

ITSE321

Software Construction

بناء البرمجيات

المحاضر: د. رضوان حسين
erudwan@yahoo.com

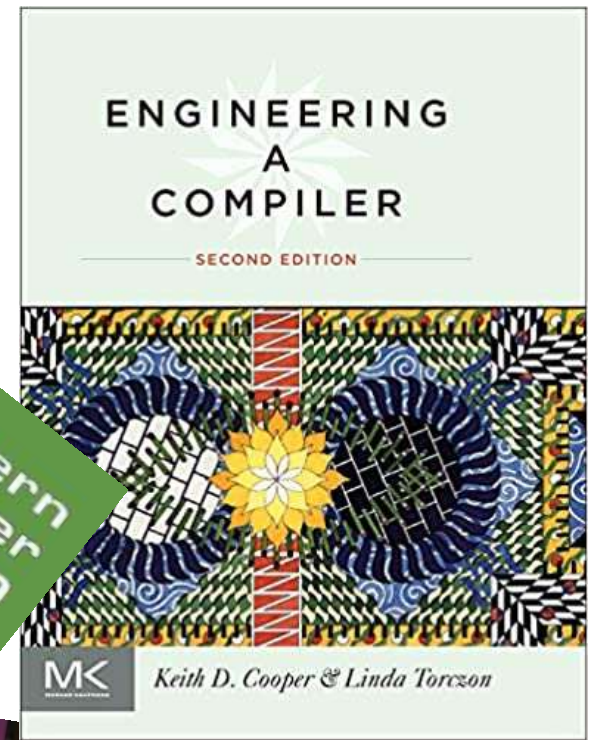
مقدمة إلى المترجمات

Introduction to Compilers

المواضيع

- لماذا ندرس المترجم
- اللغات الطبيعية Natural Languages
 - الألفاظ Lexemes أو كائنات المعاجم lexical entities , عجمة/لفظة/كلمة/مفردة Lexe
 - تركيب الجملة Syntax , ومعاني الألفاظ والجمل Semantics
- لغات الحاسوب Computer Languages
 - تحليل الألفاظ/المفردات Lexical Analysis
 - تحليل تركيب الجمل Syntax Analysis
 - تحليل المعاني Semantic Analysis
- المترجمات Compilers
 - المتطلبات الأساسية للمترجم Basic Requirements
 - عملية الترجمة Compilation Process

مرجع



لماذا ندرس المترجمات Why Compilers؟

- المترجم يعتبر حالة دراسية case study مهمة لمهندس البرمجيات
- بناء المترجم compiler construction يعلم المبرمج مهارات هندسية
- المترجم هو منتج برمجي
- تتعلم كيف تحل المسائل المعقدة
- المترجمات تجسد نظريات وتقنيات كثيرة تهم مهندس البرمجيات
- كيف يعمل الحاسوب والبرمجيات

اللغات الطبيعية

Natural Languages

- في اللغات الطبيعية
 - الجملة عبارة عن سلسلة من الكلمات
 - والكلمة word (وتسمى لفظ lexeme) عبارة عن سلسلة من الرموز (أو حتى رمز واحد)
 - فئة الرموز المستخدمة في لغة ما هي فئة محدودة **finite**
 - تعرف بالحروف الهجائية alphabet
 - فئة الجمل الممكنة للغة ما هي فئة غير محدودة **infinite**
 - القاموس يدرج كل الكلمات **lexemes** في اللغة
 - تصنف الكلمات إلى: أفعال, أسماء, حروف, ضمائر

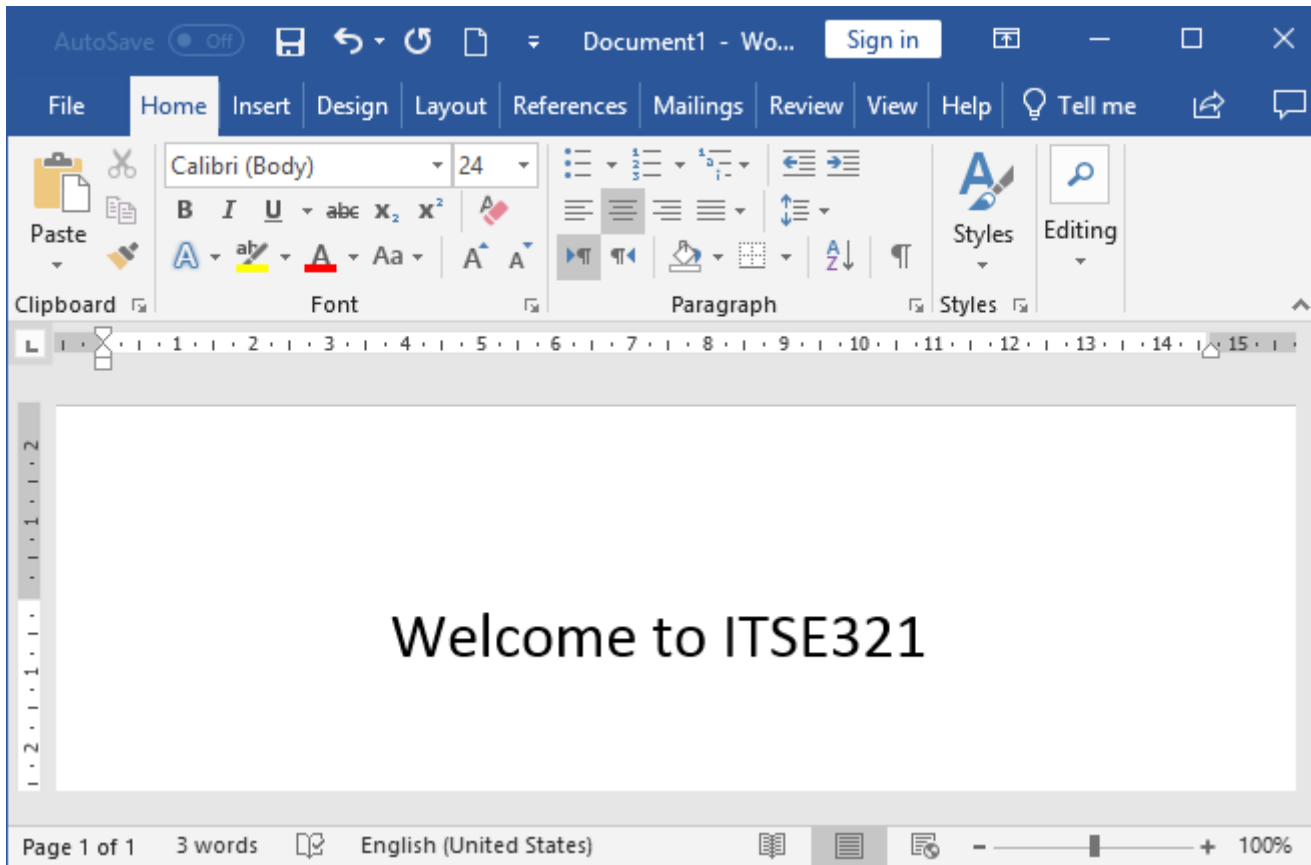
قواعد اللغات الطبيعية: النحو Grammar

- Grammar النحو يعتبر مجموعة من القواعد لصياغة نص سليم Syntax Rules
- سلسلة الكلمات (النص) (syntax) sequence of words يجب أن يكون لها بناء structure يتقيد بالقواعد النحوية grammatical rules
- الجمل المصاغة جيداً عادةً يكون لها معنى meaning يفهمه الناس

تحليل الجُمْل Analysis of Sentences

- تحليل المفردات Lexical Analysis
 - التعرف على الكلمات المكونة من الرموز
 - تصنف الكلمات إلى: أسماء, أفعال, صفات, أحوال, حروف, ضمائر, ...
- تحليل النص Syntax Analysis
 - مدى مراعاة النص لقواعد تركيب الكلمات لتشكيل جمل
- تحليل المعاني Semantic Analysis
 - يسهل على الإنسان تحليل المعنى
 - ولكن تعاني الآلات من هذا, طبعاً تتطور تقنيات معالجة اللغات
 - من مجالات البحث العلمي التي تلاقي الاهتمام **Deep Learning**

من تطبيقات الإعراب Parsing



MS Office documents are based on XML format

Raw XML files can be viewed in all major browsers.
Don't expect XML files to be displayed as HTML pages.

Viewing XML Files

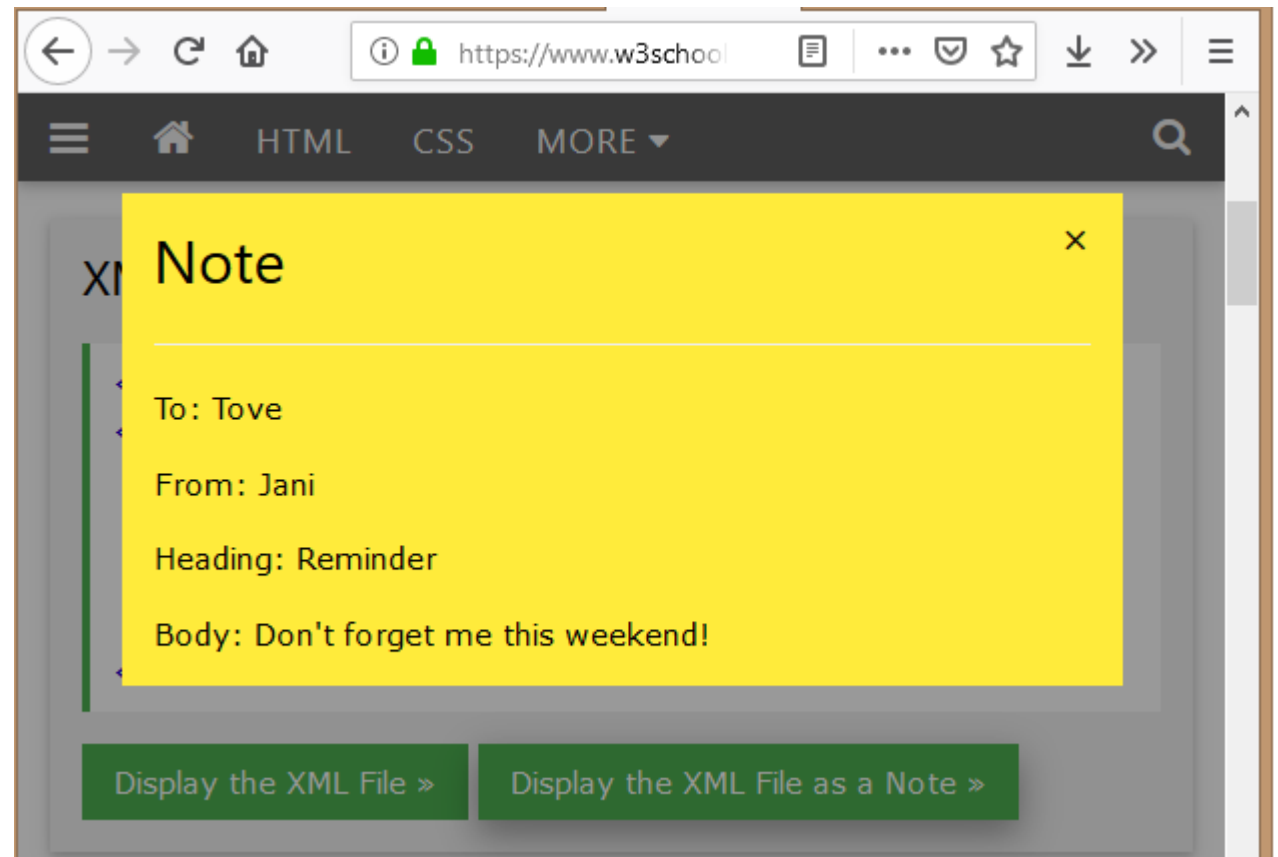
```
<?xml version="1.0" encoding="UTF-8"?>  
- <note>  
    <to>Tove</to>  
    <from>Jani</from>  
    <heading>Reminder</heading>  
    <body>Don't forget me this weekend!</body>  
</note>
```

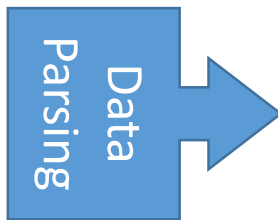
note.xml

```
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```

XML is a software- and hardware-independent tool for storing and transporting data

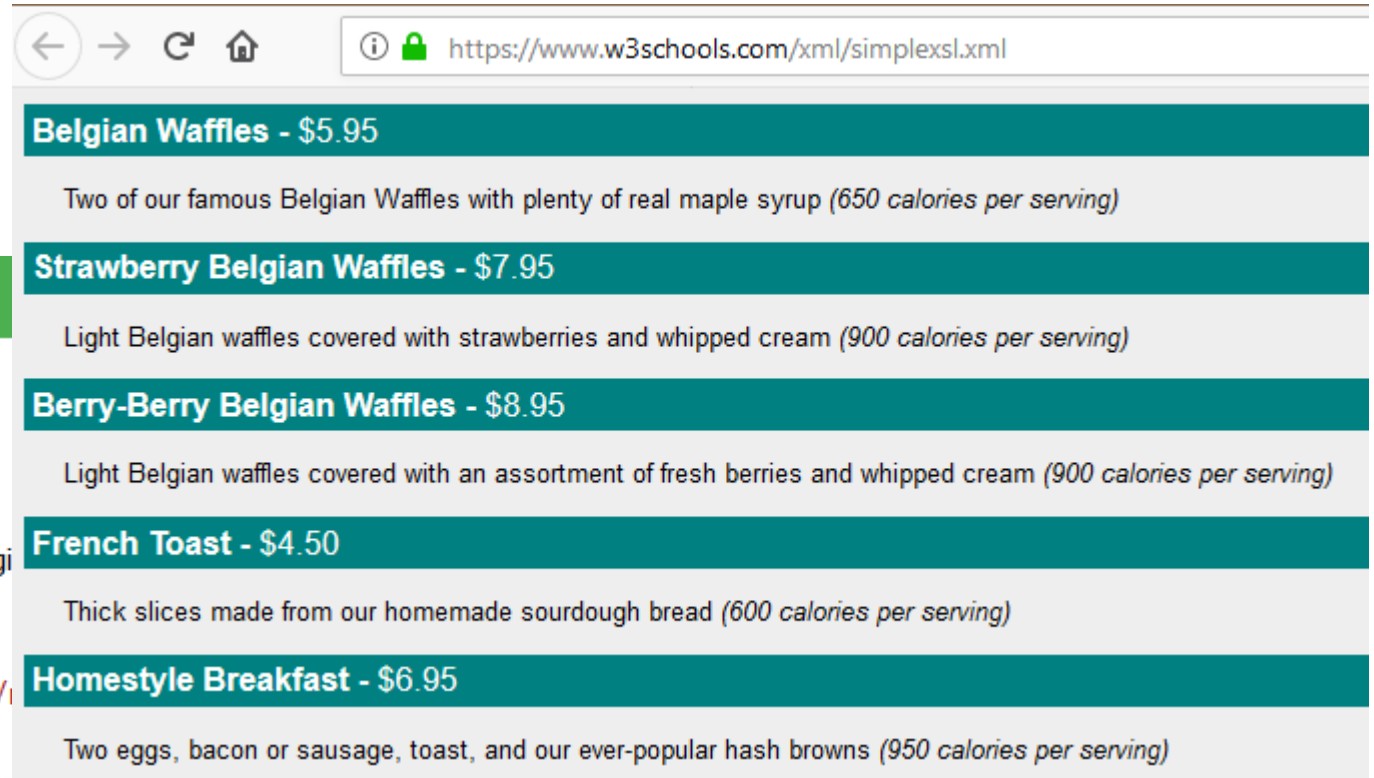
All major browsers have a built-in XML parser to access and manipulate XML.





simple.xml

```
<breakfast_menu>
  <food>
    <name>Belgian Waffles</name>
    <price>$5.95</price>
    <description>Two of our famous Belgi
    <calories>650</calories> </food>
  <food>
    <name>Strawberry Belgian Waffles</n
    <price>$7.95</price>
    <description>Light Belgian waffles covered with strawberries and whipped cream</description>
    <calories>900</calories> </food>
  <food>
    <name>Berry-Berry Belgian Waffles</name>
    <price>$8.95</price>
    <description>Light Belgian waffles covered with an assortment of fresh berries and whipped cream</description>
    <calories>900</calories> </food>
  <food>
    <name>French Toast</name>
    <price>$4.50</price>
    <description>Thick slices made from our homemade sourdough bread</description>
    <calories>600</calories> </food>
```



اللغات الطبيعية والحاسوب

- نحن نحاول نعلم الآلات لغاتنا الطبيعية

Google's human-sounding AI to answer calls at contact centers: Google's chatbots



With mixed results!!



راني قاعد في القهوة

معالجة لغات الحاسوب Computer Language Processing

- في لغات برمجيات الحاسوب computer programming languages, التعامل يكون مع البرامج التي تجسد أيضاً جملاً طويلة ونصوص
 - النص سلسلة من المفردات sequence of lexical units or lexemes
 - والمفردات تتكون من رموز Lexical unit are sequence of characters
 - قواعد المفردات Lexical Rules اللغة الحاسوب هي المفردات المحجوزة للغة
- valid lexical units**
- هناك تصنيفات عدة لمفردات لغة البرمجة: معرفات identifier, عدد number, رمز character, نص string, مشغل operator, ...

معالجة لغات الحاسوب Computer Language Processing

- قواعد النصوص Syntax Rules في لغة الحاسوب تحدد ما هو تسلسل المفردات المشكلة/المصاغة جيداً حسب لغة البرمجة

- معنى التسلسل السليم للمفردات يسمى أيضاً its semantics

- يمكن أن يكون البرنامج مشكلاً جيداً, ولكن معناه مبهم أو غير واضح

- مثال:

```
int x = 0;  
x = 1;  
x = 0;
```

- نصياً, التعبيرات صحيحة, ولكن ماذا تعني؟

معالجة لغات الحاسوب Computer Language Processing

- مترجم لغة الحاسوب compiler عليه التعرف على أخطاء المفردات والنصوص والتنبيه على تلك الأخطاء
- قد تنبهك المترجمات عن الأخطاء اللغوية الشائعة

```
public boolean test (int x){  
    boolean result;  
    if (x > 100)  
        result = true;  
    return result;  
}
```

Error message/warning:

The local variable *result* may have not been initialized

ماذا لو كانت *x* ليست أكبر من 100، ما هي
قيمة *result* ؟

- وقد يشتكى مسؤول المشروع أو العميل من باقي المشاكل المترتبة عن تنفيذ

نهائية المحاضرة,
موضوعنا التالي:

المترجم

The Compiler