

# مبادئ الأتومات والمترجمات



م. رود الأصفر

# Introduction

- Compilers are computer programs that translate one language to another.

■ المترجمات عبارة عن برامج حاسب والتي يتم فيها تحويل أحد اللغات الي لغة أخرى

- A compiler takes as its input a program written in its source language and produces an equivalent program written in its target language.

■ المترجمات : مدخلها برنامج مكتوب بلغة المنبع ويتم انتاج برنامج مناظر بلغة الهدف

# Introduction

- Usually , the source language is a high-level language , such as C or C++ ,and the target language is object code (sometimes also called machine code)

■ عادة فإن لغة المنبع هي من لغات الحاسب بينما لغة الهدف هي لغة الآلة

- We can view this process schematically as follows:-




# Introduction

- A compiler is a fairly complex program that can be anywhere from 10,000 to 1,000,000 lines of code.

■ وحقيقة فإن المترجم عبارة عن برنامج معقد به من ١٠٠٠٠ الي ١٠٠٠٠٠٠ سطر

- Writing such a program ,or even understanding it, is not a simple task . وكتابة هذا البرنامج أو فهمه ليس بالأمر الهين

- 
- Anyone professionally involved with computers should know the basic organization and operation of a compiler.

■ إن أي متخصص في الحاسبات يجب أن يكون ملماً بأساسيات التراكيب والعمل لبرامج المترجم

# Introduction

- It is the purpose of this course not only to provide knowledge about compiler and its techniques , but also to give the students all the necessary tools and practical experience to design and program an actual compiler
- إن الهدف من هذه المادة ليس فقط معرفة المترجم وتقنياته ولكن أيضا إعطاء الدارس كل أدوات والخبرات العملية لكي يقوم بتصميم أجزاء من مترجم حقيقي



# البرامج التي لها علاقة بالترجم

## ■ 1-Interpreters

- وهو لغة تحويل مثل المترجم
- ويختلف عن المترجم في أنه يقوم بتنفيذ برنامج المنبع source program مباشرة دون توليد object code
- كما في المترجم والذي يتم تنفيذه بعد استكمال كل التحويل
- ولكن في interpreter فإنه ينفذ سطر سطر مثل لغة BASIC و QUICK BASIC فإن أي سطر يتم تصحيحه أولاً بأول
- ويستخدم interpreter في برامج تطوير التعليم وكذلك برامج حجز الطيران
- ولكن المترجم يتم استخدامه عندما تكون السرعة مطلوبة لتنفيذ البرنامج ككل حيث أن السرعة = 10 - 15 أمثال interpreter

# البرامج التي لها علاقة بالترجم

## -٢- Assembler

يحول إلى لغة التجميع وإلى حاسب معين وكما هو معلوم فإن لغة التجميع عبارة عن هيئه رمزيه symbolic form إلى لغة الآله وهي سهله التحويل وفي بعض الأحيان فإن المترجم يقوم بتوليد لغة التجميع كلغة الهدف Target Language ثم تعرض على Assembler لتحويلها إلى لغة الآله .



# البرامج التي لها علاقة بالترجم

## ٣- Linkers :- الرابط

كلا من المترجم والمجمع ممكن تجميعهم في برنامج يسمى Linker لإنتاج Executable File كما يقوم الـ Linker بتوصيل الـ Object Program بالكود الخاص بـ Stander Library Function وكذلك الموارد الـ Source الموجودة في نظام التشغيل مثل I\O devices وهكذا فإن الـ Linker يقوم بعملية أساسية في المترجم أو كجزء من المترجم.

# البرامج التي لها علاقة بالمترجم

## ٤- Preprocessor :-

عبارة عن برنامج منفصل ويقوم بتنفيذه المترجم قبل أن يبدأ التحويل الحقيقي حيث يقوم المعالج بحذف

التعليقات من البرنامج COMMENTS

يحتوي على ملفات أخرى ويكون الـ macro ) ويعرف الـ macro على انه مختصر لتسلسل يتكرر في

( النص )

# البرامج التي لها علاقة بالمترجم

## ■ 5-Editor

يقوم المترجم عادة بقبول برنامج المنبع source program

والمكتوب بإستعمال أي editor والذي يقوم بإنتاج ملف قياسي

ss.cpp وحديثا فإن المترجم يتم تجميعه كحزمة مع editor والبرامج الأخرى في

## ■ IDE ( an interactive development environment )

ممكن أن يقوم بعمل تنظيم للبرنامج و و و الوان .....

■ cin cout {

■ {

■ {

# البرامج التي لها علاقة بالمترجم

■ ٦- Debugger (المصحح) :-

■ برنامج يتم استخدامه لتحديد أخطاء التنفيذ في البرنامج الذي يتم ترجمته

■ A debugger is a program that can be used to determine execution errors in a compiled program

■ كما يقوم المصحح بحفظ معلومات عن برنامج المنبع مثل أرقام السطور ، أسماء المتغيرات ، أسماء الإجراءات .

■ وغالبا ما يتم تجميعه مع المترجم في IDE

## البرامج التي لها علاقة بالمترجم

- وتختلف تجربة running a program مع وجود Debugger عنه في عدم وجوده حيث يقوم Debugger
- بحفظ معلومات برنامج المنبع مثل أرقام السطور و أسماء المتغيرات واسماء الإجراءات
- Line numbers , names of variables , names of procedures .....
- كما يقوم بإيقاف التنفيذ عند نطاق محددة Breakpoints
- ودراسته خارج نطاق المادة



# البرامج التي لها علاقة بالترجم

## ■ 7- Profiles

■ هو برنامج يقوم بتجميع إحصائيات عن سلوك برنامج object program خلال التنفيذ

■ ومثال علي الإحصائيات التي يحتاجها المبرمج من تنفيذ أي برنامج مثلا عدد مرات تنفيذ واستدعاء Procedure وكذلك النسبة المئوية لوقت التنفيذ

■ وهذه الإحصائيات مفيدة للمبرمج لتحسين سرعة تنفيذ البرنامج

■ وفي بعض الأحيان فإن المترجم يقوم باستخدام خرج profiles للتحسين الآلي لبرامج object code دون الرجوع الي المبرمج



# البرامج التي لها علاقة بالمترجم

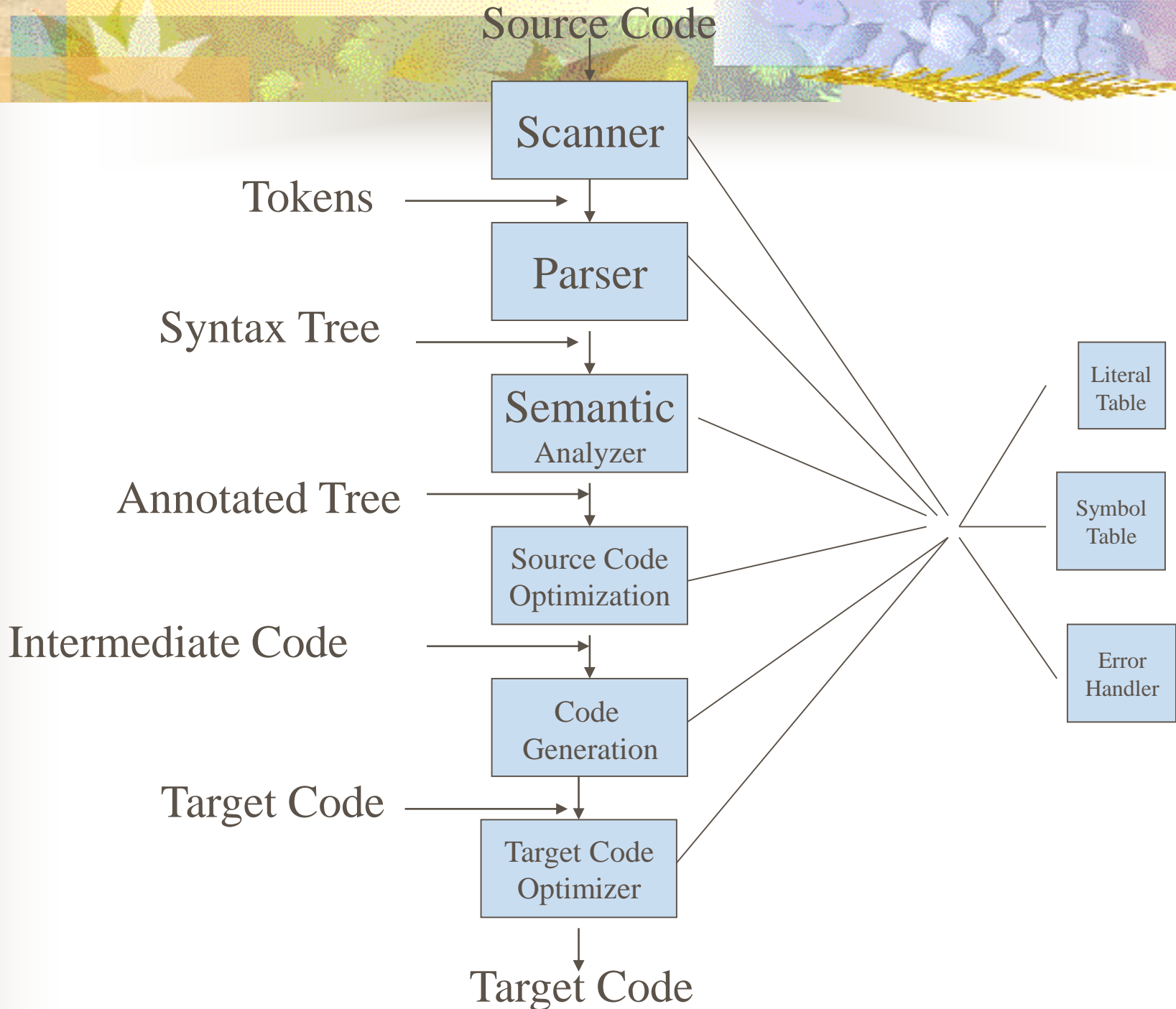
## ٨ – Project Manager :-

في مشاريع الـ Software فإنها تكون كبيرة حتى انه يقوم بتنفيذها مجموعه من المبرمجين وفي هذه الحالة فإن الملف الذي يشترك في عمل هذه المجموعة لابد من عمل تنسيق وهذه وظيفة الـ Project Manager فمثلا يقوم الـ Project Manager بتنسيق الإصدارات المختلفة لنفس الملف والذي ينتجه مبرمجين مختلفين وكذلك فإنه يحفظ تاريخ وثائقي للتغيرات التي تمت على الملف .

# 1.3 The Translation Process

## عملية التحويل

- يحتوي المترجم compiler علي مجموعة من الخطوات منفصلة (or phases) steps التي تكون عمليات منطقية .
- أنه من المفيد التفكير في هذه الخطوات علي أنها عمليات منفصلة داخل المترجم
- وقد يكونوا كذلك في المترجم كقطع منفصلة ولكن عمليا فغالبا ما يتم جمعهم مع بعضهم.
- وهذه الوجوه موجوده بالشكل التالي بالإضافة لثلاث مكونات مساعدة تتفاعل مع بعض أو كل مكونات أو خطوات المترجم كما بالشكل 1-1 .



# 1.3 The Translation Process

## عملية التحويل

- سيتم هنا وصف كل مرحلة باختصار أما الدراسة التفصيلية لتلك المراحل فسيتم عرضها خلال هذا الفصل الدراسي
- المحلل الإملائي-الماسح (lexical analyzer (scanner) الإملائي
- وفي هذه المرحلة فإن المترجم سيقوم بقراءة حقيقية لبرنامج المنبع
- Actual reading of the source program, which is usually in the form of a stream of characters.
- وبعد القراءة سيتم تجميع تتابعات الحروف إلي وحدات لها معنى وتسمى رمز أو علامة (Tokens) كمثال الكلمات في اللغات الطبيعية مثل اللغة العربية و الإنجليزية ومثل spelling

# lexical analyzer (scanner)

## المحلل الإملائي-الماسح الإملائي

■ وكمثال علي ذلك أعتبر التعبير التالي والذي يعتبر جزء من برنامج مكتوب بلغة ال C

■  $a[\text{index}] = 4 + 2$

■ إن بهذا الكود : 12 nonblank characters

■ but only 8 tokens

■ a identifier    [ left bracket    index  
identifier

■ ] right bracket    = assignment    4  
number

■ + plus sign    2 number

■ Each token consists of one or more characters that are collected into a unit before further processing takes place.



# lexical analyzer (scanner)

المحلل الإملائي-الماسح الإملائي

## مقارنة

- Lexical analysis :-
- Natural language “He wrote the program”
- Words :
- “He “ “wrote” “the” “program”
- Programming Language :-
- “ if (b = 0 ) a = b”
- Tokens “if” “(“ “b” “= =“ “0” “)”
- “a” “=“ “b”



# lexical analyzer (scanner)

## المحلل الإملائي-الماسح الإملائي

- In the lexical analysis the characters in the assignment statement :-
- $\text{position} := \text{initial} + \text{rate} * 60$
- Would be grouped into the following tokens
- 1- The identifier position      2- The assignment symbol :=
- 3 The identifier initial      4- The plus sign
- 5- The identifier rate      6- The multiplication sign
- 7- The number 60

# lexical analyzer (scanner)

## المحلل الإملائي-الماسح الإملائي

■ وقد يقوم الماسح الإملائي scanner أيضا بعمليات أخرى  
خلاف التعرف علي ال tokens مثل قد يدخل المعرفات إلي  
جدول الرموز ،،،،

- It may enter identifiers into the symbol table ,  
and it may enter literals into the literal (المعني  
اللفظي) table (Literals include numeric  
constants such as 3.1415926535 and quoted  
strings of strings of text such as “Hello,  
world”

# The Parser (Syntax analyzer)

## المحلل الإعرابي

- يستقبل المحلل الإعرابي برنامج المنبع Source (Code) علي هيئة (tokens) من ال scanner ثم يقوم بعمل التحليل الإعرابي وفيه يتم تعيين هيكل البرنامج (structure of the program) مثل التحليل النحوي علي جملة في اللغة الطبيعية.
- يقوم المحلل الإعرابي بتعيين العناصر المشتركة في عمل الهيكل للبرنامج وإيجاد العلاقة بينها

# The Parser (Syntax analyzer)

## المحلل الإعرابي

- **Syntax analysis determines the structural elements of the program as well as their relations**

■ ونتائج المحلل الإعرابي عادة ما يتم تمثيلها بشجرة الإعراب

- **The results of syntax analysis are usually represented as a parse tree or a syntax tree**