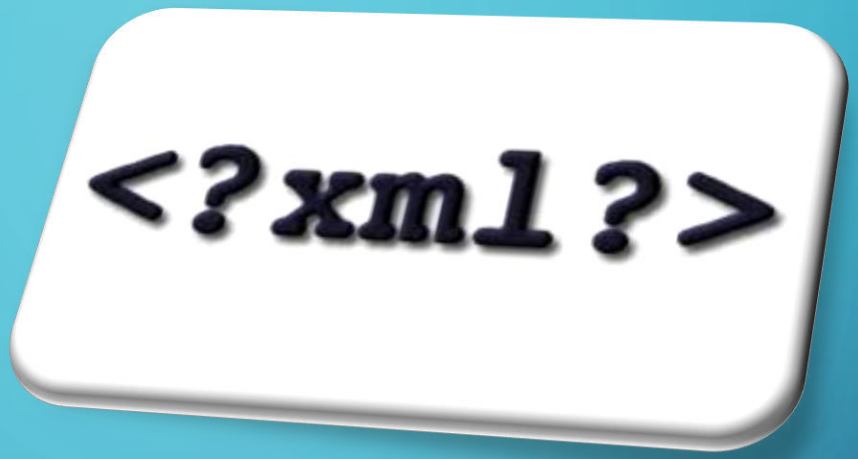

SAX AND DOM METHODOLOGIES

XML

XML Models

- DOM
  - Data Object Model: *The W3C Document Object Model (DOM) is a platform and language-neutral interface that allows programs and scripts to dynamically access and update the content, structure, and style of a document."*
- SAX
  - Simple API for XML: *SAX provides a mechanism for reading data from an XML document that is an alternative to that provided by the Document Object Model (DOM).*

Why is XML Important?

- Because it exists, and everybody uses it.
- Plain Text - you can create and edit files with anything.
- Data Identification - XML tells you what kind of data you have, not how to display it.
- Separation from style.
- Hierarchical, and easily processed.

XML Text

- XML has only one “basic” type -- text. It is bounded by tags e.g.

  <title> The Big Sleep </title>

  <year> 1935 </ year> --- 1935 is still text

- XML text is called PCDATA (for parsed character data). It uses a 16-bit encoding,

XML Structure

- XML consists of *tags* and *text*
- Tags come in pairs <date> ...</date>
- They must be properly nested

  <date> <day> ... </day> ... </date> --- good

  <date> <day> ... </date>... </day> --- bad

Nesting Tags

- Nesting tags can be used to express various structures. E.g. A tuple (record):

```
<person>
        <name> Jeff Cohen</name>
        <tel> 04-828-1345 </tel>
        <cell> 054-470-778 </cell>
        <email> jeffc@cs.technion.ac.il </email>
</person>
```

Tag Lists

- Nested tags can be part of a list too:

```
<addresses>
        <person>
                <name> Yossi Orr</name>
                <tel> 04-828-1345 </tel>
                <email> yossio@cs.technion.ac.il </email>
        </person>
        <person>
                <name> Irma Levy</name>
                <tel> 03-426-1142 </tel>
                <email>irmal@yourmail.com</email>
        </person>
</addresses>
```

XML Tree Structure