CS 345H - Programming Languages: Honors
Assignment 2

1 Programming in Prolog (60 Points)

The task is to implement a parser and evaluator for mySIMPL (my Simple IMperative Programming Language). The core concept of the language is arithmetic expressions over floating points numbers and variables. It has additional imperative constructs: declarations, assignments, and a return statement which terminates the program and returns a number.

The syntax of mySIMPL is:

\[
\begin{align*}
\text{prog} &::= \text{retStatement} \cdot | \text{declaration} \cdot \text{prog} | \text{assignment} \cdot \text{prog} \\
\text{declaration} &::= \text{var} < \text{id} > \\
\text{assignment} &::= < \text{id} > := \text{base} \\
\text{retStatement} &::= \text{return} \text{base} \\
\text{base} &::= < \text{id} > | < \text{number} > | '(\text{expression} \ ')'
\end{align*}
\]

\[
\begin{align*}
\text{expression} &::= [\text{expression} \text{addOp}] \text{term} \\
\text{term} &::= [\text{term} \text{mulOp}] \text{factor} \\
\text{factor} &::= \text{base} \\
\text{addOp} &::= '+' | '-' \\
\text{mulOp} &::= '*' | '/'
\end{align*}
\]

For simplicity, we assume that mySIMPL programs are given as token (prolog) lists. An example of a program is:

\begin{verbatim}
var x;
\ 
x := (5 * 2);
\ 
return (x + 1).
\end{verbatim}

which is represented by the token list:

\[
[\text{['var', 'x', ';', 'x', ':=', '('}, \text{5, '*', 2, ')'}, ';', \text{return}, '('}, \text{'x', '+', 1, ')'}, '.']
\]
Implement a predicate
\[ \text{parse}(\text{+TokenList, -AST}) \]
which parses the token list and returns an abstract syntax tree (no constraints on the format) for syntactically correct programs and fail for incorrect programs.

Then implement a predicate
\[ \text{evaluate}(\text{+AST, -Number}) \]
which evaluates the program’s abstract syntax tree and returns a numeric value for semantically correct program and fails for incorrect programs.

A program is syntactically correct if it can be expressed through the given grammar. A program is semantically correct if furthermore no undeclared or uninitialized variables are used in any statement.

The language is very simple and lacks fundamental concepts like control flow, types, or functions but we will extend it in future assignments.

**Additional Information and Hints**

- Prolog features DCGs for building parsers in an intuitive way.
- The grammar has left-recursive elements in it. For better parsing, a format like
  \[
  \text{expr}(E) \rightarrow \text{term}(T), [+], \text{left_assoc}(E,T,\text{plus}).
  \]
  \[
  \text{left_assoc}(	ext{expression}(\text{plus},T,T1),T,\text{plus}) \rightarrow \text{term}(T1).
  \]
  is preferable. This code produces an AST node called 'expression’ with arguments 'plus' and the two terms it has parsed around the addOp.
- The predicate \[ \text{phrase}(\text{:DCGBody, ?List}) \] can be used to parse a list of tokens against a DCG grammar.
- A crucial task of the evaluator is keeping track of variable assignments. Prolog’s associative lists can be of great help. See \text{put_assoc/4}, \text{get_assoc/3}, and \text{empty_assoc/1}.
- If you use associative lists, the different evaluation steps must keep track of the state transformations and therefore carry an input state list and an output state list which is the input list after the transformations have been applied (recall that Prolog variables cannot be redefined within an execution).
- Make sure that your parser accepts numbers in both float and integer format (i.e., 10 and 10.0).
- The return value of a successful evaluation can either be in integer or float format.
- The parser should not allow the keywords of the language (e.g., the arithmetic operators, :=, \text{var}, \text{return}) as variable identifiers.
- If you want to use any other Prolog dialect, make sure that it runs on SWIProlog. This is the platform on which we will test and grade the solutions.