The PREV'22 Programming Language

1 Lexical structure

Programs in the PREV'22 programming language are written in ASCII character set (no additional characters denoting post-alveolar consonants are allowed).

Programs in the PREV'22 programming language consist of the following lexical elements:

- Constants:
 - constant of type void: none
 - constants of type boolean: true false
 - constants of type integer:

A nonempty finite string of digits (0...9), not 0 padded but optionally preceded by a sign (+ or -).

- constants of type char:

A character with ASCII code in range {32...126} enclosed in single quotes ('); a single quote in the constant (except at the beginning and at the end) must be preceded with backslash (\).

- string constants:

A (possibly empty) string of character with ASCI codes in range $\{32...126\}$ enclosed in double quotes ("); each double quote in the constant (except at the beginning and at the end) must be preceded with a backslash (\).

- constants of pointer types: nil
- Symbols:

```
() { } [ ] . , : ; & | ! == != < > <= >= * / % + - ^ =
```

• Keywords:

bool char del do else fun if int new then typ var void where while

• Identifiers:

A nonempty finite string of letters (A...Z and a...z), digits (0...9), and underscores $(_)$ that (a) starts with either a letter or an underscore and (b) is not a keyword or a constant.

• Comments:

A string of characters starting with a hash (#) and extending to the end of line.

• White space:

Space, horizontal tab (HT), line feed (LF) and carriage return (CR). Line feed alone denotes the end of line within a source file. Horizontal tab is 8 spaces wide.

Lexical elements are recognised from left to right using the longest match approach.

2 Syntax structure

The concrete syntax of the PREV'21 programming language is defined by context free grammar with the start symbol prg and productions

```
\begin{array}{lll} (\operatorname{program}) & prg & \longrightarrow & decl \; \{\; decl \; \} \\ (\operatorname{type \; declaration}) & decl & \longrightarrow & \operatorname{typ \; identifier} = \; type \\ (\operatorname{variable \; declaration}) & decl & \longrightarrow & \operatorname{var \; identifier} \; : \; type \\ (\operatorname{function \; declaration}) & decl & \longrightarrow & \operatorname{fun \; identifier} \; (\; [\operatorname{identifier} \; : \; type \; \{\; , \operatorname{identifier} \; : \; type \; \}\; ] \; ) \; : \; type \; [= \; expr \; ] \end{array}
```

```
(atomic type)
                                        void | char | int | bool
                                         identifier
(named type)
                            type
                                         [ expr ] type
(array type)
                            type
                                         ^ type
(pointer type)
                            type
(record type)
                                         { identifier : type { , identifier : type } }
                            type
(enclosed type)
                                         (type)
                            type
(constant expression)
                            expr
                                         const
(variable access)
                                         identifier
                            expr
                                         identifier ( [expr \{, expr \}] )
(function call)
                                         (new | del) expr
(allocation expression)
                           expr
(compound expression)
                                         { stmt; { stmt; } }
(typecast expression)
                                        (expr:type)
(enclosed expression)
                                        (expr)
                                        expr (& | | | == | ! = | < | > | <= | >= | * | / | % | + | - ) expr
(infix expression)
                            expr
(prefix expression)
                                         (! | + | - | ^) expr
(postfix expression)
                            expr
                                         expr ( [ expr ] | ^{\sim} | . identifier )
(where expression)
                            expr
                                         expr where { decl { decl } }
(expression statement)
                           stmt \longrightarrow
                                         expr
(assignment statement)
                           stmt \longrightarrow
                                         expr = expr
(conditional statement)
                           stmt \longrightarrow
                                         if expr then stmt else stmt
(loop statement)
                           stmt \longrightarrow
                                         while expr do stmt
```

where const denotes any constant.

Relational operators are non-associative, all other infix operators are left associative.

The precedence of the operators is as follows:

```
postfix operators [] ^ .

prefix operators ! + - ^ new del

multiplicative operators * / %

additive operators + -

relational operators == != < > <= >=

conjunctive operator &

disjunctive operator |

declaration binder where THE LOWEST PRECEDENCE
```

In the grammar above, braces typeset as {} enclose sentential forms that can be repeated zero or more times, brackets typeset as [] enclose sentential forms that can be present or not while braces typeset as {} and brackets typeset as [] denote symbols that are a part of the program text.

3 Semantic structure

3.1 Name binding

Let function $[\cdot]_{BIND}$ bind a name to its declaration according to the rules of namespaces and scopes described below. Hence, the value of function $[\cdot]_{BIND}$ depends on the context of its argument.

Namespaces. There are two kinds of a namespaces:

- 1. Names of types, functions, variables and parameters belong to one single global namespace.
- 2. Names of record components belong to record-specific namespaces, i.e., each record defines its own namespace containing names of its components.

Scopes. A new scope is created in two ways:

1. Where-expression

$$expr$$
 where { $decl$ { $decl$ } }

creates a new scope: it starts with expr and ends with $\}$.

2. Function declaration

creates a new scope. The name of a function, the types of parameters and the type of a result belong to the outer scope while the names of parameters and the expression (if present) denoting the function body belong to the scope created by the function declaration.

All names declared within a given scope are visible in the entire scope unless hidden by a declaration in the nested scope. A name can be declared within the same scope at most once.

3.2 Type system

The set

$$\mathcal{T}_{d} = \{ \mathbf{void}, \mathbf{char}, \mathbf{int}, \mathbf{bool} \}$$

$$\cup \{ \mathbf{arr}(n \times \tau) \mid n > 0 \land \tau \in \mathcal{T}_{d} \}$$

$$\cup \{ \mathbf{rec}_{id_{1}, \dots, id_{n}}(\tau_{1}, \dots, \tau_{n}) \mid n > 0 \land \tau_{1}, \dots, \tau_{n} \in \mathcal{T}_{d} \}$$

$$\cup \{ \mathbf{ptr}(\tau) \mid \tau \in \mathcal{T}_{d} \}$$
(atomic types)
(arrays)
(records)
(pointers)

denotes the set of all data types of PREV'21. The set

$$\mathcal{T} = \mathcal{T}_d \qquad \text{(data types)}$$

$$\cup \{ (\tau_1, \dots, \tau_n) \to \tau \mid n \ge 0 \land \tau_1, \dots, \tau_n, \tau \in \mathcal{T}_d \} \qquad \text{(functions)}$$

denotes the set of all types of PREV'21.

Types τ_1 and τ_2 are equal if (a) $\tau_1 = \tau_2$ or (b) if they are type synonyms (introduced by chains of type declarations) of types τ_1' and τ_2' where $\tau_1' = \tau_2'$.

Semantic functions

$$[\![\cdot]\!]_{\mathrm{ISTYPE}}: \mathcal{P} \to \mathcal{T} \quad \mathrm{and} \quad [\![\cdot]\!]_{\mathrm{OFTYPE}}: \mathcal{P} \to \mathcal{T}$$

map syntactic phrases of PREV'21 to types. Function $[\cdot]_{ISTYPE}$ denotes the type described by a phrase, function $[\cdot]_{OFTYPE}$ denotes the type of a value described by a phrase.

The following assumptions are made in the rules below:

- Function val maps lexemes to data of the specified type.
- $\tau \in \mathcal{T}_d$ unless specified otherwise.

Type expressions.

$$\frac{\llbracket type \rrbracket_{\text{ISTYPE}} = \tau \quad \text{val}(int) = n}{0 < n \le 2^{63} - 1 \quad \tau \in \mathcal{T}_d \setminus \{\text{void}\}} \\
\boxed{\llbracket [int] type \rrbracket_{\text{ISTYPE}} = \operatorname{arr}(n \times \tau)} \tag{T2}$$

$$\frac{\forall i \in \{1 \dots n\}: \llbracket type_i \rrbracket_{\text{ISTYPE}} = \tau_i \wedge \tau_i \in \mathcal{T}_d \setminus \{\mathbf{void}\}\}}{\llbracket \{id_1: type_1, \dots, id_n: type_n\} \rrbracket_{\text{ISTYPE}} = \mathbf{rec}_{id_1, \dots, id_n}(\tau_1, \dots, \tau_n)}$$
(T3)

$$\frac{\llbracket type \rrbracket_{\text{ISTYPE}} = \tau \quad \tau \in \mathcal{T}_d}{\llbracket \uparrow type \rrbracket_{\text{ISTYPE}} = \mathbf{ptr}(\tau)}$$
(T4)

$$\frac{[type]_{ISTYPE} = \tau}{[(type)]_{ISTYPE} = \tau}$$
(T5)

Value expressions.

$$\boxed{\llbracket \mathsf{none} \rrbracket_{\mathsf{OFTYPE}} = \mathbf{void}} \quad \boxed{\llbracket \mathsf{nil} \rrbracket_{\mathsf{OFTYPE}} = \mathbf{ptr}(\mathbf{void})} \quad \boxed{\llbracket \mathit{string} \rrbracket_{\mathsf{OFTYPE}} = \mathbf{ptr}(\mathbf{char})}$$
(v1)

$$\frac{1}{\llbracket bool \rrbracket_{\text{OFTYPE}} = \mathbf{bool}} \frac{1}{\llbracket char \rrbracket_{\text{OFTYPE}} = \mathbf{char}} \frac{1}{\llbracket int \rrbracket_{\text{OFTYPE}} = \mathbf{int}}$$
(v2)

$$\frac{\llbracket expr_1 \rrbracket_{\text{OFTYPE}} = \mathbf{bool} \quad \llbracket expr_2 \rrbracket_{\text{OFTYPE}} = \mathbf{bool} \quad op \in \{\&, | \}}{\llbracket expr_1 \ op \ expr_2 \rrbracket_{\text{OFTYPE}} = \mathbf{bool}}$$

$$(v4)$$

$$\underbrace{\llbracket expr_1 \rrbracket_{\text{OFTYPE}} = \mathbf{int} \quad \llbracket expr_2 \rrbracket_{\text{OFTYPE}} = \mathbf{int} \quad op \in \{+, -, *, /, \%\}}_{\llbracket expr_1 \ op \ expr_2 \rrbracket_{\text{OFTYPE}} = \mathbf{int}}$$
(v5)

$$\frac{[expr_1]_{OFTYPE} = \tau \quad [expr_2]_{OFTYPE} = \tau}{\tau \in \{\mathbf{bool}, \mathbf{char}, \mathbf{int}\} \cup \{\mathbf{ptr}(\tau) \mid \tau \in \mathcal{T}_d\} \quad op \in \{==, !=\}}{[expr_1 \ op \ expr_2]_{OFTYPE} = \mathbf{bool}}$$
(v6)

$$\frac{[expr_1]_{\text{OFTYPE}} = \tau \quad [expr_2]_{\text{OFTYPE}} = \tau}{\tau \in \{\mathbf{char}, \mathbf{int}\} \cup \{\mathbf{ptr}(\tau) \mid \tau \in \mathcal{T}_d\} \quad op \in \{<=, >=, <, >\}}{[expr_1 op \ expr_2]_{\text{OFTYPE}} = \mathbf{bool}}$$
(v7)

$$\frac{[\![expr]\!]_{\text{OFTYPE}} = \tau}{[\![\hat{r} expr]\!]_{\text{OFTYPE}} = \mathbf{ptr}(\tau)} \qquad \frac{[\![expr]\!]_{\text{OFTYPE}} = \mathbf{ptr}(\tau)}{[\![expr]\!]_{\text{OFTYPE}} = \tau}$$
(v8)

$$\frac{[\![expr]\!]_{\text{OFTYPE}} = \mathbf{int}}{[\![new\ expr]\!]_{\text{OFTYPE}} = \mathbf{ptr}(\mathbf{void})} \qquad \frac{[\![expr]\!]_{\text{OFTYPE}} = \mathbf{ptr}(\tau)}{[\![del\ expr]\!]_{\text{OFTYPE}} = \mathbf{void}}$$
(v9)

$$\frac{\forall i \in \{1 \dots n\}: [stmt_i]_{\text{OFTYPE}} = \tau_i}{[[stmt_1; \dots stmt_n; \}]_{\text{OFTYPE}} = \tau_n} \tag{v13}$$

$$\underbrace{\llbracket expr \rrbracket_{\text{OFTYPE}} = \tau_1 \quad \llbracket type \rrbracket_{\text{ISTYPE}} = \tau_2 \quad \tau_1, \tau_2 \in \{ \mathbf{char}, \mathbf{int} \} \cup \{ \mathbf{ptr}(\tau) \mid \tau \in \mathcal{T}_d \}}_{\left[(expr : type) \right]_{\text{OFTYPE}} = \tau_2}$$
(v14)

Statements.

$$\frac{\llbracket expr_1 \rrbracket_{\text{OFTYPE}} = \tau \quad \llbracket expr_2 \rrbracket_{\text{OFTYPE}} = \tau \quad \tau \in \{ \mathbf{bool}, \mathbf{char}, \mathbf{int} \} \cup \{ \mathbf{ptr}(\tau) \mid \tau \in \mathcal{T}_d \}}{\llbracket expr_1 = expr_2 \rrbracket_{\text{OFTYPE}} = \mathbf{void}}$$
(S1)

$$\frac{[\![expr]\!]_{\text{OFTYPE}} = \mathbf{bool} \quad [\![stmts]\!]_{\text{OFTYPE}} = \tau}{[\![\text{while } expr \text{ do } stmts]\!]_{\text{OFTYPE}} = \mathbf{void}}$$
(s3)

Declarations.

$$\frac{[[identifier]]_{BIND} = typ \ identifier: type \ [[type]]_{ISTYPE} = \tau}{[[identifier]]_{ISTYPE} = \tau}$$
(D1)

3.3 Lvalues

The semantic function

$$\llbracket \cdot \rrbracket_{\mathrm{ISADDR}} : \mathcal{P} \to \{\mathbf{true}, \mathbf{false}\}$$

denotes which phrases represent lvalues.

In all other cases the value of $\llbracket \cdot \rrbracket_{\text{ISADDR}}$ equals **false**.

3.4 Operational semantics

Operational semantics is described by semantic functions

$$\begin{split} & \llbracket \cdot \rrbracket_{\mathrm{ADDR}} \, : \, \mathcal{P} \times \mathcal{M} \to \mathcal{I} \times \mathcal{M} \\ & \llbracket \cdot \rrbracket_{\mathrm{EXPR}} \, : \, \mathcal{P} \times \mathcal{M} \to \mathcal{I} \times \mathcal{M} \\ & \llbracket \cdot \rrbracket_{\mathrm{STMT}} \, : \, \mathcal{P} \times \mathcal{M} \to \mathcal{I} \times \mathcal{M} \end{split}$$

where P denotes the set of phrases of PREV'21, I denotes the set of 64-bit integers, and M denotes possible states of the memory. Unary operators and binary operators perform 64-bit signed operations (except for type **char** where operations are performed on the lower 8 bits only).

Auxiliary function addr returns either an absolute address for a static variable or a string constant or an offset for a local variable, parameter or record component. Auxiliary function size of returns the size of a type. Auxiliary function val returns the value of an integer constant or an ASCII code of a char constant.

Addresses.

$$[string]_{ADDR}^{M} = \langle addr(string), M \rangle$$
(A1)

$$\frac{\operatorname{addr}(identifier) = a}{\llbracket identifier \rrbracket_{\text{ADDR}}^{M} = \langle a, M \rangle}$$
(A2)

$$\frac{\llbracket expr_1 \rrbracket_{\text{ADDR}}^{\text{M}} = \langle n_1, \text{M}' \rangle \quad \llbracket expr_2 \rrbracket_{\text{EXPR}}^{\text{M}'} = \langle n_2, \text{M}'' \rangle \quad \llbracket expr_1 \rrbracket_{\text{OFTYPE}} = \mathbf{arr}(n \times \tau)}{\llbracket expr_1 \llbracket expr_2 \rrbracket_{\text{ADDR}}^{\text{M}} = \langle n_1 + n_2 * \text{sizeof}(\tau), \text{M}'' \rangle}$$
(A3)

$$\frac{[expr]_{ADDR}^{M} = \langle n_1, M' \rangle}{[expr.identifier]_{ADDR}^{M} = \langle n_1 + addr(identifier), M' \rangle}$$
(A4)

$$\frac{[\![expr]\!]_{\text{EXPR}}^{M} = \langle n, M' \rangle}{[\![\uparrow expr]\!]_{\text{ADDR}}^{M} = \langle n, M' \rangle}$$
(A5)

Expressions.

$$\boxed{ \text{[true]}_{\text{EXPR}}^{\text{M}} = \langle 1, \text{M} \rangle } \quad \boxed{ \text{[false]}_{\text{EXPR}}^{\text{M}} = \langle 0, \text{M} \rangle }$$

$$\frac{[[char]_{\text{EXPR}}^{\text{M}} = \langle \text{val}(char), \text{M} \rangle}{[[int]_{\text{EXPR}}^{\text{M}} = \langle \text{val}(int), \text{M} \rangle}$$
(EX3)

$$\frac{\llbracket expr \rrbracket_{\text{EXPR}}^{M} = \langle n, M' \rangle \quad \text{op} \in \{!, +, -\}}{\llbracket \text{op} \ expr \rrbracket_{\text{EXPR}}^{M} = \langle \text{op} \ n, M' \rangle}$$
(EX4)

$$\underbrace{\llbracket expr_1 \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n_1, \text{M}' \rangle \quad \llbracket expr_2 \rrbracket_{\text{EXPR}}^{\text{M}'} = \langle n_2, \text{M}'' \rangle \quad \text{op} \in \{ \mid, \&, ==, !=, <, >, <=, >=, +, -, *, /, \% \}}_{\llbracket expr_1 \text{ op} \quad expr_2 \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n_1 \text{ op} \quad n_2, \text{M}'' \rangle}$$
(EX5)

$$\frac{\llbracket expr \rrbracket_{\text{ADDR}}^{\text{M}} = \langle n, \text{M}' \rangle}{\llbracket \hat{r} expr \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n, \text{M}' \rangle} \frac{\llbracket expr \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n, \text{M}' \rangle}{\llbracket expr \hat{r} \rrbracket_{\text{EXPR}}^{\text{M}} = \langle \text{M}'[n], \text{M}' \rangle}$$
(EX6)

$$\frac{ [\![expr]\!]_{\text{EXPR}}^{\text{M}} = \langle n, \text{M}' \rangle \quad [\![new(n)]\!]_{\text{EXPR}}^{\text{M}'} = \langle a, \text{M}'' \rangle}{ [\![new \, expr]\!]_{\text{EXPR}}^{\text{M}} = \langle a, \text{M}'' \rangle}$$
(EX7)

$$\frac{ \left[expr \right]_{\text{EXPR}}^{\text{M}} = \langle a, \text{M}' \rangle \quad \left[\text{del}(a) \right]_{\text{EXPR}}^{\text{M}'} = \langle \text{undef}, \text{M}'' \rangle}{ \left[\text{del} \ expr \right]_{\text{EXPR}}^{\text{M}} = \langle \text{undef}, \text{M}'' \rangle}$$
(EX8)

$$\frac{\operatorname{addr}(identifier) = a}{\llbracket identifier \rrbracket_{\operatorname{EXPR}}^{\operatorname{M}} = \langle M[a], \operatorname{M} \rangle}$$
(EX9)

$$\frac{\left[\left[expr_{1}\left[expr_{2}\right]\right]\right]_{\text{ADDR}}^{M} = \langle a, M' \rangle}{\left[\left[expr_{1}\left[expr_{2}\right]\right]\right]_{\text{EXPR}}^{M} = \langle M'[a], M' \rangle} \tag{EX10}$$

$$\frac{[[expr.identifier]]_{ADDR}^{M} = \langle a, M' \rangle}{[[expr.identifier]]_{EXPR}^{M} = \langle M'[a], M' \rangle}$$
(EX11)

$$\frac{[expr_1]_{\text{EXPR}}^{\text{M}_0} = \langle n_1, \text{M}_1 \rangle \dots [expr_m]_{\text{EXPR}}^{\text{M}_{m-1}} = \langle n_m, \text{M}_m \rangle}{[identifier(expr_1, \dots, expr_m)]_{\text{EXPR}}^{\text{M}_0} = \langle identifier(n_1, \dots, n_m), \text{M}_m \rangle}$$
(Ex12)

$$\frac{[stmt_1]_{\text{STMT}}^{M_0} = \langle n_1, M_1 \rangle \dots [stmt_m]_{\text{STMT}}^{M_{m-1}} = \langle n_m, M_m \rangle}{[stmt_1, \dots, stmt_m]_{\text{EXPR}}^{M_0} = \langle n_m, M_m \rangle}$$
(EX13)

$$\frac{[expr]_{\text{EXPR}}^{M} = \langle n, M' \rangle}{[(expr)]_{\text{EXPR}}^{M} = \langle n, M' \rangle}$$
(EX14)

$$\frac{\llbracket expr \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n, \text{M}' \rangle \quad \llbracket type \rrbracket_{\text{ISTYPE}} \neq \mathbf{char}}{\llbracket (expr: type) \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n, \text{M}' \rangle}$$
(EX15)

$$\frac{\llbracket expr \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n, M' \rangle \quad \llbracket type \rrbracket_{\text{ISTYPE}} = \mathbf{char}}{\llbracket (expr: type) \rrbracket_{\text{EXPR}}^{\text{M}} = \langle n \mod 256, M' \rangle}$$
(EX16)

$$\frac{[expr]_{\text{EXPR}}^{M} = \langle n, M' \rangle}{[expr \text{ where } \{decls\}]_{\text{EXPR}}^{M} = \langle n, M' \rangle}$$
(EX17)

Statements.

$$\frac{[\![expr]\!]_{\text{EXPR}}^{M} = \langle n, M' \rangle}{[\![expr]\!]_{\text{STMT}}^{M} = \langle n, M' \rangle}$$
(ST1)

$$\frac{ [\![expr]\!]_{\mathrm{EXPR}}^{\mathrm{M}} = \langle \mathbf{true}, \mathbf{M}' \rangle \quad [\![stmt_1]\!]_{\mathrm{STMT}}^{\mathrm{M}'} = \langle \mathrm{undef}, \mathbf{M}'' \rangle}{ [\![\mathrm{if} \ expr \ \mathrm{then} \ stmt_1 \ \mathrm{else} \ stmt_2]\!]_{\mathrm{STMT}} = \langle \mathrm{undef}, \mathbf{M}'' \rangle}$$
 (ST3)

$$\frac{[\![expr]\!]_{\text{EXPR}}^{\text{M}} = \langle \mathbf{false}, \mathbf{M}' \rangle \quad [\![stmt_2]\!]_{\text{EXPR}}^{\mathbf{M}'} = \langle \text{undef}, \mathbf{M}'' \rangle}{[\![\![if]\!]\!] expr \text{ then } stmt_1 \text{ else } stmt_2]\!]_{\text{STMT}} = \langle \text{undef}, \mathbf{M}'' \rangle}$$
(ST4)

$$\frac{[\![\mathit{expr}]\!]_{\mathrm{EXPR}}^{\mathrm{M}} = \langle \mathbf{false}, \mathrm{M}' \rangle}{[\![\![\mathsf{while}\ \mathit{expr}\ \mathsf{do}\ \mathit{stmt}]\!]_{\mathrm{STMT}}^{\mathrm{M}} = \langle \mathrm{undef}, \mathrm{M}' \rangle}$$
(ST6)