Software Analyzers

ACSL Version 1.14
Implementation in 20.0 (Calcium)

ACSL: ANSI/ISO C Specification Language

Version 1.14 – 20.0 (Calcium)

Patrick Baudin[1], Pascal Cuoq[1], Jean-Christophe Filliâtre[4,3], Claude Marché[3,4], Benjamin Monate[1], Yannick Moy[2,4,3], Virgile Prevosto[1]

[1] CEA LIST, Software Reliability Laboratory, Saclay, F-91191
[2] France Télécom, Lannion, F-22307
[3] INRIA Saclay - Île-de-France, ProVal, Orsay, F-91893
[4] LRI, Univ Paris-Sud, CNRS, Orsay, F-91405

Sources are available at `https://github.com/acsl-language/acsl`.

This work has been supported by the ANR project CAT (ANR-05-RNTL-0030x), by the ANR CIFRE contract 2005/973, by the ANR project U3CAT (08-SEGI-021-xx), and ANR PICF project DEVICE-Soft (2009-CARN-006-01).

Contents

1 Introduction **11**
1.1 Organization of this document . . . 12
1.2 Generalities about Annotations . . . 12
1.2.1 Kinds of annotations . . . 12
1.2.2 Parsing annotations in practice . . . 13
1.2.3 About preprocessing . . . 13
1.2.4 About keywords . . . 13
1.3 Notations for grammars . . . 13

2 Specification language **15**
2.1 Lexical rules . . . 15
2.2 Logic expressions . . . 16
2.2.1 Operators precedence . . . 19
2.2.2 Semantics . . . 20
2.2.3 Typing . . . 20
2.2.4 Integer arithmetic and machine integers . . . 21
2.2.5 Real numbers and floating point numbers . . . 23
2.2.6 C arrays and pointers . . . 26
2.2.7 Structures, Unions and Arrays in logic . . . 27
2.3 Function contracts . . . 29
2.3.1 Built-in constructs \old and \result . . . 30
2.3.2 Simple function contracts . . . 31
2.3.3 Semantics of frame conditions . . . 32
2.3.4 Contracts with named behaviors . . . 32
2.3.5 Memory locations and sets of terms . . . 35
2.3.6 Default contracts, multiple contracts . . . 37
2.4 Statement annotations . . . 38
2.4.1 Assertions . . . 38
2.4.2 Loop annotations . . . 39

2.4.3 Built-in construct \at . . . 44
2.4.4 Statement contracts . . . 46
2.5 Termination . . . 47
2.5.1 Integer measures . . . 47
2.5.2 General measures . . . 48
2.5.3 Recursive function calls . . . 49
2.5.4 Non-terminating functions . . . 50
2.6 Logic specifications . . . 50
2.6.1 Predicate and function definitions . . . 50
2.6.2 Lemmas . . . 51
2.6.3 Inductive predicates . . . 52
2.6.4 Axiomatic definitions . . . 53
2.6.5 Polymorphic logic types . . . 54
2.6.6 Recursive logic definitions . . . 54
2.6.7 Higher-order logic constructions . . . 54
2.6.8 Concrete logic types . . . 56
2.6.9 Hybrid functions and predicates . . . 56
2.6.10 Memory footprint specification: reads clause . . . 59
2.6.11 Specification Modules . . . 59
2.7 Pointers and physical addressing . . . 60
2.7.1 Memory blocks and pointer dereferencing . . . 60
2.7.2 Separation . . . 62
2.7.3 Dynamic allocation and deallocation . . . 62
2.8 Sets and lists . . . 67
2.8.1 Finite sets . . . 67
2.8.2 Finite lists . . . 68
2.9 Abrupt termination . . . 69
2.10 Dependencies information . . . 71
2.11 Data invariants . . . 72
2.11.1 Semantics . . . 73
2.11.2 Model variables and model fields . . . 75
2.12 Ghost variables and statements . . . 77
2.12.1 Volatile variables . . . 80
2.13 Initialization and undefined values . . . 82
2.14 Dangling pointers . . . 82
2.15 Well-typed pointers . . . 83
2.16 Logic attribute annotations . . . 84

3 Libraries **85**
3.1 Libraries of logic specifications . . . 85
3.1.1 Real numbers . . . 85
3.1.2 Finite lists . . . 85
3.1.3 Sets and Maps . . . 85
3.2 Jessie library: logical addressing of memory blocks . . . 86
3.2.1 Abstract level of pointer validity . . . 86
3.2.2 Strings . . . 86
3.3 Memory leaks . . . 87

4 Conclusion **89**

A Appendices **91**
A.1 Glossary . . . 91
A.2 Comparison with JML . . . 91
A.2.1 Low-level language vs. inheritance-based one . . . 92
A.2.2 Deductive verification vs. RAC . . . 96
A.2.3 Syntactic differences . . . 96
A.3 Typing rules . . . 97
A.3.1 Rules for terms . . . 97
A.3.2 Typing rules for sets . . . 98
A.4 Specification Templates . . . 98
A.4.1 Accessing a C variable that is masked . . . 98
A.5 Illustrative example . . . 99
A.6 Changes . . . 106
A.6.1 Version 1.14 . . . 106
A.6.2 Version 1.13 . . . 106
A.6.3 Version 1.12 . . . 106
A.6.4 Version 1.11 . . . 106
A.6.5 Version 1.10 . . . 106
A.6.6 Version 1.9 . . . 107
A.6.7 Version 1.8 . . . 107
A.6.8 Version 1.7 . . . 107
A.6.9 Version 1.6 . . . 107
A.6.10 Version 1.5 . . . 108
A.6.11 Version 1.4 . . . 108
A.6.12 Version 1.3 . . . 108
A.6.13 Version 1.2 . . . 108

Bibliography 109

List of Figures 113

Index 115

Foreword

This document describes version 1.14 of the ANSI/ISO C Specification Language (ACSL). The language features may still evolve in the future. In particular, some features in this document are considered *experimental*, meaning that their syntax and semantics is not yet fixed. These features are marked with EXPERIMENTAL. They must also be considered advanced features, which are not supposed to be needed for a basic use of this specification language.

Acknowledgements

We gratefully thank all the people who contributed to this document: Sylvie Boldo, David Cok, Jean-Louis Colaço, Pierre Crégut, David Delmas, Catherine Dubois, Stéphane Duprat, Arnaud Gotlieb, Philippe Herrmann, Thierry Hubert, André Maroneze, Dillon Pariente, Pierre Rousseau, Julien Signoles, Jean Souyris, Asma Tafat.

Chapter 1

Introduction

This document is a reference manual for the ACSL implementation provided by the Frama-C framework [13]. ACSL is an acronym for "ANSI/ISO C Specification Language". This is a Behavioral Interface Specification Language (BISL) [15] for specifying behavioral properties of C source code.

Not all of the features mentioned in this document are currently implemented in the Frama-C kernel. Unimplemented features are signaled as in the following line:

This feature is not currently supported by Frama-C[1]

As a summary, the features that are not currently implemented in Frama-C include in particular:

- some built-in predicates and logical functions;
- definition of logical types (section 2.6);
- specification modules (section 2.6.11);
- model variables (section 2.62);
- only basic support for ghost code is provided (section 2.12);
- verification of non interference of ghost code (p. 78);
- specification of volatile variables (section 2.12.1);

The main inspiration for this language comes from the specification language of the Caduceus tool [11, 12] for deductive verification of behavioral properties of C programs. The specification language of Caduceus is itself inspired from the *Java Modeling Language* (JML [21]), which aims at similar goals for Java source code: indeed it aims both at *runtime assertion checking* and *static verification* using the OpenJML tool [6, 7], where we aim at *static verification* and *deductive verification* (see Appendix A.2 for a detailed comparison between ACSL and JML).

Going back further in history, the JML design was guided by the general *design-by-contract* principle proposed by Bertrand Meyer, originally implemented in the Eiffel language; he

[1] Additional remarks on the feature may appear in a footnote.

took his inspiration from the concepts of preconditions and postconditions on a routine, going back at least to Dijkstra, Floyd and Hoare in the late 60's and early 70's.

In this document, we assume that the reader has a good knowledge of the ISO C programming language [17, 16].

1.1 Organization of this document

In this preliminary chapter we introduce some definitions and vocabulary, and discuss generalities about this specification language. Chapter 2 presents the specification language itself. Chapter 3 presents additional information about *libraries* of specifications. The appendices provide formal rules for type-checking ACSL annotations, the relation between ACSL and JML, and specification templates. A detailed table of contents is given on page 5. A glossary is given in Appendix A.1.

1.2 Generalities about Annotations

In this document, we consider that specifications are given as annotations in comments written directly in C source files, so that source files remain compilable. Those comments must start by `/*@` or `//@` and end as usual in C.

In some contexts, it is not possible to modify the source code. It is strongly recommended that a tool that implements ACSL specifications provide technical means to store annotations separately from the source. It is not the purpose of this document to describe such means. Nevertheless, some of the specifications, namely those at a global level, can be given in separate files: logical specifications can be imported (see Section 2.6.11) and a function contract can be attached to a copy of the function profile (see Section 2.3.6).

1.2.1 Kinds of annotations

- Global annotations:
  - *function contract* : such an annotation is inserted just before the declaration or the definition of a function. See section 2.3.
  - *global invariant* : this is allowed at the level of global declarations. See section 2.11.
  - *type invariant* : this allows declaring both structure or union invariants, and invariants on type names introduced by `typedef`. See section 2.11.
  - *logic specifications* : definitions of logic functions or predicates, lemmas, axiomatizations by declaration of new logic types, logic functions, predicates with axioms they satisfy. Such an annotation is placed at the level of global declarations. See section 2.6
- Statement annotations:
  - *assertion* : these are allowed everywhere a C label is allowed, or just before a block closing brace. See section 2.4.1.
  - *loop annotation* (invariant, variant, assign clauses): is allowed immediately before a loop statement: `for` , `while` , `do` ... `while` . See Section 2.4.2.

– *statement contract*: very similar to a function contract, and placed before a statement or a block. Semantic conditions must be checked (e.g., no goto going inside, no goto going outside). See Section 2.4.4.
– *ghost code*: regular C code, only visible from the specifications, that is only allowed to modify ghost variables. See section 2.12. This includes ghost braces for enclosing blocks.

1.2.2 Parsing annotations in practice

In the original (University of Iowa) JML tools, parsing was done by simply ignoring `//@`, `/*@` and `*/` at the lexical analysis level. This technique could modify the semantics of the code, for example:

```
return x /*@ +1 */ ;
```

In our language (as in the definition of JML and current JML tools, such as OpenJML), this is forbidden. Technically, the current implementation of Frama-C isolates the comments in a first step of syntax analysis, and then parses a second time. Nevertheless, the grammar and the corresponding parser must be carefully designed to avoid interaction of annotations with the code. For example, in code such as

```
  if (c) //@ assert P;
    c=1;
```

the statement `c=1` must be understood as the branch of the `if`. This is ensured by the ACSL grammar, which states that `assert` annotations are not statements themselves, but attached to the statement that follows, like C labels.

1.2.3 About preprocessing

This document considers C source *after* preprocessing. Tools must decide how they handle preprocessing (what to do with annotations, whether macro substitution should be performed, etc.)

Preprocessing includes interpreting C digraphs and trigraphs. As these are generally deprecated and en route to removal from the C standard, ACSL does not define uses of digraphs and trigraphs. Any tool that wishes to support such alternate syntax can preprocess the tokens into conventional tokens before passing the text to ACSL tools.

1.2.4 About keywords

Additional keywords of the specification language start with a backslash, if they are used in the position of a term or a predicate (which are defined in the following). Otherwise they do not start with a backslash (like `ensures`) and they remain valid identifiers.

1.3 Notations for grammars

In this document, grammar rules are given in BNF form. In the grammar rules, we use the extra notations e^* to denote repetition of zero, one or more occurrences of e, e^+ for repetition of one or more occurrences of e, and $e^?$ for zero or one occurrence of e. For the sake of simplicity, we only describe annotations in the usual `/*@ ... */` style of comments. One-line annotations in `//@` comments are similar. Note however that two consecutive comments, regardless of their style, are considered as two independent annotations. In particular, it is not possible in general to split a multi-line annotation into several `//@` comments.

Chapter 2

Specification language

2.1 Lexical rules

Specification language text is placed inside special C comments; its lexical structure mostly follows that of ANSI/ISO C. A few differences should be noted.

- The at sign (`@`) is equivalent to a space character, except where it indicates the beginning of an ACSL annotation.
- Identifiers may start with the backslash character (\).
- Some UTF8 characters may be used in place of some constructs, as shown in the following table:

`>=`	$\geq$	0x2265
`<=`	$\leq$	0x2264
`>`	$>$	0x003E
`<`	$<$	0x003C
`\in`	$\in$	0x2208
`!=`	$\not\equiv$	0x2262
`==`	$\equiv$	0x2261
`==>`	$\Longrightarrow$	0x21D2
`<==>`	$\Longleftrightarrow$	0x21D4
`&&`	$\wedge$	0x2227
`||`	$\vee$	0x2228
`^^` (xor)	$\veebar$	0x22BB
`!`	$\neg$	0x00AC
`-` (unary minus)	$-$	0x2212
`\forall`	$\forall$	0x2200
`\exists`	$\exists$	0x2203
`integer`	$\mathbb{Z}$	0x2124
`real`	$\mathbb{R}$	0x211D
`boolean`	$\mathbb{B}$	0x1D539
`\pi`	π	0x3C0

- Comments can be put inside ACSL annotations. They use the C++ format, *i.e.* begin with `//` and extend to the end of current line.

2.2 Logic expressions

This first section presents the language of expressions one can use in annotations. These are called *logic expressions* in the following. They correspond to pure C expressions, with additional constructs that we will introduce progressively.

Figures 2.1 and 2.2 present the grammar for the basic constructs of logic expressions. In that grammar, we distinguish between *predicates* and *terms*, following the usual distinction between propositions and terms in classical first-order logic. The grammar for binders and type expressions is given separately in Figure 2.3.

With respect to C pure expressions, the additional constructs are as follows:

Additional connectives C operators `&&` (UTF8: $\wedge$), `||` (UTF8: $\vee$) and `!` (UTF8: $\neg$) are used as logical connectives. There are additional connectives `==>` (UTF8: $\Longrightarrow$) for implication, `<==>` (UTF8: $\Longleftrightarrow$) for equivalence and `^^` (UTF8: $\underline{\vee}$) for exclusive or. These logical connectives all have a bitwise counterpart, either C ones like `&`, `|`, `~` and `^`, or additional ones like bitwise implication `-->` and bitwise equivalence `<-->`.

Quantification Universal quantification is denoted by \forall τ x_1,...,x_n; e and existential quantification by \exists τ x_1,...,x_n; e.

Local binding \let x = e_1;e_2 introduces the name x for expression e_1; x can then be used in expression e_2.

Conditional c ? e_1 : e_2. There is a subtlety here: the condition may be either a boolean term or a predicate. In case of a predicate, the two branches must be also predicates, so that this construct acts as a connective with the following semantics: c ? e_1 : e_2 is equivalent to (c ==> e_1) && (! c ==> e_2).

Syntactic naming `id : e` is a term or a predicate equivalent to e. It is different from local naming with \let: the name cannot be reused in other terms or predicates. It is only for readability purposes.

Functional modifier The composite element modifier is an additional operator related to C structure field and array accessors. The expression `{ s \with .id = v }` denotes a structure value that is the same as `s`, except for the field `id`, which is equal to `v`. The equivalent expression for an array is `{ t \with [ i ] = v }`, which returns an array with the same value as `t`, except for the `i`th element whose value is `v`. See section 2.10 for an example use of these operators.

Logic functions Applications in terms and in propositions are not applications of C functions, but of logic functions or predicates; see Section 2.6 for detail.

Consecutive comparison operators The construct t_1 $relop_1$ t_2 $relop_2$ t_3 $\cdots$ t_k with several consecutive comparison operators is a shortcut for (t_1 $relop_1$ t_2) && (t_2 $relop_2$ t_3) && $\cdots$. It is required that the $relop_i$ operators must be in the same "direction", *i.e.* they must all belong either to {`<`, `<=`, `==`} or to

```
literal   ::= \true | \false                      boolean constants
            | integer                             (lexical) integer constants
            | real                                (lexical) real constants
            | string                              (lexical) string constants
            | character                           (lexical) character constants

bin-op    ::= + | - | * | / | % | << | >>
            | == | != | <= | >= | > | <
            | && | || | ^^                        boolean operations
            | & | | | --> | <--> | ^              bitwise operations

unary-op  ::= + | -                               unary plus and minus
            | !                                   boolean negation
            | ~                                   bitwise complementation
            | *                                   pointer dereferencing
            | &                                   address-of operator

term      ::= literal                             literal constants
            | poly-id                             variables, function names
            | unary-op term
            | term bin-op term
            | term [ term ]                       array access
            | { term \with [ term ] = term }      array functional modifier
            | term . id                           structure field access
            | { term \with . id = term }          field functional modifier
            | term -> id
            | ( type-expr ) term                  cast
            | poly-id ( term (, term)* )          function application
            | ( term )                            parentheses
            | term ? term : term                  ternary condition
            | \let id = term ; term               local binding
            | sizeof ( term )
            | sizeof ( C-type-name )
            | id : term                           syntactic naming
            | string : term                       syntactic naming

poly-id   ::= ident

ident     ::= id                                  lexical identifier token
```

Figure 2.1: Grammar of terms. The terminals *id*, *C-type-name*, and various literals are the same as the corresponding C lexical tokens.

{>,>=,==}. Expressions such as `x < y > z` or `x != y != z` are not allowed. Note that consecutive comparison operators are allowed only in predicate position.

A consecutive comparison as the conditional expression in a ternary operation could, according to the grammar, be either in term or predicate position. In such a case, the conditional expression is considered a predicate. As a term a consecutive comparison `x < y < z` would be parsed as `(x < y) < z` which is incorrectly typed, because in logic expressions, comparisons result in boolean values.

To enforce the same interpretation as in C expressions, one may need to add extra parentheses: `a == b < c` is equivalent to `a == b && b < c`, whereas `a == (b < c)` is equivalent to `\let x = b < c; a == x`. This situation raises some issues, as in the example below.

There is a subtlety regarding comparison operators: they are predicates when used in predicate position, and boolean functions when used in term position.

Example 2.1 *Let us consider the following example:*

```
int f(int a, int b) { return a < b; }
```

- *the obvious postcondition* `\result == a < b` *is not the right one because it is actually a shortcut for* `\result == a && a < b`.
- *adding parentheses results in a correct post-condition* `\result == (a < b)`. *Note however that there is an implicit conversion (see Sec. 2.2.3) from the* `int` *(the type of* `\result`*) to* `boolean` *(the type of* `(a<b)`*)*
- *an equivalent post-condition, which does not rely on implicit conversion, is* `(\result != 0) == (a<b)`. *Both pairs of parentheses are mandatory.*
- `\result == (integer)(a<b)` *is also acceptable because it compares two integers. The cast towards* `integer` *enforces* `a<b` *to be understood as a boolean term. Notice that a cast towards* `int` *would also be acceptable.*
- `\result != 0 <==> a < b` *is acceptable because it is an equivalence between two predicates.*

rel-op	::=	`==` | `!=` | `<=` | `>=` | `>` | `<`	
pred	::=	`\true` | `\false`	
	|	*term* (*rel-op* *term*)$^+$	comparisons (see remark)
	|	*ident* `(` *term* (`,` *term*)* `)`	predicate application
	|	`(` *pred* `)`	parentheses
	|	*pred* `&&` *pred*	conjunction
	|	*pred* `||` *pred*	disjunction
	|	*pred* `==>` *pred*	implication
	|	*pred* `<==>` *pred*	equivalence
	|	`!` *pred*	negation
	|	*pred* `^^` *pred*	exclusive or
	|	*term* `?` *pred* `:` *pred*	ternary condition
	|	*pred* `?` *pred* `:` *pred*	
	|	`\let` *id* `=` *term* `;` *pred*	local binding
	|	`\let` *id* `=` *pred* `;` *pred*	
	|	`\forall` *binders* `;` *pred*	universal quantification
	|	`\exists` *binders* `;` *pred*	existential quantification
	|	*id* `:` *pred*	syntactic naming
	|	*string* `:` *pred*	syntactic naming

Figure 2.2: Grammar of predicates

binders	::=	*binder* (, *binder*)*	
binder	::=	*type-expr* *variable-ident* (,*variable-ident*)*	
type-expr	::=	*logic-type-expr* | *C-type-name*	
logic-type-expr	::=	*built-in-logic-type*	
	|	*id*	type identifier
built-in-logic-type	::=	boolean | integer | real	
variable-ident	::=	*id*	
	|	`*` *variable-ident*	
	|	*variable-ident* `[]`	
	|	`(` *variable-ident* `)`	

Figure 2.3: Grammar of binders and type expressions

class	associativity	operators
selection	left	[···] -> .
unary	right	! ~ + - * & (cast) sizeof
multiplicative	left	* / %
additive	left	+ -
shift	left	<< >>
comparison	-	< <= > >=
comparison	-	== !=
bitwise and	left	&
list repetition	left	*^
bitwise xor/list concatenation	left	^
bitwise or	left	|
bitwise implies	right	-->
bitwise equiv	left	<-->
connective and	left	&&
connective xor	left	^^
connective or	left	||
connective implies	right	==>
connective equiv	left	<==>
ternary connective	right	···?···:···
binding	left	\forall \exists \let
naming	right	:

Figure 2.4: Operator precedence

2.2.1 Operators precedence

The precedence of C operators is conservatively extended with additional operators, as shown in Figure 2.4. In this table, operators are sorted from highest to lowest priority. Operators of same priority are presented on the same line.

Conditional expressions and labels There is a remaining ambiguity between the connective $\cdots$`?`$\cdots$`:`$\cdots$ and the labelling operator `:`. Consider for instance the expression `x?y:z:t`. The precedence table does not indicate whether this should be understood as `x?(y:z):t` or `x?y:(z:t)`. Such a case must be considered as a syntax error, and should be fixed by explicitly adding parentheses.

Labels and parsing Note also that the use of labels can subtlely change the parsing of an expression, because labeled expressions have the least binding precedence. That is, once a label is seen, the parser finds the longest valid term or predicate following the label to consider as the labeled expression. For example, `a && b ==> c && d` parses as `(a && b) ==> (c && d)`, but `a && nm: b ==> c && d` parses as `a && (nm: ( b ==> (c && d)))`.

2.2.2 Semantics

The semantics of logic expressions in ACSL is based on mathematical first-order logic [27]. In particular, it is a 2-valued logic with only total functions. Consequently, expressions are never "undefined". This is an important design choice and the specification writer should be aware of that. (For a discussion about the issues raised by such design choices, in similar specification languages such as JML, see the comprehensive list compiled by Patrice Chalin [4, 5].)

Having only total functions implies than one can write terms such as `1/0`, or `*p` when `p` is null (or more generally when it points to a non-properly allocated memory cell). In particular, the predicates `1/0 == 1/0` `*p == *p` are valid, since they are instances of the axiom $\forall x, x = x$ of first-order logic. The reader should not be alarmed, because there is no way to deduce anything useful from such terms. As usual, it is up to the specification designer to write consistent assertions. For example, when introducing the following lemma (see Section 2.6):

```
/*@ lemma div_mul_identity:
  @   \forall real x, real y; y != 0.0 ==> y*(x/y) == x;
  @*/
```

a premise is added to require `y` to be non zero.

2.2.3 Typing

The language of logic expressions is typed (as in *multi-sorted* first-order logic). Types are either C types or *logic types* defined as follows:

- "mathematical" types: `integer` for unbounded, mathematical integers, `real` for real numbers, `boolean` for booleans (with values written `\true` and `\false`);
- logic types introduced by the specification writer (see Section 2.6).

There are implicit coercions for numeric types:

- C integral types `char`, `short`, `int` and `long`, signed or unsigned, are all subtypes of type `integer`;
- `integer` is itself a subtype of type `real`;
- C types `float` and `double` are subtypes of type `real`.

Notes:

- There is a distinction between booleans and predicates. The expression `x<y` in term position is a boolean, and the same expression is also allowed in predicate position.
- Unlike in C, there is a distinction between booleans and integers. There is an implicit promotion from integers to booleans, thus one may write `x && y` instead of `x != 0 && y != 0`. If the reverse conversion is needed, an explicit cast is required, *e.g.* `(int)(x>0)+1`, where `\false` becomes 0 and `\true` becomes 1.
- Quantification can be made over any type: logic types and C types. Quantification over pointers must be used carefully, since it depends on the memory state where dereferencing is done (see Section 2.2.4 and Section 2.6.9).

Formal typing rules for terms are given in appendix A.3.

2.2.4 Integer arithmetic and machine integers

The following integer arithmetic operations apply to *mathematical integers*: addition, subtraction, multiplication, unary minus. The value of a C variable of an integral type is promoted to a mathematical integer. As a consequence, there is no "arithmetic overflow" in logic expressions.

Division and modulo are also mathematical operations, which coincide with the corresponding C operations on C machine integers, thus following the ISO C99 conventions. In particular, these are not the usual mathematical Euclidean division and remainder. Generally speaking, division rounds the result towards zero. The results are not specified if the divisor is zero; otherwise if q and r are the quotient and the remainder of n divided by d then:

- $|d \times q| \leq |n|$, and $|q|$ is maximal for this property;
- q is zero if $|n| < |d|$;
- q is positive if $|n| \geq |d|$ and n and d have the same sign;
- q is negative if $|n| \geq |d|$ and n and d have opposite signs;
- $q \times d + r = n$;
- $|r| < |d|$;
- r is zero or has the same sign as n.

Example 2.2 *The following examples illustrate the results of division and modulo depending on the sign of their arguments:*

- `5/3` *is 1 and* `5%3` *is 2;*
- `(-5)/3` *is -1 and* `(-5)%3` *is -2;*
- `5/(-3)` *is -1 and* `5%(-3)` *is 2;*
- `(-5)/(-3)` *is 1 and* `(-5)%(-3)` *is -2.*

Hexadecimal octal and binary constants

Hexadecimal, octal and binary constants are always non-negative. Suffixes `u` and `l` for C constants are allowed but meaningless.

Casts and overflows

In logic expressions, casting from mathematical integers to an integral C type `t` (such as `char`, `short`, `int`, etc.) is allowed and is interpreted as follows: the result is the unique value of the corresponding type that is congruent to the mathematical result modulo the cardinal of this type, that is $2^{8\times \mathrm{sizeof}(t)}$.

Example 2.3 `(unsigned char)1000` *is* $1000 \bmod 256$*, i.e.,* 232*; however,* `(signed char)1000` *is* $((1000+128) \bmod 256)-128$*, i.e.,* -24*.*

To express in the logic the value of a C expression, one has to add all the necessary casts. For example, the logic expression denoting the value of the C expression `x*y+z` is `(int)((int)(x*y)+z)`. Note that there is no implicit cast from integers to C integral types.

Example 2.4 *The declaration*

```
//@ logic int f(int x) = x+1 ;
```

is not allowed because `x+1`*, which is a mathematical integer, must be cast to* `int`*. One should write either*

```
//@ logic integer f(int x) = x+1 ;
```

or

```
//@ logic int f(int x) = (int)(x+1) ;
```

Quantification on C integral types

Quantification over a C integral type corresponds to integer quantification over the corresponding interval.

Example 2.5 *Thus the formula*

```
\forall char c; c <= 1000
```

is equivalent to

```
\forall integer c; CHAR_MIN <= c <= CHAR_MAX ==> c <= 1000
```

where the bounds `CHAR_MIN` *and* `CHAR_MAX` *are defined in* `limits.h`

Size of C integer types

The size of C types is architecture-dependent. ACSL does not enforce these sizes either, hence the semantics of terms involving such types is also architecture-dependent. The `sizeof` operator may be used in annotations and is consistent with its C counterpart (including that its return type is a value of type `size_t`, and in most cases a constant). For instance, it should be possible to verify the following code:

```
/*@ ensures \result <= sizeof(int); */
int f() { return sizeof(char); }
```

Constants giving maximum and minimum values of those types may be provided in a library.

Enum types

Enum types are also interpreted as mathematical integers. Casting an integer into an enum in the logic gives the same result as if the cast was performed in the C code.

Bitwise operations

Like arithmetic operations, bitwise operations apply to any mathematical integer: any mathematical integer has a unique infinite 2-complement binary representation with infinitely many zeros (for non-negative numbers) or ones (for negative numbers) on the left. Bitwise operations apply to this representation.

Example 2.6

- `7 & 12 == ···00111 & ···001100 == ···00100 == 4`
- `-8 | 5 == ···11000 | ···00101 == ···11101 == -3`
- `~5 == ~···00101 == ···111010 == -6`
- `-5 << 2 == ···11011 << 2 == ···11101100 == -20`
- `5 >> 2 == ···00101 >> 2 == ···0001 == 1`
- `-5 >> 2 == ···11011 >> 2 == ···1110 == -2`

2.2.5 Real numbers and floating point numbers

Floating-point constants and operations are interpreted as mathematical real numbers: a C variable of type float or double is implicitly promoted to a real. Integers are promoted to reals if necessary. The usual binary operations are interpreted as operators on real numbers, hence they never involve any rounding or overflow.

Example 2.7 *In an annotation,* `1e+300 * 1e+300` *is equal to* `1e+600`*, even if that last number exceeds the largest representable number in double precision: there is no "overflow".*

$2 * 0.1$ *is equal to the real number* 0.2*, and not to any floating-point approximation: there is no "rounding".*

Unlike the promotion of C integer types to mathematical integers, there are special float values that do not naturally map to a real number, namely the IEEE-754 special values for "not-a-number", $+\infty$ and $-\infty$. See below for a detailed discussion on such special values. However, remember that ACSL's logic has only total functions. Thus, there are implicit promotion functions `real_of_float` and `real_of_double` whose results on the 3 values above is left unspecified.

In logic, real literals can also be expressed under the hexadecimal form of C99: `0x`hh`.`hh`p`$\pm dd$ where h are hexadecimal digits and dd is in decimal, denotes number $hh.hh \times 2^{dd}$, e.g. `0x1.Fp-4` is $(1 + 15/16) \times 2^{-4}$.

The usual operators for comparison are also interpreted as real operators. In particular, the equality operation $\equiv$ for float (or double) expressions means equality of the real numbers they represent. Or equivalently, $x \equiv y$ for two float variables x, y means `real_of_float(x)` $\equiv$ `real_of_float(y)` with the mathematical equality of real numbers.

Special predicates are also available to express the comparison operators of float (resp. double) numbers as in C: `\eq_float`, `\gt_float`, `\ge_float`, `\le_float`, `\lt_float`, `\ne_float` (resp. for double).

Casts, infinity and NaNs

Casting from a C integer type or a float type to a float or a double is as in C: the same conversion operations apply.

Conversion of real numbers to float or double values indeed depends on various possible rounding modes defined by the IEEE 754 standard [26, 28]. These modes are defined by a logic type (see section 2.6.8):

```
/*@ type rounding_mode = \Up | \Down | \ToZero | \NearestAway | \NearestEven;
*/
```

Then rounding a real number can be done explicitly using functions

```
  logic float \round_float(rounding_mode m, real x);
  logic double \round_double(rounding_mode m, real x);
```

Cast operators `(float)` and `(double)` applied to a mathematical integer or real number x are equivalent to applying the rounding functions above with the nearest-even rounding mode (which is the default rounding mode in C programs). If the source real number is too large, this may also result in one of the special values +infinity and -infinity.

Example 2.8 *We have* `(float)0.1` $\equiv$ 13421773×2^{-27} *which is equal to* 0.100000001490116119384765625

Notice also that unlike for integers, suffixes `f` and `l` are meaningful, because they implicitly add a cast operator as above.

This semantics of casts ensures that the float result `r` of a C operation e_1 `op` e_2 on floats, if there is no overflow and if the default rounding mode is not changed in the program, has the same real value as the logic expression `(float)(`e_1 `op` e_2`)`. Notice that this is not true for the equality `\eq_float` of floats: `-0.0 + -0.0` in C is equal to the float number `-0.0`, which is not `\eq_float` to `0.0`, which is the value of the logic expression `(float)(-0.0 + -0.0)`.

Finally, additional predicates are provided that check that their argument is a finite number, an infinite one, or a NaN:

```
predicate \is_finite (double x); // is a finite double
predicate \is_plus_infinity (double x); // is equal to +infinity
predicate \is_minus_infinity (double x); // is equal to -infinity
predicate \is_infinite (double x); // is equal to +infinity or -infinity
predicate \is_NaN(double x); // is a NaN double
```

`\is_finite`, `\is_plus_infinity`, `\is_minus_infinity` and `\is_NaN` are mutually exclusive predicates. All these predicates also exist for the `float` type.

Recall that under IEEE754 rules, any comparison between two NaN values returns a false value. Consequently if a `double` variable `d` is a NaN value, then the C expression `d == d` and the logic expression `\eq_double(d,d)` will both be false, the logic expression `\real_of_double(d)` will be undefined, but the logic expression `\real_of_double(d) == \real_of_double(d)` is true, as a specific instance of the axiom `x` $\equiv$ `x`.

Sign

The sign of a non-NaN floating-point can be extracted by the function \sign:

```
/*@
 type sign = \Positive | \Negative;

  logic sign \sign(float x);
  logic sign \sign(double x);
*/
```

Quantification

Quantification over a variable of type real is of course the usual quantification over real numbers.

Quantification over float (resp. double) types is allowed too, and is supposed to range over all real numbers representable as floats (resp. doubles). In particular, this does not include NaN, +infinity and -infinity in the considered range.

Mathematical functions

Classical mathematical operations like exponential, sine, cosine, and such are built-in to ACSL:

```
integer \min(integer x, integer y) ;
integer \max(integer x, integer y) ;
real \min(real x, real y) ;
real \max(real x, real y) ;
integer \abs(integer x) ;
real \abs(real x) ;

real \sqrt(real x) ;
real \pow(real x, real y) ;

integer \ceil(real x) ;
integer \floor(real x) ;

real \e ;
real \exp(real x) ;
real \log(real x) ;
real \log10(real x) ;

real \pi ;
real \cos(real x) ;
real \sin(real x) ;
real \tan(real x) ;

real \cosh(real x) ;
real \sinh(real x) ;
real \tanh(real x) ;

real \acos(real x) ;
real \asin(real x) ;
real \atan(real x) ;

real \atan2(real y, real x) ;
```

```
real \hypot(real x, real y) ;
```

Notation `\pi` refers to the real number π and `\e` to the base of the natural logarithm: `\log(\e)==1` and `\exp(1)==\e`.

Exact computations

In order to specify properties of rounding errors, it is useful to express something about the so-called *exact* computations [3]: the computations that would be performed in an ideal mode where variables denote true real numbers.

To express such exact computations, two special constructs exist in annotations:

- `\exact`(x) denotes the value of the C variable x (or more generally any C left-value) as if the program were executed with ideal real numbers.
- `\round_error`(x) is a shortcut for $|x -$ `\exact`$(x)|$

Example 2.9 *Here is an example of a naive approximation of cosine [2].*

```
/*@ requires \abs(\exact(x)) <= 0x1p-5;
  @ requires \round_error(x) <= 0x1p-20;
  @ ensures \abs(\exact(\result) - \cos(\exact(x))) <= 0x1p-24;
  @ ensures \round_error(\result) <= \round_error(x) + 0x3p-24;
  @*/
float cosine(float x) {
  return 1.0f - x * x * 0.5f;
}
```

2.2.6 C arrays and pointers

Address operator, array access, pointer arithmetic and dereferencing

These operators are similar to their corresponding C operators.

address-of operator should be used with caution. Values in logic do not lie in C memory so it does not mean anything to talk about their "address".

Unlike in C, there is no implicit cast from an array type to a pointer type. Nevertheless, arithmetic and dereferencing over arrays lying in C memory are allowed like in C.

Example 2.10 *Dereferencing a C array is equivalent to an access to the first element of the array ; shifting it from i denotes the address of its i^{th} element.*

```
int tab[10] = { 1 } ;
int x ;
int *p = &x;

//@ requires p == &x
int main(void){
  //@ assert tab[0]==1 && *p == x;
  //@ assert *tab == 1;
  int *q = &tab[3];
  //@ assert q+1 == tab+4;
  ...
}
```

Since pointers can only refer to values lying in C memory, `p->s` is always equivalent to `(*p).s`. On the contrary, `t[i]` is not always equivalent to `*(t+i)`, especially for arrays not lying in C memory. Section 2.2.7 details the use of arrays as logic values. There are also differences between `t` and the pointer to its first element when evaluating an expression at a given program point. See Section 2.4.3 for more information.

Function pointers

Pointers to C functions are allowed in logic. The only possible use of them is to check for equality.

Example 2.11

```
int f(int x);
int g(int x);

//@ requires p == &f || p == &g;
void h(int(*p)(int)) {
...
}
```

2.2.7 Structures, Unions and Arrays in logic

Aggregate C objects (i.e. structures, unions and arrays) are also possible values for terms in logic. They can be passed as parameters to and returned from logic functions, tested for equality, etc. like any other values.

Aggregate types can be declared in logic, and their contents may be any logic types themselves. Constructing such values in logic can be performed using a syntax similar to C designated initializers.

Example 2.12 *Array types in logic may be declared either with or without an explicit non-negative length. Access to the length of a logic array can be done with* \length*.*

```
//@ type point = struct { real x; real y; };
//@ type triangle = point[3];

//@ logic point origin = { .x = 0.0 , .y = 0.0 };
/*@ logic triangle t_iso = { [0] = origin,
  @                          [1] = { .y = 2.0 , .x = 0.0 }
  @                          [2] = { .x = 2.0 , .y = 0.0 }};
  @*/

/*@ logic point centroid(triangle t) = {
  @    .x = mean3(t[0].x,t[1].x,t[2].x);
  @    .y = mean3(t[0].y,t[1].y,t[2].y);
  @ };
  @*/

//@ type polygon = point[];
/*@ logic perimeter(polygon p) =
  @   \sum(0,\length(p)-1,\lambda integer i;d(p[i],p[(i+1) % \length(p)])) ;
  @*/
```

Beware that because of the principle of only total functions in logic, t[i] can appear in ACSL annotations even if i is outside the array bounds.

Functional updates

Syntax for functional update is similar to initialization of aggregate objects.

Example 2.13 *Functional update of an array is done by*

```
{ t_iso \with [0] = { .x = 3.0, .y = 3.0 } }
```

Functional update of a structure is done by

```
{ origin \with .x = 3.0 }
```

There is no particular syntax for functional update of a union. For an `object` *of a union type, the following equality is not true*

```
{ { object \with .x = 3.0 }
          \with .y = 2.0 } == { { object \with .y = 2.0 }
                                       \with .x = 3.0 }
```

The equality predicate == applies to aggregate values, but it is required that they have the same type. Then equality amounts to recursively checking equality of fields. Equality of arrays of different lengths returns false. Beware that equality of unions is also equality of all fields.

C aggregate types

C aggregate types (struct, union or array) naturally map to logic types, by recursively mapping their fields.

Example 2.14 *There is no implicit cast to type of the updated/initialized fields.*

```
struct S { int x; float y; int t[10]; };

//@ logic integer f(struct S s) = s.t[3];
//@ logic struct S g(integer n, struct S s) = { s \with .x = (int)n };
```

Unlike in C, all fields should be initialized:

```
/*@ logic struct S h(integer n, int a[10]) = {
  @   .x = (int)n, .y = (float)0.0, .t = a
  @   };
  @*/
```

Cast and conversion

Unlike in C, there is no implicit conversion from an array type to a pointer type. On the other hand, there is an implicit conversion from an array of a given size to an array with unspecified size (but not the converse).

Example 2.15

```
//@ logic point square[4] = { origin, ... };

//@ ... perimeter(square);       // well-typed
//@ ... centroid(square);        // wrongly typed
//@ ... centroid((triangle)square); // well-typed (truncation)
```

An explicit cast from an array type to a pointer type is allowed only for arrays that lie in C memory. As in C, the result of the cast is the address of the first element of the array (see Section 2.2.6).

Conversely, an explicit cast from a pointer type to an array type acts as collecting the values it points to.

Subtyping and cast recursively apply to fields.

Example 2.16

```
struct { float u,v; } p[10];

//@ assert centroid((point[3])p) == ...

//@ assert perimeter((point[])p) == ...
```

Precisely, conversion of a pointer p of type $\tau*$ to a logic array of type $\tau[]$ returns a logic array t such that

$$\mathrm{length}(t) = (\backslash\mathrm{block\_length}(p) - \backslash\mathrm{offset}\,(p))/\,\mathrm{sizeof}\,(\tau)$$

More generally, an explicit cast from a C aggregate of type τ to another C aggregate type is allowed in order to specify such a value conversion into logical functions or function contracts without using the addressing operator `&`.

Example 2.17 *Unlike in C, conversion of an aggregate of C type* `struct` τ *to another structure type is allowed.*

```
struct long_st { int x1,y2;};
struct st { char x,y; };

//@ ensures \result == (struct st) s;
struct st from_long_st(struct long_st s) {
  return *(struct st *)&s;
}
```

2.3 Function contracts

Figure 2.5 shows a grammar for function contracts. *Location* denotes a memory location and is defined in Section 2.3.5. *Allocation-clauses* allow specifying which memory locations are dynamically allocated or deallocated by the function from the *heap*; they are defined later in Section 2.7.3.

This section is organized as follows. First, the grammar for terms is extended with two new constructs. Then Section 2.3.2 introduces *simple contracts*. Finally, Section 2.3.4 defines more general contracts involving *named behaviors*.

The `decreases` and `terminates` clauses are presented later in Section 2.5. *Abrupt-clauses* allow specifying what happens when the function does not return normally but `exits` abruptly; they are defined in Section 2.9.

function-contract [a]	::=	*requires-clause** *terminates-clause*? *decreases-clause*? *simple-clause** *named-behavior** *completeness-clause**
requires-clause	::=	`requires` *pred* `;`
terminates-clause	::=	`terminates` *pred* `;`
decreases-clause	::=	`decreases` *term* (`for` *id*)? `;`
simple-clause	::=	*assigns-clause* | *ensures-clause*
	|	*allocation-clause* | *abrupt-clause*
assigns-clause	::=	`assigns` *locations* `;`
locations	::=	*location* (`,` *location*) * | `\nothing`
location	::=	*tset*
ensures-clause	::=	`ensures` *pred* `;`
named-behavior	::=	`behavior` *id* `:` *behavior-body*
behavior-body	::=	*assumes-clause** *requires-clause** *simple-clause**
assumes-clause	::=	`assumes` *pred* `;`
completeness-clause	::=	`complete` `behaviors` (*id* (`,` *id*)*)? `;`
	|	`disjoint` `behaviors` (*id* (`,` *id*)*)? `;`

[a] empty contracts are forbidden

Figure 2.5: Grammar of function contracts

term	::=	`\old` (*term*)	old value
	|	`\result`	result of a function
pred	::=	`\old` (*pred*)	

Figure 2.6: `\old` and `\result` in terms

2.3.1 Built-in constructs `\old` and `\result`

Post-conditions usually require referring to both the function result and values in the pre-state. Thus terms are extended with the following new constructs (shown in Figure 2.6).

- `\old(e)` denotes the value of predicate or term `e` in the pre-state of the function.
- `\result` denotes the returned value of the function.

`\old(e)` can be used only in `ensures`, `assigns`, `allocates` and `frees` clauses, since the other clauses already refer to only one state, the pre-state. In addition, `\result` can not be used in the contract of a function that returns `void`.

C function parameters are obtained by value from actual parameters that mostly remain unaltered by the function calls. For that reason, formal parameters in function contracts are defined such that they always refer implicitly to their values interpreted in the pre-state. Thus, the `\old` construct is not needed (but permitted) for formal parameters (in function contracts only).

2.3.2 Simple function contracts

A simple function contract, having only simple clauses and no named behaviors, takes the following form:

```
/*@ requires P₁; requires P₂; ...
  @ assigns L₁; assigns L₂; ...
  @ ensures E₁; ensures E₂; ...
  @*/
```

The semantics of such a contract is as follows:

- The caller of the function must guarantee that it is called in a state where the property P_1 && P_2 && ... holds.
- The called function returns[1] a state where the property E_1 && E_2 && ... holds.
- All memory locations that are allocated in both the pre-state and the post-state[2] and do not belong to the set $L_1 \cup L_2 \cup \ldots$ are left unchanged in the post-state. The set $L_1 \cup L_2 \cup \ldots$ itself is interpreted in the pre-state.

Having multiple requires, assigns, or ensures clauses only improves readability since the contract above is equivalent to the following simplified one:

```
/*@ requires P₁ && P₂ && ...;
  @ assigns L₁, L₂,...;
  @ ensures E₁ && E₂ && ...;
  @*/
```

If no requires clause is given, it defaults to \true, and similarly for an omitted ensures clause. Giving no assigns clause means that locations assigned by the function are not specified, so the caller has no information at all on this function's side effects. See Section 2.3.6 for more details on default status of clauses.

Example 2.18 *The following function is given a simple contract for computing the integer square root.*

```
/*@ requires x >= 0;
  @ ensures \result >= 0;
  @ ensures \result * \result <= x;
  @ ensures x < (\result + 1) * (\result + 1);
  @*/
int isqrt(int x);
```

The contract means that the function must be called with a nonnegative argument, and returns a value satisfying the conjunction of the three ensures *clauses. Inside these* ensures *clauses, the use of the construct* \old(x) *is not necessary, even if the function modifies the formal parameter* x*, because function calls modify a copy of the effective parameters, and the effective parameters remain unaltered. In fact,* x *denotes the effective parameter of* isqrt *calls, which has the same value interpreted in the pre-state as in the post-state.*

Example 2.19 *The following function is given a contract to specify that it increments the value pointed to by the pointer given as argument.*

[1] An ensures clause does not imply that the function will necessarily return.

[2] Functions that allocate or free memory can be specified with additional clauses described in section 2.7.3.

```
/*@ requires \valid(p);
  @ assigns *p;
  @ ensures *p == \old(*p) + 1;
  @*/
void incrstar(int *p);
```

The contract means that the function must be called with a pointer `p` *that points to a safely allocated memory location (see Section 2.7 for details on the* `\valid` *built-in predicate). It does not modify any memory location but the one pointed to by* `p`*. Finally, the* `ensures` *clause specifies that the value* `*p` *is incremented by one.*

2.3.3 Semantics of frame conditions

It is worth pointing out that there are different treatments of frame conditions (assigns statements) in various specification languages. The frame condition can follow either *writes* semantics or *modifies* semantics.

- Under `writes` (or assigns) semantics, only those memory locations listed in a frame condition may be *written to*, that is, may be the target of an assignment statement. This is true whether or not the value of the memory location changes.
- Under `modifies` semantics, a memory location may be written to, as long as the value is restored (that is, not modified) by the end of the scope of the function contract. Under this semantics, a frame condition is a requirement on the relationship between two states — that a memory location has the same value in a pre-state and a post-state.

Confusion can arise because the words *assigns* and *modifies* are sometimes used interchangeably. In particular, **ACSL uses modifies semantics, even though the frame condition is introduced by the *assigns* keyword.**[3]

2.3.4 Contracts with named behaviors

The general form of a function contract may contain named behaviors (restricted to two behaviors, in the following, for readability).

```
/*@ requires P;
  @ behavior b1:
  @   assumes  A1;
  @   requires R1;
  @   assigns  L1;
  @   ensures  E1;
  @ behavior b2:
  @   assumes  A2;
  @   requires R2;
  @   assigns  L2;
  @   ensures  E2;
  @*/
```

The names of behaviors must be distinct within the given function (or statement) contract.

The semantics of such a contract is as follows:

[3] For comparison, JML and the OpenJML tool define frame conditions to have write semantics but use the keywords assigns and modifies interchangeably; however, the KeY tool for JML implements modifies semantics. Ada/SPARK's data flow contracts effectively encode write semantics.

- The caller of the function must guarantee that the call is performed in a state where the property `P && (A1 ==> R1) && (A2 ==> R2)` holds.
- The called function returns a state where the properties `\old(A`$_i$`) ==> E`$_i$ hold for each i.
- For each i, if the function is called in a pre-state where `A`$_i$ holds, then each memory location of that pre-state that does not belong to the set `L`$_i$ is left unchanged in the post-state.

`requires` clauses in the behaviors are proposed mainly to improve readability (to avoid some duplication of formulas), since the contract above is equivalent to the following simplified one:

```
/*@ requires P && (A1 ==> R1) && (A2 ==> R2);
  @ behavior b1:
  @   assumes A1;
  @   assigns L1;
  @   ensures E2;
  @ behavior b2:
  @   assumes A2;
  @   assigns L2;
  @   ensures E2;
  @*/
```

A simple contract such as

```
/*@ requires P; assigns L; ensures E; */
```

is actually equivalent to a single named behavior as follows:

```
/*@ requires P;
  @ behavior <unique name>:
  @   assumes \true;
  @   assigns L;
  @   ensures E;
  @*/
```

Similarly, global `assigns` and `ensures` clauses are equivalent to a single named behavior. More precisely, the following contract

```
/*@ requires P;
  @ assigns L;
  @ ensures E;
  @ behavior b1: ...
  @ behavior b2: ...
  @ ...
  @*/
```

is equivalent to (if b_1 and b_2 do not have requires clauses)

```
/*@ requires P;
  @ behavior <unique name>:
  @   assumes \true;
  @   assigns L;
  @   ensures E;
  @ behavior b1: ...
  @ behavior b2: ...
  @ ...
  @*/
```

Example 2.20 *In the following,* `bsearch(t,n,v)` *searches for element* `v` *in array* `t` *between indices* `0` *and* `n-1`*.*

```
/*@ requires n >= 0 && \valid(t+(0..n-1));
  @ assigns \nothing;
  @ ensures -1 <= \result <= n-1;
  @ behavior success:
  @   ensures \result >= 0 ==> t[\result] == v;
  @ behavior failure:
  @   assumes t_is_sorted : \forall integer k1, integer k2;
  @              0 <= k1 <= k2 <= n-1 ==> t[k1] <= t[k2];
  @   ensures \result == -1 ==>
  @     \forall integer k; 0 <= k < n ==> t[k] != v;
  @*/
int bsearch(double t[], int n, double v);
```

The precondition requires array `t` *to be allocated at least from indices* `0` *to* `n-1`*. The two named behaviors correspond respectively to the successful behavior and the failing behavior.*

Since the function is performing a binary search, it requires the array `t` *to be sorted in increasing order: this is the purpose of the predicate named* `t_is_sorted` *in the* assumes *clause of the behavior named* `failure`*.*

See 2.4.2 for a continuation of this example.

Example 2.21 *The following function illustrates the importance of different* assigns *clauses for each behavior.*

```
/*@ behavior p_changed:
  @   assumes n > 0;
  @    requires \valid(p);
  @    assigns *p;
  @    ensures *p == n;
  @ behavior q_changed:
  @   assumes n <= 0;
  @    requires \valid(q);
  @    assigns *q;
  @    ensures *q == n;
  @*/
void f(int n, int *p, int *q) {
  if (n > 0) *p = n; else *q = n;
}
```

Its contract means that it may modify values pointed to by `p` *or by* `q`*, conditionally on the sign of* `n`*.*

Completeness of behaviors

In a contract with named behaviors, it is not required that the disjunction of the A_i is true, *i.e.* it is not mandatory to provide a "complete" set of behaviors. If such a condition is desired, it is possible to add the following clause to a contract:

```
/*@ ...
  @ complete behaviors b_1,...,b_n;
  @*/
```

It specifies that the set of behaviors $b_1,\dots,b_n$ is complete *i.e.* that

```
R ==> (A₁ || A₂ || ... || Aₙ)
```

holds, where `R` is the precondition of the contract. The simplified version of that clause

```
/*@ ...
  @ complete behaviors;
  @*/
```

means that all behaviors given in the contract should be taken into account.

Similarly, it is not required that two distinct behaviors are disjoint. If desired, this can be specified with the following clause:

```
/*@ ...
  @ disjoint behaviors b₁,...,bₙ;
  @*/
```

It means that the given behaviors are pairwise disjoint *i.e.* that, for all distinct i and j,

```
R ==> ! (Aᵢ && Aⱼ)
```

holds. The simplified version of that clause

```
/*@ ...
  @ disjoint behaviors;
  @*/
```

means that all behaviors given in the contract should be taken into account. Multiple `complete` and `disjoint` sets of behaviors can be given for the same contract.

2.3.5 Memory locations and sets of terms

There are several places where one needs to describe a set of memory locations: for example, in `assigns` clauses of function contracts and in `loop assigns` clauses (see section 2.4.2). A *memory location* is an l-value; a *tset* is a set of values and is primarily used for sets of l-values. Moreover, the l-values contained in the argument to an `assigns` clause (a tset) must be modifiable, as described in Section A.1. More generally, we introduce syntactic constructs to denote *sets of values* (tsets) that are also useful for the `\separated` predicate (see Section 2.7.2). The terms in a tset may have any type, though the operations described below are only well-typed for certain types of tsets. For example, $s_1[s_2]$ as defined below is only well-typed if one of s_1 and s_2 is a set of arrays and the other a set of integers.

Ranges The `..` syntax for ranges of integers has the appearance of a binary operator but is not a binary operator with conventional precedence, because either or both operand is optional. A missing operand designates an open range, that is the range includes all integers in the negative (if the left operand is missing) or positive direction (if the right operand is missing). This range syntax is used only within parentheses to designate a set of integers (cf. Fig. 2.7 later) or within square brackets to designate a range of array indices, as shown in Figs. 2.1 and 2.7.

Tsets The grammar for tsets is given in Figure 2.7. The semantics is given below, where s denotes any *tset*.

- `\empty` denotes the empty set.

range	::=	*term*$^?$ `..` *term*$^?$	
tset	::=	`\empty`	empty set
	|	*tset* `->` *id*	
	|	*tset* `.` *id*	
	|	`*` *tset*	
	|	`&` *tset*	
	|	*tset* `[` *tset* `]`	
	|	*tset* `[` *range* `]`	
	|	`(` *range* `)`	a range as a set of integers
	|	`\union` `(` *tset* `(,` *tset*`)`* `)`	union of locations
	|	`\inter` `(` *tset* `(,` *tset*`)`* `)`	intersection
	|	*tset* `+` *tset*	
	|	`(` *tset* `)`	
	|	`{` *tset* `|` *binders* `(;` *pred*`)`$^?$ `}`	set comprehension
	|	`{` `(` *term* `(,` *term* `)`*`)`$^?$ `}`	explicit set
	|	*term* [a]	implicit singleton
pred	::=	`\subset` `(` *tset* `,` *tset* `)`	set inclusion
	|	*term* `\in` *tset*	set membership

[a] The given term may not itself be a set

Figure 2.7: Grammar for sets of memory locations

- a simple term denotes a singleton set.
- `s->id` denotes the set of `x->id` for each `x` $\in$ `s`.
- `s.id` denotes the set of `x.id` for each `x` $\in$ `s`.
- `*s` denotes the set of `*x` for each `x` $\in$ `s`.
- `&s` denotes the set of `&x` for each `x` $\in$ `s`.
- $s_1[s_2]$ denotes the set of $x_1[x_2]$ for each $x_1 \in s_1$ and $x_2 \in s_2$.
- t_1 `..` t_2 denotes the set of integers between t_1 and t_2, inclusive. If $t_1 > t_2$, this is the same as `\empty`
- `\union`$(s_1,\ldots,s_n)$ denotes the union of s_1,s_2,... and s_n;
- `\inter`$(s_1,\ldots,s_n)$ denotes the intersection of s_1,s_2,... and s_n;
- s_1+s_2 denotes the set of x_1+x_2 for each $x_1 \in s_1$ and $x_2 \in s_2$;
- `{` $t_1,\ldots,t_n$ `}` is the set composed of the elements t_1, ..., t_n.
- `(s)` denotes the same set as `s`;
- `{ s | b ; P }` denotes set comprehension, that is the union of the sets denoted by `s` for each value `b` of binders satisfying predicate `P` (binders `b` are bound in both `s` and `P`).
- `x \in s` holds if and only if `x` is an element of `s`. The operator has the same precedence as relational predicates (e.g., `<`).

- \subset(s_1,s_2) holds if and only if each element of s_1 is also an element of s_2 (that is, s_1 is a subset of s_2).

Note that assigns \nothing is equivalent to assigns \empty; it is left for convenience.

Example 2.22 *The following function sets each cell of an array to 0.*

```
/*@ requires \valid(t+(0..n-1));
  @ assigns t[0..n-1];
  @ assigns *(t+(0..n-1));
  @ assigns *(t+{ i | integer i ; 0 <= i < n });
  @*/
void reset_array(int t[],int n) {
  int i;
  for (i=0; i < n; i++) t[i] = 0;
}
```

It is annotated with three equivalent assigns *clauses, each one specifying that only the set of cells* {t[0],...,t[n-1]} *is potentially modified.*

Example 2.23 *The following function increments each value stored in a linked list.*

```
struct list {
  int hd;
  struct list *next;
};

// reachability in linked lists
/*@ inductive reachable{L}(struct list *root, struct list *to) {
  @   case empty{L}: \forall struct list *l; reachable(l,l) ;
  @   case non_empty{L}: \forall struct list *l1,*l2;
  @      \valid(l1) && reachable(l1->next,l2) ==> reachable(l1,l2) ;
  @ }
*/

// The requires clause forbids giving a circular list
/*@ requires reachable(p,\null);
  @ assigns { q->hd | struct list *q ; reachable(p,q) } ;
  @*/
void incr_list(struct list *p) {
  while (p) { p->hd++ ; p = p->next; }
}
```

The assigns *clause specifies that the set of possibly modified memory locations is the set of fields* q->hd *for each pointer* q *reachable from* p *following* next *fields. See Section 2.6.3 for details about the declaration of the predicate* reachable*.*

2.3.6 Default contracts, multiple contracts

A C function can be defined only once but declared several times. It is allowed to annotate each of these declarations with contracts. Those contracts are seen as a single contract with the union of the requires clauses and behaviors.

On the other hand, a function may have no contract at all, or a contract with missing clauses. Missing requires and ensures clauses default to \true. If no assigns clause is given, it remains

C-compound-statement	::=	`{` *declaration** *statement** *assertion*$^+$ `}`
C-statement	::=	*assertion* *C-statement*
assertion-kind	::=	assert | check
assertion	::=	`/*@` *assertion-kind* *pred* `;` `*/`
	|	`/*@` for *id* (`,` *id*)* `:` *assertion-kind* *pred* `;` `*/`

Figure 2.8: Grammar for assertions

unspecified. If the function under consideration has only a declaration but no body, then it means that it potentially modifies "everything", hence in practice it will be impossible to verify anything about programs calling that function; in other words giving it a contract is in practice mandatory. On the other hand, if that function has a body, giving no assigns clause means in practice that it is left to tools to compute an over-approximation of the sets of modified locations.

2.4 Statement annotations

Annotations on C statements are of three kinds:

- Assertions: allowed before any C statement or at end of blocks.
- Loop annotations: invariant , assigns clause, variant ; allowed before any loop statement: while, for, and do ... while.
- Statement contracts: allowed before any C statement, specifying their behavior in a similar manner to C function contracts.

2.4.1 Assertions

The syntax of assertions is given in Figure 2.8, as an extension of the grammar of C statements.

- assert `P` means that `P` must hold in the current state (the sequence point where the assertion occurs).
- The variant for $\texttt{id}_1$,...,$\texttt{id}_k$: assert `P` associates the assertion to the named behaviors $\texttt{id}_i$, each of them being a behavior identifier for the current function (or a behavior of an enclosing block as defined later in Section 2.4.4). It means that this assertion is only required to hold for the listed behaviors.
- Introducing the assertion with the check keyword rather than assert indicates a *non-blocking* semantics. In other words, even if the property to be checked is found invalid, the execution of the program should continue unhindered.

statement	::=	`/*@` *loop-annot* `*/` `while` `(` *C-expression* `)` *C-statement*	
	|	`/*@` *loop-annot* `*/` `for` `(` *C-expression* `;` *C-expression* `;` *C-expression* `)` *C-statement*	
	|	`/*@` *loop-annot* `*/` `do` *C-statement* `while` `(` *C-expression* `)` `;`	
loop-annot [a]	::=	*loop-clause** *loop-behavior** *loop-variant*?	
loop-clause	::=	*loop-invariant* | *loop-assigns*	
	|	*loop-allocation*	
loop-invariant	::=	`loop` `invariant` *pred* `;`	
loop-assigns	::=	`loop` `assigns` *locations* `;`	
loop-behavior	::=	`for` *id* (`,` *id*)* `:` *loop-clause*+	annotation for behavior *id*
loop-variant	::=	`loop` `variant` *term* `;`	
	|	`loop` `variant` *term* `for` *id* `;`	variant for relation *id*

[a] empty loop annotations are forbidden

Figure 2.9: Grammar for loop annotations

2.4.2 Loop annotations

The syntax of loop annotations is given in Figure 2.9, as an extension of the grammar of C statements. *Loop-allocation* clauses allow specifying which memory locations are dynamically allocated or deallocated by a loop; they are defined later in Section 2.7.3.

Loop invariants and loop assigns

The semantics of loop invariants and loop assigns is defined as follows: a simple loop annotation of the form

```
/*@ loop invariant I;
  @ loop assigns L;
  @*/
...
```

specifies that the following conditions hold.

- The predicate `I` holds before entering the loop (in the case of a `for` loop, this means right after the initialization expression).
- The predicate `I` is an inductive invariant, that is if `I` is assumed true in some state where the condition `c` is also true, and if execution of the loop body in that state ends

normally at the end of the body or with a `continue` statement, `I` is true in the resulting state. If the loop condition has side effects, these are included in the loop body in a suitable way:

- for a `while (c) s` loop, `I` must be preserved by the side-effects of `c` followed by `s`;
- for a `for(init;c;step) s` loop, `I` must be preserved by the side-effects of `c` followed by `s` followed by `step`;
- for a `do s while (c);` loop, `I` must be preserved by `s` followed by the side-effects of `c`.

Note that if `c` has side-effects, the invariant might not be true at the exit of the loop: the last "step" starts from a state where `I` holds, performs the side-effects of `c`, which in the end evaluates to false and exits the loop. Likewise, if a loop is exited through a `break` statement, `I` does not necessarily hold, as side effects may occur between the last state in which `I` was supposed to hold and the `break` statement.

- At any loop iteration, any location that was allocated before entering the loop and is not a member of `L` (interpreted in the current state, that is `LoopCurrent`) has the same value as before entering the loop (`LoopEntry`). In fact, the `loop assigns` clause specifies an inductive invariant for the locations that are not members of `L`.

Loop behaviors

A loop annotation preceded by `for id_1,...,id_k:` is similar to the above, but applies only for behaviors `id_1,...,id_k` of the current function, hence in particular holds only under the assumption of their `assumes` clauses.

Remarks

- The `\old` construct is not allowed in loop annotations. The `\at` form should be used to refer to another state (see Section 2.4.3).
- When a loop exits with `break` or `return` or `goto`, it is not required that the loop invariant holds. In such cases, locations that are not members of L can be assigned between the end of the previous iteration and the exit statement.
- If no `loop assigns` clause is given, assignments remain unspecified. It is left to tools to compute an over-approximation of the sets of assigned locations.

Example 2.24 *Here is a continuation of example 2.20. Note the use of a loop invariant associated to a function behavior.*

```
/*@ requires n >= 0 && \valid(t+(0..n-1));
  @ assigns \nothing;
  @ ensures -1 <= \result <= n-1;
  @ behavior success:
  @   ensures \result >= 0 ==> t[\result] == v;
  @ behavior failure:
  @   assumes t_is_sorted : \forall integer k1, int k2;
  @       0 <= k1 <= k2 <= n-1 ==> t[k1] <= t[k2];
  @   ensures \result == -1 ==>
  @      \forall integer k; 0 <= k < n ==> t[k] != v;
```

```
  @*/
int bsearch(double t[], int n, double v) {
  int l = 0, u = n-1;
  /*@ loop invariant 0 <= l && u <= n-1;
    @ for failure: loop invariant
    @   \forall integer k; 0 <= k < n && t[k] == v ==> l <= k <= u;
    @*/
  while (l <= u ) {
    int m = l + (u-l)/2; // better than (l+u)/2
    if (t[m] < v) l = m + 1;
    else if (t[m] > v) u = m - 1;
    else return m;
  }
  return -1;
}
```

Loop variants

Optionally, a loop annotation may include a loop variant of the form

```
/*@ loop variant m; */
```

where `m` is a term of type `integer`.

The semantics is as follows: for each loop iteration that terminates normally or with `continue`, the value of `m` at end of the iteration must be smaller than its value at the beginning of the iteration. Moreover, its value at the beginning of the iteration must be nonnegative. Note that the value of `m` at loop exit might be negative. It does not compromise termination of the loop. Here is an example:

Example 2.25

```
void f(int x) {
  //@ loop variant x;
  while (x >= 0) {
    x -= 2;
  }
}
```

It is also possible to specify termination orderings other than the usual order on integers, using the additional `for` modifier. This is explained in Section 2.5.

General inductive invariants

It is actually allowed to pose an inductive invariant anywhere inside a loop body. For example, it makes sense for a `do s while (c);` loop to contain an invariant right after statement `s`. Such an invariant is a kind of assertion, as shown in Figure 2.10.

Example 2.26 *In the following example, the natural invariant holds at this point (*`\max` *and* `\lambda` *are introduced later in Section 2.6.7). It would be less convenient to set an invariant at the beginning of the loop.*

assertion	::=	`/*@ invariant` *pred* `; */`
	|	`/*@ for` *id* `(,` *id*`)*` `: invariant` *pred* `; */`

Figure 2.10: Grammar for general inductive invariants

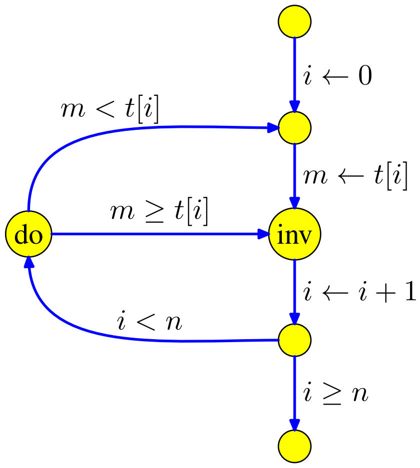
```
/*@ requires n > 0 && \valid(t+(0..n-1));
  @ ensures \result == \max(0,n-1,(\lambda integer k ; t[k]));
  @*/
double max(double t[], int n) {
  int i = 0; double m,v;
  do {
    v = t[i++];
    m = v > m ? v : m;
    /*@ invariant m == \max(0,i-1,(\lambda integer k ; t[k])); */
  } while (i < n);
  return m;
}
```

More generally, loops can be introduced by `goto`s. As a consequence, such invariants may occur anywhere inside a function's body. The meaning is that the invariant holds at that point, much like an `assert`. Moreover, the invariant must be inductive, *i.e.* it must be preserved across a loop iteration. Several invariants are allowed at different places in a loop body. These extensions are useful when dealing with complex control flows.

Example 2.27 *Here is a program annotated with an invariant inside the loop body:*

```
/*@ requires n > 0;
  @ ensures \result == \max(0,n-1,\lambda integer k; t[k]);
  @*/
double max_array(double t[], int n) {
  double m; int i=0;
  goto L;
  do {
    if (t[i] > m) { L: m = t[i]; }
    /*@ invariant
      @  0 <= i < n && m == \max(0,i,\lambda integer k; t[k]);
      @*/
    i++;
  }
  while (i < n);
  return m;
}
```

The control-flow graph of the code is as follows



The invariant is inductively preserved by the two paths that go from node "inv" to itself.

Example 2.28 *The program*

```
int x = 0;
int y = 10;

/*@ loop invariant 0 <= x < 11;
  @*/
while (y > 0) {
  x++;
  y--;
}
```

is not correctly annotated, even if it is true that `x` *remains smaller than 11 during the execution. This is because it is not true that the property* `x<11` *is preserved by the execution of* `x++ ; y--;`. *A correct loop invariant could be* `0 <= x < 11 && x+y == 10`. *It holds at loop entrance and is preserved (under the assumption of the loop condition* `y>0`*).*

Similarly, the following general invariants are not inductive:

```
int x = 0;
int y = 10;

while (y > 0) {
  x++;
  //@ invariant 0 < x < 11;
  y--;
  //@ invariant 0 <= y < 10;
}
```

since `0 <= y < 10` *is not a consequence of hypothesis* `0 < x < 11` *after executing* `y--;` *and* `0 < x < 11` *cannot be deduced from* `0 <= y < 10` *after looping back through the condition* `y>0` *and executing* `x++`. *Correct invariants could be:*

```
while (y > 0) {
  x++;
  //@ invariant 0 < x < 11 && x+y == 11;
  y--;
  //@ invariant 0 <= y < 10 && x+y == 10;
}
```

2.4.3 Built-in construct `\at`

Statement annotations usually need another additional construct `\at(e,id)` referring to the value of the expression `e` in the state at label `id`. In particular, for a C array of `int`, `t`, `\at(t,id)` is a logical array whose content is the same as the one of `t` in state at label `id`. It is thus very different from `\at((int *)t,id)`, which is a pointer to the first element of `t` (and stays the same between the state at `id` and the current state). Namely, if `t[0]` has changed since `id`, we have `\at(t,id)[0] != \at((int *)t,id)[0]`.

The label `id` can be either a regular C label or a label added within a ghost statement as described in Section 2.12. This label must be declared in the same function as the occurrence of `\at(e,id)`, but unlike `goto`s, more restrictive scoping rules must be respected:

- the label `id` must occur before the occurrence of `\at(e,id)` in the source;
- the label `id` must not be inside an inner block.

These rules are exactly the same rules as for the visibility of local variables within C statements (see [17], Section A11.1).

Default logic labels

There are seven predefined logic labels: `Pre`, `Here`, `Old`, `Post`, `LoopEntry`, `LoopCurrent` and `Init`. `\old(e)` is in fact syntactic sugar for `\at(e,Old)`.

- The label `Here` is visible in all statement annotations, where it refers to the state where the annotation appears; and in all contracts, where it refers to the pre-state for the requires, assumes, assigns, frees, decreases, terminates clauses and the post-state for ensures, allocates, and abrupt termination clauses. It is also visible in data invariants, presented in Section 2.11.
- The label `Old` is visible in assigns and ensures clauses of all contracts (both for functions and for statement contracts described below in Section 2.4.4), and refers to the pre-state of this contract.
- The label `Pre` is visible in all statement annotations, and refers to the pre-state of the function it occurs in.
- The label `Post` is visible in assigns and ensures clauses of all contracts, and it refers to the post-state.
- The label `LoopEntry` is visible in loop annotations and all annotations related to a statement enclosed in a loop. It refers to the state just before entering that loop for the first time –but after initialization took place in the case of a for loop, as for loop invariant (section 2.4.2). When `LoopEntry` is used in a statement enclosed in nested loops, it refers to the innermost loop containing that statement.
- The label `LoopCurrent` is visible in loop annotations and all other annotations related to a statement enclosed in a loop. It refers to the state at the beginning of the current step of the loop (see section 2.4.2 for more details on what constitutes a loop step in presence of side-effects in the condition). When `LoopCurrent` is used in a statement enclosed in nested loops, it refers to the innermost loop containing that statement.

term	::=	`\at` (*term* , *label-id*)
pred	::=	`\at` (*pred* , *label-id*)
label-id	::=	`Here` | `Old` | `Pre` | `Post`
	|	`LoopEntry` | `LoopCurrent` | `Init` |
	|	*id*

Figure 2.11: Grammar for `at` construct

- The label `Init` is visible in all statement annotations and contracts. It refers to the state just before the call to the `main` function, once the global data have been initialized.

Inside loop annotations, the labels `LoopCurrent` and `Here` are equivalent, except inside clauses `loop frees` (see section 2.7.3) where `Here` is equivalent to `LoopEntry`.

There is one special case regarding formal parameters. Despite any surrounding `\at` construct or the type of clause, formal parameters in a function contract are always interpreted in the pre-state.

No logic label is visible in global logic declarations such as lemmas, axioms, definition of predicate or logic functions. When such an annotation needs to refer to a given memory state, it has to be given a label binder: this is described in Section 2.6.9.

Example 2.29 *The code below implements the famous extended Euclid's algorithm for computing the greatest common divisor of two integers x and y, while computing at the same time the two* Bézout coefficients *p and q such that $p \times x + q \times y = \mathrm{gcd}(x, y)$. The loop invariant for the Bézout property needs to refer to the value of x and y in the pre-state of the function.*

```
/*@ requires x >= 0 && y >= 0;
  @ behavior bezoutProperty:
  @   ensures (*p)*x+(*q)*y == \result;
  @*/
int extended_Euclid(int x, int y, int *p, int *q) {
  int a = 1, b = 0, c = 0, d = 1;
  /*@ loop invariant x >= 0 && y >= 0 ;
    @ for bezoutProperty: loop invariant
    @    a*\at(x,Pre)+b*\at(y,Pre) == x &&
    @    c*\at(x,Pre)+d*\at(y,Pre) == y ;
    @ loop variant y;
    @*/
  while (y > 0) {
    int r = x % y;
    int q = x / y;
    int ta = a, tb = b;
    x = y; y = r;
    a = c; b = d;
    c = ta - c * q; d = tb - d * q;
  }
  *p = a; *q = b;
  return x;
}
```

Example 2.30 *Here is a toy example illustrating tricky issues with* `\at` *and labels:*

```
int i;
int t[10];

//@ ensures 0 <= \result <= 9;
int any();

/*@ assigns i,t[\at(i,Post)];
  @ ensures
  @   t[i] == \old(t[\at(i,Here)]) + 1;
  @ ensures
  @   \let j = i; t[j] == \old(t[j]) + 1;
  @*/
void f() {
  i = any();
  t[i]++;
}
```

The two `ensures` *clauses are equivalent. The simpler clause* `t[i] == \old(t[i]) + 1` *would be wrong because in* `\old(t[i])`, `i` *denotes the value of i n the pre-state.*

Also, the `assigns` *clause* `i,t[i]` *would be wrong too because again in* `t[i]`, *the value of i in the pre-state is considered.*

Example 2.31 *Here is an example illustrating the use of* `LoopEntry` *and* `LoopCurrent`

```
void f (int n) {
  for (int i = 0; i < n; i++) {
  /*@ assert \at(i,LoopEntry) == 0; */
  int j = 0;
  while (j++ < i) {
    /*@ assert \at(j,LoopEntry) == 0; */
    /*@ assert \at(j,LoopCurrent) + 1 == j; */
    }
  }
}
```

2.4.4 Statement contracts

The grammar for statement contracts is given in Figure 2.12. It is similar to function contracts, but without a `decreases` clause. Additionally, a statement contract may refer to enclosing named behaviors, with the form `for id:...`. Such contracts are only valid for the corresponding behaviors, in particular only under the corresponding `assumes` clause. Note that behaviors in statement contracts may have the same ids as enclosing function contract behaviors or enclosing statement contracts. In such cases, a use of the id (in a `for` construct) refers to the innermost behavior id.

The `ensures` clause does not constrain the post-state when the annotated statement is terminated by a `goto` jumping out of it, by any abrupt termination of the statement that is annotated. To specify such behaviors, *abrupt clauses* (described in Section 2.9) need to be used.

On the other hand, it is different with `assigns` clauses. The locations having their values modified during the path execution, starting at the beginning of the annotated statement and leading to a `goto` jumping out of it, should be part of its `assigns` clause.

statement	::=	`/*@` *statement-contract* `*/` *statement*
statement-contract [a]	::=	(`for` *id* (`,` *id*)* `:`)? *requires-clause** *simple-clause-stmt** *named-behavior-stmt** *completeness-clause**
simple-clause-stmt	::=	*simple-clause* | *abrupt-clause-stmt*
named-behavior-stmt	::=	`behavior` *id* `:` *behavior-body-stmt*
behavior-body-stmt [b]	::=	*assumes-clause** *requires-clause** *simple-clause-stmt**

[a] empty contracts are forbidden
[b] empty behavior bodies are forbidden

Figure 2.12: Grammar for statement contracts

Example 2.32 *The clause* `assigns \nothing;` *does not hold for that statement, even if the clause* `ensures x==\old(x);` *holds:*

```
/*@ assigns x;
  @ ensures x==\old(x);
  @*/
  if (c) {
    x++;
    goto L;
  }
L: ...
```

Allocation-clauses allow specifying which memory locations are dynamically allocated or deallocated by the annotated statement from the *heap*; they are defined later in Section 2.7.3.

2.5 Termination

The property of termination concerns both loops and recursive function calls. Termination is guaranteed by attaching a measure function to each loop (an aspect already addressed in Section 2.4.2) and each recursive function. By default, a measure is an integer expression, and measures are compared using the usual ordering over integers (Section 2.5.1). It is also possible to define measures using other domains and/or using a different ordering relation (Section 2.5.2).

2.5.1 Integer measures

Functions are annotated with integer measures with the syntax

```
//@ decreases e;
```

and loops are annotated similarly with the syntax

```
//@ loop variant e;
```

where the logic expression `e` has type `integer`. For recursive calls, or for loops, this expression must decrease for the relation `R` defined by

```
R(x,y) <==> x > y && x >= 0.
```

In other words, the measure must be a decreasing sequence of integers which remain nonnegative, except possibly for the last value of the sequence (See example 2.25).

Example 2.33 *The clause* `loop variant u-l;` *can be added to the loop annotations of the example 2.24. The measure* `u-l` *decreases at each iteration, and remains nonnegative, except at the last iteration where it may become negative.*

```
  @ ...
  @ loop variant u-l; */
while ...
```

2.5.2 General measures

More general measures on other types can be provided, using the keyword `for`. For functions it becomes

```
//@ decreases e for R;
```

and for loops

```
//@ loop variant e for R;
```

In those cases, the logic expression `e` has some type τ and `R` must be a relation on τ, that is a binary predicate declared (see Section 2.6 for details) as

```
//@ predicate R(τ x, τ y) ···
```

Of course, to guarantee termination, it must be proved that `R` is a well-founded relation.

Example 2.34 *The following example illustrates a variant annotation using a pair of integers, ordered lexicographically.*

```
//@ ensures \result >= 0;
int dummy();

//@ type intpair = (integer,integer);

/*@ predicate lexico(intpair p1, intpair p2) =
  @   \let (x1,y1) = p1 ;
  @   \let (x2,y2) = p2 ;
  @     x1 < x2 && 0 <= x2 ||
  @     x1 == x2 && 0 <= y2 && y1 < y2;
  @*/

//@ requires x >= 0 && y >= 0;
void f(int x,int y) {
  /*@ loop invariant x >= 0 && y >= 0;
    @ loop variant (x,y) for lexico;
    @*/
  while (x > 0 && y > 0) {

    if (dummy()) {
      x--; y = dummy();
    }
```

```
    else y--;
  }
}
```

2.5.3 Recursive function calls

The precise semantics of measures on recursive calls, especially in the general case of mutually recursive functions, is given as follows. We call a set of mutually recursive functions that is a strongly connected component of the call graph a *cluster*. Within each cluster, each function must be annotated with a `decreases` clause with the same relation `R` (syntactically). Then, in the body of any function `f` of that cluster, any recursive call to a function `g` must occur in a state where the measure attached to `g` is smaller (w.r.t `R`) than the measure of `f` in the pre-state of `f`. This also applies when `g` is `f` itself.

Example 2.35 *Here are the classical factorial and Fibonacci functions:*

```

/*@ requires n <= 12;
  @ decreases n;
  @*/
int fact(int n) {
  if (n <= 1) return 1;
  return n * fact(n-1);
}

//@ decreases n;
int fib(int n) {
  if (n <= 1) return 1;
  return fib(n-1) + fib(n-2);
}
```

Example 2.36 *This example illustrates mutual recursion:*

```
/*@
   requires n>=0;
   decreases n;
*/
int even(int n) {
  if (n == 0) return 1;
  return odd(n-1);
}

/*@
   requires x>=0;
   decreases x;
*/
int odd(int x) {
  if (x == 0) return 0;
  return even(x-1);
}
```

2.5.4 Non-terminating functions

EXPERIMENTAL

There are cases where a function is not supposed to terminate. For instance, the `main` function of a reactive program might be a `while(1)` that indefinitely waits for an event to process. More generally, a function can be expected to terminate only if some preconditions are met. In those cases, a `terminates` clause can be added to the contract of the function, using the following form:

```
//@ terminates p;
```

The semantics of such a clause is as follows: if `p` holds, then the function is guaranteed to terminate (more precisely, its termination must be proved). If such a clause is not present (and in particular if there is no function contract at all), it defaults to `terminates \true;` that is the function is supposed to always terminate, which is the expected behavior of most functions.

Note that nothing is specified for the case where `p` does not hold: the function may terminate or not. In particular, `terminates \false;` does not imply that the function loops forever. A possible specification for a function that never terminates is the following:

```
/*@ ensures \false ;
    terminates \false ;
*/
void f() { while(1); }
```

Example 2.37 *A concrete example of a function that may not always terminate is the* `incr_list` *function of example 2.23. In fact, The following contract is also acceptable for this function:*

```
// this time, the specification accepts circular lists, but does not ensure
// that the function terminates on them (as a matter of fact, it does not).
/*@ terminates reachable(p,\null);
  @ assigns { q->hd | struct list *q ; reachable(p,q) } ;
  @*/
void incr_list(struct list *p) {
  while (p) { p->hd++ ; p = p->next; }
}
```

2.6 Logic specifications

The language of logic expressions used in annotations can be extended by declarations of new logic types, and new constants, logic functions and predicates. These declarations follow the classical setting of *algebraic specifications*. The grammar for these declarations is given in Figure 2.13.

2.6.1 Predicate and function definitions

New functions and predicates can be *defined* by explicit expressions, given after an equal sign.

Example 2.38 *The following code*

C-external-declaration	::=	`/*@` *logic-def*$^+$ `*/`[a]	
logic-def	::=	*logic-const-def*	
	|	*logic-function-def*	
	|	*logic-predicate-def*	
	|	*lemma-def*	
	|	*data-inv-def*	
type-var	::=	*id*	
type-expr	::=	*type-var*	type variable
	|	*id*	
		`<` *type-expr*	
		(, *type-expr*)* `>`	polymorphic type
type-var-binders	::=	`<` *type-var*	
		(, *type-var*)* `>`	
poly-id	::=	*ident* *type-var-binders*	polymorphic object identifier
logic-const-def	::=	logic *type-expr*	
		poly-id `=` *term* `;`	
logic-function-def	::=	logic *type-expr*	
		poly-id *parameters* `=`	
		term `;`	
logic-predicate-def	::=	predicate	
		poly-id *parameters*$^?$ `=`	
		pred `;`	
parameters	::=	`(` *parameter*	
		(, *parameter*)* `)`	
parameter	::=	*type-expr* *id*	
lemma-def	::=	lemma *poly-id* `:`	
		pred `;`	

[a] These may appear as global declarations

Figure 2.13: Grammar for global logic definitions

```
//@ predicate is_positive(integer x) = x > 0;
/*@ logic integer get_sign(real x) =
  @   x > 0.0 ? 1 : ( x < 0.0 ? -1 : 0);
  @*/
```

illustrates the definition of a new predicate `is_positive` *with an integer parameter and a new logic function* `sign` *with a real parameter returning an integer.*

2.6.2 Lemmas

Lemmas are user-given propositions, a facility that might help theorem provers establish validity of ACSL specifications.

Example 2.39 *The following lemma*

logic-def	::=	*inductive-def*
inductive-def	::=	inductive *poly-id* *parameters*$^?$ { *indcase** }
indcase	::=	case *poly-id* : *pred* ;

Figure 2.14: Grammar for inductive definitions

```
//@ lemma mean_property: \forall integer x,y; x <= y ==> x <= (x+y)/2 <= y;
```

is a useful hint for a program like binary search.

Of course, a complete verification of an ACSL specification has to provide a proof for each lemma.

2.6.3 Inductive predicates

A predicate may also be defined by an inductive definition. The grammar for this style of definition is given in Figure 2.14.

In general, an inductive definition of a predicate P has the form

```
/*@ inductive P(x_1,...,x_n) {
  @   case c_1 : p_1;
...
  @   case c_k : p_k;
  @ }
  @*/
```

where each c_i is an identifier and each p_i is a proposition.

The semantics of such a definition is that P is the least fixpoint of the cases, i.e. is the smallest predicate (in the sense that it is false the most often) satisfying the propositions $p_1,\dots,p_k$. With this general form, the existence of a least fixpoint is not guaranteed, so tools might enforce syntactic conditions on the form of inductive definitions. A standard syntactic restriction could be to allow only propositions p_i of the form

```
\forall y_1,...,y_m, h_1 ==> ··· ==> h_l ==> P(t_1,...,t_n)
```

where P occurs only positively in hypotheses $h_1,\dots,h_l$ (definite Horn clauses, http://en.wikipedia.org/wiki/Horn_clause).

Example 2.40 *The following introduces a predicate* isgcd(x,y,d)*, which means that* d *is the greatest common divisor of* x *and* y*.*

```
/*@ inductive is_gcd(integer a, integer b, integer d) {
  @ case gcd_zero:
  @   \forall integer n; is_gcd(n,0,n);
  @ case gcd_succ:
  @   \forall integer a,b,d; is_gcd(b, a % b, d) ==> is_gcd(a,b,d);
  @ }
  @*/
```

This definition uses definite Horn clauses, hence is consistent.

Example 2.23 already introduced an inductive definition of reachability in linked-lists, and was also based on definite Horn clauses, and is thus consistent.

logic-def	::=	*axiomatic-decl*	
axiomatic-decl	::=	axiomatic *id* { *logic-decl** }	
logic-decl	::=	*logic-def*	
	|	*logic-type-decl*	
	|	*logic-const-decl*	
	|	*logic-predicate-decl*	
	|	*logic-function-decl*	
	|	*axiom-def*	
logic-type-decl	::=	type *logic-type* ;	
logic-type	::=	*id*	
	|	*id* *type-var-binders*	polymorphic type
logic-const-decl	::=	logic *type-expr* *poly-id* ;	
logic-function-decl	::=	logic *type-expr* *poly-id* *parameters* ;	
logic-predicate-decl	::=	predicate *poly-id* *parameters*$^?$;	
axiom-def	::=	axiom *poly-id* : *pred* ;	

Figure 2.15: Grammar for axiomatic declarations

2.6.4 Axiomatic definitions

Instead of an explicit definition, one may introduce an *axiomatic* definition for a set of types, predicates and logic functions, which amounts to declaring the expected profiles and a set of axioms. The grammar for those constructions is given in Figure 2.15.

Example 2.41 *The following axiomatization introduces a theory of finite lists of integers a la LISP.*

```
/*@ axiomatic IntList {
  @ type int_list;
  @ logic int_list nil;
  @ logic int_list cons(integer n,int_list l);
  @ logic int_list append(int_list l1,int_list l2);
  @ axiom append_nil:
  @   \forall int_list l; append(nil,l) == l;
  @ axiom append_cons:
  @   \forall integer n, int_list l1,l2;
  @     append(cons(n,l1),l2) == cons(n,append(l1,l2));
  @ }
  @*/
```

Unlike inductive definitions, there is no syntactic condition that guarantees that axiomatic definitions are consistent. It is usually up to the user to ensure that the introduction of axioms does not lead to a logical inconsistency.

Example 2.42 *The following axiomatization*

```
/*@ axiomatic sign {
  @   logic  integer  get_sign(real x);
  @   axiom sign_pos: \forall  real  x; x >= 0. ==> get_sign(x) == 1;
  @   axiom sign_neg: \forall  real  x; x <= 0. ==> get_sign(x) == -1;
  @ }
  @*/
```

is inconsistent since it implies `sign(0.0) == 1` *and* `sign(0.0) == -1`*, hence* `-1 == 1`

2.6.5 Polymorphic logic types

We consider here an algebraic specification setting based on multi-sorted logic, where types can be *polymorphic* that is parametrized by other types. For example, one may declare the type of polymorphic lists as

```
//@ type list<A>;
```

One can then consider for instance list of integers (`list <integer>`), list of pointers (e.g. `list <char*>`), list of list of reals (`list<list <real> >`[4]), etc.

The grammar of Figure 2.13 contains rules for declaring polymorphic types and using polymorphic type expressions.

2.6.6 Recursive logic definitions

Explicit definitions of logic functions and predicates can be recursive. Declarations in the same bunch of logic declarations are implicitly mutually recursive, so that mutually recursive functions are possible too.

Example 2.43 *The following logic declaration*

```
/*@ logic  integer  max_index{L}(int t[],integer n) =
  @   (n==0) ? 0 :
  @   (t[n-1]==0) ? n-1 : max_index(t, n-1);
  @*/
```

defines a logic function that returns the maximal index `i` *between* `0` *and* `n-1` *such that* `t[i]=0`*.*

There is no syntactic condition on such recursive definitions, such as limitation to primitive recursion. In essence, a recursive definition of the form `f(args) = e;` where `f` occurs in expression `e` is just a shortcut for an axiomatic declaration of `f` with an axiom `\forall args; f(args) = e`. In other words, recursive definitions are not guaranteed to be consistent, in the same way that axiomatics may introduce inconsistency. Of course, tools might provide a way to check consistency.

2.6.7 Higher-order logic constructions

Experimental

Figure 2.16 introduces new term constructs for higher-order logic.

[4] In this latter case, note that the two '>' must be separated by a space, to avoid confusion with the shift operator.

term	::=	`\lambda` *binders* `;` *term*	abstraction
	|	*ext-quantifier* `(` *term* `,` *term* `,` *term* `)`	
	|	`{` *term* `\with` `[` *range* `]` `=` *term* `}`	array slice update
ext-quantifier	::=	`\max` | `\min` | `\sum`	
	|	`\product` | `\numof`	

Figure 2.16: Grammar for higher-order constructs

Abstraction The term `\lambda` τ_1 `x`$_1$`,...,`τ_n `x`$_n$`;` `t` denotes the n-ary logic function that maps `x`$_1$,...,`x`$_n$ to `t`. It has the same precedence as `\forall` and `\exists`

Extended quantifiers Terms `\`*quant*`(t`$_1$`,t`$_2$`,t`$_3$`)` where *quant* is `max`, `min`, `sum`, `product` or `numof` are extended quantifications. `t`$_1$ and `t`$_2$ must have type `integer`, and `t`$_3$ must be a unary function with an integer argument, and a numeric value (integer or real) except for `\numof` for which it should have a boolean value. Their meanings are given as follows:

$$
\begin{array}{rcl}
\texttt{\textbackslash max(i,j,f)} & = & \max\{\texttt{f(i)}, \texttt{f(i+1)}, \ldots, \texttt{f(j)}\} \\
\texttt{\textbackslash min(i,j,f)} & = & \min\{\texttt{f(i)}, \texttt{f(i+1)}, \ldots, \texttt{f(j)}\} \\
\texttt{\textbackslash sum(i,j,f)} & = & \texttt{f(i)} + \texttt{f(i+1)} + \cdots + \texttt{f(j)} \\
\texttt{\textbackslash product(i,j,f)} & = & \texttt{f(i)} \times \texttt{f(i+1)} \times \cdots \times \texttt{f(j)} \\
\texttt{\textbackslash numof(i,j,f)} & = & \#\{\texttt{k} \mid \texttt{i} \leq \texttt{k} \leq \texttt{j} \wedge \texttt{f(k)}\} \\
 & = & \texttt{\textbackslash sum(i,j,\textbackslash lambda integer k ; f(k) ? 1 : 0)}
\end{array}
$$

If `i>j` then `\sum` and `\numof` above are 0, `\product` is 1, and `\max` and `\min` are unspecified (see Section 2.2.2).

Array slice update A term of the form `{` *a* `\with` `[` *low* `..` *up* `]` `=` *f*`}` allows updating a slice of an array. *a* must be an array of τ and *f* a unary function taking as argument an `integer` and returning a value of type τ. Such a term denotes an array a' such that:

$$
a'[i] = \begin{cases} a[i] & \text{if } i < low \\ f(i) & \text{if } low \leq i \leq up \\ a[i] & \text{if } i > up \end{cases}
$$

If *low* (resp. *up*) is missing, then all the lower (resp. upper) part of the array gets modified in a'. If both bounds are omitted, all elements of a' are computed using f.

As a special case, a term of the form `{` *a* `\with` `[` *low* `..` *up* `]` `=` *v*`}` where v is a term of type τ is equivalent to `{` *a* `\with` `[` *low* `..` *up* `]` `=` `\lambda` $\mathbb{Z}$. *v* `}`, i.e. it evaluates to an array where the relevant cells all contain the same value v.

Finally, ranges can also be used in designated initializers (see section 2.2.7), with the same semantics as above.

Example 2.44 *Function that sums the elements of an array of doubles.*

```
/*@ requires n >= 0 && \valid(t+(0..n-1)) ;
  @ ensures \result == \sum(0,n-1,\lambda integer k; t[k]);
  @*/
double array_sum(double t[],int n) {
  int i;
  double s = 0.0;
  /*@ loop invariant 0 <= i <= n;
    @ loop invariant s == \sum(0,i-1,\lambda integer k; t[k]);
    @ loop variant n-i;
```

```
  */
  for (i=0; i < n; i++) s += t[i];
  return s;
}
```

Example 2.45 *Properties of arrays initialized as a whole slice*

```
//@ type seq = integer[];

//@ logic seq init = { [ .. ] = 0 };

//@ logic seq ident = { init \with [0 .. 10] = \lambda integer i; i };

//@ lemma init_def: \forall integer i; init[i] == 0;

//@ lemma ident_def1: \forall integer i; i < 0 ==> ident[i] == 0;

//@ lemma ident_def2: \forall integer i; 0 <= i <= 10 ==> ident[i] == i;

//@ lemma ident_def3: \forall integer i; i > 10 ==> ident[i] == 0;
```

2.6.8 Concrete logic types

EXPERIMENTAL

Logic types may not only be declared but also be given a definition. Defined logic types can be either record types, or sum types. These definitions may be recursive. For record types, the field access notation $t.id$ can be used, and for sum types, a pattern-matching construction is available. Grammar rules for these additional constructions are given in Figure 2.17

Example 2.46 *The declaration*

```
//@ type list<A> = Nil | Cons(A,list<A>);
```

introduces a concrete definition of finite lists. The logic definition

```
/*@ logic integer list_length<A>(list<A> l) =
  @   \match l {
  @     case Nil : 0
  @     case Cons(h,t) : 1+list_length(t)
  @   };
  @*/
```

defines the length of a list by recursion and pattern-matching.

2.6.9 Hybrid functions and predicates

Logic functions and predicates may take arguments with either (pure) C type or logic type. Such a predicate (or function) can either be defined with the same syntax as before (or axiomatized). However, such definitions usually depend on one or more program points, because it depends upon memory states, *via* expressions such as:

- pointer dereferencing: `*p`, `p->f`;

logic-def	::=	type *logic-type* = *logic-type-def* ;	
logic-type-def	::=	*record-type* | *sum-type*	
	|	*type-expr*	type abbreviation
record-type	::=	{ *type-expr* *id* (; *type-expr* *id*)* ;? }	
sum-type	::=	|? *constructor* (| *constructor*)*	
constructor	::=	*id*	constant constructor
	|	*id* (*type-expr* (, *type-expr*)*)	non-constant constructor
type-expr	::=	*product-type*	
product-type	::=	(*type-expr* (, *type-expr*)+)	product type
term	::=	*term* . *id*	record field access
	|	\match *term* { *match-cases* }	pattern-matching
	|	(*term* (, *term*)+)	tuples
	|	{ (. *id* = *term* ;)+ }	records
	|	\let (*id* (, *id*)+) = *term* ; *term*	
match-cases	::=	*match-case*+	
match-case	::=	case *pat* : *term*	
pat	::=	*id*	constant constructor
	|	*id* (*pat* (, *pat*)*)	non-constant constructor
	|	*pat* | *pat*	or pattern
	|	_	any pattern
	|	*literal* | { (. *id* = *pat*)* }	record pattern
	|	(*pat* (, *pat*)*)	tuple pattern
	|	*pat* as *id*	pattern binding

Figure 2.17: Grammar for concrete logic types and pattern-matching

- array access: `t[i]`;
- address-of operator: `&x`;
- built-in predicate depending on memory: \valid

To make such a definition safe, it is mandatory to add after the declared identifier a set of labels, between curly braces. We then speak of a *hybrid* predicate (or function). The grammar for *ident* is extended as shown on Figure 2.18. Expressions as above must then be enclosed in an \at construct to refer to a given label. However, to ease reading of such logic expressions, it is allowed to omit a label whenever there is only one label in the context.

ident	::=	*id* *label-binders*	normal identifier with labels
label-binders	::=	{ *label-id* (, *label-id*)* }	label-id defined in Fig. 2.11

Figure 2.18: Grammar for logic declarations with labels

Example 2.47 *The following annotations declare a function that returns the number of occurrences of a given double in a memory block storing doubles between the given indexes, together with the related axioms. It should be noted that without labels, this axiomatization would be inconsistent, since the function would not depend on the values stored in t, hence the two last axioms would say both that* `a==b+1` *and* `a==b` *for some* `a` *and* `b`.

```
/*@ axiomatic NbOcc {
  @  // nb_occ(t,i,j,e) gives the number of occurrences of e in t[i..j]
  @  // (in a given memory state labelled L)
  @  logic integer nb_occ{L}(double *t, integer i, integer j,
  @                          double e);
  @ axiom nb_occ_empty{L}:
  @   \forall double *t, e, integer i, j;
  @     i > j ==> nb_occ(t,i,j,e) == 0;
  @ axiom nb_occ_true{L}:
  @   \forall double *t, e, integer i, j;
  @     i <= j && t[j] == e ==>
  @       nb_occ(t,i,j,e) == nb_occ(t,i,j-1,e) + 1;
  @ axiom nb_occ_false{L}:
  @   \forall double *t, e, integer i, j;
  @     i <= j && t[j] != e ==>
  @       nb_occ(t,i,j,e) == nb_occ(t,i,j-1,e);
  @ }
  @*/
```

Example 2.48 *This second example defines a predicate that indicates whether two memory blocks of the same size are a permutation of each other. It illustrates the use of more than a single label. Thus, the* `\at` *operator is mandatory here. Indeed the two blocks may come from two distinct memory states. Typically, one of the post conditions of a sorting function would be* `permut{Pre,Post}(t,t)`.

```
/*@ axiomatic Permut {
  @ // permut{L1,L2}(t1,t2,n) is true whenever t1[0..n-1] in state L1
  @ // is a permutation of t2[0..n-1] in state L2
  @  predicate permut{L1,L2}(double *t1, double *t2, integer n);
  @ axiom permut_refl{L}:
  @   \forall double *t, integer n; permut{L,L}(t,t,n);
  @ axiom permut_sym{L1,L2} :
  @   \forall double *t1, *t2, integer n;
  @     permut{L1,L2}(t1,t2,n) ==> permut{L2,L1}(t2,t1,n) ;
  @ axiom permut_trans{L1,L2,L3} :
  @   \forall double *t1, *t2, *t3, integer n;
  @     permut{L1,L2}(t1,t2,n) && permut{L2,L3}(t2,t3,n)
  @     ==> permut{L1,L3}(t1,t3,n) ;
  @ axiom permut_exchange{L1,L2} :
  @   \forall double *t1, *t2, integer i, j, n;
  @       \at(t1[i],L1) == \at(t2[j],L2) &&
  @       \at(t1[j],L1) == \at(t2[i],L2) &&
```

logic-function-decl	::=	logic *type-expr* *poly-id* *parameters* *reads-clause* ;
logic-predicate-decl	::=	predicate *poly-id* *parameters*$^?$ *reads-clause* ;
reads-clause	::=	reads *locations*
logic-function-def	::=	logic *type-expr* *poly-id* *parameters* *reads-clause* = *term* ;
logic-predicate-def	::=	predicate *poly-id* *parameters*$^?$ *reads-clause* = *pred* ;

Figure 2.19: Grammar for logic declarations with reads clauses

```
@       ( \forall  integer  k; 0 <= k < n && k != i && k != j ==>
@            \at(t1[k],L1) == \at(t2[k],L2))
@       ==> permut{L1,L2}(t1,t2,n);
@ }
@*/
```

2.6.10 Memory footprint specification: reads clause

Experimental

Logic declarations may be augmented with a reads clause, with the syntax given in Figure 2.19, which extends the syntax in Figure 2.13. This feature allows specifying the *footprint* of a hybrid predicate or function, that is, the set of memory locations that it depends on. From such information, one might deduce properties of the form $f\{L_1\}(args) = f\{L_2\}(args)$ if it is known that between states L_1 and L_2, the memory changes are disjoint from the declared footprint. Only mutable locations need be listed in a reads footprint: locations that hold constants that do not change in the course of a program may be omitted.

Example 2.49 *The following is the same as example 2.47 augmented with a* reads *clause.*

```
/*@ axiomatic Nb_occ {
  @ logic  integer  nb_occ{L}(double *t, integer i, integer j,
  @                          double e)
  @       reads t[i..j];
  @
  @ axiom nb_occ_empty{L}: // ...
  @
  @ // ...
  @ }
  @*/
```

If for example a piece of code between labels `L_1` *and* `L_2` *only modifies* `t[k]` *for some index* `k` *outside* `i..j`*, then one can deduce that* `nb_occ{L_1}(t,i,j,e)==nb_occ{L_2}(t,i,j,e)`.

2.6.11 Specification Modules

Experimental

Specification modules can be provided to encapsulate several logic definitions, for example

```

/*@ module List {
  @
  @   type list<A> = Nil | Cons(A , list<A>);
  @
  @   logic  integer  length<A>(list<A> l) =
  @     \match l {
  @       case Nil : 0
  @       case Cons(h,t) : 1+length(t) } ;
  @
  @   logic  A fold_right<A,B>((A -> B -> B) f, list<A> l, B acc) =
  @     \match l {
  @       case Nil : acc
  @       case Cons(h,t) : f(h,fold_right(f,t,acc)) } ;
  @
  @   logic  list<A> filter<A>((A -> boolean) f, list<A> l) =
  @     fold_right((\lambda A x, list<A> acc;
  @       f(x) ? Cons(x,acc) : acc), Nil) ;
  @
  @ }
  @*/
```

Module components are then accessible using a qualified notation like `List::length`.

Predefined algebraic specifications can be provided as libraries (see section 3), and imported using a construct like

```
//@ import List;
```

where the file `List.acsl` contains logic definitions, like the `List` module above.

2.7 Pointers and physical addressing

The grammar for terms and predicates is extended with new constructs given in Figure 2.20. *Location-address* is a term denoting a set of memory locations (a tset). It is a set of values of some common pointer type as defined in Section 2.3.5. As indicated below where necessary, many built-in functions and predicates dealing with pointers depend on the size of the referenced type. Thus, they cannot be given a pointer to void as an argument. On the other hand, a pointer referencing an incomplete type (hence having an abstract size) is possible.

2.7.1 Memory blocks and pointer dereferencing

C memory is structured into allocated blocks that can come either from a declarator or a call to one of the `calloc`, `malloc` or `realloc` functions. A block is characterized by its base address, which is the address of the declared object (the first declared object in case of an array declarator) or the pointer returned by the allocating function (when the allocation succeeds), and its length.

ACSL provides the following built-in functions to deal with allocated blocks. Each of them takes an optional label identifier as argument. The default value of that label is defined in Section 2.4.3.

term	::=	`\null`
	|	`\base_addr` *one-label*? (*term*)
	|	`\block_length` *one-label*? (*term*)
	|	`\offset` *one-label*? (*term*)
	|	`\allocation` *one-label*? (*term*)
pred	::=	`\allocable` *one-label*? (*term*)
	|	`\freeable` *one-label*? (*term*)
	|	`\fresh` *two-labels*? (*term*, *term*)
	|	`\valid` *one-label*? (*location-address*)
	|	`\valid_read` *one-label*? (*location-address*)
	|	`\separated` (*location-address* , *location-addresses*)
one-label	::=	{ *label-id* }
two-labels	::=	{ *label-id*, *label-id* }
location-addresses	::=	*location-address* (, *location-address*)*
location-address	::=	*tset*

Figure 2.20: Grammar extension of terms and predicates about memory

- `\base_addr{L}(p)` returns the base address of the allocated block containing, at the label `L`, the pointer `p`
  `\base_addr{id} : void* → char*`

- `\block_length{L}(p)` returns the length (in bytes) of the allocated block containing, at the label `L`, its argument pointer.
  `\block_length{id} : void* → size_t`

In addition, dereferencing a pointer may lead to run-time errors. A pointer `p` is said to be *valid* if `*p` is guaranteed to produce a definite value according to the C standard [16]. The following built-in predicates deal with this notion:

- `\valid` applies to a tset (see Section 2.3.5) each of whose elements has some common pointer type (other than `void*`). `\valid{L}(s)` holds if and only if dereferencing any $p \in$ s is safe at label `L`, both for reading from `*p` and writing to it. In particular, `\valid{L}(\empty)` holds for any label `L`.
  `\valid{id} : set<`α` *> → boolean`

- `\valid_read` applies to a tset of some pointer type (other than `void*`) and holds if and only if it is safe to read from all the pointers in the set
  `\valid_read{id} : set<`α` *> → boolean`

`\valid{L}(s)` implies `\valid_read{L}(s)` but the reverse is not true. In particular, it is allowed to read from a string literal, but not to write in it (see [16], 6.4.5§6).

The status of `\valid` and `\valid_read` constructs depends on the type of their argument. Namely, `\valid{L}((int *) p)` and `\valid{L}((char *)p)` are not equivalent. On the other hand, if we ignore potential alignment constraints, the following equivalence is true for any pointer `p`:

```
\valid{L}(p) <==> \valid{L}(((char *)p)+(0 .. sizeof(*p)-1))
```

allocation-clause	::=	allocates *dyn-allocation-addresses* ;
	|	frees *dyn-allocation-addresses* ;
loop-allocation	::=	loop allocates *dyn-allocation-addresses* ;
	|	loop frees *dyn-allocation-addresses* ;
dyn-allocation-addresses	::=	*location-addresses*
	|	\nothing

Figure 2.21: Grammar for dynamic allocations and deallocations

and similarly for \valid_read

```
\valid_read{L}(p) <==> \valid_read{L}(((char *)p)+(0 .. sizeof(*p)-1))
```

Some shortcuts are provided:

- \null is an extra notation for the null pointer (*i.e.* a shortcut for (void*)0). As in C itself (see [16], 6.3.2.3§3), the constant 0 can have any pointer type. Note that \valid{L}((char*)\null) and \valid_read{L}((char*)\null) are always false, for any logic label L.
- \offset{L}(p) returns the offset between p and its base address
  \offset{id} : void* → size_t
  \offset{L}(p) = (char*)p - \base_addr{L}(p)
  Again, if there are no alignment constraints, the following property holds: for any set of pointers s and label L, \valid_read{L}(s) if and only if for all p∈s:

```
    \offset{L}(p) >= 0 && \offset{L}(p) + sizeof(*p) <= \block_length{L}(p)
```

2.7.2 Separation

ACSL provides a built-in function to deal with separation of locations:

- \separated applies to tsets (see Section 2.3.5) of some common pointer type other than void*. \separated(s_1,s_2) holds for any set of pointers s_1 and s_2 if and only if for all p∈s_1 and q∈s_2:

```
  forall integer i,j; 0 <= i < sizeof(*p), 0 <= j < sizeof(*q)
     ==> (char*)p + i != (char*)q + j
```

  In fact, \separated is an n-ary predicate.
  \separated(s_1,..,s_n) means that for each $i \neq j$, \separated(s_i,s_j).

2.7.3 Dynamic allocation and deallocation

EXPERIMENTAL

Allocation-clauses allow specifying which memory locations are dynamically allocated or deallocated. The grammar for those constructions is given in Figure 2.21.

allocates \nothing and frees \nothing are respectively equivalent to allocates \empty and frees \empty; it is left for convenience like for assigns clauses.

Allocation clauses for function and statement contracts

Clauses allocates and frees are tied together. The simple contract

```
/*@ frees P1,P2,...;
  @ allocates Q1,Q2,...;
  @*/
```

means that any memory address, that does not belong to the union of sets of some pointer (having a type other than void*) P_i and Q_j, has the same *allocation status* (see below) in the post-state as in the pre-state. The only difference between allocates and frees is that sets P_i are evaluated in the pre-state, and sets Q_i are evaluated in the post-state.

The built-in type allocation_status can take the following values:

```
/*@
type allocation_status =
    \static | \register | \automatic | \dynamic | \unallocated;
*/
```

Built-in function \allocation {L}(p) returns the allocation status of the block containing, at the label L, the pointer p

\allocation {id} : void* → allocation_status

This function is such that for any pointer p and label L

```
\allocation {L}(p) == \allocation{L}(\base_addr(p))
```

and

```
\allocation {L}(p)==\unallocated ==> !\valid_read{L}((char*)p)
```

allocates $Q_1,\ldots,Q_n$ is equivalent to the postcondition

```
\forall char* p;
\separated (\union(Q1,...,Qn),p)==>
    (\base_addr{Here}(p)==\base_addr{Pre}(p)
     && \block_length{Here}(p)==\block_length{Pre}(p)
     && \valid{Here}(p)<==>\valid{Pre}(p)
     && \valid_read{Here}(p)<==>\valid_read{Pre}(p)
     && \allocation {Here}(p)==\allocation{Pre}(p))
```

In fact, just as the assigns clause does not specify precisely which memory locations are modified (just which are permitted to be modified), the *allocation-clauses* do not specify which memory locations are dynamically allocated or deallocated (just those that might be allocated or deallocated). Pre-conditions and post-conditions should be added to complete the specifications about allocations and deallocations. The following shortcuts can be used for that:

- \allocable {L}(p) holds if and only if the pointer p refers, at the label L, to the base address of an unallocated memory block.

  \allocable {id} : void* → boolean

  For any pointer p and label L

  ```
  \allocable {L}(p) <==> (\allocation{L}(p)==\unallocated
                          && (void*)p==(void*)\base_addr{L}(p))
  ```

- \freeable{L}(p) holds if and only if the pointer p refers, at the label L, to the base address of an allocated memory block that can be safely released using the C function free. Note that \freeable(\null) does not hold, despite NULL being a valid argument to the C function free.
  \freeable{id} : void* → boolean
  For any pointer p and label L

```
\freeable{L}(p) <==> (\allocation{L}(p)==\dynamic
                          && (void*)p==(void*)\base_addr{L}(p))
```

- \fresh{L_0,L_1}(p,n) indicates that p refers to the base address of an allocated memory block at label L_1, but that it is not the case at label L_0. The predicate ensures also that, at label L_1, the length (in bytes) of the block allocated dynamically equals n.
  \fresh{id,id} : void*, integer → boolean
  For any pointer p and labels L_0 and L_1

```
\fresh{L0,L1}(p,n) <==> (\allocable{L0}(p) && \freeable{L1}(p) &&
                         \block_length{L1}(p)==n &&
                         \valid{L1}((char*)p+(0 .. (n-1)))
```

Example 2.50 malloc *and* free *functions can be specified as follows.*

```

typedef unsigned long size_t;

/*@ assigns   \nothing;
  @ allocates  \result;
  @ ensures   \result==\null || \fresh{Old,Here}(\result,n);
  @*/
void *malloc(size_t n);

/*@ requires p!=\null ==> \freeable{Here}(p);
  @ assigns  \nothing;
  @ frees    p;
  @ ensures  p!=\null ==> \allocable{Here}(p);
  @*/
void free(void *p);
```

Default labels for constructs dedicated to memory are such that the logic label Here *can be omitted.*

When a behavior contains only one of the two allocation clauses, the given clause specifies the whole set of memory addresses to consider. This means that the set value for the other clause of that behavior defaults to \nothing. Now, when neither of the two allocation clauses is given, the meaning is different for anonymous behaviors and named behaviors:

- a named behavior without allocation clauses does not specify anything about allocations and deallocations. The allocated and deallocated memory blocks are in fact specified by the anonymous behavior of the contract. There is no condition to verify for these named behaviors about allocations and deallocations;

- for an anonymous behavior, the absence of allocation clauses means that there is no newly allocated nor deallocated memory block. That is equivalent to stating allocates \nothing; (which is equivalent to allocates \nothing; frees \nothing;).

These rules are such that contracts without any allocation clause should be considered as having only one `allocates \nothing;` leading to a condition to verify for each anonymous behavior.

Example 2.51 *More precise specifications can be given using named behaviors under the assumption of* `assumes` *clauses.*

```

typedef unsigned long size_t;

//@ ghost int heap_status;
/*@ axiomatic dynamic_allocation {
  @   predicate is_allocable(size_t n) // Can a block of n bytes be allocated?
  @             reads heap_status;
  @ }
  @*/

/*@ allocates \result ;
  @ behavior allocation:
  @   assumes  is_allocable(n);
  @   assigns  heap_status;
  @   ensures  \fresh ( \result ,n);
  @ behavior no_allocation:
  @   assumes  !is_allocable(n);
  @   assigns  \nothing;
  @   allocates \nothing;
  @   ensures  \result ==\null ;
  @ complete behaviors;
  @ disjoint behaviors;
  @*/
void *malloc(size_t n);

/*@ frees p;
  @ behavior deallocation:
  @   assumes  p!=\null;
  @   requires  \freeable (p);
  @   assigns  heap_status;
  @   ensures  \allocable (p);
  @ behavior no_deallocation:
  @   assumes  p==\null;
  @   assigns  \nothing;
  @   frees    \nothing;
  @ complete behaviors;
  @ disjoint behaviors;
  @*/
void free(void *p);
```

The behaviors named `allocation` *and* `deallocation` *do not need an allocation clause. For example, the allocation constraint of the* `allocation` *behavior is given by the clause* `allocates \result` *of the anonymous behavior of the* `malloc` *function contract. To set a stronger constraint into the behavior named* `no_allocation`*, the clause* `allocates \nothing` *should be given.*

Allocation clauses for loop annotations

Loop annotations may contain similar clauses allowing one to specify which memory locations are dynamically allocated or deallocated by a loop. The grammar for those constructions is given in Figure 2.21.

The clauses loop allocates and loop frees are tied together. The simple loop annotation

```
/*@ loop frees P_1,P_2,...;
  @ loop allocates Q_1,Q_2,...; */
```

means that any memory address that does not belong to the union of sets of terms P_i and Q_i has the same allocation status in the current state as before entering the loop. The only difference between these two clauses is that sets P_i are evaluated in the state before entering the loop (label LoopEntry), and Q_i are evaluated in the current loop state (label LoopCurrent).

Just as for loop assigns, the loop annotations loop frees and loop allocates define a loop invariant.

More precisely, this loop annotation

```
//@ loop allocates Q_1,...,Q_n; */
```

is equivalent to the loop invariant

```
 \forall char* p;
\separated(\union(Q_1,...,Q_n),p) ==>
   (\base_addr{Here}(p)==\base_addr{LoopEntry}(p)
    && \block_length{Here}(p)==\block_length{LoopEntry}(p)
    && (\valid{Here}(p)<==>\valid{LoopEntry}(p))
    && (\valid_read{Here}(p)<==>\valid_read{LoopEntry}(p))
    && \allocation{Here}(p)==\allocation{LoopEntry}(p))
```

Example 2.52

```
/*@ assert \forall integer j; 0<=j<n ==> \freeable(q[j]); */
/*@ loop assigns q[0..(i-1)];
  @ loop frees    q[0..\at(i-1,LoopCurrent)];
  @ loop invariant \forall integer j ;
                  0 <= j < i ==> \allocable(\at(q[j],LoopEntry));
  @ loop invariant \forall integer j ; 0 <= i <= n;
  @*/
  for (i=0; i<n; i++) {
    free(q[i]);
    q[i]=NULL;
  }

```

The addresses of locations q[0..n] *are not modified by the loop, but their values are. The clause* loop frees *catches the set of the memory blocks that may have been released by the previous loop iterations. The first* loop invariant *defines exactly these memory blocks. On the other hand,* loop frees *indicates that the remaining blocks have not been freed since the beginning of the loop. Hence, they are still* \freeable *as expressed by the initial* assert, *and* free(q[i]) *will succeed at next step.*

A *loop-clause* without an allocation clause implicitly states loop allocates \nothing. That means the *allocation status* is not modified by the loop body. A *loop-behavior* without allocation clause means that the allocated and deallocated memory blocks are in fact specified by the allocation clauses of the *loop-clauses* (Grammar of *loop-clauses* and *loop-behaviors* is given in Figure 2.9).

2.8 Sets and lists

2.8.1 Finite sets

Sets of memory locations (tsets), as defined in Section 2.3.5, can be used as first-class values in annotations. All the elements of such a set must share the same type (modulo the usual implicit conversions). Sets have the built-in type `set<A>` where `A` is the type of terms contained in the set.

In addition, it is possible to consider sets of pointers to values of different types. In this case, the set is of type `set<char*>` and each of its elements `e` is converted to `(char*)e + (0..sizeof(*e)-1)`.

Example 2.53 *The following example defines the* footprint *of a structure, that is the set of locations that can be accessed from an object of this type.*

```
struct S {
  char *x;
  int *y;
};

//@ logic set<char*> footprint(struct S s) = \union(s.x,s.y) ;

/*@ logic set<char*> footprint2(struct S s) =
  @    \union(s.x,(char*)s.y+(0..sizeof(s.y)-1)) ;
  @*/

/*@ axiomatic Conv {
    axiom conversion: \forall struct S s;
      footprint(s) == \union(s.x,(char*) s.y + (0 .. sizeof(int) - 1));
    }
*/
```

In the first definition, since the arguments of union are a `set<char*>` *and a* `set<int*>`*, the result is a* `set<char*>` *(according to typing of union). In other words, the two definitions above are equivalent.*

This logic function can be used as an argument of `\separated` *or of an* `assigns` *clause.*

Thus, the `\separated` predicate satisfies the following property (with s_1 of type `set<`τ_1`*>` and s_2 of type `set<`τ_2`*>`)

```
\separated(s1,s2) <==>
  ( \forall τ1* p; \forall τ2* q;
     \subset(p,s1) && \subset(q,s2) ==>
       ( \forall \integer i,j;
          0 <= i < \sizeof(τ1) && 0 <= j < \sizeof(τ2) ==>
            (char*)p + i != (char*)q + j))
```

and a clause `assigns` L_1`,...,`L_n is equivalent to the postcondition

```
\forall char* p; \separated(\union(&L1,...,&Ln),p) ==> *p == \old(*p)
```

term	::=	`[| |]`	empty list
	|	`[|` *term* `(,` *term*`)*` `|]`	list of elements
	|	*term* `^` *term*	list concatenation (overloading bitwise-xor operator)
	|	*term* `*^` *term*	list repetition

Figure 2.22: Notations for built-in list datatype

2.8.2 Finite lists

The built-in type `\list <A>` can be used for finite sequences of elements of the same type `A`. For constructing such homogeneous lists, built-in functions and notations are available.

The term `\Nil` denotes the empty sequence.

```
\list <A> \Nil<A>;
```

The function `\Cons` prepends an element `elt` onto a sequence `tail`

```
\list <A> \Cons<A>(<A> elt, \list<A> tail);
```

while `\concat` concatenates two sequences

```
\list <A> \concat<A>(\list<A> front, \list<A> tail);
```

and `\repeat` repeats a sequence `n` times, `n` being a positive number

```
\list <A> \repeat<A>(\list<A> seq, integer n);
```

The semantics of these functions rely on two useful functions: `\length` returns the number of elements of a sequence `seq`

```
integer \length<A>(\list<A> seq);
```

and `\nth` returns the element that is at position `n` of the given sequence `seq`. The first element is at position 0.

```
<A> \nth<A>(\list<A> seq, integer n);
```

Last but not least, the functions `\repeat` and `\nth` aren't specified for negative number `n`. The function `\nth(l)` is also unspecified for index greater than or equal to `\length(l)`.

The notation `[| |]` is just the same thing as `\Nil` and `[|1,2,3|]` is the sequence of three integers. In addition the infix operator `^` (resp. `*^`) is the same as function `\concat` (resp. `\repeat`).

Example 2.54 *The following example illustrates using such a data structure and notations in connexion with ghost code.*

```
//@ ghost int ghost_trace;
/*@ axiomatic TraceObserver {
  @   logic \list<integer> observed_trace{L} reads ghost_trace;
  @ }
  @*/

/*@ assigns ghost_trace;
  @ ensures register : observed_trace == (\old(observed_trace) ^ [| a |]);
```

abrupt-clause	::=	*exits-clause*
exits-clause	::=	`exits` *pred* ;
abrupt-clause-stmt	::=	*breaks-clause* | *continues-clause* | *returns-clause*
breaks-clause	::=	`breaks` *pred* ;
continues-clause	::=	`continues` *pred* ;
returns-clause	::=	`returns` *pred* ;
term	::=	`\exit_status`

Figure 2.23: Grammar of contracts about abrupt terminations

```
  @*/
void track(int a);

/*@ requires empty_trace: observed_trace == \Nil;
  @ assigns ghost_trace;
  @ ensures head_seq: \nth(observed_trace,0) == x;
  @ behavior shortest_trace:
  @   assumes no_loop_entrance: n<=0;
  @   ensures shortest_len: \length(observed_trace) == 2;
  @   ensures shortest_seq: observed_trace == [| x, z |];
  @ behavior longest_trace:
  @   assumes loop_entrance: n>0;
  @   ensures longest_len: \length(observed_trace) == 2+n;
  @   ensures longest_seq:
  @     observed_trace == ([| x |] ^ ([| y |] *^ n) ^ [| z |]);
  @*/
void loops(int n, int x, int y, int z) {
  int i;
  //@ ghost track(x);
  /*@ loop assigns i, ghost_trace;
    @ loop invariant idx_min: 0<=i;
    @ loop invariant idx_max: 0<=n ? i<=n : i<=0;
    @ loop invariant inv_seq:
    @   observed_trace == (\at(observed_trace,LoopEntry) ^ ([| y |] *^ i));
    @*/
  for (i=0; i<n; i++) {
    //@ ghost track(y);
    ;
  }
  //@ ghost track(z);
}
```

The function `track` *adds a value to the tail of a ghost trace variable. Calls to that function inside ghost statements allow modifying that trace; also properties of the* `observed_trace` *can be specified. Notice that the assigned ghost variable is* `ghost_trace`*.*

2.9 Abrupt termination

The `ensures` clause of function and statement contracts does not constrain the post-state when an annotated function or statement terminates abruptly. In such cases, an *abrupt clause* can be used as a *simple clause* or in a *behavior body*. The allowed constructs are shown in Figure 2.23.

The clauses breaks, continues and returns can only be found in a statement contract and state properties on the program state that hold when the annotated statement terminates abruptly with the corresponding statement (break, continue or return).

Inside these clauses, the construct \old(e) is allowed and denotes, like for statement contracts ensures, assigns and allocates, the value of e in the pre-state of the statement. More generally, the visibility in *abrupt clauses* of predefined logic labels (presented in Section 2.4.3) is the same as in ensures clauses.

For the returns case, the \result construct is allowed (if the function does not return void) and is bound to the returned value.

Example 2.55 *The following example illustrates each abrupt clause of statement contracts.*

```
int f(int x) {

  while (x > 0) {

    /*@ breaks x % 11 == 0 && x == \old(x);
      @ continues (x+1) % 11 != 0 && x % 7 == 0 && x == \old(x)-1;
      @ returns (\result+2) % 11 != 0 && (\result+1) % 7 != 0
      @         && \result % 5 == 0 && \result == \old(x)-2;
      @ ensures (x+3) % 11 != 0 && (x+2) % 7 != 0 && (x+1) % 5 != 0
      @         && x == \old(x)-3;
      @*/
    {
      if (x % 11 == 0) break;
      x--;
      if (x % 7 == 0) continue;
      x--;
      if (x % 5 == 0) return x;
      x--;
    }
  }
  return x;
}
```

The exits clause can be used in both function and statement contracts to give behavioral properties to the main function or to any function that may exit the program, *e.g.* by calling the exit function. The simple contract

```
/*@ exits E;
  @*/
```

means that, if the program terminates while executing the corresponding function (or statement), then it exits in a post-state where the property E holds. In any other termination kind, the exits clause does not constrain the post-state.

In such clauses, \old(e) is allowed and denotes the value of e in the pre-state of the function or statement, and \exit_status is bound to the return code, *e.g.* the value returned by main or the argument passed to exit. The construct \exit_status can be used only in exits, assigns and allocates clauses; \result cannot be used in exits clauses.

Example 2.56 *Here is a complete specification of the* `exit` *function, which performs an unconditional exit of the* `main` *function:*

```
/*@ assigns \nothing;
  @ ensures \false;
  @ exits   \exit_status == status;
  @*/
void exit(int status);

int status;

/*@ assigns status;
  @ exits  !cond && \exit_status == 1 && status == val;
  @*/
void may_exit(int cond, int val) {
  if (! cond) {
    status = val;
    exit(1);
    }
}
```

Note that the specification of the `may_exit` *function is incomplete since it allows modifications of the variable* `status` *when no exit is performed. Using behaviors, it is possible to distinguish between the exit case and the normal case, as in the following specification:*

```
/*@ behavior no_exit :
  @   assumes cond;
  @   assigns \nothing;
  @   exits   \false;
  @ behavior no_return :
  @   assumes !cond;
  @   assigns status;
  @   exits   \exit_status == 1 && status == val;
  @   ensures \false;
  @*/
void may_exit(int cond, int val) ;
```

In contrast to `ensures` clauses, `assigns`, `allocates` and `frees` clauses of function and statement contracts constrain the post-state even when the annotated function or statement terminates abruptly. This is shown in example 2.56 for a function contract.

2.10 Dependencies information

EXPERIMENTAL

An extended syntax of `assigns` clauses, described in Figure 2.24, allows specifying data dependencies and *functional expressions*.

Such a clause indicates that the assigned values can only depend upon the locations mentioned in the `\from` part of the clause. Again, this is an over-approximation: all of the locations involved in the computation of the modified values must be present, but some of locations might not be used in practice. If the `\from` clause is absent, all of the locations reachable at the given point of the program are supposed to be used. Moreover, for a single location, it is possible to give the precise relation between its final value and the value of its dependencies. This expression is evaluated in the pre-state of the corresponding contract.

assigs-clause	::=	assigns *location* (, *location*)* (\from *locations*)? ;
	|	assigns *term* \from *locations* = *term* ;

Figure 2.24: Grammar for dependencies information

Example 2.57 *The following example is a variation of the* `array_sum` *function in example 2.44, in which the values of the array are added to a global variable* `total`.

```
double total = 0.0;

/*@ requires n >= 0 && \valid(t+(0..n-1)) ;
  @ assigns total
      \from t[0..n-1] = total + \sum(0,n-1,\lambda int k; t[k]);
  @*/
void array_sum(double t[],int n) {
  int i;
  for(i=0; i < n; i++) total += t[i];
  return ;
}
```

Example 2.58 *The composite element modifier operators can be useful for writing such functional expressions.*

```
struct buffer { int pos ; char buf[80]; } line;

/*@ requires 80 > line.pos >= 0 ;
  @ assigns line
  @   \from line =
          { line \with .buf =
                 { line.buf \with [line.pos] = (char)'\0' } };
  @*/
void add_eol() {
  line.buf[line.pos] = '\0' ;
}
```

2.11 Data invariants

Data invariants are properties on data that are supposed to hold permanently during the lifetime of these data. In ACSL, we distinguish between:

- *global* invariants and *type* invariants: the former only apply to specified global variables, whereas the latter are associated with a static type, and apply to any variables of the corresponding type;
- *strong* invariants and *weak* invariants: strong invariants must be valid at any time during program execution (more precisely at any *sequence point* as defined in the C standard), whereas weak invariants must be valid at *function boundaries* (function entrance and exit) but can be violated in between.

The syntax for declaring data invariants is given in Figure 2.25. The strength modifier defaults to `weak`.

data-inv-def	::=	*data-invariant* | *type-invariant*
data-invariant	::=	*inv-strength*? `global` `invariant` *id* : *pred* ;
type-invariant	::=	*inv-strength*? `type` `invariant` *id* (*C-type-name* *id*) = *pred* ;
inv-strength	::=	`weak` | `strong`

Figure 2.25: Grammar for declarations of data invariants

Example 2.59 *In the following example, we declare*

1. *a weak global invariant* `a_is_positive`*, which specifies that global variable* `a` *should remain positive (weakly, so this property might be violated temporarily between function calls);*

2. *a strong type invariant for variables of type* `temperature`*;*

3. *a weak type invariant for variables of type* `struct S`*.*

```
int a;
//@ global invariant a_is_positive: a >= 0 ;

typedef double temperature;
/*@ strong type invariant temp_in_celsius(temperature t) =
  @   t >= -273.15 ;
  @*/

struct S {
  int f;
};
//@ type invariant S_f_is_positive(struct S s) = s.f >= 0 ;
```

2.11.1 Semantics

The distinction between strong and weak invariants has to do with the sequence points where the property is supposed to hold. The distinction between global and type invariants has to do with the set of values on which they are supposed to hold.

- Weak global invariants are properties that apply to global data and hold at any function entrance and function exit.

- Strong global invariants are properties that apply to global data and hold at any step during execution (starting after initialization of these data).

- A weak type invariant on type τ must hold at any function entrance and exit, and applies to any value (variable, field, array element, formal parameter, etc.) with static type τ. If the result of the function is of type τ, the result must also satisfy its weak invariant at function exit.

- A strong type invariant on type τ must hold at any step during execution, and applies to any global variable, local variable, or formal parameter with static type τ. If the result of the function has type τ, the result must also satisfy its strong invariant at function exit. Again, it says nothing about fields, array elements, memory locations, etc. of type τ.

Example 2.60 *The following example illustrates the use of a weak data invariant on a local static variable.*

```
void out_char(char c) {
  static int col = 0;
  //@ global invariant I : 0 <= col <= 79;
  col++;
  if (col >= 80) col = 0;
}
```

Example 2.61 *Here is a longer example, the famous Dijkstra's Dutch flag algorithm.*

```
typedef enum { BLUE, WHITE, RED } color;
/*@ type invariant isColor(color c) =
  @   c == BLUE || c == WHITE || c == RED ;
  @*/

/*@ predicate permut{L1,L2}(color *t1, color *t2, integer n) =
  @ \at(\valid(t1+(0..n)),L1) && \at(\valid(t2+(0..n)),L2) &&
  @ \numof(0,n,\lambda integer i; \at(t1[i],L1) == BLUE) ==
  @ \numof(0,n,\lambda integer i; \at(t2[i],L2) == BLUE)
  @   &&
  @ \numof(0,n,\lambda integer i; \at(t1[i],L1) == WHITE) ==
  @ \numof(0,n,\lambda integer i; \at(t2[i],L2) == WHITE)
  @  &&
  @ \numof(0,n,\lambda integer i; \at(t1[i],L1) == RED) ==
  @ \numof(0,n,\lambda integer i; \at(t2[i],L2) == RED);
  @*/

/*@ requires \valid(t+i) && \valid(t+j);
  @ assigns t[i],t[j];
  @ ensures t[i] == \old(t[j]) && t[j] == \old(t[i]);
  @*/
void swap(color t[], int i, int j) {
  int tmp = t[i];
  t[i] = t[j];
  t[j] = tmp;
}
typedef struct flag {
  int n;
  color *colors;
} flag;
/*@ type invariant is_colored(flag f) =
  @   f.n >= 0 && \valid(f.colors+(0..f.n-1)) &&
  @   \forall integer k; 0 <= k < f.n ==> isColor(f.colors[k]) ;
  @*/

/*@ predicate isMonochrome{L}(color *t, integer i, integer j,
```

```
  @                       color c) =
  @   \forall integer k; i <= k <= j ==> t[k] == c ;
  @*/

/*@ assigns f.colors[0..f.n-1];
  @ ensures
  @   \exists integer b, integer r;
  @     isMonochrome(f.colors,0,b-1,BLUE) &&
  @     isMonochrome(f.colors,b,r-1,WHITE) &&
  @     isMonochrome(f.colors,r,f.n-1,RED) &&
  @     permut{Old,Here}(f.colors,f.colors,f.n-1);
  @*/
void dutch_flag(flag f) {
  color *t = f.colors;
  int b = 0;
  int i = 0;
  int r = f.n;
  /*@ loop invariant
    @   (\forall integer k; 0 <= k < f.n ==> isColor(t[k])) &&
    @   0 <= b <= i <= r <= f.n &&
    @   isMonochrome(t,0,b-1,BLUE) &&
    @   isMonochrome(t,b,i-1,WHITE) &&
    @   isMonochrome(t,r,f.n-1,RED) &&
    @   permut{Pre,Here}(t,t,f.n-1);
    @ loop assigns b,i,r,t[0 .. f.n-1];
    @ loop variant r - i;
    @*/
  while (i < r) {
    switch (t[i]) {
    case BLUE:
      swap(t, b++, i++);
      break;
    case WHITE:
      i++;
      break;
    case RED:
      swap(t, --r, i);
      break;
    }
  }
}
```

Note that in this example the invariant could be declared `strong`. However, not all C `enum`s would obey a corresponding invariant, because in C, `enum` values are just `int`s and can hold values other than those listed in the declaration of the `enum` type.

2.11.2 Model variables and model fields

A *model variable* is a variable introduced in the specification with the keyword `model`. Its type must be a logic type. Analogously, types may have *model fields*. These are used to provide abstract specifications for functions whose concrete implementation must remain private.

The precise syntax for declaring model variables and fields is given in Figure 2.26. It is presented as additions to the regular C grammar for declarations

The informal semantics of model variables is as follows.

- Model variables can only appear in specifications. They are not lvalues, thus they cannot be assigned directly (unlike ghost variables, see below).
- Nevertheless, a function contract might state that a model variable is assigned, meaning that the value of the model variable may be different between the pre and post states of the contract.
- When a function contract mentions model variables:
  - the precondition is implicitly existentially quantified over those variables;
  - the postconditions are universally quantified over the old values of model variables, and existentially quantified over the new values.

Thus, in practice, the only way to prove that a function body satisfies a contract with model variables is to provide an invariant relating model variables and concrete variables, as in the example below.

Model fields behave the same, but they are attached to any value whose static type is the one of the model declaration. A model field can be attached to any C type, not only to `struct`. When it is attached to a compound type, however, it must not have the same name as a C field of that compound type. In addition, model fields are "inherited" by a `typedef` in the sense that the newly defined type has also the model fields of its parents (and can acquire more, which will not be present for the parent). For instance, in the following code, `t1` has one model field `m1`, while `t2` has two model fields, `m1` and `m2`.

```
typedef int t1;
typedef t1 t2;
/*@ model t1 { int m1 }; */
/*@ model t2 { int m2 }; */
```

Example 2.62 *Here is an example of a specification for a function that generates fresh integers. The contract is given in terms of a model variable that is intended to represent the set of "forbidden" values, e.g. the values that have already been generated.*

```
/* public interface */

//@ model set<integer> forbidden = \empty;

/*@ assigns forbidden;
  @ ensures ! (\result \in \old(forbidden))
  @   && \result \in forbidden && \subset(\old(forbidden),forbidden);
  @*/
int gen();
```

declaration	::=	*C-declaration*	
	|	`/*@` `model` *parameter* `;` `*/`	model variable
	|	`/*@` `model` *C-type-name* `{` *parameter* `;`$^?$ `}` `;` `*/`	model field

Figure 2.26: Grammar for declarations of model variables and fields

The contract is expressed abstractly, telling that

- *the forbidden set of values is modified;*
- *the value returned is not in the set of forbidden values, thus it is "fresh";*
- *the new set of forbidden values contains both the value returned and the previous forbidden values. The new set may have more values than the union of* `{ \result }` *and* `\old(forbidden)`*.*

An implementation of this function might be as follows, where a decision has been made to generate values in increasing order, so that it is sufficient to record the last value generated. This decision is made explicit by an invariant.

```
/* implementation */
int gen() {
  static int x = 0;
  /*@ global invariant I: \forall integer k;
    @   Set::mem(k,forbidden) ==> x > k;
    @*/
  return x++;
}
```

Remarks Although the syntax of model variables is close to JML model variables, they differ in the sense that the type of a model variable is a logic type, not a C type. Also, the semantics above is closer to the one of B machines [1]. It should be noticed that program verification with model variables does not have a well-established theoretical background [22, 20], so we deliberately do not provide a precise semantics in this document .

2.12 Ghost variables and statements

Ghost variables and statements are like C variables and statements, but visible only in the specifications. They are introduced by the `ghost` keyword at the beginning of the annotation (i.e. `/*@ ghost ... */` or `//@ ghost ...` for one-line ghost code, as mentioned in section 1.2). The grammar is given in Figure 2.27, in which only the first form of annotation is used. In this figure, the *C-** non-terminals refer to the corresponding grammar rules of the ISO standard, without any ACSL extension. Any non-terminal of the form *ghost-non-term* for which no definition is given in the figure represents the corresponding *C-non-term* entry, in which any *entry* is substituted by *ghost-entry*.

The variations with respect to the C grammar are the following:

- Comments within ghost code must be introduced by `//` and extend until the end of the line (the ghost code itself is placed inside a C comment, so an embedded comment of the form `/* ... */` would be incorrect C code).
- It is however possible to write multi-line annotations inside ghost code. These annotations are enclosed between `/@` and `@/` (since as indicated above, `/*@ ... */` would lead to incorrect C code). As in normal annotations, `@` characters (most commonly at the beginning of a line and at the end of an annotation, before the final `@/`) are considered to be white space. This style of annotation is only needed and permitted within ghost code. Also, ghost code may not be written within enclosing ghost code.

ghost-type-specifier	::=	*C-type-specifier*	
	|	*logic-type*	
declaration	::=	*C-declaration*	
	|	`/*@ ghost` *ghost-declaration* `*/`	
direct-declarator	::=	*C-direct-declarator*	
	|	*direct-declarator* `(` *C-parameter-type-list*? `)` `/*@ ghost` `(` *ghost-parameter-type-list* `)` `*/`	ghost args
postfix-expression	::=	*C-postfix-expression*	
	|	*postfix-expression* `(` *C-argument-expression-list*? `)` `/*@ ghost` `(` *ghost-argument-expression-list* `)` `*/`	call with ghosts
statement	::=	*C-statement*	
	|	*statements-ghost*	
statements-ghost	::=	`/*@ ghost` *ghost-statement*+ `*/`	
ghost-selection-statement	::=	*C-selection-statement*	
	|	`if` `(` *C-expression* `)` *statement* `/*@ ghost else` *ghost-statement*+ `*/`	
struct-declaration	::=	*C-struct-declaration*	
	|	`/*@ ghost` *struct-declaration* `*/`	ghost field

Figure 2.27: Grammar for ghost statements

- Logical types, such as `integer` or `real` are authorized in ghost code.
- A non-ghost function can take ghost parameters. If such a ghost clause is present in the declarator, then the list of ghost parameters must be non-empty and fixed (no vararg ghost). The call to the function must then provide the appropriate number of ghost parameters.
- Any non-ghost *if-statement* that does not have a non-ghost `else` clause can be augmented with a ghost one. Similarly, a non-ghost `switch` can have a ghost `default :` clause if it does not have a non-ghost one (there are however semantic restrictions for valid ghost labelled statements in a switch, see next paragraph for details).

Semantics of Ghost Code The question of semantics is essential for ghost code. Informally, the semantics requires that ghost statements do not change the regular program execution.[5] This implies several conditions, including e.g.:

- Ghost code cannot modify a non-ghost C variable.
- Ghost code cannot modify a non-ghost structure field.
- If `p` is a ghost pointer pointing to a non-ghost memory location, then it is forbidden to assign `*p`.
- The body of a ghost function is ghost code, and hence may not modify non-ghost variables or fields.
- If a non-ghost C function is called in ghost code, it must not modify non-ghost variables or fields.
- If a structure has ghost fields, the `sizeof` of the structure is the same as the structure without ghost fields. Also, alignment of fields remains unchanged.
- The control-flow graph of a function must not be altered by ghost statements. In particular, no ghost `return` can appear in the body of a non-ghost function. Similarly, ghost `goto`, `break`, and `continue` cannot jump outside of the innermost non-ghost enclosing block.

The semantics is specified as follows. First, the execution of a program with ghost code involves a *ghost memory heap* and a *ghost stack*, disjoint from the regular heap and stack. Ghost variables lie in the ghost heap, as do the ghost fields of structures. Thus, every memory side-effect can be classified as ghost or non-ghost. Then, the semantics is that any memory side-effects of ghost code must be only in the ghost heap or the ghost stack.

Notice that this semantics is not statically decidable. It is left to tools to provide approximations, correct in the sense that any code statically detected as ghost must be semantically ghost.

Example 2.63 *The following example shows some invalid assignments of ghost pointers:*

```

void f(int x, int *q) {
  //@ ghost int *p = q;
  //@ ghost *p = 0;
  // above assignment is wrong: it modifies *q which lies
  // in regular memory heap

  //@ ghost p = &x;
  //@ ghost *p = 0;
  // above assignment is wrong: it modifies x which lies
  // in regular memory stack

}
```

Example 2.64 *The following example shows some invalid ghost statements:*

[5] Not checked in the current implementation

```
int f (int x, int y) {
  //@ ghost int z = x + y;
  switch (x) {
  case 0: return y;
  //@ ghost case 1: z=y;
  // above statement is correct.
  //@ ghost case 2: { z++; break; }
  // invalid, would bypass the non-ghost default
  default : y++;
  }
  return y;
}

int g(int x) {
  //@ ghost int z = x;
  if (x > 0) { return x; }
  //@ ghost else { z++; return x; }
  // invalid, would bypass the non-ghost return
  return x+1;
}
```

Differences between model variables and ghost variables A ghost variable is an additional specification variable that is assigned in ghost code like any C variable. On the other hand, a model variable cannot be assigned, but one can state it is modified and can express properties about the new value, in a non-deterministic way, using logic assertions and invariants. In other words, specifications using ghost variable assignments are executable.

Example 2.65 *The example 2.62 can also be specified with a ghost variable instead of a model variable:*

```
//@ ghost set<integer> forbidden = \empty;

/*@ assigns forbidden;
  @ ensures ! \subset( \result ,\old(forbidden))
  @   && \result \in forbidden
       && \subset(\old(forbidden),forbidden);
  @*/
int gen() {
   static int x = 0;
  /*@ global invariant I: \forall integer k;
    @    k \in forbidden ==> x > k;
    @*/
  x++;
  //@ ghost forbidden = \union(x,forbidden);
  return x;
}
```

2.12.1 Volatile variables

Volatile variables can not be used in logic terms, since reading such a variable may have a side effect, in particular two successive reads may return different values.

declaration	::=	//@ volatile *locations* (reads *id*)? (writes *id*)? ; [a]

[a] only implemented for C-external-declaration

Figure 2.28: Grammar for volatile constructs

Specifying properties of a volatile variable may be done via a specific construct to attach two ghost functions to it. This construct, described by the grammar of Figure 2.28, has the following shape:

```
volatile τ x;
```

```
//@ volatile x reads f writes g;
```

where `f` and `g` are ghost functions with the following prototypes:

```
τ f(volatile τ* p);
τ g(volatile τ* p, τ v);
```

This must be understood as a special construct to instrument the C code, where each access to the variable `x` is replaced by a call to `f(&x)`, and each assignment to `x` of a value `v` is replaced by `g(&x,v)`. If a given volatile variable is only read or only written to, the unused accessor function can be omitted from the `volatile` construct.

Example 2.66 *The following code is instrumented in order to inject fixed values at each read of variable* `x`*, and collect written values.*

```
volatile int x;

//@ ghost int injector_x[3] = { 1, 2, 3 };
//@ ghost int injector_count = 0;

/*@ ghost /@ requires p == &x;
  @         assigns injector_count; @/
  @ int reads_x(volatile int *p) {
  @   if (p == &x)
  @     return injector_x[injector_count++];
  @   else
  @     return 0; // should not happen
  @ }
  @*/

//@ ghost int collector_x[3];
//@ ghost int collector_count = 0;

/*@ ghost /@ requires p == &x;
  @         assigns collector_count; @/
  @ int writes_x(volatile int *p, int v) {
  @   if (p == &x)
  @     return collector_x[collector_count++] = v;
  @   else
  @    return 0; // should not happen
  @ }
  @*/

//@ volatile x reads reads_x writes writes_x;
```

pred	::=	\initialized	*one-label*$^?$	(	*location-address*	)
	|	\dangling	*one-label*$^?$	(	*location-address*	)

Figure 2.29: Grammar extensions regarding initialized and dangling memory

```

/*@ ensures collector_count == 3 && collector_x[2] == 2;
  @ ensures \result == 6;
  @*/
int main () {
  int i, sum = 0;
  for (i=0 ; i < 3; i++) {
    sum += x;
    x = i;
  }
  return sum;
}
```

2.13 Initialization and undefined values

`\initialized {L}(p)` is a predicate taking a set of pointers (having a type other than `void*`) to l-values as argument (cf. Fig. 2.29) and means that each l-value in this set is initialized at label `L`.

`\initialized {id} : set<α*> → bool`

Example 2.67 *In the following, the assertion is true.*

```
int f(int n) {
  int x;

  if (n > 0) x = n ; else x = -n;
  //@ assert \initialized {Here}(&x);
  return x;
}
```

Default labels are such that logic label `{Here}` *can be omitted.*

2.14 Dangling pointers

`\dangling{L}(p)` is a predicate taking a set of pointers (having a type other than `void*`) to l-values as argument (cf. Fig. 2.29) and means that each l-value in this set has a *dangling content* at label `L`. That is, its value is (or contains bits of) a dangling address: either the address of a local variable referred to outside of its scope or the address of a variable that has been dynamically allocated, then deallocated.

`\dangling{id} : set<α*> → bool`

Example 2.68 *In the following, the assertion holds.*

```

int* f() {
  int a;
  return &a;
}

int* g() {
  int* p = f();
  //@ assert \dangling{Here}(&p);
  return p+1;
}
```

In most cases, the arguments to `\dangling` are pointers to l-values that themselves have type pointer, so the usual signature of `\dangling` is actually `set<`α`**>` $\rightarrow$ `bool`. The signature `set<`α`*>` $\rightarrow$ `bool` is useful to handle pointer values that have been written inside scalar variables through heterogeneous casts.

Note that `\dangling` takes a set of memory locations as its argument. The predicate is true if *all* of the memory locations contained in the argument are dangling. That semantics implies that `!\dangling(s)` is true precisely when at least one of the locations in s is not dangling. `!\dangling(s)` does *not* mean that *all* of the indicated memory locations are *not dangling*, only that *some* are.

2.15 Well-typed pointers

EXPERIMENTAL

pred	::=	`\valid_function` (*location-address*)
location-address	::=	*tset*

Figure 2.30: Grammar for predicates related to well-typedness

The predicates of Figure 2.30 are used to relate the type of a pointer to the effective type of the memory location or function that is being pointed to.

Currently, only the compatibility of a function pointer with the type of the function it points to is axiomatized, through the predicate `\valid_function`. This predicate has type `set<`α`*>` $\rightarrow$ `bool`, and `\valid_function (p)` holds if and only if

- `p` is a pointer to a function of type `t`, and
- `*p` is a function whose type is *compatible* with `t`, in the sense of [16, §6.2.7]

Example 2.69 *In the following, the assertions are true.*

```
int* f (int x);

int main() {
  int* (*p)(int) = &f;
  //@ assert \valid_function((int* (*)(int)) p); // true
  //@ assert \valid_function((int* (*)()) p); // true (see C99 6.7.5.3:15)

```

```
  //@ assert !\valid_function((void* (*)(int)) p);
    // not compatible: void* and int* are not compatible (see C99 6.7.5.1:2)

  //@ assert !\valid_function(( volatile int* (*)(int)) p);
    // not compatible: qualifiers cannot be dropped (see C99 6.7.3:9)
  return 0;
}
```

2.16 Logic attribute annotations

Chapter 3

Libraries

Disclaimer: this chapter is unfinished, it is left here to give an idea of what it will look like in the final document.

This chapter is devoted to libraries of specification, built upon the ACSL specification language. Section 3.2 describes additional predicates introduced by the Jessie plugin of Frama-C, to propose a slightly higher level of annotation.

3.1 Libraries of logic specifications

A standard library is provided, in the spirit of the List module of Section 2.6.11

3.1.1 Real numbers

A library of general purpose functions and predicates over real numbers, floats and doubles.

Includes

- abs, exp, power, log, sin, cos, atan, etc. over reals
- isFinite predicate over floats and doubles (means not NaN nor infinity)
- rounding reals to floats or doubles with specific rounding modes.

3.1.2 Finite lists

- pure functions nil, cons, append, fold, etc.
- Path, Reachable, isFiniteList, isCyclic, etc. on C linked-lists.

3.1.3 Sets and Maps

Finite sets, finite maps, in ZB-style.

3.2 Jessie library: logical addressing of memory blocks

The Jessie library is a collection of logic specifications whose semantics is well-defined only on source codes free from architecture-dependent features. In particular it is currently incompatible with pointer casts or unions (although there is ongoing work to support some of them [23]). As a consequence, a valid pointer of some type $\tau*$ necessarily points to a memory block which contains values of type τ.

3.2.1 Abstract level of pointer validity

In the particular setting described above, it is possible to introduce the following logic functions:

```
/*@
  @ logic integer \offset_min{L}<a>(a *p);
  @ logic integer \offset_max{L}<a>(a *p);
  @/
```

- `\offset_min{L}(p)` is the minimum integer `i` such that `(p+i)` is a valid pointer at label `L`.
- `\offset_max{L}(p)` is the maximum integer `i` such that `(p+i)` is a valid pointer at label `L`.

The following properties hold:

```
\offset_min{L}(p+i) == \offset_min{L}(p)-i
\offset_max{L}(p+i) == \offset_max{L}(p)-i
```

It also introduces some syntactic sugar:

```
/*@
 predicate \valid_range{L}<a>(a *p,integer i,integer j) =
   \offset_min{L}(p) <= i && \offset_max{L}(p) >= j;
*/
```

and the ACSL built-in predicate `\valid{L}(p+(a..b))` is now equivalent to `\valid_range{L}(p,a,b)`.

3.2.2 Strings

EXPERIMENTALThe logic function

```
//@ logic integer \strlen(char* p);
```

denotes the length of a 0-terminated C string. It is a total function, whose value is nonnegative if and only if the pointer in the argument is really a string.

Example 3.1 *Here is a contract for the* `strcpy` *function:*

```
/*@ // src and dest cannot overlap
  @ requires \base_addr(src) != \base_addr(dest);
  @ // src is a valid C string
  @ requires \strlen(src) >= 0 ;
  @ // dest is large enough to store a copy of src up to the 0
  @ requires \valid(dest+(0..\strlen(src)));
  @ ensures
  @   \forall integer k; 0 <= k <= \strlen(src) ==> dest[k] == src[k];
  @*/
char* strcpy(char *dest, const char *src);
```

3.3 Memory leaks

Experimental

Verification of absence of memory leak is outside the scope of the specification language. On the other hand, various models could be set up, using for example ghost variables.

Chapter 4

Conclusion

This document presents a Behavioral Interface Specification Language for ANSI C source code. It provides a common basis that can be shared among different tools. The specification language described here is intended to evolve in the future and remain open to additional constructions. One interesting possible extension regards “temporal” properties in a large sense, such as liveness properties, which can sometimes be simulated by regular specifications with ghost variables [14], or properties on evolution of data over the time, such as the history constraints of JML, or in the Lustre assertion language.

Appendix A

Appendices

A.1 Glossary

pure expressions In ACSL setting, a *pure* expression is a C expression which contains no assignments, no incrementation operator `++` or `--`, no function call, and no access to a volatile object. The set of pure expressions is a subset of the set of C expressions without side effect (C standard [17, 16], §5.1.2.3, alinea 2).

left-values A *left-value* (*lvalue* for short) is an expression which denotes some place in the memory during program execution, either on the stack, on the heap, or in the static data segment. It can be either a variable identifier or an expression of the form `*e`, `e[e]`, `e.id` or `e->id`, where `e` is any expression and `id` a field name. See C standard, §6.3.2.1 for a more detailed description of lvalues.

A *modifiable lvalue* is an lvalue allowed in the left part of an assignment. In essence, all lvalues are modifiable except variables declared as `const` or of some array type with explicit length.

pre-state and post-state For a given function call, the *pre-state* denotes the program state at the beginning of the call, including the current values for the function parameters. The *post-state* denotes the program state at the return of the call.

For a statement annotation, the *pre-state* denotes the program state just prior to the annotation statement; the *post-state* denotes the program state immediately after execution of the annotated statement (which may be a block statement).

function behavior A *function behavior* (*behavior* for short) is a set of properties relating the pre-state and the post-state for a possibly restricted set of pre-states (behavior *assumptions*).

function contract A *function contract* (*contract* for short) forms a specification of a function, consisting of the combination of a precondition (a requirement on the pre-state for any caller to that function), a collection of behaviors, and possibly a measure in case of a recursive function.

A.2 Comparison with JML

Although we took our inspiration from the Java Modeling Language (aka JML [18]), ACSL is notably different from JML in two crucial aspects:

- ACSL is a BISL for C, a low-level structured language, while JML is a BISL for Java, an object-oriented inheritance-based high-level language. Not only are the language features not the same between Java and C, but the programming styles and idioms are very different, which then entails different ways of specifying behaviors. In particular, C has no inheritance or exceptions, and no language support for the simplest properties on memory (*e.g.*, the size of an allocated memory block).
- JML also supports runtime assertion checking (RAC) when typing, static analysis and automatic deductive verification fail. The example of CCured [24, 9], which also adds strong typing to C by relying on RAC, shows that it is not possible to do it in a modular way. Indeed, it is necessary to modify the layout of C data structures for RAC, which is not modular. The follow-up project Deputy [10] thus reduces the checking power of annotations in order to preserve modularity. In contrast, we choose not to restrain the power of annotations (*e.g.*, all first order logic formulas are allowed). To that end, we rely on manual deductive verification using an interactive theorem prover (*e.g.*, Coq) when every other technique fails.

In the remainder of this chapter, we describe these differences in further detail.

A.2.1 Low-level language vs. inheritance-based one

No inherited specifications

JML has a core notion of specification inheritance, which enables support for behavioral subtyping, by applying specifications of parent methods to overriding methods. Inheritance combined with visibility and modularity account for a number of complex features in JML (*e.g.*, `spec_public` modifier, data groups, represents clauses, etc), that are necessary to express the desired inheritance-related specifications while respecting visibility and modularity. Since C has no inheritance, these intricacies are avoided in ACSL.

Error handling without exceptions

The usual way of signaling errors in Java is through exceptions. Therefore, JML specifications are tailored to express exceptional postconditions, depending on the exception raised. Since C has no exceptions, ACSL does not use exceptional specifications. Instead, C programmers typically signal errors by returning special values, as is mandated in various ways by the C standard.

Example A.1 *In §7.12.1 of the standard, it is said that functions in <math.h> signal errors as follows: "On a domain error, [...] the integer expression errno acquires the value EDOM."*

Example A.2 *In §7.19.5.1 of the standard, it is said that function fclose signals errors as follows: "The fclose function returns [...] EOF if any errors were detected."*

Example A.3 *In §7.19.6.1 of the standard, it is said that function fprintf signals errors as follows: "The fprintf function returns [...] a negative value if an output or encoding error occurred."*

Example A.4 *In §7.20.3 of the standard, it is said that memory management functions signal errors as follows: "If the space cannot be allocated, a null pointer is returned."*

As shown by these few examples, there is no unique way to signal errors in the C standard library, not to mention errors from user-defined functions. But since errors are signaled by returning special values, it is sufficient to write an appropriate postcondition:

```
/*@ ensures \result == error_value || normal_postcondition; */
```

C contracts are not Java ones

In Java, the precondition of the following function that nullifies an array of characters is always true. Even if there was a precondition on the length of array `a`, it could easily be expressed using the Java expression `a.length` that gives the dynamic length of array `a`.

```
public static void Java_nullify(char[] a) {
  if (a == null) return;
  for (int i = 0; i < a.length; ++i) {
    a[i] = 0;
  }
}
```

On the contrary, the precondition of the same function in C, whose definition follows, is more involved. First, remark that the C programmer has to add an extra argument for the size of the array, or rather a lower bound on this array size.

```
void C_nullify(char* a, unsigned int n) {
  int i;
  if (n == 0) return;
  for (i = 0; i < n; ++i) {
    a[i] = 0;
  }
}
```

A correct precondition for this function is the following:

```
/*@ requires \valid(a + 0..(n-1)); */
```

where predicate `\valid` is the one defined in Section 2.7.1. (note that `\valid(a + 0..(-1))` is the same as `\valid(\empty)` and thus is true regardless of the validity of `a` itself). When `n` is 0, `a` does not need to be valid at all, and when `n` is strictly positive, `a` must point to an array of size at least `n`. To make it more obvious, the C programmer adopted a defensive programming style, which returns immediately when `n` is 0. We can duplicate this in the specification:

```
/*@ requires n == 0 || \valid(a + 0..(n-1)); */
```

Many memory requirements are only necessary for some paths through the function, which correspond to some particular behaviors, selected according to some tests performed along the corresponding paths. Since C has no memory primitives, these tests involve other variables that the C programmer adds to track additional information, such as `n` in our example.

To make it easier, it is possible in ACSL to distinguish between the `assumes` part of a behavior, that specifies the tests that need to succeed for this behavior to apply, and the `requires` part that specifies the additional assumptions that must be true when a behavior applies. The specification for our example can then be translated into:

```
/*@ behavior n_is_null:
  @   assumes n == 0;
  @ behavior n_is_not_null:
```

```
  @   assumes n > 0;
  @   requires \valid(a + 0..(n-1));
  @*/
```

This is equivalent to the previous requirement, except here behaviors can be completed with postconditions that belong to one behavior only.

ACSL contracts vs. JML ones

In JML, the set of stated behaviors is assumed to cover all permitted uses of the function; any calling context in which none of the requires preconditions are true would be identified as an error. In ACSL, the set of behaviors for a function do not necessarily cover all cases of use for this function, as mentioned in Section 2.3.4. This allows for partial specifications. In the example above, our two behaviors are clearly mutually exclusive, and, since `n` is an `unsigned int`, they cover all the possible cases. We could have specified that as well, by adding the following lines in the contract (see Section 2.3.4).

```
  @ ...
  @ disjoint behaviors;
  @ complete behaviors;
  @*/
```

To fully understand the difference between specifications in ACSL and JML, we detail below the requirements on the pre-state and the guarantees in the post-state given by behaviors in JML and ACSL.

A JML contract is either *lightweight* or *heavyweight*. For the purpose of our comparison, it is sufficient to know that a lightweight contract is syntactic sugar for a single specific heavyweight contract; a contract can have multiple heavyweight behaviors and these can be nested. Here is a hypothetical JML contract:

```
/*@ behavior x1:
  @    requires A1;
  @    requires R1;
  @    ensures E1;
  @ behavior x2:
  @    requires A2;
  @    requires R2;
  @    ensures E2;
  @*/
```

It assumes that the pre-state satisfies the condition:

```
((A1 && R1) || (A2 && R2))
```

and guarantees that the following condition holds in post-state:

```
  (\old(A1 && R1) ==> E1) && (\old(A2 && R2) ==> E2)
```

Note particularly that the pre-state is required to satisfy the precondition of at least one behavior.

Here is now a syntactically similar ACSL specification:

```
/*@ requires P1;
  @ requires P2;
  @ ensures  Q1;
  @ ensures  Q2;
```

```
@ behavior x_1:
@   assumes A_1;
@   requires R_1;
@   ensures E_1;
@ behavior x_2:
@   assumes A_2;
@   requires R_2;
@   ensures E_2;
@*/
```

Syntactically, the only difference with the JML specification is the addition of the assumes clauses and allowing an anonymous behavior at the beginning of the contract. Rewriting the anonymous behavior with a name gives

```
/*@
@ behavior x_0:
@   assumes \true;
@   requires P_1;
@   requires P_2;
@   ensures Q_1;
@   ensures Q_2;
@ behavior x_1:
@   assumes A_1;
@   requires R_1;
@   ensures E_1;
@ behavior x_2:
@   assumes A_2;
@   requires R_2;
@   ensures E_2;
@*/
```

Its translation to assume-guarantee is however quite different than JML. It assumes the pre-state satisfies the condition

```
(\true ==> (P_1 && P_2)) && (A_1 ==> R_1) && (A_2 ==> R_2)
```

Here, it is acceptable that none of the behaviors are active (that is, that none of the assumes clauses are true, even without the unnamed behavior). In that case there is no post-condition guarantee either.

The contract guarantees that the following condition holds in the post-state:

```
(\true ==> (Q_1 && Q_2)) && (\old(A_1) ==> E_1) && (\old(A_2) ==> E_2)
```

Thus, ACSL allows distinguishing between the clauses that control which behavior is active (the assumes clauses) and the clauses that are preconditions for a particular behavior (the internal requires clauses).

In addition, as mentioned above, there is by default no requirement in ACSL for the specification to be complete. In JML an incomplete specification may cause a warning in a calling context; partial behavior is specified by an explicitly underspecified postcondition. In ACSL, an incomplete specification specifies partial behavior; a warning for a particular behavior is produced by a requires \false; clause.

A.2.2 Deductive verification vs. RAC

Sugar-free behaviors

As explained in detail in [25], JML heavyweight behaviors can be viewed as syntactic sugar that can be translated automatically into more basic contracts consisting mostly of pre- and postconditions and frame conditions. This allows complex nesting of behaviors from the user point of view, while tools only have to deal with basic contracts. In particular, older tools on JML used this desugaring process, such as the Common JML tools to do assertion checking, unit testing, etc. (see [21]) and the tool ESC/Java2 for automatic deductive verification of JML specifications (see [8]).

One issue with such a desugaring approach is the complexity of the transformations involved, as *e.g.* for desugaring assignable clauses between multiple *spec-cases* in JML [25]. Another issue is precisely that tools only see one global contract, instead of multiple independent behaviors, that could be analyzed separately in more detail. Instead, we favor the view that a function implements multiple behaviors, that can be analyzed separately if a tool feels like it. Therefore, we do not intend to provide a desugaring process. Indeed, the current JML tool, OpenJML [6, 7], also does only a partial desugaring, which at minimum is able to give more informative error messages when proof attempts fail.

Axiomatized functions in specifications

JML allows pure Java methods to be called in specifications [19]. This avoids having to write essentially duplicate logical functions that mimic Java functions. It is also useful when relying on RAC: methods called should be defined so that the runtime can call them, and they should not have side-effects in order not to pollute the program they are supposed to annotate. JML also permits model (logical) functions to be used in specifications; if the model function does not have a body, then RAC cannot be used. But for deductive verification, the properties of a model function can be specified axiomatically.

ACSL focuses on deductive verification and currently only allows calls to logical functions in specifications. These functions may be defined, like program functions, but they may also be only declared (with a suitable declaration of `reads` clause) and their behavior defined through an axiomatization. This makes for richer specifications that may be useful either in automatic or in manual deductive verification.

A.2.3 Syntactic differences

The following table summarizes the difference between JML and ACSL keywords, when the intent is the same, although minor differences might exist.

JML	ACSL
modifiable, assignable	assigns
measured_by	decreases
loop_invariant	loop invariant
decreases	loop variant
(\forall τ x ; P ; Q)	(\forall τ x ; P ==> Q)
(\exists τ x ; P ; Q)	(\exists τ x ; P && Q)
\max τ x ; a <= x <= b ; f)	\max(a,b,\lambda τ x ; f)

A.3 Typing rules

Disclaimer: this section is unfinished, it is left here just to give an idea of what it will look like in the final document.

A.3.1 Rules for terms

Integer promotion:

$$\frac{\Gamma \vdash e : \tau}{\Gamma \vdash e : \text{integer}}$$

if τ is any C integer type `char`, `short`, `int`, or `long`, whatever attribute they have, in particular signed or unsigned

Variables:

$$\frac{}{\Gamma \vdash id : \tau} \text{ if } id : \tau \in \Gamma$$

Unary integer operations:

$$\frac{\Gamma \vdash t : \text{integer}}{\Gamma \vdash op\ t : \text{integer}} \text{ if } op \in \{+, -, \sim\}$$

Boolean negation:

$$\frac{\Gamma \vdash t : \text{boolean}}{\Gamma \vdash !\ t : \text{boolean}}$$

Pointer dereferencing:

$$\frac{\Gamma \vdash t : \tau *}{\Gamma \vdash *t : \tau}$$

Address operator:

$$\frac{\Gamma \vdash t : \tau}{\Gamma \vdash \&t : \tau *}$$

Binary

$$\frac{\Gamma \vdash t_1 : \text{integer} \qquad \Gamma \vdash t_2 : \text{integer}}{\Gamma \vdash t_1\ op\ t_2 : \text{integer}} \text{ if } op \in \{+, -, *, /, \%\}$$

$$\frac{\Gamma \vdash t_1 : \text{real} \qquad \Gamma \vdash t_2 : \text{real}}{\Gamma \vdash t_1\ op\ t_2 : \text{real}} \text{ if } op \in \{+, -, *, /\}$$

$$\frac{\Gamma \vdash t_1 : \text{integer} \qquad \Gamma \vdash t_2 : \text{integer}}{\Gamma \vdash t_1\ op\ t_2 : \text{boolean}} \text{ if } op \in \{==, !=, <=, <, >=, >\}$$

$$\frac{\Gamma \vdash t_1 : \text{real} \qquad \Gamma \vdash t_2 : \text{real}}{\Gamma \vdash t_1\ op\ t_2 : \text{boolean}} \text{ if } op \in \{==, !=, <=, <, >=, >\}$$

$$\frac{\Gamma \vdash t_1 : \tau * \qquad \Gamma \vdash t_2 : \tau *}{\Gamma \vdash t_1\ op\ t_2 : \text{boolean}} \text{ if } op \in \{==, !=, <=, <, >=, >\}$$

(to be continued)

A.3.2 Typing rules for sets

We consider the typing judgement $\Gamma, \Lambda \vdash s : \tau, b$ meaning that s is a set of terms of type τ, which is moreover a set of locations if the boolean b is true. Γ is the C environment and Λ is the logic environment.

Rules:

$$\frac{}{\Gamma, \Lambda \vdash id : \tau, true} \text{ if } id : \tau \in \Gamma$$

$$\frac{}{\Gamma, \Lambda \vdash id : \tau, true} \text{ if } id : \tau \in \Lambda$$

$$\frac{\Gamma, \Lambda \vdash s : \tau*, b}{\Gamma, \Lambda \vdash *s : \tau, true}$$

$$\frac{id : \tau \quad s : set < struct\ S* >}{\vdash s- > id : set < \tau >}$$

$$\frac{\Gamma, b \cup \Lambda \vdash e : tset\tau}{\Gamma, \Lambda \vdash \{e \mid b; P\} : tset\tau}$$

$$\frac{\Gamma, \Lambda \vdash e_1 : \tau, b \quad \Gamma, \Lambda \vdash e_2 : \tau, b}{\Gamma, \Lambda \vdash e_1, e_2 : \tau, b}$$

A.4 Specification Templates

This section describes some common issues that may occur when writing an ACSL specification and proposes some solution to overcome them

A.4.1 Accessing a C variable that is masked

The situation may happen where it is necessary to refer in an annotation to a C variable that is masked at that point. For instance, a function contract may need to refer to a global variable that has the same name as a function parameter, as in the following code:

```
int x;
//@ assigns x;
int g();

int f(int x) {
  // ...
  return g();
}
```

In order to write the `assigns` clause for `f`, we must access the global variable `x`, since `f` calls `g`, which can modify `x`. This is not possible with C scoping rules, as `x` refers to the parameter of `f` in the scope of the function.

A solution is to use a ghost pointer to `x`, as shown in the following code:

```
int x;

//@ ghost int* const ghost_ptr_x = &x;

//@ assigns x;
int g();
```

```

//@ assigns *ghost_ptr_x;
int f(int x) {
  // ...
  return g();
}
```

A.5 Illustrative example

This is an attempt to define an example for ACSL, much as the Purse example in JML description papers. It is a memory allocator, whose main functions are `memory_alloc` and `memory_free`, to respectively allocate and deallocate memory. The goal is to exercise as much as possible of ACSL.

```

#include <stdlib.h>

#define DEFAULT_BLOCK_SIZE 1000

typedef enum _bool { false = 0, true = 1 } bool;

/*@ predicate finite_list<A>((A* -> A*) next_elem, A* ptr) =
  @   ptr == \null ||
  @   (\valid(ptr) && finite_list(next_elem,next_elem(ptr))) ;
  @
  @ logic integer list_length<A>((A* -> A*) next_elem, A* ptr) =
  @   (ptr == \null) ? 0 :
  @   1 + list_length(next_elem,next_elem(ptr)) ;
  @
  @
  @ predicate lower_length<A>((A* -> A*) next_elem,
  @                           A* ptr1, A* ptr2) =
  @   finite_list(next_elem, ptr1) && finite_list(next_elem, ptr2)
  @   && list_length(next_elem, ptr1) < list_length(next_elem, ptr2) ;
  @*/

// forward reference
struct _memory_slice;

/* A memory block holds a pointer to a raw block of memory allocated by
 * calling [malloc]. It is sliced into chunks, which are maintained by
 * the [slice] structure. It maintains additional information such as
 * the [size] of the memory block, the number of bytes [used] and the [next]
 * index at which to put a chunk.
 */
typedef struct _memory_block {
  //@ ghost boolean      packed;
    // ghost field [packed] is meant to be used as a guard that tells when
    // the invariant of a structure of type [memory_block] holds
  unsigned int         size;
    // size of the array [data]
  unsigned int         next;
    // next index in [data] at which to put a chunk
  unsigned int         used;
```

```
    // how many bytes are used in [data], not necessarily contiguous ones
  char*                data;
    // raw memory block allocated by [malloc]
  struct _memory_slice* slice;
    // structure that describes the slicing of a block into chunks
} memory_block;

/*@ strong type invariant inv_memory_block(memory_block mb) =
  @   mb.packed ==>
  @     (0 < mb.size && mb.used <= mb.next <= mb.size
  @     && \offset(mb.data) == 0
  @     && \block_length(mb.data) == mb.size) ;
  @
  @ predicate valid_memory_block(memory_block* mb) =
  @   \valid(mb) && mb->packed ;
  @*/

/* A memory chunk holds a pointer [data] to some part of a memory block
 * [block]. It maintains the [offset] at which it points in the block, as well
 * as the [size] of the block it is allowed to access. A field [free] tells
 * whether the chunk is used or not.
 */
typedef struct _memory_chunk {
  //@ ghost boolean packed;
    // ghost field [packed] is meant to be used as a guard that tells when
    // the invariant of a structure of type [memory_chunk] holds
  unsigned int    offset;
    // offset at which [data] points into [block->data]
  unsigned int    size;
    // size of the chunk
  bool            free;
    // true if the chunk is not used, false otherwise
  memory_block* block;
    // block of memory into which the chunk points
  char*           data;
    // shortcut for [block->data + offset]
} memory_chunk;

/*@ strong type invariant inv_memory_chunk(memory_chunk mc) =
  @   mc.packed ==>
  @     (0 < mc.size && valid_memory_block(mc.block)
  @     && mc.offset + mc.size <= mc.block->next) ;
  @
  @ predicate valid_memory_chunk(memory_chunk* mc, int s) =
  @   \valid(mc) && mc->packed && mc->size == s ;
  @
  @ predicate used_memory_chunk(memory_chunk mc) =
  @   mc.free == false ;
  @
  @ predicate freed_memory_chunk(memory_chunk mc) =
  @   mc.free == true ;
  @*/

/* A memory chunk list links memory chunks in the same memory block.
 * Newly allocated chunks are put first, so that the offset of chunks
 * decreases when following the [next] pointer. Allocated chunks should
```

```
 * fill the memory block up to its own [next] index.
 */
typedef struct _memory_chunk_list {
  memory_chunk*           chunk;
    // current list element
   struct _memory_chunk_list* next;
    // tail of the list
} memory_chunk_list;

/*@ logic memory_chunk_list* next_chunk(memory_chunk_list* ptr) =
  @   ptr->next ;
  @
  @ predicate valid_memory_chunk_list
  @                 (memory_chunk_list* mcl, memory_block* mb) =
  @    \valid(mcl) && valid_memory_chunk(mcl->chunk,mcl->chunk->size)
  @   && mcl->chunk->block == mb
  @   && (mcl->next == \null ||
  @       valid_memory_chunk_list(mcl->next, mb))
  @   && mcl->offset == mcl->chunk->offset
  @   && (
  @        // it is the last chunk in the list
  @        (mcl->next == \null && mcl->chunk->offset == 0)
  @      ||
  @        // it is a chunk in the middle of the list
  @        (mcl->next != \null
  @        && mcl->next->chunk->offset + mcl->next->chunk->size
  @           == mcl->chunk->offset)
  @      )
  @   && finite_list(next_chunk, mcl) ;
  @
  @ predicate valid_complete_chunk_list
  @                 (memory_chunk_list* mcl, memory_block* mb) =
  @   valid_memory_chunk_list(mcl,mb)
  @   && mcl->next->chunk->offset +
  @      mcl->next->chunk->size == mb->next ;
  @
  @ predicate chunk_lower_length(memory_chunk_list* ptr1,
  @                              memory_chunk_list* ptr2) =
  @   lower_length(next_chunk, ptr1, ptr2) ;
  @*/

/* A memory slice holds together a memory block [block] and a list of chunks
 * [chunks] on this memory block.
 */
typedef struct _memory_slice {
  //@ ghost boolean    packed;
    // ghost field [packed] is meant to be used as a guard that tells when
    // the invariant of a structure of type [memory_slice] holds
  memory_block*    block;
  memory_chunk_list* chunks;
} memory_slice;

/*@ strong type invariant inv_memory_slice(memory_slice* ms) =
  @   ms.packed ==>
  @     (valid_memory_block(ms->block) && ms->block->slice == ms
  @     && (ms->chunks == \null
```

```
  @        || valid_complete_chunk_list(ms->chunks, ms->block))) ;
  @
  @ predicate valid_memory_slice(memory_slice* ms) =
  @   \valid(ms) && ms->packed ;
  @*/

/* A memory slice list links memory slices, to form a memory pool.
 */
typedef struct _memory_slice_list {
  //@ ghost boolean    packed;
    // ghost field [packed] is meant to be used as a guard that tells when
    // the invariant of a structure of type [memory_slice_list] holds
  memory_slice*           slice;
    // current list element
  struct _memory_slice_list* next;
    // tail of the list
} memory_slice_list;

/*@ logic memory_slice_list* next_slice(memory_slice_list* ptr) =
  @   ptr->next ;
  @
  @ strong type invariant inv_memory_slice_list(memory_slice_list* msl) =
  @   msl.packed ==>
  @     (valid_memory_slice(msl->slice)
  @     && (msl->next == \null ||
  @          valid_memory_slice_list(msl->next))
  @     && finite_list(next_slice, msl)) ;
  @
  @ predicate valid_memory_slice_list(memory_slice_list* msl) =
  @   \valid(msl) && msl->packed ;
  @
  @ predicate slice_lower_length(memory_slice_list* ptr1,
  @                              memory_slice_list* ptr2) =
  @     lower_length(next_slice, ptr1, ptr2)
  @ } */

typedef memory_slice_list* memory_pool;

/*@ type invariant valid_memory_pool(memory_pool *mp) =
  @   \valid(mp) && valid_memory_slice_list(*mp) ;
  @*/

/*@ behavior zero_size:
  @   assumes s == 0;
  @   assigns \nothing;
  @   ensures \result == 0;
  @
  @ behavior positive_size:
  @   assumes s > 0;
  @   requires valid_memory_pool(arena);
  @   ensures \result == 0
  @     || (valid_memory_chunk(\result,s) &&
  @          used_memory_chunk(*\result));
  @ */
memory_chunk* memory_alloc(memory_pool* arena, unsigned int s) {
  memory_slice_list *msl = *arena;
```

```
memory_chunk_list *mcl;
memory_slice *ms;
memory_block *mb;
memory_chunk *mc;
unsigned int mb_size;
//@ ghost unsigned int mcl_offset;
char *mb_data;
// guard condition
if (s == 0) return 0;
// iterate through memory blocks (or slices)
/*@
  @ loop invariant valid_memory_slice_list(msl);
  @ loop variant msl for slice_lower_length;
  @ */
while (msl != 0) {
  ms = msl->slice;
  mb = ms->block;
  mcl = ms->chunks;
  // does [mb] contain enough free space?
  if (s <= mb->size - mb->next) {
    //@ ghost ms->ghost = false; // unpack the slice
    // allocate a new chunk
    mc = (memory_chunk*)malloc(sizeof(memory_chunk));
    if (mc == 0) return 0;
    mc->offset = mb->next;
    mc->size = s;
    mc->free = false;
    mc->block = mb;
    //@ ghost mc->ghost = true; // pack the chunk
    // update block accordingly
    //@ ghost mb->ghost = false; // unpack the block
    mb->next += s;
    mb->used += s;
    //@ ghost mb->ghost = true; // pack the block
    // add the new chunk to the list
    mcl = (memory_chunk_list*)malloc(sizeof(memory_chunk_list));
    if (mcl == 0) return 0;
    mcl->chunk = mc;
    mcl->next = ms->chunks;
    ms->chunks = mcl;
    //@ ghost ms->ghost = true; // pack the slice
    return mc;
  }
  // iterate through memory chunks
  /*@
    @ loop invariant valid_memory_chunk_list(mcl,mb);
    @ loop variant mcl for chunk_lower_length;
    @ */
  while (mcl != 0) {
    mc = mcl->chunk;
    // is [mc] free and large enough?
    if (mc->free && s <= mc->size) {
      mc->free = false;
      mb->used += mc->size;
      return mc;
    }
```

```
      // try next chunk
      mcl = mcl->next;
    }
    msl = msl->next;
  }
  // allocate a new block
  mb_size = (DEFAULT_BLOCK_SIZE < s) ? s : DEFAULT_BLOCK_SIZE;
  mb_data = (char*)malloc(mb_size);
  if (mb_data == 0) return 0;
  mb = (memory_block*)malloc(sizeof(memory_block));
  if (mb == 0) return 0;
  mb->size = mb_size;
  mb->next = s;
  mb->used = s;
  mb->data = mb_data;
  //@ ghost mb->ghost = true;  // pack the block
  // allocate a new chunk
  mc = (memory_chunk*)malloc(sizeof(memory_chunk));
  if (mc == 0) return 0;
  mc->offset = 0;
  mc->size = s;
  mc->free = false;
  mc->block = mb;
  //@ ghost mc->ghost = true;  // pack the chunk
  // allocate a new chunk list
  mcl = (memory_chunk_list*)malloc(sizeof(memory_chunk_list));
  if (mcl == 0) return 0;
  //@ ghost mcl->offset = 0;
  mcl->chunk = mc;
  mcl->next = 0;
  // allocate a new slice
  ms = (memory_slice*)malloc(sizeof(memory_slice));
  if (ms == 0) return 0;
  ms->block = mb;
  ms->chunks = mcl;
  //@ ghost ms->ghost = true;  // pack the slice
  // update the block accordingly
  mb->slice = ms;
  // add the new slice to the list
  msl = (memory_slice_list*)malloc(sizeof(memory_slice_list));
  if (msl == 0) return 0;
  msl->slice = ms;
  msl->next = *arena;
  //@ ghost msl->ghost = true;  // pack the slice list
  *arena = msl;
  return mc;
}

/*@ behavior null_chunk:
  @   assumes chunk == \null;
  @   assigns \nothing;
  @
  @ behavior valid_chunk:
  @   assumes chunk != \null;
  @   requires valid_memory_pool(arena);
  @   requires valid_memory_chunk(chunk,chunk->size);
```

```
  @   requires  used_memory_chunk(chunk);
  @   ensures
  @      // if it is not the last chunk in the block, mark it as free
  @      (valid_memory_chunk(chunk,chunk->size)
  @      && freed_memory_chunk(chunk))
  @     ||
  @      // if it is the last chunk in the block, deallocate the block
  @      !  \valid(chunk);
  @ */
void memory_free(memory_pool* arena, memory_chunk* chunk) {
  memory_slice_list *msl = *arena;
  memory_block *mb = chunk->block;
  memory_slice *ms = mb->slice;
  memory_chunk_list *mcl;
  memory_chunk *mc;
  // is it the last chunk in use in the block?
  if (mb->used == chunk->size) {
    // remove the corresponding slice from the memory pool
    // case it is the first slice
    if (msl->slice == ms) {
      *arena = msl->next;
      //@ ghost msl->ghost = false;  // unpack the slice list
      free(msl);
    }
    // case it is not the first slice
    while (msl != 0) {
      if (msl->next != 0 && msl->next->slice == ms) {
        memory_slice_list* msl_next = msl->next;
        msl->next = msl->next->next;
        // unpack the slice list
        //@ ghost msl_next->ghost = false;
        free(msl_next);
        break;
      }
      msl = msl->next;
    }
    //@ ghost ms->ghost = false;  // unpack the slice
    // deallocate all chunks in the block
    mcl = ms->chunks;
    // iterate through memory chunks
    /*@
      @ loop  invariant  valid_memory_chunk_list(mcl,mb);
      @ loop  variant  mcl  for  chunk_lower_length;
      @ */
    while (mcl != 0) {
      memory_chunk_list *mcl_next = mcl->next;
      mc = mcl->chunk;
      //@ ghost mc->ghost = false;  // unpack the chunk
      free(mc);
      free(mcl);
      mcl = mcl_next;
    }
    mb->next = 0;
    mb->used = 0;
    // deallocate the memory block and its data
    //@ ghost mb->ghost = false;   // unpack the block
```

```
    free(mb->data);
    free(mb);
    // deallocate the corresponding slice
    free(ms);
    return;
  }
  // mark the chunk as freed
  chunk->free = true;
  // update the block accordingly
  mb->used -= chunk->size;
  return;
}
```

A.6 Changes

A.6.1 Version 1.14

- Introduce check annotation (section 2.4.1)

A.6.2 Version 1.13

- New infix predicate \in for set membership (section 2.3.5)
- Fixes some typing error for constructs rejecting void* pointers (section 2.7.3)
- Notations added for real numbers π and e (section 2.2.5)

A.6.3 Version 1.12

- Fixes syntax rule for statement contracts in allowing completeness clauses (figure 2.12)

A.6.4 Version 1.11

- Functions related to infinites and the sign of floating-point value (section 2.2.5)
- New section for predicates related to well-typedness (section 2.15)
- Syntax for defining a set by giving explicitely its elements (section 2.3.5)
- Adding lists as first-class values (section 2.8.2)
- Change the associativity of bitwise operator --> to right, in accordance with the one of ==> operator
- Glyph used for ^^ operator (xor) fixed

A.6.5 Version 1.10

- Change keyword for importing libraries (section 2.6.11)
- Fix numerous typos reported by David Cok
- Disallow meaningless assigns \nothing \from x (section 2.10)

A.6.6 Version 1.9

- Fix typo in definition of \fresh predicate (section 2.7.3)
- Fix grammar inconsistencies
  - use proper C rules names
  - fix mismatch in non-terminal names
- Rename "Unspecified values" to "Dangling pointers" and precise it (section 2.14)

A.6.7 Version 1.8

- Mention binary literal constant typing

A.6.8 Version 1.7

- Added missing shift operators in figure 2.1
- Modified syntax for naming terms and predicates (figures 2.2 and 2.1)
- Added syntax rule for literal constants (figure 2.1)

A.6.9 Version 1.6

- Modified syntax for model fields (section 2.11.2)
- Added missing logical xor operator (figure 2.1).
- Addition of logical labels related to loops (section 2.4.3).
- Addition of labels to built-ins related to memory blocks (section 2.7.1)
- Introduction of \valid_read built-in and clarification of the notion of validity (section 2.7.1).
- Introduction of built-in \allocable , \allocation , \freeable and \fresh (section 2.7.3).
- Introduction of allocates and frees clauses (section 2.7.3).
- Clarify the semantics of assigns clauses into statement contract.
- Improvements to the volatile clause (section 2.12.1).
- Clarify the evaluation of arrays inside an at (section 2.4.3).

A.6.10 Version 1.5

- Clarify the status of loop invariant in presence of break or side-effects in the loop test.
- Introduction of \with keyword for functional updates.
- Added bnf entry for completeness of function behaviors.
- Order of clauses in statement contracts is now fixed.
- requires clauses are allowed before behaviors of statement contracts.
- Added explicit singleton construct for sets.
- Introduction of logical arrays.
- Operations over pointers and arrays have been precised.
- Predicate \initialized (section 2.13) now takes a set of pointers as argument.

A.6.11 Version 1.4

- Added UTF-8 counterparts for built-in types (integer , real , boolean).
- Fixed typos in the examples corresponding to features implemented in Frama-C.
- Order of clauses in function contracts is now fixed.
- Introduction of abrupt termination clauses.
- Introduction of axiomatic to gather predicates, logic functions, and their defining axioms.
- Added specification templates appendix for common specification issues.
- Use of sets as first-class term has been precised.
- Fixed semantics of predicate \separated .

A.6.12 Version 1.3

- Functional update of structures.
- Terminates clause in function behaviors.
- Typos reported by David Mentré.

A.6.13 Version 1.2

This is the first public release of this document.

Bibliography

[1] Jean-Raymond Abrial. *The B-Book: Assigning Programs to Meanings*. Cambridge University Press, 1996.

[2] Ali Ayad and Claude Marché. Behavioral properties of floating-point programs. Hisseo publications, 2009. http://hisseo.saclay.inria.fr/ayad09.pdf.

[3] Sylvie Boldo and Jean-Christophe Filliâtre. Formal Verification of Floating-Point Programs. In *18th IEEE International Symposium on Computer Arithmetic*, pages 187–194, Montpellier, France, June 2007.

[4] Patrice Chalin. Reassessing JML's logical foundation. In *Proceedings of the 7th Workshop on Formal Techniques for Java-like Programs (FTfJP'05)*, Glasgow, Scotland, July 2005.

[5] Patrice Chalin. A sound assertion semantics for the dependable systems evolution verifying compiler. In *Proceedings of the International Conference on Software Engineering (ICSE'07)*, pages 23–33, Los Alamitos, CA, USA, 2007. IEEE Computer Society.

[6] David R. Cok. OpenJML: JML for Java 7 by Extending OpenJDK. In Mihaela Bobaru, Klaus Havelund, Gerard Holzmann, and Rajeev Joshi, editors, *NASA Formal Methods*, volume 6617 of *Lecture Notes in Computer Science*, pages 472–479. Springer Berlin / Heidelberg, 2011. 10.1007/978-3-642-20398-5_35.

[7] David R. Cok. OpenJML: Software verification for Java 7 using JML, OpenJDK, and Eclipse. In *F-IDE*, pages 79–92, 2014.

[8] David R. Cok and Joseph R. Kiniry. ESC/Java2: Uniting ESC/Java and JML: Progress and issues in building and using ESC/Java2, including a case study involving the use of the tool to verify portions of an Internet voting tally system. In Gilles Barthe, Lilian Burdy, Marieke Huisman, Jean-Louis Lanet, and Traian Muntean, editors, *Construction and Analysis of Safe, Secure, and Interoperable Smart Devices (CASSIS 2004)*, volume 3362 of *Lecture Notes in Computer Science*, pages 108–128. Springer, 2005.

[9] Jeremy Condit, Matthew Harren, Scott McPeak, George C. Necula, and Westley Weimer. Ccured in the real world. In *PLDI '03: Proceedings of the ACM SIGPLAN 2003 conference on Programming language design and implementation*, pages 232–244, 2003.

[10] Jeremy Paul Condit, Matthew Thomas Harren, Zachary Ryan Anderson, David Gay, and George Necula. Dependent types for low-level programming. In *ESOP '07: Proceedings of the 16th European Symposium on Programming*, October 2006.

[11] Jean-Christophe Filliâtre and Claude Marché. Multi-prover verification of C programs. In Jim Davies, Wolfram Schulte, and Mike Barnett, editors, *6th International Conference*

on Formal Engineering Methods, volume 3308 of *Lecture Notes in Computer Science*, pages 15–29, Seattle, WA, USA, November 2004. Springer.

[12] Jean-Christophe Filliâtre and Claude Marché. The Why/Krakatoa/Caduceus platform for deductive program verification. In Werner Damm and Holger Hermanns, editors, *19th International Conference on Computer Aided Verification*, Lecture Notes in Computer Science, Berlin, Germany, July 2007. Springer.

[13] The Frama-C framework for analysis of C code. http://frama-c.cea.fr/.

[14] A. Giorgetti and J. Groslambert. JAG: JML Annotation Generation for verifying temporal properties. In *FASE'2006, Fundamental Approaches to Software Engineering*, volume 3922 of *LNCS*, pages 373–376, Vienna, Austria, March 2006. Springer.

[15] John Hatcliff, Gary T. Leavens, K Rustan M Leino, Peter Müller, and Matthew Parkinson. Behavioral interface specification languages. *ACM Computing Surveys (CSUR)*, 44(3):16, 2012.

[16] International Organization for Standardization (ISO). *The ANSI C standard (C99)*. http://www.open-std.org/JTC1/SC22/WG14/www/docs/n1124.pdf.

[17] Brian Kernighan and Dennis Ritchie. *The C Programming Language (2nd Ed.)*. Prentice-Hall, 1988.

[18] Gary Leavens. Jml. http://www.eecs.ucf.edu/~leavens/JML/.

[19] Gary T. Leavens, Albert L. Baker, and Clyde Ruby. Preliminary design of JML: A behavioral interface specification language for Java. Technical Report 98-06i, Iowa State University, 2000.

[20] Gary T. Leavens, K. Rustan M. Leino, and Peter Müller. Specification and verification challenges for sequential object-oriented programs. *Form. Asp. Comput.*, 19(2):159–189, 2007.

[21] Gary T. Leavens, K. Rustan M. Leino, Erik Poll, Clyde Ruby, and Bart Jacobs. JML: notations and tools supporting detailed design in Java. In *OOPSLA 2000 Companion, Minneapolis, Minnesota*, pages 105–106, 2000.

[22] Claude Marché. Towards modular algebraic specifications for pointer programs: a case study. In H. Comon-Lundh, C. Kirchner, and H. Kirchner, editors, *Rewriting, Computation and Proof*, volume 4600 of *Lecture Notes in Computer Science*, pages 235–258. Springer-Verlag, 2007.

[23] Yannick Moy. Union and cast in deductive verification. Technical Report ICIS-R07015, Radboud University Nijmegen, July 2007. http://www.lri.fr/~moy/union_and_cast/union_and_cast.pdf.

[24] George C. Necula, Scott McPeak, and Westley Weimer. CCured: Type-safe retrofitting of legacy code. In *Symposium on Principles of Programming Languages*, pages 128–139, 2002.

[25] Arun D. Raghavan and Gary T. Leavens. Desugaring JML method specifications. Technical Report 00-03a, Iowa State University, 2000.

[26] David Stevenson et al. An american national standard: IEEE standard for binary floating point arithmetic. *ACM SIGPLAN Notices*, 22(2):9–25, 1987.

[27] Wikipedia. First order logic. http://en.wikipedia.org/wiki/First_order_logic.

[28] Wikipedia. IEEE 754. http://en.wikipedia.org/wiki/IEEE_754-1985.

List of Figures

2.1 Grammar of terms. The terminals *id*, *C-type-name*, and various literals are the same as the corresponding C lexical tokens. 17
2.2 Grammar of predicates 18
2.3 Grammar of binders and type expressions 19
2.4 Operator precedence 19
2.5 Grammar of function contracts 30
2.6 \old and \result in terms 30
2.7 Grammar for sets of memory locations 36
2.8 Grammar for assertions 38
2.9 Grammar for loop annotations 39
2.10 Grammar for general inductive invariants 42
2.11 Grammar for at construct 45
2.12 Grammar for statement contracts 47
2.13 Grammar for global logic definitions 51
2.14 Grammar for inductive definitions 52
2.15 Grammar for axiomatic declarations 53
2.16 Grammar for higher-order constructs 55
2.17 Grammar for concrete logic types and pattern-matching 57
2.18 Grammar for logic declarations with labels 58
2.19 Grammar for logic declarations with reads clauses 59
2.20 Grammar extension of terms and predicates about memory 61
2.21 Grammar for dynamic allocations and deallocations 62
2.22 Notations for built-in list datatype 68
2.23 Grammar of contracts about abrupt terminations 69
2.24 Grammar for dependencies information 72
2.25 Grammar for declarations of data invariants 73
2.26 Grammar for declarations of model variables and fields 76
2.27 Grammar for ghost statements 78
2.28 Grammar for volatile constructs 81

2.29 Grammar extensions regarding initialized and dangling memory 82
2.30 Grammar for predicates related to well-typedness 83

Index

?, 17, 18
_, 57

abrupt clause, 69
\allocable, 61, 63
allocates, 62, 71
allocation, 62
\allocation, 61, 63
allocation_status, 63
annotation, 38
    loop, 38
as, 57
assert, 38
assertion, 38
assigns, 30, 31, 39, 46, 71, 72
assumes, 30
\at, 44, 45
\automatic, 63
axiom, 53
axiomatic, 53

\base_addr, 61
behavior, 30, 32, 47, 91
behaviors, 30
\block_length, 29, 61
boolean, 19, 20
breaks, 69, 70

case, 52, 57
cast, 21–24
check, 38
complete, 30
complete behaviors, 34
comprehension, 36
\concat, 68
\Cons, 68
continues, 69, 70
contract, 29, 38, 46, 91

\dangling, 82
data invariant, 72
deallocation, 62
decreases, 30, 47
dependency, 71
disjoint, 30
disjoint behaviors, 35
do, 39
\dynamic, 63
dynamic allocation, 62

else, 78
\empty, 35, 36
ensures, 30, 31, 46, 69
\eq_double, 24
\eq_float, 24
\exists, 18
\exit_status, 69, 70
exits, 69

\false, 17, 18, 20
for, 30, 38, 39, 42, 47
\forall, 18
formal parameter, 30, 31, 74
\freeable, 61, 64
frees, 62, 71
\fresh, 61, 64
\from, 72
function behavior, 32, 91
function contract, 29, 91
functional expression, 71

\ge_double, 24
\ge_float, 24
ghost, 77, 78
global, 73
global invariant, 72
grammar entries
    C-compound-statement, 38
    C-external-declaration, 51
    C-statement, 38
    abrupt-clause-stmt, 69
    abrupt-clause, 69
    allocation-clause, 62
    assertion-kind, 38

assertion, 38, 42
assigns-clause, 30, 72
assumes-clause, 30
axiom-def, 53
axiomatic-decl, 53
behavior-body-stmt, 47
behavior-body, 30
bin-op, 17
binders, 19
binder, 19
breaks-clause, 69
built-in-logic-type, 19
completeness-clause, 30
constructor, 57
continues-clause, 69
data-inv-def, 73
data-invariant, 73
declaration, 76, 78, 81
decreases-clause, 30
direct-declarator, 78
dyn-allocation-addresses, 62
ensures-clause, 30
exits-clause, 69
ext-quantifier, 55
function-contract, 30
ghost-selection-statement, 78
ghost-type-specifier, 78
ident, 17, 58
indcase, 52
inductive-def, 52
inv-strength, 73
label-binders, 58
label-id, 45
lemma-def, 51
literal, 17
location-addresses, 61
location-address, 61, 83
locations, 30
location, 30
logic-const-decl, 53
logic-const-def, 51
logic-decl, 53
logic-def, 51–53, 57
logic-function-decl, 53, 59
logic-function-def, 51, 59
logic-predicate-decl, 53, 59
logic-predicate-def, 51, 59
logic-type-decl, 53
logic-type-def, 57
logic-type-expr, 19
logic-type, 53
loop-allocation, 62
loop-annot, 39
loop-assigns, 39
loop-behavior, 39
loop-clause, 39
loop-invariant, 39
loop-variant, 39
match-cases, 57
match-case, 57
named-behavior-stmt, 47
named-behavior, 30
one-label, 61
parameters, 51
parameter, 51
pat, 57
poly-id, 17, 51
postfix-expression, 78
pred, 18, 30, 36, 45, 61, 82, 83
product-type, 57
range, 36
reads-clause, 59
record-type, 57
rel-op, 18
requires-clause, 30
returns-clause, 69
simple-clause-stmt, 47
simple-clause, 30
statement-contract, 47
statements-ghost, 78
statement, 39, 47, 78
struct-declaration, 78
sum-type, 57
terminates-clause, 30
term, 17, 30, 45, 55, 57, 61, 68, 69
tset, 36
two-labels, 61
type-expr, 19, 51, 57
type-invariant, 73
type-var-binders, 51
type-var, 51
unary-op, 17
variable-ident, 19
\gt_double, 24
\gt_float, 24

`Here`, 44, 45, 70
hybrid
function, 56

predicate, 56

if , 78
\in, 36
inductive , 52
inductive definitions, 52
inductive predicates, 52
Init, 44, 45
\initialized , 82
integer , 19, 20
\inter , 36
invariant, 41
data, 72
global, 72
loop, 39
strong, 72
type, 72
weak, 72
invariant , 39, 42, 73
\is_finite , 24
\is_infinite , 24
\is_minus_infinity , 24
\is_NaN, 24
\is_plus_infinity , 24

\lambda, 55
\le_double, 24
\le_float , 24
left-value, 91
lemma, 51
\length, 68
\let , 17, 18, 57
library, 60
\list , 68
location, 35, 67
logic , 51, 53, 59
logic specification, 50
loop
allocation, 65
annotation, 38
assigns, 39
behavior, 40
deallocation, 65
invariant, 39
variant, 41
loop, 39, 62
LoopCurrent, 44, 45
LoopEntry, 44, 45
\lt_double, 24
\lt_float , 24

lvalue, 91

\match, 57
\max, 55
\min, 55
model, 75, 76
module, 59

\ne_double, 24
\ne_float, 24
\Nil, 68
\nothing, 30
\nth, 68
\null , 61, 62
\numof, 55

\offset , 61, 62
Old, 44, 45, 70
\old, 30, 44, 70

polymorphism, 54
Post, 44, 45, 70
post-state, 91
Pre, 44, 45, 70
pre-state, 91
predicate, 16
predicate , 51, 53, 59
\product, 55
pure expression, 91

reads, 59, 81
real , 19, 20
real_of_double, 23
real_of_float, 23
recursion, 54
\register , 63
\repeat, 68
requires , 30, 31
\result , 30, 70
returns , 69, 70
\round_double, 24
\round_float, 24

\separated, 61, 62
set type, 67
\sign, 25
sizeof , 17, 22
specification, 50
statement contract, 38, 46
\static , 63
strong, 73

\subset, 36
\sum, 55

term, 16
terminates, 30, 50
termination, 41, 47
\true, 17, 18, 20
type
    concrete, 56
    polymorphic, 54
    record, 56, 57
    sum, 56, 57
type, 53, 57, 73
type invariant, 72

\unallocated, 63
\union, 36

\valid, 61
valid_function, 83
\valid_function, 83
\valid_read, 61
variant, 39, 41, 47
volatile, 80, 81

weak, 73
while, 39
\with, 17, 55
writes, 81