

Outline

1 Internet Basics

- IP addresses and URLs
- client/server and HTTP
- HTML, XML, MathML

2 Python as Web Client

- the weather forecast
- CGI Programming

MCS 260 Lecture 18
Introduction to Computer Science
Jan Verschelde, 3 July 2023

networking and the internet

1 Internet Basics

- IP addresses and URLs
- client/server and HTTP
- HTML, XML, MathML

2 Python as Web Client

- the weather forecast
- CGI Programming

The World Wide Web

WWW: historical development

- 1970s: development of TCP/IP = Transmission Control Protocol/Internet Protocol.
Main benefit: capability of electronic mail (*email*).
- mid 1980s: connections between computer facilities.
In 1989, Tim Berners-Lee of CERN developed HTML, HyperText Markup Language.
- Mosaic was the first web browser developed at NCSA, released in 1993, leading to Netscape.
Search engines originated at the end of the nineties.
- Web 2.0: publishing → participation.
Wiki is server software that allows users to freely create and edit Web pages using any Web browser.

network addresses and domain names

Each node on the internet has an *IP address*.

An IP address consists of four bytes.

Each node has a symbolic name.

For example, `people.uic.edu`.

The `edu` stands for universities. The other principal domains are `gov` (government), `mil` (military), `com` (commercial), and `org` (other organizations).

The command `nslookup people.uic.edu`
or `host people.uic.edu` returns
the numerical IP address: 128.248.156.140.

The Internet is a *Wide Area Network* (WAN), linking machines over a greater distance. A *Local Area Network* (LAN), links computers in one room or building.

Uniform Resource Locator (URL)

A URL is an addressing scheme to provide a path to an internet resource.

Example: `http://www.math.uic.edu/~jan/mcs260.html`.

The format of a URL is

`protocol://host.domain-name/path/dataname`

where

- `protocol` refers to the type of protocol to be used
- `host` refers to the server where the resource is stored
- `domain-name` contains the name and type of the domain of the server
- `path/dataname` refers to the location of the data

To preview pages offline, use the protocol `file`.

networking and the internet

1 Internet Basics

- IP addresses and URLs
- **client/server and HTTP**
- HTML, XML, MathML

2 Python as Web Client

- the weather forecast
- CGI Programming

Client/Server Networks

A *client* is a computer in the network that *requests* for access to data and services from another computer.

A *server* is a computer in the network *receives and processes* requests from clients.

Access permissions are determined by the server.

A *client/server network* consists of several computers connected in a network, acting as clients and/or servers.

Client/server computing emerged in the nineties to distribute applications (such as database administration) over a network.

Hypertext Transfer Protocol (HTTP)

exchange data between the client and the server

HTTP is based on request-response between a web browser (the client) and a web server.

A typical transaction between browser and server:

- 1 A TCP/IP connection is established between browser and server.
- 2 The browser sends a request for a web page.
- 3 The server locates the file and responds, sending the content of the requested web page.
- 4 The TCP/IP connection is closed.

networking and the internet

1 Internet Basics

- IP addresses and URLs
- client/server and HTTP
- HTML, XML, MathML

2 Python as Web Client

- the weather forecast
- CGI Programming

Markup Languages

to write web pages

To write web pages, we use

HTML hypertext markup language
written to display information, the language in which web pages are written.

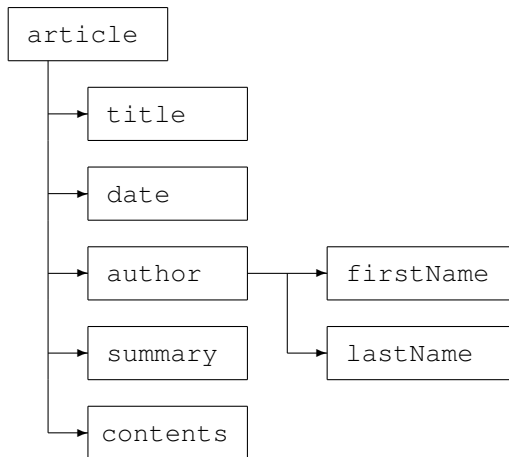
XML extensible markup language
XML is a widely supported open technology for describing data.

MathML mathematical markup language
MathML can display complex mathematical expressions. As it is created with XML, MathML is a so-called *XML vocabulary*.

The world wide web consortium (<http://www.w3c.org>) is a source for many protocol standards for the web.

Document Object Model (DOM)

The tree structure for `article.xml`:



networking and the internet

1 Internet Basics

- IP addresses and URLs
- client/server and HTTP
- HTML, XML, MathML

2 Python as Web Client

- **the weather forecast**
- CGI Programming

python forecast.py

```
$ python forecast.py  
opening http://tgftp.nws.noaa.gov/data/forecasts/state/il/ilz013.txt
```

Sun	Mon	Tue	Wed	Thu	Fri	Sat
Jul 02	Jul 03	Jul 04	Jul 05	Jul 06	Jul 07	Jul 08
Chicago Downtown						
Tstrms	Sunny	Sunny	Tstrms	Shwrs	Sunny	Sunny
66/72	63/80	69/90	72/84	65/73	62/72	62/75
70/70	20/20	20/20	30/70	70/50	40/10	10/10
Chicago O'Hare						
Tstrms	Sunny	Sunny	Tstrms	Shwrs	Sunny	Ptcldy
67/76	63/86	69/91	71/88	65/77	59/77	59/80
60/70	20/20	20/20	30/70	70/50	30/10	10/10

data from the web

The module `urllib` exports `urlopen`,
`urlopen` returns a file like object.

Template for retrieving data from web pages:

```
from urllib import urlopen
url = ' < internet address > '
f = urlopen(url)
s = f.readline()
```

networking and the internet

1 Internet Basics

- IP addresses and URLs
- client/server and HTTP
- HTML, XML, MathML

2 Python as Web Client

- the weather forecast
- CGI Programming

Python and CGI

Common Gateway Interface

Similar to a GUI (Graphical User Interface): run programs through a browser. Web interfaces are also *event driven*.

Python as scripting language for the web:
transform XML into XHTML pages.

The *Common Gateway Interface* (CGI) describes a set of protocols through which CGI programs interact with web servers and browsers.

CGI is common because it is not specific to any operating system or to any programming language or to any web server software.

We can use Python to generate dynamic web pages.

Exercises

- 1 What is the difference between `http` and `https`?
- 2 Point your browser to `https://people.uic.edu`.
Activate your site. Make your first web page.
- 3 Add a legend to the `forecast.py` script, using a dictionary to spell out the abbreviations `Tstrms`, `Ptcldy`, etc.
- 4 Write a Python script `downloadpyfile.py` which prompts a URL (for example the course web site) and the name of a `.py` file. The script downloads the content of the `.py` file from the web site into the `.py` file.