



中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY, CAS

Web编程介绍

What is HTML

- 超文本标记语言（HyperText Markup Language，简称：**HTML**）
 - 一种用于创建网页的标准标记语言
 - 标记语言是一套标记标签(tag)
 - HTML标签是由尖括号包围的关键词，如<html>
 - HTML文档 = 网页
 - 可以嵌入脚本语言，它们会影响HTML网页的行为
- 相关教程
 - <http://www.w3school.com.cn/html/index.asp>

General Structure

- HTML文档由head和body两部分组成
 - <html>和</html>之间的文本描述网页
 - head中<title>这是标题</title> 定义了浏览器的页面标题
 - <body>和</body>之间的文本为可视页面内容

```
<html>  
  <head>  
    <title>This is title</title>  
  </head>  
  <body>  
    <h1>My First Heading</h1>  
    <p>My first paragraph.</p>  
  </body>  
</html>
```

HTML Elements and Tags

- 每一个元素都包含
 - 开始标签 <tag>
 - 内容
 - 结束标签 </tag>
- 部分元素可以没有结束标签
 - <p> paragraph tag
- 有些元素不需要内容
 - <hr>: 创建水平线
 -
: 不产生新段落时换行

```
<html>
  <head>
    <title>This is title</title>
  </head>
  <body>
    <h1>My First Heading</h1>
    <p>My first paragraph.</p>
  </body>
</html>
```

Browsers Are Tolerant

- 浏览器忽略html文件中的错误
 - 忽视其无法理解的标记符号
- Internet Explorer/Chrome:
 - Html文件不要求以<html>和</html>开头和结尾
 - <head>和<body>不一定要必须出现
 - 但是会导致不确定的行为

Attributes

```
<a href=" http://www.w3school.com.cn/html/index.asp">  
This is a link  
</a>
```

- 元素中的属性(attributes)以参数形式进行传递
- 放置于开始标签中
- 赋值形式: name = value
- 属性和属性值对大小写不敏感 (标准中推荐小写)
- 属性值必须用引号包围

Comments in HTML

- Html的注释使用以下的形式:

`<!-- here is comment content -->`

- 注释不能嵌套

```
<!-- This is a comment -->  
<p>This is a paragraph.</p>  
<!-- Remember to add more information here -->
```

Table Elements

- `<table>` 标签定义表格
- 每个表格有若干行 (`<tr>` 标签定义)
- 每行有若干单元格 (`<td>` 标签定义)
 - 数据单元格可以包含文本、图片、列表、段落等
- 表头用 `<th>` 标签定义
- 相关资料
 - http://www.w3school.com.cn/html/html_tables.asp

Table Example1

两行三列:

```
<html>
```

```
<body>
```

```
<h4>两行三列: </h4>
```

```
<table border="1">
```

```
<tr><td>100</td><td>200</td><td>300</td></tr>
```

```
<tr><td>400</td><td>500</td><td>600</td></tr>
```

```
</table>
```

```
</body>
```

```
</html>
```

100	200	300
400	500	600

Table Example2

```
<html>  
<body>
```

```
<table border="1">
```

```
<th rowspan=3>row</th>
```

```
<th colspan=3>columns</th>
```

```
<tr><td>100</td><td>200</td><td>300</td></tr>
```

```
<tr><td>400</td><td>500</td><td>600</td></tr>
```

```
</table>
```

```
</body>
```

```
</html>
```

	columns		
row	100	200	300
	400	500	600

Anchor Elements

```
<a href="http://www.w3school.com.cn/html/index.asp">  
This is a link  
</a>
```

- **<a>标签定义**

- href属性指示链接目标

- **Uniform Resource Locator(URL): 统一资源定位器**

- 遵守以下语法规则

- scheme://host.domain:port/path/filename

- scheme - 因特网服务的类型 (如http)
 - host - 域主机 (如www)
 - domain - 因特网域名 (如 w3school.com.cn)
 - :port - 主机上的端口号 (http 的默认端口号是 80)
 - path - 服务器上的路径 (省略则文档必须位于网站的根目录中)
 - filename - 文档/资源的名称

Form Elements

- 搜集用户输入
- <form>元素定义表单
- 表单元素
 - 单/复选框<select>
 - 文本框、单选按钮、提交按钮<input>
 - ...
- 使用方法
 - 在HTML文件中新建form元素
 - 服务器端编写程序处理表单中的参数
- 相关资料
 - http://www.w3school.com.cn/html/html_forms.asp

```
<form>  
    <input type="text" name="value1">  
</form>
```

Form Attributes

- Action 属性

- 定义在提交表单时执行的动作
- 提交表单的通常做法是使用提交按钮

- Name 属性

- 每个输入字段必须设置一个 name 属性
- 处理程序（由action定义）通过name解析输入字段值

```
<form action="action_page.pl">  
  <input type="text" name="value1">  
</form>
```

Form Attributes

- Method 属性

- 规定在提交表单时所用的 HTTP 方法（GET 或 POST）

- GET

- 表单数据在页面地址栏中可见
 - 如搜索引擎
 - 适用情形：没有敏感信息，提交数据较少
 - 格式举例
 - `index.php?name=admin&age=44`

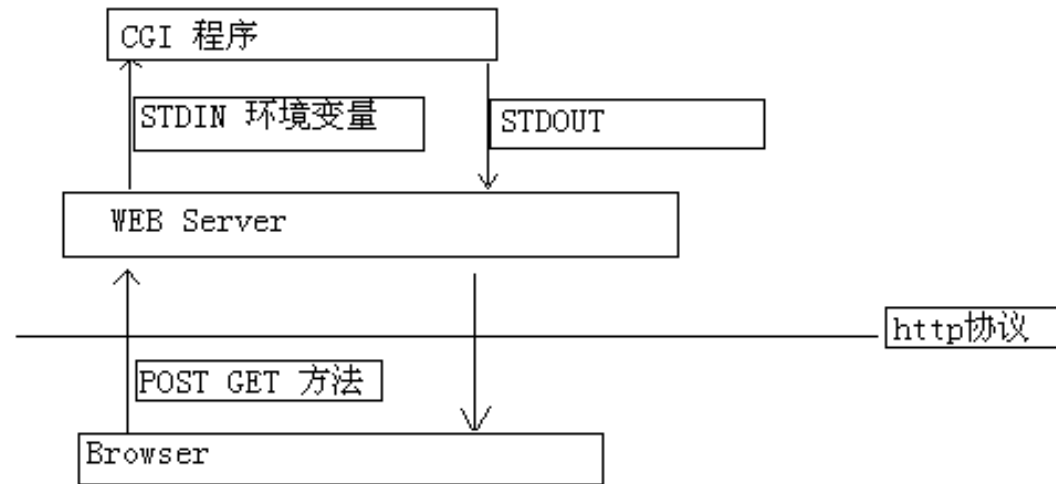
- POST

- url地址看不到请求参数，被提交的数据是不可见的

CGI

- 通用网关接口 Common Gateway Interface
- Web服务器运行时外部程序的规范
- 根据其编写程序，对客户端浏览器输入的数据进行处理，完成客户端与服务器的交互操作
- 创建动态的web文件
 - 脚本通常存储在cgi-bin文件夹中
- 相关资料

CGI Operation



CGI结构示意图

- (a) 浏览器发送访问请求到web服务器
- (b) web服务器调用CGI程序并传递参数
- (c) CGI 程序执行，从数据库操作后返回结果
- (d) 最终结果以HTML的形式返回浏览器

CGI Configuration

- /etc/apache2/conf-enabled/serve-cgi-bin.conf
- 相关资料
 - <https://httpd.apache.org/docs/2.4/howto/cgi.html>

```
<IfDefine ENABLE_USR_LIB_CGI_BIN>
    ScriptAlias /cgi-bin/ /var/www/cgi-bin/
    <Directory /var/www/cgi-bin/>
        Options Indexes FollowSymLinks MultiViews ExecCGI
        AllowOverride None
        Order allow,deny
        allow from all
        Require all granted
        AddHandler cgi-script .exe .pl .cgi
    </Directory>
</IfDefine>
```

CGI Script Input and Output

- Input

- ❑ GET method

- Data in URL

- ❑ POST method

- Data in a message body; CGI script reads from stdin

- Output

- ❑ Script sends output to stdout

Languages for CGI

- Perl
- Shell
- Python
- PHP
- C/C++
 - ❑ CGICC: C++链接库, 简化了CGI应用程序
 - ❑ <https://www.gnu.org/software/cgicc/>

CGICC example

- `<form method="get" action="/cgi-bin/cgi_get.cgi">`
- `/home/dbms/setup/code/cgi_get.cpp`

```
Cgicc formData;
```

```
... // 输出 html tags
```

```
form_iterator fvalue1 = formData.getElement("value1");  
if( !fvalue1->isEmpty() && fvalue1 != (*formData).end()) {  
    cout << "Value1: " << **fvalue1 << endl;  
}  
else  
    cout << "No text entered for value1" << endl;
```

What is PHP

- PHP 是一种通用开源脚本语言

- Personal Home Page -> PHP: Hypertext Preprocessor
- 在服务器上执行
- 将程序嵌入HTML文档中执行
- 生成动态页面内容，接收表单数据等

- 相关资料

- <https://www.w3school.com.cn/php/index.asp>
- <http://php.net/manual/zh/book.pgsql.php>
- PostgreSQL Demo
 - <http://php.net/manual/zh/pgsql.examples-basic.php>

Web编程练习

- 练习1: 将sql语句的结果输出成html文件

- ❑ select r_regionkey, r_name from region
- ❑ 参考.../setup/db-programming/libpq/myexample3.c
- ❑ 结果示例

r_regionkey	r_name
0	AFRICA
1	AMERICA
2	ASIA
3	EUROPE
4	MIDDLE EAST

Web编程练习

- 练习2: 实现一个cgi脚本，将form收集到的数据展示到网页中
 - CGICC脚本部分参考.../setup/code/cgi_get.cpp
 - HTML部分参考/var/www/html/cgi_get.html
 - 结果示例

Practice2 form

Origin Station :

Terminal Station :

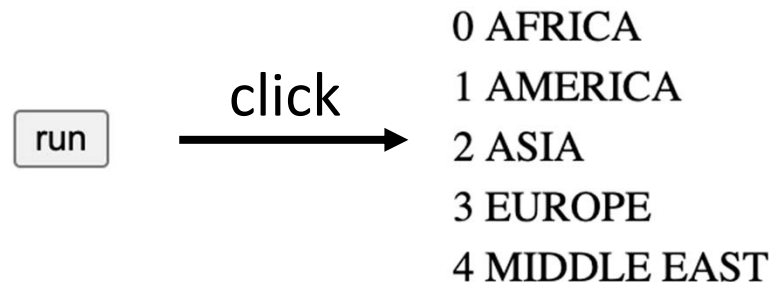


Yor are searching for route between Beijing and Tianjin

Web编程练习

- 练习3: 实现一个网页，点击按钮执行固定的sql语句，并将结果展示在html上

- select r_regionkey, r_name from region
- HTML部分通过action调用脚本
- 脚本部分有两种实现方法
 - libpq+CGICC
 - 练习1+练习2
 - PHP
 - 参考/var/www/html/db.php



谢谢！



中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES