Homework 3 (10 points)

Recursive descendent parsing

Study the BNF grammar attached to this exercise sheet and do the following tasks:

1. Check if the grammar correctly defines the syntax of the language proposed in the `sample.pas` file and discuss in the OLAT forum eventual corrections and changes;
2. Transform the grammar into a form which suitable for a recursive-descent parsing as follows:
    a. Eliminate the left recursion;
    b. Left factorize the productions;
    c. Convert the grammar into the EBNF form; **(0.5 points)**
3. Declare the set of tokens in an enumerated type and modify the lexical analyser from the Homework 1 to return the current token when a regular expression is matched; **(0.5 points)**
4. Write a recursive-descent parser that uses the lexical analyser implemented in the previous task to get the next lookahead token; **(8 points)**
5. Stop the parse at the first syntactic error with a meaningful error message including the line number. **(0.5 points)**

BNF Grammar

start → **PROGRAM IDENT ;** *varDec compStmt* **.**

varDec → **VAR** *varDecList*
| ε

varDecList → *varDecList identListType* **;**
| *identListType* **;**

identListType → *identList* **:** *type*

identList → *identList* **, IDENT**
| **IDENT**

type → *simpleType*
| **ARRAY [NUM .. NUM] OF** *simpleType*

simpleType → **INTEGER**
| **REAL**
| **BOOLEAN**

compStmt → **BEGIN** *stmtList* **END**

stmtList → *stmtList* **;** *statement*
| *statement*

statement → *assignStmt*
| *compStmt*
| *ifStmt*
| *whileStmt*

assignStmt → **IDENT :=** *expr*
| **IDENT** *index* **:=** *expr*

index → **[** *expr* **]**
| **[** *expr* **..** *expr* **]**

ifStmt → **IF** *expr* **THEN** *statement elsePart*

elsePart → **ELSE** *statement*
| ε

whileStmt → **WHILE** *expr* **DO** *statement*

forStmt → **FOR IDENT** := *expr toPart expr* **DO** *statement*

toPart → **TO** | **DOWNTO**

exprList → *exprList* **,** *expr*
| *expr*

expr → *simpleExpr relOp simpleExpr*
| *simpleExpr*

simpleExpr → *simpleExpr addOp term*
| *term*

term → *term mulOp factor*
| *factor*

factor → **NUM**
| **FALSE**
| **TRUE**
| **IDENT**
| **IDENT** *index*
| **NOT** *factor*
| **–** *factor*
| **(** *exp* **)**

relOp → **<** | **<=** | **>** | **>=** | **=** | **<>**

addOp → **+** | **–** | **OR**

mulOp → ***** | **/** | **DIV** | **MOD** | **AND**