Links and web resources

Chapter 2

Outline

- A. Web architecture
  - Components of a web page: HTML, CSS, images, …
  - URL, MIME type and HTTP
- B. Links and images
  - Relative and absolute URLs
  - <a>, style, pseudo-class, :hover
  - <img> and background-image: url()
- C. Organization of a web site
  - separated CSS style sheet

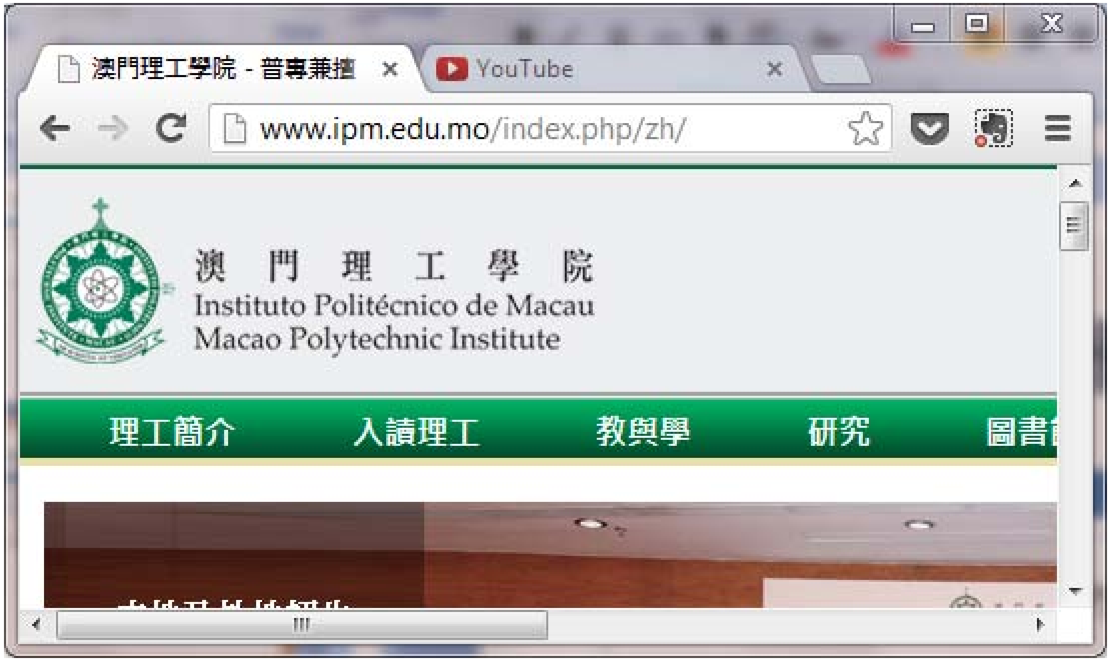
A. Web architecture

- Components of web pages
  - HTML, CSS, JavaScript
- Web architecture
  - Web resources : MIME types and URL
  - Standard protocol HTTP
- Basic HTTP interaction

Browsing the Web

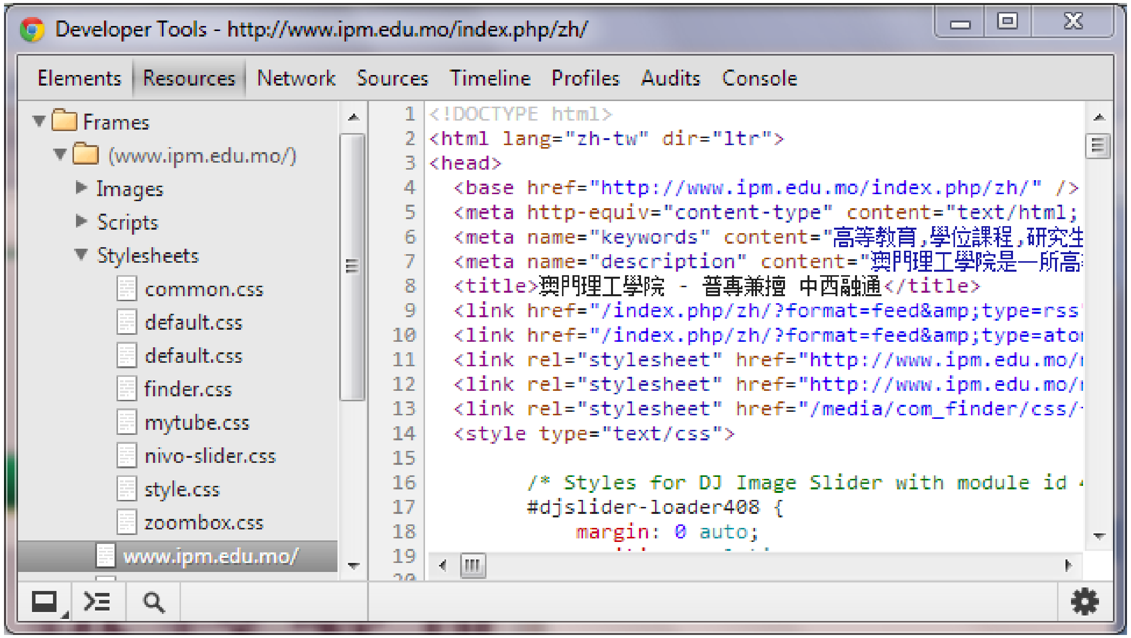
- □ Open a browser window
- □ Enter a URL in the location bar, then a web page is shown
- □ Click a hyperlink to go to another web page

1 URL = one web page?


澳門理工學院 - 普專兼擅
YouTube
www.ipm.edu.mo/index.php/zh/
澳 門 理 工 學 院
Instituto Politécnico de Macau
Macao Polytechnic Institute
理工簡介
入讀理工
教與學
研究
圖書

Components of a web page

- A web page actually consists of several kinds of resources
  - One or more HTML document
  - Some CSS style sheets
  - Some JavaScript programs
  - Images (png, jpeg, gif)
  - Objects (e.g. flash movies)
  - Font
  - … etc



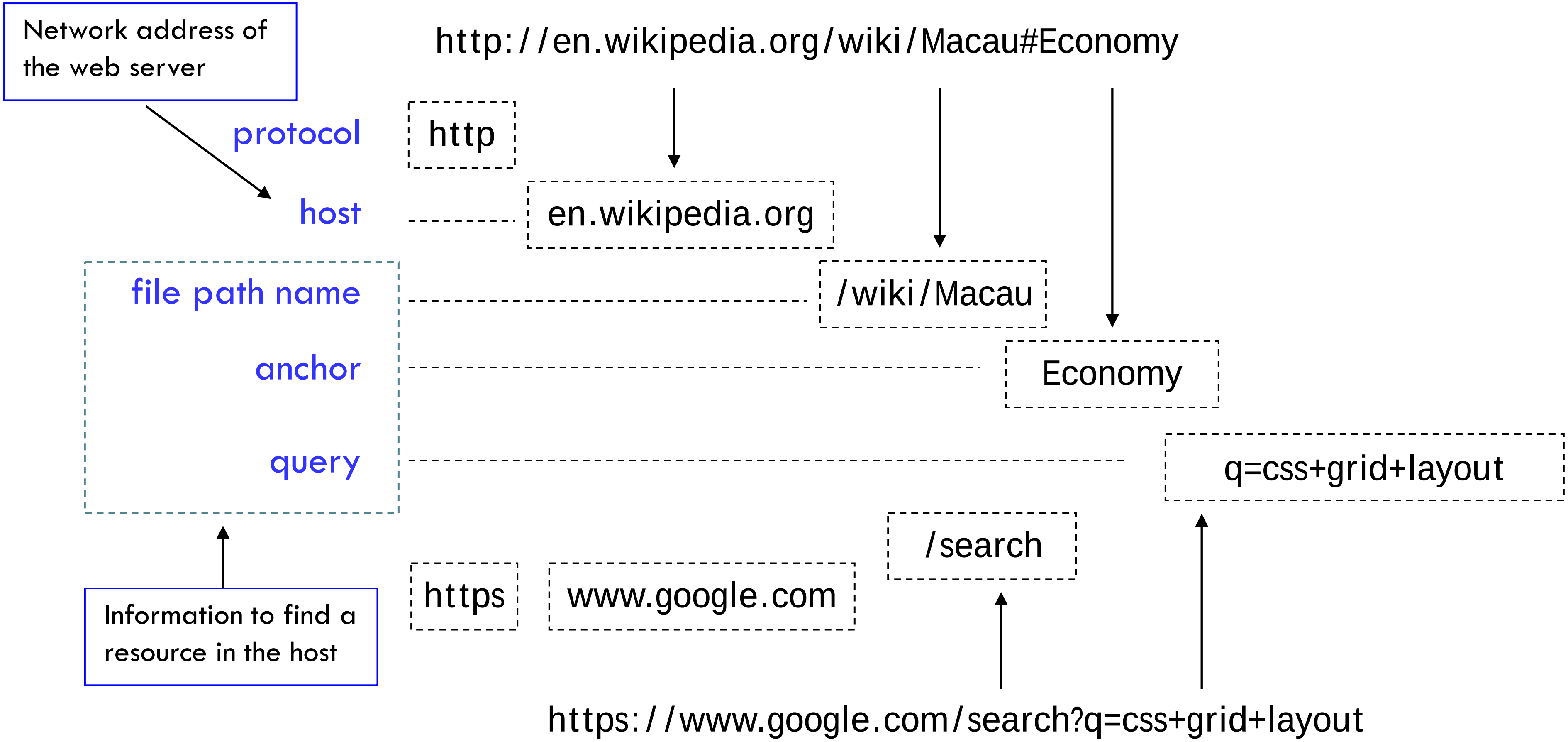
```
<!DOCTYPE html>
<html lang="zh-tw" dir="ltr">
<head>
  <base href="http://www.ipm.edu.mo/index.php/zh/" />
  <meta http-equiv="content-type" content="text/html;
  <meta name="keywords" content="高等教育,學位課程,研究生
  <meta name="description" content="澳門理工學院是一所高
  <title>澳門理工學院 - 普專兼擅 中西融通</title>
  <link href="/index.php/zh/?format=feed&amp;type=rss
  <link href="/index.php/zh/?format=feed&amp;type=ato
  <link rel="stylesheet" href="http://www.ipm.edu.mo/
  <link rel="stylesheet" href="http://www.ipm.edu.mo/
  <link rel="stylesheet" href="/media/com_finder/css/
  <style type="text/css">

        /* Styles for DJ Image Slider with module id
        #djslider-loader408 {
            margin: 0 auto;
```

How a browser obtains these resources?

- Questions:
  - Where to get a resource? (use URL)
  - How to talk with the servers? (use URL)
  - How to display? (use MIME type)
    What is the data type? Is it a png image? Or an html file?

Inside an URL …



Network address of the web server

Information to find a resource in the host

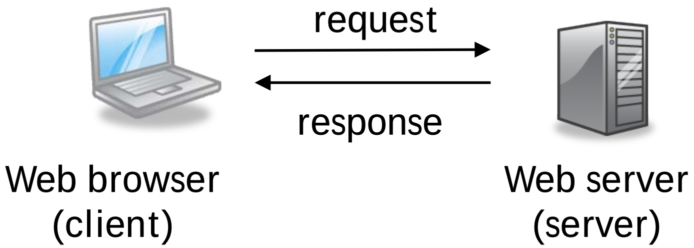
Syntax of URL

protocol://host/filepathname?query#anchor

- Each web resource has a **URL** (Universal Resource Locator).
  - Protocol is usually ‘http’ or ‘https’. https encrypts messages for better security
  - Host – network address of the web server
  - File path name (optional) – locates a resource in the host. Default is /
  - Query (optional) - submits data in the format *name1=value1&name2=value2*
  - Anchor (optional) – refers to an element with the id in an HTML file

Browsers, servers and resources

- Given an URL, a web browser has enough information to retrieve the resource.
- The browser uses the HTTP protocol to communicate with the web server
  - A browser sends a request for the resource
  - After some processing, the server returns a response which contains the requested resource.
- Each request retrieves one resource only
- The type of each resource is specified by a MIME type



Data types of web resource

- There are various types of web resources,
  - text, image, audio, video, data
- There are different formats to encode a certain type
  - an image can be in GIF, JPEG or PNG
  - a text file may be a HTML document, or a CSS file, or a plain text file
  - a video file may be in mp4, ogg, or webm
- These formats are identified by a standard called **MIME type**
- Most browsers support these formats, and can display them correctly.

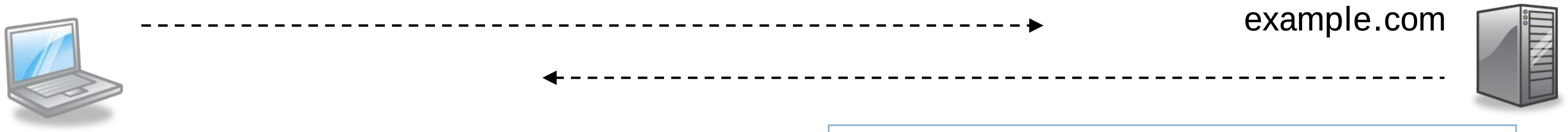
Common MIME types

category	MIME type	format
Text	text/html	HTML4.01, XHTML (.html), HTML5
	text/css	CSS (.css)
	text/plain	Plain text (.txt)
Images	image/gif	GIF (.gif)
	image/jpeg	JPEG (.jpg)
	image/png	PNG (.png) Portable Network Graphics
	image/svg+xml	SVG (.svg) Scalable Vector Graphics
Audio	audio/mp3, audio/m4a	Mp3 audio, MPEG4 audio
Video	video/mp4, video/m4v	MPEG4 video
Multipurpose	application/xml	General XML
	application/javascript	Javascript source (.js)
	application/x-shockwave-flash	Flash movie (.swf)
	application/font-woff	Web open font format
	application/octet-stream	Arbitrary binary data

HTTP message content

HTTP request

```
GET /home.html HTTP/1.1
Host: example.com
User-Agent: Mozilla/5.0 … Firefox/5.0
```



Simple HTTP retrieval of the page http://example.com/home.html. Notice how the URL of the resource is specified in the request, and how the data type is stated in the response. The server also specifies the character encoding for text resources.

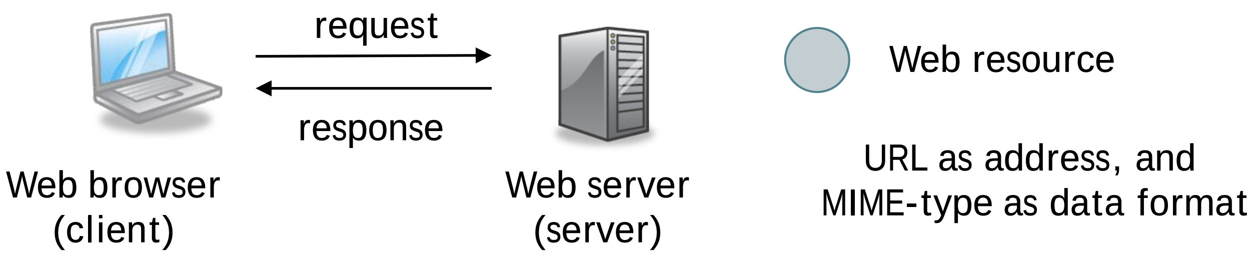
```
HTTP/1.1 200 OK
Date: Wed, 20 Jul 2011 17:52:00 GMT
Server: Apache
Content-Type: text/html; charset=utf-8

<html>… </html>
```

HTTP response

Basic retrieval

- A basic usage of HTTP is to retrieve a static file in server
  - A browser sends a request to GET a resource at a URL
  - The server receives the request, maps the URL to a file in its file system, e.g.
    http://example.com/home.html -> c:\inetpub\home.html
  - The server infers the MIME-type from the file extension, e.g. *.html -> text/html, *.png -> image/png



request

response

Web browser
(client)

Web server
(server)

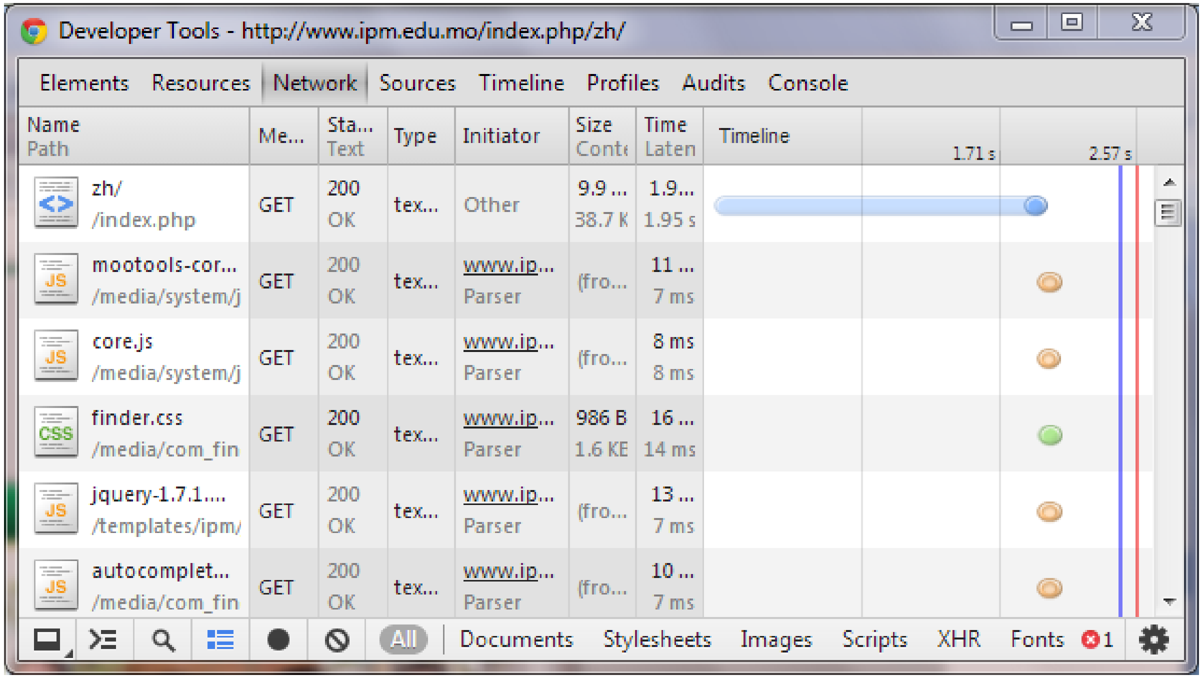
Web resource

URL as address, and
MIME-type as data format

Error handling

- A status code on the first line of a response
  - 200 – successful operation
  - 404 – page not found. The resource at the URL does not exist
  - 500 – server side error, probably a runtime error of server-side scripts
  - Many others…

Experiment



Name Path	Me...	Sta... Text	Type	Initiator	Size Conte	Time Laten	Timeline 1.71 s 2.57 s
zh/ /index.php	GET	200 OK	tex...	Other	9.9 ... 38.7 K	1.9... 1.95 s	
mootools-cor... /media/system/j	GET	200 OK	tex...	www.ip... Parser	(fro...	11 ... 7 ms	
core.js /media/system/j	GET	200 OK	tex...	www.ip... Parser	(fro...	8 ms 8 ms	
finder.css /media/com_fin	GET	200 OK	tex...	www.ip... Parser	986 B 1.6 KE	16 ... 14 ms	
jquery-1.7.1.... /templates/ipm/	GET	200 OK	tex...	www.ip... Parser	(fro...	13 ... 7 ms	
autocomplet... /media/com_fin	GET	200 OK	tex...	www.ip... Parser	(fro...	10 ... 7 ms	

- Use the Network panel of Chrome 'Developer Tools' to inspect HTTP requests and responses
- Note that loading a typical web page involves numerous HTTP transactions
- Check the status code in case the URL cannot be found

B. Links and images

- Hyperlink <a>
  - Absolute and relative URLs
  - Formatting links, pseudo-class
- Images
  - As content: inline image <img>
  - Image formats
  - For presentation only: background-image

Use of URL in HTML

- An HTML doc refers to another web resource with a URL
  - A hyperlink <a>
  - Media displayed inside HTML content: <img>, <object>, <video>
  - Background image specified in a CSS style sheet
  - External style sheet and JavaScript library for the current page
- The URL can either be absolute or relative

Hyperlink <a>

```
<a href="http://www.w3.org">WWW Consortium</a>
```

- The *required* attribute href is a URL (absolute / relative)
- The content of <a> can be text, image and some inline elements.
- When the user clicks the link, the browser goes to the referred URL.
- HTML5 allows putting block elements inside <a>
- But you cannot put an <a> inside another <a>

```
<a href=“a.html”>not allow<a href=“b.html”> to nest</a></a>
```

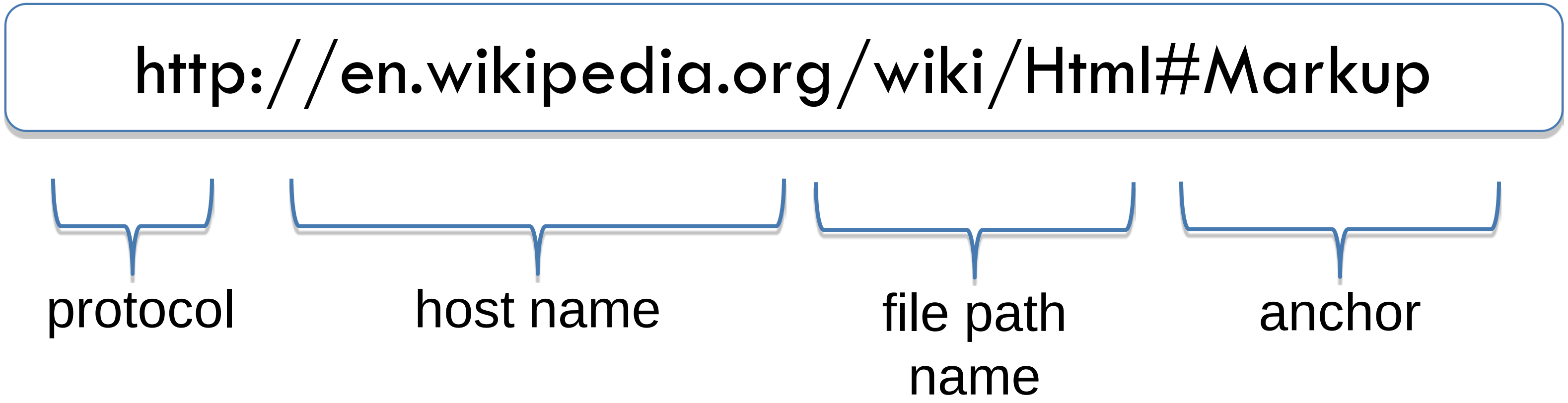
Relative URL

- If the target URL has the same prefix as the current HTML page, you can shorten the URL by removing the prefix. This produces a relative URL.

```
<!-- http://example.com/lab2-1.html -->
<html>…
<body>
  <h1>Lab 2-1</h1> <p>…</p>
  <p>When you finish, please proceed to
      <a href=“lab2-2.html”>next lab</a>.</p>
</body>
</html>
```

```
<!-- http://example.com/lab2-2.html -->
<html>…
<body>
  <h1>Lab 2-2</h1> <p>…</p>
</body>
</html>
```

Absolute URL



http://en.wikipedia.org/wiki/Html#Markup

protocol

host name

file path name

anchor

- An absolute URL must provide the protocol and host name
- If the file path name is absent, it is usually assumed to be '/'. (The front page of the site)
  - E.g. http://www.ipm.edu.mo -> http://www.ipm.edu.mo/

Relative URL

- When a web page points to another page in the same website, we may use relative URL
- Relative URLs omit the protocol, host name, and (maybe) a prefix of the file path name

```
<!-- this page is http://example.com/news/p1.html -->
<a href="p2.html">Page 2</a>
<a href="image/pic.png">a picture</a>
<a href="#top">top of this page</a>
<a href="/home.html">home</a>
```

http://example.com/news/**p2.html**
http://example.com/news/**image/pic.png**
http://example.com/news/p1.html**#top**
http://example.com**/home.html**

Anchor and the id attribute

- □ An anchor can reference the element with the specific id on the page.

```
<a href="http://example.com/a.html#Overview"> Overview
of HTML</a>
```

```
<!-- http://example.com/a.html -->
<h2 id="Overview">Overview</h2>
<p> … </p>
```

Formatting <a>

- You can use text formatting CSS properties to change the presentation of the hyperlink anchor text.
- In CSS, you can distinguish links that are visited or not. You can also change the appearance of links in hover (i.e. when the mouse cursor is over the links)
- CSS defines some pseudo-class for these states:
  - :link – a link (not useful)
  - :visited – a visited link
  - :active – a link when it is being clicked on (not useful)
  - :hover – a link when the cursor is held over it
    (:hover can also be used on other HTML elements)

The states of a link

- A link may be in several states at the same time, e.g.
  - When the cursor is over an unvisited link, the link is in both :link and :hover states
  - When you click a visited link, the link is in :link, :visited, :hover and :active states
- If there are conflicting property settings for these states, the last style rule wins.
- Rule of thumb: LoVe HAte

```
a:link { color: blue; }
a:visited { color: gray; }
a:hover { color: green; }
a:active { color: red; }
```

Example

```
a {
  text-decoration: none;
  font-weight: bold;     color: blue;
}
a:visited { color: gray; }
a:hover { text-decoration: underline; }
```

When the cursor is over a link, it is underlined. When the cursor is away, the link is not underlined.

Notice how CSS combines the styles for a and a:hover for links when the cursor is held over them. It works like …

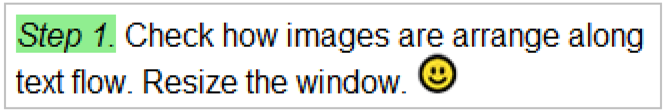
```
a:hover { text-decoration: underline;
    font-weight: bold; color: blue; }
```

Two ways to use images

- The image is part of the web page content
  - Use the HTML element
    `<img src="profile.png" alt="Peter Chan profile"/>`
- The image is for presentation only
  - Omitting it does not affect the meaning of the web page
  - Use background image property in CSS
    `p { background-image: url(bluebg.gif); }`

Image <img>

```
<img src="cat.jpg" alt="my cat" />
```



Step 1. Check how images are arrange along text flow. Resize the window. ☺

- <img> is an empty inline element
- The image file is downloaded and displayed inline with other HTML content
- The *required* attribute src is a URL to the image data file
- The *required* attribute alt provides a description of the image.
  - Browsers show it before the image is ready to display
  - A screen reader may also read it out

Example: p201.html

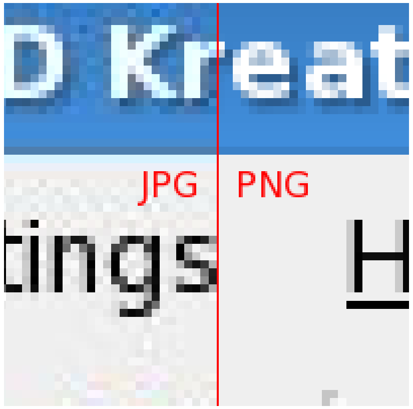
Width and height of <img>

```
<img src="cat.jpg" alt="my cat"
     width="300" height="200" />
```

- *Optional* attributes. The unit is pixel.
- Benefits: The browser can reserve enough space for the image. Layout of the page content will not change after downloading the image data
- You can also change the display size with CSS box model properties width and height
- If the intrinsic size differs from the specified size, the browser resizes the image.

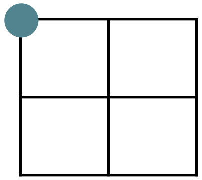
Common image formats

- GIF and PNG
  - Lossless compression, suitable for images with flat areas of plain color and text
  - GIF: 256 (8bit) colors, animated GIF
  - PNG: 48bit color, alpha transparency
  - http://en.wikipedia.org/wiki/Portable_Network_Graphics
- JPEG
  - lossy compression suitable for images such as photographs
- SVG
  - Vector graphics. Can be scaled up without any loss in quality



Background properties, 1/2

- You can set the background color and image of most HTML elements using these properties.
  - background-color:
    - Space not filled by bg image is filled with this color
    - The default is transparent
  - background-image: url( )
    - E.g background-image: url(images/bg.gif) use the image at the relative URL “images/bg.gif” as background image
  - background-position:
    - top, bottom, center, left, right, or a combination
    - Default is top left
    - Arbitrary position like 100px, 80px and 30% 30%



Background properties, 2/2

- background-repeat:
  - repeat – tile across the whole element
  - repeat-x – repeating on the x-axis, side by side
  - repeat-y – repeating on the y-axis, above and below
  - no-repeat – just show one instance of the image
- background-attachment:
  - scroll (default) – the background image will scroll with the page content
  - fixed – the background image will remain stationary when the page content scrolls

Example: p202.html

background property

- You can combine multiple background-related properties with this shorthand property

```
body {
  background: white url(images/bg.gif) top left fixed no-repeat
}
```

```
body {
  background-color: white;
  background-image: url(images/bg.gif);
  background-position: top left;
  background-attachment: fixed;
  background-repeat: no-repeat;
}
```

Using images as list bullets

- You can replace the bullets of a list with an image
  - list-style-type: none, disc, circle, square, decimal, upper-roman, lower-roman, upper-alpha, lower-alpha
  - list-style-image: url( ) uses images as bullets

Further readings

- CSS Gradient background
  - http://www.colorzilla.com/gradient-editor/
  - Reference: https://developer.mozilla.org/en-US/docs/Web/CSS/linear-gradient
- CSS3: Multiple background
  - http://snook.ca/archives/html_and_css/multiple-bg-css-gradients
  - http://www.css3.info/preview/multiple-backgrounds/
- border-image: uses images as borders
  - https://developer.mozilla.org/en-US/docs/Web/CSS/Tools/Border-image_generator
- CSS sprites: combining small images to reduce network traffic
  - http://www.alistapart.com/articles/sprites

Part C. CSS separation

- Relationship between HTML, CSS, JavaScript
- Separation of content, presentation and behavior
  - benefits
- How to link HTML to CSS

Three main Web standards

- Three main types of resources in a web page:
  - HTML document – a text file that contains text content and links to external resources through URLs
    - Examples: images, video, flash objects
  - CSS style sheet – a text file that defines the appearance of HTML content
  - JavaScript program – some programming code that defines the behavior of the web page

Hypertext Markup Language, HTML

- HTML defines the structure of a doc and the meaning of parts
  - e.g. a document contains several paragraphs, where each paragraph contains some text, images, and links.
  - e.g. <p>HTML is a <em>markup</em> language.</p>
- a W3C recommendation (http://www.w3.org/html/)
  - current version XHTML 1.0, previous version HTML 4.0
  - next version HTML 5

In this course, we'll use HTML5 as implemented in current versions of web browsers.

Cascading Style Sheet, CSS

- CSS describes how to display HTML elements
  - e.g. color of normal text, color of link, line spacing, position of text block, page break for printed copy
- a W3C recommendation (http://www.w3.org/Style/CSS/)
  - current version CSS 2.1
  - next version CSS 3
    - Many features widely implemented

In this course, we'll use CSS3 features as implemented in current versions of web browsers.

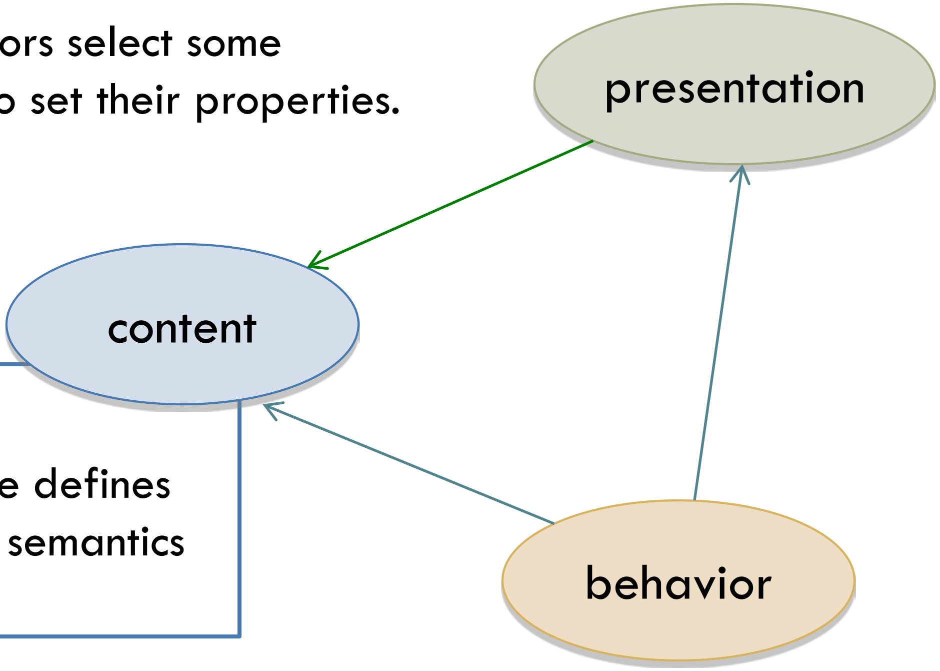
JavaScript

- JavaScript mainly runs in web browsers and adds dynamic behavior to web pages
  - The most widely used programming language on the Web
  - Standard ECMA-262, also known as ECMAScript
    - Common dialects include ActionScript (Adobe Flash)
  - looks like Java, but very different internally

We'll cover JavaScript in COMP312 Internet Programming II

Three aspects of web pages

CSS selectors select some elements to set their properties.



CSS
Style sheets define the appearance and layout of the content.

HTML
Markup language defines the structure and semantics of data.

JavaScript
JavaScript code indicates actions when users interact with the page (e.g. click a button.)

JS code may change both the content and the presentation of the content.

Adding CSS and JS to HTML file

- HTML, CSS and JavaScript are three different languages with different syntax and concerns
- Two ways to add CSS and JS to an HTML doc
  - Embed CSS style rules and JavaScript in an HTML doc
    - Convenient, but not recommended for production use
  - Separate CSS and JavaScript code in external files, and add links in the HTML doc

Example: Embed CSS and JS in HTML

```
<!DOCTYPE html>
<html>
  <head>
    <title>My First</title>
    <style> body { font-family: arial; color: black; } </style>

    <script>
      window.onload = function () { alert("finish loading"); }
    </script>
  </head>
  <body>
    <p>This HTML doc embeds CSS style rules and JavaScript code.</p>
  </body>
</html>
```

Separation

- Modern web page design encourages separation of content, presentation and behavior.
  - Save CSS style rules in a separate style sheet (usually with extension .css)
  - Save JavaScript code in a separate file (usually with extension .js)
  - Link to these external resources from the HTML doc (usually with extension .html or .htm)

Example: link to external CSS and JS

```
<!DOCTYPE html>
<html>
  <head>
    <title>My First</title>
    <link rel="stylesheet" href="style.css"/>
    <script src="action.js"></script>
  </head>
  <body>
    <p>This HTML doc links to an external CSS style sheet
        and an external JavaScript program.</p>
  </body>
</html>
```

```
/* style.css */
body { font-family: Arial; color: black; }
```

```
/* action.js */
window.onload = function () {
  alert("finish loading");
}
```

Benefits

Benefits of separating style (CSS) from content (HTML)

- Consistent design
  - All pages in a site use the same style
- Centralized control of design
  - One place to find all style rules
  - Style of all pages can be changed by modifying a common style sheet
- Smaller file sizes
  - The common style sheet is transferred to a browser only once for a site