

CGT 141/CPT 141 Lecture 19 Wk 12

Wellformedness and Validity – An XML Introduction

xHTML

- xHTML is an interim step to XML.
- Throughout the semester we have been following the rules of wellformedness, as it relates to xHTML:
 - Lowercase for all tags and attributes
 - What you will find is that xHTML and XML are case sensitive
 - Including all required structural tags (including `<!DOCTYPE>`)
 - Closing all tags (even when optional)
 - Properly nesting tags (to avoid orphaned or overlapped tags)
 - Quoting all attribute values
 - Further rules in xHTML (to make your documents xHTML):
 - Attributes always have a value:
 - Where an attribute is a single statement, such as `noshade` for `<hr>` use:
 - `noshade="noshade"`
 - Use only certain escapes sequences
 - Ampersand (`&`): `&`
 - Less than (`<`): `<`
 - Greater than (`>`): `>`
 - Apostrophes (`'`): `'`
 - Quotes (`"`): `"`
 - Close empty tags
 - `

`
 - `<hr><hr />`
 - ``
- You would also need to add the following as the first line in your document:
 - `<?xml version="1.0">`

Wellformed and Valid

- Wellformed means that a document conforms to the syntactical rules for markup.
- A valid document is a document that is well-formed and also conforms to its DTD or other type of schema
 - A DTD is a machine-readable schema that can be read by a validating parser or user agent.
 - Thus, a schema is a schema specific customized to a particular group or working agency.

What is XML?

- XML is about the classification of data in a document
 - Over time, HTML has departed from its original purpose, at least in use anyway.
 - HTML was originally designed to markup content, not format it. Thus several things have

come along to try to relieve the formatting burden from HTML, not the least of which include DHTML and CSS.

- However, HTML has always been poor as far as classification goes.

- XML is a meta markup language
 - Meta means self-describing
 - So, XML includes data about itself, that is, how it should be read.
- XML can be used to create one's own markup language
 - ChemML, MusicML, MathML are examples.
- Once an XML document is written, a schema can be defined for it.