the calculation

\[
\langle \forall i \triangleright (P_i, Q_i) \rangle_x
\]

\[
= \{ \text{ lifting } \}
\]

\[
\langle \forall i \triangleright (P_i, Q_i) \rangle_x
\]

\[
= \{ \text{ (3.9): def. of a pair predicate } \}
\]

\[
\langle \forall i \triangleright (exc.x \Rightarrow P_i.x) \land (nor.x \Rightarrow Q_i.x) \rangle
\]

\[
= \{ \text{ pred. calc. } \}
\]

\[
(exc.x \Rightarrow \langle \forall i \triangleright P_i.x \rangle) \land (nor.x \Rightarrow \langle \forall i \triangleright Q_i \rangle.x)
\]

\[
= \{ \text{ lifting } \}
\]

\[
(exc.x \Rightarrow \langle \forall i \triangleright P_i \rangle.x) \land (nor.x \Rightarrow \langle \forall i \triangleright Q_i \rangle.x)
\]

\[
= \{ \text{ (3.9): def. of a pair predicate } \}
\]

\[
\langle \forall i \triangleright P_i \rangle, \langle \forall i \triangleright Q_i \rangle \rangle_x
\]

The definition of \( wp.S.(P, Q) \) for state \( x \) is a condition on \( x \) such that every trace \( t \) of \( S \) that begins with initial state \( x \) is of finite length and satisfies \((P, Q).(last.t)\).

\[
wp.S.(P, Q)_x = \langle \forall t \mid \text{ first.t} = [x] \land t \in S \triangleright \text{ fin.t} \land (P, Q).(last.t) \rangle \tag{3.11}
\]

Above, I said that this definition needed to be designed with care. The trick is to define \( wp.S.(P, Q)_x \) for *any* \( x \), not just for normal states \( x \), despite the fact that a trace always begins with a normal state. Since \( x \) can be any state, \( [x] \) is needed in the definition (see, e.g., the second step in the calculation leading to (3.18)).

I continue with some theorems regarding \( wp \).
Theorem

\( \text{wp}.S \) is positively conjunctive. \hfill (3.12)

Proof. With \( K \) ranging over any nonempty bag of pairs, we calculate,

\[
\begin{align*}
\text{wp}.S.\{ \forall K \triangleright K \}.x \\
= \quad \{ \text{(3.11): def. of wp} \} \\
\langle \forall t \mid \text{first} \cdot t = [x] \land t \in S \triangleright \text{fin} \cdot t \land \langle \forall K \triangleright K.\text{last} \cdot t \rangle \} \\
= \quad \{ \text{lifting} \} \\
\langle \forall t \mid \text{first} \cdot t = [x] \land t \in S \triangleright \text{fin} \cdot t \land \langle \forall K \triangleright K.\text{last} \cdot t \rangle \} \\
= \quad \{ \land \text{ over } \forall, \text{ since range is nonempty} \} \\
\langle \forall t \mid \text{first} \cdot t = [x] \land t \in S \triangleright \langle \forall K \triangleright \text{fin} \cdot t \land K.\text{last} \cdot t \rangle \} \\
= \quad \{ \text{interchange of quantifications} \} \\
\langle \forall K \triangleright \langle \forall t \mid \text{first} \cdot t = [x] \land t \in S \triangleright \text{fin} \cdot t \land K.\text{last} \cdot t \rangle \} \\
= \quad \{ \text{(3.11): def. of wp} \} \\
\langle \forall K \triangleright \text{wp}.S.K.x \} \\
= \quad \{ \text{lifting} \} \\
\langle \forall K \triangleright \text{wp}.S.K \}.x \\
\square
\]

A consequence of this theorem (cf. [21]) is

\( \text{wp}.S \) is monotonic. \hfill (3.13)

Theorem

\[
\text{wp}.S.(P,Q) = \text{wp}.S.(P,\text{true}) \land \text{wp}.S.(\text{true},Q) \hfill (3.14)
\]

Proof. Follows from (3.10) and (3.12). \square

Using the "everywhere" operator, written \([.\] (cf. [21]), I can state the next theorems.

Theorem

\[
[R \Rightarrow \text{wp}.S.(P,Q)] = [R \Rightarrow \text{wp}.S.(P,\text{true})] \land [R \Rightarrow \text{wp}.S.(\text{true},Q)] \hfill (3.15)
\]

Proof. Follows from (3.14) and predicate calculus. \square

Theorem

\[
\begin{align*}
[R \Rightarrow \text{wp}.S.(P,Q)] &= \langle \exists M \triangleright [R \Rightarrow \text{wp}.S.(M,Q)] \land [M \Rightarrow P] \rangle \\
[R \Rightarrow \text{wp}.S.(P,Q)] &= \langle \exists M \triangleright [R \Rightarrow \text{wp}.S.(P,M)] \land [M \Rightarrow Q] \rangle \\
\end{align*}
\hfill (3.16)
\]

\[
\hfill (3.17)
\]
Proof.

\[ [R \Rightarrow wp.S.(P,Q)] \]
\[ = \{ \text{one-point rule} \} \]
\[ (\exists M | [M \equiv P] \triangleright [R \Rightarrow wp.S.(M,Q)]) \]
\[ \Rightarrow \{ \text{weakening} \} \]
\[ (\exists M \triangleright [M \Rightarrow P] \land [R \Rightarrow wp.S.(M,Q)]) \]
\[ \Rightarrow \{ (3.13): wp.S \text{ is monotonic} \} \]
\[ (\exists M \triangleright [R \Rightarrow wp.S.(P,Q)]) \]
\[ = \{ \text{range is nonempty} \} \]
\[ [R \Rightarrow wp.S.(P,Q)] \]

The proof of the other is similar. \qed

3.4 Calculating the weakest preconditions

I calculate \( wp \) for the various program constructs, arriving at a link with the \( wp \) definitions of Chapter 1. First, I look at \( \text{skip} \). For any state \( x \),

\[ wpskip.(P,Q).x \]
\[ = \{ (3.11): \text{def. of } wp \} \]
\[ (\forall t | \text{first} \cdot t = [x] \land t \in \text{skip} \triangleright \text{fin} \cdot t \land (P,Q).(last \cdot t)) \]
\[ = \{ (3.1): \text{def. of } \text{skip} \} \]
\[ (P,Q).[x] \]
\[ = \{ \text{nor}.[x], \text{and (3.9): def. of } (P,Q) \} \]
\[ Q.[x] \]
\[ = \{ Q \text{ is independent of } \omega \} \]
\[ Q.x \]

and hence

\[ wpskip.(P,Q) = Q \quad . \] (3.18)

By a similar calculation, we obtain, for any state \( x \),

\[ wp\text{raise}.(P,Q).x \]
\[ = \{ (3.11): \text{def. of } wp \} \]
\[ (\forall t | \text{first} \cdot t = [x] \land t \in \text{raise} \triangleright \text{fin} \cdot t \land (P,Q).(last \cdot t)) \]
\[ = \{ (3.2): \text{def. of } \text{raise} \} \]
\[ (P,Q).[x] \]
\[ = \{ [x] = [x], \text{and exc}.[x], \text{and (3.9): def. of } (P,Q) \} \]
\[ P.[x] \]
\[ = \{ P \text{ is independent of } \omega \} \]
\[ P.x \]

45
and hence
\[
wp.\text{raise}(P, Q) = P
\] (3.19)

Similarly,
\[
w_p.(v := E).(P, Q) = Q[v := E]
\] (3.20)
can be shown.

More involved are the calculations for sequential and exceptional composition. For the latter, we calculate,
\[
w_p.(S \langle T).(P, Q).x
\]
\[
= \begin{cases}
(3.11): \text{def. of wp} & \\
\{ \forall t \mid \text{first}\_t = [x] \land t \in S \langle T \triangleright \text{fin}\_t \land (P, Q).\text{(last}\_t) \} & \\
\{ (3.8): \text{def. of }< t, \text{ with } t := sy[y]\_t \text{ and } t := s \} & \\
\forall s, y, t \mid \text{first}\_y[t] = [x] \land sy \in S \land [y]\_t \in T \land \text{fin}\_s \land \text{exc}\_y \triangleright \\
\text{fin}\_y[y]\_t \land (P, Q).\text{(last}\_y[y]\_t) \} & \\
\forall s \mid \text{first}\_s = [x] \land s \in S \land (\text{inf}\_s \lor \text{nor}(\text{last}\_s)) \triangleright \text{fin}\_s \land (P, Q).\text{(last}\_s) \} & \\
\{ \text{rename } s \text{ and } t \text{ as } sy := s \text{ and } [y]\_t := t \text{ in first quantification} \} & \\
\forall s, t \mid \text{first}\_s = [x] \land s \in S \land \text{first}\_t = [\text{last}\_s] \land t \in T \land \text{fin}\_s \land \text{exc}(\text{last}\_s) \triangleright \\
\text{fin}\_s \land (P, Q).\text{(last}\_t) \} & \\
\forall s \mid \text{first}\_s = [x] \land s \in S \land (\text{inf}\_s \lor \text{nor}(\text{last}\_s)) \triangleright \text{fin}\_s \land (P, Q).\text{(last}\_s) \} & \\
\{ \text{nesting, and } \text{fin}\_s \Rightarrow (\text{fin}\_s = \text{fin}\_t) \} & \\
\forall s \mid \text{first}\_s = [x] \land s \in S \land \text{fin}\_s \land \text{exc}(\text{last}\_s) \triangleright \\
\{ \forall t \mid \text{first}\_t = [\text{last}\_s] \land t \in T \triangleright \text{fin}\_t \land (P, Q).\text{(last}\_t) \} \} & \\
\forall s \mid \text{first}\_s = [x] \land s \in S \land (\text{inf}\_s \lor \text{nor}(\text{last}\_s)) \triangleright \text{fin}\_s \land (P, Q).\text{(last}\_s) \} & \\
\{ \text{shunting twice, and (3.11): def. of wp} \} & \\
\forall s \mid \text{first}\_s = [x] \land s \in S \triangleright \text{inf}\_s \lor \text{nor}(\text{last}\_s) \lor wp.\text{T}(P, Q).\text{(last}\_s) \} \land \\
\forall s \mid \text{first}\_s = [x] \land s \in S \triangleright (\text{fin}\_s \land \text{exc}(\text{last}\_s)) \lor (\text{fin}\_s \land (P, Q).\text{(last}\_s)) \}
\]

Leaving the range first\_s = [x] \land s \in S as understood, we continue the calculation,
\[
\forall s \triangleright \text{inf}\_s \lor \text{nor}(\text{last}\_s) \lor wp.\text{T}(P, Q).\text{(last}\_s) \} \land \\
\forall s \triangleright (\text{fin}\_s \land \text{exc}(\text{last}\_s)) \lor (\text{fin}\_s \land (P, Q).\text{(last}\_s)) \}
\]
\[
= \{ \text{combine terms, and factor} \}
\]
\[
\forall s \triangleright (\text{inf}\_s \lor \text{nor}(\text{last}\_s) \lor wp.\text{T}(P, Q).\text{(last}\_s)) \land \\
\text{fin}\_s \land (\text{exc}(\text{last}\_s) \lor (P, Q).\text{(last}\_s)) \}
\]
\[
= \{ \text{absorption, and pred. calc.} \}
\]
\[
\forall s \triangleright \text{fin}\_s \land (\text{exc}(\text{last}\_s) \Rightarrow wp.\text{T}(P, Q).\text{(last}\_s)) \land \\
(\text{nor}(\text{last}\_s) \Rightarrow (P, Q).\text{(last}\_s)) \}
\]
\[
= \{ \text{nor}(\text{last}\_s) \Rightarrow ((P, Q).\text{(last}\_s) = Q.\text{(last}\_s)) \} 
\]
46
\[ \langle \forall s \triangleright \mathit{fin.s} \land (\mathit{exc.}(\mathit{last.s}) \Rightarrow \mathit{wp.T.}(P,Q).((\mathit{last.s})) \land (\mathit{nor.}(\mathit{last.s}) \Rightarrow Q.(\mathit{last.s}))) \rangle \]

\[ = \quad \{ \text{(3.9): def. of a pair} \} \]

\[ \langle \forall s \triangleright \mathit{fin.s} \land (\mathit{wp.T.}(P,Q),Q).(\mathit{last.s}) \rangle \]

\[ = \quad \{ \text{(3.11): def. of wp (recall the understood range)} \} \]

\[ \mathit{wp.S.}(\mathit{wp.T.}(P,Q),Q).x \quad , \]

and obtain

\[ \mathit{wp.}(S \triangleleft T).(P,Q) = \mathit{wp.S.}(\mathit{wp.T.}(P,Q),Q) \quad . \quad (3.21) \]

For sequential composition, we can show

\[ \mathit{wp.}(S;T).(P,Q) = \mathit{wp.S.}(P,\mathit{wp.T.}(P,Q)) \quad , \quad (3.22) \]

to be equal. Hence, I have justified the definitions of \( \mathit{wp} \) for these constructs.

\textbf{Remark 3.4.} With reference to Remark 3.2, note that in the weakest-precondition semantics, \( \mathit{raise} \) is an identity of \( \triangleleft \). We saw this not to be the case in the trace semantics because of repeated states. This comes back to the fact that the trace semantics is more concrete than the \( \mathit{wp} \) semantics, which only concerns itself with the first and last states of traces.
Chapter 4

A theorem on programming methodology

From the weakest-precondition semantics given for a programming notation, one often derives some theorems that are used in reasoning about programs. Ideally, they suggest hints for methodical program construction. An example of that is the Invariance Theorem for iterative statements (cf. [21] and Section 1.6). I give a theorem that suggests using exception handlers in a way similar to split binary semaphores (cf. [63]). The theorem is in terms of Hoare triples for normal termination, and free occurrences of raise for exceptional termination.

4.0 Hoare triples

Inspired by [36], I define the Hoare triple of exceptional programs by

$$\{ R \} S \{ Q \} = [R \Rightarrow wp.S.(true, Q)]$$

(4.0)

from which the Hoare triple for each of the basic statements can be calculated.

$$\{ R \} v := E \{ Q \} = [R \Rightarrow Q[v := E]]$$

$$\{ R \} \text{ skip } \{ Q \} = [R \Rightarrow Q]$$

$$\{ R \} \text{ raise } \{ Q \} = \text{ true}$$

$$\{ R \} S0; S1 \{ Q \} = \langle \exists M : [R \Rightarrow S0 \{ M \} \wedge \{ M \} S1 \{ Q \}] \rangle$$

(4.1)

$$\{ R \} S0 < S1 \{ Q \} = [R \Rightarrow S0 \{ Q \} \wedge \langle \exists M : [R \Rightarrow wp.S0.(M, true)] \wedge \{ M \} S1 \{ Q \} \rangle]$$

(4.2)

I prove the last two of these. For sequential composition,

$$\{ R \} S0; S1 \{ Q \}$$

$$= \{ (4.0): \text{ def. of triple } \}$$

$$[R \Rightarrow wp.(S0; S1).(true, Q)]$$

$$= \{ (1.3): wp \ of ; \}$$

$$[R \Rightarrow wp.S0.(true, wp.S1.(true, Q))]$$

$$= \{ (3.17) \text{ with } S, P, Q := S0, true, wp.S1.(true, Q) \}$$
\[ \exists M \triangleright [R \Rightarrow wp.S0.(true,M)] \land [M \Rightarrow wp.S1.(true,Q)] \]
\[ = \{ \text{(4.0): def. of triple, twice} \}
\{ R \triangleright \{R\} S0\{M\} \land \{M\} S1\{Q\} \} , \]
and for exceptional composition,
\[ \{R\} S0 \triangleleft S1\{Q\} \]
\[ = \{ \text{(4.0): def. of triple} \}
[R \Rightarrow wp.(S0 \triangleleft S1).(true,Q)] \]
\[ = \{ \text{(1.4): wp of } \triangleleft \}
[R \Rightarrow wp.S0.(wp.S1.(true,Q),Q)] \]
\[ = \{ \text{(3.16) with } S, P := S0, wp.S1.(true,Q) \}
\exists M \triangleright [R \Rightarrow wp.S0.(M,Q)] \land [M \Rightarrow wp.S1.(true,Q)] \]
\[ = \{ \text{(3.15) with } P := M \}
\exists M \triangleright [R \Rightarrow wp.S0.(M,\text{true})] \land [R \Rightarrow wp.S0.(true,Q)] \land [M \Rightarrow wp.S1.(true,Q)] \]
\[ = \{ \text{(4.0): def. of triple, twice} \}
\exists M \triangleright [R \Rightarrow wp.S0.(M,\text{true})]\land \{R\} S0\{Q\} \land \{M\} S1\{Q\} \}
\[ = \{ \land \text{ distributes over } \exists \}
\{R\} S0\{Q\} \land (\exists M \triangleright [R \Rightarrow wp.S0.(M,\text{true})] \land \{M\} S1\{Q\} \} . \]

4.1 Free occurrences of \textit{raise}

I now give the definition of \textit{free occurrence of raise}. Statement \textit{raise} occurs free in

- \textit{raise}

- \textit{S0; S1} just when it occurs free in \textit{S0} or in \textit{S1}

- \textit{S0 \triangleleft S1} just when it occurs free in \textit{S1}

- the other statements when it occurs free in one or more of their constituent statements.

If every execution of a free occurrence of \textit{raise} in a statement \textit{S} starts in a state satisfying a predicate \textit{P} whenever the execution of \textit{S} starts in a state satisfying the predicate \textit{R}, then I say "every free occurrence of \textit{raise} in \textit{S} has precondition \textit{P} in context \textit{R}". The next theorem establishes the correspondence between this informal statement and a mathematical formula.

\textbf{Theorem}

Every free (occurrence of) \textit{raise} in \textit{S} has precondition \textit{P} in context \textit{R}

\[ [R \Rightarrow wp.S.(P,\text{true})] \] (4.3)
Proof. The proof is by induction over the syntax of $S$. For assignment, we calculate,

every free $\textit{raise}$ in $\nu := E$ has precondition $P$ in context $R$

\[
= \{ \text{there are no free occurrences of $\textit{raise}$ in $\nu := E$} \}
\]

\[\text{true}\]

\[
= \{ (1.0): \text{wp of $\nu := E$; pred. calc.} \}
\]

\[ [R \Rightarrow \text{wp.} (\nu := E). (P, \text{true})] \quad . \]

The proof for $\textit{skip}$ is similar. For $\textit{raise}$,

\[
\text{every free $\textit{raise}$ in $\textit{raise}$ has precondition $P$ in context $R$}
\]

\[
= \{ \text{raise is a free occurrence of $\textit{raise}$} \}
\]

\[ [R \Rightarrow P] \]

\[
= \{ (1.2): \text{wp of $\textit{raise}$} \}
\]

\[ [R \Rightarrow \text{wp.} \textit{raise}. (P, \text{true})] \quad . \]

Now the compositions, beginning with normal,

\[
\text{every free $\textit{raise}$ in $S_0; S_1$ has precondition $P$ in context $R$}
\]

\[
= \{ \text{def. of free occurrences of $\textit{raise}$ in $S_0; S_1$, and}
\]

\[
\quad \text{notion of “context” for $;$} \quad \}
\]

\[
\text{every free $\textit{raise}$ in $S_0$ has precondition $P$ in context $R$, and}
\]

\[
\text{every free $\textit{raise}$ in $S_1$ has precondition $P$ in context $M$}
\]

\[
\quad \text{for some $M$ satisfying $\{R\} S_0 \{M\}$}
\]

\[
= \{ \text{induction hypothesis, twice} \}
\]

\[ [R \Rightarrow \text{wp.} S_0. (P, \text{true})] \land
\]

\[
\langle \exists M \triangleright [M \Rightarrow \text{wp.} S_1. (P, \text{true})] \land \{R\} S_0 \{M\} \rangle
\]

\[
= \{ (4.0): \text{def. of triple} \}
\]

\[ [R \Rightarrow \text{wp.} S_0. (P, \text{true})] \land \langle \exists M \triangleright [M \Rightarrow \text{wp.} S_1. (P, \text{true})] \land [R \Rightarrow \text{wp.} S_0. (\text{true}, M)] \rangle
\]

\[
= \{ \quad \land \quad \text{over $\exists$} \quad \}
\]

\[
\langle \exists M \triangleright [M \Rightarrow \text{wp.} S_1. (P, \text{true})] \land [R \Rightarrow \text{wp.} S_0. (P, \text{true})] \land [R \Rightarrow \text{wp.} S_0. (\text{true}, M)] \rangle
\]

\[
= \{ (3.15) \text{ with } Q := M \quad \}
\]

\[
\langle \exists M \triangleright [R \Rightarrow \text{wp.} S_0. (P, M)] \land [M \Rightarrow \text{wp.} S_1. (P, \text{true})] \rangle
\]

\[
= \{ (3.17) \text{ with } S, Q := S_0, \text{wp.} S_1. (P, \text{true}) \quad \}
\]

\[ [R \Rightarrow \text{wp.} S_0. (P, wp. S_1. (P, \text{true}))]
\]

\[
= \{ (1.3): \text{wp of $;$} \quad \}
\]

\[ [R \Rightarrow \text{wp.} (S_0; S_1). (P, \text{true})] \quad . \]

Finally, for exceptional composition,

\[
\text{every free $\textit{raise}$ in $S_0 < S_1$ has precondition $P$ in context $R$}
\]

\[
= \{ \text{def. of free occurrences of $\textit{raise}$ in $S_0 < S_1$, and}
\]

\[
\quad \text{notion of “context” for $<$} \quad \}
\]

\[
\text{every free $\textit{raise}$ in $S_1$ has precondition $P$ in the context of}
\]

\[
\quad \text{some exceptional postcondition of $S_0$ from $R$}
\]

\[
= \{ \text{induction hypothesis} \quad \}
\]
\[ \exists M \triangleright [M \Rightarrow \text{wp}.S1.(P,\text{true})] \land [R \Rightarrow \text{wp}.S0.(M,\text{true})] \]
\[ = \{ \text{(3.16) with } S, P, Q := S0, \text{wp}.S1.(P,\text{true}), \text{true} \} \]
\[ [R \Rightarrow \text{wp}.S0.(\text{wp}.S1.(P,\text{true}),\text{true})] \]
\[ = \{ \text{(1.4): wp of } \triangleleft \} \]
\[ [R \Rightarrow \text{wp}.(S0 \triangleleft S1).(P,\text{true})] \]  \hspace{1cm} \Box

4.2 Usage of exceptions

As a consequence of (4.2) and (4.3), we have

**Theorem**

\[ \{R\} S0 \triangleleft S1 \{Q\} \]
\[ = \]
\[ \{R\} S0 \{Q\} \land \]
\[ \{ \exists M \triangleright \} M \} S1 \{Q\} \land \]
\[ \text{every free occurrence of } raise \text{ in } S0 \text{ has precondition } M \text{ in context } R \}

This theorem suggests that, when constructing $S0$ in $S0 \triangleleft S1$, one should have in mind a precondition $M$ for all $raise$ statements that occur free in $S0$; the benefit is that one may assume that same precondition in the construction of $S1$.

In the next chapter, I put this theorem into action in the construction of a simple program. The theorem is also quite useful in large programs, because it shapes the way we think about using exceptions: If a particular exception is raised only when some particular global condition holds, then the theorem tells us that every handler of that exception can assume that condition upon entry. This suggests that the declaration of an exception in a programming notation provide a way to state this condition.

Finally, I remark on exception parameters, as featured, e.g., in Modula-3 and CLU [54]. The values of these parameters can be considered part of the “global condition” to which I alluded above, since they are accessible both where the exception is raised and where it is handled. Thus, the theorem suggests that these parameters be initialized, at the time the exception is being raised, to satisfy some particular condition. Like the case for other global values, the exception handler can then depend on that condition upon entry.
Chapter 5

Constructing a program with exceptions

In Chapter 4, I developed a theorem regarding the use of exceptions. In this chapter, I employ that theorem in a novel construction of a simple program.

5.0 A program derivation

Equipped with exceptions, the task in this section is to design a program that, given value $x$ and two-dimensional array $a$ of size $M \times N$, computes $b, i, j$ to satisfy

$$Q : \quad (b \Rightarrow Q0) \land (\neg b \Rightarrow Q1),$$

where

$$Q0 : \quad 0 \leq i < M \land 0 \leq j < N \land a[i, j] = x$$
$$Q1 : \quad \langle \forall m, n | 0 \leq m < M \land 0 \leq n < N \Rightarrow a[m, n] \neq x \rangle.$$

It is understood that the values of $x$ and $a$ may not be changed. As our guide, we will use the theorem from Chapter 4, which states that if, in a proof, every free occurrence of $\text{raise}$ in $S$ has precondition $K$, then $K$ can be used as the precondition for handler $T$ in a proof of $S \preceq T$.

To get us started, I suggest we make use of the fact that control sometimes flows through $T$ and sometimes not. I let these two cases correspond to the cases $b$ and $\neg b$. But which goes with which? In order to conclude that $x$ is not present in $a$, every element of $a$ needs to be tested. This can be done using two nested loops. In order to set $i$ and $j$ to a coordinate of $a$ whose value is $x$, the program first needs to find such a coordinate. This, too, can be done using two nested loops, but there is no reason to continue the search once an $x$ has been encountered. Exceptions provide a means of breaking out of such loops prematurely. For this reason, I write the first approximation of our program as

$$(S; b := \text{false}) \preceq b := \text{true}.$$
for some $S$ to be developed.

Let us now develop program $S$, whose normal postcondition we want to be $Q1$. Moreover, the theorem from Chapter 4 tells us that any raise statement in $S$ must have precondition $Q0$. Twice using the well-known technique of replacing a constant by a variable (see, e.g., [17], [29, Ch. 16], [20], [77, Ch. 8], or [5, Ch. 4]), we find invariants

\[
P0 : \quad 0 \leq i \leq M \land (\forall m, n \mid 0 \leq m < i \land 0 \leq n < N \rightarrow a[m, n] \neq x) \\
P1 : \quad P0 \land i \neq M \land 0 \leq j \leq N \land (\forall n \mid 0 \leq n < j \rightarrow a[i, n] \neq x)
\]

for the outer and inner loops, respectively, and calculate $S$ as

\[
i := 0 ; \\
\text{do } i \neq M \rightarrow j := 0 ; \\
\quad \text{do } j \neq N \rightarrow "\text{establish } a[i, j] \neq x" ; j := j + 1 \text{ od } ; \\
\quad i := i + 1 \\
\text{od .}
\]

The program segment "establish $a[i, j] \neq x" concerns us since we are not allowed to modify any of these variables at this stage. Were it not for the presence of exceptions, we would need a miracle at this time. But, since we do have exceptions at our disposal, we just need to verify that one can be raised if $a[i, j] \neq x$ does not hold. We observe that $P1 \land j \neq N \land a[i, j] = x$ implies $Q0$, so we can use a raise here. Replacing "establish $a[i, j] \neq x" with

\[
\text{if } a[i, j] = x \rightarrow \text{raise } \sqcup a[i, j] \neq x \rightarrow \text{skip fi ,}
\]

we are done, and write the entire program as

\[
(\quad i := 0 ; \\
\quad \text{do } i \neq M \rightarrow j := 0 ; \\
\quad \quad \text{do } j \neq N \rightarrow \\
\quad \quad \quad \text{if } a[i, j] = x \rightarrow \text{raise } \sqcup a[i, j] \neq x \rightarrow \text{skip fi } ; \\
\quad \quad \quad j := j + 1 \\
\quad \quad \text{od } ; \\
\quad \quad i := i + 1 \\
\quad \text{od } ; \\
\quad b := \text{false} \\
\quad ) \\
\langle \\
\quad b := \text{true} \\
\quad .
\]

5.1 Discussion

To prove the correctness of our program, we only needed to show that normal termination of the loops maintains the invariants. The rest follows from the invariants, the guards, and the theorem from Chapter 4. The proof of this program is thus simple.
Looking back at the program through operational spectacles, we see the loops followed by \( b := \text{false} \) as trying to establish the absence of \( x \) in \( a \). However, should an \( x \) be present, an exception is raised when an \( x \) is first encountered. The operation of this program is thus easy to understand.

Finally, consider a similar program that, without exceptions, uses stronger guards to facilitate exiting the loops before \( i = M \) and \( j = N \), respectively. For example, twice applying the Bounded Linear Search Theorem [17, 20], we arrive at the program

\[
\begin{align*}
b, m & := \text{false}, 0 ; \\
\text{do } & \neg b \land m \neq M \rightarrow \\
& c, n := \text{false}, 0 ; \\
\text{do } & \neg c \land n \neq N \rightarrow \\
& c, i, j, n := a[m, n] = x, m, n, n + 1 \\
\text{od} ; \\
b, m & := c, m + 1 \\
\text{od}
\end{align*}
\]

In addition to having more complicated invariants (because the invariants need to record the information to prove the postcondition for both conjuncts of \( Q \)), this program is arguably less efficient than the one that uses exceptions (because of the extra tests). Thus, we consider our program efficient. The point, however, is not to show that a structured jump can produce a more efficient program; the point is that we have an easy way of constructing such a program hand in hand with its proof.

The heated forum discussion [73] discusses programs that attempt to solve a related programming problem. The above problem can also efficiently be solved using recursion or \texttt{goto} statements, as [34] demonstrates with a nice derivation. How exception handling can simplify the structure of certain programs is also discussed in, for example, [13]. More recently, [46] shows exceptions in the construction of programs through refinements.
Chapter 6

Modeling common programming languages

In this chapter, I discuss the relation between the program semantics given in Chapter 1 and the mathematical modeling of popular features found in common programming languages like Modula-3, Ada, or C. I discuss procedures, conditional statements, checked run-time errors, and expressions.

6.0 Procedures

Imperative programming languages provide a mechanism to encapsulate and reuse code, viz., procedures. Most languages allow a procedure's declaration (or prototype) to be given separately from its implementation (or body). The purpose of the declaration is to convey the information needed to call the procedure; this information also circumscribes the implementation. In Modula-3, Ada, and C, the declaration gives the signature of the procedure, i.e., a declaration of its parameters, result values, and set of exceptions that it may raise. The separation between declaration and implementation is important when verifying programs. However, the signature does not suffice for this purpose; the declaration must also give a specification of the procedure. Calls to the procedure get their semantics from this specification, and the implementation needs to be verified to meet the specification.

6.0.0 Procedure specifications

A simple procedure is declared and specified by

\[
\text{spec } P() \text{ is spec ,}
\]
where \( P \) names the new procedure and \( \text{spec} \) is its specification. The specification is Larch-like [33] and consists of a set of clauses of the forms

\[
\begin{align*}
\text{modifies } w \\
\text{requires } \text{Pre} \\
\text{ensures } n\text{Post} \\
\text{except-ensures } e\text{Post}
\end{align*}
\]

(6.0)

where \( w \) is a list of variables, and \( \text{Pre}, n\text{Post}, e\text{Post} \) are predicates. \( n\text{Post} \) and \( e\text{Post} \) are two-state predicates, and thus may mention initial-value variables (see Section 1.7). The specification may contain any number of the clauses in any order. This is equivalent to listing all \( w \)’s in one \text{modifies} clause and taking the conjunction of \( \text{Pre} \)’s, \( n\text{Post} \)’s, and \( e\text{Post} \)’s, respectively, for the other clauses. For example, the specification

\[
\begin{align*}
\text{modifies } x \quad \text{requires } 0 \leq x \quad \text{modifies } y \\
\text{except-ensures } y_0 < y \quad \text{requires } x + y = 10
\end{align*}
\]

is the same as

\[
\begin{align*}
\text{modifies } x, y \quad \text{requires } 0 \leq x \land x + y = 10 \\
\text{ensures } \text{true} \quad \text{except-ensures } y_0 < y
\end{align*}
\]

The one exception to the given rule is that the absence of \text{except-ensures} clauses is treated as \text{except-ensures} false rather than \text{except-ensures} true. This way, a specification must explicitly advertise the fact that it may have an exceptional outcome. (Penelope [62] provides some convenient shorthands for writing specifications of exceptional behaviors.)

The meaning of specification (6.0) is that of the specification statement

\[
\begin{align*}
w : [\text{Pre}, (e\text{Post}, n\text{Post})]
\end{align*}
\]

(6.1)

(see Section 1.7).

### 6.0.1 Parameters and Result Values

Procedures can take parameters and can return result values. The specification

\[
\begin{align*}
\text{spec } r := P(x) \text{ is spec}
\end{align*}
\]

(6.2)

declares and specifies a procedure \( P \). \( P \) takes a list of parameters, here named \( x \), and returns a list of result values, here named \( r \). \( x \) and \( r \) are names that may be mentioned in \( \text{spec} \). \( x \) may be read (used) by the procedure but not written (updated). Thus, \( x \) may not appear in the \text{modifies} list in \( \text{spec} \). Each procedure call of \( P \) (described below) instantiates \( x \) with some value. \( r \), on the other hand, may be written by \( P \), but its initial value is unspecified. Consequently, the \text{modifies} list in \( \text{spec} \) is treated as always containing \( r \), and \( r \) may not appear initial-valued in the postconditions in \( \text{spec} \).

The described behavior of \( x \) and \( r \) model copy-in (or value) and copy-out parameters.
6.0.2 Procedure Calls

A procedure like (6.2) is invoked by the program statement

\[ \text{call} \; v := P(E) \]

where \( v \) is a list of variables and \( E \) is a list of expressions. The lengths of \( r \) and \( v \) are to be equal, and so are those of \( x \) and \( E \). The semantics of this procedure call is defined by

\[ [x, r \bullet x := E ; w, r : [\text{Pre}, (ePost, nPost)] ; v := r] \]

(6.4)

When a result value \( r \) is involved, I write \( \text{spec} \) for the command

\[ w, r : [\text{Pre}, (ePost, nPost)] \]

(6.5)

rather than for the command (6.1). Thus, (6.4) can be rendered as

\[ [x, r \bullet x := E ; \text{spec} ; v := r] \]

6.0.3 Procedure Implementation

Since procedures are essentially given as specification statements, we don’t expect compilers to produce executable code from them (see Sections 1.7 and 1.8). Instead, a programmer supplies an implementation of the procedure. A call statement is compiled into a subroutine call to the implementation, but the piece of code that contains the call statement is verified using the specification of the procedure. In order for this transformation to be correct, an implementation must be a refinement of the specification (see Section 1.8). Exactly then (well, see Section 6.0.4 below) do we say that the implementation meets its specification.

The notation I use for introducing the implementation of a procedure like (6.2) is

\[ \text{impl} \; r := P(x) \; \text{is gc} \]

where gc is a guarded command. The signature \( r := P(x) \) is repeated here (as opposed to just listing the name \( P \) ) to emphasize that \( x \) and \( r \) are identifiers that may be referred to in \( gc \). A procedure has exactly one implementation. \( P \) meets its specification just when

\[ w, r : [\text{Pre}, (ePost, nPost)] \sqsubseteq gc \]

(6.6)

Applying the definition of \( \sqsubseteq \) (1.19), this proof obligation is

\( (\forall P, Q \; \text{ s.t. } \ [wp.(w, r : [\text{Pre}, (ePost, nPost)]).((P, Q) \Rightarrow wp.gc.(P, Q))] \)

(6.7)

where \( P \) and \( Q \) range over all predicates on the final state. This quantification may appear overwhelming to a verification process. As a cure, I present an alternative rendering of this proof obligation in Chapter 11, where I also prove the two renderings equivalent. The alternative rendering does not quantify over all predicates and is therefore sometimes preferable to (6.7), especially when doing automatic verification.
6.0.4 Termination

I left one important detail out of the discussion of when an implementation meets its specification: termination. Since the guarded command gc may contain procedure calls, it may, for example, recursively call itself. One must therefore ensure that such recursion eventually ends, so that gc may be shown to terminate. For example, consider the following procedure specification and implementation.

\[
\text{spec} \ P() \text{ is spec} ; \\
\text{impl} \ P() \text{ is call } P()
\]

Here, we find that the implementation of \( P \) is indeed a refinement of its specification (in fact, the two are equal). Nevertheless, as programmers we know the call to \( P \) will not establish what is prescribed by \text{spec}, but will instead result in infinite recursion.

Proving, in the presence of mutually recursive procedures (and also replaceable methods, see Section 8.2), that all procedure calls in an implementation terminate can be quite tricky. This proof obligation may be dealt with separately from the proof obligation of the refinement. If termination is of no concern, \( wlp \) (as opposed to \( wp \)) can be used. Then, the refinement (6.6) is the only proof obligation for showing that a procedure implementation meets its specification. This is what I use in Part III, Chapter 11.

6.1 Alternative statements

In languages like Modula-3, Ada, and C, we find the presence of alternative statements like IF. The modeling of these is straightforward. For example, the Modula-3 statement

\[
\text{IF } b_0 \text{ THEN } S_0 \text{ ELSIF } b_1 \text{ THEN } S_1 \text{ ELSE } S_2 \text{ END}
\]

is modeled as

\[
b_0 \rightarrow S_0 \text{ } b_1 \rightarrow S_1 \text{ } S_2
\]

Since all statements in Modula-3, Ada, and C are total (cf. Section 1.4), I may write this statement as

\[
\text{if } b_0 \rightarrow S_0 \text{ } b_1 \rightarrow S_1 \text{ } S_2 \text{ fi}
\]

(cf. (1.10)).

The simpler statement

\[
\text{IF } b \text{ THEN } S \text{ END}
\]

is modeled as

\[
\text{if } b \rightarrow S \text{ } \text{skip} \text{ fi}
\]

Because this statement occurs so frequently, I introduce

\[
\text{if } b \text{ then } S \text{ fi}
\]

as a shorthand for it.
6.2 Statements that "go wrong"

When reasoning about a program, it is often useful to subdivide the notion of non-termination into cases where the program results in a failure caused by a checked run-time error like a nil dereference or an array index out-of-bounds error vs. cases where the program does not terminate because of, for example, an infinite loop. I will refer to the former as the program "going wrong".

This is important, for example, when verifying a run-time system—the execution of a run-time may never terminate, yet one wants to be certain it does not result in a checked run-time error. Note that neither wp nor wlp caters for this distinction.

The distinction can be captured by viewing programs as having three possible outcomes: normal, exceptional, and erroneous. This is a straightforward extension of having only two outcomes (cf. Section 2.6). The wlp of a program is now

\[ \text{wlp.} S(P,Q,E) \text{ is true of exactly those initial states from which execution of } S \text{ is guaranteed} \]

- to terminate exceptionally in a state satisfying \( P \), or
- to terminate normally in a state satisfying \( Q \), or
- to terminate erroneously (go wrong) in a state satisfying \( E \), or
- to not terminate at all.

The interpretation of \( wp \) is extended similarly.

With these triples, I define a third unit statement, wrong (cf. Section 1.2),

\[ \text{wp.wrong.}(P,Q,E) = E \quad \text{wlp.wrong.}(P,Q,E) = E \quad . \quad (6.8) \]

Statement wrong is used to model a run-time error. Note that execution of wrong actually does terminate.

In some programming languages, like Ada, a programmer can provide a handler for run-time errors. That calls for a third type of statement composition (the other two being ; and <1, see Section 1.2). Alternatively, if checked run-time errors are to be avoided at all costs, perhaps because they may result in the operating system aborting program execution (cf. Modula-3), no additional statement composition is needed. This is the view I take in this thesis. Hence, when computing the weakest (liberal) precondition of a statement, we are always interested in using false as the third component in the postcondition triple. To avoid cluttering formulas, I thus show only the first two components, the third component always being false implicitly.

Rewriting (6.8) with the convention of omitting the third component, we have

\[ \text{wp.wrong.}(P,Q) = \text{false} \quad \text{wlp.wrong.}(P,Q) = \text{false} \quad . \quad (6.9) \]

Remark 6.0. A desirable property of a program \( S \) is that \( \text{wlp.} S \) be universally conjunctive. Although (6.9) seems to indicate that \( \text{wp.wrong} \) is not universally conjunctive (since \( \text{wp.wrong.}(true, true) \) is not true),
one should remember that \( wlp.S.(P, Q) \) in (6.9) is only a shorthand for \( wlp.S.(P, Q, false) \). The general form (6.8) is indeed universally conjunctive, so, for example, we have

\[
wp.wrong.(true, true, true) = true
\]

With the convention of omitting the third component, all statement definitions given in Chapter 1 remain unchanged. Of those statements, only the specification statement introduces a way for a statement to go wrong. In Section 1.7, I deferred discussing the operational interpretation of the specification statement started in a state that does not satisfy the precondition. I give that interpretation now: If \( Pre \) does not hold in the state from which \( w : [Pre, (ePost, nPost)] \) is executed, the program goes wrong. Hence, the full definition of the specification statement is given by

\[
wp.(w : [Pre, (ePost, nPost)]).(P, Q, E) = (Pre \land (\forall w \lor (ePost \Rightarrow P) \land (nPost \Rightarrow Q))[v_0 := v]) \lor (\neg Pre \land E)
\]

Since the statement always terminates, its \( wlp \) coincides with its \( wp \). With \( E := false \), we have definition (1.17), as given originally.

6.2.0 ASSERT STATEMENT

The \textit{wrong} statement always goes wrong. The \texttt{assert} statement is a statement that goes wrong only under a parameterized condition. For any predicate \( b \), the assert statement is defined as

\[
\texttt{assert } b = \texttt{ if } \neg b \texttt{ then wrong fi}
\]

The weakest precondition of \texttt{assert} thus satisfies

\[
wp.(\texttt{assert } b).(P, Q) = b \land Q \tag{6.10}
\]

and similarly for its weakest liberal precondition. Interpreted operationally, \texttt{assert } \( b \) goes wrong just when executed in a state where \( b \) does not hold. If \( b \) does hold, \texttt{assert } \( b \) terminates normally and does not alter the program state. Note that \texttt{assert } \texttt{false} coincides with \texttt{wrong}.

6.3 Expressions

Expressions in many programming languages allow conveniences like \texttt{short-circuit} (or \texttt{conditional}) operators. Calls to procedures with one result value are also allowed in expressions. Another issue is that some expressions are not always defined. In this section, I treat the modeling of such expressions.
The first step in this modeling is to break complex expressions up into smaller ones. This is a well-known technique in, for example, the generation of intermediate three-address code in compiler design [0]. For example,

\[ x := y + P(z, Q(w)) \]

is elaborated into

\[
\begin{align*}
& t_0, t_1 \\
\bullet & \text{ call } t_0 := Q(w) \\
; & \text{ call } t_1 := P(z, t_0) \\
; & x := y + t_1
\end{align*}
\]

where \( t_0 \) and \( t_1 \) are temporary local variables with fresh names. This example shows how to model procedure calls in expressions. (Alternatively, one can view this as the definition of procedure calls occurring in expressions.) Note that, since procedures \( P \) and \( Q \) may have side effects, so may the evaluation of the expression \( y + P(z, Q(w)) \).

Short-circuit operators are handled similarly. For example,

\[ x := B \text{ cand } C \]

where \text{ cand } denotes the conditional-and operator, is elaborated into (or, gets its definition from)

\[
\begin{align*}
& t \bullet t := B ; \text{ if } t \text{ then } t := C \text{ fi ; } x := t
\end{align*}
\]

where \( t := B \) and \( t := C \) may require further elaboration.

Some operators are not defined for all operands. For example, division is not defined for a second argument of 0, and the array indexing operator (introduced in Chapter 7) is not defined for indices outside the index set of the array. Expressions involving such operators are called partial.

Sometimes “not defined” means the value of the expression in unspecified; more frequently, “not defined” means evaluation of the expression results in a checked run-time error. In the case of the latter, a statement

\[ x := a \text{ div } b \]

is elaborated into

\[
\begin{align*}
\text{assert } b \neq 0 ; x := a \text{ div } b
\end{align*}
\]

which incorporates the prescribed run-time check.

In the examples above, I have only shown expressions that occur in assignment statements. Expressions that occur elsewhere are handled in a similar way. For example,

\[ \text{WHILE } B \text{ DO S END} \]
is elaborated into
\[
[ \ b \ b := B \ \\
; \ do \ b \mapsto S ; b := B \ od \\
]\]
where the two occurrences of \( b := B \) may require further elaboration.

As a final concern, I discuss the order of evaluation of parameters. For example, a programming language may not specify the order of evaluation of the operands of \( + \). The order makes a difference when the operands may have side effects. Thus, a statement like \( x := A + B \) can be elaborated into
\[
[ \ t_0, t_1 \\
\bullet \ (t_0 := A \ ; \ t_1 := B \ ; \ t_0 := A) \\
; \ x := t_0 + t_1 \\
]\]
The price of this elaboration —exponentially longer formulas in its verification—compares unfavorably to the marginal value of this elaboration over one like
\[
[ \ t_0, t_1 \bullet \ t_0 := A \ ; \ t_1 := B ; x := t_0 + t_1 ]
\]
Instead, one can disallow expressions for which different permitted orders have an effect on the result of the computation. Ada, for example, defines such code to be erroneous, meaning that it causes an unchecked run-time error. It only seems fair that a verification process should catch such programming errors. To detect such possibilities, procedure specifications need to mention not only the variables a procedure writes (recall, these are given in the modifies clause), but also those it reads. Such information is facilitated, for example, by the global in construct in Penelope [62].

63
Part II

Data Structures
Data structures

In Chapter 1 of Part I, I described statements of an imperative programming notation and gave their semantics. Two of these statements, the assignment and specification statements, modify the state of a program, whereas the others deal only with the flow of control of the program. The particular types and internal structure of the variables are of no importance in Part I; I use only the fact that the variables are independent of each other, so that the update of one variable does not affect the values of other variables (see Section 1.1).

In practical settings, the state space of a program is more complicated. A program may update only some portion of a variable, e.g., an element of an array or a field of a record. A program may also have several ways of referring to the same piece of data, e.g., through two identical indices into an array or through references (pointers). Object-oriented languages provide an even fancier mechanism, known as subtyping, for organizing the data of a program.

In this Part, I describe the data structures most commonly provided by imperative languages: subranges, enumerations, arrays, records, sets, references, and objects. I show how these are modeled in the programming notation from Chapter 1.

Outline

In Chapter 7, I introduce types and describe how to model the most common data types using the constructs I define. In Chapter 8, I introduce into my programming notation the mechanisms necessary to deal with objects. Readers familiar with notions of subtyping such as [48, 56] may be in for a pleasant surprise: separation of concerns and thus simplicity. In Chapter 9, I show how to make sense of this simplicity when specifying and implementing objects. In that chapter, I present data abstraction and data refinement as these are known in the literature (cf. [38, 30, 43, 3, 11, 24]). As I show in Part III, this view of data refinement is not sound except in very restrictive modular programs. Nevertheless, Chapter 9 gives a flavor of the kinds of abstractions in which I am interested, and provides a nice coverage of the concepts used in Part III.
Data types

In this chapter, I introduce types and global variables. I then explain how to model data types from common programming languages, like arrays and records, in the programming notation from Chapter 1. I also define a construct that introduces a new type into a program, and show how it can be used in the modeling of references.

I assume some familiarity with these data types from a language like Modula-3, Ada, or C.

7.0 Types

A variable may be of a certain type. A type is a possibly empty, possibly infinite set of values. Examples of types are int, the set of integers, and bool, the set \{false, true\}. Variables can only be declared to be of nonempty types.

7.0.0 Composite Types

Types can be composed in two ways to form new types. These compositions are familiar from set theory.

For any types \( S \) and \( T \), \( S \times T \) is the type consisting of pairs of values, one from \( S \) and one from \( T \).

Much more important is the map type \( S \rightarrow T \). Type \( S \) in \( S \rightarrow T \) is called the index type (or set), and \( T \) is called the element type. Values of map types, called maps, can be applied to values of their index type, called indices. Every map is total, i.e., it can be applied to any element of its index set. Consider a variable \( a \) of type \( S \rightarrow T \). Given an expression \( i \) of type \( S \), \( a[i] \) is an expression of type \( T \). The value of this expression is the value of \( a \) applied to \( i \). Because I often think of maps and arrays as being synonymous, I often say "a at index \( i \)" or "an indexed with \( i \)" instead of "\( a \) applied to \( i \)". I will also refer to this operation as array dereferencing.

A map can be updated at a particular index using the assignment statement

\[
d[i] := E
\]  

(7.0)
This is shorthand for
\[ a := \text{subst} . a ( i : E ) \]
\[ \text{where } \text{subst} . a ( i : E ) \text{ is the map that differs from } a \text{ only in that, applied to } i, \]
\[ \text{subst} . a ( i : E ) \text{ yields } E [37, 40, 17] . \]

**Remark 7.0.** Fortunately, in spite of its shape, we know of an efficient implementation of (7.1), *viz.*, the one that (7.0) suggests: Just change the state of \( a[i] \), element \( i \) of array \( a \).

Formally, the meaning of \( \text{subst} \) is given by the following two axioms.

\[ \text{subst} . a ( i : E ) [ j ] = E \quad \text{if } i = j \]
\[ \text{subst} . a ( i : E ) [ j ] = a [ j ] \quad \text{if } i \neq j \]

Axiom (7.2) states that indexing \( \text{subst} . a ( i : E ) \) with \( i \) yields \( E \). Axiom (7.3) states that indexing \( \text{subst} . a ( i : E ) \) with \( j \), where \( i \neq j \), yields the same result as indexing \( a \) with \( j \). (See [69] on how these axioms are used in automatic theorem proving.)

### 7.0.1 Typed Variables

When writing a variable, I will sometimes write its type following the variable identifier, separated by a colon. A variable \( x \) of type \( T \) is thus denoted

\[ x : T \]

This requires that \( T \) be nonempty.

I distinguish between *local* and *global* variables. Local variables are declared by block statements (see Section 1.3). They are created at the beginning of execution of the block and perish as control leaves the block.

Global variables, on the other hand, are created at the beginning of a program execution and survive all changes of the current control point in the execution. (In a modular program, variables may or may not be *visible* at the current control point, but that doesn't play a rôle until Part III.)

I introduce each global variable with a *var* declaration, as in

\[ \text{var } x : T \]

Similarly, a local variable may be given a type at the point of declaration. As the semantics of blocks suggests, I assume the initial value of a variable to be any value of its type. (This matches Modula-3's definition, but not Ada's or C's, neither of which provides this guarantee.)

**Remark 7.1.** The Ada language [1] does not require that the language implementation initialize variables, and defines programs that use an uninitialized variable as *erroneous*. The rationale behind this may be the fact that Ada is designed to support systems programming. This
means that a variable can be mapped to a register in a machine, and
any write to such a variable may have a hardware side effect such as dis-
abling interrupts or sending a packet. However, Ada does require that the
language implementation initialize pointers to nil. So much for allowing
variables to be mapped to registers.

The C programming language [45] contains three simple types: integers,
reals, and pointers, each grouped into possibly different sizes. Since any
values of the bits that represent an integer make up some integer value,
an integer always contains a value of its type, and similarly for reals on
many machines. The bits representing a pointer, however, may denote
an invalid address, i.e., an address that cannot be dereferenced, because
of, for example, memory protection. C does not guarantee any particular
initialization of pointer variables.

7.0.2 Narrowing

Let \( v \) be a variable of type \( S \) and \( E \) an expression of type \( T \). If \( T \subseteq S \), then an
assignment

\[
\begin{align*}
v & := E \\
\end{align*}
\tag{7.4}
\]

is always legal. If \( S \subset T \), the assignment is also allowed; however, since some potential
values of \( E \) are not assignable to \( v \), a run-time check, called a narrow check, is
necessitated. Assignment (7.4) is thus treated as

\[
\left[ t : T \bullet t := E ; \text{assert } t \in S ; v := t \right]
\]

where \( t \) denotes a temporary variable with a fresh name (cf. partial expressions,
Section 6.3).

Remember that parameters and result values of procedures are defined via assign-
ments (Section 6.0); thus, narrowing applies there, too. For example, for \( S \subseteq T \),

\[
\text{spec } P(x : S) \text{ is} \\
\text{requires } Q
\]

is equivalent to

\[
\text{spec } P(x : T) \text{ is} \\
\text{requires } x \in S \land Q
\]

7.1 Types in common programming languages

In this section, I describe the correspondence between the previous section and data
structures in programming languages like Modula-3.
7.1.0 SUBRANGES

In Modula-3, a subrange is a type like \([M..N]\), where \(M\) and \(N\) are integer constants. It contains the (possibly empty) inclusive range of integers from \(M\) to \(N\).

Subranges need not be treated as separate types in my formalism. Consider a variable \(x\) of type \([M..N]\). Like all other variables, \(x\) is initialized to a value of its type. Values assigned to \(x\) (i.e., \(E\) in \(x := E\)) are checked at compile-time to be integers. However, the restriction of \(E\) being in the correct range of the integers is checked at run-time by narrowing. Hence, \(x := E\) is treated as

\[
[t : \text{int} \cdot t := E \ ; \ \text{assert} \ M \leq t \land t \leq N \ ; \ x := t]
\]

Consequently, \(M \leq x \land x \leq N\) is an invariant of the program. This fact can then be used in proofs, as required, for example, in showing that

\[
[y : [2\cdotM..2\cdotN] \cdot y := 2\cdot x]
\]
does not go wrong.

7.1.1 ENUMERATIONS

In Modula-3,

\[
\text{TYPE } E = \{\text{egg, sugar, flour}\};
\]

is an example of a declaration of an enumeration type \(E\). This particular type has three elements, written \(E.\text{egg}\), \(E.\text{sugar}\), and \(E.\text{flour}\). These names may have a meaning to a programmer. Mathematically, however, I treat them simply as an alternate notation for 0, 1, and 2, respectively. This renders unnecessary the introduction of a new domain of values and an order thereon. Hence, the type \(E\) is treated as the subrange \([0..2]\).

A programming language may impose additional restrictions regarding the use of these types. For example, a variable of type \(E\) cannot be assigned to a variable of type \text{int}. These restrictions enforce a disciplined use of the types and do not pose any problem in the theory.

7.1.2 ARRAYS

An array type

\[
\text{ARRAY } S \text{ OF } T
\]

is treated as the map \(S \rightarrow T\). The expression \(a[i]\) is partial (Section 6.3) and can only be evaluated if \(i\) is in the index set of \(a\), something that in general needs a narrow check at run-time (for example, if \(S\) is a subrange and \(i\) is an integer).

Commonly, programming languages like Modula-3 define

\[
\text{ARRAY } S0, S1 \text{ OF } T
\]
to be a shorthand for

\[
\text{ARRAY } S0 \text{ OF ARRAY } S1 \text{ OF } T,
\]

(or, in the theory, \( S0 \rightarrow (S1 \rightarrow T) \)), and similarly \( a[i, j] \) for \( a[i][j] \). Thus,

\[
a[i, j] := E
= \begin{cases} \text{Modula-3 shorthand} & \end{cases}
\]

\[
a[i][j] := E
= \begin{cases} \text{(7.1): array update shorthand} & \end{cases}
\]

\[
a[i] := \text{subst.}(a[i])(j : E)
= \begin{cases} \text{(7.1): array update shorthand} & \end{cases}
\]

\[
a := \text{subst.a.}(i : \text{subst.}(a[i])(j : E)).
\]

This allows \( a[i] \) to be treated as an array in its own right. If that feature is not needed, one may prefer to treat this array type as

\[
(S0 \times S1) \rightarrow T
\]

Then, an index is a pair, and thus \( a[i, j] := E \) is simply

\[
a := \text{subst.a.}((i, j) : E)
\]

that is, \( a \) gets \( a \) in which element \((i, j)\) has been replaced by \( E \).

### 7.1.3 Records

An example Modula-3 record type is

\[
\text{TYPE } R = \text{RECORD } f0 : T0 ; f1 : T1 \text{ END;}
\]

\( f0 \) and \( f1 \) are distinct identifiers known as the fields of \( R \). They have the respective types \( T0 \) and \( T1 \). For a value \( r : R \), \( r.f0 \), called a field dereference, denotes field \( f0 \) of \( r \).

This record type can be thought of as an array type with index set \( \{f0, f1\} \). Thus, \( r.f0 \) means \( r[f0] \). By treating records as arrays, no additional theory is required; the array axioms suffice.

Note that \( r[f0] \) and \( r[f1] \) have types \( T0 \) and \( T1 \), respectively, and that these types may differ. This causes no problem, even in languages that do typing at compile-time, because the only way to index \( r \) is by the constants \( f0 \) and \( f1 \) themselves. The type of a field dereference is thus immediately available.
7.1.4 Sets

Modula-3 writes a set of elements of a type $T$ as

$$\text{SET OF } T$$

It, too, can be treated as a map, viz.,

$$T \to \text{bool}$$

Hence, a set operation like

$$t \text{ IN } s$$

where $s$ is of type SET OF $T$ and $t$ is of type $T$, is taken to be the boolean expression $s[t]$.

7.2 Declaring new types

Just as variables can be declared by `var` declarations, my programming notation allows new types to be defined by `type` declarations, as in

$$\text{type } R$$

This introduces a new name $R$, and declares it to be a type. $R$ contains an infinite number of elements, all but one of which differ from the elements of other types declared by `type`. The one exception is a special constant called `nil`, which is part of every type declared by `type`.

I will use the names `reference` type or `object` type when referring to a type declared by `type`. Similarly, I will call elements of such a type `references` or `objects`. The inspiration for these names is discussed below and in the next chapter.

7.2.0 References

In common imperative languages, a `reference` type is sometimes called a `pointer` type. (Ada calls them `access` types.) Every reference type has a `referent` type, that is, the type to which the reference type is a reference. A reference is usually implemented as the address of its `referent`, a piece of data residing in the heap of a program. In Modula-3,

$$\text{REF } T$$

is a reference type whose referent type is $T$.

A reference $r$ is dereferenced by $r^\land$. This maps $r$ to its referent, a fact that reveals the need for a map of type $(\text{REF } T) \to T$ [40].
Let $R$ be a unique name for a particular reference type with referent type $T$. $R$ is then modeled by

```plaintext
type $R$
var $map^R : R^- \rightarrow T$
```

where $R^-$ denotes $R \setminus \{\text{nil}\}$ and $^*$ is a reserved character so that $map^R$ is a name uniquely determined by $R$. $map^R$ is called a dereference map (or collection). A dereference $r^*$ is then simply the array dereference $map^R[r]$. Note that the index type of $map^R$ is $R^-$, so $map^R[r]$ is a partial expression that can be evaluated only for references $r$ other than $\text{nil}$ (see Section 6.3).

Remark 7.2. Dereference maps are explicitly declared and manipulated in the programming language Euclid [47] and in an early version of Pascal [79].

The crux with modeling pointers in a theory is the aliasing they introduce. Using maps (arrays) to model references reduces the problem to the aliasing problem of indices into arrays; this, in turn, is handled by the axioms (7.2) and (7.3). This assumes that all access of referents go via references, i.e., there is no way other than $^*$ to alias a referent. This is true in Modula-3 and Ada, where all references point into the heap, but not in C or C++, where one can take the address of variables (and of just about everything else). Thus, only a disciplined subset of C and C++ can be modeled directly by the techniques I have described here.

### 7.2.1 Allocation and Deallocation

Languages like Modula-3, Ada, and C++ that provide references also provide a mechanism to create new references and referents. In Modula-3,

```plaintext
NEW(R)
```

where $R$ is a reference type, returns a new reference of type $R$. In effect, NEW also allocates a referent in the program heap to which the new reference points. This can be modeled by introducing a map

```plaintext
var allocated^R : R^- \rightarrow \text{bool}
```

initialized so that

```plaintext
(\forall r : R^- \Rightarrow \neg\text{allocated}^R[r])
```

NEW(R) is the only construct that modifies allocated^R. The specification of NEW(R) as a procedure is given as

```plaintext
spec r : R^- := NEW(R) is
modifies allocated^R
ensures \neg\text{allocated}^R_0[r] \land \text{allocated}^R[r] \land
\langle \forall s : R^- \mid s \neq r \Rightarrow \text{allocated}^R_0[s] = \text{allocated}^R[s] \rangle
```

74
Remark 7.3. Recall from Section 1.7 that a variable subscripted with 0 refers to the initial value of that variable. So, \( \text{allocated} \times R_0 \) refers to the value of \( \text{allocated} \times R \) on entry to the procedure.

This assumes there is always enough memory for new referents. Deallocation of a referent is modeled similarly. A map

\[
\text{var deallocated} \times R : R^- \to \text{bool}
\]

initialized like \( \text{allocated} \times R \), is introduced. Then, a procedure \( \text{FREE} \) is defined.

\[
\text{spec \ FREE}(r : R^-) \text{ is}
\begin{align*}
\text{modifies} & \ \text{deallocated} \times R \\
\text{requires} & \ \text{allocated} \times R[r] \land \neg \text{deallocated} \times R[r] \\
\text{ensures} & \ \text{deallocated} \times R[r] \land \\
& (\forall s : R^- \land s \neq r \lor \text{deallocated} \times R_0[s] = \text{deallocated} \times R[s])
\end{align*}
\]

This introduces another requirement on evaluating the expression \( \text{map} \times R[r] \), viz.,

\[
\neg \text{deallocated} \times R[r]
\]

Because \( \text{NEW}(R) \) and \( \text{FREE} \) are the only procedures that modify \( \text{allocated} \times R \) and \( \text{deallocated} \times R \) — these map variables are not accessible like regular program variables — , we can prove

\[
(\forall r : R^- \lor \text{deallocated} \times R[r] \Rightarrow \text{allocated} \times R[r])
\]

to be an invariant of any program execution.

The reason for providing a procedure like \( \text{FREE} \) is to gain (if “gain” is really the right word) programmer-defined control of storage efficiency. A pleasant alternative is for the run-time system to assume control of this, an effort realized by a garbage collector. The run-time system then reclaims the storage of referents to which no reference exists. This choice is pursued by, for example, Modula-3, thus rendering \( \text{deallocated} \times R \) and \( \text{FREE} \) unnecessary.

7.3 Maps and specifications

Consider a simple procedure, call it \( \text{Update} \), whose effect is

\[
a[i] := y
\]

for some global map variable \( a \) and appropriately typed parameters \( i \) and \( y \). Since \( \text{Update} \) updates \( a \),

\[
\text{modifies} \ a
\]

is part of \( \text{Update} \)'s specification. The final value of \( a[i] \) is specified by

\[
\text{ensures} \ a[i] = y
\]

75
Since no other elements of $a$ are updated, the clause
\[
\text{ensures } \left( \forall k \mid k \neq i \Rightarrow a_0[k] = a[k] \right)
\]
is also part of Update's specification.

Updating a map variable at one given index is quite common, as examples throughout the rest of this thesis show. Therefore, I introduce the shorthand
\[
\text{modifies } a[i]
\]
Roughly speaking, this means
\[
\text{modifies } a \text{ ensures } \left( \forall k \mid k \neq i \Rightarrow a_0[k] = a[k] \right)
\]
I say "roughly" because of some subtle and messy details, described next.

Consider a procedure \textit{Swap}(i,j) that swaps $a[i]$ and $a[j]$. It would be convenient to be able to write its specification as
\[
\text{spec } \textit{Swap}(i,j) \text{ is }
\begin{align*}
\text{modifies } a[i], a[j] \\
\text{ensures } a[i] &= a_0[j] \land a[j] = a_0[i]
\end{align*}
\]
However, applying the rough formulation of the shorthand, we get
\[
\text{spec } \textit{Swap}(i,j) \text{ is }
\begin{align*}
\text{modifies } a, a \\
\text{ensures } a[i] &= a_0[j] \land a[j] = a_0[i] \\
&\quad \left( \forall k \mid k \neq i \Rightarrow a_0[k] = a[k] \right) \\
&\quad \left( \forall k \mid k \neq j \Rightarrow a_0[k] = a[k] \right)
\end{align*}
\]
For $i = j$, this \textbf{ensures} condition simplifies to
\[
\text{ensures } a_0 = a 
\]
and for $i \neq j$, it simplifies to
\[
\text{ensures } a_0 = a \land a_0[i] = a_0[j]
\]
which makes the specification a partial command (Section 1.4)—if $i \neq j \land a[i] \neq a[j]$ holds initially, \textit{Swap} needs a miracle to establish $a_0[i] = a_0[j]$. This is not the intended specification, a fact for which I blame the formulation of the shorthand. Instead,
\[
\text{modifies } a \text{ ensures } \left( \forall k \mid k \neq i \land k \neq j \Rightarrow a_0[k] = a[k] \right)
\]
does the trick. I leave it to the reader to write the straightforward but messy generalization of this rule.

There is another detail to be discussed. In the examples above, $i$ and $j$ are expressions whose values are unchanged by the procedure. Consider a variation of
procedure Update where \( i \) is not a parameter but a global variable. Then consider the specification

\[
\text{spec UpdateAndAdvance}(y) \text{ is}
\begin{align*}
\text{modifies } & a[i], i \\
\text{ensures } & a[i_0] = y \land i = i_0 + 1
\end{align*}
\]

This specifies UpdateAndAdvance to write \( y \) at \( a[i] \) and then to increment \( i \) by 1. Here, the intention is that the modifies clause be a shorthand for

\[
\text{modifies } a, i \text{ ensures } (\forall k \mid k \neq i_0 \Rightarrow a_0[k] = a[k])
\]

as opposed to

\[
\text{modifies } a, i \text{ ensures } (\forall k \mid k \neq i \Rightarrow a_0[k] = a[k]) \tag{7.5}
\]

The same holds true for many other examples, but it is conceivable that one may sometimes want (7.5). Yet another choice is to allow the shorthand only for constant indices.
Objects

In this chapter, I discuss object types. These are similar to reference types, but cannot be dereferenced using ^, and thus lack the map*R map that every reference type R has. Instead of that one map, object types can have several maps, called data fields. In addition, object types feature subtyping and methods. I describe each of these features, and conclude by relating objects in my notation to those in common programming languages.

Remark 8.0. Since C++ terminology of objects differs from the usual ones [28, 64], readers familiar with C++ but with no other object-oriented language may want to read Section 8.6.1 before reading this chapter from the start.

8.0 Subtypes

Every object type has an infinite number of elements, one of which is nil (see Section 7.2). S is called a subtype of an object type T just when S, too, is an object type and S is a subset of T. T is then called a supertype of S.

As described in the previous chapter, an object type T is declared by

\[ \text{type } T \]

I now define a way to declare one object type from another. For any object type T,

\[ \text{type } S <: T \]

introduces a new name S, and declares it to be a subtype of T different from T itself, i.e., a proper subtype of T. T is called the immediate supertype of S (see also Remark 8.1).

If T0 is a subtype of T, then any T0 object, t say, is also a T object (because \( t \in T0 \subseteq T \)). The smallest subtype Tn such that \( t \in Tn \) is called the dynamic or allocated type of t. For any object type T, NEW(T) is guaranteed to return an object whose allocated type is T.
8.1 Data fields

For any object type $T$, I call a map

$$\text{var } x : T^\rightarrow X$$

a data field of type $T$ (or of some $T$ object). $x$ (or $x$ applied to some $T$ object) can be thought of as an attribute or property of $T$ (or of that $T$ object).

Note that a subtype shares (or inherits) the properties of its supertypes: For $T_0$ a subtype of $T$ and $t$ a (non-nil) $T_0$ object, $x$ can be applied to $t$, because $t$ is in the index set of $x$ ($x \in T_0^\rightarrow \subseteq T^\rightarrow$).

8.2 Methods

Like values of any other type, objects can be passed as parameters to procedures. In addition, objects have special procedures, called methods, that can be applied to them. An object type $T$ declares a method $m$ by

$$\text{method } t : T \text{ spec } r := m(x) \text{ is spec } .$$  \hspace{1cm} (8.0)

This is similar to a procedure specification (6.2) except for the prefixed “method $t : T$”. The “method...” shows that the method is declared for object type $T$. $T$ is called the declaring type of $m$. The “$t :$” introduces a name for a special parameter; this parameter is called the receiver and is often referred to as self [28, 64] or this [54, 23]. Consequently, $t$ abides by the same rules as $x$, with regard to being mentioned in spec (see Section 6.0).

An invocation of a method like (8.0) is written

$$\text{call } v := o.m(E) \hspace{1cm} ,$$  \hspace{1cm} (8.1)

where $v$ is a variable, $o$ is an expression of type $T$, and $E$ is a list of expressions (cf. (6.3)). This invokes method $m$ on object $o$. The semantics of this method invocation (cf. (6.4)) is

$$[ t, x, r \bullet t, x := o, E ; w, r : [\text{Pre}, (ePost, nPost)] ; v := r ] \hspace{1cm} ,$$

where $w, r : [\text{Pre}, (ePost, nPost)]$ is the method specification spec interpreted as a specification statement (cf. (6.1,6.5)).

8.3 Method implementations

Unlike procedures, which are restricted to one implementation, a method can have one implementation per subtype of the declaring type. Let $T$ denote the allocated type of $o$. Then, the method invocation (8.1) is implemented as a subroutine call to $T$’s implementation of $m$. 

79
If no explicit implementation is given for a particular subtype in a program, the method implementation defaults to that of the immediate supertype. For simplicity, I assume this definition leads to some implementation for every method invocation of every execution of a program.

**Remark 8.1.** This is the only place where the notion of an immediate supertype comes into play. Everywhere else,

\[
\text{type } S <: T
\]

can be interpreted as declaring \( S \) as any proper subtype of \( T \) — \( T \) need not be the immediate supertype. Before one can link and execute such a program, the name of the \( S \)'s immediate supertype must be given, so that methods can be implemented as described above. Modula-3 features so-called *partially opaque types*, which take advantage of this.

Let \( T_0 \) be a subtype of \( T \). Then, the notation used for associating with \( T_0 \) an implementation of method \( m \) (8.0) is

\[
\text{method } t : T_0 \text{ impl } r := m(x) \text{ is gc }
\]

\( t, r, x \) are identifiers that may be used in the guarded command \( \text{gc} \). For the implementation to meet its specification, the condition

\[
w, r : [\text{Pre}, (\text{ePost}, \text{nPost})] \subseteq \text{gc}
\]

must be established for \( t \) of type \( T_0 \).

### 8.4 Object simplicity

This is all there is to objects. Note that the notion of a subtype is as simple as the notion of a subset. Unlike [48, 56], the data fields have nothing to do with the subtype relation. Each data field is introduced independently, and I never mention that the fields of a subtype *data refine* or *simulate* those of a subtype. These concepts have to do with abstraction (see next chapter), not subtyping.

And nowhere do I need to consider the complete set of methods of an object type (though, admittedly, unlike [56], I am only considering properties of sequential programs). Instead, methods are declared independently of each other. A method implementation for a particular subtype has very little to do with the subtype itself. Instead, refinement is what matters. Notice also that the refinement is always between an implementation and the (one and only) specification of a method, not between the implementation at one type and the implementation at the immediate supertype.

Finally, note that the semantics of a method invocation depends only on the method specification. Thus, this, too, is independent of subtypes.

Both [48] and [56] seem to treat the issue of what it means for a collection of data fields and methods to form a data type, and the relation between such types. Thus,
they consider all properties of the behavior of such types. When reasoning about
the correctness of a program (as opposed to reasoning about properties of a type),
different properties of a type can be considered in isolation—for example, reasoning
about the effect of one method is orthogonal to the effects of other methods—; in
fact, considering them in isolation may be preferable, because it may simplify the
verification conditions. Nevertheless, techniques like those in [48, 56] provide some
utility in the design of object types and object-type hierarchies.

In summary, by separating the concepts of subtyping and abstraction, I achieve
simplicity.

8.5 Language implementations of objects

The fact that data fields are introduced separately and are maps whose index sets are
infinite poses no problems in language implementations. At any point in a program
execution, only a finite number of objects have been allocated, and only the allocated
portion of an index set will ever be used to dereference a map. Instead of placing a
data field next in memory to the same data field for each object (as the map notation
may suggest), all data fields for one object are placed next to each other. Thus, for
a data field $x$ and an object $t$, $x[t]$ is stored at some offset—a function of the name
$x$— into $t$’s data record, not at index $t$ of array $x$ as the notation may suggest.

Most object-oriented programming languages use a notation like $t.x$ instead of
$x[t]$. This has appeal, especially to object-orientation buffs—"$t$'s attribute $x"$ puts
the stress on $t$, whereas "$x$ at $t$" puts $x$ in the spotlight. My focus, however, is on
the mathematical meaning of objects.

An implementation makes sure that accessible from an object is an identification
of its allocated type. This type identification makes it possible to dispatch to the
right method implementation in a method invocation.

8.6 Objects in common programming languages

In this section, I describe the correspondence between the object types presented here
and those in common programming languages. I focus on Modula-3 and C++; other
languages are similar.

8.6.0 Modula-3

In Modula-3, an object type $T$ is declared by

\begin{verbatim}
TYPE T = SuperT OBJECT
    fieldlist
METHODS
    methodlist
\end{verbatim}

81
OVERRIDES

overridelist

END;

SuperT gives the name of the immediate supertype of T. Hence, in my notation, T is declared by

type T <: SuperT

fieldlist is a list of field declarations. In my notation, each such field x of type X is declared by

var x : T -> X

Similarly, every method in the method list is declared with a method T spec declaration. Note, though, that my notation requires a specification, whereas Modula-3 provides no way to state a specification (except, of course, informally as a comment).

In Modula-3, an implementation P for a method m is specified by appending ":= P" to m's declaration in methodlist for T, if T is the declaring type of m. If m is declared in a proper supertype of T, then m := P is given as an element in overridelist. This P must name a procedure. For a method m specified by (8.0), the specification of P must have the form

spec r := P(t : T0', x) is pspec

where T0' names some supertype of T0. (The names r, t, and x are allowed to differ in spec and pspec, since they are local to each specification. For simplicity, I assume them to be the same.) For P to be a valid implementation of m, one must then prove, for t of type T0,

spec ⊑ call r := P(t,x)

or stated differently,

spec ⊑ pspec

Since Modula-3 uses structural equivalence among types, a unique name for T should be used when translating to my notation. For a discussion of Modula-3's partially opaque types, see Remark 8.1.

8.6.1 C++

Modeling C++ in my notation is similar to modeling Modula-3. However, C++ uses some different terminology that is worth explaining in order to avoid confusion.

In C++, an object type, called a class, is a record type (a "struct") with extra features. Thus, what I consider an object is, in C++, a pointer to a C++ class. Then, my objects are like C++ class pointers as long as all such pointers are obtained via new. A class declared as a local variable is modeled as a record. But taking the
address of such a variable, like taking the address of anything else, introduces aliasing that I don’t handle (see Section 7.2.0).

So, as my objects correspond to pointers to records in C++, why do I not model objects and data fields that way, too? Subtypes may increase the number of data fields that an object has, and different subtypes may have different data fields. Using a map for each data field provides the flexibility required for this task—data fields can be added arbitrarily and can be added only to those subtypes for which they exist. Simplicity is another reason, because having the index set of the data record vary as a function of the object that is used to get to the data record becomes clumsy and awkward.

What I have called a method, C++ calls a virtual method. What C++ calls a (non-virtual) method, I simply view as a procedure, because it cannot be replaced in subtypes.

Finally, a note on protection levels. C++ features three data field protection levels, private, protected, and public. These levels enforce an access discipline and do not affect the semantics of the data fields and methods that can be accessed.
Abstraction

In this chapter, I introduce the concept of *data abstraction*, first for general variables and later for data fields. In Section 9.1, I introduce *data refinement*, an introduction that stays close to the way it was done originally by Hoare [38]. In Part III, however, where programs consist of modules, the techniques discussed in Section 9.1 will not suffice. Nevertheless, this chapter provides a flavor of what data abstraction and refinement are all about and why they are of interest.

9.0 Abstract variables

So far, each variable we have seen has represented one coordinate in the program state space. Let us now consider functions over these variables. For example, if \( x \) and \( y \) are program variables, we may consider the function \( x + 2 \cdot y \). We may name such a function by introducing an *abstract* (or *specification*) variable.

The declaration

\[
\text{spec var } z
\]

introduces an identifier \( z \), and declares it to be an abstract variable.

**Remark 9.0.** I only introduce *global* abstract variables, because they are the ones that are needed when writing modular specifications (see Part III). All concepts apply to local abstract variables as well—all that’s needed is a notation for declaring such variables.

An abstract variable is not a coordinate in the state space that is changed independently of other variables (cf. Section 1.1). Rather, it is simply a function of regular variables (the latter hereinafter called *program* or *concrete* variables). When the values of these program variables change, so does the value of the abstract variable, and *vice versa*.

**Remark 9.1.** A program variable, too, is a function. It abstracts a meaningful value from the bits that physical machines have. However,
since the program variables are independent of each other, one such bit represents part of only one program variable. (Each bit, also, is a function of more concrete entities such as the components of a virtual memory system, and they, in turn, are functions of voltages, and so on; in this thesis, I do not get more concrete than program variables.)

To specify what the value of \( z \) is, with respect to other variables, a \textit{rep} declaration is used. It has the form

\[
\text{rep } z \text{ is } R,
\]

where \( z \) names an abstract variable and \( R \), called the \textit{representation} of \( z \), is a predicate involving \( z \). The representation specifies the value of \( z \). For example,

\[
\text{rep } z \text{ is } z = x + 2 \cdot y
\]
defines \( z \) to be the function \( x + 2 \cdot y \).

For expressiveness, \( R \) is a predicate rather than the function itself. This allows a declaration like

\[
\text{rep } F \text{ is } 9 \cdot (C + 40) = 5 \cdot (F + 40)
\]

instead of forcing the formula to be written as

\[
F = \frac{9}{5} \cdot C + 32
\]

Nevertheless, I require that \( R \) specify \( z \) uniquely from the program variables. The jargon is that \( R \) is an \textit{abstraction function}, as opposed to an \textit{abstraction relation} (or \textit{nondeterministic function} [72]), which would only specify \( z \) down to a set of possible values. I restrict my attention to abstraction functions so as to avoid issues such as, "Does \( z = z \) hold?".

**Remark 9.2.** Interestingly enough, the formal proof of soundness in Chapter 12 does not rely on representations being abstraction functions; abstraction relations work just as well. However, with abstraction relations, Chapter 11 needs revamping.

It is common, however, for several values of the concrete variables to yield the same value of the abstract variable. This justifies the name \textit{abstract} variable, since it abstracts away from some details of particular states.

**Remark 9.3.** The names \textit{specification} variable \( \textit{vs. program} \) variable stem from the fact that the former typically exists only in specifications whereas the latter is compiled and takes up memory like the rest of the program.
I also assume that \( R \) is \textit{total} in the variables that represent \( z \). That is, I assume that for every state that ever occurs in an execution of a program, the value of \( z \) is defined.

An abstract variable need not be defined only in terms of program variables; it can also be defined in terms of other abstract variables. For example, for abstract variables \( a \) and \( b \),

\[
\begin{align*}
\text{rep } a & \text{ is } a = 3 \cdot b + 2 \cdot x ; \\
\text{rep } b & \text{ is } b = x + y
\end{align*}
\]

in effect defines \( a \) to satisfy

\[
a = 5 \cdot x + 3 \cdot y
\]

9.1 Abstract variables and refinement

In this section, I treat refinements involving abstract variables. Consider the following example.

\[
\begin{align*}
\text{var } c ; \\
\text{spec var } a ; \\
\text{rep } a & \text{ is } a = c^2
\end{align*}
\]

The statement \( a := 9 \) thus has the effect of setting \( c \) to either \(-3\) or \(3\). Therefore, the statement \( c := 3 \), which always sets \( c \) to \(3\) and never to \(-3\), is a refinement of \( a := 9 \) \((\text{cf. Section 1.8})\).

A refinement like this is called a \textit{data refinement} \([38]\). I proceed to show how such a refinement is often proven \((\text{cf. [38, 30, 43, 3]}\). Then, I briefly mention why data abstraction and refinement are important, and why the classical view of data refinement is too restrictive for doing abstraction in modular programs.

9.1.0 Classical data refinement

The idea is to consider two state spaces, one containing the abstract variables and the other containing the concrete variables. Within each state space, the respective variables are considered to be independent coordinates. The two state spaces are related by the predicate that defines the representation of the abstract variables.

A new command, \texttt{SwitchToAbstract}, is introduced. It is defined by the following weakest precondition. (The presence of exceptions is tangential to this discussion, so I assume programs whose outcome is always normal.)

\[
wp.\texttt{SwitchToAbstract}.Q = (\forall a \mid R \triangleright Q)
\]

where \( a \) is the list of abstract variables whose representation is prescribed by \( R \), and \( Q \) is a predicate over the abstract variables.
Informally, \textit{SwitchToAbstract} switches from the concrete partition to the abstract partition, and thus sets the values of the abstract variables according the values of the concrete variables and the representation predicate \( R \). \textit{SwitchToAbstract} establishes (abstract) postcondition \( Q \) from those (concrete) initial states in which, for each value of \( a \) that satisfies \( R \), \( Q \) holds. Since I assume that \( R \) determines \( a \) uniquely, there is at most one such value for \( a \). Moreover, since I assume \( R \) to be total in the representation of \( a \), there is at least one value for \( a \) that satisfies \( R \). Thus, the interpretation of \textit{SwitchToAbstract} can be reformulated: \textit{SwitchToAbstract} establishes \( Q \) from those initial states in which \( Q \) holds for the value of \( a \) that satisfies \( R \).

A concrete program \( C \) is defined to data refine an abstract program \( A \) just when

\[
\text{SwitchToAbstract} \; ; \; A \sqsubseteq C \; ; \; \text{SwitchToAbstract},
\]

where \( \sqsubseteq \) is defined as in Section 1.8. Both sides of \( \sqsubseteq \) show a program that starts in the concrete state space and ends in the abstract state space. The refinement thus compares the left- and right-hand sides with respect to their outcomes in the abstract state space. In other words, the details of the concrete state space matter not; only the abstract representation of these states do.

Remark 9.4. Unlike the data refinement found in the literature, the \textit{SwitchToAbstract} command does not appear in Part III, despite the fact that (or maybe, because) a more general form of abstraction is considered there.

Let’s now prove that \( c := 3 \) refines \( a := 9 \) in our example. We need to show

\[
\text{SwitchToAbstract} \; ; \; a := 9 \sqsubseteq c := 3 \; ; \; \text{SwitchToAbstract}.
\]

For any predicate \( Q \), we calculate,

\[
\begin{align*}
\text{wp.(c := 3 ; SwitchToAbstract).Q} &= \{ \; ; \text{and SwitchToAbstract} \; \} \\
&= \text{wp.}(c := 3).\{\forall a \mid a = c^2 \triangleright Q\} \\
&= \{ \text{one-point rule} \} \\
&= \text{wp.(c := 3).}(Q[a := c^2]) \\
&= \{ \; := \; \} \\
&= Q[a := c^2][c := 3] \\
&= \{ \text{substitution, since } Q \text{ is a predicate over the abstract variables and thus does not contain } c \} \\
&= Q[a := 9],
\end{align*}
\]

and

\[
\begin{align*}
\text{wp.(SwitchToAbstract ; a := 9).Q} &= \{ \; ; \text{and :=} \; \} \\
&= \text{wp.}\text{SwitchToAbstract.(Q[a := 9])} \\
&= \{ \text{SwitchToAbstract} \}
\end{align*}
\]
\( \forall a \ | \ a = c^2 \vdash Q[a := 9] \)
\[ = \{ \text{one-point rule, since } a \text{ does not occur free in } Q[a := 9] \} \]
\[ Q[a := 9] \]

This proves the data refinement.

9.1.1 Advantages and Shortcomings

Data abstraction and refinement are important because they allow us to introduce abstract variables to describe the abstract behavior of the procedures and methods in a module. They also admit a way to represent the properties of objects abstractly, independently of the details of the particular subtypes (see Section 9.2 below).

For example, consider the procedures of a module that implements a file system. Abstract variables are introduced to be used in the procedure specifications that describe the intended behavior of the file system. Later, in the implementation, design decisions are made as to how to represent the file system. At that time, concrete variables are introduced and the representation for each abstract variable is given.

Notice that the abstract variables, which are functions of the concrete variables, are introduced before the exact function definitions are given. In fact, the exact functions of the concrete variables that the abstract variables make up are not known at the time the abstract variables are introduced. The exact mapping is given only later; nevertheless, the mere naming of the abstract variables allows specifications to be written in terms of them.

Also note that any mapping to concrete variables (that admits an implementation—and even those that do not [66]) will do, because the details of the concrete representation do not matter to the callers of the procedures. This allows an implementation to change without radiating a new abstract behavior.

Classical data refinement does not permit reasoning about programs where abstract and concrete variables exist together. For example, neither of the programs

\[ a := 9 \ ; \ c := c + 1 \]

and

\[ a := 4 \cdot c \]

makes any sense in the classical data refinement model.

Remark 9.5. When doing data refinement via so-called auxiliary variables, statements like these are allowed (see, e.g., [65]).

Instead, the world is divided: either the abstract variables are present or the concrete variables are, but never both at the same time. This gives the view of the program as having exactly two modules, one with access only to abstract variables, the other with access only to concrete ones. My view of a modular program in Part III allows
any number of modules. Furthermore, my modules are not separated from each other; rather, a module can include (import) others. (This corresponds to the use of interfaces and modules in Modula-3 and Modula-2, and to the use of packages in Ada.) This means that abstract and concrete variables do exist together, a situation that yearns for a solution.

Before entering Part III, let me discuss abstract data fields and partial representations.

9.2 Abstract data fields

In Section 9.0, I introduced abstract variables without saying anything about types. Now, I consider a specific kind of abstract variables—those whose types make them data fields (see Section 8.1).

Consider an object type $T$ and an abstract data field declared by

$$\text{spec var } a : T^- \rightarrow A$$

for some type $A$. Let $T_0$ and $T_1$ be two subtypes of $T$, neither a subtype of the other. I allow the representation of $a$ to be different for these two subtypes. For example, if

$$\text{var } p : T^- \rightarrow A$$

is a data field of $T$ and $A$ denotes some numeric type, then $a$ can be represented by $a = 1.12 \cdot p$ for $T_0$ objects and $a = 1.08 \cdot p$ for $T_1$ objects. Each of these is called a partial representation of $a$. They are written

$$\text{rep } a[t : T_0] \text{ is } a[t] = 1.12 \cdot p[t] ;$$

$$\text{rep } a[t : T_1] \text{ is } a[t] = 1.08 \cdot p[t] \quad (9.0)$$

The "[..T0]" shows that the rep declaration only gives the representation of $a$ for $T_0$ objects, and the "t :" introduces a name (of type $T_0$) that can be used in the representing expression, and similarly for $[t : T_1]$.

**Remark 9.6.** It would be more accurate to write "[t : T0-]" instead of "[t : T0]", but since the index set of $a$ is $T^-$, i.e., $T \setminus \{\text{nil}\}$, I take "[t : T0]" to mean "[t : T0-]", which simplifies the notation slightly.

**Remark 9.7.** An alternative way of viewing the representation of $a$ (9.0) is

$$\text{rep } a \text{ is } (\forall t : T^- \rightarrow a[t]) = \left\{ \begin{array}{ll} 1.12 \cdot p[t] & \text{if } t \in T_0 \\ 1.08 \cdot p[t] & \text{if } t \in T_1 \\ \vdots & \end{array} \right. \ .$$

89
I say that \( a[i] = 1.12 \cdot p[i] \) is the representation of \( a \) at type \( T0 \).

The representation of \( a \) can also be in terms of the data fields of the respective subtypes. For example, if

\[
\begin{align*}
\textbf{var} \; x &: T0^- \rightarrow A \; ; \\
\textbf{var} \; y &: T1^- \rightarrow A
\end{align*}
\]

are data fields of \( T0 \) and \( T1 \), respectively, and \( A \) denotes the integers, then \( a \) can be represented by \( a = 2 \cdot x \) for \( T0 \) objects and \( a = y - 7 \) for \( T1 \) objects.

\[
\begin{align*}
\textbf{rep} \; a[t : T0] & \text{ is } a[i] = 2 \cdot x[i] \; ; \\
\textbf{rep} \; a[t : T1] & \text{ is } a[i] = y[i] - 7
\end{align*}
\]

The flexibility of stating different representations for different subtypes comes with a restriction. If a subtype \( T0 \) provides a representation, then no subtype of \( T0 \) may provide a different representation. That is, unlike methods, the subtypes of \( T0 \) may not override the representation provided at \( T0 \). This can be illustrated with a picture. In Figure 9.0, types are represented as nodes in a tree. Edges connect a subtype with its immediate supertype, and the convention is that supertypes are placed above their subtypes in the picture. Circled nodes indicate that a representation is given at this node. Thus, the shaded regions show subtypes that may not provide their own representation.

This restriction is pronounced in the interest of modular verification (see Part III), where all information about a program may not be available at the time of verification. The restriction then prevents the verification process (be it performed by a human or by a machine) from using the wrong representation. For example, if both \( T0 \) and \( Tk \) \((\text{cf. Figure 9.0})\) were to provide representations, then for a scope where types \( T \), \( T0 \), and \( Tn \) are visible, and only \( T0 \)'s (and not \( Tk \)'s) representation is visible, the verification process might use the representation provided at \( T0 \) as the prevailing representation for \( Tn \), whereas with full information about the program, the representation \( Tk \) would be used.

The object model in [56] does not have this restriction. There, a subtype provides all of its own data fields, \( i.e., \) it does not inherit any data fields — and, in particular,
inherits no abstract data fields— from its supertypes. The model herein comes closer to objects provided by languages like Modula-3 and C++.

The restriction does not so much restrict as it does provide a guiding methodology for the construction of object-type hierarchies. Consider a type $T$ that declares a data field $c$. Some supertype of $T$ declares an abstract data field $a$, and $T$ gives its representation of $a$ in terms of $c$. If, for some reason, it is expected that subtypes of $T$ would want to provide different representations of $a$, then one can often split type $T$ into two types, $T'$ and $Timpl$, say, where $Timpl <: T'$. The representation of $a$ previously given at $T$ is now given at $Timpl$. This usually means that $c$ should be declared at $Timpl$, but declaring it at $T'$ is sometimes also a possibility. The aforementioned subtypes of $T$ are now declared subtypes of $T'$, which means they are free to give their own representations of $a$.

To recap, subtypes of a type that declares an abstract data field can provide their own representations of that field. Each such representation is called a partial representation. Usage of partial representations is restricted to avoid that an object would have more than one representation of some data field. This restriction provides guidance in the design of object-type hierarchies.
Modularity
Modules and modular verification

In Parts I and II, I treated the semantics of control structures of programs and of data structures that those programs manipulate. In this Part, I consider making the verification of large programs feasible.

The verification of a program consists of the verification of a set of refinements. Each such refinement can be verified independently. In what I have presented, it is always specifications that are refined. To make the mathematics work for us in a large program, it is therefore crucial that we know how to write specifications in a large program. For example, just because one procedure calls several others, need the specification of one involve all the details of the others being called? We hope not, because if this were the case, writing the specification for a procedure in a program with tens or hundreds of thousands of lines of code would be too complex a task for us to manage. Luckily, we are armed with the tool of abstraction (see Chapter 9), with which there is hope to hide the complexity of the implementation details at various levels.

A mechanism used as an aid in abstraction and data hiding is modularity. By organizing the code into separate modules, we are able to hide implementation details of that module from other parts of the program.

Modularity lends itself to separate compilation of modules. Not only can this be a time-saving device when a small change is made in a program, but it also allows code to be written by different groups or vendors and collected into libraries without these groups knowing the details of the code in the other groups' modules.

Similar to the concept of separate compilation of modules is the concept of modular verification—that each module can be verified in isolation from the verification of other modules. Modular verification not only saves time, but is essential to enable libraries to be verified without knowing the details of the programs into which they eventually will be linked. To perform modular verification, we need to learn how to write specifications for procedures that appear in modules.

In this Part, I deal with specifications and modular verification. Having these under control is vital to making the verification of large programs feasible.

Outline

In Chapter 10, I describe the basic problem with abstraction within a setting of an arbitrary number of modules when modular verification is important. I also give a solution to this problem, and discuss related work. In Chapter 11, I give a formal description of this solution for a simple language with modules. In Chapter 12, I prove that the solution in Chapter 11 is sound with respect to modular verification, i.e., that a module verified to be correct using the detailed technique shown in Chapter 11 would also be verified to be correct had all information about the program been taken into account in the verification. In Chapter 13, I discuss some remaining problems.

Chapters 10 and 13 are of interest to those readers concerned with abstraction,
multiple modules, and modular verification. Chapter 11 is of interest to those readers who want to apply the methods of Chapter 10 in practice, for example by building a formal verification system that is unforgiving when it comes to having all the details. Chapter 12 targets a smaller audience, and is provided for completeness. It is intended for semanticists with a thirst for detailed understanding, possibly because they are facing similar proof obligations, and for those who like formal proofs in their own right. However, even the readers with an interest in only Chapters 10 and 13 may be curious to read Section 12.5, which comments on the shape of the proof.
Chapter 10

Specifications in modular programs

In this chapter, I describe a specification problem that arises in the context of modular verification of programs with many modules. I proceed concretely by showing a programming example in which the problem surfaces. Since the goal in this Part is to make program verification feasible in practice, I draw my example from an attempt to write the specification for a real library of input/output streams [9], a library written in a modular, object-oriented style.

After some motivation, I describe the problem, introducing the necessary concepts along the way. Then, I describe a solution to the problem, followed by some discussion. The solution can be viewed as a generalization of classical data refinement [38]. I conclude this chapter by making a connection between the two, and by comparing my solution with other specification languages.

I try to keep the discussion at as high a level as possible while exposing enough details to reveal the problem. A precise description of the solution is given in the next chapter. I assume some familiarity with the concepts of modules and interfaces (or packages) from languages like Modula-3 and Modula-2 (or Ada, respectively).

10.0 Motivation

Programs written in a programming language like Modula-3 are divided into modules and interfaces. A module or interface imports another interface in order to gain access to the entities declared in that interface.

A procedure is declared in an interface, and its implementation is given in a module. This hides the private data of the module from the clients of the interface. The implementation of a procedure declared in an interface may have an effect on the private data in the module. I use data abstraction to combat this problem: The interface describes an abstract view of the behavior of the procedure; the module provides the implementation and prescribes the relation between the abstract view and the concrete one.
As described in the preface to Part III, we are interested in modular verification. That means that one should be able to verify an implementation given only its module and the module's imported interfaces. Having the entire program in view at one time simplifies verification, but is unreasonable to require, because, for example, then a library could not be verified until it were linked with a complete program.

Specifications play a central rôle in this chapter. Recall from Sections 6.0.0 (and 1.7) that a specification includes a frame (given by a modifies clause), which lists those variables that are allowed to be modified. Without modifies clauses, a procedure would be able to modify anything, just as long as the postcondition were met. For example, consider the specification given by $Q, Q_0, Q_1$ in Chapter 5. Without any notion of what the program to be developed is allowed to modify, a perfectly valid implementation would be

$$a[0,0], x, i, j, b := 0, 0, 0, 0, true$$

or, if $M$ and $N$ were not given as constants,

$$M, N, b := 0, 0, false$$

The lack of a construct like modifies yields specifications that are too weak to be useful in the setting of an entire program. Hence, modifies clauses are important.

**Remark 10.0.** The specification language Anna [57] does not feature modifies clauses. When specifications are interpreted by a person (rather than by an unforgiving machine), several conventions are understood (or misunderstood, as it may be). For example, maybe if Swap from Section 7.3 were specified only by

$$\text{ensures } a[i] = a_0[i] \land a[j] = a_0[j]$$

a programmer "knows" not to modify $a$ at any index other than $i$ and $j$. Thus, Anna is useful for formal documentation of a program. Anna annotations admit translation into checks that can be executed at runtime. These are an important and useful aid in finding errors already introduced into a program. They do not, however, establish the absence of errors in a program. For that, verification is needed, so for that, Anna is not, without further restrictions, suitable.

### 10.1 Problem

The problem I'm about to show arises in the presence of three things: data abstraction, *friends* interfaces —which I will describe below—, and modular verification. I won't discriminate between modules and interfaces, because their distinguishing characteristics are not material to the discussion. Instead, I will refer to either as a unit.
10.1.0 Writer Example

The example consists of four units. The first, named \textit{Wr}, shows the declaration of a \textit{writer} class. A writer is an output stream. Examples of writers—that is, of writer subtypes—are \textit{file writers}, which write their output stream to a file in a file system, and \textit{text writers}, which write their output to a text string in memory.

Here’s the first unit.

\begin{verbatim}
unit Wr is
  type T;
  spec var target : T → seq[char];
  spec PutChar(wr : T ; ch : char) is
    modifies target[wr]
    ensures target[wr] = target[wr] ++ ch
end

Unit \textit{Wr} introduces a type \(T\) (see Section 7.2), and declares a specification (abstract) variable \textit{target} (see Chapter 9). \textit{target} is of type \(T → seq[char]\); hence, it designates an (abstract) data field (see Sections 8.1 and 9.2).

\textit{seq[char]} is a type whose values are sequences of characters. Readers familiar with Modula-3 may think of \textit{seq[char]} as Modula-3’s \texttt{TEXT} type, but the exact nature of \textit{seq[char]} is not central to the discussion. I use ++ to denote concatenation.

Unit \textit{Wr} also shows the specification of a procedure, \textit{PutChar}, which takes as parameters a writer and a character. The procedure modifies the target of the writer and ensures that the target, upon termination of the procedure, equals the initial target extended with the given character.

Here’s the second unit.

\begin{verbatim}
unit WrFriends import Wr is
  var buff : Wr.T → seq[char]
  (* Wr.target[wr] = flushed characters of wr ++ buff[wr] *)
end

Unit \textit{WrFriends} is what is known as a \textit{friends interface}. It gets that name from the fact that the unit is intended for import by other units with a close tie to the implementation of writers—a tie only “good friends” are thought to have. (In [9], \textit{WrFriends} is called \textit{WrClass}.)

Unit \textit{WrFriends} imports unit \textit{Wr} to make the declarations in \textit{Wr} visible in \textit{WrFriends}. The import relation is transitive—that is, by importing \textit{Wr}, \textit{WrFriends} also imports all units imported by \textit{Wr} (if there were any). The \textit{scope} of a unit is the set of declarations given in that unit and the units that it imports.

\textit{WrFriends} declares a data field \textit{buff}. Note that I have prefixed type \(T\) from unit \textit{Wr} with “\textit{Wr}.” since it is imported from another unit, a practice common among languages with modules. Note also that \textit{buff} is not a specification variable, but a program variable.

98
The unit ends with a comment describing to programmers the intended usage of field \texttt{buff}. This comment is vital to the discussion; in a sense, the problem I am describing is the problem of formalizing this comment. The idea is the following. The target of each writer is made up of a \textit{flushed} portion and a \textit{buffered} portion. Different writers store their flushed portion in different ways. For example, a file writer keeps the flushed portion on disk, whereas a text writer keeps it in a string in memory. This is the essence of object-oriented programming; different subtypes define their own ways of dealing with this aspect of being a writer. All subtypes share a mechanism for the buffered portion. For that purpose, \texttt{buff} has been introduced. Procedure \texttt{PutChar}, then, simply adds the given character to the end of \texttt{buff} for the given writer. If \texttt{buff} becomes too large as the result of such an operation, \texttt{PutChar} calls out to the particular subtype to perform a flush, i.e., to copy the value of \texttt{buff} into the flushed portion of the writer and clear \texttt{buff}. The call to the particular subtype is done via a method invocation to some \textit{flush} method, not shown here.

Now, let’s move on to a unit declaring a particular kind of writer: text writers.

\begin{verbatim}
unit TextWr import Wr is
  type T <: Wr.T;
  spec Init(wr : T) is
    modifies Wr.target[wr]
    ensures Wr.target[wr] = "";
  spec result : seq[char] ::= Target(wr : T) is
    ensures result = Wr.target[wr]
end

Unit TextWr declares a type \texttt{T} as a subtype of \texttt{Wr.T} (see Section 8.0). That is, objects of type \texttt{TextWr.T} make up some of the objects of type \texttt{Wr.T}.

Unit \texttt{TextWr} also declares two procedures, \texttt{Init} and \texttt{Target}. \texttt{Init} clears the target of a text writer. Its specification states that the target of the given text writer is modified to ensure that it equals the empty string upon termination. Procedure \texttt{Target} returns the target for a given text writer. It does not modify anything in the process.

The fourth and final unit shows the implementation of text writers. In a language like Modula-3, this unit would be a module. However, since I do not distinguish between modules and interfaces, this is simply another unit, whose name I choose to be \texttt{TextWrImpl}.

\begin{verbatim}
unit TextWrImpl import Wr, WrFriends, TextWr is
  var flushed : TextWr.T" -> seq[char];
  rep Wr.target[wr : TextWr.T] is
    Wr.target[wr] = flushed[wr] ++ WrFriends.buff[wr];
  impl TextWr.Init(wr : TextWr.T) is
\end{verbatim}

99
impl result : seq[char] := TextWr.Target(wr : TextWr.T) is
\[ \text{result} := \text{flushed}[\text{wr}] \rightarrow \text{WrFriends.buff}[\text{wr}] \]
end

This unit imports all of the previously introduced units. It declares a variable \textit{flushed}, whose purpose is to contain the flushed portion of text writers. Hence, the representation of \textit{target} for text writers can be given, as stated by the \textit{rep} clause (see Section 9.0): \textit{target} of a text writer is the concatenation of \textit{flushed} and \textit{buff} for that writer.

Unit \textit{TextWrImpl} also gives the implementation of procedures \textit{Init} and \textit{Target}. \textit{Init}, which is supposed to set \textit{target} to the empty string, sets both \textit{flushed} and \textit{buff} to the empty string for that writer, the concatenation of which is the empty string. Procedure \textit{Target}, which is supposed to return \textit{target} of the given text writer, simply returns the concatenation of \textit{flushed} and \textit{buff} for that writer.

10.1.1 Frames and Abstract Variables

Now that I've shown the example, I am able to raise a question. Why is it that procedure \textit{Init}, which is specified to only modify \textit{Wr.target}, is allowed to modify variables \textit{flushed} and \textit{WrFriends.buff}? From Chapter 9, we answer, "Because \textit{flushed} and \textit{WrFriends.buff} are part of the representation of \textit{Wr.target}.

Having decided that, we find ourselves at the edge of the problem. Consider the following client unit, which imports \textit{Wr}, \textit{WrFriends}, and \textit{TextWr}—that is, all the units seen so far except the text writer implementation.

\texttt{unit FaultyClient import Wr, WrFriends, TextWr is}
\texttt{\ :}
\texttt{\text{TextWr.Init(wr); (* Wr.target[wr] = "" *)}}
\texttt{\text{WrFriends.buff}[wr] := \ldots ;}}
\texttt{\text{if TextWr.Target(wr) \neq "" then wrong fi}}
\texttt{\ :}
\texttt{end}

This faulty client calls \textit{TextWr.Init} for some text writer \textit{wr}. After that call, we can conclude that the target of this writer is the empty string. To remind ourselves of that, I have shown this as a comment in the code. Although it is not visible in this scope, the fact that the representation of \textit{target} equals, for text writers, the concatenation of \textit{flushed} and \textit{buff} means that \textit{flushed} and \textit{buff} each equals the empty string at this point in any execution.

The next statement mucks with this writer's \textit{buff} field. Hence, this statement actually affects the value of \textit{target[wr]}. But this fact goes unnoticed to the verification process, to which \textit{target}'s \textit{rep} clause is not visible. Therefore, the modular-verification process treats the update of \textit{buff} as having no effect on \textit{target}.

The last line of the code compares the result value of \textit{TextWr.Target(wr)}, \textit{i.e.}, the value \textit{target[wr]}, with the empty string. If the update of \textit{buff} has no effect on \textit{target},
which, recall, is what a modular-verification process would conclude from the given information, then \texttt{target[or]} and \texttt{“”} will be equal, and the branch that goes wrong is not taken. However, full information about the program—in particular, having full information about the representation of \texttt{target} for text writers—reveals that the update of \texttt{buff} does affect the value of \texttt{target}, and thus the program will go wrong.

We have reached the climax of the exposition. The question we’re facing is: What, in a modular verification of this program, should prevent the program from being verified to be correct?

10.2 Solution

Now that we understand what the problem is, let me move on to its solution.

The solution is to introduce a new specification construct called \texttt{depends}. This will allow dependencies between variables—that one variable is part of the representation of another—to be revealed. The clause

\[
\texttt{depends } a \texttt{ on } c
\]

reveals that abstract variable \( a \) may be represented in terms of variable \( c \).

The \texttt{depends} construct allows a programmer to give part of the representation of an abstract variable. \texttt{depends} does not state what the representation is, but reveals a variable on which it depends.

**Remark 10.1.** Modula-3 programmers familiar with partially opaque types may find that the relation between partial and full type revelations is similar to that of \texttt{depends} and \texttt{rep} clauses.

The use of an abstract variable in a frame can now be defined. I define an abstract variable listed in a \texttt{modifies} clause to be a shorthand for also listing the variables on which the abstract variable depends. In other words, the actual frame is the reflexive transitive closure of the frame given by the programmer. It is sometimes convenient to call this closure the \texttt{downward} closure, to indicate that the closure goes toward the more concrete (or down-to-earth) representation.

I require that all variables on which an abstract variable’s representation depends be given in \texttt{depends} clauses. Thus, in order to mention a variable \( c \) in the \texttt{rep} of an abstract variable \( a \),

\[
\texttt{depends } a \texttt{ on } c
\]

must be visible (or deducible by transitivity) in the scope in which the \texttt{rep} appears.

Finally, a predicate \texttt{Pred} that mentions an abstract variable \( a \) (here written \texttt{Pred}(a)) is interpreted as the same predicate with \( a \) replaced by a function instantiation. The function, which here I shall call \( a' \), is a function of its dependencies. Thus, \texttt{Pred}(a) is defined to mean

\[
\texttt{Pred}( a'(c,\ldots) )
\]

101
Remark 10.2.  There is an important detail, coined *residues*, that plays a rôle here. I postpone discussing this detail until Chapter 11.

10.2.0  Correcting the example

Let me illustrate how depends solves the problem in the example I introduced earlier. Since the representation of Wr.target involves WrFriends.buf, i.e., Wr.target depends on WrFriends.buf, we add the line

\[
\text{depends Wr.target}[t : \text{Wr.T}] \text{ on } \text{buf}[t]
\]

in unit WrFriends. This states that, for every writer \( t \), target[t] depends on buf[t].

Similarly, in unit TextWrImpl, we add the line

\[
\text{depends Wr.target}[t : \text{TextWr.T}] \text{ on } \text{flushed}[t]
\]

It discloses that, for every text writer \( t \), target[t] depends on flushed[t].

As these dependencies are in the scope of TextWrImpl, the rep clause there is permitted to mention buf and flushed. Furthermore, because of the depends clause in WrFriends, the faulty client’s update of buf is perceived as an update of target. The precise effect on target is, however, unknown in the scope of TextWrImpl—all that is known is that an update of buf may cause the value of target to change. Hence, the faulty client will no longer verify.

10.2.1  Visibility Requirement

Let us consider some restrictions that apply in the use of depends clauses. Certainly, there must be some restriction, because otherwise all depends clauses could be written in some distant unit that almost never is imported, and then these clauses would do no good given our goal of modular verification.

To withstand this problem, I require that the dependency between two variables be visible wherever both of those two variables are. I call this rule the visibility requirement.

10.2.2  Benevolent side effects

Consider the following couple of units.

\[
\begin{align*}
\text{unit } D \text{ is} & \quad \text{unit } D^\text{impl} \text{ import } D \text{ is} \\
\text{spec var valid} ; & \quad \text{var hash_table, start, } n, \ldots ; \\
\text{spec var state} ; & \quad \text{depends valid on hash_table, start, } n, \ldots ; \\
\text{spec } P() \text{ is} & \quad \text{depends state on hash_table, start, } n, \ldots ; \\
\text{requires valid} & \quad \text{rep} \ldots ; \\
\text{modifies state} & \quad \text{impl } P() \text{ is} \ldots \\
\text{end} & \quad \text{end}
\end{align*}
\]
These two units show a common paradigm (see also Chapter 13), viz., using, in addition to specification variables describing the values provided by interface \( D \) (here, simply and generally called \( \text{state} \)), a specification variable \( \text{valid} \). \( \text{valid} \) is true just when the values of the internal implementation are in a state that "makes sense", i.e., a state which represents a value of \( \text{state} \). An example of when this might not be the case is before that state is initialized. \( \text{valid} \) is set by some initialization procedure (not shown here) and is required as a precondition by all procedures in the interface.

**Remark 10.3.** Some consider \( \text{valid} \) the object or module invariant and make it implicit (cf. [38, 64, 57, 56]). In the presence of operations that involve many objects, it is not always clear when the object invariant is supposed to hold. At the expense of verbosity, giving \( \text{valid} \) explicitly makes it clear at what points the invariant must hold.

Let's discuss the value of \( \text{valid} \) upon exit from \( P \). The specification of \( P \) requires that \( \text{valid} \) hold upon entry to \( P \); it also states that only \( \text{state} \), not \( \text{valid} \), is modified. Hence, we conclude that \( \text{valid} \) holds upon exit from \( P \).

Very well, let's now focus on the implementation in \( D\text{impl} \). Here, the \text{modifies} clause of \( P \) is interpreted by taking the downward closure of \( \text{state} \), which shows that the variables \( \text{hash_table, start, n, } \ldots \) are allowed to be modified. Note that these variables are the representation also of \( \text{valid} \), and the value of \( \text{valid} \) is not allowed to be changed. This means that the implementation of \( P \) is constrained to modify \( \text{hash_table, start, n, } \ldots \) only in such ways that the value of \( \text{valid} \) is preserved. In the parlance, \( P \)'s side effects must be benevolent. A way to view this is to add

\[
\text{ensures } \text{valid}_0 = \text{valid}
\]

to the specification of \( P \).

### 10.2.3 Authentic Abstractions and Variables

Now that I've introduced benevolent side effects, consider the following program unit.

```plaintext
unit AuthenticityProblem import Wr, WrFriends is
spec var a;
depends a on WrFriends.buff;

(* a = A *) Wr.PutChar(wr, ch) (* a = A ? *)

end
```

This unit declares a specification variable called \( a \), and reveals that \( a \) depends on \( \text{buff} \). For simplicity, I have left subscripts off.

From the specification of \( \text{PutChar} \) (Section 10.1.0), we determine that any side effect on \( a \) is benevolent. That is, despite the fact that both \( \text{target} \) (which is in the
modifies clause of \textit{Wr.PutChar} and \( a \) (which is not) depend on \textit{buff}, \textit{PutChar} does not alter the value of \( a \). But to ensure this, \( a \) (and its representation) must be available at the time the implementation of \textit{Wr.PutChar} is verified. How is this to be guaranteed?

In general, such a guarantee cannot be made. However, if \( a \) is visible anywhere \textit{buff} is, then the implementation (if it modifies \textit{buff} at all) has both \textit{buff} and \( a \) in scope. Due to the visibility requirement, the dependency of \( a \) on \textit{buff} is thus also visible.

\textbf{Remark 10.4.} We don’t need to worry about whether or not \( a \)’s representation is in scope; if it is not, verification of code that modifies \textit{buff} but that must not modify the value of \( a \) will not go through.

If \( a \) is indeed visible anywhere \textit{buff} is, then I say that this dependency is an \textit{authentic abstraction}. A specification variable is \textit{authentic} just when all its (downward) dependencies are authentic abstractions. A variable that is authentic is also said to satisfy the \textit{authenticity property}.

Hence, the condition of benevolent side effects can be guaranteed only for authentic variables. Distinguishing between authentic and unauthentic variables can be done (see next section); however, the distinction is maybe more easily detected by a machine than by a programmer. An option is therefore to simply rule out unauthentic variables, a rule I give the name \textit{authenticity requirement}. That is, the authenticity requirement states that all variables satisfy the authenticity property.

\textbf{Remark 10.5.} The soundness proof in Chapter 12 uses the authenticity requirement. However, the proof can easily be modified to also allow unauthentic variables. To do this, the definition of benevolent side effects in Section 11.1.4 is changed to project not to specification variables but to \textit{authentic} specification variables. The proof is then modified in Sections 12.3.1 and 12.4.1 to use, instead of the authenticity requirement, the fact that all variables in \textit{bb} satisfy the authenticity property.

10.3 Enforcing the requirements

Having posed two requirements on modular programs, a natural concern is: Can these requirements be enforced? The answer is simple: By following a simple convention, both of the requirements follow.

Let me recap the two requirements and state them with respect to two variables \( a \) and \( c \).

\textbf{Visibility requirement} If \( a \) depends on \( c \), then this dependency must be visible anywhere both \( a \) and \( c \) are.

\textbf{Authenticity requirement} If \( a \) depends on \( c \), then \( a \) must be visible anywhere \( c \) is.
Observe that, if the dependency of \( a \) on \( c \) is declared either in the unit that declares \( a \) or in the unit that declares \( c \), then the visibility requirement follows—then, any unit that imports both \( a \)'s unit and \( c \)'s unit also imports the dependency. If declared in \( a \)'s unit, the dependency is not authentic; if declared in \( c \)'s unit, the dependency is authentic. Thus, in conjunction, the two requirements can be stated as one.

Declaration \textbf{depends} \( a \) \textbf{on} \( c \) is placed in the unit that declares \( c \).

This is a condition that is easy to check using the declarations only of the unit being verified and its imports.

\textbf{Remark 10.6.} The convention is a sufficient condition for the two requirements to hold. However, in languages where cyclic imports are allowed (this excludes, \textit{e.g.}, Modula-3), the requirements can be satisfied without adhering to the convention.

\textbf{Remark 10.7.} Not only does the convention capture the two requirements concisely, it is also easier to teach to a programmer than the two requirements are separately. However, I do not regret having introduced the two requirements separately, because of the now manifest opportunity to replace each independently (see, \textit{e.g.}, Remark 10.8) and because of their separate rôles in the soundness proof (see Section 12.5 and Remark 10.5).

### 10.4 Soundness of modular verification

I introduced \textbf{depends} and motivated two requirements for its use. This invites the question: Are these two requirements enough? To answer that question positively, we must prove the \textit{soundness} of modular verification with respect to these requirements. This means that if the verification of each unit goes through, then the verification of the whole program would go through, provided the program satisfies the two requirements.

Such a result allows the verification of a procedure implementation with respect to its specification to be performed in the scope of the unit in which the implementation occurs, \textit{i.e.}, using only the information from that unit and its imports.

Since soundness does hold (see Chapter 12), one may wonder about completeness. Completeness means that if the program could be verified correct given \textit{all} its units at once, then each unit can be verified correct by itself. There is no hope of achieving this. For example, in the faulty client example I showed, if the line that mucks with \texttt{buff} actually sets \texttt{buff} to the empty string, then \texttt{target} remains unchanged. But the only way to determine that \texttt{target} indeed remains unchanged is to have \( a \)'s \texttt{rep} clause in scope, and requiring that violates the essence of data hiding.

So, it is not completeness in which we're interested. Instead, we're interested in \textit{adequacy}. That is, we want to be able to specify and verify programs we care about. I revisit the issue of adequacy in Chapter 13.
**Remark 10.8.** For the record, this remark discusses a predecessor of the authenticity requirement that proved to be inadequate.

Before inventing the authenticity requirement, I was using what I called the *forest requirement*. Consider the graph whose vertices are the variables and whose directed edges correspond to the given dependencies among variables. Then, the forest requirement states that this graph is acyclic and that the transitive reduction of the graph is a forest. The transitive reduction of a directed graph $G$ is the graph with the fewest edges among those graphs whose transitive closure equals the transitive closure of $G$.

**Remark 10.9.** A transitive reduction of a graph is unique if the graph is acyclic.

A forest is a set of trees, and a tree is a directed graph in which every vertex has in-degree at most one (see, *e.g.*, [53]).

For example, Figure 10.0 shows three graphs. Graph (ii) is the transitive reduction of graph (i), and graph (iii) is its own transitive reduction. Graphs (i) and (ii) satisfy the forest requirement, whereas graph (iii) does not.

I had to give up the forest requirement, because it proved inadequate. For example, the dependencies of the common paradigm explained in Section 10.2.2 take the form of Figure 10.0(iii). (The specification language Aspect [41], in some sense, can only handle dependencies satisfying the forest requirement—see its Section 8.1.)

10.5 A generalization of classical data refinement

In this section, I compare my *depends* solution with classical data refinement [38] (see also Section 9.1).

10.5.0 MODELING CLASSICAL DATA REFINEMENT

Any program specified and refined with the techniques of [38] can also be specified and refined in my model, because [38] is essentially just a restriction on my model:
There are just two units, one for the specification and one for the implementation. Call these $S$ and $M$, respectively. $S$ then declares all abstract variables, and gives the specifications in terms of those. $M$ imports $S$, declares the concrete variables, and specifies the representation of the abstract variables. In this simple world, the dependencies can be inferred from the rep clause. Finally, the implementation only refers to concrete variables.

10.5.1 Explicit Functions

Like [38], I treat abstract variables as functions. However, an important distinction is that I prove refinements by making these functions explicit. Classical data refinement instead introduces the program $\text{SwitchToAbstract}$ and does the dance with two state spaces. It is also crucial that the visible dependencies are shown explicitly. (The fact that dependencies that are not in scope need not be present is proven by the soundness proof in Chapter 12.)

For example, consider variables $a$ and $c$, where $a$ depends on $c$. If dependencies are not shown explicitly, then calculating the weakest precondition of a statement $c := 2$ with respect to a postcondition $a \geq 0$ is done as follows. (For simplicity, I leave off the exceptional postcondition in this discussion.)

$$\begin{align*}
wp.(c := 2).(a \geq 0) &= \{ := \} \\
&= (a \geq 0)[c := 2] \\
&= \{ \text{substitution} \} \\
&= a \geq 0
\end{align*}$$

Notice that, without any further information, the substitution in the last step is performed on the basis that $c$ does not appear the expression $a \geq 0$. The last line of the calculation is not the desired result, because an update of $c$ affects the value of $a$. Thus, proving a refinement this way calls for some other measure, like the business with $\text{SwitchToAbstract}$.

In contrast, translating $a$ into a function $a'$ with dependencies shown explicitly yields the following calculation.

$$\begin{align*}
wp.(c := 2).(a'(c, \ldots) \geq 0) &= \{ := \} \\
&= (a'(c, \ldots) \geq 0)[c := 2] \\
&= \{ \text{substitution} \} \\
&= a'(2, \ldots) \geq 0
\end{align*}$$

Note that not only is the last line the expression we want, but both the weakest precondition of an assignment and the rules for substitution are the same as those we would use to prove programs without abstraction.

So, in summary, instead of changing the notion of refinement, predicates are converted with respect to visible dependencies. Then, refinement, assignment, and substitution are those that we are used to.
10.5.2 Inferring dependencies

Languages like Modula-3, Ada, and Modula-2 provide a more flexible model of units than Simula [14], on which [38] is based. These languages allow the state of an implementation to be distributed across multiple units. Only when all of these units and the representation function of an abstract variable are visible can the dependencies of the abstract variable be inferred. Since a unit, in general, imports but a proper subset of these units, automatically inferring dependencies is no longer possible. The depends construct solves this problem.

10.6 Other specification languages

In this section, I discuss how some existing specification and programming languages deal with the problem I have described.

10.6.0 ANNA

The Anna specification language [57] does not provide a construct like modifies; hence, it is not equipped, as is, to handle the kind of verification that I have described (see Remark 10.0).

Moreover, Anna allows the body of a procedure to use a different specification than the so-called visible declaration given at the procedure declaration. The motivation for this is that the procedure declaration appears in the package declaration and the procedure body appears in the package body. Since the scope of the body contains the private data of the package, one may want in the package body to extend the visible specification so that it also specifies the behavior in terms of the private data. Anna lets the conjunction of the postconditions appearing in these two specifications be the prevailing postcondition. This is sound, because the implementation is allowed to provide a more specific behavior than originally specified. Stated differently, this is sound because a specification, as a predicate transformer, is antimonotonic in its postcondition. However, Anna also takes the prevailing precondition to be the conjunction of the two given preconditions. That is not sound, because the specification is monotonic in its precondition, not antimonotonic—external clients of the procedure have no way to establish the strengthened postcondition, let alone know what it is.

Finally, Anna provides a mechanism called package states. It is used as an abstraction mechanism. Anna requires that package states be used only for packages that satisfy the Hidden State Principle, which essentially states that the package implementation cannot make use of the program state outside the package (this includes using, directly or indirectly, global variables declared in other packages). Thus, this kind of data abstraction is like that in classical data refinement.

Any one of these three shortcomings means that Anna does not provide a solution for the problem I describe in this chapter.
10.6.1 Penelope

Penelope [32, 62] is an interactive environment for developing and verifying Ada programs. It is based on Larch [33] but borrows much of its syntax from Anna. Penelope's verification is sound across modules. However, in addition to not supporting pointer types (and Ada doesn't feature object types), Penelope does not support abstraction and thus ducks the problem altogether.

10.6.2 CLU

Abstraction in CLU (see [55]) is done under the assumption that the implementation does not "share" its values (so-called objects) with clients of the interface. Such sharing is referred to as rep exposure. That phenomenon is easily produced, however, and then verification is no longer sound. Similar issues are treated in Section 13.2.

CLU, like Simula and classical data refinement, does not provide language features that allow an implementation to be distributed across modules; hence, friends interfaces are not even a possibility.

10.6.3 Summary

Both Anna and Penelope dodge the problem I have described, because they do not even provide the necessary specification features in the presence of which the problem arises. Both Anna and Penelope are, of course, useful in their own rights, but neither can claim to provide an answer to specifying and verifying modular programs.

What both Anna and Penelope provide, however, which I have not discussed, is the theory packages (see [33]) one needs to write elaborate specifications, e.g., a library of axioms and theorems regarding tree structures. These issues, and the issues of how theorems are proven automatically in the presence of these, are orthogonal to my discussions. Both Anna and Penelope aim at achieving full specifications. With extended static checking as one's goal, theory packages play a diminished rôle.
In this chapter, I formalize the idea of \textit{depends}. I do so by first introducing a simple programming notation for units (modules and interfaces). Then, I define the relation \textit{Refine}, which is true just when a program implements (refines) its specification. The Prolog-style definition of \textit{Refine} suggests a precise operational way to generate the formulas, known as \textit{verification conditions}, that need to be proven in order to establish the correctness of the refinement.

The previous chapter provides an informal discussion of what this chapter formalizes, but with one notable exception: \textit{residues}. Residues are swept under the rug in Chapter 10, because they would have diverted attention from the central ideas in that chapter. In this chapter, I do include residues; in fact, I conclude this chapter by showing the importance of residues. The next chapter, which proves the soundness of modular verification, builds on the definitions presented in this chapter.

The reader will notice that this chapter changes gears to more formality from the previous chapter. In that sense, this chapter also serves as a preparation for the next chapter, which contains the formal proof of soundness. Another change from the previous chapter is that, in order to focus on the relevant details, I have left out many bells and whistles from the richer notation used in Chapter 10.

11.0 \hspace{1em} A notation for modular programs

In this section, I present the syntax of a programming notation with units. I also give some definitions, many of which are review from previous chapters. The reason for showing the syntax of the notation is that this makes everything explicit and lends itself to inductive definitions of some of the relations defined in the rest of the chapter.

A program consists of a number of \textit{units}.

\textbf{Remark 11.0.} Programming languages like Modula-3 typically provide two kinds of units, modules and interfaces. Since I do not distinguish between the two, a Modula-3 program is a restriction of what I present.
The syntax of each unit is given by the grammar

\[
<\text{Unit}> ::= \text{unit} \ <id> \ [\text{import} \ <\text{idlist}>] \ {\{\text{Decl}\}} \end
\]

\[
<\text{Decl}> ::= \\
\quad \text{var} \ <\text{idlist}> \\
\quad | \text{spec var} \ <\text{idlist}> \\
\quad | \text{spec} \ <id> \text{ is modifies} \ <\text{idlist}> \text{ requires} \ <\text{Pred}> \text{ ensures} \ <\text{Pred}> \\
\quad | \text{impl} \ <id> \text{ is} \ <\text{Command}> \\
\quad | \text{depends} \ <id> \text{ on} \ <\text{idlist}> \\
\quad | \text{rep} \ <id> \text{ is} \ <\text{Pred}> \\
\]

where \{<\text{Decl}>\} means any number of occurrences of \(<\text{Decl}>\), separated by semicolons.

An \(<id>\) is an identifier, and an \(<\text{idlist}>\) is a nonempty, comma-delimited list of \(<id>\)’s. The declarations \text{unit}, \text{var}, \text{spec var}, and \text{spec} introduce new \(<id>\)’s. The declared \(<id>\)’s in a program must be unique.

\textbf{Remark 11.1.} Modular programming languages typically require that top-level identifiers within one unit be unique. Identifiers declared in imported units are then prefixed by the name of the unit in which they are declared, like “Wr.” in the examples from the previous chapter. For local variables, variables are distinguished by scope rules. In this chapter, however, I assume that such resolutions have already been done, and that all identifiers are unique.

\text{unit} declares the succeeding identifier to be a unit. The \(<\text{idlist}>\) after \text{import} must list only units. The set of units reachable via imports from a unit \(A\) is called the import closure of \(A\), denoted ImportClosure(\(A\)). Note that \(A \in \text{ImportClosure}(A)\).

Every identifier mentioned in a unit \(A\) must be \text{visible} in \(A\), meaning that it has its declaration in some unit in ImportClosure(\(A\)).

Identifiers introduced by \text{var} or \text{spec var} are called \text{variables}. I distinguish between the two kinds of variables by referring to the former as program variables and the latter as specification variables (see Section 9.0). Identifiers introduced by \text{spec} are called \text{procedures}.

The \(<\text{idlist}>\) in \text{spec} and \text{depends} must list only variables. The \(<id>\) in the \text{impl} clause must be a procedure, and the \(<id>\)’s in the \text{depends} and \text{rep} clauses must be specification variables.

The \(<\text{idlist}>\) in the \text{spec} clause is called the \text{frame}, and the two \(<\text{Pred}>\)’s are called the \text{precondition} and \text{postcondition}, respectively. The postcondition may mention a variable subscripted with 0, an \text{initial-value} variable. This refers to the value of the variable upon entry to the procedure (see Section 1.7).

The \text{impl} declaration gives an implementation of a procedure. A procedure has exactly one implementation. Note that the name of a procedure is declared by a \text{spec} clause, whereas \text{impl} just associates an implementation with an already declared
procedure identifier. Consequently, the specification of a procedure must be visible in the unit that gives the implementation.

The **depends** and **rep** declarations are described later. For each of these, \(<id>\) must be a specification variable. Only one **rep** clause per specification variable is allowed. Note that, as with **impl**, the (**spec var**) declaration of the identifier to which depends and rep pertain must be visible in the unit that declares the depends or rep clause.

As in Chapter 7, I reserve the character * in identifiers, so that new unique identifiers needed in the proof can easily be constructed.

Because exceptions are orthogonal to the present discussion, I omit them and consider programs with only one postcondition.

**Remark 11.2.** Using as that postcondition a partitioned predicate (see [61] or Remark 3.3), the discussion also applies to exceptional programs. The addition of **raise** and **<** then only affects the proof of property (12.12), found in Section 12.4.2.

A guarded command has the following syntax.

\[
<\text{Command}> ::= \\
\hspace{1em}\text{skip} \\
\hspace{1em}\text{wrong} \\
\hspace{1em}<id> \text{ " := " } <\text{expr}> \\
\hspace{1em}\text{call } <id> \\
\hspace{1em}<\text{Command}> \text{ " ; " } <\text{Command}> \\
\hspace{1em}<\text{Command}> \text{ " [] " } <\text{Command}> \\
\hspace{1em}<\text{bool-exp}> \text{ " \rightarrow " } <\text{Command}> \\
\hspace{1em} \text{[ " } <id> \text{ " * " } <\text{Command}> \text{ " ] " }
\]

The **<id>** in the assignment statement and the identifiers mentioned in **<expr>** and **<bool-exp>** must be program variables. The **<id>** in the procedure call must denote a procedure. The **<id>** in the block statement introduces a new identifier that can be used in the subsequent **<Command>**. This **<id>** must be distinct from all other **<id>**'s in the program.

**Remark 11.3.** One of the simplifications made from the richer notation previously presented is the absence of loops. Note that programmers can still declare and call tail-recursive procedures.

### 11.1 Definitions

I now give all the definitions necessary to define **Refine**. I start bottom-up, and will end with the definition of **Refine** itself.

I define an environment to be the set of declarations in a set of units closed under imports. Stated differently, an environment is the union of a set of units closed under imports. Most of the definitions take an environment as a parameter.
I treat sets and lists of variables to be synonymous, and do not distinguish between single variables and lists of variables. So, for example, for a list of variables \( w \), the expression \( w_0 = w \) is a shorthand for
\[
(\forall v \mid v \in w \Rightarrow v_0 = v)
\]

### 11.1.0 Dependencies

I start by defining the dependency relation, named \( \text{Depends} \). For any variables \( a \) and \( c \) and environment \( E \),

\[
\text{Depends}(a, c, E)
\]

holds just when \( a \) depends on \( c \) in \( E \). This means that \( E \) contains enough \text{depends} declarations so that the dependency of \( a \) on \( c \) is deducible by reflexivity and transitivity.

Note that every variable depends on itself. Also, note that \( \text{Depends} \) is monotonic in its last argument, i.e., for any \( a, c, E, E' \),

\[
E \subseteq E' \Rightarrow (\text{Depends}(a, c, E) \Rightarrow \text{Depends}(a, c, E'))
\]  

(11.0)

### 11.1.1 Resolve

\( \text{Resolve} \) of a list of variables \( w \) and an environment \( E \) is the downward closure (or resolve set) of the variables in \( w \), as is visible in \( E \).

\[
\text{Resolve}(w, E) = \\
\{ v, x \mid v \in w \land \text{Depends}(v, x, E) \Rightarrow x \} \cup \\
\{ v, x \mid v_0 \in w \land \text{Depends}(v, x, E) \Rightarrow x_0 \}
\]

It will be convenient to allow \( \text{Resolve}(w_0 = w, E) \) as a shorthand for

\[
\text{Resolve}(w_0, E) = \text{Resolve}(w, E)
\]

i.e., the formula that states that the value of each variable in \( \text{Resolve}(w, E) \) is unchanged.

**Properties of \( \text{Resolve} \)**

From the monotonicity of \( \text{Depends} \) (11.0), we have that \( \text{Resolve} \) is monotonic in both arguments, i.e., for any \( w, w', E, E' \),

\[
E \subseteq E' \land w \subseteq w' \Rightarrow \text{Resolve}(w, E) \subseteq \text{Resolve}(w', E')
\]

(11.1)

In rewriting an expression describing an element of a resolve set, the following property will come in handy. For any \( c, w, E \),

\[
(c \in \text{Resolve}(w, E)) = \langle \exists a \mid a \in w \Rightarrow \text{Depends}(a, c, E) \rangle
\]

(11.2)

113
We also have, for any $w$ visible in some environment $E$,

$$w \subseteq \text{Resolve}(w, E) \subseteq E \quad .$$

The fact that \text{Resolve} is a closure is reflected in the following property. For any $a, c, w, E$,

$$\text{Depends}(a, c, E) \land a \in \text{Resolve}(w, E) \Rightarrow c \in \text{Resolve}(w, E) \quad .$$

(11.4)

11.1.2 \hspace{1em} \textbf{Functionalize}

I now formally explain the interpretation of specification variables. A specification variable is a function on the program variables. For every specification variable $a$, I introduce a function $f*a$. (Recall that $*$ is a special character, so that $f*a$ is just an identifier. This identifier is distinct from all programmer-declared identifiers and is uniquely determined from the name $a$.)

\textbf{Example}

Let me start with an example. Consider the following declarations.

\begin{quote}
\texttt{spec var a ; var c ; depends a on c ; rep a is a = c^2}
\end{quote}

Let's ponder the interpretation of expression $a = 9$. The $a$ in this expression is \textit{functionalized} into

$$f*a(a, c)$$

and $a$'s \texttt{rep} declaration is functionalized into

$$\forall a, c \triangleright f*a(a, c) = c^2 \} \quad .$$

Thus, the expression $a = 9$ is shorthand for $f*a(a, c) = 9$, which, using the functionalized representation of $a$, simplifies to $c^2 = 9$. Hence, a guarded command $c := 3$ will establish $a = 9$, and so will $c := -3$.

Notice the occurrence of the symbol $a$ as a parameter in the instantiation of $f*a$. This $a$ is called the \textit{residue} of the specification variable $a$. The inclusion of residues is important for soundness; they must not be left out, as I show in Section 11.3.

\textbf{Definition}

\textbf{Functionalize} distributes over all connectives. I show its effect on atomic formulas, quantifiers, and substitution functions.

\textbf{Functionalize} has no effect on constants or program variables. Hence, for any constant $c$ and (possibly initial-value) program variable $v$, we have the following rules.

$$\begin{align*}
\text{Functionalize}(c, E) &= c
\end{align*}$$
\[ \text{Functionalize}(v, E) = v \]

For specification variables \( a \), \text{Functionalize} is more interesting.

\[ \text{Functionalize}(a, E) = f \star a(\text{Resolve}(a, E)) \]

\[ \text{Functionalize}(a_0, E) = f \star a(\text{Resolve}(a_0, E)) \]

The order of the arguments in \( f \star a \)'s argument list is a function of \( E \). That is, if specification variable \( a \) is functionalized more than once with respect to the same environment, the order of the arguments of function \( f \star a \) remains the same. For a different environment, however, the order may be different, and the number of arguments may also differ since different environments provide different visibility of dependencies. In the next chapter, I consider the functionalization with respect to two environments, one of which is a subset of the other. I then assume that the common dependencies are listed first, and are listed in the same order for both environments.

Now for quantifiers.

\[ \text{Functionalize}(\{ Q \ w \mid R \triangleright T \}, E) = \]
\[ \{ Q \ \text{Resolve}(w, E) \mid \text{Functionalize}(R, E) \triangleright \text{Functionalize}(T, E) \} \]

Finally, for the substitution function,

\[ \text{Functionalize}(Q[x := y], E) = \]
\[ \text{Functionalize}(Q, E)[\text{Resolve}(x, E) := \text{Resolve}(y, E)] \]

where the first of the two occurrences of \( \text{Resolve} \) takes a list of variables \( x \) and produces \( x \) in which each variable is replaced by its list of dependencies, and similarly for \( \text{Resolve} \) of the list of expressions \( y \). Thus, for example, given

\[ \text{spec } \text{var } a, b \quad ; \quad \text{var } c, d, e \quad ; \quad \text{depends } a \text{ on } c, d \quad ; \quad \text{depends } b \text{ on } d, e \]

in \( E \), \text{Functionalize}(Q[a_0, b_0 := a, b], E) simplifies to

\[ \text{Functionalize}(Q)[a_0, c_0, d_0, b_0, d_0, e_0 := a, c, d, b, d, e] \]

Note that this expression contains two occurrences of \( d_0 := d \). Instead of complicating the definition to avoid this, I simply take its meaning to be

\[ \text{Functionalize}(Q)[a_0, c_0, b_0, d_0, e_0 := a, c, b, d, e] \]

that is, duplicate substitutions are ignored. Note that, provided the original substitution contains no duplicate symbols on the left-hand side, the functionalized substitution will contain no conflicting substitutions (by a conflicting substitution, I mean, e.g., \( x, x := y, z \), where \( y \) and \( z \) differ).
**Functionalize and substitutions**

Recall from the meaning of units and imports, that if a unit $A$ imports another unit $B$, then $A$ may reference identifiers declared in $B$. Conversely, if $A$ does not import $B$, then $A$ cannot reference any identifiers declared in $B$. I use this fact implicitly.

In a similar way, I implicitly use the fact that the representation of a specification variable $a$ refers only to those variables on which $a$ depends. It is this fact that justifies the correctness of the definition of Functionalize: Since a specification variable $a$ is replaced by an expression that explicitly shows all visible variables on which $a$ depends, substitution in functionalized expressions works properly.

To illustrate my point, let me offer an example. Consider the expression $a = 9$ appearing in the example earlier in this section. The functionalization of that expression is $f \ast a(a, c) = 9$. Applying the substitution $c := 3$ (i.e., calculating the weakest precondition of $c := 3$ with respect to postcondition $f \ast a(a, c) = 9$), we then get $f \ast a(a, 3) = 9$. Subsequently applying the representation of $a$, we get $3^2 = 9$, which simplifies to true.

Now consider what happens if, hypothetically, the representation of $a$ also mentioned a variable $d$, e.g.,

$$\text{rep } a \text{ is } a = c^2 + d$$

where $a$ is not declared to depend on $d$. Then, the functionalization of $a = 9$ would still be $f \ast a(a, c) = 9$. Thus, a substitution $d := d + 1$ has no effect on this expression, since $d$ does not occur in it. However, the update of $d$ does have an effect on the value of $a$, so here, substitution in the functionalized expression does not work properly.

In short, the restriction that a rep clause for a variable $a$ can only refer to those variables on which $a$ depends, justifies why performing substitutions in functionalized expressions is correct.

**Leibniz**

Note that for any environment $E$ that contains a representation of a specification variable $a$,

$$[\text{Functionalize}(a_0 = a, E) \Leftarrow \text{Resolve}(a_0 = a, E)] \quad (11.5)$$

That is, if all variables on which $a$ depends are unchanged, then the value of $a$ functionalized is unchanged, too. Applied to the example above, this line reads

$$f \ast a(a_0, c_0) = f \ast a(a, c) \Leftarrow a_0 = a \land c_0 = c$$

This phenomenon is known as Leibniz's Rule.
11.1.3 RepPreds

RepPreds of an environment \( E \) is the conjunction of "axioms" formed from the rep declarations in \( E \). For each

\[
\text{rep } a \text{ is } P
\]

occurring in \( E \), the corresponding axiom is

\[
\langle \forall w \triangleright \text{Functionalize}(P, E) \rangle,
\]

where \( w = \text{Resolve}(a, E) \).

We can rewrite this expression as follows.

\[
\langle \forall \text{Resolve}(a, E) \triangleright \text{Functionalize}(P, E) \rangle = \text{Functionalize}(\langle \forall a \triangleright P \rangle, E)
\]

Let the first argument to Functionalize in the last line be denoted by \( \text{RepAxiom}(a, E) \), and let \( \text{RepAxioms}(E) \) be the conjunction of \( \text{RepAxiom}(a, E) \) for each \( a \) for which a rep is declared in \( E \). Since Functionalize distributes over conjunction, we thus have

\[
\text{RepPreds}(E) = \text{Functionalize}(\text{RepAxioms}(E), E).
\]

11.1.4 GetSpec

Given a procedure \( id \) and an environment \( E \), GetSpec\((id, E)\) denotes the specification of \( id \), desugared according to \( E \).

\[
\text{GetSpec}(id, E) = w : [\text{Pre, Post}]
\]

where

\[
(\text{modifies frame requires } P \text{ ensures } Q) = \text{Lookup}(id, E) \land
\]

\[
w = \text{Resolve}(\text{frame}, E) \land
\]

\[
\text{Pre} = \text{Functionalize}(P, E) \land
\]

\[
\text{Post} = \text{Functionalize}(Q \land \text{BenSideEffects}(w, E), E)
\]

For a procedure \( id \), Lookup\((id, E)\) returns the specification of \( id \) as declared in \( E \).

BenSideEffects is the conjunct added to the postcondition when desugaring a specification. This ensures that the procedure's effect on specification variables not listed in the resolved frame can be described as benevolent side effects (see Section 10.2.2), meaning their values do not change.

Before giving the exact definition of BenSideEffects, let me illustrate the idea with an example. Consider the declarations

\[
\text{spec } \text{var } a, b ; \ \text{var } c ; \ \text{depends } a \text{ on } c ; \ \text{depends } b \text{ on } c
\]

117
and the procedure specification

\textbf{spec $p$ is modifies $a$ requires true ensures $a = 9$} .

GetSpec of $p$ in this environment desugars the specification of $p$ into

\[ a, c : [\text{true, Functionalize}(a = 9 \land b_0 = b, E)] \]

which simplifies to

\[ a, c : [\text{true, } f\star a(a,c) = 9 \land f\star b(b_0,c_0) = f\star b(b,c)] \]

Here, $a$ has been replaced by its resolve set $a, c$, and the pre- and postconditions have been functionalized. Moreover, the postcondition has an extra conjunct which, in effect, states that the procedure may only modify $c$ in such a way as to preserve the value of $b$, that is, preserve the value of $f\star b(b,c)$.

Now for the definition of BenSideEffects.

\[ \text{BenSideEffects}(w, E) = \]
\[ b_0 = b \]
\[ \text{where} \]
\[ b = \text{SpecVarProjection(BenSideSet}(w, E)) \]

In words, BenSideEffects states that $b$ is unchanged, where $b$ denotes the list of variables prescribed by BenSideSet. In the following, $v$ ranges over variables.

\[ \text{BenSideSet}(w, E) = \]
\[ \{ v \mid v \in E \setminus w \Rightarrow v \} \]

Stated differently, if $V$ is the set of variables in $E$, then BenSideSet$(w, E)$ is simply $V \setminus w$.

\textbf{Remark 11.4.} The set BenSideSet can be quite large. Its size can be reduced by using (11.9). Then, BenSideSet$(w, E)$ can be simply those variables whose downward closure overlaps with $w$. For an already resolved list $w$, the definition is then written

\[ \text{BenSideSet}(w, E) = \]
\[ \{ v \mid v \in E \setminus w \land w \cap rv \neq \{ \} \Rightarrow v \} \]
\[ \text{where} \]
\[ rv = \text{Resolve}(v, E) \]

Finally, SpecVarProjection of a set $s$ is the set of specification variables in $s$.

\[ \text{SpecVarProjection}(s) = \]
\[ \{ a \mid a \in s \land a \text{ is a specification variable } \Rightarrow a \} \]

Note that SpecVarProjection, being a projection, is monotonic in its argument, i.e., for any sets of variables $s$ and $s'$,

\[ s \subseteq s' \Rightarrow \text{SpecVarProjection}(s) \subseteq \text{SpecVarProjection}(s') \]

118
11.1.5 Weakest liberal precondition

Commands are modeled by their weakest liberal preconditions. The variables mentioned in program statements are all program variables, so their interpretation does not depend on the environment. However, procedure calls are defined in terms of their specifications, which may refer to specification variables. Therefore, the interpretation of procedure calls is affected by the environment.

For any command gc and predicate Q written in an environment E, \( wlp(gc, Q, E) \) gives the weakest liberal precondition of the command interpreted in E.

\[
\begin{align*}
  wlp(gc, Q, E) &= \ \\
  wlp.Env(gc, E).Q
\end{align*}
\]

where the second occurrence of \( wlp \) denotes the weakest liberal precondition from Chapter 1 but dropping the exceptional postcondition (i.e., implicitly letting it be false; cf. Section 6.2).

\( Env(gc, E) \) is the interpretation of a command \( gc \) in an environment \( E \). \( Env(gc, E) \) is defined as \( gc \) in which all occurrences of procedure calls call \( p \) are replaced by the specification statement \( w : [P, Q] \), where \( w : [P, Q] = GetSpec(p, E) \).

From the definition of \( wlp \) for each statement, we can show that \( wlp \) is positively finitely conjunctive (in its second argument), i.e., for any predicates \( P \) and \( Q \) and environment \( E \),

\[
  wlp(gc, P \land Q, E) = wlp(gc, P, E) \land wlp(gc, Q, E)
\]

(\( wlp \) is actually universally conjunctive in general (see also Remark 6.0), but we only need conjunctivity for nonempty finite bags of predicates.) Note that conjunctivity implies monotonicity [21].

11.1.6 Target

The set of targets of a command is the set of all variables that, according to a naïve syntactic inspection of the command, are potentially updated by an execution of the command.

In \( Target(gc, E) \), \( gc \) is a guarded command and \( E \) is an environment. It equals the resolve set of \( RawTarget(gc, E) \).

\[
\begin{align*}
  Target(gc, E) &= \ \\
  Resolve(RawTarget(gc, E), E)
\end{align*}
\]

On account of (11.3), we thus have

\[
  RawTarget(gc, E) \subseteq Target(gc, E)
\]

\( RawTarget \) is defined inductively over the syntax of the command. Only two commands update variables: the assignment statement and the procedure call.
\textit{RawTarget} is the unfunctionalized set of variables that are possibly updated by a command.

\begin{align*}
\text{RawTarget}(v := e, E) &= \\
&= \{v\}
\end{align*}

\begin{align*}
\text{RawTarget}(\text{call } p, E) &= \\
&= fr \\
\text{where} \\
(\text{modifies } fr \text{ requires } P \text{ ensures } Q) &= \text{Lookup}(p, E)
\end{align*}

I also show the definition of \textit{RawTarget} for a block statement.

\begin{align*}
\text{RawTarget}(\left[ x \bullet s \right], E) &= \\
&= \text{RawTarget}(s, E) \setminus \{x\}
\end{align*}

For all other primitive statements, \textit{RawTarget} equals the empty set. \textit{RawTarget} of other statement compositions equals the union of \textit{RawTarget} for the constituent statements.

Note that since \textit{RawTarget} does not perform functionalization (the environment is used only to retrieve the raw specification of procedures), \textit{RawTarget}(gc, E) is the same for all environments \( E \) in which gc can be written.

### 11.1.7 Refine

Finally, I get to the definition of \textit{Refine}.

\begin{align*}
\text{Refine}(\text{impl id is gc}, E) &= \\
&= \text{Rep} \Rightarrow (w : [Pre, Post] \sqsubseteq \text{impl}) \\
\text{where} \\
&= w : [Pre, Post] = \text{GetSpec}(id, E) \land \\
\text{impl} = \text{Env}(gc, E) \land \\
\text{Rep} = \text{RepPre}ds(E)
\end{align*}

Remember from (1.18) that \( s \sqsubseteq t \) (here for \( wlp \)) is defined as

\[
(\forall R \triangleright [wlp.s.R \Rightarrow wlp.t.R])
\]

where \( R \) ranges over all predicates. This formula may look a bit intimidating, since it involves a quantification over all predicates. Using a result from the next section, I now transform this expression into an equivalent one. This alternative formulation may be preferable to work with in certain circumstances, \textit{e.g.}, in automated theorem proving. It is also the formulation I use in the proof in Chapter 12.

First, note that \textit{Rep} has no free variables, \textit{i.e.}, it is a boolean scalar. We calculate,

\begin{align*}
\text{Rep} \Rightarrow (w : [Pre, Post] \sqsubseteq \text{impl}) &= \\
&= \left\{ (1.18): \text{def. of } \sqsubseteq \right\}
\end{align*}
\[ \text{Rep} \Rightarrow \langle \forall R \triangleright \{ w : \{ \text{Pre}, \text{Post} \} \rangle, R \Rightarrow \text{wp}(R) \rangle \]

\[ = \begin{cases} (1.16): \text{wp} \text{ of specification statement, where } v_0 \text{ denotes the } \\
\text{initial-value variables in } \text{Post} \end{cases} \]

\[ = \begin{cases} \Rightarrow \text{ over } \forall, \text{ since } \text{Rep} \text{ is a boolean scalar} \end{cases} \]

\[ = \begin{cases} \Rightarrow \text{ over } \{ \}, \text{ since } \text{Rep} \text{ is a boolean scalar} \end{cases} \]

\[ = \begin{cases} \text{ pred. calc.} \end{cases} \]

\[ = \begin{cases} (1.16): \text{wp} \text{ of specification statement} \end{cases} \]

\[ = \begin{cases} (1.18): \text{def. of } \sqsubseteq \end{cases} \]

\[ = \text{w} : [\text{Rep} \land \text{Pre}, \text{Post}] \sqsubseteq \text{impl} \]

Define \( t \) and \( m \) by

\[ t = \text{Target(impl, E)} \land m = t \setminus \text{w} \]

and let \( z \) denote the list of all variables. \( m \) is the list of variables that, from a syntactic inspection of \( gc \), are possible targets of \( gc \), but which, according to the procedure specification, are not allowed to be modified. We then have,

\[ \text{w} : [\text{Rep} \land \text{Pre}, \text{Post}] \sqsubseteq \text{impl} \]

\[ = \begin{cases} (11.9), \text{ since } \text{w} \text{ and } m \text{ are disjoint} \end{cases} \]

\[ \text{w}, \text{m} : [\text{Rep} \land \text{Pre}, \text{Post} \land m_0 = m] \sqsubseteq \text{impl} \]

\[ = \begin{cases} \text{w and m partition } t \end{cases} \]

\[ t : [\text{Rep} \land \text{Pre}, \text{Post} \land m_0 = m] \sqsubseteq \text{impl} \]

\[ = \begin{cases} t \text{ is the set of targets of } \text{impl}, \text{i.e., } t \text{ is a conservative} \\
\text{estimate of the variables that are possibly modified by } \text{impl} \end{cases} \]

\[ z : [\text{Rep} \land \text{Pre}, \text{Post} \land m_0 = m] \sqsubseteq \text{impl} \]

Let \( v_0 \) denote a list of initial-value variables, such that \( v_0 \) is a superset of the initial-value variables in \( \text{Post} \land m_0 = m \). Then, from the above and (11.10), we have

\[ \text{Refine(impl id is gc, E)} = [\text{Rep} \land \text{Pre} \land v_0 = v \Rightarrow \text{wp}(\text{Post} \land m_0 = m)] (11.8) \]

In the above calculation, I used the fact that, for any disjoint lists of variables \( w \) and \( m \), and predicates \( P \) and \( Q \),

\[ w : [P, Q] = w, m : [P, Q \land m_0 = m] \]

(11.9)

I end this section by discharging that proof obligation.

Recall that the weakest (liberal) precondition of a specification \( w : [P, Q] \) with respect to a postcondition \( R \) is

\[ P \land \langle \forall w \mid Q \triangleright R \rangle[v_0 := v] \]
where $v_0$ is the list of initial-value variables that occur in $Q$. Since neither $w$ nor $R$ contain initial-value variables—initial-value variables can only be written in the postcondition of specifications—, this expression equals

$$P \land (\forall w \mid Q \triangleright R)[z_0 := z],$$

where $z$ is any superset of $v$ (cf. Section 1.7).

Let $z$ denote any superset of $v \cup m$ (which, of course, is also a superset of $v$). Then, we calculate,

$$wlp.(w, m : [P, Q \land m_0 = m]).R$$

$$= \{ \text{(1.16): $wlp$ of specification, using $z$ instead of $v$ } \}$$

$$P \land (\forall w, m \mid Q \land m_0 = m \triangleright R)[z_0 := z]$$

$$= \{ \text{one-point rule, since $m$ and $m_0$ are disjoint } \}$$

$$P \land (\forall w \mid Q[m := m_0] \triangleright R[m := m_0])[z_0 := z]$$

$$= \{ \text{substitution distributes over $\forall$, since $m$ and $w$ are disjoint } \}$$

$$P \land (\forall w \mid Q \triangleright R)[m := m_0][z_0 := z]$$

$$= \{ \text{substitution, since $m \subseteq z$ and thus $m_0 \subseteq z_0$, and since $m$ and $z_0$ are disjoint } \}$$

$$P \land (\forall w \mid Q \triangleright R)[m := m][z_0 := z]$$

$$= \{ \text{[(1.16): $wlp$ of specification, using $z$ instead of $v$ } \} \}$$

$$wlp.(w : [P, Q]).R \ .$$

### 11.2 Proving refinements

In this section, I give a proof of

$$(z : [\text{Pre, Post}] \subseteq S) = [\text{Pre} \land v_0 = v \Rightarrow wlp.S.Post] \ ,$$

(11.10)

where $z$ denotes the list of all variables and $v_0$ is a list of initial-value variables such that $v_0$ is a superset of the set of initial-value variables in $\text{Post}$. This property allows a refinement of this form to be proven using a simple implication.

I define a $z$-predicate to be a predicate over a state space whose variables are $z$. For the $S$ in (11.10), $wlp.S$ is thus a $z$-predicate transformer, i.e., a function from one $z$-predicate to another. Hence, for any $z$-predicate $R$, $wlp.S.R$ is a $z$-predicate.

We calculate,

$$z : [\text{Pre, Post}] \subseteq S$$

$$= \{ \text{(1.18): def. of $\subseteq$, and (1.16): $wlp$ of specification statement } \}$$

$$\langle \forall R \triangleright [\text{Pre} \land (\forall z \mid \text{Post} \triangleright R)[v_0 := v] \Rightarrow wlp.S.R] \rangle$$

$$= \{ \text{substitution, since $v_0$ does not occur free in $\text{Pre}$ or $wlp.S.R$ } \}$$

$$\langle \forall R \triangleright [(\text{Pre} \land (\forall z \mid \text{Post} \triangleright R) \Rightarrow wlp.S.R)[v_0 := v]] \rangle$$

$$= \{ \text{below } \}$$

122
\[(\text{Pre} \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]\]
\[= \{ \text{substitution, since } v_0 \text{ does not occur free in } \text{Pre} \} \]
\[\{ \text{one-point rule, since } v_0 \text{ does not appear in} \]
\[\text{Pre} \Rightarrow \text{wp}.\text{S.Post}[v_0 := v]\} \]
\[\{ \text{equality and substitution} \} \]
\[\{ \text{Pre} \land v_0 = v \Rightarrow \text{wp}.\text{S.Post}[v_0 := v] \} \]

The deferred step is proved by mutual implication. For one direction,

\[\forall R \triangleright [(\text{Pre} \land (\forall z \triangleright \text{Post} \triangleright R) \Rightarrow \text{wp}.\text{S.R})[v_0 := v]]\]
\[\Leftrightarrow \{ \text{monotonicity, since } \forall z \text{ quantifies over all free variables of } R, \]
\[\text{and } \text{wp}.\text{S is a } z\text{-predicate transformer} \} \]
\[\forall R \triangleright [(\text{Pre} \land (\forall z \triangleright \text{Post} \triangleright R) \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]]\]
\[\Leftrightarrow \{ \text{monotonicity} \} \]
\[\forall R \triangleright [(\text{Pre} \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]]\]
\[\{ \text{range in nonempty} \} \]
\[\text{[(Pre} \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]]\]

and the other,

\[\forall R \triangleright [(\text{Pre} \land (\forall z \triangleright \text{Post} \triangleright R) \Rightarrow \text{wp}.\text{S.R})[v_0 := v]]\]
\[= \{ [ ] \text{ over } \forall \} \]
\[\forall R \triangleright (\text{Pre} \land (\forall z \triangleright \text{Post} \triangleright R) \Rightarrow \text{wp}.\text{S.R})[v_0 := v] \]
\[\Rightarrow \{ \text{ instantiate with } R := \text{Post} \} \]
\[\text{[(Pre} \land (\forall z \triangleright \text{Post} \triangleright \text{Post}) \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]]\]
\[\Rightarrow \{ \text{ pred. calc.} \} \]
\[\text{[(Pre} \Rightarrow \text{wp}.\text{S.Post})[v_0 := v]]\]

**Remark 11.5.** A proof of (11.10) appears in [67]. However, that proof contains a step whose hint is rather vague. In my notation, using symbols similar to those above, and with \(\text{wp}\) instead of \(\text{wp}\) (but remember that \(\text{wp}\) and \(\text{wp}\) coincide for the specification statement—see Sections 1.7 and 6.2), that step is

\[\text{Pre} \Rightarrow \text{wp}.\text{S.Post} \]
\[\Rightarrow \{ \text{by distributivity of } \Rightarrow \text{ over weakest preconditions} \} \]
\[\text{Pre} \land (\forall z \triangleright \text{Post} \Rightarrow R) \Rightarrow \text{wp}.\text{S.R} \]

where \(R\) is an arbitrary predicate. The step is correct, but the hint seems to suggest that a step like

\[\text{Pre} \Rightarrow \text{wp}.\text{S.Post} \]
\[\Rightarrow \{ \text{by distributivity of } \Rightarrow \text{ over weakest preconditions} \} \]
\[\text{Pre} \land (\text{Post} \Rightarrow R) \Rightarrow \text{wp}.\text{S.R} \]
would also be correct, which it is not! \( wp.S \) is a monotonic function from \( z \)-predicates to \( z \)-predicates. This monotonicity is used in the proof step. However, in order to apply monotonicity, the two predicates involved (\( Post \) and \( R \) above) must be ordered by implication at every value of \( z \). This subtle point is not visible from the hint in \[67\].

In my first proof of the above theorem, which I constructed with Paul Sivilotti not knowing of the proof in \[67\], we fell into the trap of writing a vague hint like the one in \[67\]. Only later did I realize the subtlety described here.

As another remark on this step, \( (\forall z \mid Post \Rightarrow R) \) can be written with everywhere brackets as \( [Post \Rightarrow R] \), where the everywhere brackets quantify over all values of \( z \). Then the hint in my proof can be stated as

\[
[Post \Rightarrow R] \Rightarrow [wp.S.Post \Rightarrow wp.S.R],
\]

which clearly expresses the property being used. However, there is another set of everywhere brackets in the proof, and those everywhere brackets quantify over all values of \( z \) and \( z_0 \). To avoid confusion between the two different pairs of everywhere brackets, I kept the explicit quantification over \( z \).

### 11.3 The importance of residues

I conclude this chapter by giving an example that shows the importance of the presence of residues. Consider the following unit.

```plaintext
unit A is
  spec var a ; var c ; depends a on c ;
  spec pc is modifies c requires true ensures true ;
  spec pa is modifies a requires true ensures c_0 = c ;
end
```

Procedure \( pc \) is specified to play havoc with \( c \), but with no other variable and, in particular, without affecting the value of \( a \). Procedure \( pa \) plays havoc with \( a \) (i.e., it plays havoc with the resolve set of \( a \)), but ensures that the value of \( c \) is unchanged.

Contemplate the following (incorrect) argument.

\( pc \) is specified to only modify \( c \), whereas \( pa \) can modify \( a \). In the scope of \( A \), the downward closure of \( a \) contains only one concrete variable, viz., \( c \). Thus, both \( pc \) and \( pa \) are constrained to modifying only \( c \). Furthermore, \( pc \) can change \( c \) to any value, whereas \( pa \) must leave its value unchanged. Hence, \( pa \) refines \( pc \).
As I show in this section, \( pa \) does not refine \( pc \), so the above argument is incorrect. It is incorrect because the downward closure of \( a \) may contain concrete variables other than \( c \)—concrete variables that are not visible in \( A \). Residues are what catch the erroneous reasoning demonstrated in the above argument. Let’s take a look at how this works.

Consider another unit.

\[
\begin{align*}
\text{unit } B \text{ import } A \text{ is} \\
\quad \text{var } d ; & \quad \text{depends } a \text{ on } d ; \\
\quad \text{impl } pa \text{ is } d := d + 1 ; \\
\quad \text{end}
\end{align*}
\]

\( \cdots \ d := 0 ; \ call \ pc ; \ assert \ d = 0 \ \cdots \quad (11.11) \)

This unit introduces another of \( a \)'s dependencies, \( d \). The unit defines the implementation of \( pa \), which I will show to be a correct refinement of its specification. In the unit, I show another little code segment, (11.11) I will demonstrate that this code segment does not go wrong.

First, the refinement of \( pa \).

\[
\begin{align*}
\text{Refine (impl } pa \text{ is } d := d + 1 , B) \\
= \\
\quad \{ \text{(11.8): Refine } \} \\
\quad [ c_0 = c \Rightarrow \ wlp.(d := d + 1).(c_0 = c)] \\
= \\
\quad \{ \text{(1.0): wlp of } := \} \\
\quad [ c_0 = c \Rightarrow c_0 = c] \\
= \\
\quad \{ \text{pred. calc. } \} \\
\quad \text{true}
\end{align*}
\]

This shows that \( pa \) is indeed implemented correctly. Note that, as suggested immediately following the incorrect argument above, this shows that \( pa \) really does have an effect on a concrete variable other than \( c \). (Thus far, we have not used the presence of residues.)

Let’s prove that the code segment (11.11) in unit \( B \) does not go wrong.

\[
\begin{align*}
\text{wlp}(d := 0 ; \ call \ pc ; \ assert \ d = 0 , \ true , B) \\
= \\
\quad \{ \text{(1.3,6.10): wlp of ; and assert } \} \\
\quad \text{wlp.(d := 0).(wlp.Env(call pc,B).(d = 0))} \\
= \\
\quad \{ \text{Env , using Remark 11.4 } \} \\
\quad \text{wlp.(d := 0).(wlp.(c : [true, true]).(d = 0))} \\
= \\
\quad \{ \text{(1.14): wlp of specification statement } \} \\
\quad \text{wlp.(d := 0).}(\forall c \triangleright d = 0) \\
= \\
\quad \{ \text{pred. calc. } \} \\
\quad \text{wlp.(d := 0).(d = 0)} \\
= \\
\quad \{ \text{(1.0): wlp of := , and pred. calc. } \} \\
\quad \text{true}
\end{align*}
\]

125
(This calculation, too, is independent of the presence of residues.) We conclude that \( p_a \) is correctly implemented and that (11.11) does not go wrong regardless of whether or not residues are included.

Now, let’s look at one more unit. This unit provides the implementation of \( p_c \).

\[
\text{unit } C \text{ import } A \text{ is } \\
\quad \text{impl } p_c \text{ is call } p_a \\
\text{end}
\]

Does this implementation of \( p_c \) meet \( p_c \)'s specification? It would if \( p_a \) were a refinement of \( p_c \). Let’s find out.

\[
\text{Refine(impl } p_c \text{ is call } p_a, C) \quad (11.12)
= \quad \{ \text{(11.8): Refine } \}
\quad [a_0 = a \land c_0 = c \implies \text{wlp.}(\text{call } p_a, C). (f \ast a(a_0, c_0) = f \ast a(a, c))]
= \quad \{ \text{Env } \}
\quad [a_0 = a \land c_0 = c \implies \text{wlp.}(a, c : [\text{true}, c_0 = c]). (f \ast a(a_0, c_0) = f \ast a(a, c))]
= \quad \{ \text{(11.9), with } w, m := a, c \}
\quad [a_0 = a \land c_0 = c \implies \text{wlp.}(a : [\text{true}, \text{true}]). (f \ast a(a_0, c_0) = f \ast a(a, c))]
= \quad \{ \text{(1.14): wlp of specification statement } \}
\quad [a_0 = a \land c_0 = c \implies (\forall a \triangleright f \ast a(a_0, c_0) = f \ast a(a, c))]
\]

Since nothing is known about function \( f \ast a \), we say it is \textit{uninterpreted}; thus, we understand \( f \ast a \) as being universally quantified over all functions. By instantiating \( f \ast a \) with \(+\), i.e., \( f \ast a(x, y) = x + y \) for all \( x \) and \( y \), the right-hand side of \( \implies \) reads

\[
(\forall a \triangleright a_0 + c_0 = a + c)
\]

which simplifies to \textit{false}. We conclude that, with residues, we are not able to establish that unit \( C \) provides a correct implementation of \( p_c \). This is good, because by tracing the procedure call in that implementation, we find that \( p_c \) has the effect of incrementing \( a \), even though its specification constrains it to modify only \( c \). Hence, the code segment (11.11) \textit{will} go wrong, and we are pleased that we were not able to conclude otherwise.

Let’s then compare what happens if residues are not used. Not using residues means that the downward closure of \( a \) does not contain the residue \( a \). The calculation embarked on at (11.12) then arrives at the expression

\[
[c_0 = c \implies f \ast a(c_0) = f \ast a(c)]
\]

which simplifies to \textit{true} on account of Leibniz. So, without residues, a verification process would incorrectly verify the implementation of \( p_c \) to be correct.

In modular verification, one unit can never be certain that it has information about all dependencies. Therefore, the downward closure of a variable includes a residue that represents all dependencies that are not in scope.

126
Only if the representation of an abstract variable is in scope is there a way to eliminate residues from expressions. For example, the representation

\texttt{rep a is a = c + d}

is functionalized (see Section 11.1.3) into the axiom

\( \langle \forall a, c, d \triangleright f\ast a(a, c, d) = c + d \rangle \).

Note that, despite the presence of \( a \)'s \texttt{rep} clause, \( a \) may have dependencies that are not in scope—nothing prevents another unit from declaring other (unused) dependencies. Since such other dependencies do not affect the value of abstract variable \( a \), it is sound to use the above axiom.

In conclusion, I showed in this section that without residues, modular verification would not be sound. In the next chapter, I show that with residues, and with the visibility and authenticity requirements, modular verification is sound.
Chapter 12

Soundness of modular verification

In this chapter, I give a formal proof of the soundness of modular verification with respect to my `depends` solution. Using the notation and definitions from the previous chapter, I define the visibility and authenticity requirements formally. Then, I state and prove the soundness theorem. I close the chapter with some remarks on the proof.

12.0 Requirements

An environment `Prog` satisfies the *visibility requirement* exactly when for every environment `E` that is a subset of `Prog`, and for every two variables `a` and `c`,

\[
\langle \exists E' \mid E' \subseteq \text{Prog} \triangleright \text{Depends}(a, c, E') \rangle \land a \in E \land c \in E \Rightarrow \text{Depends}(a, c, E)
\]

In words, the visibility requirement states that if `a` depends on `c`, and both `a` and `c` are visible in `E`, then the dependency of `a` on `c` is visible in `E`.

An environment `Prog` satisfies the *authenticity requirement* exactly when for every environment `E` that is a subset of `Prog`, and for every two variables `a` and `c`,

\[
\langle \exists E' \mid E' \subseteq \text{Prog} \triangleright \text{Depends}(a, c, E') \rangle \land c \in E \Rightarrow a \in E
\]

That is, a variable `a` that depends on a variable `c` is visible anywhere `c` is.

12.1 Soundness

I now state the soundness theorem.

\[
\text{Refine}(:, \text{impl id is gc}, E) \Rightarrow \text{Refine}(:, \text{impl id is gc}, \text{Prog})
\]

\[
\iff
\]

*impl id is gc* \in E \land E \subseteq \text{Prog} \land

`E` is an environment \land

`Prog` is an environment that satisfies the

visibility and authenticity requirements
The theorem states that $\text{impl } id \text{ is gc}$ is deemed a correct implementation (of procedure $id$ with respect to its specification) in the environment $E$ only if it would be deemed a correct implementation in the entire program $Prog$. That is, the implementation of $id$ is verified to be correct in light of the information contained in $E$, a subset of $Prog$, only if it would be verified to be correct in light of all information in $Prog$.

This theorem is what allows a verification process to make use of only the information in the scope of a unit $U$, i.e., $E := \text{ImportClosure}(U)$, when verifying that an implementation in $U$ refines its specification.

Stated differently, to show that a procedure implementation meets its specification, one needs to establish

$$\text{Refine}(\text{impl } id \text{ is gc}, \ Prog)$$  \hspace{1cm} (12.0)

Since this involves the entire program $Prog$, and the entire program may not be available at the time the procedure is to be verified, this proof obligation seems difficult to meet. The soundness theorem states that $\text{Refine}$ is monotonic with respect to its second argument, the environment. Hence, to show \((12.0)\), it suffices to show

$$\text{Refine}(\text{impl } id \text{ is gc}, E)$$  \hspace{1cm} (12.1)

for any subset $E$ of $Prog$. $E$ can then be picked as the import closure of the unit that contains the procedure implementation, and the proof of \((12.1)\) can be carried out as a modular verification.

In the next few sections, I give the proof of this theorem.

### 12.2 Proof outline

Let $id, gc, E, Prog$ satisfy the antecedent of the theorem.

$$(\text{impl } id \text{ is gc}) \in E \land E \subseteq Prog \land$$

$E$ is an environment $\land$

$Prog$ is an environment that satisfies the

visibility and authenticity requirements

We need to prove

$$\text{Refine}(\text{impl } id \text{ is gc}, E) \Rightarrow \text{Refine}(\text{impl } id \text{ is gc}, Prog)$$

I introduce symbols $\text{frame, pre, post, b, b', m, m', rep, reprep}$, satisfying (cf. Section 11.1)

$$(\text{modifies } \text{frame requires pre ensures post}) = \text{Lookup}(id, E)$$

$$b = \text{SpecVarProjection}(\text{BenSideSet}(\text{Resolve}(\text{frame, E}), E))$$

$$b' = \text{SpecVarProjection}(\text{BenSideSet}(\text{Resolve}(\text{frame, Prog}), Prog))$$

$$m = \text{Target}(gc, E) \setminus \text{Resolve}(\text{frame, E})$$

$$m' = \text{Target}(gc, Prog) \setminus \text{Resolve}(\text{frame, Prog})$$

129
rep = RepAxioms(E)
reprep = RepAxioms(Prog \ E)

We thus have

\[ \text{RepPreds}(E) = \text{Functionize}(\text{rep}, E) \]
\[ \text{RepPreds}(\text{Prog}) = \text{Functionize}(\text{rep}, \text{Prog}) \land \text{Functionize}(\text{reprep}, \text{Prog}) \]

Let \( r^p \) be a shorthand for \( \text{rep} \land \text{pre} \).

For brevity, I will use \( R \) synonymously with \( \text{Resolve} \) and define \( F \) and \( F' \) as follows.

\[ F(Q) = \text{Functionize}(Q, E) \]
\[ F'(Q) = \text{Functionize}(Q, \text{Prog}) \]

In the sequel, I will implicitly use the fact that \( F \) and \( F' \) distribute over logical connectives.

Let \( z \) denote the list of all variables in \( \text{Prog} \). In terms of the above shorthands, we now have (cf. 11.8),

\[ \text{Refine(impl id is gc, E)} = \]
\[ \quad \quad [F(\text{rep}) \land F(\text{pre}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad wlp(gc, F(\text{post} \land b_0 = b) \land m_0 = m, E)] \]
\[ \text{Refine(impl id is gc, Prog)} = \]
\[ \quad \quad [F'(\text{rep}) \land F'(\text{reprep}) \land F'(\text{pre}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad wlp(gc, F'(\text{post} \land b'_0 = b') \land m'_0 = m', \text{Prog})] \]

I will make use of the property

\[ b \subseteq b' \]

which I will prove later. In light of this, I let \( bb \) denote \( b' \setminus b \), and partition \( bb \) into \( bx \) and \( by \). The exact nature of this partition won’t be revealed until Section 12.3.1.

We calculate,

\[ \text{Refine(impl id is gc, Prog)} \]
\[ = \]
\[ \quad \quad \quad \quad \{ \text{(12.3)} \} \]
\[ \quad \quad \quad \quad [F'(\text{rep}) \land F'(\text{reprep}) \land F'(\text{pre}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad \quad \quad \quad wlp(gc, F'(\text{post} \land b'_0 = b') \land m'_0 = m', \text{Prog})] \]
\[ \Leftarrow \]
\[ \quad \quad \quad \quad \{ \text{strengthen, } r^p, \text{ and distribute } F' \} \]
\[ \quad \quad \quad \quad [F'(\text{rp}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad \quad \quad \quad wlp(gc, F'(\text{post} \land b_0 = b) \land m_0 = m', \text{Prog})] \]
\[ \quad \quad \quad \quad \{ \text{partition } b' \text{ into } b, bx, by \text{ (see (12.4) above)} \} \]
\[ \quad \quad \quad \quad [F'(\text{rp}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad \quad \quad \quad wlp(gc, F'(\text{post} \land b_0 = b) \land F'(bx_0 = bx) \land F'(by_0 = by) \land m'_0 = m', \text{Prog})] \]
\[ \quad \quad \quad \quad \{ \text{(11.6): } wlp \text{ is conjunctive, and pred. calc.} \} \]
\[ \quad \quad \quad \quad [F'(\text{rp}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad \quad \quad \quad \quad wlp(gc, F'(\text{post} \land b_0 = b) \land F'(bx_0 = bx) \land m'_0 = m', \text{Prog})] \]
\[ \quad \quad \quad \quad \quad \land \]
\[ \quad \quad \quad \quad \quad \land [F'(\text{rp}) \land z_0 = z \Rightarrow \]
\[ \quad \quad \quad \quad \quad \quad \quad \quad wlp(gc, F'(by_0 = by), \text{Prog})] \].
The latter of these conjuncts follows from

\[ z_0 = z \Rightarrow wlp(gc, \ F'(by_0 = by), \ Prog) \]

which I will prove later.

For the former, I will use a function \( X \). The idea is that \( X \) syntactically translates from \( E \)-expressions to \( \text{Prog} \)-expressions. Applied to a predicate, \( X \) provides a way to re-functionalize the predicate with respect to a larger environment (see (12.6)); applied to a list of variables (cf. (12.11)), \( X \) provides a way to re-resolve that list. For any predicate \( Q \) over the variables in \( E \), and any guarded command \( S \) in \( E \), \( X \) enjoys the following properties.

\[
F'(Q) = X(F(Q))
\]

\[
z = X(z)
\]

\( X \) distributes over logical connectives

\[
[X(Q)] \leftarrow [Q]
\]

\[
\text{Resolve}(bx, \ Prog) \subseteq X(m)
\]

\[
m' \subseteq X(m)
\]

\[
wlp(S, X(Q), \ Prog) \leftarrow X(wlp(S, Q, E))
\]

I postpone the precise definition of \( X \) and the proofs that show that \( X \) does have these properties.

Note that \( X \) is monotonic: For any predicates \( Q_0 \) and \( Q_1 \) over the variables of \( E \),

\[
[X(Q_0) \Rightarrow X(Q_1)]
\]

\[
\Rightarrow \quad \{ \ (12.8) \ \}
\]

\[
[X(Q_0 \Rightarrow Q_1)]
\]

\[
\Rightarrow \quad \{ \ (12.9) \ \}
\]

\[
[Q_0 \Rightarrow Q_1]
\]

We calculate,

\[
F'(post \land b_0 = b) \land F'(bx_0 = bx) \land m_0 = m'
\]

\[
\Rightarrow \quad \{ \ (12.11) \ \}
\]

\[
F'(post \land b_0 = b) \land F'(bx_0 = bx) \land X(m_0 = m)
\]

\[
\Rightarrow \quad \{ \ \text{Leibniz (11.5): } F'(bx_0 = bx) \Leftarrow R(bx_0 = bx, \ Prog) \ \}
\]

\[
F'(post \land b_0 = b) \land R(bx_0 = bx, \ Prog) \land X(m_0 = m)
\]

\[
\Rightarrow \quad \{ \ (12.10) \ \}
\]

\[
F'(post \land b_0 = b) \land X(m_0 = m)
\]

Now, we have,

\[
[F'(rp) \land z_0 = z \Rightarrow wlp(gc, \ F'(post \land b_0 = b) \land F'(bx_0 = bx) \land m_0 = m', \ Prog)]
\]

\[
\Rightarrow \quad \{ \ \text{above calculation, and } wlp \text{ is monotonic} \ \}
\]

\[
[F'(rp) \land z_0 = z \Rightarrow wlp(gc, \ F'(post \land b_0 = b) \land X(m_0 = m), \ Prog)]
\]

\[
\Rightarrow \quad \{ \ (12.6), \text{ and } (12.7), \text{ and } (12.8) \ \}
\]
\[ [X(F(\rho) \land z_0 = z) \Rightarrow wlp(gc, X(F(\text{post} \land b_0 = b) \land m_0 = m), \text{Prog})] \]
\[ \Leftarrow \{ \begin{array}{l}
\{ \text{(12.12) } \}
[ X(F(\rho) \land z_0 = z) \Rightarrow X(wlp(gc, F(\text{post} \land b_0 = b) \land m_0 = m, E))] \end{array} \}
\]
\[ \Leftarrow \{ \begin{array}{l}
\{ \text{ } X \text{ is monotonic } \}
F(\rho) \land z_0 = z \Rightarrow wlp(gc, F(\text{post} \land b_0 = b) \land m_0 = m, E) \end{array} \}
\]
\[ = \{ \begin{array}{l}
\{ \text{(12.2) } \}
\text{Refine}(\text{impl id is gc}, E) \end{array} \} . \]

That was the soundness proof! Well, only an outline thereof—I still need to produce the proofs of the properties (12.4)–(12.12) that were used in this proof outline. I will do these in order.

12.3 All side effects are benevolent

In this section, I discharge proof obligations (12.4) and (12.5).

12.3.0 \( b \subseteq b' \)

I now prove (12.4), that is, \( b \subseteq b' \).

\[
\begin{align*}
\{ b \subseteq b' \} \\
&= \{ \text{ } b \text{ and } b' \} \\
&\quad \text{SpecVarProjection}(\text{BenSideSet}(R(\text{frame}, E), E)) \\
&\subseteq \{ \text{ SpecVarProjection } \text{ is monotonic } \} \\
&\quad \text{BenSideSet}(R(\text{frame}, E), E) \subseteq \text{BenSideSet}(R(\text{frame}, \text{Prog}), \text{Prog}) \\
&\Leftarrow \{ \begin{array}{l}
\text{SpecVarProjection} \text{ is monotonic } \\
\text{BenSideSet}(R(\text{frame}, E), E) \text{ is monotonic is } E 
\end{array} \} \quad (12.13)
\end{align*}
\]

I prove this monotonicity. Let \( E \) and \( E' \) be two environments such that \( E \subseteq E' \), and let \( w \) be any list of variables visible in \( E \). Letting \( v \) range over variables, we calculate,

\[
\begin{align*}
\text{BenSideSet}(R(w, E), E) &\subseteq \text{BenSideSet}(R(w, E'), E') \\
&= \{ \text{ } \text{BenSideSet} \} \\
&\quad \{ v \in E \setminus R(w, E) \Rightarrow v \} \subseteq \{ v \in E' \setminus R(w, E') \Rightarrow v \} \\
&= \{ \begin{array}{l}
\text{ sets } \\
\forall v \in E \land v \notin R(w, E) \Rightarrow v \in E' \land v \notin R(w, E') 
\end{array} \} \\
&= \{ \begin{array}{l}
\text{ } E \subseteq E' \\
\forall v \in E \land v \notin R(w, E) \Rightarrow v \notin R(w, E') 
\end{array} \} \\
&= \{ \text{ } \text{trading} \} \\
&\quad \{ \forall v \in E \land v \in R(w, E') \Rightarrow v \in R(w, E) \} .
\end{align*}
\]

For any variable \( v \), we calculate,
\[ v \in E \land v \in \text{Resolve}(w, E') \]
\[ \Rightarrow \{ \text{(11.2): property of Resolve} \} \]
\[ v \in E \land (\exists a \mid a \in w \Rightarrow \text{Depends}(a, v, E')) \]
\[ \Rightarrow \{ \text{visibility requirement,} \]
\[ \text{since } a \text{ is visible in } E \ (a \in w \subseteq E), \]
\[ \text{and } v \text{ is visible in } E \ (v \in E) \} \]
\[ v \in E \land (\exists a \mid a \in w \Rightarrow \text{Depends}(a, v, E)) \]
\[ \Rightarrow \{ \text{(11.2): property of Resolve} \} \]
\[ v \in \text{Resolve}(w, E) \]
which concludes the proof.

12.3.1 \ bx AND by

In this section, I define \ bx and \ by. As advertised before, \ bx and \ by partition \ bb. From the definition of \ bx, I calculate the necessary ingredient in the upcoming proof of (12.10). From the definition of \ by, I prove (12.5).

For any specification variable \( d \), we calculate,

\[ d \in bb \]
\[ = \{ \ b b , \ \ \ \land \ \} \]
\[ d \in b \land b \notin b' \]
\[ = \{ \ b , \ b' \} \]
\[ d \in \text{Prog} \land R(\text{frame}, \text{Prog}) \land d \notin E \land R(\text{frame}, E) \]
\[ \Rightarrow \{ \ \land, \text{twice} \} \]
\[ d \notin R(\text{frame}, \text{Prog}) \land (d \notin E \lor d \in R(\text{frame}, E)) \]
\[ = \{ \text{(11.1): Resolve is monotonic, since } E \subseteq \text{Prog} \} \]
\[ d \notin R(\text{frame}, \text{Prog}) \land d \notin E \] \hspace{1cm} (12.14)

Treatment of \ bx

I now define \ bx, a subset of \ bb: For every \( d \) in \ bb, \( d \in bx \) exactly when there is a procedure call \ call p \ in \ gc \ and a \( g \) in the frame of the specification of \ p \ such that

\[ \text{Depends}(g, d, \text{Prog}) \]

Thus,

\[ \text{true} \]
\[ = \{ \ g \text{ is in frame of specification of } p \ \} \]
\[ g \in \text{RawTarget}(\text{call } p, E) \]
\[ \Rightarrow \{ \ \text{call } p \text{ is a substatement of } gc \ \} \]
\[ g \in \text{RawTarget}(gc, E) \]
\[ \Rightarrow \{ \text{(11.7): RawTarget}(gc, E) \subseteq \text{Target}(gc, E) \} \]
\[ g \in \text{Target}(gc, E) \]
We calculate,

\[
\begin{align*}
\text{Depends}(g, d, \text{Prog}) \\
= \\
\{ (12.14) \} \\
\text{Depends}(g, d, \text{Prog}) \land d \notin R(\text{frame}, \text{Prog}) \Rightarrow \\
\{ (11.4): \text{Depends}(g, d, \text{Prog}) \land g \in R(\text{frame}, \text{Prog}) \Rightarrow d \in R(\text{frame}, \text{Prog}) \} \\
\therefore g \notin R(\text{frame}, \text{Prog}) \\
\{ (11.1): \text{Resolve is monotonic, since } E \subseteq \text{Prog} \} \\
\therefore g \notin R(\text{frame}, E)
\end{align*}
\]

From the last two calculations, we conclude,

\[
g \in \text{Target}(g, E) \setminus R(\text{frame}, E)
\]

i.e., \(g \in m\). From this and from (12.14), every element \(d\) of \(bx\) enjoys the property

\[
d \notin E \land (\exists g \mid g \in m \Rightarrow \text{Depends}(g, d, \text{Prog}))
\]

(12.15)

Treatment of \(by\)

I define \(by\) as \(bb \setminus bx\). Hence, \(bx\) and \(by\) do indeed partition \(bb\). For every \(d\) in \(by\), no procedure call in \(gc\) has a frame that mentions a \(g\) such that \(\text{Depends}(g, d, \text{Prog})\). Thus, no procedure call in \(gc\) modifies any variable which depends on \(d\); stated differently, \(d\) does not appear in the downward closure of the frame of the specification of any procedure call in \(gc\). So, no procedure call in \(gc\) has any effect on the value of \(d\).

Let \(y\) be a variable updated by an assignment statement in \(gc\). \(y\) is visible in \(E\), because \(gc\) is contained in \(E\). We calculate,

\[
y \in E \\
= \\
\{ (12.14) \}
\]

\[
y \in E \land d \notin E \\
\Rightarrow \\
\neg \text{Depends}(d, y, \text{Prog})
\]

Since the only statements that may have an effect on the state are assignment statements and procedure calls, we conclude that no statement in \(gc\) has an effect on the value of \(d\). This is so for every \(d\) in \(by\), and thus,

\[
\text{wlp}(gc, F'(by_0 = by), \text{Prog}) = F'(by_0 = by)
\]

the right-hand side of which, by Leibniz, follows from

\[
R(by_0 = by, \text{Prog})
\]

Since \(R(by, \text{Prog})\) is a subset of \(z\), the set of all variables, (12.5) follows.
12.4 \( X \)

I define function \( X \), which implicitly is a function of \( E \) and \( \text{Prog} \).

\[
X(Q) =
\]
\[
Q \text{ in which every occurrence (not just free occurrences) of } x \text{ is replaced by } y
\]
where
\[
x \text{ is the list of specification variables in } E \text{ and their initial-value forms, and}
\]
for each \( a \) in \( x \), the corresponding term in \( y \) is:
\[
a \text{ and the symbols in } \text{Resolve}(a, \text{Prog}) \setminus \text{Resolve}(a, E)
\]

This substitution is a rather curious one in that one identifier may be replaced by a list of identifiers. This causes no problem textually, but one may question the meaning of the resulting expression. If \( Q \) is a functionalized expression, then all occurrences in \( Q \) of the identifiers in list \( x \) appear as residues (I assume that identifiers are not renamed from the way they were produced as per Chapter 11). This means that they appear as parameters to \( f \ast \) functions or as dummies of quantifications. Replacing an identifier of the first kind with a list of identifiers thus alters the arity of the function. It is important that \textit{every} occurrence of \( x \), not just free occurrences, be replaced, so that the arity of each \( f \ast \) function is changed consistently; also replacing the identifiers of the second kind takes care of making \( y \) bound whenever \( x \) was.

Let me illustrate with an example. Let \( E \) be the environment

\[
\text{spec var } a \ ; \ \text{var } c \ ; \ \text{depends } a \text{ on } c ,
\]
and \( \text{Prog} \) the environment \( E \) extended with

\[
\text{var } d \ ; \ \text{depends } a \text{ on } d .
\]

Then,

\[
X(\text{Functionalize}(\forall a, c \mid f \ast a(a, c) \leq 0 \triangleright f \ast a(a, c) = 0), E))
\]
\[
= \{ \text{Functionalize with respect to } E \}
\]
\[
X((\forall a, c \mid f \ast a(a, c) \leq 0 \triangleright f \ast a(a, c) = 0))
\]
\[
= \{ X \}
\]
\[
((\forall a, c \mid f \ast a(a, c) \leq 0 \triangleright f \ast a(a, c) = 0))[a := (a, d)]
\]
\[
= \{ \text{substitution} \}
\]
\[
(\forall a, c, d \mid f \ast a(a, c, d) \leq 0 \triangleright f \ast a(a, c, d) = 0)
\]
\[
(12.16)
\]

I assume the order of the symbols returned by \text{Resolve} ensures that \text{Resolve}(a, E) is a prefix of \text{Resolve}(a, \text{Prog}). I also assume operator \( \setminus \) to preserve ordering. If the insertion of new symbols is then always done appropriately at the end of lists (as the example shows), then, for any list of variables \( w \) visible in \( E \), we have

\[
\text{Resolve}(w, \text{Prog}) = X(\text{Resolve}(w, E))
\]
\[
(12.17)
\]

135
Consequently, we have, for any $Q$ built from symbols visible in $E$,

$$\text{Functionalize}(Q, \text{Prog}) = X(\text{Functionalize}(Q, E))$$

This shows (12.6). For example, note that (12.16) equals

$$\text{Functionalize}(\langle \forall a \mid a \leq 0 \Rightarrow a = 0 \rangle, \text{Prog})$$

As another consequence of (12.17), we have, for any variable $v$ and list of variables $w$,

$$\langle \exists a \mid a \in w \Rightarrow v \in R(a, \text{Prog}) \setminus R(a, E) \rangle \Rightarrow v \in X(w)$$

(12.18)

For any expression $e$ in which every variable is a program variable, $e = X(v)$. $X$ does not introduce identifiers not found in $\text{Prog}$; hence, for the set $z$ of all variables, $z = X(z)$, justifying (12.7). Note that $X$, like regular substitution, distributes over logical connectives, justifying (12.8). For quantifications, we have

$$X(\langle Q \mid R \Rightarrow T \rangle) = \langle Q \rangle X(w) \mid X(R) \Rightarrow X(T)$$

(12.19)

For the substitution function,

$$X(Q[w_0 := w]) = X(Q)[X(w_0) := X(w)]$$

(12.20)

provided $Q$ does not contain any occurrences of $X(w_0) \setminus w_0$.

### 12.4.0 Substitution Theorem

A well-known rule in the predicate calculus, referred to as instantiation, is

$$[T] \Rightarrow [T[s := t]]$$

where $s$ denotes a list of program variables and $t$ denotes a list of expressions. The substitution function $[s := t]$ only replaces free occurrences of $s$. If we instead let it replace bound occurrences, too, we will still find the implication shown above.

So what about the substitution function $X$? We could apply the above observations to $X$, if it weren’t for the fact that $X$ may replace one symbol with a list of other symbols. Hence, the following argument.

Focusing back on (12.9) and the definition of $X$, note that no symbol in $y \setminus x$ appears in $Q$. Moreover, since $x$ is a list of residues, $Q$ contains no information about the values of the identifiers in $x$; the identifiers in $x$ occur only as parameters to $f*$ functions or as bound variables in $Q$. Thus, the value of $[Q]$ does not depend on the exact values of $x$; the variables in $x$ can be thought of as just place holders. In the same way, $y$ will just be place holders in $[X(Q)]$, place holders appearing in the same places in $[X(Q)]$ that $x$ does in $[Q]$.

It thus follows that $[X(Q)]$ is true if $[Q]$ is, proving (12.9).
12.4.1 \( X \) AND \( m \)

In this section, I prove (12.10) and (12.11).

**Treatment of** \( \text{Resolve}(bx, \text{Prog}) \subseteq X(m) \)

I prove (12.10) from (12.15). We calculate, for any \( x \),

\[
\begin{align*}
  x & \in \text{Resolve}(bx, \text{Prog}) \\
  \Rightarrow & \quad \{ \text{(11.2): property of Resolve} \} \\
  & \quad \{ \text{(12.15): property of elements of } bx \} \\
  & \quad \{ \text{by authenticity requirement, } \text{Depends}(d, x, \text{Prog}) \land x \in E \Rightarrow d \in E \} \\
  & \quad \{ \text{(12.15): property of elements of } bx \} \\
  & \quad \{ \text{transitivity of Depends} \} \\
  & \quad \{ \text{(11.2): Depends}(g, d, \text{Prog}) \equiv x \in \text{Resolve}(g, \text{Prog}) \} \\
  & \quad \{ \text{Resolve}(g, E) \subseteq E \} \\
  & \quad \{ x \in X(m) \} \\
  x & \in X(m)
\end{align*}
\]

**Treatment of** \( m' \subseteq X(m) \)

I let \( rt \) stand for \( \text{RawTarget}(gc, E) \); then, \( rt \) also equals \( \text{RawTarget}(gc, \text{Prog}) \). Thus, \( m \) and \( m' \) can be written as

\[
\begin{align*}
  m & = R(rt, E) \setminus R(\text{frame}, E) \\
  m' & = R(rt, \text{Prog}) \setminus R(\text{frame}, \text{Prog})
\end{align*}
\]

Since we have \( m \in X(m) \), consider any \( v \in m' \setminus m \); I will show \( v \in X(m) \). We have

\[
\begin{align*}
  v & \in m' \\
  \Rightarrow & \quad \{ m' \text{ and } \setminus \} \\
  v & \notin R(\text{frame}, \text{Prog}) \\
  \Rightarrow & \quad \{ \text{(11.1): Resolve is monotonic, since } E \subseteq \text{Prog} \} \\
  v & \notin R(\text{frame}, E) \quad \text{(12.21)}
\end{align*}
\]

and

\[
\begin{align*}
  v & \notin m \\
  = & \quad \{ m \text{ and } \setminus \} \\
  v & \notin R(rt, E) \lor v \in R(\text{frame}, E) \\
  = & \quad \{ \text{(12.22)} \} \\
  v & \notin R(rt, E) \quad \text{(12.23)}
\end{align*}
\]
I complete the proof with the following calculation.

\[ v \in m' \]
\[ \Rightarrow \{ m' \text{ and } \\} \]
\[ v \in R(rt, Prog) \]
\[ = \{ (11.2): \text{property of } Resolve \} \]
\[ (\exists a : a \in rt \triangleright Depends(a, v, Prog)) \]
\[ = \{ (11.3): rt \subseteq R(rt, E) \} \]
\[ (\exists a : a \in rt \triangleright a \in R(rt, E) \land Depends(a, v, Prog)) \]
\[ = \{ (12.24) \text{ below: } Depends(a, v, Prog) \Rightarrow a \not\in R(frame, Prog) \} \]
\[ (\exists a : a \in rt \triangleright a \in R(rt, E) \land a \not\in R(frame, Prog) \land Depends(a, v, Prog)) \]
\[ \Rightarrow \{ (11.1): \text{Resolve is monotonic} \} \]
\[ (\exists a : a \in rt \triangleright a \in m \land Depends(a, v, Prog)) \]
\[ \Rightarrow \{ (12.25) \text{ below: } a \in rt \land Depends(a, v, Prog) \Rightarrow v \in R(a, Prog) \setminus R(a, E) \} \]
\[ (\exists a \triangleright a \in m \land v \in R(a, Prog) \setminus R(a, E)) \]
\[ \Rightarrow \{ (12.18): \text{property of } X \} \]
\[ v \in X(m) \]

The deferred proof obligations are shown by

\[ \text{true} \]
\[ = \{ (11.4): \text{closure property of } Resolve \} \]
\[ Depends(a, v, Prog) \land a \in R(frame, Prog) \Rightarrow v \in R(frame, Prog) \]
\[ = \{ \text{shunting} \} \]
\[ Depends(a, v, Prog) \Rightarrow a \not\in R(frame, Prog) \lor v \in R(frame, Prog) \]
\[ = \{ (12.21) \} \]
\[ Depends(a, v, Prog) \Rightarrow a \not\in R(frame, Prog) \] \hspace{1cm} (12.24)

and

\[ v \in R(a, Prog) \setminus R(a, E) \]
\[ = \{ \\} \]
\[ v \in R(a, Prog) \land v \not\in R(a, E) \]
\[ = \{ (11.2): v \in Resolve(a, Prog) \equiv Depends(a, v, Prog) \} \]
\[ Depends(a, v, Prog) \land v \not\in R(a, E) \]
\[ \Leftarrow \{ (11.1): \text{Resolve is monotonic} \} \]
\[ Depends(a, v, Prog) \land a \in rt \land v \not\in R(rt, E) \]
\[ = \{ (12.23) \} \]
\[ Depends(a, v, Prog) \land a \in rt \]

**Remark 12.0.** For the record, here is an example that shows that equality between \( m' \) and \( X(m) \) does not hold. Let \( rt \) consist of specification variables \( a, b \), and let \( frame \) be \( b \). Let these be the only things visible in \( E \). In \( Prog \), let both \( a \) and \( b \) depend on \( c \).
Then, \( \text{Resolve}(rt, E) \setminus \text{Resolve}(frame, E) \) is \( a \), and \( X(a) \) is \( a, c \). However, 
\( \text{Resolve}(rt, Prog) \setminus \text{Resolve}(frame, Prog) \) is simply \( a \).

12.4.2 \( X \) AND \( wlp \)

Coping with the last proof obligation, I will now show (12.12), that is,

\[
\text{wlp}(gc, X(Q), Prog) \iff X(\text{wlp}(gc, Q, E))
\]

by induction over the structure of \( gc \).

For \( \text{skip} \),

\[
\text{wlp}(\text{skip}, X(Q), Prog) = \begin{cases} \text{skip} \\ X(Q) = \begin{cases} \text{skip} \\ X(\text{wlp}(\text{skip}, Q, E)) \end{cases} \end{cases}
\]

and for \( \text{wrong} \),

\[
\text{wlp}(\text{wrong}, X(Q), Prog) = \begin{cases} \text{wrong} \\ \text{false} = \begin{cases} X \\ X(\text{false}) = \begin{cases} \text{wrong} \\ X(\text{wlp}(\text{wrong}, Q, E)) \end{cases} \end{cases} \end{cases}
\]

For assignment,

\[
\text{wlp}(v := expr, X(Q), Prog) = \begin{cases} \vdots \\ X(Q)[v := expr] = \begin{cases} \text{substitution involving only program variables distributes} \\
\text{over } X, \text{since } v := expr \text{ can be written in } E \end{cases} \\ X(Q[v := expr]) = \begin{cases} \vdots \\ X(\text{wlp}(v := expr, Q, E)) \end{cases} \end{cases}
\]

Next, for the hardest one, the procedure call. Let \( p \) denote a procedure declared (in \( E \)) by

\[
\text{spec } p \text{ is modifies } fr \text{ requires } Pr \text{ ensures } Po
\]
Let $t, t', u, u', B, B'$ satisfy

\[ t = \text{Resolve}(fr, E) \]
\[ t' = \text{Resolve}(fr, Prog) \]
\[ u = \text{list of initial-value variables from } F(Po) \]
\[ u' = \text{list of initial-value variables from } F'(Po) \]
\[ bs = \text{BenSideEffects}(t, E) \]
\[ bs' = \text{BenSideEffects}(t', Prog) \]

Then, realize that $t' = X(t)$ (from (12.17)), $u' = X(u)$ (from (12.6)), and $bs \leftarrow bs'$ (from (12.13)).

\[
wlp(call\ p, X(Q), Prog) =
\begin{array}{l}
\{ \text{call } p \} \\
F'(Pr) \land (\forall \text{Resolve}(fr, Prog) \mid F'(bs' \land Po) \triangleright X(Q)) \{ u'_0 := u' \}
\end{array}
\]
\[
\Leftrightarrow
\begin{array}{l}
\{ bs \leftarrow bs' \text{, and monotonicity } \} \\
F'(Pr) \land (\forall \text{Resolve}(fr, Prog) \mid F'(bs \land Po) \triangleright X(Q)) \{ u'_0 := u' \}
\end{array}
\]
\[
= 
\begin{array}{l}
(12.6), \text{ and } t' \\
X(F(Pr)) \land (\forall t' \mid X(F(bs \land Po)) \triangleright X(Q)) \{ u'_0 := u' \}
\end{array}
\]
\[
= 
\begin{array}{l}
\{ \text{realizations about } t \text{ and } u \} \\
X(F(Pr)) \land (\forall X(t) \mid X(F(bs \land Po)) \triangleright X(Q)) \{ X(u_0) := X(u) \}
\end{array}
\]
\[
= 
\begin{array}{l}
(12.19): \text{distribution of } X \text{ over quantification} \\
X(F(Pr)) \land X((\forall t \mid F(bs \land Po) \triangleright Q)) \{ X(u_0) := X(u) \}
\end{array}
\]
\[
= 
\begin{array}{l}
\{ \text{within the quantification, only } F(bs \land Po) \text{ can contain initial-value variables, and} \\
F(bs \land Po) \text{ contains no reference to any variable in } u' \land u, \text{ and} \\
(12.20): \text{distribution of } X \text{ over substitution} \} \\
X(F(Pr)) \land X((\forall t \mid F(bs \land Po) \triangleright Q)) \{ u_0 := u \}
\end{array}
\]
\[
= 
\begin{array}{l}
\{ \text{distribution of } X \text{ over } \land, \text{ and } t \} \\
X(F(Pr)) \land (\forall \text{Resolve}(fr, E) \mid F(bs \land Po) \triangleright Q) \{ u_0 := u \}
\end{array}
\]
\[
= 
\begin{array}{l}
\{ \text{call } p \} \\
X(wlp(call\ p, Q, E))
\end{array}
\]

Now for sequential composition,

\[
wlp(s; t, X(Q), Prog) =
\begin{array}{l}
\{ ; \} \\
wlp(s, wlp(t, X(Q), Prog), Prog)
\end{array}
\]
\[
\Leftrightarrow
\begin{array}{l}
\{ \text{induction hypothesis, and } wlp \text{ is monotonic } \} \\
wlp(s, X(wlp(t, Q, E)), Prog)
\end{array}
\]
\[
\Leftrightarrow
\begin{array}{l}
\{ \text{induction hypothesis, with } Q := wlp(t, Q, E) \} \\
X(wlp(s, wlp(t, Q, E), E)
\end{array}
\]
\[
= 
\begin{array}{l}
\{ ; \} \\
X(wlp(s; t, Q, E))
\end{array}
\]

For choice,
\[
\text{wlp}(s \circ t, X(Q), \text{Prog}) = \\
\{ \circ \} \\
\text{wlp}(s, X(Q), \text{Prog}) \land \text{wlp}(t, X(Q), \text{Prog}) \\
\iff \\
\{ \text{induction hypothesis, twice} \} \\
X(\text{wlp}(s, Q, E)) \land X(\text{wlp}(t, Q, E)) \\
= \\
\{ \circ \} \\
X(\text{wlp}(s \circ t, Q, E)).
\]

For the guard statement,
\[
\text{wlp}(g \rightarrow s, X(Q), \text{Prog}) = \\
\{ \rightarrow \} \\
g \Rightarrow \text{wlp}(s, X(Q), \text{Prog}) \\
\iff \\
\{ \text{induction hypothesis} \} \\
g \Rightarrow X(\text{wlp}(s, Q, E)) \\
= \\
\{ \rightarrow \} \\
X(\text{wlp}(g \rightarrow s, Q, E)).
\]

Finally, for block,
\[
\text{wlp}([y \cdot s], X(Q), \text{Prog}) = \\
\{ [\cdot] \} \\
(\forall y \triangleright \text{wlp}(s, X(Q), \text{Prog})) \\
\iff \\
\{ \text{induction hypothesis} \} \\
(\forall y \triangleright X(\text{wlp}(s, Q, E))) \\
= \\
\{ X \text{ distributes over universal quantification where dummy is not name of specification variable} \} \\
X((\forall y \triangleright \text{wlp}(s, Q, E))) \\
= \\
\{ [\cdot] \} \\
X(\text{wlp}([y \cdot s], Q, E)).
\]

It's a wrap.

12.5 Epilogue

Now that I have proven the soundness theorem, let’s look back at the proof to see what is used where.

Firstly, note that residues are used with every occurrence of $X$, and thus appear everywhere in the proof. If instead of residues some extra requirements were used,
the proof that those requirements ensure soundness would require quite a different proof.

Secondly, as foretold by Remark 9.2, the proof does not make use of the structure of representations of specification variables. In particular, the choice of abstraction functions vs. abstraction relations does not surface as being important to the proof. Realize that the soundness proof just provides a bridge between a "one"-module proof and an all-modules proof; hence, Remark 9.2 does not imply anything positive about the effects of using abstraction relations, thus letting functions be nondeterministic, in a "one"-module proof.

Thirdly, the proof does not rely on whether or not dependencies are cyclic. It is not clear, however, for what purpose cyclic dependencies are useful.

**Remark 12.1.** If imports are acyclic, then the visibility and authenticity requirements dictate that all cycles among dependencies be placed in one unit.

Finally, the visibility requirement plays a role in proving the monotonicity property (12.13). This property is used in the proofs of \( b \subseteq b' \) (12.4) and in the distribution of \( X \) over \( \text{wp} \) for procedure calls (12.12).

The authenticity requirement comes into play in showing that program variables updated by assignment statements in gc have only benevolent side effects, i.e., do not alter the value of, any specification variable in \( b' \setminus b \). The authenticity requirement is applied twice in the proof, once for each partition of \( b' \setminus b \), by and \( \text{fix} \).

Notice that the visibility and authenticity requirements are used in separate places in the proof. Hence, for the proof, it is preferable to have a division between visibility and authenticity requirements, rather than making use of the convention presented in Section 10.3.
depends in perspective

In this chapter, I give a flavor of the successes and shortcomings of \texttt{depends}. I start by showing how \texttt{depends} can be used to specify \textit{consumer} objects. Then, I discuss some problems for which \texttt{depends} is not the whole solution. I don’t solve these problems here, but I conclude the chapter with some insights into the problems.

13.0 Specification of a consumer

In this section, I describe a difficult specification problem for which I then show a solution in my formalism.

13.0.0 Consumers

Consider the following unit, which provides a type of so-called \textit{consumer} objects.

\begin{verbatim}
unit Consumer is
  type T ;
  spec Consume(s : seq[char] ; t : T) is
    (* for each character in s, invoke t.consume *) ;
  method t : T spec consume(ch : char) is
    (* . . . *)
end
\end{verbatim}

The idea is to call \texttt{Consume} with a character sequence and an object \texttt{t} of some subtype of \texttt{T}. This invokes \texttt{t.consume} for each character of the given string. (In functional programming, an operation like \texttt{Consume} is called a \textit{map} (see, \textit{e.g.}, [6]).)

Another example of a consumer takes a \textit{reader}, \textit{i.e.}, an input stream (\textit{cf.} [9] and Chapter 10), and passes character sequences read from the reader to \texttt{t.consume} in arbitrary-size pieces.

The task is to write the specification of procedure \texttt{Consume} and method \texttt{consume}. The problem is —and it is for this reason that I earlier referred to this problem as being difficult— doing so without knowing what \texttt{t.consume} requires, modifies, or ensures for all subtypes of \texttt{T}.
**Remark 13.0.** In his thesis, Jackson says for Aspect that "it is not clear how to solve this problem" [41, Sec. 8.3], and leaves the problem without a solution.

The specification must admit a provably correct implementation and must be flexible enough to be useful for subtypes of $T$.

### 13.0.1 A Solution

I present a solution to the problem. For simplicity, I let the postcondition of the consume method be $true$. Other postconditions can be handled similarly to the way I handle the precondition.

**Remark 13.1.** For the purposes of extended static checking (see Section 0.0), using a postcondition of $true$ in cases like this is quite common. The precondition, and the fact that some invariant properties are preserved, are more important.

I introduce two properties of consumers, $valid$ and $state$.

```plaintext
spec var valid : $T^- \rightarrow bool ;
spec var state : $T^- \rightarrow any
```

The type $any$ indicates that the particular value of $state[t]$ is not relevant. Consequently, no $rep$ is given for $state[t]$, but $depends$ clauses are still used to declare the concrete variables on which $state[t]$ depends. This allows subtypes of $T$ to declare the variables that make up their states. These concrete variables can then be modified when $state$ is allowed to be modified. This important mechanism is not available in classical data refinement [38].

I write a specification of the consume method in terms of the properties $valid$ and $state$.

```plaintext
method t : $T$ spec consume(ch : char) is
  requires $valid[t]
  modifies state[t]
```

An invocation of consume requires that $t$ be valid. This essentially means that $t$ satisfies its object invariant (cf. [64] or type constraints in [57]). Unlike [64, 57], I make it explicit when the invariant should hold by explicitly stating $valid[t]$ as a precondition. Since $valid[t]$ is not mentioned in the $modifies$ clause, the value of $valid[t]$ must not be changed by the method; hence, the object will be valid upon exit, too. This makes object invariants a convention rather than a canned feature of the specification language (see also Remark 10.3).

The specification of consume allows the state of object $t$ to be modified. The only restriction on this modification is that the values of properties like $valid$ that are not listed in the $modifies$ clause are unchanged.
The specification of \textit{Consume} can now be written. As it turns out, this specification coincides with that of method \textit{consume}.

\begin{verbatim}
spec Consume(s : seq[char] ; t : T) is
  requires valid[t]
  modifies state[t]

impl Consume(s : seq[char] ; t : T) is
  if s \neq "" then t.consume(first(s)) ; Consume(rest(s), t) fi
\end{verbatim}

These specifications allow an implementation of \textit{Consume} to iterate over the characters of \textit{s}, passing each one, in sequence, to \textit{t.consume}.

Without stating a stronger postcondition, this particular specification is not strong enough for full verification. For example, an implementation of \textit{Consume} meets its specification even if it invokes the method twice for some characters of \textit{s}, never for others, and the invocations of \textit{consume} for different characters of \textit{s} are not in the same order as the order in which they appear in \textit{s}. Other correct implementations may invoke \textit{consume} with characters not in \textit{s}, or may never call \textit{consume} at all. However, for extended static checking, this specification may still be strong enough, depending on the assumptions made about the consumer after the call to \textit{Consume}.

Each subtype defines for itself what it means for one of its object to be valid and what the state of one of its objects is. Hence, the solution works for any consumer subtype.

13.1 Shortcomings of \textit{depends}

From what we have seen, \textit{depends} is promising, but it is not the end of the story. There are more problems to be solved, because the visibility and authenticity requirements are too strong for the solution to be adequate.

The solution is adequate as long as there is only one level of specification variables, \textit{i.e.}, so long as no specification variable depends on another specification variable. However, when one module is implemented in terms of another, there is usually more than one level.

Let me give an example of a problem where the visibility and authenticity requirements are too strict. The \textit{TextWrlmpl} unit in Chapter 10 declares the \textit{flushed} data field of each text writer to be of type \texttt{seq[char]}. The \texttt{seq[char]} type available through the programming language at hand (\textit{e.g.}, \texttt{TEXT} in Modula-3) may not provide the most efficient implementation for the task in \textit{TextWrlmpl}, or the language may not directly provide this type at all. We may then want to use a custom-programmed implementation of sequences of characters.
Consider a unit \textit{FastSeq}, declared as follows.

\textbf{unit FastSeq is}
\begin{verbatim}
type T;
spec contents : T → seq[char];
spec init(t : T) is
  modifies contents[t]
  ensures contents[t] = "" |
spec result : seq[char] := Contents(t : T) is
  ensures result = contents[t] |
spec Append(t : T ; ch : char) is
  modifies contents[t]
  ensures contents[t] = contents0[t] ++ ch
\end{verbatim}
\textbf{end}

Such a unit allows a change of unit \textit{TextWrImpl} to

\textbf{unit TextWrImpl import Wr, WrFriends, TextWr, FastSeq is}
\begin{verbatim}
var flushed : TextWr.T → FastSeq.T |
depends Wr.target[t : TextWr.T] on FastSeq.contents[flushed[t]] |
rep Wr.target[t : TextWr.T] is
  Wr.target[t] = FastSeq.contents[flushed[t]] ++ Wr.buf[t] |
\end{verbatim}
\textbf{end}

Here, \textit{Wr.target} is declared to depend on \textit{FastSeq.contents}. But according to the (convention suggested by the) two requirements, this dependency must be declared in unit \textit{FastSeq}. That means that \textit{FastSeq} needs to import \textit{Wr} and give the dependency. The same holds for any other similar client of \textit{FastSeq}. Thus, \textit{FastSeq} needs to import all of these clients! That is unreasonable, because the implementor of \textit{FastSeq} cannot anticipate all of \textit{FastSeq}'s clients, so with every new client, \textit{FastSeq} needs to be updated. This makes it practically impossible to put \textit{FastSeq} in a library.

Furthermore, in order for \textit{FastSeq} to be able to declare the dependency shown in \textit{TextWrImpl} above, the data field \textit{flushed} must be visible in \textit{FastSeq}. This goes beyond the notion of "friends" interfaces and violates the notion of data hiding.

The problem described also surfaces, for example, in the consumer example in Section 13.0 if a consumer subtype depends on a data field like \textit{Wr.target} or \textit{FastSeq.contents}.

Let us take a closer look at the shape of the \textbf{depends} clause that's causing problems. In examples, I have freely used \textbf{depends} clauses of the form
\begin{verbatim}
depends a[t : T] on c[t] |
\end{verbatim}
I call this a \textit{pointwise dependency}, because \textit{a} at a particular index depends on \textit{c} at some particular index. Here, the two indices are the same. Chapters 11 and 12 explore dependencies only between entire variables, as in
\begin{verbatim}
depends a on c |
\end{verbatim}
Understanding pointwise dependencies and developing a methodology for their use are on the path to a solution to the problem I've described in this section.

Another closer look at the shape of the `depends` clause that presents the challenge shows

```
depends a[t : T] on c[b[i]]
```

That is, `a` depends on `c` pointwise, but not at the same points. Instead, `b` serves as a function from the point of `a` to the dependent point of `c`. All three of `a, b, c` need to be visible in order to state the dependency. This reminds us of the visibility requirements. Furthermore, the intermediary variable `b` is declared in the same unit as the dependency itself. This leaves us with a scent of the authenticity requirement. These observations give us hope that there may be a way to modify the visibility and authenticity requirements to fit the needs described here. This is an open problem. In the next section, I discuss some considerations of its solution.

### 13.2 Private values

In the example presented in the previous section, the `FastSeq.T` value of `flushed[i]` can be thought of as being "private" to the unit (or to the object `t`). In this section, I develop the idea of private values further. I do not present a solution to the problem of writing specifications involving private values, but I discuss some issues that a solution should take into account.

Values may be "private" not just to modules, but also to procedures (as I show later in this chapter) and possibly also to objects or object types (not shown here).

#### 13.2.0 A GIVEN SPECIFICATION

Consider a unit `A` and its friends interface `AFriends`.

```
unit A is
  type T ;
  spec var valid : T\rightarrow bool ;
  spec var state : T\rightarrow any ;
  spec Init(t : T) is
    modifies state[t], valid[t]
    ensures valid[t] ;
  spec Update(t : T) is
    requires valid[t]
    modifies state[t] ;
  spec Destroy(t : T) is
    modifies state[t], valid[t]
end
```

148
unit AFriends import A is
    var previous : A.T ;
    depends A.state on previous ;
    depends A.valid on previous
end

Also consider B, a unit declared as follows.

unit B is
    type T ;
    spec var valid : T → bool ;
    spec var state : T → any ;
    spec Init(t : T) is
        modifies state[t], valid[t]
        ensures valid[t] ;
    spec Update(t : T) is
        requires valid[t]
        modifies state[t]
end

The implementation of B reveals one of type B.T’s data fields.

unit Bimpl import B, A is
    private var x : B.T → A.T ;
    ...
end

Notice that B is a regular client of A that does not import AFriends.

The keyword private is intended to indicate that field x is never “imported” into or “leaked” from this module. (I intentionally omit the exact definitions of those terms—the inclusion of these words is just supposed to convey a flavor of the full meaning.) The motivation for this is as follows. The state and validity of a B.T object depends on the state and validity of the fields of a B.T. Thus, Bimpl gives the following information about its specification variables.

    depends B.valid[t : B.T] on A.valid[x[t]] ;
    depends B.state[t : B.T] on A.valid[x[t]] ;
    rep B.valid[t : B.T] is valid[t] = A.valid[x[t]]

But this violates the visibility and authenticity requirements. Our first reaction is that the violation is for silly reasons; if B.T happens to use an A.T in its private implementation, why does this need to be advertised in A, which knows nothing about B? The private keyword used in conjunction with the declaration of field x is an attempt at allowing the implementation to depart from the visibility and authenticity requirements. The details of private are still under experimentation.
Unit Bimpl also includes the implementation of the procedures declared in \( B \).

\[
\text{impl } B.\text{Init}(t : B.T) \text{ is} \\
\text{call } A.\text{init}(x[t]) ; \\
\text{impl } B.\text{Update}(t : B.T) \text{ is} \\
\text{call } A.\text{Update}(x[t])
\]

13.2.1 A VIOLATION OF SOUNDNESS

Let’s take a look at how a client may use \( A \) and \( B \).

\[
\text{unit } Client \text{ import } A, AFriends, B \text{ is} \\
\text{spec } P() \text{ is} \\
\text{modifies } A.\text{state}, A.\text{valid}, B.\text{valid}, B.\text{state} ; \\
\text{impl } P() \text{ is} \\
\begin{array}{c}
[ \\
b : B.T \bullet b := \text{new}(B.T) \\
; B.\text{Init}(b) \\
; \text{if } \text{AFriends.previous} \neq \text{nil} \text{ then } A.\text{Destroy}(\text{AFriends.previous}) \text{ fi} \\
; B.\text{Update}(b) \\
\end{array}
\]
\]

The question here is: Does \( P \)'s call to \( B.\text{Update} \) meet its required precondition? From the information given and the rules from Chapter 11, the verification would indeed verify \( P \) as having the required precondition of \( B.\text{Update} \). But consider an implementation of \( A.\text{init}(t) \) that, in addition to initializing \( t \), sets \( \text{AFriends.previous} \) to \( t \). Then we do not want the verification process to deem the implementation of procedure \( P \) correct, because there is no guarantee that the call to \( B.\text{Update} \) actually meets its precondition.

The problem seems to be that although Bimpl vowed (by using the keyword private) not to leak a value of an \( x \) field, Bimpl passed such a value to the procedures of \( A \), and \( A \) never made a promise not to leak such values. Hence, we don’t want Bimpl to get away with declaring \( x \) as private unless \( A \) guarantees that instances of \( A.T \) are “privatizable”.

To recap, I am discussing the circumstances under which the declaration of dependencies of a specification variable are allowed to depart from the visibility and authenticity requirements. In the above, in order for \( B.\text{valid}[t] \) to be declared to depend on \( A.\text{valid}[x[t]] \), \( A.\text{valid} \) must have been declared to be privatizable. If \( A.\text{valid}[t] \) is privatizable, its potential dependencies are restricted. A research goal is to find a good set of rules for this. For example, one may require that \( A.\text{valid}[t] \) must only depend on entities of the form \( w[i] \) where \( w \) is also privatizable. As a base case of the definition of privatizable, all program (i.e., non-abstract) data fields of a type are privatizable.

150
13.2.2 PRIVATE VALUES AND new

I present an issue related to the one above. Consider a unit S and its implementation.

```
unit S is
  spec kip() is
    modifies (* nothing *)
  end

unit Simpl import S is
  impl kip() is
    [ a • a := new(A.T) ; A.Init(a) ]
  end
```

A question to ask is: Does Simpl contain a correct refinement of S.kip? Since it has an effect on A.state and A.valid, the implementation would not be considered correct by the rules from Chapter 11. Nevertheless, one may be inclined to answer, “Yes, this is a correct refinement, because the value of a is conceived within this procedure (i.e., it is not imported, and since the procedure has no return value, a is also not leaked from the procedure) and should thus be considered a private value of the procedure”.

In response to this answer, let me define another unit.

```
unit M import S, A, AFriends is
  spec P() is
    ensures true ;
  end

  impl P() is
    [ p : A.T • p := AFriends.previous
      ; S.kip()
      ; if p ≠ AFriends.previous then wrong fi
    ]
  end
```

Using the rules from Chapter 11, the implementation of M.P is considered correct. Thus, also validating S.kip would not be sound, because then a trace of M.P would actually go wrong.

Therefore, the quoted claim above is flawed. The problem, as with B and Bimpl above, is that A.state and A.valid are not guaranteed to be privatizable.

13.2.3 SUMMARY

In this section, I have argued for a notion of “privatizable”. If w is privatizable and t is a private value of a procedure, then modification of w[t] is allowed, even if w is not mentioned in the frame. Similarly, if t is a private value of a unit, then dependencies on w[t] declared in that unit are allowed, even if w is declared elsewhere. The exact details of “private values” and “privatizable” are still under experimentation, and soundness will be a function of these details.
Part IV

Epilogue
Chapter 14

Summary

In summary, this thesis is about the correctness of programs, and is in the direction of making the specification and verification of large programs —of the kinds that are written in practice— feasible. In particular, I deal with sequential programs that can raise and handle exceptions; programs whose expressions can be partial, can contain short-circuit boolean operators, and can have side effects; programs whose data structures include arrays, sets, records, references (pointers), and objects with methods and inheritance; and programs that achieve data hiding by being organized into modules.

The thesis explores several aspects of the semantics of programs with exceptions, and unveils the algebraic cosmos on which the foundation of this semantics rests. The thesis also takes a fresh look at object-oriented programs, and proposes a simple notion of their mathematical meaning.

Striving toward making large programs more reliable, I confront the specification and verification of modular programs, with the goal of achieving sound modular verification. This thesis solves one of the problems and reports on others that are still open.

There is more work to be done before practical tools that aid in the construction of large programs that are correct are used routinely by programmers. Type-checking compilers have become commonplace, and it seems only natural that a compiler or other tool for extended static checking will be next. Through something like extended static checking, it is my hope that it won’t be too long before the science of program correctness will become a practical aid in the everyday life of programmers, so that we all will write correct programs more often.
Bibliography


159


Index

; , see sequential composition
[ := ] , see substitution
[ ] , see everywhere operator
∧ (conjunction), 5
⊙ , see box operator
⊕ (concatenation), 98
⊗ , see else operator
· (multiplication), 7
¬ (negation), 4
∨ (disjunction), 5
⟨ I ⊳ ⟩ , see quantifications
⊑ , see refinement
{ I ⊳ }, see set constructor
<: (subtype relation), 78
⟨⟩ , see double composition
⇝ , 21
[ ] , see ceiling
⟨⟩ , see left composition
[ ] , see floor
⟨⟩ , see right composition
∼ , see transposition
× , 68
→ (guard), 18
→ (map), 68
<1 , see exceptional composition
* , 74, 112, 114
− shorthand , 74
0 , subscripting a variable by, see initial-value variables

abort , 22, 42
abstract interpretation, 22
abstract variables , 84, 111
and assignment ( := ) , 107
and refinement, 86
definition of usage in frames, 101
interpretation of, in predicates, 101
interpreted as explicit functions, 107
local, 84
mixed with program variables, 88
origin of name, 85
representation of, 85
abstraction, see data abstraction
abstraction functions, 85, 142
and subtyping, 89
axioms resulting from, 117
in scope vs. out of scope, 104
partial, 89
totality of, 86
vs. abstraction relations , 85, 142
abstraction relations, 85, 142
acknowledgements, v
Ada [1], 8, 12, 56, 50, 60, 68, 69, 74, 89, 96, 108
initialization of variables in, 69
references in, 69, 74
algebra over functions of two arguments, 12, 17, 26–35
applications of, 17, 35
generalization of, 35
motivation for, 17
unit elements in, 27
aliasing, 74, 83
allocated type, see dynamic type
allocation, 74, 150, 151
alternative statements ( IF THEN ), 59
Anna [57], 97, 108, 109
arity
change of, 135
array dereferencing, 68
array index out-of-bounds errors, see
run-time errors
array types, 71
arrays, 66
  multi-dimensional, 72
Aspect [41], 106, 145
assert statement (assert), 19, 61
assertions
  failing, see run-time errors
assignment (:=), 16, 38, 112
  and abstract variables, 107
narrowing in, 70
  shorthand for maps (arrays), 68
AT&T crash, 1990, 0
authentic abstractions, 103
authentic variables, 103
authenticity property, 104
authenticity requirement, 104, 106,
  128, 134, 137, 146
  deviating from, 150
  dispensing with, 104
  enforcing, 104
  generalizing, 150
AuthenticityProblem unit, 103
auxiliary variables, 88

Back, R.-J.R., 24
backtracking, 18
benevolent side effects, 102, 103, 117
  examples, 102, 117
BenSideEffects, 117
BenSideSet, 118
  reducing size of, 118
binary semaphores, 48
binding powers, 4
bits, 84
blatant hand waving,
block statement ([•]), 17, 112
body
  of block statement, 17
  using guard statement in, 19
  of iterative statement, 22
  of procedure, 56
bool, 68
boolean scalar, 6, 120
bound function, 22
box operator (☐), 20, 112
buff data field, 98–102
  comment describing usage of, 99
C++ [23], 8, 12, 74, 78, 81, 82, 91
  references in, 74
  terminology, 78, 82–83
calculations
  format of, see proof format
call, 58, 112
Caltech, v
cand, 62
ceiling ( ⌈ ⌉ )
  of functions, 30
  of programs, 34
  of states, 38
Chandy, K. Mani, v
chapter dependencies, 3
choice compositions, 20
choice operators, 20
Civitype, v
classical data refinement, 2, 3, 86
  and Simula, 108
  generalization of, 87, 106
  shortcomings of, 88
client, 8, 96, 100, 102, 105, 108, 109,
  147, 149, 150
CLU [54], 51, 109
code reuse, 56
collaboration, 4
collections, 74
command, 8
compositions
  duality between, 33
  exceptional, 17, 42
  normal, 16, 39
conditional operators, see short-circuit
  operators
conjunction ( ∧ ), 5
conjunctivity, 6
  monotonicity implied by, 6, 44, 119
  of wlp, 119
  of wp, 44

163
consume method, 144
Consume procedure, 144
consumer example, 144
problem with, 147
Consumer unit, 144
consumers, 144
contents (of FastSeq.T), 147
contributions, see thesis contributions
control structures, 14
convention, see requirements, convention
correctness
partial, 15
total, 15
cyclic index entries, see index entries, cyclic
C [45], 8, 12, 56, 59, 68–70
initialization of variables in, 70
references in, 74
types in, 70
data abstraction, 66, 84
importance of, 88
in modular programs, 96
data fields, 78
abstract, 89
in record types, 72
of objects, 79
on notation of, 81
protection levels, 83
data hiding, 94, 96, 105, 147
data refinement, 66, 84, 86
and partial commands, 88
classical, see classical data refinement
in modular programs, 66
in subtypes, 80
data structures, 66
data types, see also types
declaring new, 73
in common programming languages, 70
meaning of, 80
relation between, 80
deallocation, 74
DEC SRC, see Digital
declaring type, 79
demonic choice, see box operator
Denver International Airport, 0
dependencies, 101
cyclic, 142
inferring, 108
pointwise, 147
dependency graph, 106
dependency relation, 113
Depends, 113
depends, 3, 101, 112
formalization of, 110
placement of, 102
restrictions on usage of, 102
enforcing, 104
shortcomings of, 146
successes of, 144
derference (^), 73, 78
derference map, 74
Detlefs, David, v
Digital
SRC Extended Static Checker, v
Systems Research Center, v, 4
Dijkstra, Edsger W., iii, 1, 12, 24, 35, 43
disjunction (V), 5
disjunctivity, 6
distributivity, 6
double composition (/, 28
downward closure, 101, 113
dynamic type, 78
element type, 68
ELSE, 59
else operator (if), 20
ELSIF, 59
ensures, 57
entry variables, see initial-value variables
enumerations, 66, 71
Env, 119
environment, 112
erroneous (Ada lingo), 63
errors
  in software, see software errors
Euclid [47], 74
everywhere operator ([ ]), 5
  for different state spaces, 124
except-ensures, 57
exception
  definition, 2
exception handler, see exceptional composition
exceptional composition (<|>, 17
exceptional programs
  definition, 3
exceptional projection, see ceiling
exceptional semantics
  previous work, 35
exceptions, 12
  in program construction, 52, 54
  more than one, 35, 38, 60
  theorem regarding usage of, 51
Excluded Middle, Law of, 18
explicit functions, 107
expressions, 61–63
  and compilers, 62
  evaluation of, 16, 61–63
extended static checking, v, 1, 2, 145,
  154
  vs. full verification, 1
  idea of, 1
f* functions, 114
false, 6
  outcome component implicitly being, 60, 119
FastSeq unit, 147
FaultyClient unit, 100
feasible statements, see partial commands
Feijen, Wim H.J., 7
fields, see data fields
file writers, see writers, file
fin, 38
fixed points, 21
floor ([ ])
  of functions, 30
  of programs, 34
  of states, 38
flush method, 99
flushed data field, 100, 146
forest requirement, 106
frame, see specifications, frame of
FREE, 75
friends interfaces, 97, 98, 147
function application, 4
function compositions, 26
  programs as, 33
functionalization, 114
Functionalize, 114
functions, 4
  nondeterministic, 85
  of two arguments, 17, 26
  uninterpreted, 126
GetSpec, 117
global in, 63
global variables, see variables, global
van der Goot, Marcel R., v
guard function, 19
guard of a command, 19
guard statement (→), 18, 112
guarded commands, 8, 12, 14
Harley, Robert J., v
heuristic for program construction, 52
Heydon, Allan, v
Hoare triples, 1, 48
Hoare, C.A.R., iii, 1, 84
Hofstee, H. Peter, v
identifiers
  naming of, 111
if fi brackets, 20
IF THEN , 59
if then fi, 59
impl, 58, 111
import, 96
import, 111
import closure, 111
ImportClosure, 111
imported values, 149
index entries
cyclic, see cyclic index entries
index set, 68
infinite, 81
index type, 68
indices of maps, 68
inf, 38
ingrown toenails?, see physician
inheritance, 79
initial-value variables, 22, 23, 57, 75, 111, 121
input stream, 144
input/output streams, 96, 98, 144
instantiation, 136
int, 68
interfaces, 89, 96, 97, 110
intermediate code, 62
Invariance Theorem, 21, 48
iteration, 21
iterative statement (do ⇒ od), 21
absence of, 112
Jackson, Daniel, 145
joke (getting to know my readers), vi
junctivity, 6
Juno-2 [35], 20
Knaster-Tarski Theorem, 21
L projection, 17, 27
Larch [33], 57
last, 38
LATEX, v
lattices, 12, 24
leaked values, 149
left composition (⊙), 17, 27
Leibniz’s Rule, 116, 126, 131, 134
Leino, India Jane, vi
libraries, 94
of axioms, 109
lifting, 5, 28
LIM [10], 18
logical variables, see initial-value variables
Lookup, 117
loop, see iterative statement
infinite, 22
Lukkien, Johan J., 37
Manasse, Mark S., 36
Manohar, Rajit, v
map (in functional programming), 144
map types, 68
map* maps, 74, 78
maps, 68
axioms for, 69
data fields, see data fields
implementation of, 69, 81
in specifications, 75
totality of, 68
updating, 68
Martin, Alain J., v
Massingill, Berna L., v
mathematical variables, see initial-value variables
method, 79, 80
impl, 80
spec, 79
methods, 78, 79
C++ lingo for, 83
declaration of, 79
implementations, 79
invocation of, 79
protection levels, 83
virtual (C++ lingo), 83
Microsoft, v, 2
personal experiences at, 2
miracle
invoking, 18
miraculous statements, see partial commands
modifies, 57
abstract variables listed by, 101
and maps, 75
in modular specifications, 97
Modula-2 [81], 8, 12, 89, 96, 108
Modula-3 [71], 8, 12, 34, 51, 56, 59, 60,
   68–74, 80–82, 89, 91, 96, 98, 99,
   101, 105, 108, 110, 146
object types in, 81–82
partially opaque types in, 101
references in, 74
structural equivalence in, 82
modular programs
   and SwitchToAbstract, 87
definition, 2
   precise interpretation of, 3
modular specifications
   problem of writing, 97
modular verification, 2, 3, 94, 97
   soundness of, 3
   adequacy of, 105
   completeness of, 105
   example of inadequacy of, 106
   problem in, 94
   soundness of, 94, 105, 128
modularity, 94
module invariant, 103
modules, 89, 96, 97, 110
monotonicity, 6, 15, 21, 25, 44
Moye, S.G., v

name qualifier (:), 98, 111
narrowing, 70, 71
   in abstract interpretation, 22
Nelson, Greg, v, 4, 36
new, 74, 150, 151
nil, 73, 78
nil -dereferencing, see run-time errors
nondeterminism, 20, 24
nondeterministic functions, 85
normal composition, see sequential composition
normal projection, see floor notation
   explanation of, 4
   how it inspires us, 31
   on the choice of, 35

object invariant, 103, 145
object simplicity, 80
object types, 73, 78
object-oriented programming
   essence of, 99
objects, 66
   attributes of, 79
   creating new, 74
   implementation of, 81
   properties of, 79
   simplicity in modeling, 66
oc, see outcome coordinate
operators
   binding powers, see binding powers
originality, 4
outcome, 15, 38, 60
   erroneous, 60
   exceptional, 15
outcome coordinate (oc), 37
outline
   of Part I, 12
   of Part II, 66
   of Part III, 94
   of thesis, 3
output stream, 98
packages (Ada lingo), 89
parameters
   evaluation of, 63
   narrowing of, 70
   of methods, 79
   of procedures, 57
   special, 79
partial commands, 18
   examples of, 19
   operational interpretation of, 18
partial correctness, see correctness, partial
partial representations, 89
   reason for restriction on, 90
   restriction on, 90
partially opaque types, 80, 101
partitioned predicates, 35, 43, 112
Pascal [42], 8, 12
early version of [79], 74
Penelope [62], 57, 63, 109
pointers, see references
pointwise dependencies, 147
postcondition
in specifications, 22
specifying for procedure, see ensures or except-ensures
POSTSCRIPT, v
precondition
in specifications, 22
specifying for procedure, see requires
weakest, see weakest preconditions
weakest liberal, see weakest liberal preconditions
predicate transformers, 5, 12
predicates, 5, 12
interpretation of abstract variables in, 101
over two states, see two-state predicates
prime factor example, 7
private, 149, 150
private, 83
private values, 148, 151
privatizable, 151
procedures, 56, 111
calls to, 58
implementation of, 58
parameters, 57
evaluation of, 63
narrowing of, 70
result values, 57
narrowing of, 70
signature of, 56
specification of, 56
tail-recursive, 112
termination of, 59
program constructs
implementation of, 81
implementing, 24, 36
program derivation, 12, 52
program state space, 16, 37
program statement, 8
program variables, 84, 111
as functions, 84
mixed with abstract variables, 88
programming methodology, 12, 81, 91, 148
programming notation, 14, 56, 110
for programs with units, 110
programming tools, 154
projections
L and R, 27
on states, 38
proof format, 7
proof hints, 7
vague, 123
protected, 83
protection levels, 83
public, 83
PutChar procedure, 98
quantifications, 5
dummy of, 5
existential (∃), 5
notation for, 5
range of dummy in, 5
range omitted in, 5, 7
set constructor, see set constructor
term of, 5
union (∪), 5
universal (∀), 5
R projection, 17, 27
raise, 16, 39
free occurrences of, 49
RawTarget, 119
readers, 144
receiver, 79
record types, 66, 72
REF, 73
reference types, 73, 78
references, 66, 73
creating new, 74
crux with modeling, 74
dereferencing, 73
Refine, 110, 112, 120
monotonicity of, 129
refinement (\(\subseteq\)), 15, 24, 58, 94, 120
and abstract variables, 86
data, see data refinement
proving, 122
rep, 85, 112
RepAxiom, 117
RepAxioms, 117
RepPreds, 117
representation function, see abstraction function
requirements, 102
convention, 104, 105, 142
necessity of, 105
deviating from, 150
enforcing, 104
explaining to programmers, 105
generalizing, 150
too strict, 146
requires, 57
reserved character, see *
residues, 102, 110, 124
importance of, 124
Resolve, 113
shorthand for equalities, 113
resolve set, see downward closure
Rifkin, Adam F., v
right composition (\(\circ\)), 17, 27
run-time checks, 62
run-time errors, 60
checked, 1, 62
proving absence of, 1
checks for, 62
unchecked, 63
S.kip procedure, 151
Sanders, Beverly A., v
saving variables
observation regarding, 23
scope, 98
self, 79
semantics, 1, 3, 12
concreteness of, 36, 47
correctness of, 36
influence of, 1
operational, 3, 37
separate compilation, 94
seq[char], 98, 146
sequential composition (\(\cdot\)), 16, 112
set constructor, 6
set difference (\(\setminus\)), 6
set types, 73
sets
as data types, 66
explanation of notation, 6
short-circuit operators, 61
side effects
benevolent, see benevolent side effects
in expressions, 16
signature of procedures, 56
simulation of types, 80
Simula [14], 108, 109
Sivilotti, Paolo A.G., v, 124
skip, 16, 39, 112
van de Snepscheut, Jan L.A., 1953–1994, v, 4
software
today, iii, 0
software errors, 0
possible causes of, 0
soundness
violation of, 100, 103, 105, 150, 151
soundness of modular verification, see modular verification
soundness proof
in retrospect, 141
soundness theorem, see modular verification, soundness of
spec, 56, 111
spec var, 84, 111
specification languages
related work, 108
specification statement, 22, 57, 79
being partial, 23, 24
implementation of, 24
operational interpretation of, 22
specification variables, see abstract variables
specifications, 1, 97
frame of, 22, 97, 111
full, 1
lack of understanding of writing, 3
modular, 96
  basic problem with, 100
  solution that problem, 101
of a consumer, 144
of modular programs, 3
of procedures, 56
of streams library, 96
postcondition in, 22, 111
precondition in, 22, 111
SpecVarProjection, 118
state, see program state space
state, 102, 145
state space, 66
statement, 8
statement compositions, 16
  asymmetry between, 34
statements
  that "go wrong", 60
states
  normal and exceptional, 38
streams library, 96
strict commands, 18
subrange types, 66, 71
subscripting a variable by 0, see initial-value variables
subst, 69
substitution ([ := ]), 4
X, see X
substitution theorem, 136
substitutions
  duplicate, 115
subtype relation ( <: ), 78, 80, 99
subtypes, 78
  proper, 78
purpose of, 99
subtyping, 66, 78, 80
  considered independently of abstraction, 80
  methodology, 81, 91
  notion of, 66
summary, 154
supertype, 78
  immediate, 78, 80
SwitchToAbstract, 86, 107
  and modular programs, 87
tail recursion, 112
target, 98
  buffered portion of, 99
  flushed portion of, 99
  representation of, 100
Target (function), 119
Target (procedure), 99
target set, 119
temperature unit conversion example, 85
termination, 14, 15, 21, 22, 59, 60
  normal, 14
  proving, 15, 22
testing, 0
TEXT, 98, 146
text writers, see writers, text
TextWrl, 99
TextWr.T, 99
TextWrImpl, 99, 146
then, 59
theorem provers (automatic), 1
theorem proving
  automatic, 1
tree theory packages, 109
Therac-25, 0
thesis contributions, 2, 4, 154
this, 79
Thornley, John, v
three-address code, 62
tools, 154
total commands, 18
total correctness, see correctness, total
trace, 38
trace semantics, 36
trace sets, 37
  programs as, 38
transitive reduction, 106
transposition (\(\sim\)), 32
  implementation of, 34
true, 6
  as postcondition, 145
  range, omitted, 7
TRY FINALLY END, 34
two-state predicates, 23, 57, 124
type
  declaring, see declaring type
type, 73, 78
type checking, 2, 154
type constraints, 145
types, 68
  allocated, see dynamic type
  composite, 68
  data, see data types
dynamic, see dynamic type
  element, see element type
  enumeration, see enumerations
  index, see index type
  map, see map types
nonempty, 68
of variables, 68, 69
record, see record types
set, see set types
structural equivalence among, 82
subrange, see subrange types
subtyping, see subtyping

unit, 111
unit statements, 16, 39, 60, 112
  raise, see raise
  skip, see skip
  wrong, see wrong
units, 97
  convention of naming, 99

valid paradigm, 103, 145
  implicit vs. explicit, 103, 145

var, 69, 111
variables, 16, 111
  abstract, see abstract variables
  authentic, 103
  auxiliary, 88
  concrete, see program variables
dependencies between, see dependencies
functions over, 84
global, 68, 69
  abstract, 84
local, 17, 69, 84
  abstract, 84
  scope rules for, 111
  program, see program variables
structure of, 66
visibility of, 69

variant function, see bound function
verification, 0
  and refinements, 94
  challenges of, 1
  feasibility of, 2, 94, 96
  full, 1, 146
verification conditions, 110
  generating, 110
verification process, 8
verification system
  building of, 95
virtual memory, 85
visibility requirement, 102, 104, 128,
  133, 142, 146
deviating from, 150
  enforcing, 104
  generalizing, 150
visible declaration (Anna lingo), 108
voltages, 85

weakest liberal preconditions, 12, 14,
  15, 60, 119
weakest preconditions, 1, 3, 12, 14, 43,
  45, 60
  in trace semantics, 43, 45
widening, 22
Wilson, Richard M., v
Wirth, Niklaus, 24

wp, see weakest liberal preconditions
  universal conjunctivity of, 60, 119

wp, see weakest preconditions
  positive conjunctivity of, 44

Wr, 98

Wr.T, 98

WrClass, 98

WrFriends, 98

writer example, 98

writers, 98
  file, 98, 99
  idea behind subtypes of, 99
  text, 98, 99

wrong, 60, 112

X, 131, 135–141