



## CENTER FOR TEXTUAL STUDIES AND DIGITAL HUMANITIES

# DIGH 401 - Introduction to Computing

Fall Semester 2013

Week 5

## Today's Class

- Programming tools
- XML

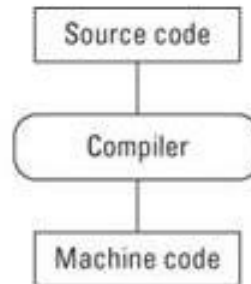
## Programming Tools

- an editor and compiler or interpreter
- compiler converts source code into machine code
- machine code saved as an executable
- debugger and profiler
- disassembler

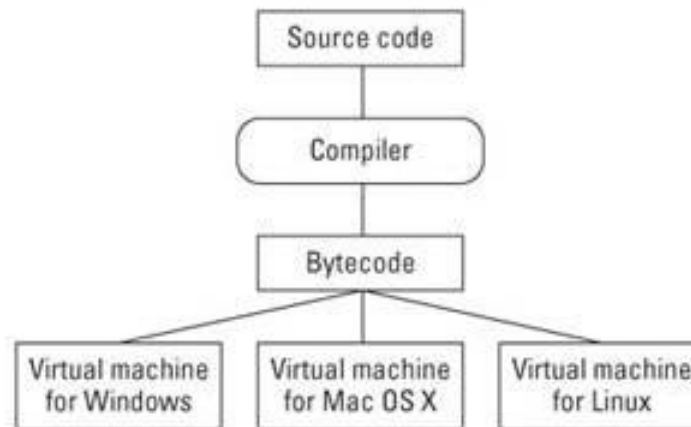
## Programming Tools - Virtual Machine

- compilers are often difficult to make for multiple OSs and processors
- interpreters need the source code to run
  - often unsuitable for broadly distributing software
- virtual machine was designed to address these issues
- bytecode or pseudocode (p-code)

# DIGH 401 - Introduction to Computing



A compiler normally converts source code directly into machine code for a specific type of processor.



When compiled to bytecode, a program can run on any operating system that has the bytecode virtual machine installed.

# DIGH 401 - Introduction to Computing

## Programming Tools - Virtual Machine

- reverse engineer from bytecode to source code
- Java currently most popular language to use a VM
- Java compiles source code into bytecode and runs using a VM

## Programming Tools - Sandbox

- testing environment that isolates untested code changes, experiments...
- protects live and active servers and their data
- replicate at least minimal functionality to enable testing and development...
- version control software such as Mercurial and Subversion

[Google Code](#) | [GitHub](#)

## Introduction to XML

- XML = EXtensible Markup Language
- markup language similar to HTML
- designed to carry data but not display data
- XML tags are not pre-defined, you must define your own
- designed to be self-descriptive
- XML is a W3C recommendation



## Introduction to XML

```
<calendar>  
<date>20th September 2012</date>  
<title>DIGH 401</title>  
<location>Room 201 Cudahy Library</location>  
</calendar>
```

## Introduction to XML

- separation of data from display
- simplifies data sharing and transport
- simplifies and eases transition to new platforms and upgrades
- XML becomes more accessible
- XML has also been used to create new languages

## XML Structure

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<calendar type="personal">  
  <date>20th September 2012</date>  
  <title>DIGH 401</title>  
  <location>Room 201 Cudahy Library</location>  
</calendar>
```

## XML Structure

### Syntax

- there must be closing tags

eg: `<p>a new paragraph...</p>` and not  
`<p>a new paragraph...`

- tags are case-sensitive
- elements must be properly nested

eg: `<bold><italic>a new phrase...</italic></bold>` and not  
`<bold><italic>a new phrase...</bold></italic>`

- must have a root element
- attribute values must be quoted
- entity references

eg: using a character such as `<` instead of `&lt;`;

## XML Structure

### Entity References

<	&lt;	less than
>	&gt;	greater than
&	&amp;	ampersand
'	&apos;	apostrophe
"	&quot;	quotation mark

## XML Structure

### Syntax

- comments

`<!-- this is a comment in XML -->`

## XML Structure

### Elements

```
<library>
  <book category="fiction">
    <title>To the Lighthouse</title>
    <author>Virginia Woolf</author>
    <year>1927</year>
  </book>
</library>
```

- everything from the start to end tag of an element, inclusive of the tags
- can contain
  - other elements
  - text
  - attributes
  - or a mix of all of the above...

## XML Structure

### Elements

- Naming rules
  - elements can contain letters, numbers, and other characters
  - elements cannot start with a number or punctuation character
  - elements cannot start with the letters xml (or XML, or Xml, etc)
  - elements cannot contain spaces
- no words are reserved considering the above rules
- best practice examples for element naming
- XML elements are extensible



## XML Structure

### Attributes

```
<book type="print">Hannibal's Footsteps</book>
```

- additional information about the element
- attribute values must be quoted
- double or single quotes
- no strict rules when to use attribute and when to use element
- some issues with attributes are:
  - cannot contain multiple values (elements can)
  - cannot contain tree structures (elements can)
  - not easily expandable (for future changes)
- attribute values for metadata

```
<book id="4024">A Good Year</book>
```

## XML Tests

Think about how you might encode the following information:

- a car
- a sports team
- 2 Musical CDs including each song per album per CD (each album has 5 songs)

# DIGH 401 - Introduction to Computing

## To the Lighthouse

(Here Mr. Carmichael, who was reading Virgil, blew out his candle.)

### 3

But what after all is one night? A short space, especially when the darkness dims so soon, and so soon a bird sings, a cock crows loudly, or a faint green quickens, like a turning leaf, in the hollow of the wave.

It seemed now as if, touched by human penitence and all its tool, divine goodness had parted the curtain and displayed behind it, single, distinct, the hare erect; the wave falling; the boat rocking, which, did we deserve them, should be