

INFO-F-403 : Language theory and compiling
Rapport projet partie 3 - Génération de code

Simon Picard
Arnaud Rosette

13 février 2015

Table des matières

1	Introduction	3
2	Transformation de la grammaire donnée	3
2.1	Opérateurs binaires et unaires	3
2.2	Priorité et associativité des opérateurs	3
2.3	Suppression des récursions à gauche	3
2.4	Left factoring	3
2.5	Les variables <code><Instruction></code> et <code><InstructionList></code>	3
2.6	Fonctions	4
2.7	Instructions commençant par un identifieur	4
2.8	Automatisation	4
3	Grammaire	4
4	Ensembles First et Follow	7
5	Action Table	15
6	Génération de code	17
6.1	Introduction	17
6.2	Structures conditionnelles	17
6.3	Boucles	17
6.4	Fonctionnalités du langage implémentées	17

1 Introduction

Dans cette partie du projet, il nous est demandé de modifier la grammaire donnée dans l'énoncé afin de rendre celle-ci LL(1). Il faut donc appliquer toute une série de règles et d'algorithmes pour rendre cette grammaire moins ambiguë et LL(1), dans le but de construire "l'action table" à partir des ensembles first et follow. L'inclusion des fonctions dans la grammaire fait partie d'un bonus que nous avons décidé de réaliser.

2 Transformation de la grammaire donnée

2.1 Opérateurs binaires et unaires

La première étape pour rendre la grammaire LL(1) fut la distinction des deux types d'opérateurs : les opérateurs binaires qui agissent sur deux expressions l'une située à gauche et l'autre à droite de ces opérateurs et les opérateurs unaires qui agissent sur une seule expression située à droite de ces opérateurs. Il faut différencier les deux types d'opérateurs afin de ne pas avoir des expressions du style : $\gg 4, *2, 6|, \dots$. Les opérateurs unaires sont au nombre de quatre : ! (negation), \sim (not bit à bit), + et -. Les opérateurs + et - sont également des opérateurs binaires.

2.2 Priorité et associativité des opérateurs

Lors de cette étape, nous avons modifié la grammaire afin de fixer la priorité et l'associativité des différents opérateurs, ceci afin de rendre la grammaire moins ambiguë. Nous avons remarqué que les opérateurs unaires étaient plus prioritaires que les opérateurs binaires et que les opérateurs unaires étaient tous associatifs à droite, tandis que les opérateurs binaires l'étaient tous à gauche. Pour fixer la priorité, nous avons mis les opérateurs les moins prioritaires le plus "haut" possible dans la grammaire et les plus prioritaires le plus "bas" possible. Dans la grammaire ci-dessous où le "start symbol" est E, l'opérateur "+" est considéré comme plus "haut" que l'opérateur "*" dans la grammaire car il est possible d'obtenir l'opérateur "+" en utilisant moins de règles de production à partir du "start symbol" que pour obtenir l'opérateur "*".

```
E->E + T
  ->T
T->T * F
  ->F
F->ID
  ->( E )
```

2.3 Suppression des récursions à gauche

L'étape suivante fut l'étape de suppression des récursions à gauche qui sont incompatibles avec les "top-down parser" car cela fait entrer ce genre de parser dans une boucle ou récursion infinie. Cette étape rend tout opérateur associatif à droite si l'on ne fait rien pour résoudre ce problème lors de l'analyse sémantique.

2.4 Left factoring

Cette étape rassemble les règles de production d'une variable qui ont un préfixe commun en une seule règle de production qui contient ce préfixe commun et une nouvelle variable. Cette nouvelle variable possède des règles de production vers les différents suffixes qui étaient présents à l'origine dans les différentes règles de production qui ont été mises en commun. Ceci est nécessaire car le parser qu'on va créer est LL(1), il faut donc qu'il puisse choisir la bonne règle de production en regardant seulement le prochain token qui est en entrée.

2.5 Les variables <Instruction> et <InstructionList>

Dans la grammaire donnée, à chaque fois que la variable <Instruction> se trouvait dans la partie droite d'une règle de production excepté lorsque la partie gauche de la règle était <InstructionList>,

il y avait une autre règle de production pour cette même partie gauche où `<Instruction>` était remplacé par `<InstructionList>`. Par exemple, `<If>→<Expression> <Empty> <InstructionList> <IfEnd>` et `<If>→<Expression> <Empty> <Instruction> <IfEnd>`. Ceci permettant de ne pas mettre de `END_OF_INSTRUCTION` à la fin d'une `<Instruction>` lorsque le corps d'un block (ici, le corps du "if") ne contenait qu'une seule `<Instruction>`. Le problème posé par le doublement systématique des règles de production contenant des instructions (une règle pour `<Instruction>` et une autre pour `<InstructionList>`) est que ce n'est pas factorisé à gauche et que le first de `<InstructionList>` peut être `<Instruction>`. Afin de résoudre ce problème, nous avons décidé de ne plus utiliser que la variable `<InstructionList>` dans les autres règles de production. Ainsi, on évite le doublement des règles de production. Une `<InstructionList>` est une liste d'`<Instruction>` séparées par des `END_OF_INSTRUCTION`, avec la dernière `<Instruction>` qui n'est pas nécessairement suivie d'un `END_OF_INSTRUCTION`. Ceci permet d'avoir un programme du style : `if(a>b) ;a=10 end ;` Donc seule une `<InstructionList>` permet de produire des `<Instruction>`. De plus, vu qu'une `<InstructionList>` peut contenir des instructions vides (une instruction contenant seulement `END_OF_INSTRUCTION`) la variable `<Empty>` n'est plus nécessaire car celle-ci servait uniquement à indiquer qu'il fallait au moins un terminal `END_OF_INSTRUCTION`.

2.6 Fonctions

Les fonctions amènent deux types d'instructions en plus : les définitions de fonction et les appels de fonction. Les appels sont considérés comme des expressions atomiques afin de pouvoir mettre des appels de fonction au sein d'une expression. Par exemple, `a=foo()+5 ;` De plus, il faut qu'une fonction puisse être appelée sans avoir besoin de faire une assignation ou un autre type d'instruction car une fonction peut ne rien retourner et simplement avoir un effet de bord. Par exemple, la fonction `println` est une fonction qui ne retourne rien mais produit un effet de bord qui est l'affichage dans le terminal. Il faut donc pouvoir faire, par exemple, `println(a)`. C'est pour cela qu'il faut qu'un appel de fonction soit une instruction à part entière.

2.7 Instructions commençant par un identifiant

Plusieurs instructions commencent par un identifiant. Il s'agit des assignations, des déclarations de variables et des appels de fonctions. Afin que la grammaire soit LL(1) il faut factoriser ces instructions. C'est pour cela que la variable `<IdentifierInstruction>` a été créée. Celle-ci représente une instruction qui commence par un identifiant. La variable `<IdentifierInstructionTail>` a été également créée afin de distinguer les différents types d'instructions qui commencent par un identifiant. La distinction est après aisément faite car lorsque le symbole suivant l'identifiant est "=", on sait que c'est une assignation, lorsque ce symbole est ":", on sait que c'est une déclaration de variable et lorsque que ce symbole est "(", on sait que c'est un appel de fonction.

2.8 Automatisation

Nous avons développé une série d'algorithmes qui permettent de supprimer la récursion à gauche, de factoriser à gauche, de supprimer les symboles inutiles et de générer une table d'action ce qui nous permet de vérifier si une grammaire est bien LL(1) via l'action table ou de nous aider à la transformer en grammaire LL(1) grâce aux autres algorithmes.

3 Grammaire

Voici la grammaire donnée dans l'énoncé et transformée en grammaire LL(1). A chaque règle de production correspond un numéro qui identifie cette règle.

[1]	<code><Program></code>	→	<code><InstructionList></code>
[2]	<code><Instruction></code>	→	<code><IdentifierInstruction></code>
[3]		→	<code><ConstDefinition></code>
[4]		→	<code><Block></code>
[5]		→	<code><Loop></code>
[6]		→	<code><BuiltInFunctionCall></code>
[7]		→	<code><FunctionDefinition></code>
[8]	<code><InstructionList></code>	→	<code><Instruction> <InstructionListTail></code>

[9]		→	<InstructionListTail>
[10]	<InstructionListTail>	→	END_OF_INSTRUCTION <InstructionList>
[11]		→	ε
[12]	<IdentifierInstruction>	→	IDENTIFIER <IdentifierInstructionTail>
[13]	<IdentifierInstructionTail>	→	<AssignmentTail>
[14]		→	TYPE_DEFINITION <Type>
[15]		→	<FunctionCallTail>
[16]	<AssignmentTail>	→	ASSIGNATION <Expression>
[17]		→	COMMA IDENTIFIER <AssignmentTail> COMMA <Expression>
[18]	<ConstDefinition>	→	CONST IDENTIFIER <AssignmentTail>
[19]	<Block>	→	LET IDENTIFIER <AssignmentTail> END_OF_INSTRUCTION <InstructionList> END
[20]	<Loop>	→	<If>
[21]		→	WHILE <Expression> END_OF_INSTRUCTION <InstructionList> END
[22]		→	FOR IDENTIFIER ASSIGNATION <Expression> TERNARY_ELSE <Expression> <ForTail>
[23]	<ForTail>	→	END_OF_INSTRUCTION <InstructionList> END
[24]		→	TERNARY_ELSE <Expression> END_OF_INSTRUCTION <InstructionList> END
[25]	<Type>	→	BOOLEAN_TYPE
[26]		→	REAL_TYPE
[27]		→	INTEGER_TYPE
[28]	<Expression>	→	<BinaryExpression> <TernaryIfExpression>
[29]	<TernaryIfExpression>	→	TERNARY_IF <Expression> <TernaryElseExpression>
[30]		→	ε
[31]	<TernaryElseExpression>	→	TERNARY_ELSE <Expression>
[32]	<AtomicExpression>	→	<AtomicIdentifierExpression>
[33]		→	INTEGER
[34]		→	REAL
[35]		→	BOOLEAN
[36]		→	<BuiltInFunctionCall>
[37]	<AtomicIdentifierExpression>	→	IDENTIFIER <AtomicIdentifierExpressionTail>
[38]	<AtomicIdentifierExpressionTail>	→	<FunctionCallTail>
[39]		→	ε
[40]	<UnaryExpression>	→	NEGATION <UnaryExpression>
[41]		→	<UnaryBitwiseNotExpression>
[42]	<UnaryBitwiseNotExpression>	→	BITWISE_NOT <UnaryBitwiseNotExpression>
[43]		→	<UnaryMinusPlusExpression>
[44]	<UnaryMinusPlusExpression>	→	MINUS <UnaryMinusPlusExpression>
[45]		→	PLUS <UnaryMinusPlusExpression>
[46]		→	<UnaryAtomicExpression>
[47]	<UnaryAtomicExpression>	→	<AtomicExpression>
[48]		→	LEFT_PARENTHESIS <Expression> RIGHT_PARENTHESIS
[49]	<BinaryExpression>	→	<BinaryLazyOrExpression> <BinaryExpression'>
[50]	<BinaryExpression'>	→	LAZY_OR <BinaryLazyOrExpression> <BinaryExpression'>
[51]		→	ε
[52]	<BinaryLazyOrExpression>	→	<BinaryLazyAndExpression> <BinaryLazyOrExpression'>
[53]	<BinaryLazyOrExpression'>	→	LAZY_AND <BinaryLazyAndExpression> <BinaryLazyOrExpression'>
[54]		→	ε

[55]	<BinaryLazyAndExpression>	→	<BinaryNumericExpression> <BinaryLazyAndExpression'>
[56]	<BinaryLazyAndExpression'>	→	GREATER_THAN <BinaryNumericExpression> <BinaryLazyAndExpression'>
[57]		→	LESS_THAN <BinaryNumericExpression> <BinaryLazyAndExpression'>
[58]		→	GREATER_OR_EQUALS_THAN <BinaryNumericExpression> <BinaryLazyAndExpression'>
[59]		→	LESS_OR_EQUALS_THAN <BinaryNumericExpression> <BinaryLazyAndExpression'>
[60]		→	EQUALITY <BinaryNumericExpression> <BinaryLazyAndExpression'>
[61]		→	INEQUALITY <BinaryNumericExpression> <BinaryLazyAndExpression'>
[62]		→	€
[63]	<BinaryNumericExpression>	→	<BinaryTermExpression> <BinaryNumericExpression'>
[64]	<BinaryNumericExpression'>	→	PLUS <BinaryTermExpression> <BinaryNumericExpression'>
[65]		→	MINUS <BinaryTermExpression> <BinaryNumericExpression'>
[66]		→	BITWISE_OR <BinaryTermExpression> <BinaryNumericExpression'>
[67]		→	BITWISE_XOR <BinaryTermExpression> <BinaryNumericExpression'>
[68]		→	€
[69]	<BinaryTermExpression>	→	<BinaryShiftedExpression> <BinaryTermExpression'>
[70]	<BinaryTermExpression'>	→	ARITHMETIC_SHIFT_LEFT <BinaryShiftedExpression> <BinaryTermExpression'>
[71]		→	ARITHMETIC_SHIFT_RIGHT <BinaryShiftedExpression> <BinaryTermExpression'>
[72]		→	€
[73]	<BinaryShiftedExpression>	→	<BinaryFactorExpression> <BinaryShiftedExpression'>
[74]	<BinaryShiftedExpression'>	→	TIMES <BinaryFactorExpression> <BinaryShiftedExpression'>
[75]		→	DIVIDE <BinaryFactorExpression> <BinaryShiftedExpression'>
[76]		→	REMAINDER <BinaryFactorExpression> <BinaryShiftedExpression'>
[77]		→	BITWISE_AND <BinaryFactorExpression> <BinaryShiftedExpression'>
[78]		→	INVERSE_DIVIDE <BinaryFactorExpression> <BinaryShiftedExpression'>
[79]		→	€
[80]	<BinaryFactorExpression>	→	<UnaryExpression> <BinaryFactorExpression'>
[81]	<BinaryFactorExpression'>	→	POWER <UnaryExpression> <BinaryFactorExpression'>
[82]		→	€
[83]	<If>	→	IF <Expression> END_OF_INSTRUCTION <InstructionList> <IfEnd>

[84]	<IfEnd>	→ ELSE_IF <Expression> END_OF_INSTRUCTION <InstructionList> <IfEnd>
[85]		→ ELSE <InstructionList> END
[86]		→ END
[87]	<BuiltInFunctionCall>	→ READ_REAL LEFT_PARENTHESIS RIGHT_PARENTHESIS
[88]		→ READ_INTEGER LEFT_PARENTHESIS RIGHT_PARENTHESIS
[89]		→ INTEGER_CAST LEFT_PARENTHESIS <Expression> RIGHT_PARENTHESIS
[90]		→ REAL_CAST LEFT_PARENTHESIS <Expression> RIGHT_PARENTHESIS
[91]		→ BOOLEAN_CAST LEFT_PARENTHESIS <Expression> RIGHT_PARENTHESIS
[92]		→ PRINTLN LEFT_PARENTHESIS <Expression> RIGHT_PARENTHESIS
[93]	<FunctionCallTail>	→ LEFT_PARENTHESIS <Parameter> RIGHT_PARENTHESIS
[94]	<Parameter>	→ <Expression> <ParameterTail>
[95]		→ ϵ
[96]	<ParameterTail>	→ COMMA <Expression> <ParameterTail>
[97]		→ ϵ
[98]	<FunctionDefinition>	→ FUNCTION IDENTIFIER LEFT_PARENTHESIS <Argument> RIGHT_PARENTHESIS <InstructionList> <FunctionDefinitionEnd>
[99]	<FunctionDefinitionEnd>	→ RETURN <Expression> END
[100]		→ END
[101]	<Argument>	→ IDENTIFIER TYPE_DEFINITION <Type> <ArgumentTail>
[102]		→ ϵ
[103]	<ArgumentTail>	→ COMMA IDENTIFIER TYPE_DEFINITION <Type> <ArgumentTail>
[104]		→ ϵ

4 Ensembles First et Follow

Dans le tableau suivant, EPSILON_VALUE $\equiv \epsilon$.

Variable	First	Follow
<Program>	FUNCTION, WHILE, READ_REAL EPSILON_VALUE IDENTIFIER, CONST BOOLEAN_CAST, PRINTLN END_OF_INSTRUCTION READ_INTEGER, FOR INTEGER_CAST, LET, IF REAL_CAST	
<Instruction>	FUNCTION, READ_INTEGER FOR, WHILE, READ_REAL INTEGER_CAST BOOLEAN_CAST, CONST IDENTIFIER, PRINTLN, LET, IF REAL_CAST	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN

<InstructionList>	FUNCTION, WHILE, READ_REAL EPSILON_VALUE BOOLEAN_CAST, IDENTIFIER CONST, PRINTLN END_OF_INSTRUCTION READ_INTEGER, FOR INTEGER_CAST, LET, IF REAL_CAST	END, ELSE_IF, ELSE, RETURN
<InstructionListTail>	EPSILON_VALUE END_OF_INSTRUCTION	END, ELSE_IF, ELSE, RETURN
<IdentifierInstruction>	IDENTIFIER	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<IdentifierInstructionTail>	ASSIGNATION TYPE_DEFINITION, COMMA LEFT_PARENTHESIS	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<AssignmentTail>	ASSIGNATION, COMMA	END, COMMA END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<ConstDefinition>	CONST	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<Block>	LET	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<Loop>	FOR, WHILE, IF	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<ForTail>	END_OF_INSTRUCTION TERNARY_ELSE	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<Type>	REAL_TYPE, BOOLEAN_TYPE INTEGER_TYPE	RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<Expression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST	RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<TernaryIfExpression>	TERNARY_IF EPSILON_VALUE	RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<TernaryElseExpression>	TERNARY_ELSE	RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN

<AtomicExpression>	READ_INTEGER, INTEGER REAL, BOOLEAN, READ_REAL INTEGER_CAST BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN
<AtomicIdentifierExpression>	IDENTIFIER	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN
<AtomicIdentifierExpressionTail>	EPSILON_VALUE LEFT_PARENTHESIS	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN

<UnaryExpression>	<p> INTEGER, REAL, BOOLEAN NEGATION, BITWISE_NOT READ_REAL LEFT_PARENTHESIS IDENTIFIER, BOOLEAN_CAST PRINTLN, READ_INTEGER PLUS, MINUS, INTEGER_CAST REAL_CAST </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN </p>
<UnaryBitwiseNotExpression>	<p> READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN </p>
<UnaryMinusPlusExpression>	<p> READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN, MINUS READ_REAL, INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN </p>

<UnaryAtomicExpression>	READ_INTEGER, INTEGER REAL, BOOLEAN, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN
<BinaryExpression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST BOOLEAN_CAST LEFT_PARENTHESIS IDENTIFIER, PRINTLN REAL_CAST	TERNARY_IF RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<BinaryExpression'>	LAZY_OR, EPSILON_VALUE	TERNARY_IF RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<BinaryLazyOrExpression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST	TERNARY_IF, LAZY_OR RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<BinaryLazyOrExpression'>	LAZY_AND, EPSILON_VALUE	TERNARY_IF, LAZY_OR RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<BinaryLazyAndExpression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST BOOLEAN_CAST LEFT_PARENTHESIS IDENTIFIER, PRINTLN REAL_CAST	LAZY_AND, TERNARY_IF LAZY_OR RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN

<BinaryLazyAndExpression'>	EQUALITY, INEQUALITY GREATER_THAN EPSILON_VALUE LESS_OR_EQUALS_THAN GREATER_OR_EQUALS_THAN LESS_THAN	LAZY_AND, TERNARY_IF LAZY_OR RIGHT_PARENTHESIS, END COMMA END_OF_INSTRUCTION ELSE_IF, ELSE TERNARY_ELSE, RETURN
<BinaryNumericExpression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST	EQUALITY, TERNARY_IF GREATER_THAN RIGHT_PARENTHESIS, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, LAZY_OR, END ELSE_IF GREATER_OR_EQUALS_THAN
<BinaryNumericExpression'>	BITWISE_OR, PLUS, MINUS BITWISE_XOR EPSILON_VALUE	EQUALITY, TERNARY_IF GREATER_THAN RIGHT_PARENTHESIS, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, LAZY_OR, END ELSE_IF GREATER_OR_EQUALS_THAN
<BinaryTermExpression>	READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST BOOLEAN_CAST LEFT_PARENTHESIS IDENTIFIER, PRINTLN REAL_CAST	EQUALITY, TERNARY_IF RIGHT_PARENTHESIS GREATER_THAN BITWISE_XOR, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, BITWISE_OR PLUS, LAZY_OR, MINUS, END ELSE_IF GREATER_OR_EQUALS_THAN
<BinaryTermExpression'>	ARITHMETIC_SHIFT_RIGHT ARITHMETIC_SHIFT_LEFT EPSILON_VALUE	EQUALITY, TERNARY_IF RIGHT_PARENTHESIS GREATER_THAN BITWISE_XOR, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, BITWISE_OR PLUS, LAZY_OR, MINUS, END ELSE_IF GREATER_OR_EQUALS_THAN

<BinaryShiftedExpression>	<p> READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST LEFT_PARENTHESIS BOOLEAN_CAST, IDENTIFIER PRINTLN, REAL_CAST </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, TERNARY_IF ARITHMETIC_SHIFT_LEFT RIGHT_PARENTHESIS GREATER_THAN BITWISE_XOR, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, PLUS BITWISE_OR, LAZY_OR, MINUS END, ELSE_IF GREATER_OR_EQUALS_THAN </p>
<BinaryShiftedExpression'>	<p> BITWISE_AND INVERSE_DIVIDE REMAINDER, TIMES EPSILON_VALUE, DIVIDE </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, TERNARY_IF ARITHMETIC_SHIFT_LEFT RIGHT_PARENTHESIS GREATER_THAN BITWISE_XOR, COMMA END_OF_INSTRUCTION LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE, RETURN LESS_THAN, LAZY_AND INEQUALITY, PLUS BITWISE_OR, LAZY_OR, MINUS END, ELSE_IF GREATER_OR_EQUALS_THAN </p>
<BinaryFactorExpression>	<p> READ_INTEGER, INTEGER REAL, PLUS, BOOLEAN NEGATION, MINUS BITWISE_NOT, READ_REAL INTEGER_CAST BOOLEAN_CAST LEFT_PARENTHESIS IDENTIFIER, PRINTLN REAL_CAST </p>	<p> ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN </p>

<BinaryFactorExpression'>	POWER, EPSILON_VALUE	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND ARITHMETIC_SHIFT_LEFT TERNARY_IF, GREATER_THAN LESS_OR_EQUALS_THAN ELSE, TERNARY_ELSE INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS BITWISE_XOR, REMAINDER COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, END, TIMES ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN
<If>	IF	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<IfEnd>	END, ELSE_IF, ELSE	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<BuiltInFunctionCall>	READ_INTEGER, READ_REAL INTEGER_CAST BOOLEAN_CAST, PRINTLN REAL_CAST	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND TERNARY_IF ARITHMETIC_SHIFT_LEFT GREATER_THAN LESS_OR_EQUALS_THAN TERNARY_ELSE, ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS REMAINDER, BITWISE_XOR COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, TIMES, END ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN
<FunctionCallTail>	LEFT_PARENTHESIS	ARITHMETIC_SHIFT_RIGHT EQUALITY, BITWISE_AND TERNARY_IF ARITHMETIC_SHIFT_LEFT GREATER_THAN LESS_OR_EQUALS_THAN TERNARY_ELSE, ELSE, POWER INEQUALITY, BITWISE_OR MINUS, RIGHT_PARENTHESIS REMAINDER, BITWISE_XOR COMMA END_OF_INSTRUCTION RETURN, LESS_THAN LAZY_AND, PLUS, LAZY_OR INVERSE_DIVIDE, TIMES, END ELSE_IF, DIVIDE GREATER_OR_EQUALS_THAN

<Parameter>	INTEGER, REAL, BOOLEAN NEGATION, BITWISE_NOT READ_REAL, EPSILON_VALUE BOOLEAN_CAST, IDENTIFIER LEFT_PARENTHESIS PRINTLN, READ_INTEGER PLUS, MINUS, INTEGER_CAST REAL_CAST	RIGHT_PARENTHESIS
<ParameterTail>	COMMA, EPSILON_VALUE	RIGHT_PARENTHESIS
<FunctionDefinition>	FUNCTION	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<FunctionDefinitionEnd>	END, RETURN	END, END_OF_INSTRUCTION ELSE_IF, ELSE, RETURN
<Argument>	EPSILON_VALUE IDENTIFIER	RIGHT_PARENTHESIS
<ArgumentTail>	COMMA, EPSILON_VALUE	RIGHT_PARENTHESIS

5 Action Table

Les lignes de “l’action table” représentent les variables et les colonnes représentent les terminaux. Une cellule de cette table contient le numéro d’une règle de production correspondant au numéro repris dans la section grammaire.

Dans le tableau suivant, $\text{EPSILON_VALUE} \equiv \epsilon$.

6 Génération de code

6.1 Introduction

Dans la génération de code, la prise en charge du type réel et de la création de fonction n'a pas été faite car non requise.

Pour la génération de code nous sommes passés par un parser récursif ordinaire, ce dernier appelle des méthodes de la classe *LlvmCodeGenerate* qui génèrent le code correspondant.

LlvmCodeGenerate possède un stack global qui permet d'y stocker des expressions, donc des variables, des valeurs entières ou booléennes.

Ce stack est indispensable car toute instruction combinée doit être divisée en une série d'actions binaires, par exemple :

$a = 1 + 1$

Il faut commencer par évaluer $1 + 1$, stocker le résultat dans une variable temporaire et puis assigner cette variable temporaire à a .

Le stack permet ici de transférer la variable temporaire.

6.2 Structures conditionnelles

Le *if else* est déjà implémenté dans LLVM grâce à une fonction qui permet de sauter à une ancre ou une autre en fonction de la valeur d'une expression booléenne.

Exemple :

```
def i32 compareTo ( i32 %a, i32 %b){
    entry :
        % cond = icmp slt i32 %a ,%b
        br i1 %cond , label %lower , label % greaterOrEquals
    lower :
        ret i32 -1
    greaterOrEquals :
        %1 = sub i32 %a ,%b
        ret i32 %1
}
```

Le *else if* est une cascade de *if else* où le *else* est une ancre vers la prochaine *if*.

Le *if* ternaire n'est qu'une autre façon d'écrire un *if*.

6.3 Boucles

La boucle *while* fonctionne en utilisant trois ancres, une pour le test de la condition, une pour le cœur de la boucle et une pour sortir de la boucle.

On commence par tester la condition, si elle est fausse, on saute à la fin de la boucle, si elle est vraie, on saute au cœur de la boucle, une fois celui-ci terminé, on saute au test de la condition.

La boucle *for* est une autre façon d'écrire une boucle *while* en définissant une variable avec une valeur de base qui permettra de définir la condition de la boucle, celle-ci est que cette valeur de base doit être inférieure à un maximum défini par l'utilisateur.

Enfin, après avoir exécuté le cœur de la boucle il faut incrémenter la valeur de base par un nombre défini lui aussi par l'utilisateur et puis sauter au test de la condition de la boucle.

6.4 Fonctionnalités du langage implémentées

Comme mentionné ci-dessus, la grammaire contient tous les éléments présents dans le langage, en ce compris les fonctions. De ce fait, le parser est capable de reconnaître la syntaxe de tous les éléments du langage Iulius.

Pour ce qui est de l'analyse sémantique et la génération de code, seuls deux types de données sont manipulables : les booléens et les entiers. Les fonctions ne sont également pas disponibles pour la génération de code. Le reste des fonctionnalités est quand à lui disponible lors de la génération de code. Ces fonctionnalités sont, par exemple : la déclaration de variables, de constantes (vérification qu'il n'y a

pas d'assignation après l'assignation d'origine), l'assignation (et l'assignation multiple), la conversion, les opérations arithmétiques de base (avec vérification des types de chaque opérande afin de vérifier qu'ils sont compatibles), les opérations bits à bits, les boucles (while et for), les branchements conditionnels (if, elseif, else), l'affichage, la récupération de caractères de l'input, la création de bloc. La gestion du scope des variables est également réalisé.