Milestone1 - SER 502 - Team 26

Team Members:

- Hardik Sakhuja (hsakhuja)
- Manisha Deshpande (mdeshp10)
- Sahithya Cherukuri (scheru20)
- Sameer Mungole (smungole)
- Saudamini Khare (skhare10)

Language Name: Saanp (.hiss) - Inspired by Python Programming Language

Language Design:

The language design defines a syntax for a programming language that includes a set of rules for constructing programs. The language design aims to provide a flexible and expressive syntax that allows programmers to write programs that can perform a wide range of tasks.

- Data Types:

- INTEGER
  (A number is a sequence of digits that may be positive or negative.)
- STRING
  (A string is a sequence of characters enclosed in double quotes or single quotes.)
- BOOLEAN
  (True | False)

- Identifiers:

An ID is a variable name that starts with a lowercase letter/underscore and may contain lowercase letters, digits, or underscores. It cannot be a reserved keyword of the language.

- Operators:

Logical operators - [and, or, not]

Comparison operators: [<, >, ==, !=]

Integer arithmetic operators - [+, -, *, /]

- *Assignment Operators:*
    - **=** operator is used to assigning a value to an identifier.

Identifiers: [a-z_]+[a-z0-9_]* - (True|False|if|else|endif|for|endfor|while|endwhile|range|print)
[An ID is a variable name that starts with a lowercase letter/underscore and may contain lowercase letters, digits, or underscores. It cannot be a reserved keyword of the language.]

- *Conditional Constructs:*
    - Assigning value to an identifier from a ternary operation. (?:)
    - If-else: An if-else statement checks a condition and executes a block of code if the condition is true; otherwise, it executes another block of code.

- *Looping Structures:*
    - Traditional for loop: A for loop iterates over a range of numbers and executes a block of code for each number.
    - While loop: A while loop executes a block of code repeatedly if a condition is true.
    - Enhanced Loop: A while loop executes a block of code repeatedly as long as a condition is true.

- *Print:*
    - A print statement outputs the value of a variable to the console.

Language Grammar

```
PROG ::= BLK
BLK ::= DEC | IFE | FOR | WHILE | EFOR | PRINT

DEC ::= ID = EXP. | ID = TER. | DEC, BLK

IFE ::= if LOG: BLK endif | if LOG: BLK else: BLK endif | IFE, BLK
TER ::= LOG? EXP : EXP

FOR ::= for ID = NUM, LOG, INC: BLK endfor | FOR, BLK
INC ::= ID = EXP

WHILE ::= while LOG: BLK endwhile | WHILE, BLK

EFOR ::= for ID in range(NUM, NUM): BLK endfor | EFOR, BLK

PRINT ::= print(ID).
PRINT ::= print(STR).
PRINT ::= print(NUM).
PRINT ::= print(BOOL).
PRINT ::= PRINT, BLK

LOG ::= CMP and CMP | CMP or CMP | not CMP | CMP
CMP ::= EXP == EXP | EXP != EXP | EXP < EXP | EXP > EXP | ID | BOOL

EXP ::= TERM | TERM + EXP | TERM - EXP | STR | BOOL
TERM ::= FACTOR | FACTOR * TERM | FACTOR / TERM
FACTOR ::= ID | NUM | (EXP)

ID ::= [a-z]+[a-z0-9_]* - {True|False|if|else|endif|for|endfor|while|endwhile|range|print}
STR ::= "[^"]*"
NUM ::= [0-9]+ | -[0-9]+
BOOL ::= True | False
```

Updated version changes:

- Introduced LOG (Logical expression)
- Made BOOL a primitive data type.
- Added support for printing BOOL, STR, and NUM.
- Removed STMT rule and implemented right recursion for DEC | IFE | FOR | WHILE | EFOR to accept consecutive statements.

Language Information:

Tokenizer:

- *Python*

Creates a list of tokens in a temporary file called “saanp” which is deleted after the execution of the program.

Parser:

- *Prolog*

Reads the list of tokens from the temporary file and generates a parse tree.

Parsing Technique:

- *Prolog*
- *Data Structures: List*
- *Parsing Technique: Top-down*
- *Grammar: DCG*

Parsing is the process of analyzing a string of symbols according to formal grammar rules. It involves breaking down the input into its constituent parts, determining its structure and relationships, and generating a parse tree or abstract syntax tree that represents the meaning of the input.

Evaluator:

- *Prolog*

Takes the generated parsed tree from the previous step and evaluates the program.