XHTML

Fulvio Corno, Dario Bonino
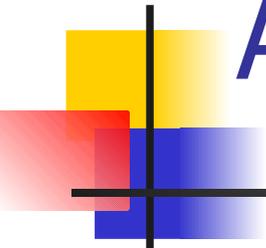
Agenda

What will be done?

- HTML / XHTML
- Practical experiments

What will be achieved?

- Capability to correctly write basic XHTML pages (your web page as an example)
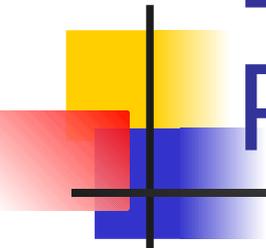
Document creation and Publishing on the Web

It is possible to create WWW pages and to make them public and accessible from the Internet.

A simple text editor is sufficient.

Many browsers already integrate a web editor.

Pages are written either in HTML (old) or XHTML.
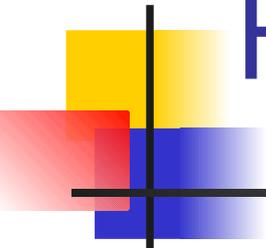
HTML

HTML: HyperText Markup Language

Evolving standard

Provides a structured description of a hypertext

It is "software" independent: shall be understandable by different browsers (with the same graphic effects)

Text based: it uses the first 127 characters of the ASCII code
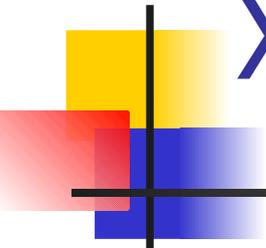
XHTML

XHTML: eXtensible HyperText Markup Language

XHTML replaces HTML

Pretty similar to HTML 4.01

A more strict and clean version of HTML

XHTML stems from XML

W3C Recommendation of January, the 26th, 2000
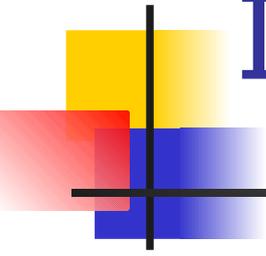
Introduction

To visualize HTML or XHTML specific programs must be used, called "browsers"

The most known web browsers are:

- Mozilla (Navigator / Firefox)
- Netscape Navigator
- Opera
- Microsoft Internet Explorer
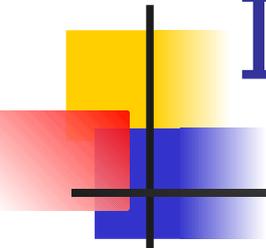
Introduction (2)

To write HTML or XHTML code you only need:

- A text editor
- Or a specific program
  - Ultraedit / EditPlus (Win32), Emacs (Linux/Unix), jEdit (os-independent)
  - Dreamweaver, HotDog, 1dtPage (win32)
  - Quanta, BlueFish (Linux)
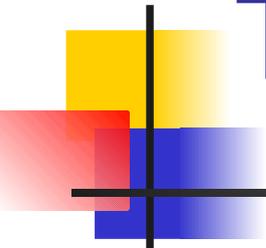
Tags & tags

HTML and XHTML tag a document allowing to specify different parts composing the doc:

- Headings
- Lists
- Tables
- ...

The document is tagged using "tags"

The browsers read the tags and visualize the hypertext properly
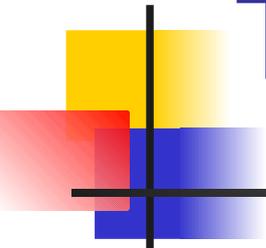
Tags?

Expressions / constructs always comprised between a "less than" sign (<) and a "greater than" sign (>)

Tags are in couples: a start tag and an end tag that delimits the tagged text

- **Ex:** `<h1>this is a heading</h1>`
- The end tag is always like the start tag but it starts with a / so the end tag for <p> is </p>
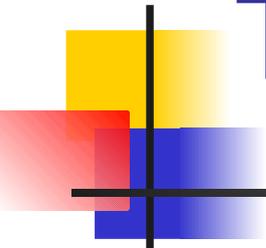
Tags? (2)

Tags are not isolated (in XHTML), however some of them can be "self-closing":

- `<br/>`

- In this case the start tag and the end tag are collapsed into a single self-closing tag

In HTML tags are case-insensitive; in XHTML, conversely, they are strictly case sensitive and <u>must be written in lower case</u>

- `<h1></h1>` (OK)

- `<H1></h1>` (HTML= OK, XHTML=KO)=KO
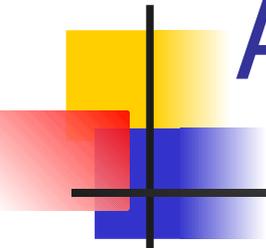
Attributes

Attributes allow to better specify what a given tag does

- Example: specify the destination of a link, the alternative text of an image, etc.

They are basically a named variable to which a value is assigned

- **Example:** `<a href="www.google.com">...</a>`

The value shall always be in between two quotation marks "" (XHTML)
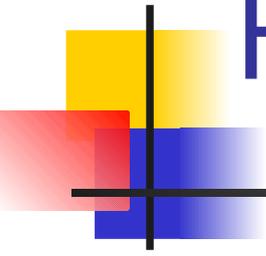
HTML vs XHTML

In XHTML tags shall be nested properly

In XHTML documents shall be well formed

In XHTML tags shall always be in lower case

In XHTML tags shall always be closed (or self-closing)
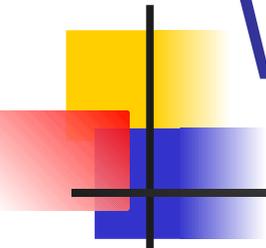
Well formed or not?

What means "well formed"?

- All the elements (tags) shall be nested in between their root tag (<html> </html> for XHTML/HTML)
- Elements can have children and descendants
- The sub-elements shall be in couples and shall be nested properly into the father element
  - Example:

     first item

     second item
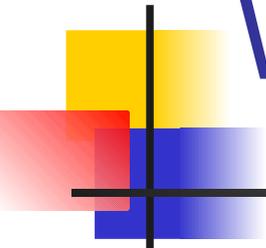
Which XHTML?

There are three versions of XHTML:

- Strict
- Transitional
- Frameset

Each version is defined by a DTD (document type definition)

A DTD basically specifies what can be written in a XHTML document (and also in a XML document or in a SGML document)
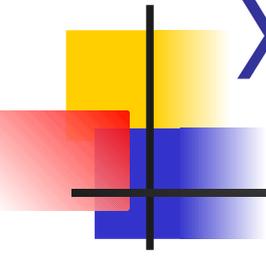
XHTML DTDs

Strict:

- <pre><code><!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0
Strict//EN" "http://www.w3.org/TR/
xhtml1/DTD/xhtml1-strict.dtd"></code></pre>

- Strict and clean: clear separation between contents and presentation

- Presentation ☐ CSS (style sheets)
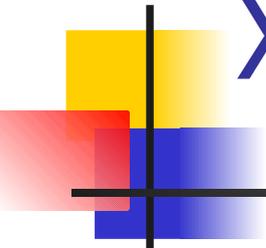
XHTML DTDs (2)

Transitional

- ```
  <!DOCTYPE html PUBLIC "-//W3C//DTD
  XHTML 1.0 Transitional//EN" "http://
  www.w3.org/TR/xhtml1/DTD/
  xhtml1-transitional.dtd">
  ```

- It is nearly equivalent (in functionalities) to HTML4.01

- HTML presentation tags can be used (but their use is discouraged)

- Can be used for browsers not supporting CSS (old browsers)
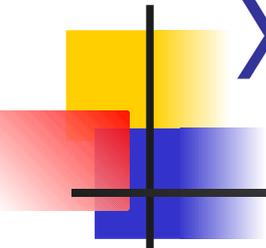
XHTML DTDs (3)

Frameset:

- ```
  <!DOCTYPE html PUBLIC "-//W3C//DTD
  XHTML 1.0 Frameset//EN"
  "http://www.w3.org/TR/xhtml1/DTD/
  xhtml1-frameset.dtd">
  ```

- For frame-based sites

- DO NOT USE IT: FRAMES ARE NOT USABLE / ACCESSIBLE

  - Unaddressable resources: URL does not change as a user navigates his way through the site
  - The focus is on layout, not structure
  - Totally incompatible (NOFRAMES element as ALT)
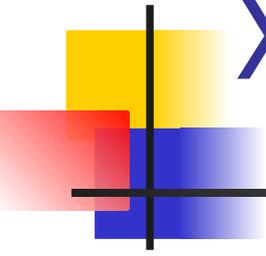
XHTML documents

We work with XHTML transitional

No frames

Accessibility is an issue that shall be tackled

Style is completely separated from content

XHTML for structuring content

CSS for formatting

XHTML documents (2)

The basic structure is:

```
<?xml version="1.0" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0
Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-
transitional.dtd">
```

→ XML and DTD declaration

```
<html xmlns="http://www.w3.org/1999/xhtml">
    <head>
    .....
    .....
    </head>
```

→ Head: meta-information

```
    <body>
    ......
    ......
    </body>
</html>
```

→ Body: actual information / content
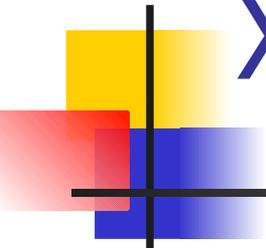
XHTML documents (3)

Some more information:

- The body of the hypertext is located between the `<body>` and `</body>` tags
- Spaces, end of lines (LF) and carriage return (CR):
  - Multiple spaces are equivalent to a single space (they are ignored except the first one)
  - Carriage returns (or line feeds) are completely ignored
- **Comments:** `<!-- this is a comment -->`

Title

```
<?xml version="1.0" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//
EN" "
http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">

  <head>
    <title> Put here the title of the document </title>
  </head>


  <body>

  </body>
</html>
```

The title that appears on the "title bar" of the web browser
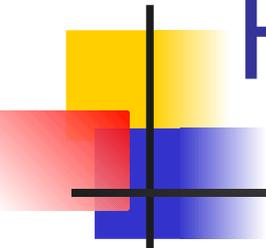
Headings

There are 6 different headings available: **they are meant for structuring documents**, not for presentation.

```
<h1> heading1 </h1>
<h2> heading2 </h2>
<h3> heading3 </h3>
<h4> heading4 </h4>
<h5> heading5 </h5>
<h6> heading6 </h6>
```

Chapter 1

Section 2.3

Sub-Section 2.3.1
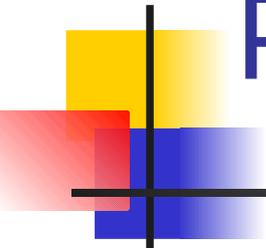
Paragraphs

`<p> Paragraph </p>`

- Indicates a text paragraph

`<br/>`

- Indicates a line break

`<hr/>`

- Indicates an horizontal rule

`<center>....</center>`

- Indicates that the content between tags shall be centered
- DO NOT USE IT: it is a presentation tag!!! (CSS)
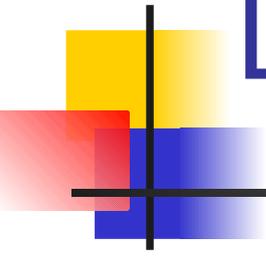
Lists

Lists can be ordered lists (with numbers or letters)

Or unordered lists (with points, squares and so on)

List delimiters define the beginning and the end of a list:

- `<ul></ul>` **or** `<ol></ol>`

List item delimiters define list items:

- `<li>...list item ...</li>`
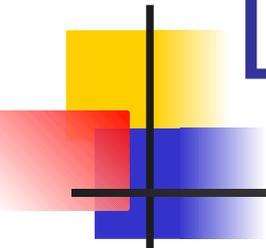
Lists (2)

Examples:

```
<ol>
    <li> 1st list item </li>
    <li> 2nd list item </li>
    <li> 3rd list item </li>
    <li> 4th list item </li>
</ol>
```

1. 1st list item
2. 2nd list item
3. 3rd list item
4. 4th list item

```
<ul>
    <li> 1st list item </li>
    <li> 2nd list item </li>
    <li> 3rd list item </li>
    <li> 4th list item </li>
</ul>
```

- 1st list item
- 2nd list item
- 3rd list item
- 4th list item
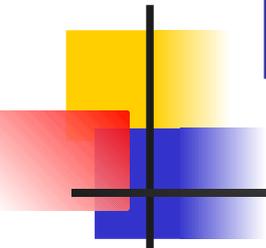
Lists (3)

Definition List

```
<dl>
    <dt> Term 1 </dt>
    <dd> definition 1 </dd>
    <dt> Term 2 </dt>
    <dd> definition 2 </dd>
    ...................
</dl>
```

Term 1
    definition 1
Term2
    definition2

It is possible to nest lists into other lists...
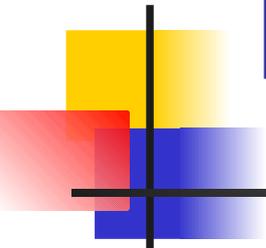
Lists (4)

It is possible to change the style of ordered lists (AVOID IT there are CSS)

```
<li type="1">    1.   First item
<li type="A">    A.   First item
<li type="a">    a.   First item
<li type="I">    I.   First item
<li type="i">    i.   First item
```

It also possible to change the starting number

```
<ol start="3">
```
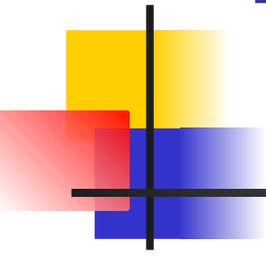
Text Formatting

Logic styles or Presentation styles can be used (logic style are OK, presentation styles KO)

A single portion of a text can be in fact classified by:

- Its role in the document: logic style
- The external aspect it should have: presentation style (or physical style)

Physical styles (Forget Them!!)

```
<b> Bold </b>
<i> Italic  </i>
<tt> TType - fixed spacing font </tt>
```

W3C strongly discourages the use use of these tags and recommends to use CSS styles (also they are not valid in XHTML strict)

Logic styles (good ones)

```
<dfn> definition </dfn>
<cite> citation </cite>
<code> code (programs..) </code>
<em> emphasis </em>
<kbd> keyboard-like printing </kbd>
<samp> example </samp>
<strong> strong </strong>
<var> variable </var>
```
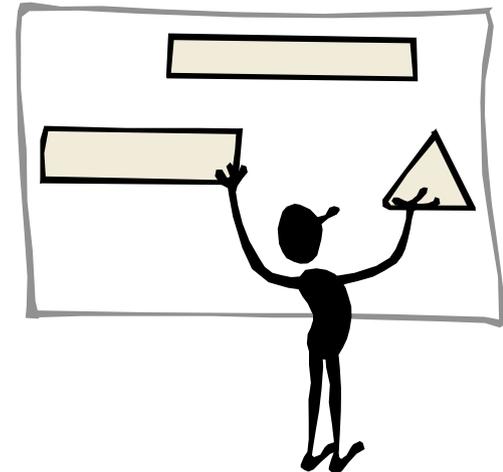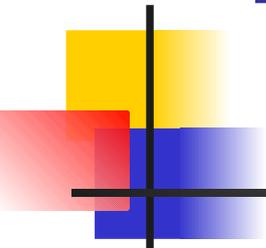
Text block formatting

```
<address> ... </address>
```

- For printing addresses

```
<blockquote> ... </blockquote>
```

- Citations on many lines

```
<pre> ... </pre>
```

- pre-formatted text
- Usually adopted for showing program codes
- Spaces and line breaks are preserved
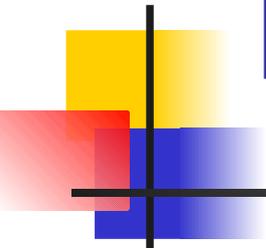
Entities

- Entities as well as tags are case sensitive

Result	Entity
<	<
>	>
&	&
"	"
è	è
é	é
ò	ò
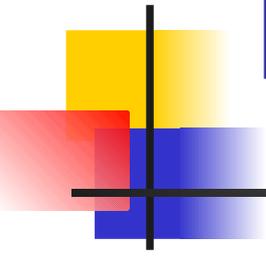
Links

Links, technically known as "anchors" allow to link different pages of a hypertext.

Hypertext navigation is supported by links

A link (Hyperlink) can be created using a <a> tag

The <a> tag requires the attribute href to indicate the destination of the link (i.e., the page that will be loaded when the anchor is clicked by the user)
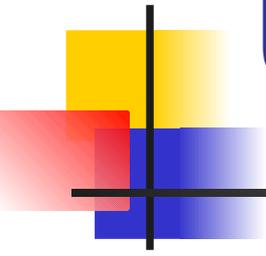
URLs

URLs can be either absolute or relative

Absolute

- href="http://elite.polito.it/courses/01gor"

Relative

- href="index.html"
- href="../index.html"
- href="../courses/images/logo.gif"

Links

**Link without specific target points**

Title

Address

**Title doc01**

Address

Title doc02

Address
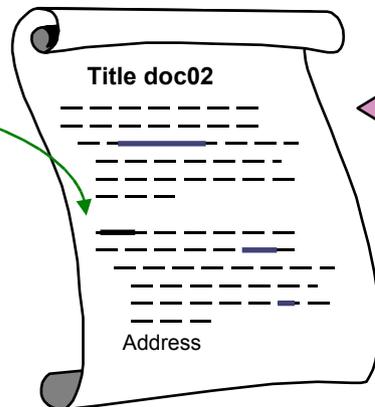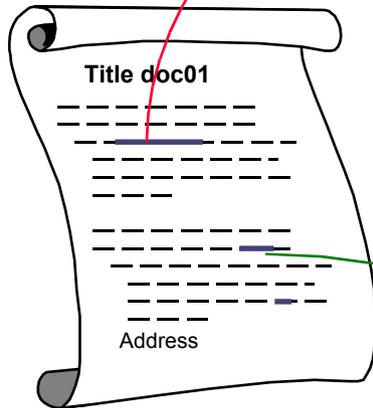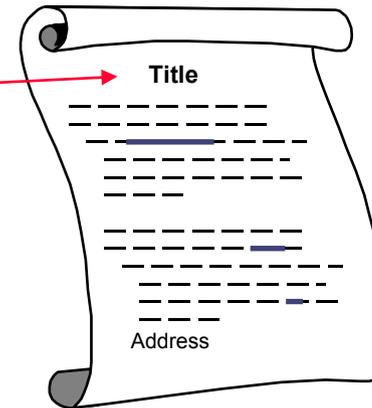
Link to a specific target point in a page

Links (2)

Link without specific target points

Title

Address

doc5.html

Title doc01

Address

`<a href= "doc5.html"> go to the next document </a>`
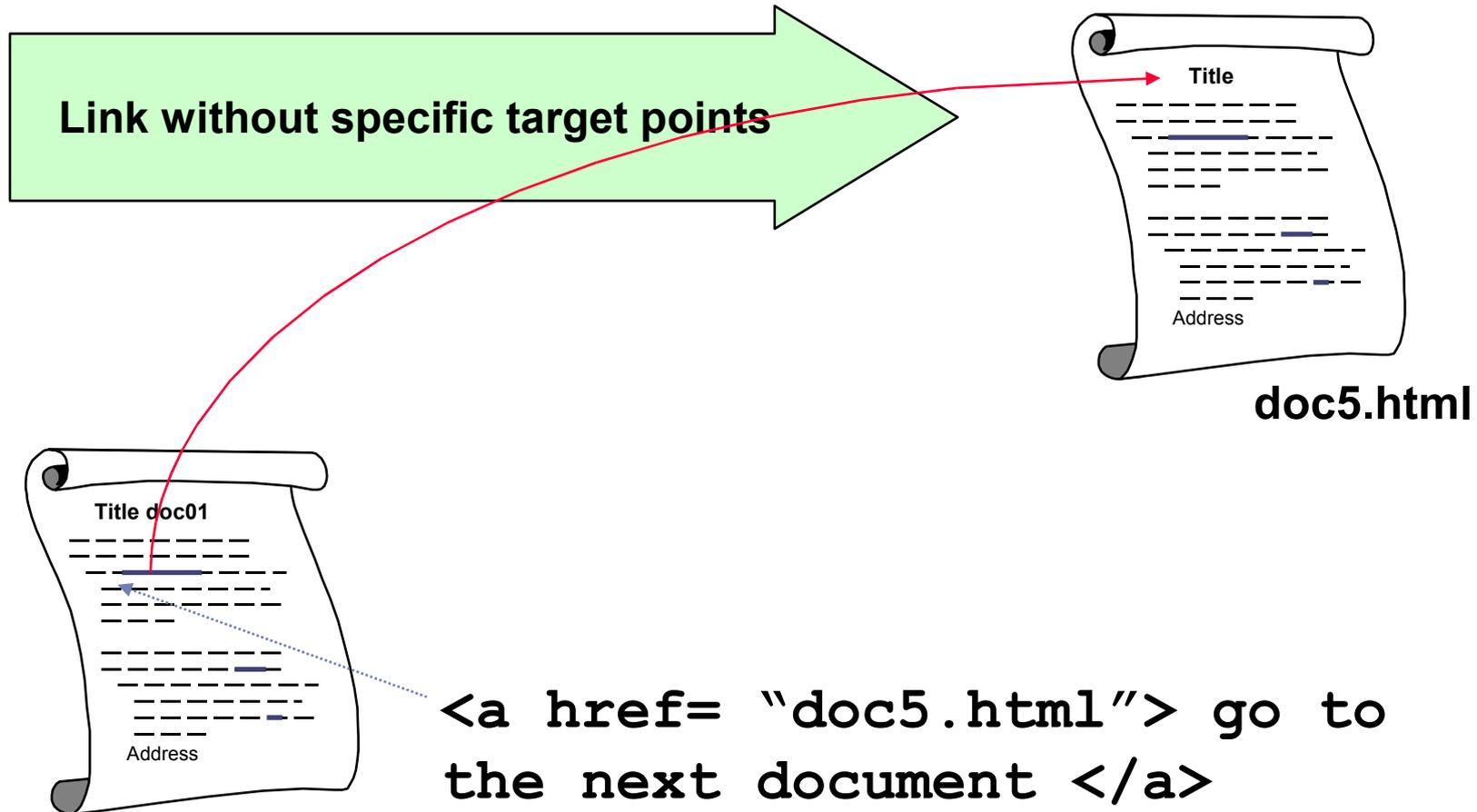
Links (3)

`<a href="doc5.html#point3">go to the`
`point3 in the document</a>`

`<a id="point3">`

Title doc01

Address

Title doc02

Address

doc5.html

Link to a specific target point in a page
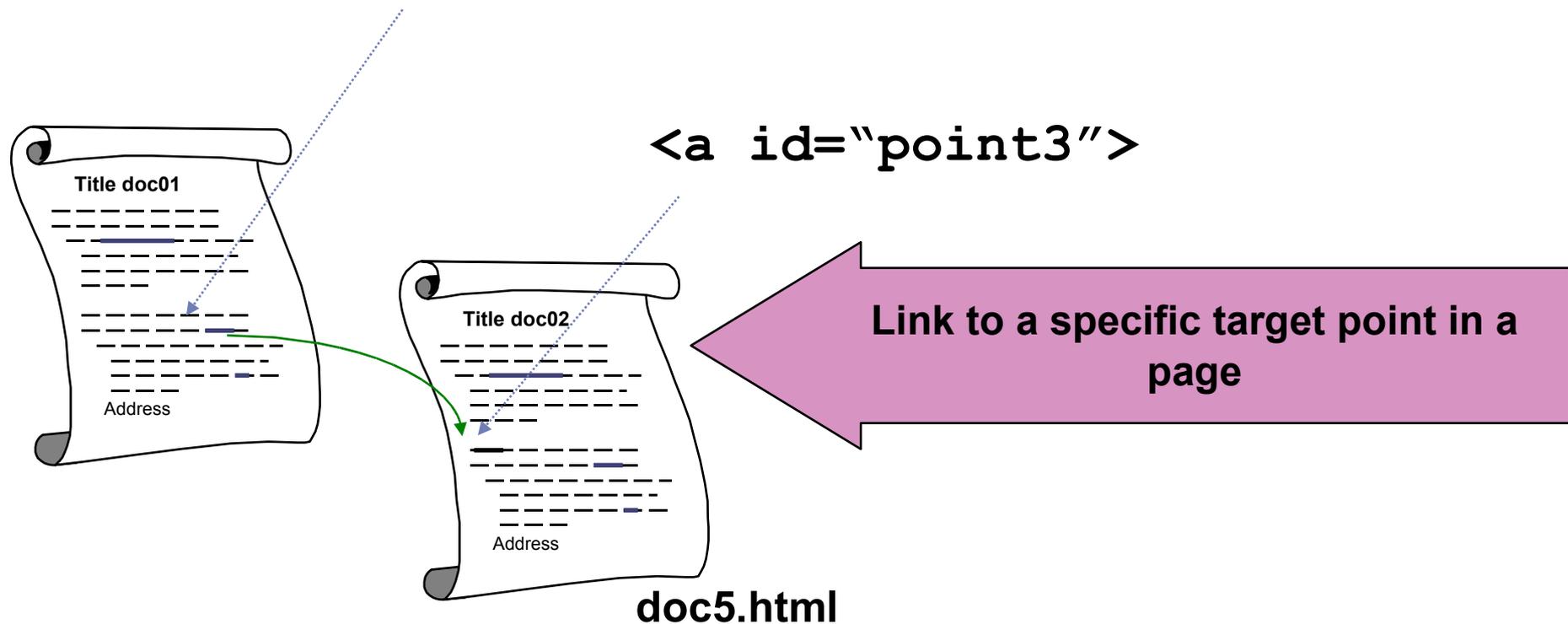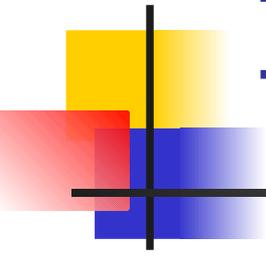
Images

Images can be of any format, the ones currently supported by all browsers are

- gif, jpeg and png
- `<img src="img.gif" />` Inserts the image "img.gif" into the web page

MUST NOT BE CONFUSED WITH

`<a href="img.gif">`, which creates a link to a new page that displays only the image named img.gif

Image attributes

```
<img src="img.gif" alt="picture"
title="my picture"/>
```

- Inserts into the web page the image img.gif
- If the browser does not support graphics "picture" is visualized instead.
- When the user interacts with the image (for example by moving the mouse on it) the title "my picture" is shown (IE uses ALT if title not available:NOK Firefox displays nothing: OK)

18/02/08
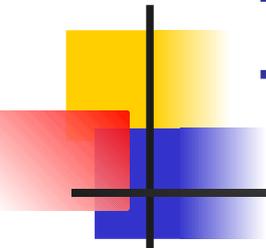
Images (2)

In HTML (and thus in XHTML Transitional, but not in XHTML strict)

- Images can be inserted in different positions with respect to the text usign the align attribute
  - `<img align="left" src="immagine.gif" />`
  - `<img align="right" src="immagine.gif" />`

In XHTML, the same and more is obtained through CSS

18/02/08
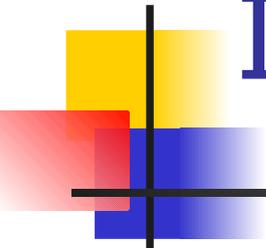
Images (3)

The attributes width and height allow to modify / change the size of an image

- `<img width="120" height="80" />`
- Can aid the browser to allocate the right space for the image before starting to load the image itself.
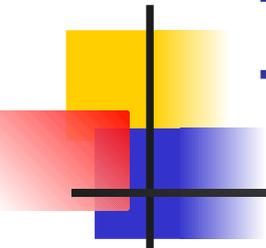
Images

The border attribute causes the browser to draw a border around the image

- `<img border="4" src="xxx.gif" />`

- `<img border="0" src="xxx.gif" />`

It is useful to eliminate image borders, especially when images are also links or when images lay side by side.

Remember CSS do this and more so, hold on and do not use these attributes…
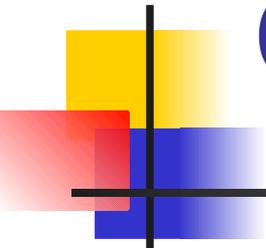
Clickable images

An image can be used as an "hot word" for navigating a hypertext, this could be done by inserting the image into an anchor

- `<a href="url"><img src="file.gif"></a>`

Or, if needed, single portions of a image can be links to different pages or points in pages using "image maps"
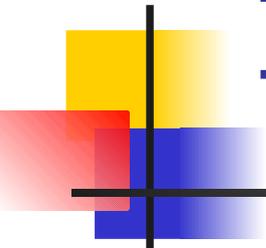
Image maps

```
<img src="img.gif" usemap="#img"/>
<map id="img">
<area shape="rect" coords="338,114,510,190"
        href="eolo.htm"/>
<area shape ="circle" coords ="260,110,50"
        href="pisolo.htm" />
    <area shape ="default" nohref="nohref"/>
</map>
```

Tables

```
<table border="1">
  <tr>
    <th>a</th>
    <th>b</th>
    <th>c</th>
  </tr>
  <tr>
    <td>d</td>
    <td>e</td>
    <td>f</td>
  </tr>
</table>
```

a	b	c
d	e	f
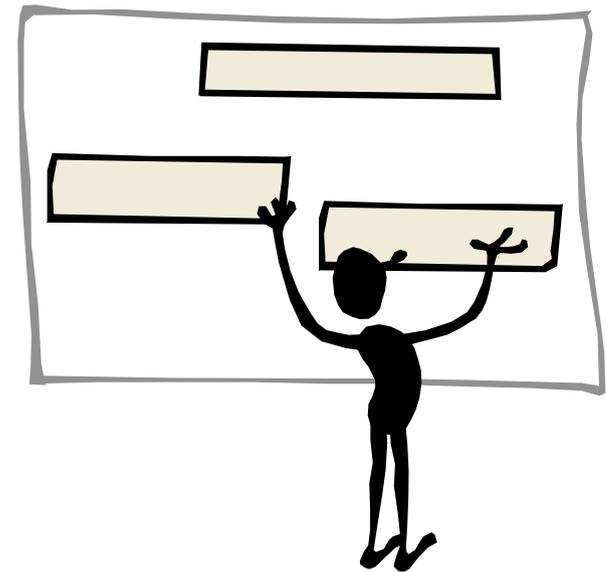
Forms

XHTML and HTML allows to create form for gathering data but do not allow data elaboration

- There is a so-called application server that is charged to perform the required elaboration

- Techniques:
  - CGI (Common Gateway Inteface)
  - Server Side Scripting (ASP, PHP, JSP)
  - Java Servlet

PERSONAL ZONE

user:

pass:

@ libero vai

» MailBox da 15MB Gratis!
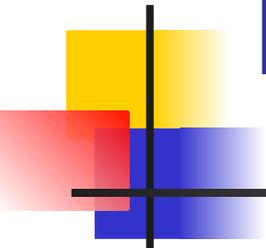
Forms

```
<form method="GET"
 action="/cgi-bin/example.cgi">
   input elements
</form>
```

METHOD: protocol adopted for passing
parameters to the server-side
program GET (default) or POST

ACTION: URI (Uniform Resource
Identifier) of the CGI program
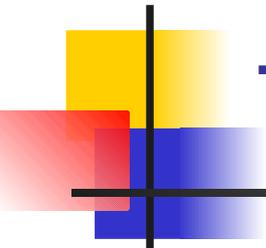
Input elements

```
<input type="text" name="sign"
 value="Name Surname" size=40 />

<input type="checkbox" name="checkbox”
 checked="checked” />

<input type="submit" value="Send" />

<input type="hidden" name="secret" />

<input type="password" name="password"
 value="keyword" />

<input type="image" name="face"
 src="faccia.gif" />
```
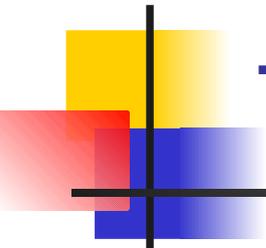
Input elements

```
<input type="radio" name="zoom"
value="2" checked="checked" />x2
<input type="radio" name="zoom"
value="4" />x4
<input type="radio" name="zoom"
value="6" />x6

<input type="reset" />

<select name="fruit">
  <option>Strawberry </option>
  <option selected="selected">Arancia</option>
  <option>Cilegia</option>
</select>
```

Input elements

Nome Cognome

☑ Invia *******

◉ x2 ○ x4 ○ x6

Reset Arancia ▾

Fragola
Arancia
Cilegia
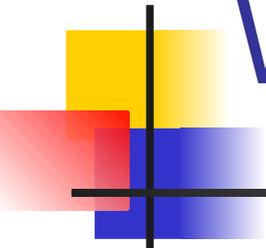
What should not be done

The following 5 slides show how, in HTML, the page content can be formatted for a good look and feel

In XHTML Transitional this is STRONGLY discouraged and in XHTML strict this is PROHIBITED

CSS allow to the same things and much, much, more.....
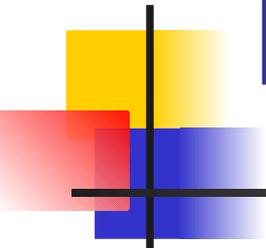
Background

It is possible to set up the background color of pages

- `<body bgcolor="#00FFA5">`
- The number following the #(sharp) sign encodes the color in RGB (8bit RED, 8bit GREEN, 8bit BLUE), and it is expressed as an hexadecimal code
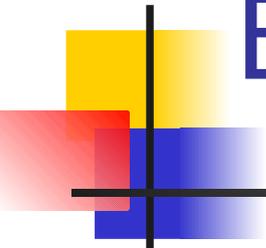
Background

It is also possible to use an image as page background

- `<body background="img.gif">`

- However, if the image is too small, it will be repeated until it will cover all the page surface

- CSS allow, for example, to disable the repetition and to center the background image with respect to the page
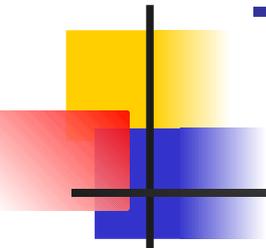
Text Color

It is possible, in HTML, to change the color of all text elements and of all links

- ```
  <body bgcolor="#00FFA5" text=
  "#CCCCCC"
  ling= "#0000DD" alink= "#FF0000"
  vlink= "#000033">
  ```

- text: text color

- link: not visited links color

- alink: activated links color

- vlink: visited links color
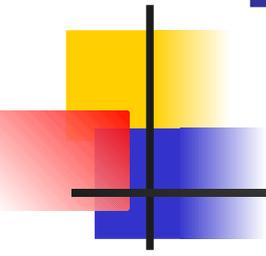
Text color (2)

Usability guidelines suggest to not change colors in order to not disorientate users

At least, XHTML imposes that CSS shall be used instead of these attributes

18/02/08
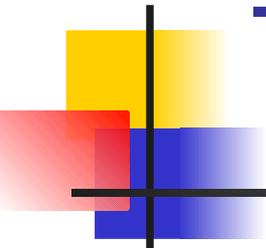
Text Attributes

It is possible to change the color of a set of words

- `<font color="#FF00FF"> I've changed the text color </font>`

It is also possible to change the Font of a set of words

- `<font face="Arial"> I've changed font</font>`

Remember!!! not all fonts are installed on every machine (Comic for example is usually not available on Linux machines)

Links...

Web sites

- http://www.w3schools.com/html
- http://www.w3c.org
- http://www.w3.org/TR/xhtml1/
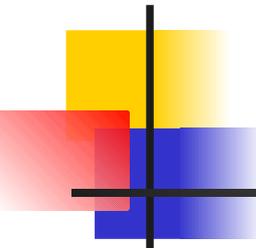
Books

- Elizabeth Castro, "HTML 4 for the World Wide Web", 4th edition, Peachpit Press, 2000
- HTML and XHTML Pocket Reference (Pocket Reference (O'Reilly)) by Jennifer Niederst Robbins (Paperback - May 8, 2006)
- CSS Cookbook by Christopher Schmitt (Paperback - Oct 1, 2006)