CpSc 4620/6620: Database Management Systems (DBMS) (TEXNH Approach)

HTML4 vs HTML5

James Wang

Why HTML5?

- **HTML has been in continuous evolution since it was introduced to the Internet in the early 1990s.**
- **HTML4 became a W3C Recommendation in 1997.**
  - While HTML4 continues to serve as a rough guide to many of the core features of HTML, it does not provide enough information to build implementations that interoperate with each other and, more importantly, with Web content. The same goes for XHTML1, which defines an XML serialization for HTML4, and DOM Level 2 HTML, which defines JavaScript APIs for both HTML and XHTML.
- **The HTML5 specification, started in 2004:**
  - Defines a single language called HTML which can be written in HTML syntax and in XML syntax.
  - Defines detailed processing models to foster interoperable implementations.
  - Improves markup for documents.
  - Introduces markup and APIs for emerging idioms, such as Web applications.

2

HTML5 vs HTML4: DOCTYPE

- **The HTML syntax requires a doctype to be specified to ensure that the browser renders the page in standards mode. The doctype has no other purpose. The doctype declaration for the HTML syntax is case-insensitive.**

```
<!DOCTYPE html>
```

- **Doctypes from earlier versions of HTML were longer because the HTML language was SGML-based and therefore required a reference to a DTD.**

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
"http://www.w3.org/TR/html4/frameset.dtd">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

- **Looking at these different types of DOCTYPE declaration, you know why a simple HTML5 DOCTYPE is needed.**

3

HTML5 vs HTML4: New Elements

- **section represents a generic document or application section. It should be used together with the h1, h2, h3, h4, h5, and h6 elements to indicate the document structure.**
- **article represents an independent piece of content of a document, such as a blog entry or newspaper article.**
- **main represents the main content of the body of a document or application.**
- **aside represents a piece of content that is only slightly related to the rest of the page.**
- **header represents a group of introductory or navigational aids.**
- **footer represents a footer for a section and can contain information about the author, copyright information, etc.**
- **nav represents a section of the document intended for navigation..**

4

HTML5 New Elements

- **figure represents a piece of self-contained flow content, typically referenced as a single unit from the main flow of the document.**

```
<figure>
 <video src="example.webm" controls></video>
 <figcaption>Example</figcaption>
</figure>
```

- **figcaption can be used as caption (it is optional).**
- **template can be used to declare fragments of HTML that can be cloned and inserted in the document by script.**
- **video and audio for multimedia content. Both provide an API so application Web developers can script their own user interface, but there is also a way to trigger a user interface provided by the user agent. source elements are used together with these elements if there are multiple streams available of different types.**

5

HTML5 New Elements

- **track provides text tracks for the video element.**
- **embed is used for plugin content.**
- **mark represents a run of text in one document marked or highlighted for reference purposes, due to its relevance in another context.**
- **progress represents a completion of a task, such as downloading or when performing a series of expensive operations.**
- **meter represents a measurement, such as disk usage.**
- **time represents a date and/or time.**
- **ruby, rt, and rp allow for marking up ruby annotations.**
- **bdi represents a span of text that is to be isolated from its surroundings for the purposes of bidirectional text formatting.**
- **wbr represents a line break opportunity.**

6

HTML5 New Elements

- canvas is used for rendering dynamic bitmap graphics on the fly, such as graphs or games.
- datalist together with the a new list attribute for input can be used to make comboboxes:

```
<input list="browsers">
<datalist id="browsers">
 <option value="Safari">
 <option value="Internet Explorer">
 <option value="Opera">
 <option value="Firefox">
</datalist>
```

- keygen represents control for key pair generation.
- output represents some type of output, such as from a calculation done through scripting.

7

HTML Video and Audio

- video

```
<video width="320" height="240" controls>
 <source src="myvideo.mp4" type="video/mp4">
 <source src="myvideo.ogg" type="video/ogg">
Your browser does not support the video tag.
</video>
```

- audio

```
<audio controls>
 <source src="myaudio.ogg" type="audio/ogg">
 <source src="myaudio.mp3" type="audio/mpeg">
Your browser does not support the audio element.
</audio>
```

8

HTML5 New Input Type

- The input element's type attribute now has the following new values:
  - tel
  - search
  - url
  - email
  - date
  - time
  - number
  - range
  - color
- The idea of these new types is that the user agent can provide the user interface, such as a calendar date picker or integration with the user's address book, and submit a defined format to the server. It gives the user a better experience as his input is checked before sending it to the server meaning there is less time to wait for feedback.

9

HTML5 vs HTML4: Changed Elements

- The address element is now scoped by the nearest ancestor article or body element.
- The b element now represents a span of text to which attention is being drawn for utilitarian purposes without conveying any extra importance and with no implication of an alternate voice or mood.
- The blockquote element still represents content that is quoted from another source but now also allows including a citation in a footer or cite element as well as inline changes such as annotations and abbreviations.
- The dl element now represents an association list of name-value groups, and is no longer said to be appropriate for dialogue.
- The hr element now represents a paragraph-level thematic break.

10

HTML5 vs HTML4: Changed Elements

- The i element now represents a span of text in an alternate voice or mood, or otherwise offset from the normal prose in a manner indicating a different quality of text.
- For the label element the browser should no longer move focus from the label to the control unless such behavior is standard for the underlying platform user interface.
- The noscript element is no longer said to be rendered when the user agent doesn't support a scripting language invoked by a script element earlier in the document.
- The s element now represents contents that are no longer accurate or no longer relevant.
- The script element can now be used for scripts or for custom data blocks.
- The small element now represents side comments such as small print.

11

HTML5 vs HTML4: Changed Elements

- The strong element now represents importance rather than strong emphasis.
- The u element now represents a span of text with an unarticulated, though explicitly rendered, non-textual annotation, such as labeling the text as being a proper name in Chinese text (a Chinese proper name mark), or labeling the text as being misspelt.

12

HTML5 vs HTML4: Changed Attributes

- The alt attribute on img has more elaborate requirements, and in some cases can also be omitted.
- The accept attribute on input now allows the values audio/*, video/* and image/*.
- The accesskey global attribute now allows multiple characters to be specified, which the user agent can choose from.
- The action attribute on form is no longer allowed to have an empty URL.
- The border attribute on table only allows the values "1" and the empty string.
- The colspan attribute on td and th now has to be greater than zero.
- The coords attribute on area no longer allows a percentage value of the radius when the element is in the circle state.

13

HTML5 vs HTML4: Changed Attributes

- The data attribute on object is no longer said to be relative to the codebase attribute.
- The defer attribute on script now explicitly makes the script execute when the page has finished parsing.
- The dir global attribute now allows the value auto.
- The enctype attribute on form now supports the value text/plain.
- The width and height attributes on img, iframe and object are no longer allowed to contain percentages. They are also not allowed to be used to stretch the image to a different aspect ratio than its intrinsic aspect ratio.
- The href attribute on link is no longer allowed to have an empty URL.
- The href attribute on base is now allowed to contain a relative URL.

14

HTML5 vs HTML4: Changed Attributes

- All attributes that take URLs, e.g. href on the a element, now support IRIs if the document's encoding is UTF-8 or UTF-16.
- The http-equiv attribute on meta is no longer said to be used by HTTP servers to create HTTP headers in the HTTP response. Instead, it is said to be a pragma directive to be used by the user agent.
- The id global attribute is now allowed to have any value, as long as it is unique, is not the empty string, and does not contain space characters.
- The lang global attribute takes the empty string in addition to a valid language identifier, just like xml:lang does in XML.
- The media attribute on link now accepts a media query list and defaults to "all".
- The event handler attributes (e.g. onclick) now always use JavaScript as the scripting language.

15

HTML5 vs HTML4: Changed Attributes

- The value attribute of the li element is no longer deprecated as it is not presentational. The same goes for the start and type attributes of the ol element.
- The style global attribute now always uses CSS as the styling language.
- The tabindex global attribute now allows negative values which indicate that the element can receive focus but cannot be tabbed to.
- The target attribute of the a and area elements is no longer deprecated, as it is useful in Web applications, e.g. in conjunction with iframe.
- The type attribute on script and style is no longer required if the scripting language is JavaScript and the styling language is CSS, respectively.
- The usemap attribute on img no longer takes a URL, but instead takes a valid hash-name reference to a map element.

16

HTML5 vs HTML4: Obsolete Elements

- The following elements are not in HTML5 because their effect is purely presentational and their function is better handled by CSS:
  - basefont
  - big
  - center
  - font
  - strike
  - tt
- The following elements are not in HTML5 because using them damages usability and accessibility:
  - frame
  - frameset
  - noframes

17

HTML5 vs HTML4: Obsolete Elements

- The following elements are not included because they have not been used often, created confusion, or their function can be handled by other elements:
  - acronym is not included because it has created a lot of confusion. Web developers are to use abbr for abbreviations.
  - applet has been obsoleted in favor of object.
  - isindex usage can be replaced by usage of form controls.
  - dir has been obsoleted in favor of ul.
- Finally the noscript element is only conforming in the HTML syntax. It is not allowed in the XML syntax. This is because in order to not only hide visually but also prevent the content to run scripts, apply style sheets, have submittable form controls, load resources, and so forth, the HTML parser parses the content of the noscript element as plain text. The same is not possible with an XML parser.

18

HTML5 vs HTML4: Obsolete Attributes

- Some attributes from HTML4 are no longer allowed in HTML5.
  - rev and charset attributes on link and a.
  - shape and coords attributes on a.
  - longdesc attribute on iframe.
  - target attribute on link.
  - nohref attribute on area.
  - profile attribute on head.
  - version attribute on html.
  - name attribute on img (use id instead).
  - scheme attribute on meta.
  - archive, classid, codebase, codetype, declare and standby attributes on object.
  - valuetype and type attributes on param.
  - axis attribute on td and th.
  - abbr and scope attributes on td.
  - summary attribute on table.
  - accept attribute on form.
  - usemap attribute on input.

19

HTML5 vs HTML4: Obsolete Attributes

- All presentational attributes are obsoleted in HTML5 that were in HTML4. Use CSS:
- align attribute on caption, iframe, img, input, object, legend, table, hr, div, h1, h2, h3, h4, h5, h6, p, col, colgroup, tbody, td, tfoot, th, thead and tr.
- alink, link, text and vlink attributes on body.
- background attribute on body.
- bgcolor attribute on table, tr, td, th and body.
- border attribute on object.
- cellpadding and cellspacing attributes on table.
- char and charoff attributes on col, colgroup, tbody, td, tfoot, th, thead and tr.
- clear attribute on br.
- compact attribute on dl, ol and ul.
- frame attribute on table.

20

HTML5 vs HTML4: Obsolete Attributes

- More obsoleted presentational attributes:
- frameborder attribute on iframe.
- height attribute on td and th.
- hspace and vspace attributes on img and object.
- marginheight and marginwidth attributes on iframe.
- noshade attribute on hr.
- nowrap attribute on td and th.
- rules attribute on table.
- scrolling attribute on iframe.
- size attribute on hr.
- type attribute on li, and ul.
- valign attribute on col, colgroup, tbody, td, tfoot, th, thead and tr.
- width attribute on hr, table, td, th, col, colgroup and pre.

21

HTML5 vs HTML4: Obsolete Attributes

- The following attributes are allowed but Web developers are discouraged from using them and instead strongly encouraged to use an alternative solution:
- The border attribute on img. It is required to have the value "0" when present. Web developers can use CSS instead.
- The language attribute on script. It is required to have the value "JavaScript" (case-insensitive) when present and cannot conflict with the type attribute. Web developers can simply omit it as it has no useful function.
- The name attribute on a. Web developers can use the id attribute instead.

22

HTML5 New APIs

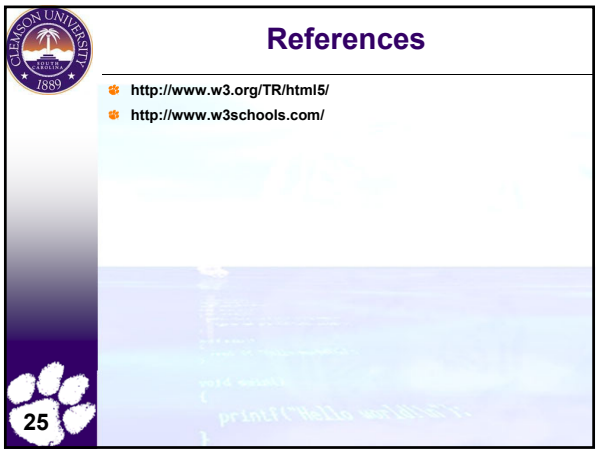
- HTML introduces a number of APIs that help in creating Web applications:
- Media elements (video and audio) have APIs for controlling playback, syncronising multiple media elements, and timed text tracks (e.g. subtitles).
- An API for form constraint validation (e.g. the setCustomValidity() method).
- An API that enables offline Web applications, with an application cache.
- An API that allows a Web application to register itself for certain protocols or media types, using registerProtocolHandler() and registerContentHandler().
- Editing API in combination with a new global contenteditable attribute.
- An API that exposes the components of the document's URL and allows scripts to navigate, redirect and reload (the Location interface).

23

HTML5 New APIs

- An API that exposes the session history and allows scripts to update the document's URL without actually navigating, so that applications don't need to abuse the fragment component for "Ajax-style" navigation (the History interface).
- An API for base64 conversion (atob() and btoa() methods).
- An API to schedule timer-based callbacks (setTimeout() and setInterval()).
- An API to prompt the user (alert(), confirm(), prompt()).
- An API for printing the document (print()).
- An API for handling search providers (AddSearchProvider() and IsSearchProviderInstalled()).
- The Window, Navigator and External interfaces have been defined.

24


References
http://www.w3.org/TR/html5/
http://www.w3schools.com/
25