ALGONQUIN COLLEGE

MAMBA Language LANGUAGE SPECIFICATION[1]

General View

This document focusses on MAMBA LS (Language Specification) that is based on combination between PLATYPUS language, originally created by Prof. Svillen Ranev for Algonquin College and Python.

Grammar, which knows how to control even kings . . .
—Molière, *Les Femmes Savantes* (1672), Act II, scene vi

Note 1

Please change this template, replacing any "BOA LANGUAGE" reference by your language name. Remember that this document is using the professor's language and you need to adapt the BNF to your own language. This time, you just need to define the grammar, using **white boxes**. It is not necessary to solve problems such as **LR** (Left Recursion) and **LF** (Left Factoring). You don't need to define the FIRST set (that will be used later in the implementation).

A context-free grammar is used to define the lexical and syntactical parts of the **MAMBA LANGUAGE** language and the lexical and syntactic structure of a **MAMBA LANGUAGE** program.

1. The MAMBA LANGUAGE Lexical Specification

1.1. White Space

White space is defined as the ASCII space, horizontal and vertical tabs, and form feed characters, as well as line terminators. White space is discarded by the scanner.

<white space> → *one of* { SPACE, TAB, FF, NL, CR, NLCR }

1.2. Comments

[1] SOFIA (from Greek, "Wisdom") is also the name of the Bulgarian capital city, homeland form prof. Svillen Ranev, professor from Compilers for several years in the Algonquin College.

MAMBA LANGUAGE supports only single-line comments: all the text from the ASCII characters **!!** to the end of the line is ignored by the scanner.

<comments> → # { sequence of ASCII chars } #
<comments> → # { sequence of ASCII chars } \n

1.3. Variable Identifiers

The following variable identifier (VID) tokens are produced by the scanner: two kinds of arithmetic tokens: **IVID_T** (integer) and one kind of strings: **SVID_T**.

<variable identifier> → VID_T

1.4. Keywords

The scanner produces a single token: **KW_T**. The type of the keyword is defined by the attribute of the token (the index of the keywordTable []). Remember that the list of keywords in MAMBA LANGUAGE is given by:

main, int, def, string, if, then, else, while, do, sapla

1.5. Integer Literals

The scanner produces a single token: **INL_T** with an integer value as an attribute.

<integer_literal> → INTL_T

1.6. String Literals

STR_T token is produced by the scanner.

<string_literal> → STR_T

1.7. Separators

<separator> → *one of* { ( ) , : }

Seven different tokens are produced by the scanner - **LPR_T**, **RPR_T**, COMMA_T, COLON_T

1.8. Operators

```
ArithmeticOperators { OP_ADD, OP_SUB, OP_MUL, OP_DIV }
```

<separator> → *one of* { -, +, *, / }

A single token is produced by the scanner: `ART_OP_T` . The type of the operator is defined by the attribute of the token.

<arithmetic operator> → *one of* { +, -, *, / }

A single token is produced by the scanner: **SCC_OP_T**.

<string concatenation operator> → +

A single token is produced by the scanner: **REL_OP_T**. The type of the operator is defined by the attribute of the token.

<relational operator> → *one of* { >, <, ==, !=, >=, <= }

A single token is produced by the scanner: **LOG_OP_T**. The type of the operator is defined by the attribute of the token.

<logical operator> → *one of* { **and** , **or** , **not** }

A single token is produced by the scanner: **ASS_OP_T**.

<assignment operator> → =

2. The MAMBA LANGUAGE Syntactic Specification

2.1. MAMBA LANGUAGE Program

2.1.1. Program

MAMBA LANGUAGE program is composed by one special function: "def main():" (Method name) defined as follows.

<program> → **def main():**
<data_session>
<code_session>

2.1.2. DATA

The first part (**data**) is the place we declare the variables:

<opt_varlist_declarations> = <opt_varlist_declarations>

Variable Lists

The optional variable list declarations is used to define several datatype declarations:

<opt_varlist_declarations> → <varlist_declarations> | ϵ

Variable Declarations

<varlist_declarations> → <varlist_declaration>
| <varlist_declarations><varlist_declaration>

- **PROBLEM DETECTED: Left recursion – SOLVING FOR YOU:**

New Grammar

<varlist_declarations> → <varlist_declaration> <varlist_declarationsPrime>
<varlist_declarationsPrime> → <varlist_declaration> <varlist_declarationsPrime> | ϵ

Each variable declaration can be done as follows:

<varlist_declaration> → < varlist_declaration>
| < varlist_declaration>
| < varlist_declaration>

2.1.3. Declaration of Lists:

The variables list declaration is defined here:

< varlist_declaration> → < variable_list>;

2.1.4. List of Variables:

The list of variables is defined here:

Integers:

<variable_list> → <variable>
| < variable_list>, < variable>

<variable> → VID_T

Strings:

< variable_list> → <string_VALUE>

		| <string_variable>
<string_variable>	→	SVID_T

2.1.5. CODE session:

The second part (CODE) is the place we have statements:

def main():

<code_session> → <opt_statements>

Optional Statements:

<opt_statements> → <statements> | ϵ

2.1.6. Statements

<statements> → <statement> | <statements> <statement>

2.2. Statement

<statement> → <assignment statement> | <selection statement> | <iteration statement>
| <input statement> | <output statement>

2.2.1. Assignment Statement

<assignment statement> → <assignment expression>

2.2.2. Assignment Expression

<assignment expression>	→	< variable> = <arithmetic expression>
		| <variable> = <arithmetic expression>
		| <variable>= <string expression>

2.2.3. Selection Statement (if statement)

<selection statement> → **if** (<conditional expression>) :	
	Indentation <opt_statements>
Then:	
	{ **Indentation** <opt_statements> }
Else:	
	{ **Indentation** <opt_statements> }

2.2.4. Iteration Statement (the loop statement)

<iteration statement> → **while** (<conditional expression>):

do { **Indentation** <statements>};

2.2.5. Input Statement

<input statement> → **input&** (<variable list>);

Variable List:

<variable list> → <variable identifier> | <variable list>,<variable identifier>

Variable Identifier:

<variable identifier> → <variable>

2.2.6. Output Statement

<output statement> → **print**(<opt_variable list>); | WRITE (VID_T);

Optional Variable List:

<opt_variable list> → <variable list> | ϵ

2.3. Expressions

2.3.1. Arithmetic Expression

<arithmetic expression> → <unary arithmetic expression> | <additive arithmetic expression>

Unary Arithmetic Expression:

<unary arithmetic expression> → - <primary arithmetic expression>
| + <primary arithmetic expression>

Additive Arithmetic Expression:

<additive arithmetic expression> →
<additive arithmetic expression> + <multiplicative arithmetic expression>
| <additive arithmetic expression> - <multiplicative arithmetic expression>
| <multiplicative arithmetic expression>

Multiplicative Arithmetic Expression:

<multiplicative arithmetic expression> →
<multiplicative arithmetic expression> * <primary arithmetic expression>
| <multiplicative arithmetic expression> / <primary arithmetic expression>
| <primary arithmetic expression>

Primary Arithmetic Expression:

<primary arithmetic expression> → <variable>

| INL_T

| (<arithmetic expression>)

2.3.2. String Expression

<string expression> →
<primary string expression> + <string expression> + <primary string expression>

Primary String Expression:

<primary string expression> → <string_variable> | STR_T

2.3.3. Conditional Expression

<conditional expression> → <logical OR expression>

Logical OR Expression:

<logical OR expression> → <logical AND expression>
| <logical OR expression> **or** <logical AND expression>

Logical AND Expression:

<logical AND expression> → <logical NOT expression>
| <logical AND expression> **and** <logical NOT expression>

Logical NOT Expression:

<logical NOT expression> → **not** <relational expression>
| <relational expression>

2.3.4. Relational Expression

<relational expression> →
<relational a_expression> | <relational s_expression>

Relational Arithmetic Expression:

<relational a_expression> →
<primary a_relational expression> == <primary a_relational expression>
| <primary a_relational expression> => <primary a_relational expression>
| <primary a_relational expression> =< <primary a_relational expression>
| <primary a_relational expression> != <primary a_relational expression>
| <primary a_relational expression> > <primary a_relational expression>
| <primary a_relational expression> < <primary a_relational expression>

Relational String Expression:

<relational s_expression> →

<primary s_relational expression> == <primary s_relational expression>
| <primary s_relational expression> <> <primary s_relational expression>
| <primary s_relational expression> > <primary s_relational expression>
| <primary s_relational expression> < <primary s_relational expression>

Primary Arithmetic Relational Expression:

<primary a_relational expression> → <variable> | VID_T

<primary s_relational expression> → <primary string expression>

Good luck with Assignment 3.1!