

Histoire (Construction des compilateurs)

Autrefois " un mystère ", aujourd'hui l'une des branches les plus maîtrisées
En informatique.

1957	Fortran	Premiers compilateurs (Expressions, instructions, procédures)
1960.	Algol	Première définition formelle d'un langage (grammaires sous forme de Backus-Naur, bloc, récursivité, ...)
1970	Pascal	types, machines virtuelles (P-code)
1985	C++	Orientation objets, exceptions
1995	Java	Portabilité

Le cours concerne uniquement les langages impératifs (procéduraux)
langages fonctionnels (Lisp) et langages logiques (Prolog) exigent d'autres techniques.

Structure d'un compilateur

Programme source v a l = 1 0 * v a l + i



Analyse lexicale



Unités lexicales

1	3	2	4	1	5	1
(ident)	(assign)	(number)	(times)	(ident)	(plus)	(ident)
"val"	-	10	-	"val"	-	"i"

Code de l'unité

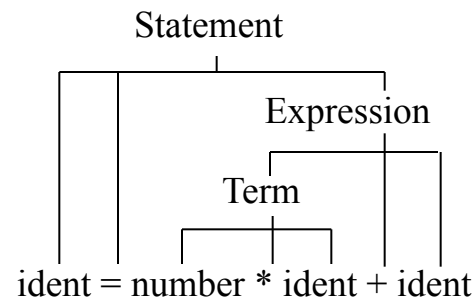
Valeur de l'unité



Analyse syntaxique

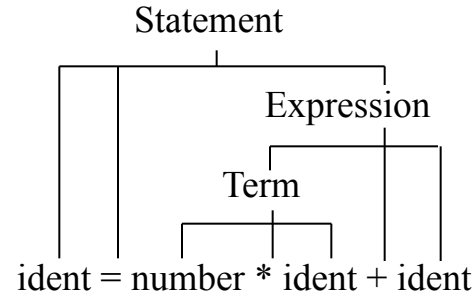


Arbre syntaxique



Structure d'un compilateur

Arbre syntaxique



Analyse sémantique



*Représentation
intermédiaire*

Arbre syntaxique, table des symboles, ...



Optimisation



Génération de code

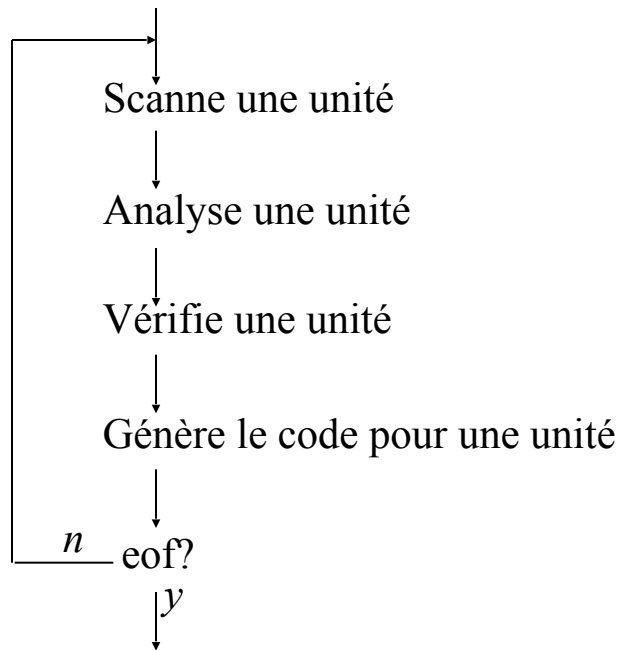


Code machine

00101110000
01101101101
00011111010
...

Compilateurs à une passe

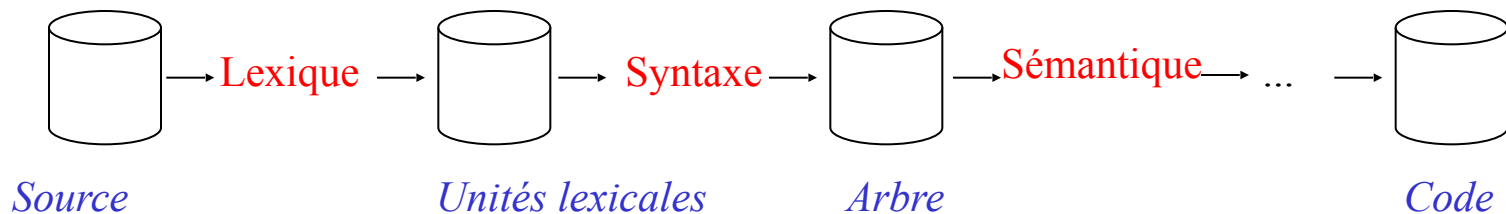
Fonctionnement simultané des phases



Le programme objet est généré en même temps que le programme source est lu.

Compilateurs à plusieurs passes

Les phases sont des programmes séparés qui s'exécutent séquentiellement

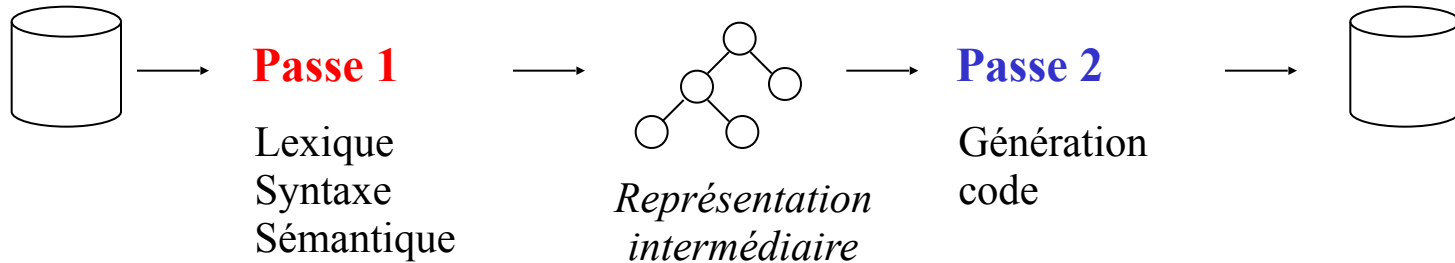


Chaque phase lit à partir d'un fichier et écrit sur un nouveau fichier

Pourquoi plusieurs passes?

- Langage complexe
- Portabilité importante

En général: Compilateurs à deux passes



Dépendant du langage

Java

C

Pascal

Dépendant de la machine

Pentium

PowerPC

SPARC

Toute combinaison possible

Avantages

- Meilleure portabilité
- Combiner les techniques entre les deux passes
- Optimisations plus simples sur la représentation intermédiaire que sur le code source

Inconvénients

- Lenteur
- Plus de mémoire