An algebraic specification for the static semantics of Pascal

A. van Deursen

Computer Science/Department of Software Technology

CS-R9129 1991
An Algebraic Specification
for the Static Semantics of Pascal

A. van Deursen

CWI, Department of Software Technology,
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands, email: arie@cwi.nl

Abstract

Over the last few years, several formal specifications for the static semantics of Pascal have been given. Thus far, however, no algebraic specification has been published. In this document we discuss an algebraic specification we have made for the complete static semantics of Pascal as defined by the International Standardization Organization (ISO). We explain how the specification has been set up, and how several details have been dealt with in a convenient way. Finally, we relate the specification to algebraic specifications for other aspects of (other) programming languages, and we briefly compare the algebraic approach with the other formal specifications for Pascal.

Key Words & Phrases: Formal specifications, algebraic specifications, Pascal, programming language semantics, static semantics, type checking.


Note: Partial support has been received from the European Communities under ESPRIT project 2177 (Generation of Interactive Programming Environments II - GIPE II) and from the Netherlands Organization for Scientific Research – NWO, project Incremental Program Generators.
Contents

1 Introduction 3

2 Algebraic Specifications 3
   2.1 The Signature 4
   2.2 Terms 4
   2.3 Equations 4
   2.4 Term Rewriting Systems 5
   2.5 The Initial Algebra 5
   2.6 Conditions and Negations 6
   2.7 Modularization 6
   2.8 Syntactic Freedom 6
   2.9 Signatures as Grammars 7

3 ASF+SDF 7
   3.1 Lexical Syntax 7
   3.2 List functions 8
   3.3 Disambiguation 8
   3.4 Sections 8

4 Pascal 9

5 Pascal in ASF+SDF 10
   5.1 Points of departure 10
   5.2 Overview 10

6 Basic Structures 10
   6.1 Syntax 10
   6.2 Contexts 11
   6.3 Type Representation 11

7 Constraints 13
   7.1 Error Messages 13
   7.2 Communication 13
   7.3 Constraints 13
   7.4 An Example 14

8 Type check functions 14
   8.1 Declarations 14
   8.2 Expressions 16
   8.3 Statements 18

9 Related Work 18

10 Concluding Remarks 19

References 22

A Import Graphs 23
   A.1 Syntax 23
   A.2 Miscellaneous 23
   A.3 Contexts 23
   A.4 Environments 24
   A.5 Type checking 24

B Module Descriptions 25
   B.1 Syntax 25
   B.2 Miscellaneous 25
   B.3 Contexts 25
   B.4 Environments 26
   B.5 Type checking 26

C Text of the Specification 27
   C.1 Syntax 27
   C.2 Miscellaneous 32
   C.3 Contexts 37
   C.4 Environments 50
   C.5 Type Checking 60

D An Example 92
1 Introduction

A formal specification [LB79, Win90] of a programming language precisely describes the syntax and semantics of the language in question. The syntax of a language can easily be specified, but describing the semantics precisely may prove more difficult. Formal language specifications are important for the design and description of new programming languages, and for the automatic generation of tools (parsers, compilers).

Formal specifications for various languages have been published, including several specifications for Pascal [HW73, BJ82, KHZ82, Duk87], but no algebraic specification for Pascal has been presented. We have made an algebraic specification for the static semantics of Pascal, in accordance with the definition of the International Standardization Organization [ISO83]. We discuss the specification, show how it has been set up, explain how difficult details have been handled conveniently, and describe the measures taken to keep the specification simple and easy to read, write, and adapt. We conclude the paper with a brief comparison of the Pascal specification with other algebraic language specifications, and with Pascal specifications in other formalisms.

We assume the reader is familiar with formal specifications, and is interested in language specifications. We introduce algebraic specifications in Section 2. The language used in this document, the ASF+SDF formalism, is explained in Section 3. Some features of Pascal are covered in Section 4.

The formal specification for the static semantics of Pascal is presented in detail in Sections 5, 6, 7 and 8. Finally, related work is discussed in Section 9, and some conclusions are drawn in Section 10. Detailed descriptions of each module are given in the appendices, just as the complete text of the specification.

We conclude the paper with a demonstration how a simple Pascal program is type checked.

We have constructed the specification because of our interest in programming environments. We aim at generating a programming environment from an algebraic specification for the relevant aspects of the programming language. With the generator built so far [Kli90, Hen91] it is possible to generate, in an incremental way, a parser, syntax-directed editor, and type checker for Pascal from the specification presented in this document.

2 Algebraic Specifications

Algebraic specifications can be classified by the fact that the object to be specified is an algebra [GTW78]. An algebra is nothing more than a set of values (called the carrier set), and a set of operations (or functions) over these values. As an example, the algebra of integers has all integers as its values, and operations like + and – as its functions.

More complex algebras may need several sorts of values; for instance an algebra for a “stack” requires values of sort STACK, sort INTEGER, and perhaps of sort BOOLEAN. We therefore introduce a many-sorted algebra which has several carriers, and a unique name (the sort) for each carrier. Note that sorts in algebras are similar to types in programming languages.

The problem is, of course, how to describe the algebra we need. Algebras may be infinite objects (infinitely many integers), so we need some compact notation. We will present such a notation using equations, and we will call a description in such a notation an algebraic specification.

An algebraic specification consists of a signature and a set of equations.
2.1 The Signature

In the signature a list of the sorts is given, and the functions involved are declared.

For example, we may have a signature declaring the sort BOOLEAN \(^1\) and declaring a function and over this sort, having two arguments of sort BOOLEAN and returning a BOOLEAN value:

sorts
  BOOLEAN
functions
  and: BOOLEAN # BOOLEAN -> BOOLEAN

By considering constant values as functions without arguments, we can introduce values T and F:

sorts
  BOOLEAN
functions
  T: -> BOOLEAN
  F: -> BOOLEAN
  and: BOOLEAN # BOOLEAN -> BOOLEAN

We may have different sorts, and the functions may have arguments of different sorts. As an example, consider the signature for a stack, containing declarations for one constant, the empty-stack, and a function top, returning the top element, a function pop, removing the top element, and a function push to add an element.

sorts
  ELEMENT
  STACK
functions
  empty-stack: -> STACK
  top: STACK -> ELEMENT
  pop: STACK -> STACK
  push: STACK # ELEMENT -> STACK

2.2 Terms

If we want to "use" a function, we will have to write something like \(\text{and}(T, F)\), or \(\text{and}(T, \text{and}(F, T))\). We will call such usages closed terms. In general, infinitely many closed terms can be constructed over a signature, e.g., over the Booleans we can construct: \(\text{and}(T, F)\), \(\text{and}(T, \text{and}(T, F))\), \(\text{and}(\text{and}(T, F), \text{and}(F, T))\), etc.

If we allow the declaration of variables, we can construct even more terms:

sorts
  BOOLEAN
functions
  T: -> BOOLEAN
  F: -> BOOLEAN
  and: BOOLEAN # BOOLEAN -> BOOLEAN
variables
  p, q: -> BOOLEAN

Terms with variables, like \(\text{and}(p, q)\), are called open terms. Open terms and closed terms together are called just terms.

2.3 Equations

From a signature we can construct terms, but the meaning of the functions is not yet clear. How do we know, for instance, that the result of \(\text{and}(T, F)\) should be equal to \(F\) (which is of course the behavior of \(\text{and}\) we expect)?

To that end we introduce equations:

sorts
  BOOLEAN
functions
  T: -> BOOLEAN
  F: -> BOOLEAN
  and: BOOLEAN # BOOLEAN -> BOOLEAN
variables
  p, q: -> BOOLEAN
equations:
  [1] \(\text{and}(F, p) = F\)
  [2] \(\text{and}(T, p) = p\)
  [3] \(\text{and}(p, q) = \text{and}(q, p)\)

An equation relates two terms over a signature. Equation [1] states that if one of the arguments of the \(\text{and}\) function is \(F\), the entire term equals \(F\). Equation [2] states that if one argument is \(T\), the result is equal to the value of
the other argument. Equation [3] states that the and operator is commutative.

By substituting closed terms for the variables, we may “compute” complex closed terms. Consider for instance the term \( \text{and}(T, \text{and}(F,T)) \). By equation [2] this is equal to \( \text{and}(F,T) \), which in turn is equal to \( F \) by equation [1].

### 2.4 Term Rewriting Systems

It would be nice if there were some mechanical way to “compute” complex closed terms. The previous example already indicates that such an automatic way may be hard to find. Especially equation [3] \( \text{and}(p,q) = \text{and}(q,p) \) causes danger, because it can be applied infinitely many times. Indeed, it can be proved that the question how to “compute” terms is undecidable: There is no mechanical way to test whether two terms are equal.

There is a more restricted alternative, which nevertheless is attractive. We can view each equation as a rewrite rule explaining how a term can be rewritten into another term. We “direct” equations from left to right, and take care that we rewrite left-hand sides to “smaller” right-hand sides. This automatically excludes infinite loops like those caused by equation [3] \( \text{and}(p,q) = \text{and}(q,p) \). If we take these precautions, we can mechanically reduce closed terms to smaller ones by applying rewrite rules.

In this example the term \( \text{and}(\text{or}(T,F),F) \) can be reduced by applying rewrite rule [3] and [2]: \( \text{and}(\text{or}(T,F),F) \rightarrow \text{and}(T,F) \rightarrow F \). There is no way to reduce \( F \) any further: such terms are called normal forms.

A system of rewrite rules is called a term rewriting system. Term rewriting systems provide means to “execute” an algebraic specification. The “meaning” of a term is its normal form. Although it is possible that this normal form is not unique, we will require the rewrite rules to be defined in such a way that any normal form is unique (i.e., we require the term rewriting system to be confluent).

For a more thorough treatment of term rewriting systems in general, see [Klo91].

### 2.5 The Initial Algebra

A term rewriting system provides an intuitively clear meaning for an algebraic specification. The term rewriting interpretation, however, puts limitations on the equations. In this section we will return to our point of departure, and explain how an algebra can be derived from the specification rather than a term rewriting system, without imposing limitations on the equations. Such an algebra is called a model of the specification.

The equations are now undirected again, and can be applied in both directions. Consequently, a term is not reduced to one single normal form, but to an equivalence class of terms that are all equal to each other. All these equal terms are put into an equivalence class. In the term rewriting interpretation, with every term a normal form is associated, while in the algebraic interpretation with every term an equivalence class is associated.

The elements of the algebra over the specification are precisely these equivalence classes. The functions of the initial algebra are the functions defined in the signature. These operate on equivalence classes of closed terms.
Since it can be proved that the equations not only define an equivalence relation, but also a congruence (the operators respect the equivalence classes) we can use the equations as definitions for the behavior of the functions.

Various models differ in their definition for equal terms. In the initial model two terms are equal if and only if their equality can be proven from the equations, whereas in the final model two terms are equal unless their inequality can be proved. We will use the initial algebra approach.

The initial algebra has two important properties: there are no other elements in it than the terms over the signature (no junk), and elements are only equal if this can be proved from the equations (no confusion).

Thus, we can obtain an (infinite) algebra from a (finite) algebraic specification consisting of a signature and a number of equations. The values in the algebra correspond to terms over the signature, and the functions in the algebra are the functions in the signature operating on equivalence classes of terms.

A far more detailed and formal treatment of initial algebras can be found in [GTW78].

2.6 Conditions and Negations

We can obtain a more powerful (i.e., needing less equations to obtain the same specified object) specification by adding conditional equations. In that case, equations of the form \( T_1 = T_2 \) when \( T_3 = T_4 \) or \( T_1 = T_2 \) when \( T_3 \neq T_4 \) become possible.

For instance, suppose we specify the natural numbers in the classical algebraic way: 0, succ(0), succ(succ(0)), etc., then we may need a check (i.e., a function returning \( T \) or \( F \)) whether a number is zero or not. We can specify this function using a negative condition:

\[
\begin{align*}
T_t & \rightarrow \text{BOOLEAN} \\
F_t & \rightarrow \text{BOOLEAN} \\
o & \rightarrow \text{NAT} \\
\text{succ} : \text{NAT} & \rightarrow \text{NAT} \\
\text{is-null} : \text{NAT} & \rightarrow \text{BOOLEAN} \\
\text{add} : \text{NAT} \# \text{NAT} & \rightarrow \text{BOOLEAN} \\
\text{variables} \\x & , \ y & \rightarrow \text{NAT} \\
\text{equations:} \\[1\] \text{add}(x, 0) = x \\
[2\] \text{add}(x, \text{succ}(y)) = \text{succ}(\text{add}(x, y)) \\
[3\] \text{is-null}(0) = T \\
[4\] \text{is-null}(x) = F \text{ when } x \neq 0
\end{align*}
\]

2.7 Modularization

In order to keep large specifications manageable, modularization operations are needed for splitting large specification into several small ones.

This allows to distinguish sorts and functions that are exported by the module from the hidden sorts and functions that are only used to define some properties of the exported functions.

2.8 Syntactic Freedom

In large specifications, we do not want to be annoyed by clumsy notations. For instance, we do not want to write \( \text{and}(T,F) \) but \( T \) and \( F \), and we do not want to write \( \text{push}(S,I) \) but something like \( \text{push _on } S \). Thus, we do not want to be restricted to prefix operations, but would like to have postfix or mixfix as well. We want to have as much notational freedom as possible.

To allow this, we will use another notation for signatures:

\[
\begin{align*}
\text{sorts} & \text{ELEMENT} \ \ \ \text{STACK} \\
\text{functions} & [\text{"["} \ \ \text{STACK} \\
& [\text{"pop from"} \ \text{STACK} \\
& [\text{"what is the top of"} \ \text{STACK} \ [\text{"?"}] \ \rightarrow \ \text{ELEMENT} \\
& [\text{"push"} \ \text{ELEMENT} \ [\text{"on"} \ \text{STACK} \ \rightarrow \ \text{STACK}
\end{align*}
\]
Although the use of free syntax may have been a little exaggerated in this example, it will be clear that this notational freedom does provide benefits.

We can construct terms over such a signature in the same way, we can write down equations over the signature, and we can eventually derive a term rewriting system from a signature with equations. Thus, we have obtained notational freedom without introducing any problems.

2.9 Signatures as Grammars

An interesting result is obtained when reading these free-syntax function declarations as \( X := \alpha \) instead of \( \alpha \rightarrow X \), and when considering the sorts as nonterminals. In that case each function declaration turns out to be a grammar rule of a BNF grammar for the language of closed terms over the signature.

3 ASF+SDF

Thus far, we only introduced general theory concerning algebraic specifications. By now we have developed enough machinery to introduce the ASF+SDF [Hen91] algebraic specification formalism smoothly.

Originally only ASF, an acronym for Algebraic Specification Formalism existed [BHK89, Hen88]. ASF is a formalism supporting modularization and conditional equations. A system, the ASF system, was able to generate rewriting systems from the specification, and to execute them.

At almost the same time, SDF, Syntax Definition Formalism, had been developed [HHKR89]. SDF supports the definition of lexical, context-free and abstract syntax at the same time, and incremental parsers can be generated from it.

These two formalisms have been combined into one algebraic specification formalism called ASF+SDF. The ASF part takes care of the (conditional) equations, while the SDF part provides complete syntactic freedom for the specifier.

Specifying in the ASF+SDF formalism is supported by the ASF+SDF system [Kli90] [Hen91, Chapter 5]. This system is able to generate parsers from ASF+SDF specifications and to derive term rewriting systems for specifications in ASF+SDF. Moreover, it can generate syntax-directed editors for both modules in the specification as well as terms over the signature. The system is able to perform several static and semantic checks on the specifications, and supports testing of specifications.

3.1 Lexical Syntax

SDF allows the definition of lexical syntax, i.e., the definition of the elementary "words" of the syntax. There are two important corollaries for the specifier: First of all the layout for the equations must be defined, i.e., the specifier himself can (and must) define what symbols constitute white space, and how comments can be recognized. A typical layout definition is:

```plaintext
exports
lexical syntax
[ \t\n]   -> LAYOUT
"\%" \[\t\n\] \"\%"   -> LAYOUT
```

A space, tab, or new line is a layout symbol, as well as everything between two percent signs and a new line. If tokens of sort LAYOUT are detected in a text, they are ignored.

Secondly, variable declarations are treated as declarations of lexical syntax. This implies that any constructs allowed in the lexical syntax definition are allowed in the variable definition section as well. Consequently it is possible to define the variables \( p_1, p_2, p_3, \ldots \) all at once:
variables
"p"[0-9]*\s*⇒\s*SORTNAME

This declares all words starting with a p followed by zero or more characters in the range 0-9 to be variables of sort SORTNAME.

3.2 List functions

An important feature of the ASF+SDF formalism is the existence of list functions and list variables. List functions have a variable number of arguments, and list variables may range over any number of arguments of a list function.

As an example, suppose we would like to have a function \[\] for the empty set, \[[e_1]\] for a set with one element, \[[e_1, e_2]\] for a set with two elements, and so on.

The way to define this in ASF+SDF is as follows:

```plaintext
exports sorts SET ELEMENT
context-free syntax
"[" \{ELEMENT ","\}/* "]" ⇒ SET
```

The asterisk * says that we want zero or more ELEMENTs, while the comma says that these should be separated by commas. Thus, a set consists of ELEMENTs, separated by commas and delimited by [ and ].

This list notation is simply an abbreviation for the declaration of infinitely many functions \[[ \ldots \] \], each with a different number of arguments. Likewise, the same (concrete) syntax could have been obtained by the following “normal” BNF grammar rules:

```plaintext
context-free syntax
"[" ELEMENTS "]" ⇒ SET
"[" "]" ⇒ SET
ELEMENT ⇒ ELEMENTS
ELEMENTS "," ELEMENT ⇒ ELEMENTS
```

In order to define equations over list functions, we need list variables:

```plaintext
variables
"Elts"[1,3] ⇒ ELEMENT
"Elt" ⇒ ELEMENT
```

```plaintext
equations
[1] [Elts₁, Elt, Elts₂, Elt, Elts₃] =
[Elts₁, Elt, Elts₂, Elts₃]
```

Elts₁, Elts₂, and Elts₃ are list variables, ranging over list of zero or more ELEMENTs separated by commas.

Here we have specified in one single equation that elements of sets do not have multiplicity: Any set containing element Elt at least two times is equal to the set containing one occurrence less of Elt.

3.3 Disambiguation

SDF supports priority and associativity declarations, which are used to disambiguate terms which can otherwise be parsed in several ways. For instance,

```plaintext
exports sorts BOOL
category syntax
"T" ⇒ BOOL
BOOL "and" BOOL ⇒ BOOL \{left\}
BOOL "or" BOOL ⇒ BOOL \{left\}
```

Here and has a higher priority than or, thus stating that, for instance, the term \(T \text{ or } F\) and \(F\) should be interpreted as \(T \text{ and } (F \text{ or } F)\) rather than as \((T \text{ or } F) \text{ and } F\).

Moreover, both and and or are defined to be left associative, that is \(T \text{ and } T\) and \(T\) is interpreted as \((T \text{ and } T) \text{ and } T\) rather than as \(T \text{ and } (T \text{ and } T)\).

3.4 Sections

A module in the ASF+SDF formalism may contain the following keywords:
• imports, listing the names of imported modules;
• exports or hidden, listing the exported or hidden items;
• sorts, listing the sorts;
• lexical syntax, giving lexical syntax;
• context-free syntax, listing the function declarations. Attributes, between \{\} brackets, may be associated with functions, stating for instance that the function is left of right associative;
• priorities, giving priorities between functions;
• variables, declaring the variables used in the equations;
• equations, listing the (conditional) equations of the module.

4 Pascal

Pascal [JW86] is one of the best known and most influential programming languages. It has been standardized by the International Standardization Organization ISO [ISO83]. The Pascal specification described in this document conforms to that standard, except for conformant array parameters and gotos.

We loosely define the static semantics of a Pascal program to be all properties of the program that can be checked without executing it. Most of these properties deal with the type and scope rules of Pascal. We will assume that the reader is more or less familiar with these rules. We will, however, briefly mention some aspects of the static semantics that may be less familiar, yet have significant influence on the design of the specification.

First of all, each identifier, denoting for instance a type, procedure, or function, must be declared before it can be used. The so-called required identifiers (eof, char, integer, etc.) are declared implicitly. There is, however, one exception to this rule. The component of a pointer type definition may be a type identifier that has not yet been declared:

```pascal
type CellPtr = 'Cell;
Cell = record
  Info: InfoType;
  Next: CellPtr;
end;
```

In the similar situation for procedures (or functions), the identifiers may be used before the complete declaration if there is a forward declaration:

```pascal
function ParseTerm(var S:T): Boolean; forward;
function ParseFactor(var S:T): Boolean;
begin
  \{ \ldots call ParseTerm \ldots \}
  \end
end
```

```pascal
function ParseTerm;
begin
  \{ \ldots call ParseFactor \ldots \}
  \end
```

Furthermore, type equivalence in Pascal is not based on structure. Each occurrence of a type constructor denotes a type that is distinct from any other type denoted by a constructor [ISO83, p.12]. Consequently, A1 and A2 are not of the same type, whereas B1 and B2 are:

```pascal
type BType = array [1..10] of char;
var A1: array [1..10] of char;
A2: array [1..10] of char;
B1: BType;
B2: BType;
```

Finally, redefinition of identifiers, including the required identifiers, is always possible. The old definitions of symbols using an identifier Id within their definition should not change, however, if the definition of Id changes.
5 Pascal in ASF+SDF

5.1 Points of departure

While constructing the specification for the static semantics of Pascal, the following points of departure played an important role.

- Use the official document of the International Standardization Organization (ISO) defining Pascal [ISO83] as the source reference for Pascal, rather than a definition of some Pascal dialect,
- Optimize the specification with respect to readability, i.e., concentrate on ease of reading, writing and adapting the specification rather than on execution speed.
- Do not only specify whether the static semantics of Pascal program are legal or illegal, but also specify what is wrong with an incorrect program (i.e., the specification should produce error messages).

5.2 Overview

The ultimate goal is a function mapping a Pascal program to a possibly empty list of error messages. The static semantics of a program is correct if and only if this list is empty. Type checking does not stop after the first error has been found. A very simple error recovery mechanism is used.

The specification can be divided into three layers.

The top layer (Section 8) covers the “real” type check functions, i.e., the functions that operate on some syntactic construct and call the appropriate functions to check the constraints for that particular construct.

The bottom layer (Section 6) defines several basic properties of Pascal programs, like the syntax, contexts for representing declared variables, and several context-sensitive properties (what is the type of this identifier, is this type compatible with that other type, and so on).

The middle layer (Section 7) defines functions checking the constraints required by the top layer, using the properties defined in the bottom layer. The functions check some property, and if the check fails they generate an error message. This allows a separate treatment of error handling; neither the top nor the bottom level need to know that error messages are produced.

In order to keep the specification manageable, it has been split up into 34 modules. While discussing the specification, we will not cover the details of this modularization. More detailed descriptions of single modules can be found in the appendices.

6 Basic Structures

6.1 Syntax

First, we define what Pascal programs are. To do so, we define terms over the sort \texttt{PROGRAM}, which are constructed using terms over several other sorts, like \texttt{EXPR}, \texttt{STATEMENT}, and \texttt{BLOCK}.

These terms are the syntactically correct Pascal programs. They are defined by translating the BNF grammar rules given in the ISO standard directly into corresponding SDF rules in the signature. In some cases, we slightly changed the grammar, since the ISO grammar is not entirely unambiguous, and since SDF allows more elegant notations using, for instance, priority declarations.

Terms over these syntactic sorts will be the input for several type check functions.
6.2 Contexts

A Pascal programmer may define several entities (such as types, variables, or procedure) and may assign symbolic names (identifiers) to them. At each point in a Pascal program, some symbols have been declared, and only these symbols may be used. This list of declared symbols may, of course, be different at different points in the program. We will call such a list a *context* of a point in the program.

The context contains both the *required* identifiers (such as integer or eof) and the identifiers declared by the programmer.

The entries for the definitions of the user-defined identifiers are copies of the program texts giving their declarations. Thus, a context simply is a list of Pascal declarations.

The required identifiers that can be described using the conventional Pascal declaration constructs, like type Boolean = (false, true), are represented in that way in the context. Those that cannot, like the procedure readln which may have any number of parameters, are described by the entries predefined-function, predefined-procedure or by the type definition type <required-type-name> = <required-type-name>.

As an example, consider the following skeleton of a part of a Pascal program:

```pascal
procedure proc1(par1, par2: t1);
  var v1: t1;
  var v2: t2;
  procedure subproc(var p3: t3);
  const MAX = 30;
  begin
    ... (* P1 *)
  end;
begin
  ...;
end;
```

The context at point P1 is shown below. Definitions for some of the required identifiers are also given:

```plaintext
[ predefined-procedure readln,
  type integer = integer,
  type Boolean = (false, true),
  function sin(x: real): real,
  procedure proc1(par1, par2: t1),
  block-mark,
  var v1: t1,
  var v2: t2,
  procedure subproc(var p3: t3),
  block-mark,
  const MAX = 30 ]
```

To allow easy processing of scope rules, the block-mark entry is introduced. It is used to define a series of nested blocks. If a new block is entered, a block-mark is added to the context. If the end of a block is reached, all symbols added to the context since the most recent block-mark, as well as the block-mark itself, can be removed.

Note that it is not strictly necessary to define contexts. All information in the constructed CONTEXT term can be retrieved from the PROGRAM term itself. We introduced contexts because they are more practical. Searching the declaration of an identifier in a a PROGRAM term is much more complicated than looking it up in the corresponding CONTEXT term.

6.3 Type Representation

It should be clear by now that the CONTEXT is an essential structure in the specification. Moreover, we will use CONTEXTS as the representation for the various data types in the program.

A considerable part of checking the static semantics of a Pascal program is concerned with checking types. In an expression, the type of the operands of * must be integer or real, and the types of the actual parameters must match the types of the formal parameters of a procedure. In order to check these requirements, some representation of types is needed.

Single type declarations cannot be used as the representation for types, since type equivalence is not based on structure (see Section 4).
The identifiers used in a type declaration may be redefined later on. Moreover, textual expansion of the type identifiers is not possible either, since the type definitions for pointers may be recursive.

Type equivalence depends on the place where the type was defined, or, to say the same, depends on the context in which the type was defined. Therefore, we have chosen to represent a type by a `CONTEXT`. If a `CONTEXT` is used as a type, the entry entered last should be a type definition, and the entries before it constitute the context of the type. As a consequence, equivalence of Pascal types can be checked simply by the `=` operator applied to terms of the sort `CONTEXT`.

In order to allow practical processing of types, several simple abstract functions have been defined on contexts. There are functions to classify types (is a type an array, a subrange, a numerical type, etc.), or to retrieve information from types (what is the component type of an array, of a set, of a pointer, etc.). All information needed to answer these questions is available in the type representation (the `CONTEXT`) itself. The only type operation that needs extra information is the retrieval of the component of a pointer type, since the definition of the component type need not be given before the definition of the pointer type itself.

We illustrate definitions of these functions by two examples. A context denotes a subrange type if its most recent entry is a subrange type definition:

```plaintext
context-free syntax
"is-subrange-type?" "{ CONTEXT "}" -> BOOL
variables
"_const"[12]* -> CONST
"_id" -> IDENT
"Prefix" -> { ENTRY "," }*

equations
is-subrange-type?([Prefix,
```

```plaintext
type _id1 = _const1 .. _const2 ])
=
TRUE

Secondly, in order to obtain the type representation for the index-type of an array, the definition for the index-type subterm should be retrieved in the preceding entries of the context:

```plaintext
context-free syntax
CONTEXT "." "index-type" -> CONTEXT
variables
[12]"-type" -> TYPE-DESCRIPTOR
"_id1" -> IDENT
"Prefix" -> { ENTRY "," }*

equations
[Prefix, type _id1 = array [I-type] of C-type] . index-type = get-type of I-type in [Prefix]
```

As an example of the practical use of these abstract functions, consider the following part of the specification of type compatibility. The equation is the algebraic alternative to the ISO sentence “Types T1 and T2 shall be designated compatible if both T1 and T2 are subranges of the same host type” [ISO83, p.20] Abstract functions `is-subrange-type` and `.host-type` are used:

```plaintext
context-free syntax
CONTEXT "and" CONTEXT "are compatible?" -> BOOL
variables
"T"[12]* -> CONTEXT

equations
is-subrange-type?(T1) = TRUE,
is-subrange-type?(T2) = TRUE,
T1.host-type the same as T2.host-type? = TRUE

---------------------------------------------------------------------
T1 and T2 are compatible? = TRUE
```

These functions provide an abstract interface to retrieve all information necessary from Pascal data types.
7 Constraints

7.1 Error Messages

Error messages are terms over a signature of function names that are normal English sentences (using the syntactic freedom of the formalism). These functions may be parameterized by language constructs. For example:

```
context-free syntax
"Identifier" IDENT "not declared."  -> ERROR
"Parameter lists are not congruous."  -> ERROR
```

If a more detailed description of the place where the error occurred is desired, the language construct in which the error occurred can be set. The signature for this is shown, including the specification of lists of errors.

```
context-free syntax
CONSTRUCT "::" ERROR       -> CONSTR-ERROR
CONSTR-ERROR*  -> ERRORS
```

7.2 Communication

Many tc-functions will be needed to check all kinds of Pascal constructs. These tc-functions will have to communicate information. In order to provide communication between tc-functions in a uniform way, we introduce environments.

The environment provides the information needed to check a particular construct, and it provides a way to represent results of a check.

The most important fields of an environment are the context field, representing the CONTEXT of the construct in question, the errors field, representing the list of errors encountered so far, and the result field, which can be used to pass results of a type check, for example the type of an expression. Operations to manipulate these environments are provided.

Thus, a typical tc-function will retrieve its information from the (context of the) incoming environment, and will return its information by an adapted environment (either an error added to the errors, or a result in the result field).

7.3 Constraints

Many tc-functions will have to check constraints on a construct or context of a construct. If that constraint is met, no special action needs to be taken, but if it is not met, an error message should be added to the environment. Different tc-functions may have to check the same constraints.

We will call the class of functions checking these constraints should-be functions. The names of these functions are like normal English sentences, and typically have the words "should be" in their name.

The should-be functions operate on an environment. This environment is returned unchanged if the constraint is met, but it is returned with an error message added to it if the constraint is not met. The constraint is typically tested by a Boolean function.

We give an example of the specification of a typical should-be function:

```
context-free syntax
  CONTEXT "should be ordinal in" ENV  -> ENV
variables
  "T"  -> CONTEXT
  "Env"  -> ENV
equations
  is-ordinal-type?(T) = TRUE
  T should be ordinal in Env = Env
  is-ordinal-type?(T) ! = TRUE
  T should be ordinal in Env =
  add error [Ordinal type expected.] to Env
```
7.4 An Example

We illustrate the use of environments and should-be functions by a typical tc-function, the function stat-tc checking the Pascal if-statement.

context-free syntax
"stat-tc(" STATEMENT "," ENV ")" -> ENV

variables
"_stat" -> STATEMENT
"_expr" -> EXPR
"Env"[1-4] -> ENV

equations
expr-tc( _expr, Env1 ) = Env2,
stat-tc( _stat, Env2 ) = Env3,
Env2.result should be Boolean in Env3 = Env4

stat-tc( if _expr then _stat, Env1 ) = Env4

The tc-functions stat-tc and expr-tc have an environment as argument and return a (possibly changed) environment. The should-be function is used to check the constraint that the expression is of type Boolean. It takes the result of type checking an expression, that is the result field of the environment Env2, and returns an adapted version of Env3.

8 Type check functions

The contexts, errors, environments, and should-be functions provide enough machinery to specify the functions that really check a Pascal language construct (called the tc-functions).

The tc-functions together perform a single pass over the syntax, and a uniform way of checking the static semantic of the program. A tc-function may check a language construct by calling one or more should-be functions in order to check constraints, or by calling other tc-functions for the syntactic parts of the construct, or by a combination of these. The tc-functions operate on a language construct and an env structure, and return a (possibly changed) env structure.

Thus, checking an if statement, (Section 7.4) implies calling the tc-functions for EXPR and STAT, and checking the constraint that the expression should be Boolean.

The top tc-function checks an entire Pascal program. It is shown below. This function operates on a program only (not on an env structure), and returns a list of errors.

context-free syntax
"tc(" PROGRAM ")" -> ERRORS
"header-tc(" PROG-HEADER "," ENV ")" -> ENV
"block-tc(" BLOCK "," ENV ")" -> ENV

variables
"Env"[123] -> ENV
"Header" -> PROG-HEADER
"Block" -> BLOCK

equations
header-tc( Header, new-env) = Env1,
block-tc( Block, Env1) = Env2

stat-tc( if _expr then _stat, Env1 ) = Env4

c( Header; Block; ) = Env2.errors

8.1 Declarations

Checking the declarations (labels, constants, types, variables, procedures, and functions) of a Pascal program involves (1) checking that the declarations themselves are correct, and (2) representing the declarations in the context field of the environment, so that the proper use of the defined identifiers can be checked if they appear later on in the program.

The most important checks of the declarations themselves deal with the correct use of identifiers. Within one block, no identifier should be declared more than once. If a declaration uses another identifier, this identifier should have been declared already, and it should be used according to its definition. For instance:

context-free syntax
"const-tc(" CONST-DEF "," ENV ")" -> ENV
variables
"_id[1-2]" -> IDENT
"E[1-3]" -> ENV
"Sign" -> SIGN
"Type" -> CONTEXT

equations
_id[1] should not be declared in block of E1 = E2
_id[2] should be a constant in E2 = E3,
const-type of _id[2] in E2.context = Type,
Type should be a number in E3 = E4

const-tc( _id[1] = Sign _id[2], E1 ) =
  add decl "const _id[1] = Sign _id[2]" to E4

Thus, for the definition const _id[1] = _id[2]; it should be checked that _id[1] has not yet been declared within the current block, and that _id[2] has been defined as a numerical constant. The should-be functions required for these constraints can be defined easily.

Constants are the only items in a Pascal program for which values are maintained during type checking. For variables or expressions this is not needed, since their values depend on the execution of the program. For constants of an ordinal type (integer, char, or enumeration), the ordinal values, as defined by the ISO standard, are used to check subranges. Empty subranges, like 1..-1 are forbidden.

Correct handling of procedures (functions) is obtained by first adding the procedure-heading to the context (so that the procedure is known and can be called), then adding a block-mark to the context, followed by the addition of all formal parameters as local variables of the procedure. This allows easy checking of the proper use of the parameters in the body of the procedure. As soon as the end of the procedure body is reached, all declarations within this block, including the block-mark, are removed from the context. The procedure heading, however, entered before the block-mark, is not removed, because the procedure still can be called by other procedures.

We give an example of the functions to check a type denoter. The checks for a type denoter have to be specified for each type constructor (array, record, enumeration, etc.). We show the checks for the array declaration. Multidimensional arrays can be defined in two ways, as array [i_1..i_2] of c, or as array [i_1] of array [i_2] of c. These types are equivalent according to the standard [ISO83, p.14]. The first equation (below) makes this equivalence explicit. The tc-function for the arrays (in the second equation) then only has to check the one-dimensional case, which is done by checking the index type and the component type of the array, and by checking that the index type is an ordinal type. Notice the use of the variable name i_1..i_2 a common mathematical notation for one or more items separated by comma's.

equations
type-denot-tc( I-type, Env1 ) = Env2,
type-denot-tc( C-type, Env2 ) = Env3,
Env2.result should be ordinal in Env3 = Env4
type-denot-tc( array [I-type] of C-type,
             Env1 ) = Env4

Both pointer component identifiers and forwarded procedures or functions (see Section 4) deal with delayed definitions for identifiers. Therefore, in spite of several differences, they are handled in a similar way. We will briefly describe the pointer case.

For each pointer type definition, add the identifier in the component to a list of pointer-component-identifiers. Check at the end of the
checks for the type definition part, whether a definition has been given for all identifiers in that list. Consequently, a list of pointer identifiers must be passed between type definition check functions. This list can be made part of the environment.

8.2 Expressions

Type checking an expression (constants, variables, function calls, arithmetic, Boolean, etc.) always results in a type for that expression, which is returned in the result field of the environment. This allows to check easily whether the expression has been used in a legal way. Remember for example the tc-function for the if statement (Section 7.4), which involved checking whether the type of the expression was Boolean.

It is easy to find a type for most of the simple expressions (numbers, strings, constants, etc.) of Pascal. For the empty set \[\{\}\] this is more difficult. Since the empty set should be compatible with each other set, it is impossible to associate a type already available in Pascal with the empty set. Therefore, we introduce a type not available in Pascal itself, which we call \texttt{empty-set}, and we define the \texttt{is-compatible?} function in such a way that \texttt{empty-set} is compatible with any set. For similar reasons we have introduced the \texttt{nil-pointer} as a new type. Type checking simple expressions is illustrated below:

\begin{verbatim}
context-free syntax
"sim-expr-tc(" EXPR "," ENV ")" -> ENV

variables
"_id" -> IDENT
"Env"[1-3]* -> ENV
"Type" -> CONTEXT

equations
_id should be a constant in Env1 = Env2.
const-type of _id in E1.context = Type
sim-expr-tc( _id, Env1 ) = set-result of Env2 to Type
sim-expr-tc( nil, Env ) =
set-result of Env to [nil-pointer]
sim-expr-tc( [], Env ) =
set-result of Env to [empty-set]
\end{verbatim}

The type of an expression denoting a variable access is simply the type of the variable as it can be found in the context. Retrieval of components of variables of structured types (e.g. \texttt{a[3]}, \texttt{a^i} or \texttt{a[i]}) results in the type of the component of the variable. First, of course, it is checked that the type of the variable allows the retrieval of a component.

The type of an expression denoting a function call is the type returned by the function, which is found in the context given the function identifier. The most complex part of checking a function call is checking that the actual parameters match the formal parameters. Since parameters of functions and procedures are checked in a similar way, we will discuss both here, even though procedures are not expressions in Pascal.

The requirements for the normal, user-defined procedures are easily caught by the should-be approach. The constraints include that there must be neither too few, nor too many actual parameters, that the types of the actual parameters must be compatible (value parameter), or assignment-compatible (var parameter) with the formal parameters, and that for procedure or function parameters, the parameter lists must be congruous. In addition, for \texttt{var} parameters, the actual parameters should not be components or fields of \texttt{packed} structured types. Finally, the context-free syntax allows all actual parameters to be postfixed by a format, e.g. the \texttt{:3:4} in the required procedure \texttt{write(a:3:4)}. This is only allowed for the procedure \texttt{write}. Since it is impossible to "wire" this in the context-free syntax (the programmer may redefine the identi-
The correct use of the format has to be checked for each actual parameter.

Slightly different is the specification for the required procedures, a part of which is shown below. Many input/output procedures have an optional file parameter. If this parameter is omitted, the procedures operate on standard input or output. Besides, any number of parameters may be given. The ISO standard specifies these procedures by summarizing the equivalences between the various cases, and by giving the constraints for one of them. For example, the call `readln(a,b)` is equivalent to (if `a` is not a file) `read(input,a); read(input,b); readln(input`). The requirement for `readln` with only one parameter then is that the parameter should denote a file. In the algebraic specification we follow the same approach, the specification for calls of the procedure `readln` is shown below.

**context-free syntax**
```
"std-ioc(" IDENT PAR-LIST "," ENV ")" \(\Rightarrow\) ENV
"file-ioc(" IDENT PAR-LIST "," ENV ")" \(\Rightarrow\) ENV
```
**variables**
```
"Env" \(\Rightarrow\) ENV
"par" \(\Rightarrow\) ACTUAL-PAR
"file" \(\Rightarrow\) ACTUAL-PAR
"pars" \(\Rightarrow\) { ACTUAL-PAR "," }*
"par+" \(\Rightarrow\) { ACTUAL-PAR "," }+
```
**equations**
```
std-ioc( readln, Env ) =
  file-ioc( readln(input), Env )
```
```
file-ioc( readln(file, par), Env ) =
  file-ioc( readln(file),
          file-ioc( readln(file), Env ) )
```

It would be possible to give a specification in should-be style for each arithmetic operator (+, and, < etc.). However, there are 16 operators which all may have operands of several types, and this may result in many large and unreadable equations.

ISO [ISO83, p.39] solves the problem by giving tables summarizing the various operand types allowed for each operator. The question whether `real` and `integer` are allowed as operands for the + operator, can be answered simply by looking in the table for the + and checking the operand types in that entry.

But if this is the most practical way to specify operand types, why not do it that way in our algebraic specification?

**sorts**
```
TABLE-HEADER TABLE-ENTRY TABLE
```
**context-free syntax**
```
"Operator | Operand1 | Operand2 | Result
-----------|----------|----------|----------
+           | int      | int      | int
+           | int      | real     | real
+           | real     | int      | real
+           | real     | real     | real
+           | set      | set      | set
...         
```
```
{ similar for \(-\), \(*\), \div\), \mod\), \not\), \and\)
```
**sorts**
```
arithmetic-op-table
```
**equations**
```
arithmetic-op-table =
  Operator | Operand1 | Operand2 | Result
-----------|----------|----------|----------
in         | ordinal  | set      | boolean
<          | string   | string   | boolean
<          | simple   | simple   | boolean
```
```
relational-op-table
```
**equations**
```
relational-op-table =
  Operator | Operand1 | Operand2 | Result
-----------|----------|----------|----------
in         | ordinal  | set      | boolean
<          | string   | string   | boolean
<          | simple   | simple   | boolean
```
8.3 Statements

We have already seen how procedure statements (Section 8.2) and if statements are checked. Most of the tc-functions for statements (if then else, while do, repeat until, assignment, and begin ... end), follow the same approach as the tc-function for the if statement. For the case statement the uniqueness of the case constants has to be checked. For a with statement a new block can be created, in which all fields of the record in question are entered as local variables.

We will only describe the for statement, for which the type check equation is given in the next figure. It is the most complex statement to be checked. Moreover, it is the most complex equation of the entire specification, and we think it is interesting to see how even the most complex equation is still manageable.

Type checking the for statement mainly deals with checking properties of the control variable. The constraints for the control variable, illustrated clearly in the equation itself, are: the control variable should be declared in the same block in which the for statement is used (line 9), the control variable should possess an ordinal type (line 15), and both the initial (expr1) and the final (expr2) should be assignment-compatible with the type of the control variable (lines 18 and 21).

The notions of lines 10 and 11, control and threatening variables, may need an explanation. In Pascal [ISO83, p.46], the loop variable of a for statement (called the control variable) should not be at the same time a control variable of an outer for statement. This is checked in line 10.

Moreover, the control variable should not be threatened by assignments to it within the body of the for statement. This is checked in line 11.

To implement the functions of line 10 and 11, we have used an idea of Welsh and Hay [WH86, Chapter 10]. The definitions of variable declarations has been changed to allow marking variables. Marks can be possibly-threatening, control-variable, or no mark at all. During the checks for the body of the for statement, the control variable is marked control-variable. Variables that may become control variables, and are assigned some where are marked possibly-threatening. If a possibly-threatening variable is used as a control variable, or if a control-variable is assigned somewhere, an error is raised.

The mark is set in line 13, and reset in line 24.

9 Related Work

Thanks to the importance of Pascal and to the influence of the language on other languages, several formal specifications for Pascal have been made over the past few years. All recent Pascal specifications are based on the ISO standard [ISO83], which is written in formal English but not in a formalism with mathematically defined semantics.

Attribute grammar specifications are given for the synthesizer generator in the Synthesizer Specification Language [RT89] and for the GAG system in the ALADIN formalism [KHZ82]. From these, just as from ASF+SDF specifications, parsers, syntax-directed editors, and type checkers can be derived. Both the attribute grammars and our algebraic specification take about 60 pages, but the ASF+SDF specification seems to be better readable. Moreover, it does not depend on predefined notions like integers or strings, but
defines them all explicitly.

Watt’s extended attribute grammar for Pascal [Wat79] is very compact, but not yet based on the ISO standard. Duke [Duk87] gives a specification in predicate rules [DJR87], an approach similar to attribute grammars in a notation akin to the Z specification language. Neither of these specifications can be executed.

A very early formal specification of the semantics of Pascal is given by Hoare and Wirth [HW73], but they “consider such topics as rules about the scope or validity of names as belonging to the realm of syntax” [HW73, p.336], so they hardly cover the static semantics of Pascal.

A denotational specification, also according to the ISO standard, of both the static and dynamic semantics for Pascal is given by Andrews and Henhapl [BJ82, Chapter 7], in 75 pages using the VDM formalism.

10 Concluding Remarks

We have given the first algebraic specification for the static semantics of Pascal. We aimed at readability and clearness by carefully choosing
the names of sorts, functions, and variables, by providing abstract interfaces, like the functions operating on types, by looking for a good modularization, and by striving for as much uniformity as possible.

The uniform way of dealing with constraints and errors, using the should-be functions, raises the question whether the processing of errors can be mechanized. Can clear error messages be derived from the names of the functions checking particular constraints? If so (as suggested, among others, also by [BLM89, JJWW90]) we would have an algebraic specification of the correct static semantics of a language, and a system (typically an adapted term rewriting system) capable of producing adequate error messages for the incorrect programs analogous to the generation of a syntax checker from a BNF grammar. This promises to be an interesting area for further research.

Another question concerns the style of the specification. Inspired by the objective of readability, a software engineering specification style has been used, resulting sometimes in verbose names or superfluous functions. It would also be interesting to construct a specification using a more mathematically oriented style, e.g., with short names and the absolute minimum number of functions.

We have constructed the algebraic specification in software-engineering style with validation against the ISO standard in mind. The names of functions and structures in the algebraic specification precisely match the names of the ISO concepts they intend to formalize.

Alternatively, we could have written the specification in a more mathematical style if we had intended to verify the specification with another formal definition (e.g., VDM, [BJ82, Chapter 7]) of Pascal. Although this would be very interesting, proving some mathematical relation or equivalence between the algebraic and denotational specification probably would be quite complex.

Another interesting question is which concepts or modules can be reused in specifications for other programming languages. It may be clear that for the static semantics of other block structured languages like Modula-3, Ada, or C, many aspects of the Pascal specification can be reused. Many of them probably can be reused for the dynamic semantics as well; candidates are the environments, functions on types, and the should-be approach for error handling.

Algebraic specifications for programming languages may become more and more important, especially when they will be used as the input for programming environment generators. To allow easy construction of large specifications, generic modules for important structures, development methods, and specification heuristics are needed. The only way to obtain these is by writing, reading, and exchanging many complex and realistic algebraic specifications.

Acknowledgements

Paul Hendriks and Emma van der Meulen introduced me to ASF+SDF very patiently. Jasper Kamperman, Emma van der Meulen, Jan Rekers, and Machted Vork commented on (parts of) the manuscript. Jan Heering and Paul Klint carefully read the entire paper, suggesting many useful improvements. Thank you all.
References


A Import Graphs

The import graph for the specification is given. Several modules are grouped in 5 groups: Syntax, Miscellaneous, Contexts, Environments, and Type Checking. An arrow from module \( M_1 \) to \( M_2 \) implies that \( M_1 \) is imported by \( M_2 \).

A.1 Syntax

A.2 Miscellaneous

A.3 Contexts
A.4 Environments

A.5 Type checking
Module Descriptions

Each module in the specification is briefly described, in the order in which they appear in the import graphs. This appendix is intended for reference purposes.

B.1 Syntax

SyntaxLayout Define the lexical syntax of Pascal layout symbols, e.g. white space and comment brackets { and }.

SyntaxNumbers Define the lexical syntax of Pascal unsigned numbers, e.g. 123, 123.45, or 123e-2.

SyntaxConst Define the syntax of basic constant values, i.e., signed numbers, strings and labels.

SyntaxTypes Define the syntax of the type constructors (array, record, etc.) of Pascal.

SyntaxHeaders Define the syntax of procedure and function headers, and of the various sections (label, const, type, var, procedure or function) in a block.

SyntaxExpr Define the syntax of expressions, e.g. constants, variable accesses, Boolean operators, sets, arithmetic operators, relational operators, etc.

SyntaxStat Define the syntax of all statements (while, repeat, if, etc.).

SyntaxProgram Define the syntax of the program header and of the Pascal block.

B.2 Miscellaneous

Layout Define the Layout as used in the equations.

Booleans Define the syntax and semantics of the Booleans as needed within the specification. These Booleans are not used within a Pascal program, in which the Booleans simply are a predefined enumeration type. In order to avoid confusion, the values TRUE and FALSE in the specification are CAPITALIZED.

Integers Define the syntax and semantics of the Integers. This module imports the module SyntaxNumbers, thus allowing to perform arithmetic on numbers used in a Pascal program.

StringFunctions Define string functions, at the moment only a function to determine the length of a string.

ASCII Define the ordinal values of elements of the Pascal predefined data type “char”. This module has been generated by a Pascal program.

B.3 Contexts

Remember that the context plays a major role in the specification, since it is used as the representation for types.

Contexts Define the contexts as tables of declarations, and define the initial block containing the definitions for the required identifiers.

ContextDefs Define functions investigating how the identifier in the top entry of the context is defined, e.g. as a type, as a variable, as a function, etc.

ContextSearch Define functions searching a context, i.e., returning the most recent definition for a given identifier.

ConstantAnalysis Though ISO defines constants as single lexical entities, they still
may have properties depending on the contexts (e.g. the type of \texttt{const a = -b};
depends on the declaration of \texttt{b}). These context operations are defined in this
module.

\textbf{Context-sen-fun} In this module, all context sensitive definitions of the ISO standard
are given. These include classification of several types, retrieval of components of
types, etc.

\textbf{Compatible} Define type compatibility.

\textbf{B.4 Environments}

\textbf{OperatorTables} Define the tables and functions on them, used when type checking
expressions. Redefinition of operator names is necessary, since only lexical items can appear in associativity and priority declarations.

\textbf{Errors} Define the syntax of the error messages. Modules providing sorts needed in
the messages are imported.

\textbf{Environments} Define the structure for the environments and the basic operations on
environments.

\textbf{ShouldBe} Define the should-be functions

\textbf{B.5 Type checking}

\textbf{TC-defs} Define how to type check the declarations of a block.

\textbf{Expr-misc} Define several functions needed when type checking expressions.

\textbf{TC-expr-simple} Define how to type check simple (constants) expressions.

\textbf{TC-expr-vars} Define how to type check variable accesses.

\textbf{TC-expr-calls} Define how to type check user-defined function and procedure calls.

\textbf{TC-spec-funs} Define how to type check calls to predefined functions (eof, eoln, ord, etc.)

\textbf{TC-spec-procs} Define how to type check calls to predefined procedures (readln, write, pack, etc.).

\textbf{TC-expr} Merge all expression type check functions.

\textbf{TC-stat} Define how to type check statements.

\textbf{TC-block} Define how to type check an entire block, including forwarded functions and procedures.

\textbf{TC-prog} Define how to type check an entire program, including the program header.
C    Text of the Specification

In this appendix the complete text of the specification for the static semantics of Pascal is given.

C.1  Syntax

module SyntaxLayOut
XX Define layout symbols of Pascal programs. Note that in this specification
XX no " between comment delimiters (" and ") is allowed.
exports
lexical syntax
   [ \t\n ]    -> LAYOUT
   "(*) [ ]* "    -> LAYOUT
   "(+) [ ]* "    -> LAYOUT
end module SyntaxLayOut

module SyntaxNumbers
XX Specify the syntax of real and integer numbers
exports
sorts UNS-NUMBER UNS-INT UNS-REAL DIGIT
lexical syntax
   [ \t\n ]    -> LAYOUT
   [0-9]    -> DIGIT
   DIGIT*    -> UNS-INT
   UNS-INT "," UNS-INT    -> UNS-REAL
   UNS-INT "," UNS-INT [eE] UNS-INT    -> UNS-REAL
   UNS-INT "," UNS-INT [+-] UNS-INT    -> UNS-REAL
   UNS-INT [eE] UNS-INT    -> UNS-REAL
   UNS-INT [eE] [+-] UNS-INT    -> UNS-REAL
context-free syntax
   UNS-INT    -> UNS-NUMBER
   UNS-REAL    -> UNS-NUMBER
variables
   "_uns-int"[0-9]*    -> UNS-INT
   "_uns-real"[0-9]*    -> UNS-REAL
   "_uns-number"[0-9]*    -> UNS-NUMBER
end module SyntaxNumbers

module SyntaxConsts
XX Define constants and values, ISO section 6.1.3, 6.1.5 6.1.7 and 6.3.
XX Empty strings are not allowed in ISO, but are tolerated in this definition.
imports SyntaxLayOut SyntaxNumbers
exports
sorts IDENT STRING STRING-PART CONST NEW-CONST SIGN LABEL
lexical syntax

```
" " \"[\"\\-]\" -> STRING-PART
STRING-PART+ -> STRING
DIGIT -> LABEL
DIGIT DIGIT -> LABEL
DIGIT DIGIT DIGIT -> LABEL
DIGIT DIGIT DIGIT DIGIT -> LABEL
```

case-free syntax

```
"+" -> SIGN
"-" -> SIGN
STRING -> NEW-CONST
UNS-NMumber -> NEW-CONST
SIGN UNS-NMumber -> NEW-CONST
SIGN IDENT -> CONST
IDENT -> CONST
NEW-CONST -> CONST
```

variables

```
"_const\[0-9]+" -> CONST "_const\[0-9]+" -> {CONST ","}+
"_id\[0-9]+" -> IDENT "_id\[0-9]+" -> {IDENT ","}+
"_sign\[0-9]+" -> SIGN "_id\[0-9]+" -> {IDENT ","}+
"_new-const\[0-9]+" -> NEW-CONST "_string\[0-9]+" -> STRING
"_label\[0-9]+" -> LABEL "_label\[0-9]+" -> {LABEL ","}+
"_const\[0-9]+" -> {CONST ","}+
```

end module SyntaxConsts

module SyntaxTypes

%% The definitions in this module are according to the syntax as
%% specified in section 6.4.1, 6.4.2 and 6.4.3 in the ISO standard.
imports SyntaxConsts

exports

sorts NEW-TYPE TYPE-DENOTER STRUCTURED-TYPE PACKED-STATUS FIELD-LIST
FIXED-PART VAR-PART VARIANT RECORD-SECTION OPT-SEMI-COLON SELECTOR
FIELDS-DESCRIPTION

context-free syntax

```
packed -> PACKED-STATUS
IDENT -> TYPE-DENOTER
NEW-TYPE -> TYPE-DENOTER
CONST "," CONST -> NEW-TYPE
"" { IDENT "," }+ ")" -> NEW-TYPE
"" IDENT -> NEW-TYPE
PACKED-STATUS STRUCTURED-TYPE -> NEW-TYPE
array "[ \{ TYPE-DENOTER "," }+ "]" of TYPE-DENOTER -> STRUCTURED-TYPE
record FIELD-LIST end -> STRUCTURED-TYPE
set of TYPE-DENOTER -> STRUCTURED-TYPE
file of TYPE-DENOTER -> STRUCTURED-TYPE
```

FIELDS-DESCRIPTION OPT-SEMI-COLON -> FIELD-LIST

end module SyntaxTypes

28
module SyntaxTypes

"." -> OPT-SEMI-COLON
";" -> OPT-SEMI-COLON
FIXED-PART -> FIELDS-DESCRIPTION
VAR-PART -> FIELDS-DESCRIPTION
FIXED-PART ":;" VAR-PART -> FIELDS-DESCRIPTION
{ RECORD-SECTION ";:" }+ -> FIXED-PART
{ IDENT "," }+ ":;" TYPE-DENOTER -> RECORD-SECTION
case SELECTOR of { VARIANT "," }+ -> VAR-PART
IDENT -> SELECTOR
IDENT ":;" IDENT -> SELECTOR
{ CONST "," }+ ":;=" (" FIELD-LIST ")" -> VARIANT

variables
"_type-den" [0-9]* -> TYPE-DENOTER
"_type-den" [0-9]* ":=" -> {TYPE-DENOTER ",;" }+
"_type-den" [0-9]* ":=" -> {TYPE-DENOTER ",;" }+
"_new-type" [0-9]* -> NEW-TYPE "PS" -> PACKED-STATUS
"_fixed" [0-9]* -> FIXED-PART "_var-part" -> VAR-PART
"_variant" [0-9]* -> VARIANT "_variant+" -> {VARIANT ",;" }+
"_selector" -> SELECTOR "_rec-section" -> RECORD-SECTION
"_field-list" -> FIELD-LIST "_rec-section+" -> {RECORD-SECTION ";:" }+
"opt;" -> OPT-SEMI-COLON

end module SyntaxTypes

module SyntaxHeaders

XX Define Syntax formal parameters, ISO section 6.6.3.
XX (i.e. level 0).
imports SyntaxTypes

exports
sorts FORMAL-PAR-LIST FORMAL-PAR-SECTION HEADING PROCEDURE-HEADING
FUNCTION-HEADING FUNCTION-IDENTIFICATION LABELS CONST-DEF CONSTS
TYPES

context-free syntax
PROCEDURE-HEADING -> HEADING
FUNCTION-HEADING -> HEADING
procedure IDENT FORMAL-PAR-LIST -> PROCEDURE-HEADING
function IDENT FORMAL-PAR-LIST ":;" IDENT -> FUNCTION-HEADING
function IDENT -> FUNCTION-IDENTIFICATION
( ( {FORMAL-PAR-SECTION ";:" }+ ")") -> FORMAL-PAR-LIST
{ IDENT "," }+ ":;" IDENT -> FORMAL-PAR-LIST
var { IDENT "," }+ ":;" IDENT -> FORMAL-PAR-LIST
PROCEDURE-HEADING -> PROCEDURE-HEADING
FUNCTION-HEADING -> FUNCTION-HEADING

label { LABEL "," }+ ":;" -> LABELS
{ CONST-DEF "," }+ ":;" -> CONSTS
IDENT "=" CONST -> CONST-DEF
TYPES

29
module SyntaxHeaders
end module SyntaxHeaders

module SyntaxExpr
XX Specify Pascal Expressions, ISO section 6.7.
XX A slighty more general specification is given, allowing expressions like
XX 1 * -1 (instead of " 1 * (-1) " in ISO) as well.
imports SyntaxConsts
exports
sorts VARIABLE-ACCESS MEMBER-DESIGNATOR SET-CONSTRUCTOR
ACTUAL-PAR-LIST NON-EMPTY-ACT-PAR-LIST ACTUAL-PAR EXPR FORMAT

current-free syntax
   IDENT -> VARIABLE-ACCESS
   VARIABLE-ACCESS [" { EXPR "," }]+ -> VARIABLE-ACCESS
   VARIABLE-ACCESS "." IDENT -> VARIABLE-ACCESS
   VARIABLE-ACCESS "." -> VARIABLE-ACCESS
   EXPR -> MEMBER-DESIGNATOR
   EXPR "." EXPR -> MEMBER-DESIGNATOR
   "." {MEMBER-DESIGNATOR ","}+ "." -> SET-CONSTRUCTOR
   "." EXPR -> FORMAT
   "." EXPR "." EXPR -> FORMAT
   EXPR FORMAT -> ACTUAL-PAR
   "+" {ACTUAL-PAR ","}+ "+" -> NON-EMPTY-ACT-PAR-LIST
   NON-EMPTY-ACT-PAR-LIST -> ACTUAL-PAR-LIST
   -> ACTUAL-PAR-LIST
   VARIABLE-ACCESS -> EXPR
   UNS-NUMBER -> EXPR
   STRING -> EXPR
   null -> EXPR
   IDENT NON-EMPTY-ACT-PAR-LIST -> EXPR
   SET-CONSTRUCTOR -> EXPR
   "+" EXPR ")" -> EXPR bracket
not EXPR -> EXPR
"+-" EXPR -> EXPR
"--" EXPR -> EXPR

30
module SyntaxExpr

variables
"_expr"[0-9]* -> EXPR
"_var-acc"[0-9]* -> VARIABLE-ACCESS
"_format"[0-9]* -> FORMAT
"_member"[0-9]* -> MEMBER-DESIGNATOR
"_act-par"[0-9]* -> ACTUAL-PAR
"_act-par-list"[0-9]* -> ACTUAL-PAR-LIST
"_NE-act-par-list"[0-9]* -> NON-EMPTY-ACT-PAR-LIST
"_member"[0-9]* -> MEMBER-DESIGNATOR

priorities
{ not, "+" EXPR -> EXPR , "-" EXPR -> EXPR } >
{ left: "+", "/", div, mod, and } >
{ non-assoc: "+", "/”, "<<", ">>>", ">=" } >

end module SyntaxExpr

module SyntaxStats

XX Specify statements of Pascal, ISO section 6.8.
XX module SyntaxStats
imports SyntaxExpr

exports
sorts UNL-STAT SIMPLE-STAT STRUCT-STAT STATEMENT
CASE-LIST-ELT COMP-STAT DOWN-TO
context-free syntax
LABEL ":" UNL-STAT
UNL-STAT
SIMPLE-STAT
STRUCT-STAT

VARIABLE-ACCESS ":=" EXPR
IDENT ACTUAL-PAR-LIST
goto LABEL

31
// STRUCT-STAT
begin { STATEMENT ";" )+ end -> STRUCT-STAT
if EXPR then STATEMENT -> COMP-STAT
if EXPR then STATEMENT else STATEMENT -> STRUCT-STAT

\{ CASE-LIST-ELT ":;" \)+ OPT-SEMI-COLON end -> STRUCT-STAT
\{ CONST ",," )+ ";" STATEMENT -> CASE-LIST-ELT
repeat \{ STATEMENT ":;" \)+ until EXPR -> STRUCT-STAT
while EXPR do STATEMENT -> STRUCT-STAT
"to" -> DOWN-TO
downto -> DOWN-TO
for IDENT ":=" EXPR DOWN-TO EXPR do STATEMENT -> STRUCT-STAT
with \{ VARIABLE-ACCESS ",," \)+ do STATEMENT -> STRUCT-STAT

priorities
if EXPR then STATEMENT else STATEMENT -> STRUCT-STAT
>
if EXPR then STATEMENT -> STRUCT-STAT
end module SyntaxStats

module SyntaxProgram
%% Define programs and blocks, according to ISO section 6.2.1 and 6.10
imports SyntaxHeaders SyntaxStats

exports
sorts PROC-OR-FUN PROCS-OR-FUNS DEFS DECLS BLOCK PROGRAM PROGRAM-HEADER
PROGRAM-PARAMS
context-free syntax
HEADING ":;" forward -> PROC-OR-FUN
HEADING ":;" BLOCK -> PROC-OR-FUN
FUNCTION-IDENTIFICATION ":;" BLOCK -> PROC-OR-FUN

\{ PROC-OR-FUN ":;" \)+ ";" -> PROC-OR-FUN

DEFS DECLS COMP-STAT -> BLOCK
LABELS CONSTS TYPES -> DEFS
VARS PROC-OR-FUNS -> DECLS

"(" \{ IDENT ",," \)+ ")" -> PROGRAM-PARAMS
program IDENT PROGRAM-PARAMS -> PROGRAM-HEADER
PROGRAM-HEADER ":;" BLOCK "." -> PROGRAM

end module SyntaxProgram

C.2 Miscellaneous

module LayOut
exports
lexsynt
module Booleans
imports LayOut

exports
sorts BOOLEAN
context-free syntax
  "TRUE" -> BOOLEAN
  "FALSE" -> BOOLEAN
  BOOLEAN "|" BOOLEAN -> BOOLEAN assoc
  BOOLEAN "&" BOOLEAN -> BOOLEAN assoc
  " (BOOLEAN )" -> BOOLEAN bracket

priorities "|" < "&"

hiddens
variables b -> BOOLEAN

equations
  [1] TRUE | b = TRUE
  [2] FALSE | b = b
  [3] b | TRUE = TRUE
  [4] b | FALSE = FALSE
  [5] TRUE & b = b
  [7] b & TRUE = b
  [8] b & FALSE = FALSE
end module Booleans

module Integers
imports Booleans SyntaxNumbers

exports
sorts NAT INT
context-free syntax
  UNS-INT -> NAT
  NAT -> INT
  "+-" NAT -> INT
  "-" NAT -> INT
  INT "+-" INT -> INT {left}
  INT "-" INT -> INT {left}
  INT "<" INT -> BOOLEAN
  " (" INT ")" -> INT {bracket}
  "(" INT ")" -> INT {bracket}

hiddens
sorts UNS-INT-LIST
context-free syntax
  " { UNS-INT "," } ")" -> UNS-INT-LIST
  UNS-INT-LIST "right" DIGIT -> UNS-INT-LIST
  UNS-INT-LIST "left" DIGIT -> UNS-INT-LIST
  pred "(DIGIT ")" -> DIGIT
  non-zero "(" NAT ")" -> BOOLEAN

variables
  n [0-9]* -> UNS-INT
  uns-ints [0-9]* -> (UNS-INT ",") *
  c [0-9]* -> CHAR
  d [0-9]* -> DIGIT
Nat [0-9]* -> NAT
Int [0-9]* -> INT

priorities
\{left: INT "+" INT -> INT, INT "-" INT -> INT \} < "*

equations

[01] pred(0) = 0
[02] pred(1) = 0
[03] pred(2) = 1
[04] pred(3) = 2
[05] pred(4) = 3
[06] pred(5) = 4
[07] pred(6) = 5
[08] pred(7) = 6
[09] pred(8) = 7
[10] pred(9) = 8

% -- left and right rotation --
[11] d = 0 => \{uns-ints\} left d = \{uns-ints\}
[12] d ! = 0 => \{n, uns-ints\} left d = \{uns-ints, n\} left pred(d)
[13] d = 0 => \{uns-ints\} right d = \{uns-ints\}
[14] d ! = 0 => \{uns-ints, n\} right d = \{n, uns-ints\} right pred(d)
[15] n ! = 0 => non-zero(n) = \text{T}
[16] n = 0 => non-zero(n) = \text{F}
[17] \text{uns-int("0" y)} = \text{uns-int(y)}

% -- addition --
[18] \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18\}
  \left-digit(c1) \left-digit(c2) = \{n, uns-ints\}
  \text{uns-int(c1) + uns-int(c2) = n}

[19] \text{uns-int(c1) + uns-int(c2) = uns-int(x c),}
  \text{uns-int("0" x1) + uns-int("0" x2) + uns-int("0" x) = uns-int(y)}
  \text{uns-int(x1 c1) + uns-int(x2 c2) = uns-int(y c)}

% -- multiplication --
[20] n * 0 = 0
[21] pred(digit(c1)) = digit(c2)
  \text{uns-int(y1) * uns-int(y2) = uns-int(y)}
  \text{uns-int(y1) * uns-int(y2 c) = uns-int(y "0") + uns-int(y1) * uns-int(c)}

% -- greater than --
[22] \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 0, 0, 0, 0, 0, 0, 0\}
  \left-digit(c1) \right-digit(c2) = \{n, uns-ints\}
  \text{uns-int(c1) > uns-int(c2) = non-zero(n)}

[23] \text{uns-int(y1 c1) > uns-int(c2) = \text{T}}
[24] \text{uns-int(c1) > uns-int(y2 c2) = \text{F}}
[25] \text{uns-int(y c1) > uns-int(y c2) = uns-int(c1) > uns-int(c2)}

34
uns-int(y1) > uns-int(y2) = TRUE
uns-int(y1 c1) > uns-int(y2 c2) = TRUE

uns-int(y1) > uns-int(y2) = FALSE, uns-int(y1) != uns-int(y2)
uns-int(y1 c1) > uns-int(y2 c2) = FALSE

% less than or equal --

Nat1 <= Nat2 = TRUE when Nat2 > Nat1 = TRUE
Nat1 <= Nat1 = TRUE

% subtraction of naturals --

\{0,1,2,3,4,5,6,7,8,9,0,0,0,0,0,0,0,0\}
left digit(c1) right digit(c2) = \{n, uns-ints\}
uns-int("1" c1) - uns-int(c2) = n
uns-int(x1 c1) > uns-int(x2 c2) = TRUE,
uns-int("0" c1) - uns-int(c2) = uns-int(x c),
(uns-int("0" x1) + uns-int("0" x1)) - uns-int("0" x2) - 1 = uns-int(y)
uns-int(x1 c1) - uns-int(x2 c2) = uns-int(y c)

Nat1 > Nat2 = FALSE, Nat1 - Nat2 = Nat  =>  Nat1 - Nat2 = -Nat

% addition and subtraction of integers

Int1 = +Nat1  =>  Int1 = Nat1
Int1 = Nat1, Int2 = Nat2  =>  Int1 + Int2 = Nat1 + Nat2
Int1 = Nat1, Int2 = -Nat2  =>  Int1 + Int2 = Nat1 - Nat2
Int1 = -Nat1, Int2 = Nat2  =>  Int1 + Int2 = Nat2 - Nat1
Int1 = -Nat1, Int2 = -Nat2, Nat1 + Nat2 = Nat  =>  Int1 + Int2 = -Nat
Int1 = Nat1, Int2 = -Nat2  =>  Int1 - Int2 = Nat1 - Nat2
Int1 = -Nat1, Int2 = -Nat2  =>  Int1 - Int2 = Nat1 + Nat2
Int1 = -Nat1, Int2 = Nat2, Nat1 + Nat2 = Nat  =>  Int1 - Int2 = -Nat
Int1 = -Nat1, Int2 = -Nat2  =>  Int1 - Int2 = Nat2 - Nat1
Int1 = Nat1, Int2 = Nat2  =>  Int1 * Int2 = Nat1 * Nat2
Int1 = Nat1, Int2 = -Nat2, Nat1 + Nat2 = Nat  =>  Int1 * Int2 = -Nat
Int1 = -Nat1, Int2 = Nat2, Nat1 + Nat2 = Nat  =>  Int1 * Int2 = -Nat
Int1 = -Nat1, Int2 = -Nat2  =>  Int1 * Int2 = Nat1 * Nat2

end module Integers
module StringFunctions
imports Integers SyntaxConsts

exports
context-free syntax
  length "(" STRING ")" -> INT

hiddens
variables
  a-char -> CHAR
  chars -> CHAR*
equations
  [s1] length (""") = 0
  [s2] length (string (""") chars "") = length (string (""") chars "") + 1
  [s3] string ("" a-char ") != string ("" a-char chars "")
                                      =
                                      length (string ("" a-char chars "") )
                                      =
                                      length (string ("" chars "") ) + 1
end module StringFunctions

module ASCII
imports Integers SyntaxConsts StringFunctions

exports
context-free syntax
  ascii-val "(" STRING ")" -> INT
equations
  { Ordinal numbers for characters. }
  % var i: integer;
  % begin
  % for i := 32 to 127 do
  % begin
  %   write(" [i:3,'', ascii-val('', chr(i), ''')=', i:3);
  %   if (i+1) mod 3 = 0 then writeln;
  % end
  % end.
  [ 32] ascii-val (')') = 32 [ 33] ascii-val (')') = 33 [ 34] ascii-val (')') = 34
  [ 38] ascii-val ("') = 38 [ 39] ascii-val ("') = 39 [ 40] ascii-val ("') = 40
end module ASCII

C.3 Contexts

module Contexts

%% Specify the context containing all declarations found so far.
imports SyntaxHeaders

exports

sorts CONTEXT ENTRY SYMBOL VAR-STATUS

context-free syntax
   possibly-threatening -> VAR-STATUS
   control-variable -> VAR-STATUS

label LABEL -> ENTRY
const CONST-DEF -> ENTRY
type TYPE-DEF -> ENTRY
var VAR-DECL VAR-STATUS
HEADING -> ENTRY
predefined-function IDENT -> ENTRY
predefined-procedure IDENT -> ENTRY
empty-set IDENT -> ENTRY
nil-pointer IDENT -> ENTRY
block-mark  ->  ENTRY
error-entry  ->  ENTRY
IDENT       ->  SYMBOL
LABEL       ->  SYMBOL
block-mark  ->  SYMBOL
"[" { ENTRY "," }* "]"  ->  CONTEXT
"(" CONTEXT ")"  ->  CONTEXT
predefined-block  ->  CONTEXT
init-block  ->  CONTEXT
type integer  ->  CONTEXT
type real  ->  CONTEXT
type char  ->  CONTEXT
type Boolean  ->  CONTEXT
type text  ->  file of char

variables
Context[0-9]*  ->  CONTEXT
Entry[0-9]*  ->  ENTRY
[TC][0-9]*  ->  CONTEXT
Entry[0-9]*"*"  ->  ENTRY
"_symbol"[0-9]*  ->  SYMBOL
Prefix[0-9]*  ->  ENTRY
Postfix[0-9]*  ->  ENTRY

equations
[i0]  init-block = []

[i1]  predefined-block =
    [block-mark,
     predefined-procedure rewrite, predefined-procedure put,
     predefined-procedure reset, predefined-procedure get,
     predefined-procedure read, predefined-procedure write,
     predefined-procedure readln, predefined-procedure writeln,
     predefined-procedure page, predefined-procedure dispose,
     predefined-procedure unpack,
     predefined-function abs, predefined-function sqr,
     predefined-function trunc, predefined-function round,
     predefined-function ord, predefined-function succ,
     predefined-function pred, predefined-function eof,
     predefined-function chr,
     type integer = integer, type real = real,
     type char = char, type Boolean = (false, true),
     type text = file of char,
     function sin (x: real): real,
     function cos (x: real): real,
     function exp (x: real): real,
     function ln (x: real): real,
     function sqrt (x: real): real,
     function arctan (x: real): real,
     function chr (x: integer): char,
     function odd (x: integer): Boolean,
     const maxint = 32768,
    ]
]end module Contexts
module Context-defs
imports Contexts Booleans
exports
sorts ENUM-OUT FIXED-OUT VAR-OUT
context-free syntax
  CONTEXT is definition-point for SYMBOL "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a label "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a constant "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a type "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a variable "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a function "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a procedure "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a normal constant "?" -> BOOLEAN
  CONTEXT defines SYMBOL as an enumeration constant "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a normal function "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a predefined function "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a normal procedure "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a predefined procedure "?" -> BOOLEAN
  CONTEXT defines SYMBOL as a block-mark "?" -> BOOLEAN
  "<" BOOLEAN "," ENTRY ">" -> ENUM-OUT
  CONTEXT defines SYMBOL as an enum const -> ENUM-OUT
  VAR-OUT defines SYMBOL as an enum const in CONTEXT -> ENUM-OUT
  FIXED-OUT defines SYMBOL as an enum const in CONTEXT -> ENUM-OUT
  ENUM-OUT "]" ENUM-OUT -> ENUM-OUT {left}
  is-record "?" "(" CONTEXT ")" -> BOOLEAN
  CONTEXT ":" fixed -> FIXED-OUT
  CONTEXT ":" variant -> VAR-OUT
  FIXED-PART -> FIXED-OUT
  VAR-PART -> VAR-OUT
hiddens
variables
  Enum[?] -> BOOLEAN E-out -> ENUM-OUT
equations
[dp1] C is definition-point for _symbol? =
[de1] [Prefix, label _label] defines _label as a label? = TRUE
[de2] [Prefix, const _id = _const] defines _id as a normal constant? = TRUE
[de3] [Prefix, type _id = _type-den] defines _id as a type? = TRUE
[de4] [Prefix, var _id : _type-den] defines _id as a variable? = TRUE
[de5] [Prefix, function _id _form-par-list: _id2]
  defines _id as a normal function? = TRUE
defines _id as a predefined function? = TRUE

C defines _id as a function?
C defines _id as a normal function? |
C defines _id as a predefined function?

defines _id as a procedure? = TRUE
defines _id as a predefined procedure?

C defines _id as a normal procedure? |
C defines _id as a predefined procedure?

C defines _id as a constant?
C defines _id as a normal constant? |
C defines _id as an enumeration constant?

C defines _id as an enum const = <Enum?, Entry>

C defines _id as an enumeration constant? = Enum?

defines _id2 as an enum const = <TRUE, type _id1 = (_id1*, _id2, _id2*)>

defines _id1 as an enum const =

defines _id1 as an enum const =

defines _id1 as an enum const =

defines _id1 as an enum const =

defines _id1 as an enum const =

defines _id1 as an enum const =

defines _id1 as an enum const =

T1 defines _id as an enum const =
T1.fixed defines _id as an enum const in T1 | T1.variant defines _id as an enum const in T1 when is-record?(T1) = TRUE
module Context-search
defines_search_functions_on_contexts.
import
exports
sorts COMPLETE-CONTEXT
context-free
syntax
"< CONTEXT " ENTRY " CONTEXT >" -> COMPLETE-CONTEXT
add ENTRY "to" CONTEXT -> CONTEXT
delete inner block of CONTEXT -> CONTEXT
find SYMBOL in CONTEXT -> CONTEXT
SYMBOL exists in CONTEXT "?" -> BOOLEAN
SYMBOL exists in inner block of CONTEXT "?" -> BOOLEAN
equations
[al] add Entry to [Prefix] = [Prefix, Entry]
[cf0] complete find_symbol in C, error-entry, [] = <C1, Entry, C2>,
<Entry> is definition-point for _symbol? = TRUE
complete find_symbol in C = <C1, Entry, C2>
{ Avoid having to take the predefined identifiers always with you,
only look in predefined identifiers if normal search fails. )

[cf0'] complete find _symbol in <[Entry1*], error-entry, []> = 
<[], error-entry, [Entry1*]>
complete find _symbol in predefined-block = 
<[Entry2*], Entry, [Entry3*]>
==============================================================================
complete find _symbol in [Entry1*] = 
<[Entry2*], Entry, [Entry3*, Entry1*]> 

[cf1] [Entry] is definition-point for _symbol? = TRUE
==============================================================================
complete find _symbol in <[], Entry, C2> = <[], Entry, C2>

[cf2] [Entry] is definition-point for _symbol? != TRUE
==============================================================================
complete find _symbol in <[], Entry, C2> = <[], error-entry, C2>

[cf3] [Prefix, Entry] is definition-point for _symbol? = TRUE
==============================================================================
complete find _symbol in <[Prefix], Entry, [Postfix]> = 
<[Prefix], Entry, [Postfix]>

[cf4] [Prefix, Entry2, Entry1] is definition-point for _symbol? != TRUE
==============================================================================
complete find _symbol in <[Prefix, Entry2], Entry1, [Postfix]> = 
complete find _symbol in <[Prefix, Entry1, Postfix]>

[f1] complete find _symbol in C = <[Prefix], Entry, C2>
==============================================================================
find _symbol in C = [Prefix, Entry]

[d1] complete find block-mark in C = <[Prefix], Entry, C2>
==============================================================================
delete inner block of C = [Prefix]

[f11] complete find block-mark in C1 = <[Prefix], block-mark, C2>,
find _symbol in C2 = [Entry]
==============================================================================
find _symbol in inner block of C1 = [Prefix, block-mark, Entry*]

[ex1] find _symbol in C = [Prefix, Entry].
Entry != error-entry
elligence
_symbol exists in C? = TRUE

[ex2] find _symbol in inner block of C = [Prefix, Entry].
Entry != error-entry
elligence
_symbol exists in inner block of C? = TRUE

[pd1] integer-type = find integer in init-block
[pd2] boolean-type = find Boolean in init-block
[pd3] char-type = find char in init-block
module ConstantAnalysis

%% Analyze elementary constants.
imports Context-search ASCII

exports
context-free syntax
new-const-type of NEW-CONST -> CONTEXT
new-const-val of NEW-CONST -> INT
id-number of IDENT in "(" { IDENT ","+ ")" )" -> INT
is-string-const "?" "(" CONST ")" -> BOOLEAN

hiddens
context-free syntax
id-number "(" "(" IDENT ","+ ")" ")" IDENT "," INT ")" -> INT

equations
[v1] new-const-val of _uns-int = _uns-int
[v2] new-const-val of + _uns-int = _uns-int
[v3] new-const-val of - _uns-int = -1 * _uns-int
[v4] new-const-val of _uns-real = -1
[v5] new-const-val of + _uns-real = -1
[v6] new-const-val of - _uns-real = -1
[v7] new-const-val of _string = ascii-val(_string)
[t1] new-const-type of _uns-int = integer-type
[t2] new-const-type of _sign _uns-int = integer-type
[t3] new-const-type of _uns-real = real-type
[t4] new-const-type of _sign _uns-real = real-type
[t5] new-const-type of _string = char-type when length(_string) = 1

[t6] length(_string) = _uns-int.
  _uns-int != 1
  type Generated = packed array [ 1 .. _uns-int ] of char
to init-block

[i1] id-number of _id in (_id1, _id*) = id-number((_id1, _id*), _id, 0)
[i2] id-number((), _id, _uns-int) = -1
[i3] id-number((_id, _id*), _id, _uns-int) = _uns-int
[i4] id-number((_id1, _id*), _id, _uns-int) =
    id-number((_id*, _id, _uns-int + 1 )
    when _id != _id1

[i5] is-string-const?(_string) = TRUE
end module ConstantAnalysis

module Context-sen-funs

43
variables

Index-type -> TYPE_DENOTER  Comp-type -> TYPE_DENOTER
Struct-type -> STRUCTURED-TYPE  VS[0-9] -> VAR-STATUS

equations

{ Classification of new types }
[st1] is-subrange-type?([Prefix, type_id = _const1 .. _const2]) = TRUE
[st2] is-integer-type?([Prefix, type_id = (_id*))] = TRUE
[st3] is-array-type?([Prefix, type_id = _PS array[,] of _type-den]) = TRUE
[st4] is-non-empty-set-type?([Prefix, type_id = _PS set of _type-den]) = TRUE
[st5] is-empty-set-type?([Prefix, empty-set]) = TRUE
[st6] is-set-type?(T) = is-non-empty-set-type?(T) | is-empty-set-type?(T)
[st7] is-nil-pointer-type?([Prefix, type_id = _id]) = TRUE
[st8] is-nil-pointer-type?([Prefix, nil-pointer]) = TRUE
[st9] is-pointer-type?(T) = is-nil-pointer-type?(T) | is-non-nil-pointer-type?(T)
[ip1] is-packed?([Prefix, type_id = packed Struct-type]) = TRUE

{ Accessing components of types: }
[ac1] [Prefix, type_id = _PS array[,] of Comp-type].index-type =
  get-type of Index-type in [Prefix]
[ac2] [Prefix, type_id = _PS array[,] of Comp-type].comp-type =
  get-type of Comp-type in [Prefix]
[ac3] [Prefix, type_id = _PS set of Comp-type].comp-type =
  get-type of Comp-type in [Prefix]
[ac4] [Prefix, type_id = _PS file of Comp-type].comp-type =
  get-type of Comp-type in [Prefix]
[ac5] [Prefix, type_id = _const1 .. _const2].host-type =
  const-type of _const1 in [Prefix]
[ac6] pointer-component of [Prefix, type_id = _id1 .. _id2] in Context =
  get-type of _id2 in Context
[ce] [Prefix, Entry].entry = Entry

[ct1] T.canonical = T.host-type when is-subrange-type?(T) = TRUE
[ct2] T.canonical = T when is-subrange-type?(T) != TRUE

  when T = integer-type
[nol] [Prefix, type_id = _const1 .. _const2] . nr-of-elements =
  const-val of _const2 in [Prefix] -
  const-val of _const1 in [Prefix] + 1
[nol] [Prefix, type_id = _id1 .. _id2] . nr-of-elements = 1
[nol] [Prefix, type_id = (_id2, _id*)] . nr-of-elements =
  1 + [Prefix, type_id = (_id*)] . nr-of-elements
[nol] T.nr-of-elements = ascii-val(’ ’) - ascii-val(’ ’) + 1
  when T = char-type

{ Grouping types }

45
isSimpleType(T) = TRUE when T = real-type
isSimpleType(T) = TRUE when T = integer-type
isSimpleType(T) = TRUE when T = char-type
isSimpleType(T) = TRUE when T = boolean-type
isSimpleType(T) = TRUE when isSubrangeType(T) = TRUE
isSimpleType(T) = TRUE when isEnumerationType(T) = TRUE
isNumberType(T) = TRUE when T = integer-type
isStringType(T) = TRUE when isArrayType(T) = TRUE, isPacked(T) = TRUE, T.host-type = char-type
hasVariantPart([Prefix, type _id= record _var-part opt; end]) = TRUE
hasVariantPart([Prefix, type _id= record _fixed ; _var-part opt; end])
selectorField(T1, _id1) = TRUE
when T1.variant = case _id1:_id2 of _variant+
Prefix, Entry].record-of-variant-part =
[Prefix, type Generated = record _var-part end]
when [Prefix, Entry].variant = _var-part
T1.selector-type = find _id2 in T1
when T1.variant = case _id1:_id2 of _variant+
T1.selector-type = find _id in T1
when T1.variant = case _id of _variant+

get-type of _new-type in Context = [Prefix, type Generated = _new-type]
find _id in Context = [Prefix, type _id = _new-type]
get-type of _id in Context = [Prefix, type _id = _new-type]
find _id1 in Context = [Prefix, type _id1 = _id2], _id1 != _id2
get-type of _id1 in Context = get-type of _id2 in [Prefix]
find _id1 in Context = [Prefix, type _id1 = _id1]
get-type of _id1 in Context = [Prefix, type _id1 = _id1]

constType of _new-const in Context = new-const-type of _new-const
constType of _sign _id in Context = const-type of _id in Context
find _id in Context = [Prefix, const _id = _const]
const-type of _id in Context = const-type of _const in [Prefix]
find _id in Context = [Prefix, Entry].
[Prefix, Entry] defines _id as an enum const = <TRUE, type _id2 = (_id*)>
const-type of _id in Context = [Prefix, type _id2 = (_id*)]
[ctv3] const-val of + _id in Context = const-val of + id in Context

[ctv4] find _id in Context = [Prefix, const _id = _const]
  const-val of _id in Context = const-val of _const in [Prefix]

[ctv5] find _id in Context = [Prefix, Entry],
  [Prefix, Entry] defines _id as an enum const = <TRUE, type _id2= (_id*)>
  const-val of _id in Context = id-number of _id in (_id*)

[ctv6] [Prefix, const _id = _const].type = const-type of _const in [Prefix]
[ctv7] [Prefix, const _id = _const].value = const-val of _const in [Prefix]

{ Operations on variables }
[ov1] [Prefix, var _id: _type-den VS].type =
  get-type of _type-den in [Prefix]
[ov2] type-to-var([Prefix, type _id = _type-den]) =
  [Prefix, var Generated: _type-den]
[ov3] is-control-variable?(_id)
  [Prefix, var _id: _type-den control-variable] = TRUE
[ov4] is-possibly-threatening?(_id)
  [Prefix, var _id: _type-den possibly-threatening] = TRUE
[ov5] set var-status of _id in Context to VS1 =
  [Prefix, var _id: _type-den VS1, Postfix]
  when complete find _id in Context =
  <![Prefix], var _id: _type-den VS1, [Postfix]>

{ Operations on functions and procedures }
[ft1] [Prefix, function _id _form-par-list: _id2] . formal-parameters =
  _form-par-list
[pt1] [Prefix, procedure _id _form-par-list] . formal-parameters =
  _form-par-list
[ft2] [Prefix, function _id _form-par-list: _id2] . type =
  get-type of _id2 in [Prefix]
[ft1] find block-mark in Context = [Prefix, Entry, block-mark],
  [Prefix, Entry] defines _id as a function? = TRUE
  is-id the name of the inner function-block in Context? = TRUE

[pfl1] not-permissable-as-file-comp?(T) = TRUE
  when is-file-type?(T) = TRUE
[pfl2] not-permissable-as-file-comp?(T) = TRUE
when is-array-type?(T) = TRUE,
   is-file-type?(T.comp-type) = TRUE

[cs1] is-empty-set-type?(T2) = TRUE
    --------------------------
    T1 is the set-component of T2? = TRUE

[cs1] is-empty-set-type?(T2) = TRUE,
    T2.comp-type is the same as T1? = TRUE
    --------------------------
    T1 is the set-component of T2? = TRUE

{ Both types should be set-of-T. Empty set matches all other sets. }
[st1] is-empty-set-type?(T1) = TRUE,
    is-set-type?(T2) = TRUE
    ------------------
    same-set-of-T?(T1, T2) = TRUE

[st2] is-empty-set-type?(T2) = TRUE,
    is-set-type?(T1) = TRUE
    ------------------
    same-set-of-T?(T1, T2) = TRUE

[st3] is-non-empty-set-type?(T1) = TRUE,
    is-non-empty-set-type?(T2) = TRUE,
    T1.comp-type.canonical is the same as T2.comp-type.canonical? = TRUE
    -----------------------------------------------
    same-set-of-T?(T1, T2) = TRUE

{ Types are the same if they are constructed by the same NEW-TYPE operator. Two types are constructed by the same NEW-TYPE operator, if this NEW-TYPE is defined in the same place in both contexts of the type. Context equivalence can be checked by checking whether the CONTEXT-term is the same. In addition, the empty-set is the same as any other set, and the nil-pointer is the same as any other pointer. }

[sa1] T1 is the same as T1? = TRUE

[sa2] is-empty-set-type?(T1) = TRUE,
    is-set-type?(T2) = TRUE
    ------------------
    T1 is the same as T2? = TRUE

[sa3] is-empty-set-type?(T2) = TRUE,
    is-set-type?(T1) = TRUE
    ------------------
    T1 is the same as T2? = TRUE

[sa4] is-nil-pointer-type?(T1) = TRUE,
    is-pointer-type?(T2) = TRUE
    ------------------


T₁ is the same as T₂? = TRUE

[ss6] is-nil-pointer-type?(T₂) = TRUE,
     is-pointer-type?(T₁) = TRUE

T₁ is the same as T₂? = TRUE
end module Context-sem-funs

module Compatible
 XX Define compatible types.
 imports Context-sem-funs

equations

[c1] T₁ is the same as T₂? = TRUE
    T₁ and T₂ are compatible? = TRUE

[c2] is-subrange-type?(T₁) = TRUE,
     T₁.host-type is the same as T₂? = TRUE
    T₁ and T₂ are compatible? = TRUE

[c3] is-subrange-type?(T₂) = TRUE,
     T₂.host-type is the same as T₁? = TRUE
    T₁ and T₂ are compatible? = TRUE

[c4] is-subrange-type?(T₁) = TRUE,
     is-subrange-type?(T₂) = TRUE,
     T₁.host-type is the same as T₂.host-type? = TRUE
    T₁ and T₂ are compatible? = TRUE

[c5] is-non-empty-set-type?(T₁) = TRUE,
     is-non-empty-set-type?(T₂) = TRUE,
     is-packed?(T₁) = is-packed?(T₂),
     T₁.comp-type and T₂.comp-type are compatible? = TRUE
    T₁ and T₂ are compatible? = TRUE

[c6] is-empty-set-type?(T₁) = TRUE,
     is-set-type?(T₂) = TRUE
    T₁ and T₂ are compatible? = TRUE

[c7] is-empty-set-type?(T₂) = TRUE,
     is-set-type?(T₁) = TRUE
    T₁ and T₂ are compatible? = TRUE
T1 and T2 are compatible? = TRUE

[a9] is-string-type?(T1) = TRUE,  
is-string-type?(T2) = TRUE,  
T1.index-type.nr-of-elements = T2.index-type.nr-of-elements

T1 and T2 are compatible? = TRUE

T1 is the same as T2? = TRUE,  
not-permissible-as-file-comp?(T1) ! = TRUE

T2 is assignment-compatible with T1? = TRUE

T1 = real-type,  
T2 = integer-type

T1 is assignment-compatible with T2? = TRUE

T1 and T2 are compatible? = TRUE

T2 is assignment-compatible with T2? = TRUE

T2 is set-type? = TRUE,  
is-set-type?(T2) = TRUE,  
T1 and T2 are compatible? = TRUE

T2 is assignment-compatible with T2? = TRUE

C.4 Environments

module OperatorTables
XX Specify Operator Tables
imports Booleans

exports  
sorts OPR-TYPE BAR OPERATOR TABLE-HEADER TABLE-ENTRY LOOK-FOR-OUT  
OPERATOR-TABLE

lexical syntax  
[\-]+ “+” [\-]+ “+” [\-]+ “+” [\-]+  
-> BAR

c context-free syntax
"not"  ->  OPERATOR  "+,"  ->  OPERATOR  "-"  ->  OPERATOR
"*"  ->  OPERATOR  "/"  ->  OPERATOR  "div"  ->  OPERATOR
"mod"  ->  OPERATOR  "and"  ->  OPERATOR  "or"  ->  OPERATOR
"="  ->  OPERATOR  "<"  ->  OPERATOR  "<"  ->  OPERATOR
">"  ->  OPERATOR  "<="  ->  OPERATOR  ">="  ->  OPERATOR
"in"  ->  OPERATOR
int  ->  GP-TYPE  real  ->  GP-TYPE  boolean  ->  GP-TYPE
undefined  ->  OP-TYPE  set  ->  OP-TYPE  ordinal  ->  OP-TYPE
string  ->  OP-TYPE  simple  ->  OP-TYPE  pointer  ->  OP-TYPE

"Operator"  "|"  "Operand1"  "|"  "Operand2"  "|"  "Result"
BAR  ->  TABLE-HEADER
OPERATOR  "|"  GP-TYPE  "|"  GP-TYPE  "|"  GP-TYPE  ->  TABLE-ENTRY
TABLE-HEADER  TABLE-ENTRY*  ":;"  ->  OPERATOR-TABLE

look-for  "("  OPERATOR  ","
GP-TYPE  ","  GP-TYPE  ","  (OPERATOR-TABLE ")"
LOOK-FOR-OUT  ->  OperatorTable

"<"  BOOLEAN  ","  GP-TYPE  ">"
LOOK-FOR-OUT  ->  LOOK-FOR-OUT

hiddens
context-free syntax
entry-matches  "?"  "("  OPERATOR  ","  (GP-TYPE  ","  GP-TYPE  ","
OPERATOR  ","  GP-TYPE  ","  GP-TYPE  ")"
BOOLEAN  ->  BOOLEAN
variables
[[]Optor[12]*  ->  OPERATOR  [[]Op[12]*  ->  OP-TYPE
[[]Res  ->  OP-TYPE  [[]Header  ->  TABLE-HEADER
[[]OpEntries  ->  TABLE-ENTRY*

equations
[lf1]  look-for(_Optor, _Op1, _Op2, _Header; ) = <FALSE, undef>


[zal]  entry-matches?(_Optor1, _Op1, _Op2, _Optor1, _Op1, _Op2) = TRUE

end module OperatorTables

module Errors
XXX Module Errors
imports SyntaxProgram OperatorTables
exports
sorts ERROR ERRORS CONSTRUCT CONSTR-ERROR IDENT-LIST

context-free syntax
TYPE-DEF  ->  CONSTRUCT
CONST-DEF  ->  CONSTRUCT

51
VAR DECL → CONSTRUCT
HEADING → CONSTRUCT
STATEMENT → CONSTRUCT
program IDENT PROGRAM-PARAMS → CONSTRUCT

CONSTRUCT ":: ERROR → CONSTR-ERROR
CONSTR-ERROR* → ERRORS
"[ { IDENT "," ]}* ]" → IDENT-LIST

"Identifier" SYMBOL not "declared." → ERR OR
"Identifier" IDENT multiple "declared." → ERR OR
"Identifier" IDENT is not a "constant." → ERR OR
"Identifier" IDENT is not a type "." → ERR OR
"Identifier" IDENT is not a "variable." → ERR OR
"Identifier" IDENT is not a "function." → ERR OR
"Identifier" IDENT is not a "procedure." → ERR OR
"Identifier" IDENT is not a "constant," "variable," or "function." → ERR OR
"Empty string not allowed." → ERR OR
"Operator" "+" only allowed on pointer or file "variables." → ERR OR
"Non-number" types not "allowed." → ERR OR
"Types" are not the "same." → ERR OR
"Types" are not "assignment-compatible." → ERR OR
"Type must be ordinal." → ERR OR
"Type must be an array-type." → ERR OR
"Type must be a packed array." → ERR OR
"Type must be an unpacked array." → ERR OR
"Type must be a pointer." → ERR OR
"Type must be a file-type." → ERR OR
"Type must" be Boolean." → ERR OR
"Type must be a record." → ERR OR
"Type must be a variant record." → ERR OR
"Variants must be given for all possible values." → ERR OR
"Variant selector cannot be passed as var-parameter." → ERR OR
"Constants in case statement or record variant should be unique." → ERR OR
"Real expected." → ERR OR
"Integer expected." → ERR OR
"Empty" range not "allowed." → ERR OR
"No" body given for forwarded identifiers ":" IDENT-LIST "." → ERR OR
"Case" constants should be "unique." → ERR OR
"Illegal" use of control variable IDENT "." → ERR OR
"Too" many actual "parameters." → ERR OR
"Too" few actual "parameters." → ERR OR
"Reference" "To" packed component in VARIABLE-ACCESS not "allowed." → ERR OR
"Illegal use of formatted write parameter." → ERR OR
"Parameter" lists are not "congruous." → ERR OR
ACTUAL-PAR is not a file "variable." → ERR OR
"Number-type or character-type expected." → ERR OR
"Number, character, boolean or string type expected." → ERR OR
"Illegal" pointer type "definition." → ERR OR
"Illegal" file "component." → ERR OR
"Illegal" operand types for operator OPERATOR "." → ERR OR
"Type" incompatibility in "set-component." → ERR OR
"Incompatible set-types." → ERR OR
ACTUAL-PAR "is not a text file." → ERR OR
ACTUAL-PAR "is not a file variable." → ERR OR
ACTUAL-PAR ",:" must be a variable "access." → ERROR
ACTUAL-PAR ",: Function" identifier "expected." → ERROR
ACTUAL-PAR ",: Procedure" identifier "expected." → ERROR
"Types" of OPERATOR are not "compatible," "numbers," or of the same set-of-T type ",:" → ERROR

variables
Err[0-9]* → CONSTR-ERROR Err[0-9]*"*" → CONSTR-ERR*

equations
[Err] Err* Err Err2* Err Err3* = Err1* Err Err2* Err3*
end module Errors

module Environments
imports Context-search Errors

exports
sorts ENV FIELD
context-free syntax
(" context ",:" CONTEXT ")" → FIELD
(" errors ",:" ERRORS ")" → FIELD
(" result ",:" CONTEXT ")" → FIELD
(" construct ",:" CONSTRUCT ")" → FIELD
(" [ FIELD ",:" ] "* ")" → ENV
(" [ ENV ]")" → ENV bracket
new-env → ENV
ENV ",:" result → CONTEXT
ENV ",:" context → CONTEXT
ENV ",:" construct → CONSTRUCT
ENV ",:" errors → ERRORS
ENV "+:" ENTRY → ENV
leave inner block of ENV → ENV
add-error "[ " ERROR "] " to " ENV → ENV
set-result of ENV " to " CONTEXT → ENV
set-construct of ENV " to " \" CONSTRUCT \" → ENV
set-context of ENV " to " CONTEXT → ENV

variables
Env[0-9]* → ENV E[0-9]* → ENV

hiddens
variables
Field → FIELD Fields[12]* → \{FIELD ",:"\}*
[\_]\{constr\}[12]* → CONSTRUCT Error → ERROR

equations
[d2] [ Fields1. (errors, Err*), Fields2 ].errors = Err*
[d3] [ Fields1. (result, C), Fields2 ].result = C
[d4] [ Fields1. (context, C), Fields2 ].context = C
[d5] [ Fields1. (construct, _c), Fields2 ].construct = _c
[a1] set-result of Env1 to [error-entry] = Env2.
 Env2 = [Fields1, (errors, Err*), Fields2]

add-error [Error] to Env1 =
 [Fields1, (errors, Err* Env1.construct : Error), Fields2]

[a2] leave inner block of [Fields1, (context, C), Fields2] =
 [Fields1, (context, delete inner block of C), Fields2]

[a3] set-result of [Fields1, (result, C1), Fields2] to C2 =
 [Fields1, (result, C2), Fields2]

[a4] set-context of [Fields1, (context, C), Fields2] to "_c2" =
 [Fields1, (context, _c2), Fields2]

[a5] set-context of [Fields1, (context, C), Fields2] to C2 =
 [Fields1, (context, C2), Fields2]

[c1] new-env = [(construct,), (context.init-block), (errors), (result, [])] end module Environments

module ShouldBe
XX Define several basic should-be functions
XX These functions operate on an Environment. With each should-be function
XX a boolean condition is associated.
XX If the boolean condition is met, an unchanged environment is returned.
XX Otherwise, the environment with an error message added to it is returned.
imports Environments Compatible
exports
sorts VAL-LIST
context-free syntax
IDENT "should not be declared in inner block of" ENV -> ENV
IDENT "should be declared in inner block of" ENV -> ENV
SYMBOL "should be declared in" ENV -> ENV
IDENT "should be a constant in" ENV -> ENV
IDENT "should be a type in" ENV -> ENV
IDENT "should be a type in inner block of" ENV -> ENV
IDENT "should be a variable in" ENV -> ENV
IDENT "should be a variable in inner block of" ENV -> ENV
IDENT "should be a function in" ENV -> ENV
IDENT "should be a function in inner block of" ENV -> ENV
IDENT "should be a procedure in" ENV -> ENV
IDENT "should be a procedure in inner block of" ENV -> ENV
STRING "should not be an empty string in" ENV -> ENV
CONTEXT "should be allowed as file-component in" ENV -> ENV
CONTEXT "should be a number in" ENV -> ENV
CONTEXT "should be integer in" ENV -> ENV
CONTEXT "should be real in" ENV -> ENV
CONTEXT "should be Boolean in" ENV -> ENV
"(" { IDENT "," } ")" "should not have double names in" ENV -> ENV
CONTEXT and CONTEXT "should be the same type in" ENV -> ENV
CONTEXT "should be assignment-compatible with" CONTEXT in ENV -> ENV
CONTEXT "should be ordinal in" ENV -> ENV

54
CONTEXT "should be an array in" ENV -> ENV
CONTEXT "should be a packed array in" ENV -> ENV
CONTEXT "should be an unpacked array in" ENV -> ENV
CONTEXT "should be a pointer in" ENV -> ENV
CONTEXT "should be a file in" ENV -> ENV
INT "should be less than" INT in ENV -> ENV
CONTEXT "should be the set-component of" CONTEXT in ENV -> ENV
CONTEXT and CONTEXT "should be the same set-of-T type in" ENV -> ENV
CONTEXT OPERATOR CONTEXT "should be compatible," "numbers,"
  "or of the same set in" ENV -> ENV
CONTEXT should be a pointer or a file in ENV -> ENV
CONTEXT "should be a record in" ENV -> ENV
CONTEXT "should be a variant record in" ENV -> ENV
all elements of CONTEXT should occur in VAL-LIST in ENV -> ENV
INT should not be a member of VAL-LIST in ENV -> ENV
"[ { INT ",," }* ]" -> VAL-LIST

variables
Value[0-9]* -> INT Value[0-9]*"*" -> { INT ",," }*
Val-list[0-9]* -> VAL-LIST

hiddens
context-free syntax
check double names of "( { IDENT ",," }* )" in ENV -> ENV
length "( VAL-LIST ")" -> INT
INT is member of VAL-LIST "?" -> BOOLEAN

variables
Optor -> OPERATOR V[0-9]* -> INT

equations

[md1] _id exists in inner block of Env.context? !=" TRUE

                        ============================== 
                        _id should not be declared in inner block of Env = Env

[md2] _id exists in inner block of Env.context? = " TRUE

                        ============================== 
                        _id should not be declared in inner block of Env = Env
                        add-error [Identifier _id multiple declared.] to Env

[md3] _id exists in inner block of Env.context? = " TRUE

                        ============================== 
                        _id should be declared in inner block of Env = Env

[md4] _id exists in inner block of Env.context? !=" TRUE

                        ============================== 
                        _id should be declared in inner block of Env = Env
                        add-error [Identifier _id not declared.] to Env

{ If _id is not yet defined, raise error "Identifier not declared".
  Otherwise set result to corresponding entry. }
[f1] find _symbol in Env.context = [Prefix, Entry],
    Entry != error-entry
    _symbol should be declared in Env =
    set-result of Env to [Prefix, Entry]

[f2] find _symbol in Env.context = [error-entry]
    _symbol should be declared in Env =
    add-error [Identifier _symbol not declared.] to Env

[co1] _id should be declared in E1 = E2,
    E2.result defines _id as a constant? = TRUE
    _id should be a constant in E1 = E2

[co2] _id should be declared in E1 = E2,
    E2.result defines _id as a constant? != TRUE
    _id should be a constant in E1 =
    add-error [Identifier IDENT is not a constant.] to E2

[ty1] _id should be declared in E1 = E2,
    E2.result defines _id as a type? = TRUE
    _id should be a type in E1 = E2

[ty2] _id should be declared in E1 = E2,
    E2.result defines _id as a type? != TRUE
    _id should be a type in E1 =
    add-error [Identifier IDENT is not a type.] to E2

[ty3] _id should be a type in inner block of E =
    _id should be a type in (_id should be declared in inner block of E)

[va1] _id should be declared in E1 = E2,
    E2.result defines _id as a variable? = TRUE
    _id should be a variable in E1 = E2

[va2] _id should be declared in E1 = E2,
    E2.result defines _id as a variable? != TRUE
    _id should be a variable in E1 =
    add-error [Identifier IDENT is not a variable.] to E2

[va3] _id should be a variable in inner block of E =
    _id should be a variable in (_id should be declared in inner block of E)

[fu1] _id should be declared in E1 = E2,
    E2.result defines _id as a function? = TRUE
_id should be a function in E1 = E2

[fu2] _id should be declared in E1 = E2,
E2.result defines _id as a function? != TRUE
===============================================
_id should be a function in E1 =
add-error [Identifier IDENT is not a function.] to E2

[fu3] _id should be a function in inner block of E =
_id should be a function in (_id should be declared in inner block of E)

[pr1] _id should be declared in E1 = E2,
E2.result defines _id as a procedure? != TRUE
===============================================
_id should be a procedure in E1 = E2

[pr2] _id should be declared in E1 = E2,
E2.result defines _id as a procedure? != TRUE
===============================================
_id should be a procedure in E1 =
add-error [Identifier IDENT is not a procedure.] to E2

[pr3] _id should be a procedure in inner block of E =
_id should be a procedure in (_id should be declared in inner block of E)

[ft1] not-permissible-as-file-comp?(T) != TRUE
===============================================
T should be allowed as file-component in E = E

[ft2] not-permissible-as-file-comp?(T) = TRUE
===============================================
T should be allowed as file-component in E =
add-error [Illegal file component.] to E

[cm1] T should be a number in E = E when is-number-type?(T) = TRUE
[cm2] T should be a number in E =
add-error [Non-number types not allowed.] to E
when is-number-type?(T) != TRUE

[bt1] T should be Boolean in E = E when T = boolean-type
[bt2] T should be Boolean in E = add-error [Type must be Boolean.] to E
when T != boolean-type

[in1] T should be integer in E = E when T = integer-type
[in2] T should be integer in E = add-error [Integer expected.] to E
when T != integer-type

[ra1] T should be real in E = E when T = real-type
[ra2] T should be real in E = add-error [Real expected.] to E
when T != real-type

[dn1] E1 + block-mark = E2,
check-double-names of (_id*) in E2 = E3,
leave inner block of E3 = E4
===============================================
57
(_id*) should not have double names in E1 = E4

[dn2] check-double-names of (_) in E = E
[dn3] check-double-names of (_id, _id*) in E1 = E
    check-double-names of (_id*) in
    ((_id should not be declared in inner block of E1) + var _id: dummy)

[ts1] T1 and T2 should be the same type in E = E
    when T1 is the same as T2? = TRUE
[ts2] T1 and T2 should be the same type in E =
    add-error [Types are not the same.] to E
    when T1 is the same as T2? != TRUE

[ac1] T1 should be assignment-compatible with T2 in E = E when
    T1 is assignment-compatible with T2? = TRUE
[ac2] T1 should be assignment-compatible with T2 in E =
    add-error [Types are not assignment-compatible.] to E
    when T1 is assignment-compatible with T2? != TRUE

[or1] T should be ordinal in E = E when is-ordinal-type?(T) = TRUE
[or2] T should be ordinal in E = add-error [Type must be ordinal.] to E
    when is-ordinal-type?(T) != TRUE

[ar1] T should be an array in E = E when is-array-type?(T) = TRUE
[ar2] T should be an array in E = add-error [Type must be an array-type.] to E
    when is-array-type?(T) != TRUE
[ar3] T should be a packed array in E = E
    when is-array-type?(T) & is-packed?(T) = TRUE
[ar4] T should be a packed array in E =
    add-error [Type must be a packed array.] to E
    when is-array-type?(T) & is-packed?(T) != TRUE
[ar5] T should be an unpacked array in E = E
    when is-array-type?(T) = TRUE, is-packed?(T) != TRUE
[ar6] T should be an unpacked array in E =
    add-error [Type must be an unpacked array.] to E
    when is-array-type?(T) = TRUE, is-packed?(T) != TRUE

[re1] T should be a record in E = E when is-record?(T) = TRUE
[re2] T should be a record in E = add-error [Type must be a record.] to E
    when is-record?(T) != TRUE
[re3] T should be a variant record in E = E when has-variant-part?(T) = TRUE
[re4] T should be a variant record in E =
    add-error [Type must be a variant record.] to E
    when has-variant-part?(T) != TRUE

[fi1] T should be a file in E = E when is-file-type?(T) = TRUE
[fi2] T should be a file in E = add-error [Type must be a file-type.] to E
    when is-file-type?(T) != TRUE

[fi3] T should be a pointer in E = E when is-pointer-type?(T) = TRUE
[fi4] T should be a pointer in E = add-error [Type must be a pointer.] to E
    when is-pointer-type?(T) != TRUE
V1 should be less than V2 in E = E when V1 <= V2 = TRUE

V1 should be less than V2 in E =
    add-error [Empty range not allowed.] to E
    when V1 <= V2 != TRUE

T1 should be the set-component of T2 in E = E
    when T1 is the set-component of T2? = TRUE

T1 should be the set-component of T2 in E =
    add-error [Type incompatibility in set-component.] to E
    when T1 is the set-component of T2? != TRUE

T1 and T2 should be the same set-of-T type in E = E
    when same-set-of-T?(T1, T2) = TRUE

T1 and T2 should be the same set-of-T type in E =
    add-error [Incompatible set-types.] to E
    when same-set-of-T?(T1, T2) != TRUE

T1 and T2 are compatible?
    ( is-number-type?(T1) & is-number-type?(T2) ) |
    same-set-of-T?(T1, T2) = TRUE
====
T1 Optor T2 should be compatible, numbers, or of the same set in E = E

T1 and T2 are compatible?
    ( is-number-type?(T1) & is-number-type?(T2) ) |
    same-set-of-T?(T1, T2) != TRUE
====
T1 Optor T2 should be compatible, numbers, or of the same set in E =
    add-error [Types of Optor are not compatible, numbers, or of the
              same set-of-T type.] to E

is-pointer-type?(T1) & is-file-type?(T1) = TRUE
====
T1 should be a pointer or a file in E = E

is-pointer-type?(T1) & is-file-type?(T1) != TRUE
====
T1 should be a pointer or a file in E =
    add-error [Operator only allowed on pointer or file variables.] to E

length(_string) != 0
====
_string should not be an empty string in E = E

length(_string) = 0
====
_string should not be an empty string in E =
    add-error [Empty string not allowed.] to E

Value is member of [Value1*, Value, Value2*] = TRUE

length([]) = 0
[value length([Value, Value*]) = length([Value*]) + 1

Value should not be a member of Val-list in E = E
    when Value is member of Val-list? != TRUE
Value should not be a member of Val-list in E = add-error
 Constants in case statement or record variant should be unique.] to E

all elements of T1 should occur in Val-list in E = E
when length(Val-list) = T1.nr-of-elements

all elements of T1 should occur in Val-list in E = add-error
Variants must be given for all possible values.] to E
when length(Val-list) ! T1.nr-of-elements

end module ShouldBe

C.5 Type Checking

module TC-defs
%% Type check the definitions and declarations of a Pascal Program.
imports ShouldBe

exports
sorts VAL-OUT
context-free syntax
label-decs-tc "(" {LABEL","}*,", ENV ")" -> ENV
const-decs-tc "(" {CONST-DEF ";",}*,", ENV ")" -> ENV
const-tc "(" CONST ",", ENV ")" -> ENV
type-tc "(" {TYPE-DEF ";",}*,", ENV ")" -> ENV
var-tc "(" VAR-DECL ";",ENV ")" -> ENV
new-block-wth-fields "(" CONTEXT ",", ENV ")" -> ENV
const-val-list-tc "(" {CONST ",",}*,", CONTEXT ",", VAL-LIST ",", ENV ")"
"<" VAL-LIST ",", ENV ")" -> VAL-OUT
"<" VAL-LIST ",", ENV ")" -> VAL-OUT

hiddens
sorts PTR-OUT ID-LIST PTR-VAL-OUT
context-free syntax
"(" { IDENT ",",}*,")" -> ID-LIST
"<" ID-LIST ",", ENV ")" -> PTR-OUT
type-ptr-tc "(" { TYPE-DEF ";",}*,", ENV ")" -> PTR-OUT
type-denot-tc "(" TYPE-DENOTER ",", ENV ")" -> PTR-OUT
fields-tc "(" FIELD-LIST ",", ENV ")" -> PTR-OUT
fixed-tc "(" FIXED-PART ",", ENV ")" -> PTR-OUT
variant-tc "(" VAR-PART ",", ENV ")" -> PTR-OUT
"<" ID-LIST ",", VAL-LIST ",", ENV ")" -> PTR-VAL-OUT
variant-tc "(" {VARIANT ",",}*,", CONTEXT ",", VAL-LIST ",", ENV ")" -> PTR-VAL-OUT
selector-tc "(" SELECTOR ",", ENV ")" -> ENV
check-pointer-identifiers "(" ID-LIST ",", ENV ")" -> ENV
ID-LIST should be empty in ENV -> ENV

variables
V[1..23]* -> INT Component-type -> TYPE-DENOTER
Ptr-list[0..9] -> ID-LIST Index-type[1..23]* -> TYPE-DENOTER
equations

{ Label declarations }
[l1] label-decls-tc(E) = E
[l2] label-decls-tc(_label, _label*, E) =
    label-decls-tc(_label*, label-decls-tc(_label, E))
[l3] label-decls-tc(_label, E) = E + label _label

{ Constant declarations }
[c1] const-defs-tc(E) = E
[c2] const-defs-tc(Const-def, Const-defs, E) =
    const-defs-tc(Const-def, const-defs-tc(Const-def, E))
[c3] set-construct of E1 to "_id = _const" = E2,
    _id should not be declared in inner block of E2 = E3,
    const-tc(_const, E3) = E4
    -----------------------------------------------
    const-defs-tc(_id = _const, E1) = E4 + const _id = _const
[c4] const-tc(_string, E) = _string should not be an empty string in E
[c5] const-tc(_new-const, E) = E when is-string-const(_new-const) != TRUE
[c6] const-tc(_id, E) = _id should be a constant in E
[c7] _id should be a constant in E1 = E2,
    (const-type of _id in E2, context) should be a number in E2 = E3
    -----------------------------------------------
    const-tc(_sign _id, E1) = E3

{ Type declarations }
[t1] type-ptrs-tc(Type-defs, E1) = <Ptr-list, E2>,
    check-pointer-identifiers(Ptr-list, E2) = E3
    -----------------------------------------------
    type-defs-tc(Type-defs, E1) = E3
[t2] type-ptrs-tc(, E) = <[], E>
[t3] type-ptrs-tc( Type-def, E1 ) = <[Ptr-ids], E2>,
    type-ptrs-tc( Type-defs, E2 ) = <[Ptr-ids2], E3>
    -----------------------------------------------
    type-ptrs-tc( Type-def, Type-defs, E1 ) = <[Ptr-ids1, Ptr-ids2], E3>
[t4] set-construct of E1 to "_id = _type-den" = E2,
    _id should not be declared in inner block of E2 = E3,
    type-denot-tc(_type-den, E3) = <Ptr-list, E4>
    -----------------------------------------------

61
type-ptrs-tc(_id = _type-den, E1) =
<Ptr-list, E4 + type_id = _type-den>

{ Identifier as type-denoter }
[t5] type-denot-tc(_id, E) = <[], _id should be a type in E>

{ Enumeration type constructor }
[te1] type-denot-tc(_id+), E1) = <[], (_id+) should not have double names in E1>

{ Subrange type constructor }
[ts1] const-tc(_const1, E1) = E2,
    const-tc(_const2, E2) = E3,
    const-type of _const1 in E3.context = T1,
    const-type of _const2 in E3.context = T2,
    T1 and T2 should be the same type in E3 = E4.
    T1 should be ordinal in E4 = E5.
    (const-val of _const1 in E3.context) should be less than
    (const-val of _const2 in E3.context) in E5 = E6
    const-type of _const1 .. _const2, E1) = <[], E6>

{ Array type constructor }
[ar1] _PS array [Index-type1, Index-types+] of Component-type =
    _PS array [Index-type1] of
    _PS array [Index-types+] of Component-type

[ar2] type-tc( Index-type, E1) = E2,
    type-denot-tc( Component-type, E2 ) = <Ptr-list2, E3>,
    E2.result should be ordinal in E3 = E4
    type-denot-tc( _PS array [ Index-type ] of Component-type , E1 ) =
<Ptr-list2, E4>

{ Set type constructor }
[se1] type-tc( Component-type, E1 ) = E2,
    E2.result should be ordinal in E2 = E3
    type-denot-tc( _PS set of Component-type, E1) = <[], E3>

{ File type constructor }
[fi1] type-denot-tc( Component-type, E1) = <Ptr-list, E2>,
    E1.result should be allowed as file-component in E2 = E3
    type-denot-tc( _PS file of Component-type, E1) = <Ptr-list, E3>

{ Pointer type constructor }
[pt1] (find _id in E1.context) defines _id as a type? = TRUE
    type-denot-tc( ~ _id, E1) = <[], E1>

[pt1] (find _id in E1.context) defines _id as a type? != TRUE
    type-denot-tc( ~ _id, E1) = <[_id], E1>

62
{ Record type constructor }

[re1] fields-tc(_field-list, E1 + block-mark ) = <Ptr-list, E2>,
    leave inner block of E2 = E3
==================================================================================================================================
type-denot-tc(_PS record _field-list end, E1 ) = <Ptr-list, E3>

[re2] fields-tc(_field-list, E1 + block-mark ) = <Ptr-list, E2>
==================================================================================================================================
new-block-with-fields([Prefix, type _id = record _field-list end], E1) = E2

[re3] fields-tc(, E) = <[], E>
[re4] fields-tc(_fixed opt;:, E) = fixed-tc(_fixed, E)
[re5] fields-tc(_var-part opt;:, E) = variant-tc(_var-part, E)
[re6] fixed-tc(_fixed, E1 ) = {[Ptr-ids1], E2},
    variant-tc(_var-part, E2 ) = {[Ptr-ids2], E3}
==================================================================================================================================
fields-tc(_fixed; _var-part opt;:, E1 ) = {[Ptr-ids1], E2},

[fx1] type-denot-tc(_type-den, E1 ) = <Ptr-list, E2>,
    var-decls-tc(_id+: _type-den, E2 ) = E3
==================================================================================================================================
fixed-tc(_id+: _type-den, E1 ) = <Ptr-list, E3>

[fx2] fixed-tc(_rec-section, E1 ) = {[Ptr-ids1], E2},
    fixed-tc(_rec-section+, E2 ) = {[Ptr-ids2], E3}
==================================================================================================================================
fixed-tc(_rec-section; _rec-section+, E1 ) = {[Ptr-ids1, Ptr-ids2], E3}

[vp1] selector-tc(_selector, E1 ) = E2,
    E2.result.type should be ordinal in E2 = E3,
    variant-tc(_variant*, E2.result.type, [], E3) = <Ptr-list, Val-list, E4>,
    all elements of E2.result.type should occur in Val-list in E4 = E5
==================================================================================================================================
variant-tc( case _selector of _variant*, E1 ) = <Ptr-list, E5>

[vp2] selector-tc(_id, E ) = type-tc(_id, E )
[vp3] selector-tc(_id : _id2, E ) = var-decls-tc(_id : _id2, E )

[val1] const-val-list-tc(_const+, T1, Val-list1, E1 ) = <Val-list2, E2>,
    fields-tc(_field-list, E2 ) = <Ptr-list, E3>
==================================================================================================================================
variant-tc(_const+: _field-list, T1, Val-list1, E1 ) = <Ptr-list, Val-list2, E3>

[val2] variant-tc(_variant, T1, Val-list1, E1 ) = {[Ptr-ids1], Val-list2, E2},
    variant-tc(_variant*, T1, Val-list2, E2 ) = {[Ptr-ids2], Val-list3, E3}
==================================================================================================================================
variant-tc(_variant; _variant*, T1, Val-list1, E1 ) =
    {[Ptr-ids1, Ptr-ids2], Val-list3, E3}

{ type-tc may be called when no not-yet-existing pointer identifiers are assumed to be used within in the type-denoter, e.g. in the var-decls. }
[ty1] type-denot-tc(_type-den, E1 ) = <Ptr-list, E2>,

63
Ptr-list should be empty in \( E_2 = E_3 \)

\[ \text{type-tc}(_\text{type-den}, E_1) = \text{set-result of } E_3 \text{ to get-type of } _\text{type-den} \text{ in } E_3.\text{context} \]

[pP1] \([]\) should be empty in \( E = E \)

[pP2] \([\_id\*]\) should be empty in \( E = \text{add-error} \) [Illegal pointer type definition.] to \( E \)

[cp1] check-pointer-identifiers([], E) = E

[cp2] check-pointer-identifiers([\_id], E) =

  \_id should be a type in inner block of \( E \)

[cp3] check-pointer-identifiers([\_id, \_id\*], E) =

  check-pointer-identifiers([\_id],
  check-pointer-identifiers([\_id\*], E))

{ Variable declarations }

[v1] var-decls-tc(, E) = E

[v2] var-decls-tc(Var-decl; Var-decls, E) =

  var-decls-tc(Var-decls, var-decls-tc(Var-decl, E))

[v3] var-decls-tc(\_id, \_id\*; _\text{type-den}, E) =

  var-decls-tc(\_id; _\text{type-den}; \_id\*; _\text{type-den}, E)

[v4] set-construct of \( E_1 \) to "\_id : _\text{type-den}" = \( E_2 \),

  \_id should not be declared in inner block of \( E_2 = E_3 \),

  \text{type-tc}(_\text{type-den}, E_3) = E_4

  \text{-----------------------------}

  var-decls-tc(\_id; _\text{type-den}, E_1) = E_4 + \_id : _\text{type-den}

[cv1] const-tc(\_const, E_1) = E_2,

  \( T_1 \) and \( E_2.\text{result.type} \) should be the same type in \( E_2 = E_3 \),

  \( E_2.\text{result.value} \) should not be a member of [Value\*] in \( E_3 = E_4 \)

  \text{------------------------------}

  const-val-list-tc(\_const, \_const; \_const; T_1, Val-list1, E_1) = <Val-list2, E_2>,

  const-val-list-tc(\_const; \_const; T_1, Val-list2, E_2) = <Val-list3, E_3>

  \text{------------------------------}

  const-val-list-tc(\_const, \_const; T_1, Val-list1, E_1) = <Val-list3, E_3>

end module TC-defs

module Expr-misc

% Several general functions needed for typechecking expressions.

imports TC-defs

exports

context-free syntax

expr-tc "(" EXPR "\" ENV ")" -> ENV
exports

sorts DESIGNATOR

context-free syntax

value → DESIGNATOR
var → DESIGNATOR
procedure → DESIGNATOR
function → DESIGNATOR

FORMAT should be an empty format in ENV → ENV

act-par-tc "{ DESIGNATOR ACTUAL-PAR "," ENV "}" → ENV

act-par-tc "{ DESIGNATOR ACTUAL-PAR "," ENV "}" → ENV

context-free syntax

value → DESIGNATOR
var → DESIGNATOR
procedure → DESIGNATOR
function → DESIGNATOR

FORMAT should be an empty format in ENV → ENV

act-par-tc "{ DESIGNATOR ACTUAL-PAR "," ENV "}" → ENV

is-simple"expr" "{ IDENT "," VAR-STATUS "," ENV "}" → ENV

is-var"access" "{ ACTUAL-PAR "," CONTEXT "}" → BOOLEAN

is-function"call" "{ EXPR "," CONTEXT "}" → BOOLEAN

is-a-function-identifier "{ IDENT "," CONTEXT "}" → BOOLEAN

is-a-procedure-identifier "{ IDENT "," CONTEXT "}" → BOOLEAN

is-packed "{ ACTUAL-PAR "," CONTEXT "}" → BOOLEAN

is-just-an-identifier "{ ACTUAL-PAR "," CONTEXT "}" → BOOLEAN

is-file-variable "{ ACTUAL-PAR "," CONTEXT "}" → BOOLEAN

is-text-file "{ ACTUAL-PAR "," ENV "}" → BOOLEAN

is-char-or-number "{ CONTEXT "}" → BOOLEAN

is-bool-or-string "{ CONTEXT "}" → BOOLEAN

act-par-is-selector "{ ACTUAL-PAR "," ENV "}" → BOOLEAN

is-file-variable "{ ACTUAL-PAR "," CONTEXT "}" → BOOLEAN

is-text-file "{ ACTUAL-PAR "," ENV "}" → BOOLEAN

is-char-or-number "{ CONTEXT "}" → BOOLEAN

is-bool-or-string "{ CONTEXT "}" → BOOLEAN

no { ACTUAL-PAR "," ENV }* should be given in ENV → ENV

hiddens

context-free syntax

is-potential-control-variable "{ IDENT "," CONTEXT "}" → BOOLEAN

is-formatted"par-tc" "{ DESIGNATOR EXPR "," ENV "}" → ENV

variables

Definition → CONTEXT Status → VAR-STATUS Designator → DESIGNATOR

equations

{ Dealing with control variables. }

65
[tr0] take care of possibly threatening variable \_act-par in E = E when is-just-an-identifier?(_act-par) != TRUE

[tr1] is-potential-control-variable?(_id, E1.context) != TRUE
  ================
  take care of possibly threatening variable _id in E1 = E1

[tr2] is-potential-control-variable?(_id, E1.context) = TRUE,
  _id should not be a control variable in E1 = E2,
  mark-variable(_id, possibly-threatening, E2) = E3
  ================
  take care of possibly threatening variable _id in E1 = E3

[pc1] _id exists in inner block of Context? != TRUE,
  find _id in Context = Definition,
  Definition defines _id as a variable? = TRUE,
  is-ordinal-type?(Definition.type) = TRUE
  ================
  is-potential-control-variable?(_id, Context) = TRUE

[ic1] find _id in E.context = Definition,
  is-control-variable?(Definition) = TRUE
  ================
  _id should not be a control variable in E =
  add-error [Illegal use of control variable _id.] to E

[ic2] find _id in E.context = Definition,
  is-control-variable?(Definition) != TRUE
  ================
  _id should not be a control variable in E = E

[ic5] find _id in E.context = Definition,
  is-possibly-threatening?(Definition) = TRUE
  ================
  _id should not be a possibly threatening variable in E =
  add-error [Illegal use of control variable _id.] to E

[ic6] find _id in E.context = Definition,
  is-possibly-threatening?(Definition) = TRUE
  ================
  _id should not be a possibly threatening variable in E = E

[st1] mark-variable(_id, Status, E) =
  set-context of E to (set var-status of _id in E.context to Status)

{ Actual parameters }
[ap1] non-formatted-par-tc( Designator _expr, E1 ) = E2,
  _format1 should be an empty format in E2 = E3
  ================
  act-par-tc( Designator _expr _format1 , E1 ) = E3

66
non-formatted-par-tc( value _expr, E1 ) = expr-tc( _expr, E1 )

var-access-tc( _expr, E1 ) = E2
non-formatted-par-tc( var _expr, E1 ) = set-result of E2 to E2.result.type

non-formatted-par-tc( procedure _expr, E ) =
_expr should be a procedure identifier in E

non-formatted-par-tc( function _expr, E ) =
_expr should be a function identifier in E

should be an empty format in E = E
_format should be an empty format in E =
add-error [Illegal use of formatted write parameter.] to E
when _format !=

{ classification of expressions }

is-just-an-identifier?( _id ) = TRUE

is-simple-exp?( _uns-number, Context ) = TRUE
is-simple-exp?( _string, Context ) = TRUE
is-simple-exp?( _nil, Context ) = TRUE
is-simple-exp?( _member*, Context ) = TRUE
is-simple-exp?( _id, Context ) = TRUE
when (find _id in Context) defines _id as a constant? = TRUE

is-var-access?( _id, Context ) = TRUE
when (find _id in Context) defines _id as a variable? = TRUE
is-var-access?( _var-acc, Context ) = TRUE
when is-just-an-identifier?( _var-acc ) != TRUE

_expr should be a variable-access in E = E
when is-var-access?( _expr, E.context ) = TRUE

_expr should be a variable-access in E =
add-error [Expr: must be a variable access.] to E
when is-var-access?( _expr, E.context ) != TRUE

is-function-call?( _id, Context ) = TRUE
when (find _id in Context) defines _id as a function? = TRUE

is-a-procedure-identifier?(_id, Context) = TRUE
when (find _id in Context) defines _id as a procedure? = TRUE
is-a-function-identifier?(_id, Context) = TRUE
when (find _id in Context) defines _id as a function? = TRUE

_id should be a procedure identifier in E =
set-result of E to (find _id in E.context)
when is-a-procedure-identifier?(_id, E.context) = TRUE

_expr should be a procedure identifier in E =
add-error [Expr: Procedure identifier expected.] to E
when is-a-procedure-identifier?(_expr, E.context) != TRUE

67
_id should be a function identifier in E =
set-result of E to (find _id in E.context)
when is-a-procedure-identifier?(_id, E.context) = TRUE

_expr should be a function identifier in E =
add-error [Expr: Function identifier expected.] to E
when is-a-procedure-identifier?(_expr, E.context) != TRUE

is-packed?(_id, Context) = is-packed?(find _id in Context)
is-packed?(_var-acc [_expr ] , Context ) =
is-packed?(_var-acc . _id, Context ) =
is-packed?(_var-acc , Context )

component-of-packed?(_var-acc , _expr , Context ) =
is-packed?(_var-acc , Context )

component-of-packed?(_var-acc . _id, Context ) =
is-packed?(_var-acc , Context )

component-of-packed?(_var-acc , Context ) =
is-packed?(_var-acc , Context )

component-of-packed?(_act-par, E.context ) != TRUE

_act-par should not be a packed component in E = E

component-of-packed?(_var-acc, E.context ) = TRUE

_var-acc should not be a packed component in E =
add-error [Reference to packed component in _var-acc not allowed.] to E

no should be given in E = E

no _act-par* should be given in E =
add-error [Too many actual parameters.] to E

act-par-tc(var _act-par, E1) = E2,
E2.result should be a file in E2 = E3

_act-par should be a file variable in E1 = E3

is-file-variable?(_act-par, E) =
is-var-access?(_act-par, E.context) &
is-file-type?act-par-tc(var _act-par, E).result

char-or-number?(T) = TRUE when T = char-type
char-or-number?(T) = TRUE when is-number-type?(T) = TRUE

T should be a character or a number in E = E
when is-char-or-number?(T) = TRUE

T should be a character or a number in E =
add-error [Number-type or character-type expected.] to E
when is-char-or-number?(T) != TRUE

is-bool-or-string?(T) = TRUE when T = boolean-type
is-bool-or-string?(T) = TRUE when is-string-type?(T) = TRUE

T should be a number, character, bool, or string in E = E when is-bool-or-string?(T) | is-char-or-number?(T) = TRUE

T should be a number, character, bool, or string in E = add-error [Number, character, boolean or string type expected.] to E when is-bool-or-string?(T) | is-char-or-number?(T) = TRUE

is-text-file?(_act-par, E) = TRUE when is-var-access?(_act-par, E.context) = TRUE.
act-par-tc(var _act-par, E).result = text-type

_act-par should be a text file variable in E = E when is-text-file?(_act-par, E) = TRUE

_act-par should be a text file variable in E = add-error [_act-par is not a text file.] to E when is-text-file?(_act-par, E) != TRUE

var-access-tc(_var-acc, E1) = E2,
selector-field?(E2.result, _id1) = TRUE

_act-par should not be a variant selector field in Env = Env when act-par-is-selector?(_act-par, Env) != TRUE

_act-par should not be a variant selector field in Env = add-error [Variant selector cannot be passed as var-parameter.] to Env when act-par-is-selector?(_act-par, Env) = TRUE

end module Expr-misc

module TC-expr-simple
XX Type check simple expressions: Numbers, Strings, Nil, Set-constructor
imports Expr-misc

hiddens
  context-free syntax
  member-tc "(" MEMBER-DESIGNATOR "," ENV ")" -> ENV

equations

sim-expr-tc(_uns-number, E) = set-result of E to new-const-type of _uns-number

sim-expr-tc(_string, E) = set-result of E to new-const-type of _string

_id should be a constant in E1 = E2

sim-expr-tc(_id, E1) = set-result of E2 to const-type of _id in E1.context

sim-expr-tc(nil, E) = set-result of E to [nil-pointer]

sim-expr-tc([], E) = set-result of E to [empty-set]
member-tc([_member], E1) = E2,
sim-expr-tc([_member*], E2) = E3,
E2.result and E3.result should be the same type in E3 = E4

sim-expr-tc([_member].[_member*], E1) =
  set-result of E4 to E2.result

member-tc(_expr, E) = expr-tc(_expr, E)

expr-tc(_expr1, E1) = E2,
expr-tc(_expr2, E2) = E3,
E2.result and E3.result should be the same type in E3 = E4,
E2.result should be ordinal in E4 = E5

member-tc(_expr1 .. _expr2, E1) = E5

end module TC-expr-simple

module TC-expr-vars

%% Typecheck a variable access.

imports Expr-misc

hiddens

context-free syntax

var-arrow-tc(" VARIABLE-ACCESS ", " ENV " ) -> ENV

equations

{ var-access-tc returns an environment. The result of the environment is set to a variable definition of the appropriate type.
Type checking an expression consisting of only that variable-access will transform that variable definition into a type definition. }

var-access-tc(_id, E) = _id should be a variable in E

var-access-tc(_var-acc[_expr, _expr+], E) =
  var-access-tc(_var-acc[_expr] [ _expr+ ], E)

var-access-tc(_var-acc, E1) = E2,
expr-tc(_expr, E2) = E3,
E2.result.type should be an array in E3 = E4,
E3.result should be assignment-compatible with E2.result.type.index-type in E4 = E5

var-access-tc(_var-acc [ _expr], E1) =
  set-result of E5 to type-to-var(E2.result.type.comp-type)

var-access-tc(_var-acc, E1) = E2,
  is-file-type?( E2.result.type ) = TRUE

var-access-tc(_var-acc, E1) =
  set-result of E2 to type-to-var(E2.result.type.comp-type)

var-access-tc(_var-acc, E1) = E2,
  is-pointer-type?( E2.result.type ) = TRUE

end module
module TC-expr-vars

% Type check procedure and function variables.

hiddens

context-free syntax

context-free syntax

variables

equations

[p1] find _id in E1.context = Definition.
    Definition defines _id as a predefined procedure? = TRUE,
    spec-proc-tc(_id _act-par-list, E1) = E2
    set-result of E2 to E1.result.type
    proc-call-tc(_id _act-par-list, E1) = E2

[p2] find _id in E1.context = Definition.
    Definition defines _id as a normal procedure? = TRUE,
    call-tc( Definition.formal-parameters, _act-par-list, E1) = E2
    set-result of E2 to type-to-var( pointer-component of E2.result.type in E1.context)
    proc-call-tc(_id _act-par-list, E1) = E2

[f1] find _id in E1.context = Definition.
    Definition defines _id as a predefined function? = TRUE,
    spec-fun-tc(_id, E1) = E2

end module TC-expr-vars
fun-call-tc(_id, E1) = E2

[f2] find _id in E1.context = Definition,
    Definition defines _id as a predefined function? = TRUE,
    spec-fun-tc(_id _NE-act-par-list, E1) = E2
    ============
    fun-call-tc(_id _NE-act-par-list, E1) = E2

[f3] find _id in E1.context = Definition,
    Definition defines _id as a normal function? = TRUE,
    call-tc( Definition.formal-parameters, _NE-act-par-list, E1) = E2
    ============
    fun-call-tc(_id _NE-act-par-list, E1) =
        set-result of E2 to Definition.type.canonical

[f4] find _id in E1.context = Definition,
    Definition defines _id as a normal function? = TRUE,
    call-tc( Definition.formal-parameters, , E1) = E2
    ============
    fun-call-tc(_id, E1) = set-result of E2 to Definition.type.canonical

[c0] call-tc( , E ) = E

[c1] call-tc( (_form-par), , E) = add-error
    [Too few actual parameters] to E

[c2] call-tc( , (_act-par+), E) = add-error
    [Too many actual parameters] to E

[c3] call-tc( (_form-par; _form-par), (_act-par, _act-par), E) =
    call-tc( (_form-par), (_act-par), E )
    call-tc( (_form-par), (_act-par), E )

[c4] act-par-tc(value _act-par, E1) = E2,
    type-tc(_id2, E2) = E3,
    E2.result should be assignment-compatible with E3.result in E3 = E4
    ============
    call-tc( (_id1: _id2), (_act-par), E1 ) = E4

[c5] act-par-tc(var _act-par, E1) = E2,
    _act-par should not be a packed component in E2 = E3,
    take care of possibly threatening variable _act-par in E3 = E4,
    ( _expr should not be a variant selector field in E4 = E5)
    type-tc(_id2, E4) = E5,
    E2.result and E5.result should be the same type in E5 = E6
    ============
    call-tc( (var _id1: _id2), (_act-par), E1 ) = E6

[c6] act-par-tc(procedure _act-par, E1) = E2,
    E2.result.formal-parameters and _form-par-list
    should be congruous in E2 = E3
    ============
    call-tc( (procedure _id1 _form-par-list), (_act-par), E1 ) = E3

72
\[ \text{act-par-tc}(\text{function } \_act-par, E1) = E2, \]
\[ E2.\text{result formal-parameters and } \_\text{form-par-list} \]
\[ \text{should be congruous in } E2 = E3 \]
\[ \text{call-tc( (function \_id1 \_form-par-list: \_id2), (\_act-par), E1 ) = E3} \]

{ Two parameter lists are congruous if they have the same number of arguments, and if each parameter pair refers to the same type. }

\[ \text{congruous?}(.,E) = \text{TRUE} \]
\[ \text{congruous?}((\_form-par1; \_form-par1+), (\_form-par2; \_form-par2+), E) = \]
\[ \text{congruous?}((\_form-par1), (\_form-par2), E) \& \]
\[ \text{congruous?}((\_form-par1+), (\_form-par2+), E) \]

\[ \text{type-tc(\_id1, E1)} = E2, \]
\[ \text{type-tc(\_id2, E2)} = E3, \]
\[ E2.\text{result is the same as } E3.\text{result?} = \text{TRUE} \]
\[ \text{congruous?}((\_id11: \_id1), (\_id22: \_id2), E1) = \text{TRUE} \]

\[ \text{type-tc(\_id1, E1)} = E2, \]
\[ \text{type-tc(\_id2, E2)} = E3, \]
\[ E2.\text{result is the same as } E3.\text{result?} = \text{TRUE} \]
\[ \text{congruous?}((\text{var } \_id11: \_id1), (\text{var } \_id22: \_id2), E1) = \text{TRUE} \]

\[ \text{congruous?}((\text{procedure } \_id1 \_form-par-list1), \]
\[ (\text{procedure } \_id2 \_form-par-list2), E) = \]
\[ \text{congruous?}((\_form-par-list1, \_form-par-list2), E) \]

\[ \text{type-tc(\_id1, E1)} = E2, \]
\[ \text{type-tc(\_id2, E2)} = E3, \]
\[ E2.\text{result is the same as } E3.\text{result?} = \text{TRUE} \]
\[ \text{congruous?}((\text{function } \_id11 \_form-par-list1 : \_id1), \]
\[ (\text{function } \_id22 \_form-par-list2 : \_id2), E1) = \]
\[ \text{congruous?}((\_form-par-list1 \_form-par-list2), E1) \]

\[ \text{congruous?}((\_form-par-list1, \_form-par-list2, E) = \text{TRUE} \]
\[ \text{congruous?}((\_form-par-list1, \_form-par-list2, E) != \text{TRUE} \]
\[ \text{congruous?}((\_form-par-list1, \_form-par-list2, E) != \text{TRUE} \]
\[ \text{add-error [Parameter lists are not congruous.] to } E \]

end module TC-expr-calls

module TC-spec-funs
\% Type check special, predefined, functions.
imports Expr-misc

73
equations

[sm1] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result and real-type should be the same type in E3 = E4
                           -----------------------------------
    spec-fun-tc( trunc(_act-par, _act-par*), E1 ) =
                     set-result of E4 to real-type

[sm2] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result and real-type should be the same type in E3 = E4
                           -----------------------------------
    spec-fun-tc( round(_act-par, _act-par*), E1 ) =
                     set-result of E4 to real-type

[sm3] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be a number in E3 = E4
                           -----------------------------------
    spec-fun-tc( abs(_act-par, _act-par*), E1 ) =
                     set-result of E4 to E2.result

[sm4] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be a number in E3 = E4
                           -----------------------------------
    spec-fun-tc( sqr(_act-par, _act-par*), E1 ) =
                     set-result of E4 to E2.result

[sm5] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be ordinal in E3 = E4
                           -----------------------------------
    spec-fun-tc( ord(_act-par, _act-par*), E1 ) =
                     set-result of E4 to integer-type

[sm6] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be ordinal in E3 = E4
                           -----------------------------------
    spec-fun-tc( succ(_act-par, _act-par*), E1 ) =
                     set-result of E4 to E2.result

[sm7] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be ordinal in E3 = E4
                           -----------------------------------
    spec-fun-tc( pred(_act-par, _act-par*), E1 ) =
                     set-result of E4 to E2.result

[sm8] act-par-tc( var _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3,
    E2.result should be a file in E3 = E4
                           -----------------------------------
spec-fun-tc( eoln(_act-par, _act-par*), E1 ) =
    set-result of E4 to boolean-type
    [sf9] act-par-tc( value _act-par, E1 ) = E2,
    no _act-par* should be given in E2 = E3.
    E2.result should be a file in E3 = E4
    ----------------------------------------------
    spec-fun-tc( eof(_act-par, _act-par*), E1 ) =
    set-result of E4 to boolean-type
    [eol1] spec-fun-tc( eof, E) = spec-fun-tc( eof(input), E)
    [eol2] spec-fun-tc( eoln,E) = spec-fun-tc(eoln(input), E)
end module TC-spec-funs

module TC-spec-procs
%% Check the predefined procedures: input/output, dynamic allocation, (un)pack
imports Expr-misc

hiddens
context-free syntax
    stc (" IDENT ACTUAL-PAR-LIST "," ENV ") ) ) -> ENV
    tcf (" IDENT ACTUAL-PAR-LIST "," ENV ") ) ) -> ENV
    "write-par-tc " (" ACTUAL-PAR "," ENV ") ) ) -> ENV
    tag-constants-tc " (" (ACTUAL-PAR ",")+ "," CONTEXT ,", ENV " ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) }
insert file variable "input". If the file parameter is of type text, (ISO section 6.9.1), the actual parameters shall be of char-type, integer-type, or real-type. If the file parameter is not of type text, (ISO section 6.6.5.2) the type of the actual parameter should be the same as the component type of the file. 

[ea1] stc( read, E ) = add-error [Too few actual parameters.] to E

[ea2] stc( read(_par), E ) = add-error [Too few actual parameters.] to E
    when is-file-variable?(_par,E) = TRUE

[ea3] stc( read(_par, _par+), E ) = tcf(read(_par, _par+), E)
    when is-file-variable?(_par,E) = TRUE

[ea4] stc( read(_par, _par*), E ) = tcf(read(input,_par, _par*), E)
    when is-file-variable?(_par,E) = TRUE

[ea5] tcf( read(_file, _par+), E ) =
    tcf( read(_file, _par), tcf( read(_file, _par+), E))

[ea6] act-par-tc(var _file, E1) = E2,
    E2.result.type = text-type,
    act-par-tc(var _par, E2) = E3,
    E3.result should be a character or a number in E3 = E4
    tcf( read(_file, _par), E1) = E4

[ea7] act-par-tc(var _file, E1) = E2,
    E2.result.type /= text-type,
    act-par-tc(var _par, E2) = E3,
    E3.result should be assignment-compatible with
    E2.result.type.comp-type in E3 = E4
    tcf( read(_file, _par), E1) = E4

{ The procedure readln, ISO section 6.9.2 }

[ln1] stc( readln, E ) = tcf( readln(input), E)

[ln2] stc( readln(_par, _par*), E ) = tcf(readln(input, _par, _par*), E)
    when is-file-variable?(_par, E) /= TRUE

[ln3] stc( readln(_par, _par*), E ) = tcf(readln(_par, _par*), E)
    when is-file-variable?(_par, E) = TRUE

[ln4] tcf( readln(_file, _par+), E ) =
    tcf( readln(_file), tcf( readln(_file, _par+), E))

[ln5] tcf( readln(_file), E ) = _file should be a text file variable in E

{ The procedure write, ISO section 6.9.3, 6.6.5.2. 
 Eventually insert file-variable output. Distinguish between text-files and non-text files. Check formats. }
[wr1] stc( write, E) = add-error [Too few actual parameters.] to E

[wr2] stc( write(_par), E) = add-error [Too few actual parameters.] to E
when is-file-variable?(_par, E) = TRUE

[wr3] stc( write(_par, _par+), E) = tcf(write(_par, _par+), E)
when is-file-variable?(_par, E) = TRUE

[wr4] stc( write(_par, _par*), E) = tcf(write(output, _par, _par*), E)
when is-file-variable?(_par, E) != TRUE

[wr5] tcf( write(_file, _par, _par+), E) =
tcf (write(_file, _par), tcf(write(_file, _par+), E))

[wr6] act-par-tc(var _file, E1) = E2,
E2.result.type = text-type,
write-par-tc(_par, E2) = E3,
E3.result should be a number, character, bool, or string in E3 = E4

==---------------------------------------------------------------------=
tcf (write(_file, _par), E1) = E4

[wr7] act-par-tc(var _file, E1) = E2,
E2.result.type != text-type,
act-par-tc(value _par, E2) = E3,
E3.result should be assignment-compatible with
E2.result.type.comptype in E3 = E4

==---------------------------------------------------------------------=
tcf (write(_file, _par), E1) = E4

{ The procedure writeln, ISO section 6.9.4. }
[ln6] stc( writeln, E) = tcf( writeln(output), E )

[ln7] stc( writeln(_par, _par*), E) = tcf(writeln(output, _par, _par*), E)
when is-file-variable?(_par, E) != TRUE

[ln8] stc( writeln(_par, _par*), E) = tcf(writeln(_par, _par*), E)
when is-file-variable?(_par, E) = TRUE

[ln9] tcf( writeln(_file, _par+), E) =
tcf (writeln(_file), tcf(writeln(_file, _par+), E))

[ln0] tcf( writeln(_file), E) = _file should be a text file variable in E

{ The procedure page, ISO section 6.9.5. }
[pa1] stc( page, E ) = stc( page(output), E )

[pa2] stc( page(_file), E) = _file should be a text file variable in E

[pa3] stc( page(_file, _par+), E) = add-error [Too many actual parameters.] to E

[exp1] write-par-tc(_expr, E) = expr-tc(_expr, E)
\[ \text{expr-\text{tc}}( \text{\_expr2}, \text{E1} ) = \text{E2}, \]
\[ \text{expr-\text{tc}}( \text{\_expr1}, \text{E2} ) = \text{E3}, \]
\[ \text{E3.result should be integer in E3 = E4} \]

\[ \text{write-par-\text{tc}}( \text{\_expr1 : \_expr2}, \text{E1} ) = \text{E4} \]

\[ \text{expr-\text{tc}}( \text{\_expr3}, \text{E1} ) = \text{E2}, \]
\[ \text{expr-\text{tc}}( \text{\_expr2}, \text{E2} ) = \text{E3}, \]
\[ \text{expr-\text{tc}}( \text{\_expr1}, \text{E3} ) = \text{E4}, \]
\[ \text{E2.result should be integer in E4 = E5}, \]
\[ \text{E3.result should be integer in E5 = E6}, \]
\[ \text{E4.result should be real in E6 = E7} \]

\[ \text{write-par-\text{tc}}( \text{\_expr1 : \_expr2 : \_expr3}, \text{E1} ) = \text{E5} \]

\{ Dynamic allocation procedures new and dispose, ISO section 6.6.5.3. \}
\[ \text{act-par-\text{tc}}( \text{\_par, E1} ) = \text{E2}, \]
\[ \text{E2.result.type should be a pointer in E2 = E3} \]

\[ \text{stc}( \text{new(\_p), E1} ) = \text{E3} \]

\[ \text{stc}( \text{new(\_p), E1} ) = \text{E2}, \]
\[ \text{tag-constants-\text{tc}}( \text{\_par+, E2.result.type}, \text{E2} ) = \text{E3} \]

\[ \text{stc}( \text{new(\_p, \_par+)}, \text{E1} ) = \text{E3} \]

\[ \text{stc}( \text{\_par, E1} ) = \text{E2}, \]
\[ \text{tag-constants-\text{tc}}( \text{\_par, T1, E1} ) = \text{E3} \]

\[ \text{tag-constants-\text{tc}}( \text{\_par+, T1.result-of-variant-part}, \text{E2} ) = \text{E3} \]

\[ \text{tag-constants-\text{tc}}( \text{\_par+, \_par+, T1, E1} ) = \text{E3} \]

\[ \text{stc}( \text{\_par, E1} ) = \text{add-error [Too few actual parameters.]} \text{ to E} \]
\[ \text{stc}( \text{dispose _act-par-list, E} ) = \text{stc( new _act-par-list, E) } \]

\{ Transfer procedures, pack(a,i,z), unpack(z,a,j), ISO section 6.6.5.4. \}
\[ \text{act-par-\text{tc}}( \text{\_par, E1} ) = \text{E2}, \]
\[ \text{act-par-\text{tc}}( \text{\_a, E2} ) = \text{E3}, \]
\[ \text{act-par-\text{tc}}( \text{\_i, E3} ) = \text{E4}, \]
\[ \text{E2.result.type should be an unpacked array in E4 = E5}, \]
\[ \text{E3.result.type should be a packed array in E5 = E6}, \]
\[ \text{E2.result.type.comp-type and E3.result.type.comp-type should be the same type in E6 = E7}, \]
\[ \text{E4.result should be assignment-compatible with E3.result.type.index-type in E7 = E8} \]

\[ \text{stc}( \text{pack(\_a, \_i, \_z), E1} ) = \text{E8} \]
module TC-expr
imports TC-expr-vars TC-expr-simple TC-expr-calls TC-spec-funs TC-spec-procs

hiddens
context-free syntax
arithm-op-table -> OPERATOR-TABLE
relational-op-table -> OPERATOR-TABLE
unop-tc "(" OPERATOR "," EXPR "," ENV ")" -> ENV
binop-tc "(" OPERATOR "," EXPR "," EXPR "," ENV ")" -> ENV
relop-tc "(" OPERATOR "," EXPR "," EXPR "," ENV ")" -> ENV
match-ops "(" OPERATOR-TABLE "," OPERATOR "," OP-TYPE "," OP-TYPE "," ENV ")" -> ENV
check-sets "(" CONTEXT "," CONTEXT ")" in ENV -> ENV
get-result "(" CONTEXT "," CONTEXT ")" -> CONTEXT
map "(" CONTEXT ")" -> OP-TYPE
map "(" OP-TYPE ")" -> CONTEXT
rmap "(" CONTEXT ")" -> OP-TYPE
omap "(" CONTEXT ")" -> OP-TYPE
is-just-an-identifier "," "(" EXPR ")" -> BOOLEAN

variables
[\_\_]\*Op\-[era/\]/*tor -> OPERATOR Op\-[0-9]\* -> OP-TYPE
Result-type -> OP-TYPE OpTable -> OPERATOR-TABLE

equations
{ Direct the basic expressions: constants and simple expressions,
variables accesses, and function calls. }

[e1] is-simple-expr?\(_expr\,,\Env\.context\) = TRUE
                                  =sim-expr-tc\(_expr\,,\Env\)

[e2] is-var-access?\(_expr\,,\Env1\.context\) = TRUE,
                           var-access-tc\(_expr\,,\Env1\) = Env2
                                  =set-result of Env2 to Env2.result.type.canonical
is-function-call?( _expr, Env1.context ) = TRUE

expr-tc( _expr, Env1 ) = fun-call-tc(_expr, Env1 )

is-simple-expr?( _id, E.context ) != TRUE,
is-var-access?( _id, E.context ) != TRUE,
is-function-call?( _id, E.context ) != TRUE

expr-tc( _id, E ) = add-error

[Identifier _id is not a constant, variable, or function.] to E

{ Type check expressions constructed of arithmetic and relational operators. }

[ott1] arithm-op-table =

<table>
<thead>
<tr>
<th>Operator</th>
<th>Operand1</th>
<th>Operand2</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>+</td>
<td>int</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>+</td>
<td>real</td>
<td>int</td>
<td>real</td>
</tr>
<tr>
<td>+</td>
<td>real</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>+</td>
<td>int</td>
<td>undef</td>
<td>int</td>
</tr>
<tr>
<td>+</td>
<td>real</td>
<td>undef</td>
<td>real</td>
</tr>
<tr>
<td>+</td>
<td>set</td>
<td>set</td>
<td>set</td>
</tr>
<tr>
<td>-</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>-</td>
<td>int</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>-</td>
<td>real</td>
<td>int</td>
<td>real</td>
</tr>
<tr>
<td>-</td>
<td>real</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>-</td>
<td>int</td>
<td>undef</td>
<td>int</td>
</tr>
<tr>
<td>-</td>
<td>real</td>
<td>undef</td>
<td>real</td>
</tr>
<tr>
<td>-</td>
<td>set</td>
<td>set</td>
<td>set</td>
</tr>
<tr>
<td>*</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>*</td>
<td>int</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>*</td>
<td>real</td>
<td>int</td>
<td>real</td>
</tr>
<tr>
<td>*</td>
<td>real</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>*</td>
<td>set</td>
<td>set</td>
<td>set</td>
</tr>
<tr>
<td>/</td>
<td>int</td>
<td>int</td>
<td>real</td>
</tr>
<tr>
<td>/</td>
<td>int</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>/</td>
<td>real</td>
<td>int</td>
<td>real</td>
</tr>
<tr>
<td>/</td>
<td>real</td>
<td>real</td>
<td>real</td>
</tr>
<tr>
<td>div</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>div</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>mod</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>mod</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>and</td>
<td>boolean</td>
<td>boolean</td>
<td>boolean</td>
</tr>
<tr>
<td>or</td>
<td>boolean</td>
<td>boolean</td>
<td>boolean</td>
</tr>
<tr>
<td>not</td>
<td>boolean</td>
<td>undef</td>
<td>boolean</td>
</tr>
</tbody>
</table>
### relational-op-table

<table>
<thead>
<tr>
<th>Operator</th>
<th>Operand 1</th>
<th>Operand 2</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>ordinal</td>
<td>set</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;</td>
<td>string</td>
<td>string</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;=</td>
<td>simple</td>
<td>simple</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;=</td>
<td>string</td>
<td>string</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;=</td>
<td>set</td>
<td>set</td>
<td>boolean</td>
</tr>
<tr>
<td>&gt;</td>
<td>string</td>
<td>string</td>
<td>boolean</td>
</tr>
<tr>
<td>&gt;=</td>
<td>simple</td>
<td>simple</td>
<td>boolean</td>
</tr>
<tr>
<td>&gt;=</td>
<td>simple</td>
<td>string</td>
<td>boolean</td>
</tr>
<tr>
<td>&gt;=</td>
<td>set</td>
<td>set</td>
<td>boolean</td>
</tr>
<tr>
<td>=</td>
<td>simple</td>
<td>simple</td>
<td>boolean</td>
</tr>
<tr>
<td>=</td>
<td>pointer</td>
<td>pointer</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;&gt;</td>
<td>simple</td>
<td>simple</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;&gt;</td>
<td>pointer</td>
<td>pointer</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;&gt;</td>
<td>set</td>
<td>set</td>
<td>boolean</td>
</tr>
<tr>
<td>&lt;&gt;</td>
<td>string</td>
<td>string</td>
<td>boolean</td>
</tr>
</tbody>
</table>

{ Type check the operators in the expressions. }
{ To do this, find the types of the operators, and check in a type-table whether these operator types can be matched in that type-table. If not raise an error. }

[u1] \[ \text{expr-tc}(+ \_expr, E) = \text{unop-tc}(+, \_expr, E) \]
[u2] \[ \text{expr-tc}(- \_expr, E) = \text{unop-tc}(-, \_expr, E) \]
[u3] \[ \text{expr-tc}(\text{not} \_expr, E) = \text{unop-tc}(\text{not}, \_expr, E) \]

[w4] \[ \text{expr-tc}(\_expr, E_1) = E_2, \text{match-ops}(\text{arithm-op-table}, \_Operator, \text{map(E2.result)}, \text{undef}, E_2) = E_3 \]

[b1] \[ \text{expr-tc}(\_expr1 / \_expr2, E) = \text{binop-tc}(\text{div}, \_expr1, \_expr2, E) \]
[b2] \[ \text{expr-tc}(\_expr1 \text{ div} \_expr2, E) = \text{binop-tc}(\text{div}, \_expr1, \_expr2, E) \]
[b3] \[ \text{expr-tc}(\_expr1 \text{ mod} \_expr2, E) = \text{binop-tc}(\text{mod}, \_expr1, \_expr2, E) \]
[b4] \[ \text{expr-tc}(\_expr1 \text{ mod} \_expr2, E) = \text{binop-tc}(\text{mod}, \_expr1, \_expr2, E) \]
[b5] \[ \text{expr-tc}(\_expr1 \text{ or} \_expr2, E) = \text{binop-tc}(\text{or}, \_expr1, \_expr2, E) \]
[b6] \[ \text{expr-tc}(\_expr1 \text{ or} \_expr2, E) = \text{binop-tc}(\text{or}, \_expr1, \_expr2, E) \]
[b7] \[ \text{expr-tc}(\_expr1 \text{ and} \_expr2, E) = \text{binop-tc}(\text{and}, \_expr1, \_expr2, E) \]
[b8] \[ \text{expr-tc}(\_expr1 \text{ and} \_expr2, E) = \text{binop-tc}(\text{and}, \_expr1, \_expr2, E) \]
[b9] \[ \text{expr-tc}(\_expr1, E_1) = E_2, \text{expr-tc}(\_expr2, E_2) = E_3, \text{match-ops}(\text{arithm-op-table}, \_Operator, \text{map(E2.result)}, \text{map(E3.result)}, E_3) = E_4, \text{check-sets}(E2.result, E3.result) \text{ in } E_4 = E_5 \]
\[
\text{binop-tc}(\_\text{Operator}, \_\text{expr}1, \_\text{expr}2, E1) = E5
\]

\[
\begin{align*}
[r1] \text{expr-tc}(\_\text{expr}1 = \_\text{expr}2, E) &= \text{relop-tc}(=, \_\text{expr}1, \_\text{expr}2, E) \\
[r2] \text{expr-tc}(\_\text{expr}1 <\!> \_\text{expr}2, E) &= \text{relop-tc}(<?, \_\text{expr}1, \_\text{expr}2, E) \\
[r3] \text{expr-tc}(\_\text{expr}1 \leq \_\text{expr}2, E) &= \text{relop-tc}(\leq, \_\text{expr}1, \_\text{expr}2, E) \\
[r4] \text{expr-tc}(\_\text{expr}1 \geq \_\text{expr}2, E) &= \text{relop-tc}(\geq, \_\text{expr}1, \_\text{expr}2, E) \\
[r5] \text{expr-tc}(\_\text{expr}1 > \_\text{expr}2, E) &= \text{relop-tc}(>, \_\text{expr}1, \_\text{expr}2, E) \\
[r6] \text{expr-tc}(\_\text{expr}1 \neq \_\text{expr}2, E) &= \text{relop-tc}(\neq, \_\text{expr}1, \_\text{expr}2, E) \\
\end{align*}
\]

\[
\begin{align*}
[r8] \text{expr-tc}(\_\text{expr}1, E1) &= E2, \\
\text{expr-tc}(\_\text{expr}2, E2) &= E3, \\
E2.\text{result }\_\text{Operator} E3.\text{result should be compatible, numbers,} & \\
or \text{of the same set in } E3 = E4, \\
\text{match-ops(\text{relational-op-table,}} & \\
\_\text{Operator}, \text{\_Operator, rmap(E2.result), rmap(E3.result), E3) = E4} \\
\text{relop-tc(\_\text{Operator}, \_\text{expr}1, \_\text{expr}2, E1) = E4} \\
\end{align*}
\]

\[
\begin{align*}
[r9] \text{expr-tc}(\_\text{expr}1, E1) &= E2, \\
\text{expr-tc}(\_\text{expr}2, E2) &= E3, \\
\text{match-ops(\text{relational-op-table, in,}} & \\
\_\text{Operator}, \text{\_Operator, omap(E2.result), rmap(E3.result), E3) = E4,} \\
E2.\text{result should be the set-component of E3.result in E4 = E5} \\
\text{expr-tc(\_\text{expr}1 in \_\text{expr}2, E1) = E5} \\
\end{align*}
\]

{ If the operators of +, -, or * where sets, extra checks need to be done. The sets must be of the same component-type. The result should be set (if possible) to the non-empty list argument. }

\[
\begin{align*}
[c0] \text{is-set-type?(T1)} &| \text{is-set-type?(T2)} | = \text{TRUE} \\
\text{check-sets(T1, T2) in E1 = E1} \\
\end{align*}
\]

\[
\begin{align*}
[c1] \text{is-set-type?(T1)} &| \text{is-set-type?(T2)} = \text{TRUE}, \\
\text{T1 and T2 should be the same set of T type in E1 = E2} \\
\text{check-sets(T1, T2) in E1 = set-result of E2 to get-result(T1, T2)} \\
\end{align*}
\]

\[
\begin{align*}
[c2] \text{get-result}(T1, T2) &= T1 \text{ when is-non-empty-set-type?(T1)} = \text{TRUE} \\
[c3] \text{get-result}(T1, T2) &= T2 \text{ when is-non-empty-set-type?(T2)} = \text{TRUE} \\
[c4] \text{get-result}(T1, T2) &= T1 \text{ when is-empty-set-type?(T1)} &| \text{is-empty-set-type?(T2)} = \text{TRUE} \\
\end{align*}
\]

\[
\begin{align*}
[f1] \text{look-for(Op\_tor, Op1, Op2, Op\_Table)} &= <\text{TRUE, Result-type}> \\
\text{match-ops(Op\_Table, Op\_tor, Op1, Op2, E)} &= \text{set-result of E to map(Result-type)} \\
\end{align*}
\]

\[
\begin{align*}
[f2] \text{look-for(Op\_tor, Op1, Op2, Op\_Table)} &= <\text{FALSE, Result-type}> \\
\end{align*}
\]
match-ops(OpTable, Optor, Op1, Op2, E) =
  add-error [Illegal operand types for operator Optor.] to E

[ma1] map(T) = int when T = integer-type
[ma2] map(T) = boolean when T = boolean-type
[ma3] map(T) = real when T = real-type
[ma4] map(T) = set when is-set-type?(T) = TRUE

[mb1] map(int) = integer-type
[mb2] map(real) = real-type
[mb3] map(boolean) = boolean-type

[rm1] rmap(T) = simple when is-simple-type?(T) = TRUE
[rm2] rmap(T) = set when is-set-type?(T) = TRUE
[rm3] rmap(T) = pointer when is-pointer-type?(T) = TRUE
[rm4] rmap(T) = string when is-string-type?(T) = TRUE

[om1] omap(T) = ordinal when is-ordinal-type?(T) = TRUE

end module TC-expr

module TC-stat

%% Type check statements; Goto statements are not dealt with.
imports TC-expr
exports
  context-free syntax
    stat-tc "(" {STATEMENT ";"+ "," ENV ")" -> ENV

hiddens
sorts CASE-OUT

context-free syntax
  lhs-tc "(" VARIABLE-ACCESS "," ENV ")" -> ENV
  case-lists-tc "(" {CASE-LIST-ELT ";"+ "," CONTEXT "," VAL-LIST "," ENV ")" -> VAL-OUT

variables
  [,_stat[1:2]*/ -> STATEMENT [,_"stat*" -> { STATEMENT ";"*]
  [,_unl-stat* -> UML-STAT FunDef -> CONTEXT
  VarDef -> CONTEXT "Cl-elt"* -> {CASE-LIST-ELT ";"+}
  Ct-elt -> CASE-LIST-ELT "(Down)To" -> DOWN-TO

equations
  [s2] stat-tc(, E) = E
  [s3] stat-tc(_stat; _stat*, E) =
      stat-tc(_stat*, stat-tc(_stat, E))

  { Procedure call }
  [s4] set-construct of E0 to "_id _act-par-list" = E1, 
      proc-call-tc(_id _act-par-list, E1 ) = E2
      ==============================================================
      stat-tc(_id _act-par-list, E0 ) = E2

83
{ Compound statement }
[as6] \( \text{stat-tc}(\text{begin} \_\text{stat}+ \text{end}, E) = \text{stat-tc}(\_\text{stat}+, E) \)

{ Assignment statement }
[as1] set-construct of \( E_0 \) to "\_\text{var} := \_\text{expr} = E_1, \)
\( \text{lhs-tc}(\_\text{var} := \_\text{expr}, E_1) = E_2, \)
expr-tc(\_\text{expr}, E_2) = E_3, \)
\( E_2.\text{result should be assignment-compatible with } E_3.\text{result in } E_3 = E_4 \)

\( \text{stat-tc}(\_\text{var} := \_\text{expr} \), E_0) = E_4 \)

[as2] \_\text{id} is the name of the inner function-block in \( E.\text{context} = \) TRUE, \)
\( \text{find } \_\text{id} \text{ in } E.\text{context} = \text{FunDef} \)
\( \text{lhs-tc}(\_\text{id}, E) = \text{set-result of } E \text{ to } \text{FunDef.\text{type}.canonical} \)

[as3] \_\text{id} is the name of the inner function-block in \( E_1.\text{context} \) != TRUE, \)
\( \text{expr-tc}(\_\text{id}, E_1) = E_2, \)
\( E_2.\text{result defines } \_\text{id} \text{ as a variable?} = \text{TRUE}, \)
\( \text{take care of possibly threatening variable } \_\text{id} \text{ in } E_2 = E_3 \)
\( \text{lhs-tc}(\_\text{id}, E_1) = \text{set-result of } E_3 \text{ to } E_2.\text{result.\text{type}.canonical} \)

[as4] \text{is-just-\text{var}-\text{identifier}(\_\text{var} = \_\text{acc}) } != \text{TRUE} \)
\( \text{lhs-tc}(\_\text{var} = \_\text{acc}, E_1) = \text{var-access-tc}(\_\text{var} = \_\text{acc}, E_1) \)

{ If statement }
[is1] set-construct of \( E_0 \) to "if \_\text{expr} then \langle \text{STATEMENT} \rangle \" = E_1, \)
\( \text{expr-tc}(\_\text{expr}, E_1) = E_2, \)
\( \text{stat-tc}(\_\text{stat}, E_2) = E_3, \)
\( E_2.\text{result should be } \text{Boolean in } E_3 = E_4 \)
\( \text{stat-tc}(\text{if } \_\text{expr} \text{ then } \_\text{stat}, E_0) = E_4 \)

[is2] set-construct of \( E_0 \) to "if \_\text{expr} then \langle \text{STATEMENT} \rangle \text{ else } \langle \text{STATEMENT} \rangle \" = E_1, \)
\( \text{expr-tc}(\_\text{expr}, E_1) = E_2, \)
\( \text{stat-tc}(\_\text{stat1}, E_2) = E_3, \)
\( \text{stat-tc}(\_\text{stat2}, E_3) = E_4, \)
\( E_2.\text{result should be } \text{Boolean in } E_4 = E_5 \)
\( \text{stat-tc}(\text{if } \_\text{expr} \text{ then } \_\text{stat1} \text{ else } \_\text{stat2}, E_0) = E_5 \)

{ With statement }
[ws1] \text{stat-tc}(\text{with } \_\text{var} = \_\text{var}, \_\text{var}+ \text{do } \_\text{stat}, E) = \text{stat-tc}(\text{with } \_\text{var} = \_\text{var}+ \text{do } \_\text{var} \text{ do } \_\text{stat}, E) \)

[ws2] set-construct of \( E_0 \) to "with \_\text{var} = \_\text{var} \text{ do } \langle \text{STATEMENT} \rangle \" = E_1, \)
\( \text{var-access-tc}(\_\text{var}, E_1) = E_2, \)
\( E_2.\text{result.\text{type} should be a record in } E_2 = E_3, \)
\( \text{new-block-with-fields}(E_2.\text{result.\text{type}}, E_3) = E_4, \)
\( \text{stat-tc}(\_\text{stat}, E_4) = E_5, \)
\( \text{leave inner block of } E_5 = E_6 \)
The text contains a formal specification of the type checking (``tc``) process for case statements and repetitive statements in a programming language. The text includes examples and conditions for validation. The document is structured into several sections, each specifying different aspects of the type checking process.

### Section 1: Case Statement

- **Case Statement:**
  - The case statement constructs a `case _expr of Cl-elt+ end` where `Cl-elt+` is the case expression.
  - The result of the expression `E1` should be ordinal.

### Section 2: Repetitive Statements

- **Repetitive Statements:**
  - The repetitive statements include `repeat` and `while` constructs.
  - The `repeat` statement should have a Boolean result.
  - The `while` statement should also have a Boolean result.

### Additional Conditions

- **Variable Access:**
  - Variables used in the case statements must be within the same context.
  - Control variables are not allowed.

- **Expression Types:**
  - Expressions used in the case statement must be of the same type and have different values.

The document concludes with examples illustrating the type checking process for each section.
E₆.result should be ordinal in E₁₀ = E₁₁.
E₇.result should be assignment-compatible with E₅.result in E₁₁ = E₁₂,
E₈.result should be assignment-compatible with E₅.result in E₁₂ = E₁₃

stat-tc(for _id := _expr1 (Down)To _expr2 do _stat, E₀) = E₁₃

{ It is only checked whether labels are declared indeed. }
[ls1] _label should be declared in E₁ = E₂,
  stat-tc(_ unl-stat, E₂) = E₃

[ls₂] stat-tc(goto _label, E) = _label should be declared in E

end module TC-stat

module TC-block
  %% Type check a block of a Pascal Program (ISO section 6.2.1.)
  %% Most of the work is done in the checks for procedure and function
  %% declarations:
  %% The heading must be correct (all parameter names different, all type
  %% identifiers legal), and the block must be correct.
  %% All forwarded headings should be followed by an identification
  %% associating a body with that heading.

imports TC-stat

exports
  context-free syntax
    block-tc "(" BLOCK "," ENV ")" -> ENV

hiddens
  sorts FORW-DECLS PROFUN-OUT
  context-free syntax
    labels-tc "(" LABELS "," ENV ")" -> ENV
    const-tc "(" CONSTS "," ENV ")" -> ENV
    types-tc "(" TYPES "," ENV ")" -> ENV
    var-tc "(" VARS "," ENV ")" -> ENV
    procs-or-funs-tc "(" PROCS-OR-FUNS "," ENV ")" -> ENV
    profuns-tc "(" PROC-OR-FUN "\n", ENV ")" -> ENV
    profuns-tc "(" PROC-OR-FUN "\n", FORW-DECLS "," ENV ")" -> PROFUN-OUT
    block-heading-tc "(" HEADING "," FORW-DECLS "," ENV ")" -> PROFUN-OUT
    forwarded-profun-tc "(" PROCEDURE-HEADING ","
      FORW-DECLS "," ENV ")" -> PROFUN-OUT
    forwarded-profun-tc "(" FUNCTION-IDENTIFICATION ","
      FORW-DECLS "," ENV ")" -> PROFUN-OUT
    forwarded-heading-tc "(" HEADING "," ENV ")" -> ENV
    param-list-tc "(" FORVAL-PAR-LIST "," ENV ")" -> ENV
    "[" { HEADING "\n", "}\n" -> FORW-DECLS
    "<= FORW-DECLS "," ENV ">" -> PROFUN-OUT

86
remove IDENT from FORW-DECLS
get-id-list "(" FORW-DECLS ")" → IDENT-LIST
is-forward-list-empty "?" "(" FORW-DECLS ")" → BOOLEAN
IDENT is a forwarded procedure in FORW-DECLS "?" → BOOLEAN
forward list FORW-DECLS should be empty in ENV → ENV

variables
Id-list → IDENT-LIST Block → BLOCK
Labels → LABELS_consts → CONSTS
Const-defs → {CONST-DEF ";";}* Types → TYPES
Concat-defs → {TYPE-DEF ";";}*Vars → VARS
Var-defs → {VAR-DECL ";";}* Procs-or-funs → PROCS-OR-FUNS
Profuns → { PROC-OR-FUN ";";}* Profun-decl → PROC-OR-FUN
Comp-stat → COMP-STAT P(forw-list)*[0-9]* → FORW-DECLS
Forwards → { HEADING ",";}*Heading → HEADING
Headings[0-9]* → { HEADING ",";}* Heading-plus → { HEADING ","; }+
Fun-heading → FUNCTION-HEADING Proc-heading → PROCEDURE-HEADING

equations
[b1] block-tc (Labels_consts_Types Vars Procs-or-funs Comp-stat, E ) =
    stat-tc ( Comp-stat,
             procs-or-funs-tc ( Procs-or-funs,
                            vars-tc ( Vars,
                                      types-tc ( Types,
                                                consts-tc ( Consts,
                                                            labels-tc ( Labels, E ))) )))
[b2] labels-tc ( , E ) = E
[b3] labels-tc (label _label*, , E ) = label-decls-tc ( _label*, E )
[b4] consts-tc ( , E ) = E
[b5] consts-tc ( const Const-defs; , E ) = const-defs-tc ( Const-defs, E )
[b6] types-tc ( , E ) = E
[b7] types-tc ( type Type-defs; , E ) = type-defs-tc ( Type-defs, E )
[b8] vars-tc ( , E ) = E
[b9] vars-tc ( var Var-defs; , E ) = var-decls-tc ( Var-defs, E )
[b10] vars-tc ( forw list Forwards; , E ) = procs-or-funs-tc ( Procs-or-funs, E )

{ Type check procedure and function declarations }
[pf1] profuns-tc ( Profuns; [], E1 ) = <Forw-list, E2>,
    forward list Forw-list should be empty in E2 = E3
    procs-or-funs-tc ( Profuns, E1 ) = E3
[pf2] profuns-tc ( Profun-decl, Forw-list1, E1 ) = <Forw-list2, E2>,
    profuns-tc ( Profuns, Forw-list2, E2 ) = <Forw-list3, E3>
    profuns-tc ( Profun-decl, Profuns, Forw-list1, E1 ) = <Forw-list3, E3>
[pf3] profuns-tc ( F, E ) = <F, E>

{ A normally defined procedure or function; }
[pf6] block-heading-tc ( Heading, Forw-list1, E1 ) = <Forw-list2, E2>,

87
block-tc(Block, E2) = E3,
leave inner block of E3 = E4

profuns-tc(Heading, Block, Forw-list1, E1) = <Forw-list2, E4>

{ A declaration with a forward directive; add heading to Forw-list and to environment. }
[pf4] forw-heading-tc(Heading, E1) = E2

profuns-tc(Heading, forward, [Forwards], E1) =
< [Forwards, Heading], E2 + Heading>

{ A function already declared forward, whose body is given now. }
[pf5] forwarded-profun-tc(function _id, Forw-list1, E1) = <Forw-list2, E2>,
block-tc(Block, E2) = E3,
leave inner block of E3 = E4

profuns-tc(function _id; Block, Forw-list1, E1) = <Forw-list2, E4>

{ Type check a heading of a forwarded function or procedure. 
Just check the correctness of the parameters and their types. }
[fh1] _id1 should not be declared in inner block of E1 = E2,
E2 + block-mark = E3,
param-list-tc(_form-par-list, E3) = E4,
_id2 should be a type in E4 = E5,
leave inner block of E5 = E6

forw-heading-tc(function _id1 _form-par-list: _id2, E1) = E6

[fh2] _id1 should not be declared in inner block of E1 = E2,
E2 + block-mark = E3,
param-list-tc(_form-par-list, E3) = E4,
leave inner block of E4 = E5

forw-heading-tc(procedure _id1 _form-par-list, E1) = E5

{ Type check a heading of a function or procedure with a block. 
The normal case is a new procedure or function, possibly with parameters. 
In the "procedure _id: block" case, the _id may either refer to an 
already forwarded procedure, or to a new procedure without parameters. }
[bh1] _id should not be declared in inner block of E1 = E2,
E2 + function _id1 _form-par-list : _id2 + block-mark = E3,
param-list-tc(_form-par-list, E3) = E4,
_id2 should be a type in E4 = E5

block-heading-tc(function _id1 _form-par-list: _id2, Forw-list, E1) = <Forw-list, E5>

[bh2] _id is a forwarded procedure in Forw-list? ! = TRUE,
_id should not be declared in inner block of E1 = E2,
E2 + procedure _id _form-par-list + block-mark = E3,
param-list-tc(_form-par-list, E3) = E4

block-heading-tc(procedure _id _form-par-list, Forw-list, E1) = <Forw-list, E4>

88
_id is a forwarded procedure in Forw-list? = TRUE

block-heading-tc(procedure _id, Forw-list, E) =
    forwarded-profun-tc(procedure _id, Forw-list, E)

{ Check a function or procedure identification; i.e. a function or
procedure already forwarded, and whose body is given now. }

[fw1] remove _id from Forw-list1 = Forw-list2,
    _id should be a function in inner block of E1 = E2,
    Fun-heading = E2.result.entry,
    _form-par-list = E2.result.formal-parameters,
    E2 + block-mark + Fun-heading = E3,
    param-list-tc(_form-par-list, E3) = E4
    forwarded-profun-tc(function _id, Forw-list1, E1) = <Forw-list2, E4>

[fw1] remove _id from Forw-list1 = Forw-list2,
    _id should be a procedure in inner block of E1 = E2,
    Proc-heading = E2.result.entry,
    _form-par-list = E2.result.formal-parameters,
    E2 + block-mark + Proc-heading = E3,
    param-list-tc(_form-par-list, E3) = E4
    forwarded-profun-tc(function _id, Forw-list1, E1) = <Forw-list2, E4>

{ Check a formal parameter list.
Add the parameters as variables to the block: guarantees the uniqueness
of the names. Check whether the types are OK. }

[pr1] (_form-par*: _id1, _id+: _id3, _form-par2*) =
    (_form-par*: _id1: _id3, _id+: _id3, _form-par2*)
[pr2] (_form-par*: var _id1, _id+: _id3, _form-par2*) =
    (_form-par*: var _id1: _id3, var _id+: _id3, _form-par2*)

[pl1] param-list-tc(E) = E
[pl2] param-list-tc(_form-par: _form-par+), E) =
    param-list-tc(_form-par+), param-list-tc(_form-par, E))

[pl3] _id1 should not be declared in inner block of E1 = E2,
    _id2 should be a type in E2 = E3,
    E3.result should be allowed as file-component in E3 = E4
    param-list-tc(_id1: _id2), E1) = E3

[pl4] _id1 should not be declared in inner block of E1 = E2,
    _id2 should be a type in E2 = E3
    param-list-tc(var _id1: _id2), E1) = E3

[pl5] param-list-tc(Fun-heading), E) = forw-heading-tc(Fun-heading, E)
[pl6] param-list-tc(Proc-heading), E) = forw-heading-tc(Proc-heading, E)
{ Operations on lists of forwarded headings: }

[fl1] _id is a forwarded procedure in
   [Headings1, procedure _id _form-par-list, Headings2] = TRUE
[fl2] remove _id from [Headings1, procedure _id _form-par-list, Headings2] =
    [Headings1, Headings2]
[fl3] remove _id1 from [Headings1, function _id1 _form-par-list:_id2, Headings2] = 
    [Headings1, Headings2]
[fl4] is-forward-list-empty([[]]) = TRUE
[fl5] is-forward-list-empty?(Forw-list) = TRUE,

forward list Forw-list should be empty in E
E
[fl6] is-forward-list-empty?(Forw-list) != TRUE,
   Id-list = get-id-list(Forw-list)

forward list Forw-list should be empty in E
   add-error [No body given for forwarded identifiers: Id-list.] to E

[fl7] get-id-list([procedure _id _form-par-list]) = [_id]
[fl8] get-id-list([function _id _form-par-list : _id2]) = [_id]
[fl9] get-id-list([Heading]) = [_id].
   get-id-list([Heading-plus])= [_id+]
   get-id-list([Heading, Heading-plus]) = [_id, _id+]
[fl10] get-id-list([]) = []

end module TC-block

module TC-program
XX Typecheck Pascal programs, according to ISO section 6.2.1. and 6.10.
imports TC-block
exports
   context-free syntax
     tc "(" PROGRAM ")" -> ERRORS
hiddens
   context-free syntax
     param-tc-before-block "(" PROGRAM-HEADER "," ENV ")" -> ENV
     param-tc-after-block "(" PROGRAM-HEADER "," ENV ")" -> ENV
     param-tc-before-block "(" IDENT ",")* "," ENV ")" -> ENV
     param-tc-after-block "(" IDENT ",")* "," ENV ")" -> ENV
     "check-input/output" "(" IDENT ",")+ "," ENV ")" -> ENV
     "is-input/output-parameter" "(" IDENT ",")+ "," IDENT ")" -> BOOLEAN
] variables
   Program -> PROGRAM
   Prog-params -> PROGRAM-PARAMS
   Block -> BLOCK
   [_]pars -> { IDENT ","}*
   [_]par -> IDENT
90
equations

[pr1] params-tc-before-block( _header, new-env ) = E1,
  block-tc( Block, E1 ) = E2,
  params-tc-after-block( _header, E2 ) = E3

==

[pr1] tc( _header, Block ) = E3.errors

{ Check the parameter list of the program header. }

[pb1] set-construct of E1 to "program _id (_pars)" = E2,
  (_pars) should not have double names in E2 = E3,
  check-input/output( _pars, E3 ) = E4

==

params-tc-before-block( program _id (_pars), E1 ) = E4

[pb2] params-tc-before-block( program _id, E ) = E

{ If input or output are given as program parameters, they are supposed
  to be implicitly defined as variables of type char. Make this explicit
  by adding them to the environment. }

[io1] check-input/output( _par, _pars, E ) =
  check-input/output( _par, check-input/output( _par, E ) )
[io2] check-input/output( input, E ) = E + var input: text
[io3] check-input/output( output, E ) = E + var output: text
[io4] check-input/output( _par, E ) =
  when is-input/output-parameter?( _par ) != TRUE
[io5] check-input/output( , E ) = E
[io6] is-input/output-parameter?( input ) = TRUE
[io7] is-input/output-parameter?( output ) = TRUE

{ All program parameters should be defined as variables in the
  program block. This is checked after the program block has been checked. }

[pa0] set-construct of E1 to "program _id (_pars)" = E2,
  params-tc-after-block( _pars, E2 ) = E3

==

params-tc-after-block( program _id (_pars), E1 ) = E3

[pa1] params-tc-after-block( _par, _pars, E ) =
  params-tc-after-block( _pars, params-tc-after-block( _par, E ) )

[pa2] params-tc-after-block( _par, E ) = _par should be a variable in E
[pa3] params-tc-after-block( , E ) = E

end module TC-program
D An Example

There is no way to give a formal proof of the correctness of the algebraic specification in this document with respect to the ISO specification in natural language. However, in order to strengthen the idea that our specification is correct, we will give a small example of how it works on a Pascal program containing some static semantic errors. We give a program to compute the greatest common divisor of two natural numbers using Euclid’s algorithm (see, e.g., [Dij76]).

\{
  1 \} \text{program Euclid (output);}
  \{ 2 \}
  \{ 3 \} \text{type positive = 1..maxint;}
  \{ 4 \} \text{var x,y: positive;}
  \{ 5 \}
  \{ 6 \} \begin{array}{}{\begin{array}{}
  \{ 7 \} \text{writeln('Enter two naturals: ');} \\
  \{ 8 \} \text{readln(x, y);} \\
  \{ 9 \}
  \{10\} \text{while x <> y do}
  \{11\} \text{if x < y then x := x - 'y'} \\
  \{12\} \text{else y := y - x;}
  \{13\}
  \{14\} \text{writeln('Greatest Common Divisor: ', z);} \\
  \{15\} \text{end.}
\end{array} \end{array}
\}

Note that the programmer has made three errors: in line 1 he forgot mentioning the program parameter input although he is using the function readln; in line 11 he by accident quoted y in x := x - 'y'; and in line 14 he typed writeln(... z) instead of writeln(... x).

Putting the function tc around this program, and passing the term to the term rewriting system that can be derived from the entire specification, will have the following result. In each paragraph, we mention the module name, and the tags of the most important equations involved.

**TC-program.** [pr1], [ts2]: Since the header contains input as a program parameter, the declaration var input: text is added to the environment.

**TC-block.** [bl1], [bl6], [be8]: The checks for the block are split up into consecutive checks of the type definitions (line 3), the variable declarations (line 4), and the statements (line 6-15).

**TC-defs.** [ts4]: The type definition is broken up into checks for the left and right-hand side, after which the type definition type positive = 1..maxint is added to the environment.

**ShouldBe.** [md1] and **Context-search.** [ex2]. [fi1]: Since the identifier positive is not yet declared in this block, the left hand side is correct.

**TC-defs.** [ts1], [cs5], [cs6]: The right-hand side involves checking a subrange type declaration. Both constants (1 and maxint) are checked, and maxint is found in the part of the context of the environment containing the predefined identifiers. Since the types of both constants are the same and ordinal, and since the value of maxint, 32768, is more than the value of 1, 1, the subrange declaration is correct.

**TC-defs.** [ts3], [ts4]: The checks for the variable declaration of line 4 are split up into checks for left and right hand side of var x: positive and var y: positive. Since the declaration for positive is found in the context of the environment, and neither x nor y are declared, all these checks succeed, and the declarations for x and y are added to the environment.

**TC-stat.** [se6], [se8]: The compound statement is thus checked in an environment containing the definitions of the program. The checks for the compound statements are split up into checks for the individual statements.

**TC-stat.** [se5] and **TC-exp-calls.** [p1]: The first statement to be checked is a procedure call, and since the identifier writeln is found and detected as a predefined identifier in the
environment, the type check function for special procedures is called.

**TC-spec-procs [ln7], [ln9], [wr6]:** Since the first argument of writeln is not a file variable, writeln( '...' ) is rewritten to writeln(output, '...' ), which in turn is broken up into writeln(output) and write(output, '...' ). Because output is a variable of type text, and the argument of write is a string, it is concluded that the statement is correct. Consequently the environment is kept unchanged.

**TC-spec-procs [ln2], [ln4] and [es7]:** The readln of line 8 is checked in a similar way: This time, however, variable input is not declared (since it was not mentioned in the program header). Therefore, the checks of the actual parameters produce the messages readln(x,y): Identifier input not declared. and readln(x,y): Identifier input is not a variable. Moreover, no type can be found for the file argument, producing readln(x,y): Types are not assignment-compatible. Finally, the checks for a single readln with one arguments raises [ln5] will add the message readln(x,y): input is not a text file.

So line 8 of the program causes 4 error messages to be added to the environment. Note that these messages are prefixed by the construct field of the environment, which is set in the checks for the statements.

**TC-stat [ws1]:** For the while statement, the condition and body are checked, after which it is confirmed that the type of the expression is of type Boolean.

**TC-expr [r2], [es2], [es1], [ms1]:** For the expression x -> y, first the two arguments are checked, detecting that they are expressions of type integer. Thus, the arguments are simple types, and the entry -> | simple | simple | boolean is found. Consequently, it is concluded that the expression is correct and of type Boolean.

**TC-stat [ms2] [as1] [as3], TC-expr-vars [v1] and TC-expr [r2]:** The body of the while statement is the if statement of lines 11 and 12. The condition of the if statement is checked similar to the condition of the while statement, just as the expressions on the right hand sides of the assignment statements in the conclusions of the if statement. However, x = 'y' results in trying to lookup an entry that does not exist, so an error message is added: x := x = 'y': Illegal operand types for operator '='. Consequently, it cannot be concluded that the types of the first assignment are compatible, resulting in: x := x = 'y': Types are not assignment-compatible.

**TC-spec-proc, [wr6], and TC-expr, [es4] in the last writeln statement (the check of which is, of course, similar to the check of the writeln of line 7), the identifier z is not known.** The several constraints thus raise messages Identifier z not declared. and Identifier z is not a constant, variable or function. The check for the write does not know one of its argument types and complains with Number, character, boolean, or string type expected.

Finally, since in module Errors equation [es1] eliminates double error messages, the total list of errors resulting from equation [pr1] is:

- readln(x,y): Identifier input not declared.
- readln(x,y): Identifier input is not a variable.
- readln(x,y): Types are not assignment-compatible.
- readln(x,y): input is not a text file.

- x := x = 'y': Illegal operand types for operator '='.
- x := x = 'y': Types are not assignment-compatible.

- writeln( 'Greatest Common Divisor: ', z): Identifier z not declared.
- writeln( 'Greatest Common Divisor: ', z): Identifier z is not a constant, variable, or function.
- writeln( 'Greatest Common Divisor: ', z): Number, character, boolean, or string type expected.