**The ZINC Programming Language**

***COSC 4316, Spring 2018***

**Introduction**

ZINC (for **Z**INC **I**s **N**ot **C**) is a language that is similar to Ada. It is designed to provide you the experience writing a compiler without getting burdened by all of the complexities and details of a complete, standard programming language. (ZINC is adapted from an instructional language formerly used by Michael Franz in his CS241 course at University of California, Irvine.)

**Example ZINC Program**

**1 program**

**2 var N : integer;**

**3 var SQRT : integer;**

**4 begin**

**5 N := readInt;**

**6 SQRT := 0;**

**7**

**8 -- go until SQRT exceeds the square root of N }**

**9 while SQRT \* SQRT <= N loop**

**10 SQRT := SQRT + 1;**

**11 endwhile;**

**12**

**13 SQRT := SQRT - 1; -- subtract one SQRT is <= sqrt(N)**

**14**

**15 writeInt(SQRT);**

**16**

**17 end**

**Language Overview**

**Lexical Features**

The ZINC language is lexically simple. All identifiers start with a capital letter, and may contain only numbers, capital letters, and underscores. All keywords start with lower-case letters or are a symbol. The symbols "(", ")", ":=", ";","\*\*", "\*", "div", "mod", "+", "-", "=", "<>", "<", "<=", ">", ">=" are used.

**Data Types**

ZINC supports 32-bit integers ("integer"). Variables are always declared to be of a particular type, but the only type is integer.

**Operators**

ZINC has several infix binary operators that work on integer operands. The multiplication "\*", division "div", modulus "mod", addition "+", and subtraction "-" produce integer results. The comparison operators (i.e., equals "=", not equal "<>", less than "<", less-than or equal-to "<=", greater than ">", and greater-than or equal-to ">=") all produce Boolean (0 or 1) values as results.

**Control Structures**

ZINC is a structured programming language. The only control structures supported are "if" and "while" statements. Both take a boolean expression that guards the body of the control structure. In the case of an "if" statement, the statements after the "then" are executed if the expression is true, and the statements after the "else" (if there is one) are executed if the expression is false. In the case of the "while" statement, the loop is exited if the expression false; otherwise if the expression is true, the body will be executed, and then the expression will be re-evaluated.

**Assignment**

Assignments are a kind of statement rather than a kind of operator. The ":=" keyword is used to separate the left hand side (which is the variable being assigned to) from the right hand side, which is an expression that must be of the same type as the left hand side.

**Built-in Procedures**

ZINC does not support user-defined functions or procedures, but it does support one built-in procedures "writeInt" that outputs an integer and a new-line to the console (respectively), and one user-defined function, "readInt" that reads an integer from the console. The syntax for these is hard-coded into ZINC's BNF grammar.

**Syntax**

**Comments**

The occurrence of the character pair "--" (two consecutive hyphens) on a line denotes the start of a comment that extends to the end of that line. For purposes of determining the lexical elements of the source file, the entire comment will be treated as if it were whitespace.

**Lexical Elements**

If the definition of a lexical element is in quotes, then it is meant to match exactly, the contained string. Otherwise, it is a regular expression. Square brackets in regular expressions are used as an abbreviation for matching ranges of letters. For example, [0-9] matches any digit, and [a-zA-Z] matches any English letter in capital or lower case.

**Numbers, Literals, and Identifiers:**

* **num = (+|-)?[1-9][0-9]\*|0**

The regular expression allows all natural numbers, but since we are using 32-bit integers, only   
-2147483648 through 2147483647 are valid integer constants. Integer constants outside of that range should be flagged as illegal.

* **ident = [A-Z][\_A-Z0-9]\***

**Symbols and Operators:**

* **LP = "("**
* **RP = ")"**
* **ASGN = ":="**
* **SC = ";"**
* **COLON = ":"**
* **POWER = "\*\*"**
* **MULTIPLICATIVE = "\*" | "div" | "mod"**
* **ADDITIVE = "+" | "-"**
* **COMPARE = "=" | "<>" | "<" | ">" | "<=" | ">="**

**Keywords:**

* **IF = "if"**
* **THEN = "then"**
* **ELSE = "else"**
* **BEGIN = "begin"**
* **END = "end"**
* **ENDIF = "endif"**
* **ENDWHILE = "endwhile"**
* **WHILE = "while"**
* **LOOP = "loop"**
* **PROGRAM = "program"**
* **VAR = "var"**
* **INT = "integer"**

**Built-in Procedures:**

* **WRITEINT = "writeInt"**
* **READINT = "readInt"**

**BNF Grammar**

**<program> ::= PROGRAM <declarations> BEGIN <statementSequence> END**

**<declarations> ::= VAR ident COLON <type> SC <declarations>**

**| ε**

**<type> ::= INT**

**<statementSequence> ::= <statement> SC <statementSequence>**

**| ε**

**<statement> ::= <assignment>**

**| <ifStatement>**

**| <whileStatement>**

**| <writeInt>**

**| ε**

**<assignment> ::= ident ASGN <expression>**

**| ident ASGN READINT**

**<ifStatement> ::= IF <expression> THEN <statementSequence> <elseClause> ENDIF**

**<elseClause> ::= ELSE <statementSequence>**

**| ε**

**<whileStatement> ::= WHILE <expression> LOOP <statementSequence> ENDWHILE**

**<writeInt> ::= WRITEINT <expression>**

**<expression> ::= <simpleExpression>**

**| <simpleExpression> COMPARE <expression>**

**<simpleExpression> ::= <term> ADDITIVE <simpleExpression>**

**| <term>**

**<term> ::= <factor> MULTIPLICATIVE <term>**

**| <factor>**

**<factor> ::= <primary> POWER <factor> | <primary>**

**<primary> ::= ident**

**| num**

**| LP <expression> RP**

**Operator Precedence and Associativity**

The order of precedence among the operators is:

1. The POWER operator.
2. The MULTIPLICATIVE operators.
3. The ADDITIVE operators.
4. The COMPARE operators.

All binary operators are left-associative.

**Type Rules**

1. The operands of all MULTIPLICATIVE, ADDITIVE, and COMPARE operators must be integers
2. The MULTIPLICATIVE and ADDITIVE operators create an integer result.
3. The COMPARE operators create **integer** results (0 = “false”, non-zero = “true”);
4. All variables must be declared, and may be declared only once.
5. The left-hand of assignment must be a variable, and the right-hand side must be an expression of the variable's type.

**Semantics**

* Only those variables which have been declared can be assigned to or used.
  + All integer variables and array elements are considered to have initial values of "0".
* All binary operators operate on signed integer operands:
  + "x \* y" results in the product of x and y.
  + "x div y" which results in the integer quotient of x divided by y.

The behavior is not defined if y is 0.

* + "x mod y" results in the remainder of x divided by y when x is non-negative and y is positive.   
     Otherwise, the result is undefined.
  + “x \*\* y” results in x being raised to the power of y when y is non-negative. Otherwise, the result is undefined.
  + "x + y" results in the sum of x and y.
  + "x - y" is the difference of y subtracted from x.
  + "x = y" is true if x and y are the same, otherwise it is false.
  + "x <> y" is false if x and y are the same, otherwise it is true.
  + "x < y" is true if x is less than y, otherwise it is false.
  + "x > y" is true if x is greater than y, otherwise it is false.
  + "x <= y" is true if x is less than or equal to y, otherwise it is false.
  + "x >= y" is true if x is greater than or equal to y, otherwise it is false.
  + Computations on ZINC integers should be done using a 32-bit 2's complement representation. Overflowing computations should simply "wrap around"; that is, the result of all integer operations should be the integer that is not less than -231, not more than 231-1, and congruent modulus 232 with the result of performing the normal mathematical operation on actual mathematical integers.
* "if" statements evaluate their expression; if the expression is true, then the "then-statements" are executed, if it is false, the "else-statements" are executed.
* "while" statements first evaluates its expression. If it is false, execution continues after the end of the "while" statement. If it is true, the statements in the body of the "while" loop are executed. After they finish executing, the expression is re-evaluated. As long as the expression is true, the process repeats itself, alternatively evaluating the expression and executing the statements in the body. Once the expression is false, execution continues after the end of the "while" loop.
* "writeInt" evaluates its expression and outputs the result to the console and causes the cursor to move to the beginning of the next line.
* "readInt" reads an integer from the console and updates an integer variable to hold that value.