



UNIVERSITÉ LIBRE DE BRUXELLES

INFO-F-403  
INTRODUCTION TO LANGUAGE THEORY AND COMPILING

---

## Assignement 1 : The scanner

---

Maxime DESCLEFS - 362626

Julian SCHEMBRI - 380446

2015 - 2016

# 1 Introduction

Dans le cadre du cours d'introduction à la théorie du langage et de compilation (INFO-F-403), il a été demandé d'implémenter et d'écrire un compilateur pour le langage SUPRALGOL<sup>2016</sup>, une variante du langage Algol68.

La première partie consiste en la création du scanner. Ce scanner analyse et découpe le code en différents tokens différenciant chacun des éléments du langage.

# 2 Tokens

Voici le tableau qui reprend l'ensemble des tokens du langage Supralgol. Chaque ligne de celui-ci est composée du nom du token, le type dans l'enum LexicalUnit qui lui est associé et de l'expression régulière qui correspond à ce token.

Nom	Type	Expression régulière
Comment	-	co.*co
LeftParenthesis (()	LEFT_PARENTHESIS	(
RightParenthesis ())	RIGHT_PARENTHESIS	)
Beg	BEGIN	begin
End	END	end
Semicolon (;)	SEMICOLON	;
Assign (:=)	ASSING	:=
Minus (-)	MINUS	-
Plus (+)	PLUS	+
Times (*)	TIMES	*
Divide (/)	DIVIDE	/
If	IF	if
Then	THEN	then
Fi	FI	fi
Else	ELSE	else
Not	NOT	not
Or	OR	or
And	AND	and
Equal (=)	EQUAL	=
GreaterEqual (≥)	GREATER_EQUAL	>=
Greater (>)	GREATER	>
SmallerEqual (≤)	SMALLER_EQUAL	<=
Smaller (<)	SMALLER	<
Different (≠)	DIFFERENT	!=
While	WHILE	while
Do	DO	do
Od	OD	od
For	FOR	for
From	FROM	from
By	BY	by
To	TO	to
Print	PRINT	print
Read	READ	read
Number	NUMBER	{[0-9]}*
VarName	VARNAME	{{[A-Z]} {[a-z]}}>{{[A-Z]} {[a-z]} {[0-9]}}*

## 3 Implémentation

### 3.1 Mauvaise variable

Pour éviter les cas où des noms de variables commencent par un numéro, un type *BadVarName* a été implémenté pour retourner une erreur lorsque le cas se produit.

### 3.2 Compilation et exécution

Pour lancer le projet, il faut d'abord lancer l'exécutable */more/compile.sh* qui se chargera de créer un *.jar* rendant l'exécution compatible avec la machine.