HTTP

1 HTTP

1.0.1 Possible responses from a web request

Responses include

- A plain text file
- A web page: some HTML
- An image file (jpeg, gif, png)
- A document (pdf, word)
- Some data (XML, JSON)
- A CSS file
- A javascript program
- A flash movie
- A redirection (in headers)
- A cookie value (in headers)
- An error
- A combination of the above

How might requests be generated?

1.0.2 Sources of requests

- URL typed in by user
- Hyperlink followed
- Form submitted
- Clicking in an image map
- Image included in source file
- Frameset or iframe in HTML source (can be recursive)
- Following a redirection (including 301 error)
- Javascript execution (triggered by mouseover etc)
- Flash execution (or other plug-in e.g. pdf)
- From a server (e.g. curl, robot, web service request)

Response to request may be used to update or replace some or all of a web page.

1.0.3 Hypertext Transfer Protocol (HTTP)

- Underlies many aspects of the web
- Based around sockets (usually port 80 for web pages)
- Fairly stable:
  - HTTP 0.9 (1991)
  - HTTP 1.0 (1996)
  - HTTP 1.1 (1997)
  - HTTP 2.0 (2015)
- Commonly accepted extensions: cookies
- HTTP 2 approved in 2015, including compression, push, pipelining and multiplexing</po
- For full details see `http://www.w3c.org/Protocols`
- For tutorial see `http://www.jmarshall.com/easy/http/`
- Some knowledge important for web apps
- Not just for HTML, but general resource (uRl)

1.0.4 Overview

- Client/Server: (usually) no response without request
- Requests and responses have similar format:
  - **Request/Status Line** including HTTP version and Status Codes for response
  - **Headers** including the host in HTTP 1.1, allowing for multiple sites on same IP
  - **Blank Line**
- Can run manually using telnet

1.0.5 telnet requests

At a Linux prompt:

```
telnet community.dur.ac.uk 80
GET /s.p.bradley/teaching/WP/lecture_http/ HTTP/1.1
Host: community.dur.ac.uk
```

Some sites require https (e.g. www.dur.ac.uk)

1.0.6 Request

- **GET** most common
- **POST** for some forms
- **HEAD** to check if a page exists
- **PUT** rarely used outside web services
- **DELETE** rarely used outside web services

Headers can include cookie values

1.0.7 Response

Response Codes

- **100-199** Informational (e.g. continue). Client should respond
- **200-299** Successful
- **300-399** File has moved (permanently or temporarily)
- **400-499** Client error (401 Unauthorised, 403 Forbidden, 404 Not Found)
- **500-599** Server error