

# Informatica Documentale

## Modulo II

Marino Miculan

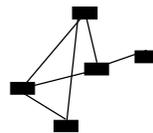
[miculan@dimi.uniud.it](mailto:miculan@dimi.uniud.it)

## Computazione distribuita

- collegamento in rete di calcolatori ed utenti, che permette la condivisione di risorse (programmi, dati, periferiche)
  - non più tramite un unico elaboratore centrale cui tutti sono collegati
  - ma attraverso la distribuzione delle risorse su calcolatori diversi, all'occorrenza accessibili da utenti diversi
- vantaggi:
  - flessibilità
  - efficienza
  - economicità
  - possibilità di estendere le reti a livello internazionale

## Alcune definizioni

- **Rete** di calcolatori (elaboratori):
  - insieme di calcolatori collegati tra di loro
  - è costituita da un insieme di **nodi** (calcolatori)
  - ed un insieme di **archi** (i collegamenti tra gli elaboratori)
- i collegamenti possono essere effettuati in vari modi, normalmente tramite **cavi** su cui "passano" i dati
- cavi diversi hanno caratteristiche diverse, da cui dipende soprattutto la velocità di trasmissione, detta **banda** (bit al secondo)



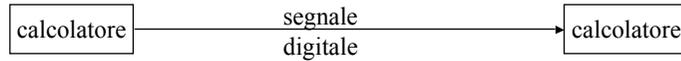
## Alcune categorie di reti

- caratteristiche della codifica:
  - trasmissione digitale/analogica
- continuità della connessione:
  - linea dedicata (leased line)/commutata
- contemporaneità della trasmissione:
  - trasmissione seriale/parallela/a pacchetti
- distanza tra i nodi
  - reti locali (LAN, Local Area Network)
  - reti metropolitane (MAN, Metropolitan Area Network)
  - reti geografiche (WAN, Wide Area Network)
- Forma (topologia) delle reti

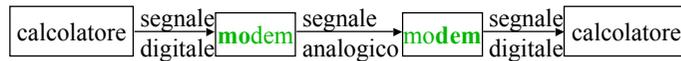
## Modalità di trasmissione

- le informazioni sono sempre memorizzate digitalmente;
- non sempre i mezzi di trasmissione permettono l'invio di segnali digitali, perché usano altre codifiche;
- quindi esistono due modalità di trasmissione:

– digitale



– analogica



InfoDoc 04-05 - Modulo II

5

## Esempi (trasm. seriale)

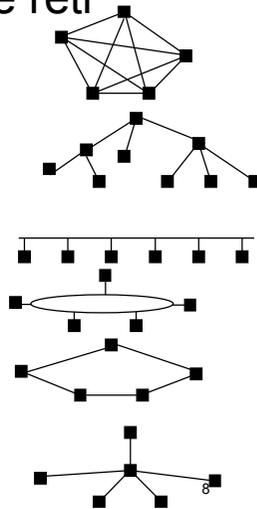
- linee telefoniche tradizionali:
  - trasmissione analogica
  - linea commutata
  - banda: fino a 57.6 Kb/s
- ISDN (Integrated Digital Number Network)
  - trasmissione digitale
  - linea commutata
  - banda: 128Kb/s (basic rate) -> 2Mb/s (primary)
- ADSL (Asymmetric Digital Subscriber Line)
  - Trasmissione analogica
  - Linea commutata
  - Banda: fino a 2Mb/s in download, 640kbit in upload
- linee dedicate
  - trasmissione analogica/digitale
  - linea dedicata
  - banda: 64 Kb/s -> 34Mb/s ed oltre

InfoDoc 04-05 - Modulo II

7

## Topologie delle reti

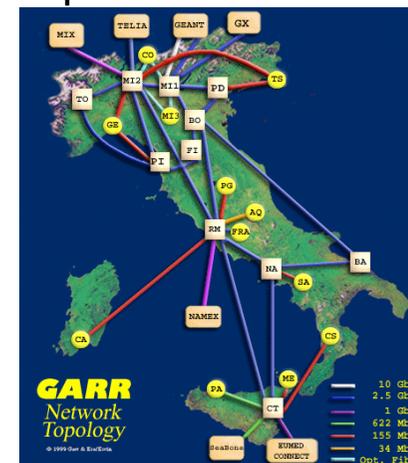
- **Punto-punto**: un collegamento per ogni coppia di nodi nella rete. Molto efficiente e ridondante, ma costosissima ( $N*(N-1)/2$  archi)
- Rete **gerarchica**: nodi collegati ad albero secondo una gerarchia. Basso costo, non molto efficiente
- Rete **lineare**: unico cavo (**dorsale**) cui sono collegati tutti i nodi. Affidabile, l'efficienza dipende dalla dorsale. Se richiusa alle estremità: rete **circolare**
- Rete ad **anello**: nodi collegati l'uno all'altro ad anello. Non molto veloce né affidabile.
- Rete **a stella**: nodo centrale collegato a tutti gli altri nodi. Molto veloce, non troppo costosa, il nodo centrale è il collo di bottiglia



InfoDoc 04-05 - Modulo II

8

## Esempio di WAN: il GARR



InfoDoc 04-05 - Modulo II

9

## Perché comunicare? E come?

- la comunicazione tra i nodi può riguardare lo scambio di dati e di servizi
- in generale, i nodi possono essere considerati come delle entità che si scambiano **messaggi**, composti da:
  - un **mittente**;
  - un **destinatario** (di solito non si hanno collegamenti punto-punto);
  - il **corpo** del messaggio, a sua volta costituito dai dati scambiati e/o dalla descrizione del servizio richiesto
- Esempio: sulla rete ci sono un calcolatore ed una stampante, ed il calcolatore richiede il servizio di stampa di una serie di dati via rete
- Affinché la comunicazione avvenga, bisogna accordarsi su come scambiarsi le informazioni tramite un **protocollo di comunicazione**

## Protocollo di comunicazione

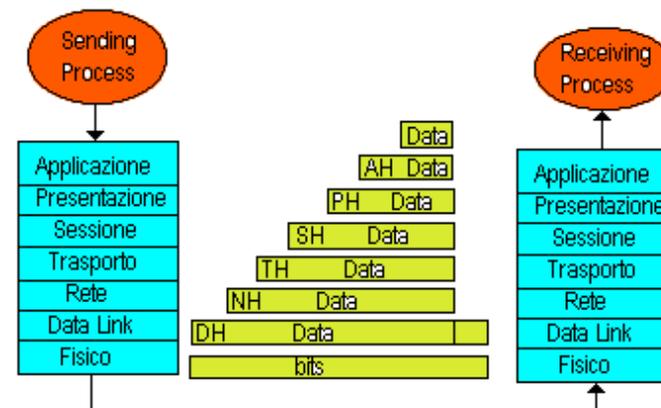
- insieme di regole che determina le operazioni da compiere per scambiarsi i messaggi:
  - come indirizzare i messaggi,
  - eventualmente attraverso quali nodi intermediari,
  - come codificare le informazioni che fanno parte del messaggio,
  - come codificare i servizi eventualmente richiesti,
  - come assicurare che le informazioni non vengano alterate,
  - come accertarsi che le informazioni siano arrivate,
  - ...
- ci sono molti protocolli diversi, e di solito sono organizzati a strati (come il S.O.)

## I protocolli di comunicazione

- poiché ci sono diversi livelli di astrazione nei problemi e nella loro gestione (dal controllo fisico dei cavi al servizio specificamente reso all'utente), di solito anche i protocolli di comunicazione sono stratificati, cioè con una struttura a cipolla come i S.O.
- due modelli stratificati principali:
  - **ISO/OSI**, standard nato nel 1978, molto completo ma poco usato
  - **TCP/IP**: standard de facto costituito dai protocolli di Internet

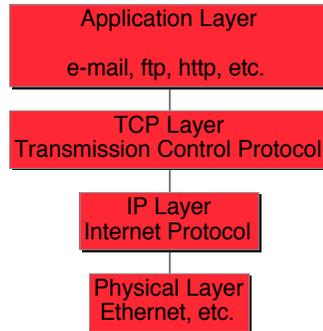
## ISO/OSI

- 7 strati di protocolli



# TCP/IP

Communication Layers



- † I protocolli applicativi definiscono modalità e formati per lo scambio di messaggi di ogni dimensione e scopo
- † TCP manipola flussi di dati orientati alla connessione, suddividendo i flussi in datagrammi, inviandoli tramite IP e ri assemblando i flussi dall'altro lato, ritrasmettendo dati persi.
- † IP è responsabile dell'invio di pacchetti di dati individuali (datagrammi) senza connessioni durature
- † interfaccia comune ai collegamenti fisici tra i vari nodi (per esempio, Ethernet, modem, etc)

InfoDoc 04-05 - Modulo II

14

# Principi di base

## •Identificazione dei nodi:

- ogni nodo in rete è identificato da un numero a 32 bit (**indirizzo IP**), per comodità scritto come i 4 valori dei byte che lo compongono, separati da punti
- all'indirizzo IP può essere associato un nome logico costituito da stringhe separate da punti. La tabella è mantenuta da nodi speciali detti **Name Server**

## •Identificazione degli utenti:

- nome.logico@computer.su.cui.risiede

## •Identificazione delle connessioni:

- un nodo può utilizzare più connessioni nello stesso tempo, ognuna identificata da una **porta** (numero a 16 bit)
- per comodità, alcune porte sono utilizzate sempre dagli stessi protocolli applicativi (**standard sockets**)

InfoDoc 04-05 - Modulo II

16

# Indirizzi

## • Esempio:

- indirizzo del computer principale del CdC: 158.110.1.2
- suo nome logico: hydrus.cc.uniud.it = *macchina specifica, al Centro di Calcolo dell'Università di Udine in Italia*

## • in questo modo, a livello logico si gestiscono le reti gerarchicamente

## • Alcune radici sono predefinite:

- .com, .edu, .org, .gov, .mil, .sigla di uno stato (.it, .fr,...)

## • Utenti: *paolino.paperino@hydrus.cc.uniud.it*, che da un punto di vista logico potrebbe essere anche *paolino.paperino@uniud.it*

## • alcuni utenti sono predefiniti (es *postmaster*)

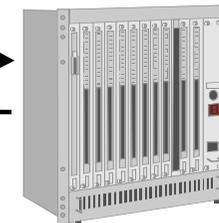
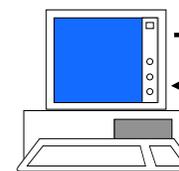
InfoDoc 04-05 - Modulo II

17

# Il paradigma client/server

## CLIENT

## SERVER



richieste di servizi, dati

servizi, dati

**protocollo di comunicazione**

## programma **client**:

- quando necessario, si connette al server (calcolatore) sulla porta specifica associata al server (programma)
- invia dei messaggi composti secondo il protocollo di comunicazione
- aspetta i risultati

## programma **server**:

- è sempre attivo (**daemon**)
- "ascolta" i messaggi in arrivo su una porta
- li interpreta (grazie al protocollo) ed effettua il servizio richiesto
- rispedisce indietro i risultati

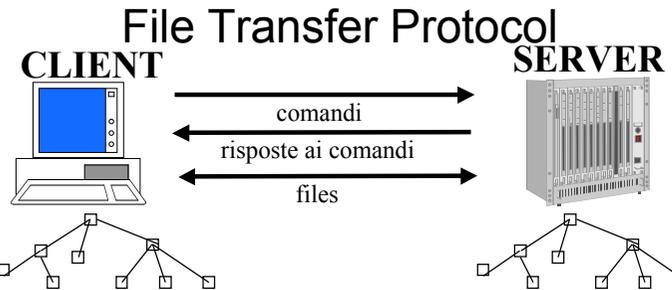
InfoDoc 04-05 - Modulo II

## I protocolli di Internet

- come accennato, sono “stratificati”
- i protocolli applicativi (ma non solo) rispettano tutti lo stesso principio client/server
- definiscono:
  - il formato dei messaggi scambiati
  - il significato di alcune parti del messaggio (es. richiesta servizi)
- sono descritti negli "RFC" (Request for Comments), liberamente disponibili su Internet

InfoDoc 04-05 - Modulo II

19



- storicamente, si accede ad un server ftp tramite una shell di comandi simile ad una shell Unix; si accede tramite autenticazione utente, e poi si hanno comandi usuali:  
**cd, pwd, del, mkdir, ls, ...**
- copia files sul sistema remoto: **put**
- copia files dal sistema remoto: **get**
- tre tipi di trasmissione: file di testo, binari, a record

InfoDoc 04-05 - Modulo II

28

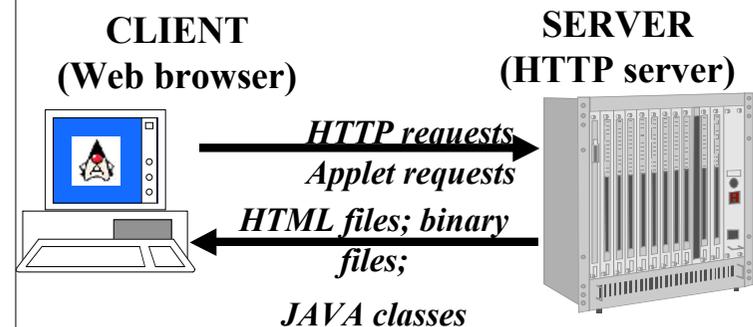
## Un altro protocollo applicativo di Internet: HTTP

- principale protocollo utilizzato nel World Wide Web
- utilizzato per la realizzazione di sistemi ipermediali distribuiti
- è essenzialmente un protocollo per il trasferimento di file
- porta standard: 80

InfoDoc 04-05 - Modulo II

30

## HTTP: client/server



InfoDoc 04-05 - Modulo II

31

## Il browser

- funge da interfaccia uniforme ad ogni servizio di rete
- gestisce i protocolli indicati negli schemi degli URL, autonomamente o tramite programmi esterni detti **helper**
- visualizza dinamicamente il testo ricevuto dai server secondo le istruzioni di formattazione specificate dai tag di HTML
- visualizza almeno una serie di formati multimediali standard, autonomamente o tramite helper
- interpreta le richieste dell'utente di selezione dei link, nascondendo ogni passo necessario alla selezione del protocollo, alla ricerca del server ed alla richiesta della risorsa
- Esempi: Mosaic, Netscape, Explorer, Lynx, Firefox, Safari, Konqueror...

## I server

- ogni calcolatore che funge da server ha in genere in funzione un programma per ogni protocollo servito (http, ftp, ...)
- le risorse distribuite da HTTP di solito consistono in
  - files (HTML, JPEG, GIF, MPEG, ...) presenti sul file system locale in appositi sottoalberi
  - a volte programmi di cui viene distribuito l'output (es. database, ed in generale CGI)
  - a volte programmi che vanno ad essere eseguiti sul calcolatore client (classi JAVA)

## HTTP

- HyperText Transfer Protocol
- è un protocollo molto semplice, senza stati e senza connessione (è però in cambiamento)
- i messaggi scambiati tra client e server seguono alcune delle convenzioni MIME
- i file trasferiti possono essere di qualsiasi tipo; comunque, un loro corretto trattamento prevede che il server conosca il tipo/sottotipo MIME da associare
- il formato comunemente trasferito è HTML, che permette la composizione di documenti ipertestuali

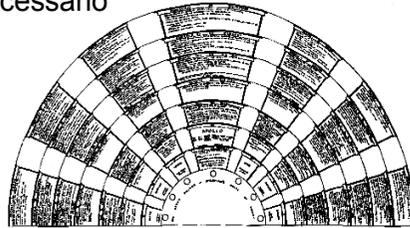
## Ipertesti

- un ipertesto è un testo accessibile anche in modo non lineare
  - Esempio: l'enciclopedia
- possiede una struttura attiva di collegamenti che permettono all'utente di spostarsi da una parte all'altra delle informazioni, agevolando la consultazione non lineare
- V. Bush (1945):
  - “La mente umana non lavora in questo modo. Essa opera in modo associativo. Avendo afferrato un concetto, essa salta istantaneamente al prossimo che viene suggerito dall'associazione di idee, in accordo con qualche intricata ragnatela di percorsi tracciata dalle cellule del cervello”
- Cfr. *Ulysses* di J. Joyce (1922): linearizzazione di un flusso mentale non lineare

## Breve storia degli ipertesti

1550 Giulio Camillo: *Theatro della Memoria*

- per meccanizzare la memorizzazione di grandi quantità di informazioni, e il richiamo dalla memoria di frammenti di esse quando necessario

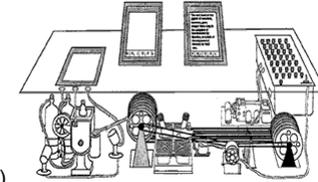


InfoDoc 04-05 - Modulo II

36

## Breve storia degli ipertesti (2)

- *MEMEX*, V.Bush (1945)
  - MEMory Extension



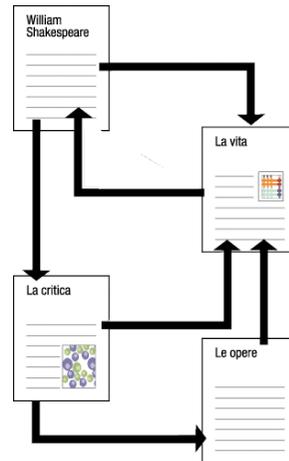
- *AUGMENT*, D.Engelbart (1963)
- *XANADU*, T.Nelson (1967) - conio la parola *ipertesto*
- ...
- Berners-Lee (1989): World Wide Web, HTML e HTTP
- Etc. etc. Si veda <http://www2.polito.it/didattica/polymath/ICT/Htmls/Argomenti/Apunti/Storialpertesto/Storialpertesto.htm>

InfoDoc 04-05 - Modulo II

37

## Ipertesto: definizioni

- **Nodo**: raggruppamento significativo di dati. A volte si può distinguere il tipo di nodo, in base alla sua struttura
- **Link**: collegamento tra nodi o frammenti di nodi
- tramite link, i nodi possono essere organizzati in **unità semantiche**, ovvero legate da qualche vincolo di significato, a prescindere dalla memorizzazione fisica



InfoDoc 04-05 - Modulo II

38

## I link

- in generale, un link ha una sua rappresentazione grafica attivabile dall'utente per passare da un nodo all'altro
- Finalità dei link:
  - organizzazione dell'informazione
  - connessione di frammenti di testo
  - connessione di riferimenti ai documenti riferiti
  - connessione di commenti o annotazioni
  - connessione di celle di tabelle a descrizioni o altre tabelle
  - esecuzione di programmi esterni
- In base alle finalità, abbiamo tre funzioni:
  - **link di riferimento** per l'organizzazione non gerarchica dell'ipertesto
  - **link organizzativi**, per la strutturazione gerarchica delle informazioni
  - **link a parola chiave**: link impliciti che connettono un frammento ad uno o più frammenti, tramite meccanismi di ricerca per chiave

InfoDoc 04-05 - Modulo II

39

## Multimedia/Ipermedia

- i documenti multimediali comprendono informazioni provenienti da sorgenti diverse: testo, immagini, audio, video, ...
- i documenti ipermediali sono documenti multimediali con capacità ipertestuali, ovvero con possibilità di accesso non lineare
- un nodo di un documento ipermediale sarà quindi costituito da testo, immagini, etc., ed un link potrà puntare a frammenti costituiti da questi particolari tipi di documenti

## World Wide Web (Web, WWW, w<sup>3</sup>)

- è un unico ipertesto distribuito sulla rete Internet
- si basa su diversi formati di memorizzazione delle informazioni, e qualsiasi tipo di informazione
  - quello che dà le caratteristiche ipertestuali è HTML
- utilizza diversi protocolli applicativi di Internet (potenzialmente tutti)
  - quello principalmente usato è HTTP
- Storicamente:
  - nasce nel 1989 al CERN per la collaborazione su progetti di ricerca internazionali di fisica, con condivisione di documenti (di testo)
  - 1991: prima demo pubblica, solo testo
  - 1993: prima interfaccia grafica (NCSA Mosaic) -> successo
  - 1994: nasce Netscape, CERN e MIT fondano il W3 Consortium
  - 1995: Netscape, sebbene in rosso, raccoglie investimenti...

## WWW: cos'è? (per l'utente...)

- collezione di documenti (testo, immagini, etc) che risiede su calcolatori sparsi in giro, e connessi tramite Internet
- I documenti vengono chiamati **pagine**
- ogni pagina può contenere dei collegamenti (**link**) ad altre pagine correlate, situate ovunque
- l'utente può seguire questi collegamenti, cliccando sui link e muovendosi di pagina in pagina
- Le pagine sono viste grazie ad un programma detto **browser**, che recupera e mostra le pagine, interpreta le richieste dell'utente, etc.

## WWW: cosa c'è sotto?



## WWW: ipermedia distribuiti

- WWW è un documento ipermediale distribuito:
  - distribuito perché lo spazio di memorizzazione fisica è costituito dai file system di computer siti ovunque e connessi ad Internet
  - ipermediale perché i nodi possono essere costituiti da qualsiasi tipo di documento, il cui "tessuto" ipertestuale è dato da HTML
  - i nodi ed il modo di accesso sono identificati da nomi appositi (**URL**)
- si basa sul principio client/server
  - il client si occupa della visualizzazione dei nodi, e dell'interpretazione delle richieste di navigazione dell'utente
  - i dati sono invece mantenuti e distribuiti da appositi server in modalità diverse

## URL - Uniform Resource Locator

- in generale, i nodi di un ipertesto devono essere identificati in qualche modo
- se l'ipertesto è anche distribuito, allora i problemi sono:
  - dare un nome al nodo
  - identificare dove il nodo è memorizzato
  - indicare come accedere al nodo
- l'URL serve a questo scopo, ed è fatto così:  
schema://indirizzo.su.Internet/identificatore/locale/della/pagina
  - **schema** è il protocollo per l'accesso (può essere anche locale)
  - **l'indirizzo** identifica il calcolatore su Internet ove risiede il nodo (pagina)
  - **l'identificatore locale** è usualmente la posizione del nodo nel file system locale, o sua abbreviazione

## URL

- Gli schemi possibili sono i più svariati, e corrispondono ai protocolli applicativi di Internet (http, ftp, file, news, gopher, mailto, telnet,...)
- Gli indirizzi possono essere ulteriormente specificati dalla porta su cui connettersi, se non viene utilizzata quella standard per il protocollo
- L'identificatore locale può essere ulteriormente specificato per indicare anche uno specifico frammento all'interno del nodo
- Esempi:  
`http://www.uniud.it:80/drrm/anpat/home.html#frammento`  
`mailto:miculan@dimi.uniud.it`  
`news:it.hobby.motociclismo`  
`ftp://ftp.unipd.it/mirror/rfc/rfc-index.html`  
`file:///Users/miculan/Sites/DiMI/index.html`

## La componente ipertestuale: HTML

- **HTML** (HyperText Markup Language) è un linguaggio di markup, cioè i cui comandi sono inseriti esplicitamente all'interno del testo
- è un'applicazione di SGML Standard Generalized Markup Language, ISO 8879)
- i comandi permettono la formattazione del testo similmente a quel che accade in un word processor, e però anche la creazione di link
- in aggiunta a ciò, alcuni comandi hanno anche valore semantico, in quanto indicano porzioni di documento di significato specifico (es. titolo, intestazioni, etc.).
- è comunque un formato testuale

# HTML

- I comandi di HTML si chiamano **TAG**.
- possono essere annidati
- Assumono queste forme:
  - `<nome_tag>`
  - `testo`
  - `</nome_tag>`
  - `<nome_tag attributo1=val1 ... attributoN=valN>`
  - `testo </nome_tag>`
  - `<nome_tag>testo`

## Esempio campione

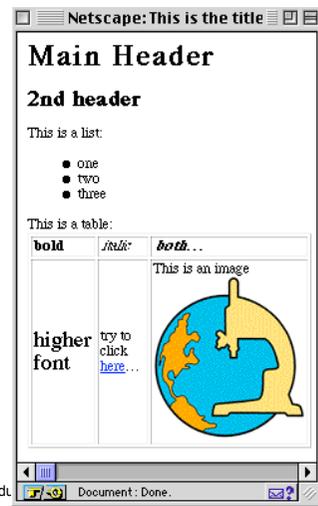
```
<html>
<head>
<title>This is an
HTML document
</title>
</head>
<body>
...
</body>
</html>
```

# I tag

- semantici: `<HEAD>`, `<BODY>`, `<TITLE>`, `<H1>...<Hn>`, `<ADDRESS>`, `<P>`, `<UL>`, `<OL>`, `<LI>`, `<TABLE>`, `<TR>`, `<TD>`,
- di formattazione del testo: `<I>`, `<B>`, `<BR>`, `<FONT>`, ...
- definizione di link (detti **ancore**): `<A HREF="nodo puntato">frammento puntante</A>` (di solito i link appaiono **sottolineati**)
- inserimento di multimedia: `<IMG SRC="risorsa">`, ...
- input utente: `<FORM>`, `<INPUT>`, `<SELECT>`, ...
- tante altre cose in continua evoluzione

# Un esempio

```
<HTML><HEAD>
<TITLE>This is the title</TITLE>
</HEAD>
<BODY BGCOLOR=White>
<H1>Main Header</H1>
<H2>2nd header</H2>
This is a list:
<UL><LI> one </LI> two
<LI> three</UL>
This is a table:
<TABLE BORDER>
<TR><TD><B>bold</B>
<TD><I>italic</I>
<TD><I><B>both...</B></I></TR>
<TD><FONT SIZE=5>higher
font</FONT>
<TD>try to click <A
HREF="link.html">here</A>...
<TD>This is an image <BR>
<IMG SRC="img.gif"></TABLE>
</BODY></HTML>
```



# HTML

- Il linguaggio di specifica per le pagine Web è l'HTML (HyperText Markup Language).
- Si tratta di un linguaggio che utilizza dei marcatori (detti anche istruzioni) inseriti opportunamente nel testo "normale" in modo da conferire ad un documento una natura ipertestuale (i.e., la possibilità di una lettura non necessariamente sequenziale) e multimediale (inclusione di immagini, suoni, filmati).
- Quindi le pagine HTML non sono altro che dei file di testo in cui sono presenti dei marcatori.
- La visualizzazione di una pagina web è il risultato finale dell'interpretazione dei marcatori HTML da parte del browser (tale processo viene denominato talvolta "rendering").

## Caratteri speciali

Nelle pagine HTML, bisogna utilizzare dei codici appositi per essere sicuri che certi caratteri vengano visualizzati correttamente da tutti i browser:

Entità carattere	Carattere corrispondente	Descrizione
&lt;	<	Segno di minore
&gt;	>	Segno di maggiore
&amp;	&	Ampersand
&quot;	"	Apice doppio
&nbsp;		Spazio unificatore
&copy;	©	Copyright
&reg;	®	Marchio registrato
&#126;	~	Carattere corrispondente al numero 126 nel set di caratteri di default (ISO Latin-1)

InfoDoc 04-05 - Modulo II

53

## Intestazioni

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Esempi di intestazioni </TITLE>
</HEAD>

<BODY>

<H1>Intestazione di livello 1</H1>
<H2>Intestazione di livello 2</H2>
<H3>Intestazione di livello 3</H3>
<H4>Intestazione di livello 4</H4>
<H5>Intestazione di livello 5</H5>
<H6>Intestazione di livello 6</H6>

</BODY>
</HTML>
```

**Intestazione di livello 1**

**Intestazione di livello 2**

**Intestazione di livello 3**

**Intestazione di livello 4**

**Intestazione di livello 5**

Intestazione di livello 6

InfoDoc 04-05 - Modulo II

54

## Esempio

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Declaration of independence </TITLE>
</HEAD>

<BODY>
<H1 ALIGN="center">Declaration of Independence</H1>
<H5 ALIGN="right">(Adopted in Congress 4 July 1776)</H5>
<H3>The Unanimous declaration of the Thirteen United States of America</H3>
<P><H4 ALIGN="center">When, in the course of human events, it becomes necessary
for one people to dissolve the political bonds which have connected them with another,
and to assume among the powers of the earth, the separate and equal station to which
the laws of nature and of nature's God entitle them, <BR> a decent respect to the
opinions of mankind requires that they should declare the causes which impel them
to the separation.</H4></P>
</BODY>
</HTML>
```

### Declaration of Independence

(Adopted in Congress 4 July 1776)

#### The Unanimous declaration of the Thirteen United States of America

When, in the course of human events, it becomes necessary for one people to dissolve the political bonds which have connected them with another, and to assume among the powers of the earth, the separate and equal station to which the laws of nature and of nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

Inf

## Testo preformattato

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0
Transitional//EN">
<HTML>
<HEAD>
<TITLE> Blocchi di testo </TITLE>
</HEAD>

<BODY>

In HTML per andare a capo
bisogna usare &lt;br>&lt;br> in questo modo.

<PRE>
oppure bisogna racchiudere
il testo all'interno l'istruzione PRE.
</PRE>

</BODY>
</HTML>
```

**Risultato:**

In HTML per andare a capo bisogna usare <br> in questo modo.

oppure bisogna racchiudere il testo all'interno l'istruzione PRE.

InfoDoc 04-05 - Modulo II

56

## Blocchi di testo

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Blocchi di testo </TITLE>
</HEAD>
```

```
<BODY>
```

```
<P ALIGN="center">Esempio di paragrafo allineato <STRONG>centralmente</STRONG></P>
```

```
Citazione:
```

```
<BLOCKQUOTE>
Il flusso del tempo che corre veloce porta il tormento dell'intuire.
</BLOCKQUOTE>
```

Risultato:

```
<ADDRESS>
H.P. Lovecraft<BR>
Providence, Rhode Island
</ADDRESS>
```

```
</BODY>
</HTML>
```

Esempio di paragrafo allineato <b>centralmente</b>
Citazione:
Il flusso del tempo che corre veloce porta il tormento dell'intuire.
<i>H.P. Lovecraft</i> <i>Providence, Rhode Island</i>

## Tabelle

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Tabelle </TITLE>
</HEAD>
<BODY>
<TABLE CELLSPACING="5" BORDER="2">
<CAPTION ALIGN="bottom">I pianeti interni del Sistema Solare</CAPTION>
<TR>
<TH colspan="2"></TH>
<TH colspan="2">Distanza dal Sole</TH>
<TH rowspan="2">Durata<BR>dell'anno</TH>
<TH rowspan="2">Durata<BR>del giorno</TH>
</TR>
<TR>
<TH>Km</TH>
<TH>UA</TH>
</TR>
<TR>
<TH>Mercurio</TH>
<TD>57.900.000</TD>
<TD>0,38</TD>
<TD>88 giorni</TD>
<TD>59 giorni</TD>
</TR>
<TR>
<TH>Venere</TH>
<TD>108.200.000</TD>
<TD>0,72</TD>
<TD>225 giorni</TD>
<TD>243 giorni</TD>
</TR>
<TR>
<TH>Terra</TH>
<TD>149.600.000</TD>
<TD>1,0</TD>
<TD>365 giorni</TD>
<TD>24 ore</TD>
</TR>
<TR>
<TH>Marte</TH>
<TD>227.900.000</TD>
<TD>1,5</TD>
<TD>687 giorni</TD>
<TD>24,6 ore</TD>
</TR>
</TABLE>
</BODY>
</HTML>
```

continua a [laboroDoc 04-05 - Modulo II](#)

58

## Tabelle

	Distanza dal Sole		Durata dell'anno	Durata del giorno
	Km	UA		
Mercurio	57.900.000	0,38	88 giorni	59 giorni
Venere	108.200.000	0,72	225 giorni	243 giorni
Terra	149.600.000	1,0	365 giorni	24 ore
Marte	227.900.000	1,5	687 giorni	24,6 ore

I pianeti interni del Sistema Solare

59

## Righe orizzontali

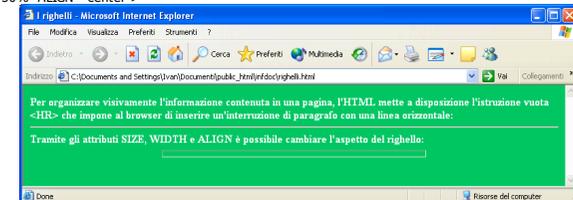
```
<HTML>
<HEAD>
<TITLE> I righelli </TITLE>
</HEAD>
```

```
<BODY BGCOLOR="#00CC66">
<FONT COLOR="#FFFFFF">
<STRONG>
```

Per organizzare visivamente l'informazione contenuta in una pagina, l'HTML mette a disposizione l'istruzione vuota <HR> che impone al browser di inserire un'interruzione di paragrafo con una linea orizzontale. Tramite gli attributi SIZE, WIDTH e ALIGN &grave; possibile cambiare l'aspetto del righello:

```
<HR SIZE="10" WIDTH="50%" ALIGN="center">
```

```
</STRONG>
</FONT>
</BODY>
</HTML>
```



## Le immagini (sorgente)

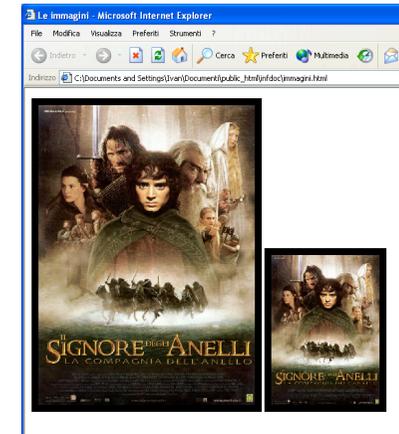
```
<HTML>
<HEAD>
<TITLE> Le immagini </TITLE>
</HEAD>

<BODY>
<IMG SRC="C:\Documents and
Settings\Ivan\Documenti\Immagini\compagnia0.jpg"
  ALT="Il Signore degli Anelli: La Compagnia dell'Anello" BORDER="10"
  WIDTH="30%"
  HEIGHT="70%">
<IMG SRC="..\..\Immagini\compagnia0.jpg"
  ALT="Il Signore degli Anelli: La Compagnia dell'Anello" BORDER="10"
  WIDTH="15%"
  HEIGHT="35%">
</BODY>
</HTML>
```

InfoDoc 04-05 - Modulo II

61

## Le immagini (risultato visivo)



62

## I collegamenti

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> I collegamenti </TITLE>
</HEAD>

<BODY>
<A NAME="top"><H1>Inizio pagina</H1></A>
<P>Link ad un documento esterno su un altro web server (indirizzamento assoluto):<BR>
<A HREF="http://www.isignoredeglianeli.it" >Il Signore degli Anelli: Sito ufficiale</A></P>
<P>Link ad un documento sullo stesso web server in cui si trova questa pagina (indirizzamento relativo):<BR>
<A HREF="immagini.html"> Il Signore degli Anelli: La compagnia dell'Anello</A></P>
<TABLE>
<TR><TH>Il Signore degli Anelli: Le due Torri</TH></TR>
<TR><TD><IMG SRC="..\..\Immagini\duetorri0.jpg" WIDTH="562" HEIGHT="600" BORDER="0"></TD></TR>
</TABLE>

<P>
Link ad un altro punto in questo stesso documento:<BR>
<A HREF="#top">Inizio pagina</A>
</P>

</BODY>
</HTML>
```

InfoDoc 04-05 - Modulo II

63

## I collegamenti



64

## I collegamenti



65

## Gli stili del testo

- Le istruzioni di stile cambiano l'aspetto tipografico del testo.
- Si possono individuare due categorie di stili:
  - Stili logici: rendono lo stile del testo relativo, dicendo, ad esempio, che una certa parola deve essere enfattizzata rispetto al contesto.
  - Stili fisici: permettono un controllo più preciso di come viene reso visivamente il testo dal browser.
- Esempi di istruzioni di stili logici:
  - `<EM></EM>` (enfasi)
  - `<STRONG></STRONG>` (maggiore enfasi)
  - `<TT></TT>` (Typewriter Text)
  - `<BIG></BIG>`, `<SMALL></SMALL>` (testo un po' più grande/piccolo)
  - `<STRIKE></STRIKE>` (testo barrato)
  - `<SUB></SUB>`, `<SUP></SUP>` (pedici e apici)
- Esempi di istruzioni di stili fisici:
  - `<B></B>` (grassetto)
  - `<i></i>` (corsivo)
  - `<U></U>` (sottolineato)

InfoDoc 04-05 - Modulo II

66

## Elenchi ordinati e non ordinati

```
<HTML>
<HEAD>
<TITLE> Elenchi </TITLE>
</HEAD>
```

```
<BODY>
<H1 ALIGN="center">Corso di HTML</H1>
<HR>
<H2>Contenuti</H2>
<H3>Capitoli</H3>
<OL>
<LI>Introduzione</LI>
<LI>Il linguaggio</LI>
<UL>
<LI>Sintassi</LI>
<LI>Formattazione</LI>
</UL>
<LI>Scrittura di ipertesti</LI>
</OL>
<HR>
</BODY>
</HTML>
```



Info

Done

Risorse del computer

## Attributi degli elenchi

- Elenchi ordinati:
  - TYPE="1" (numerazione normale: default)
  - TYPE="A" (lettere maiuscole: A, B, C, D, ...)
  - TYPE="a" (lettere minuscole: a, b, c, d, ...)
  - TYPE="I" (numeri romani maiuscoli: I, II, III, IV, ...)
  - TYPE="i" (numeri romani minuscoli: i, ii, iii, iv, ...)
- Elenchi non ordinati:
  - TYPE="circle" (cerchi vuoti: default)
  - TYPE="square" (quadrati pieni)
  - TYPE="disc" (cerchi pieni)

InfoDoc 04-05 - Modulo II

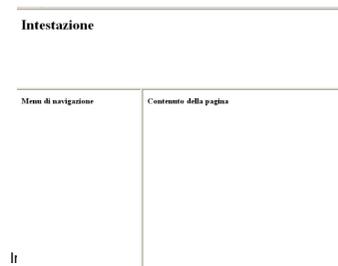
68

## Struttura di una pagina HTML

- Al fine di facilitare la “navigazione” dell'utente e di rendere facilmente riconoscibile il punto del documento ipertestuale/sito in cui ci si trova, le pagine HTML solitamente hanno una struttura standard che solitamente comprende quanto segue:
  - l'**intestazione** (titolo, logo ecc.) del sito;
  - il **menu di navigazione** (contenente i link alle varie sezioni del sito);
  - il **contenuto** della pagina corrente (questa parte è quella a cui viene dedicato maggior spazio).
- Ovviamente, l'intestazione ed il menu di navigazione sono parti fisse che si ripetono in ogni pagina del documento ipertestuale/sito.
- Quindi in ogni pagina è necessario ricopiare il codice HTML per riprodurre l'intestazione ed il menu di navigazione.

## I frame

- Per evitare di ricopiare il codice HTML relativo a parti che rimangono fisse in ogni pagina, sono stati introdotti i **frame** (quadri).
- I frame consentono di suddividere la finestra di un browser in vari quadri all'interno dei quali è possibile caricare una pagina HTML distinta.



## I frame

- Per realizzare il risultato del lucido precedente (formato da 3 quadri), sono necessari 4 file:
  - il primo descrive la struttura (ovvero, come viene suddivisa la finestra);
  - i file rimanenti sono delle pagine HTML convenzionali che vengono caricate nei vari quadri seguendo l'impostazione strutturale del primo file.
- Nel caso dell'esempio del lucido precedente, abbiamo quindi 4 file:
  - frames.html (file che descrive la struttura dei quadri);
  - intestazione.html (pagina caricata nel frame in alto nella prima riga);
  - menu.html (pagina caricata nel frame a sinistra nella seconda riga);
  - contenuto.html (pagina caricata nel frame a destra nella seconda riga).

## frames.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Esempio di utilizzo dei frame </TITLE>
</HEAD>

<FRAMESET ROWS="20%,*">
  <FRAME SRC="intestazione.html" NAME="intestazione">
  <FRAMESET COLS="20%,*">
    <FRAME SRC="menu.html" NAME="menu">
    <FRAME SRC="contenuto.html" NAME="contenuto">
  </FRAMESET>
</FRAMESET>

</HTML>
```

## intestazione.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Intestazione </TITLE>
</HEAD>

<BODY>
<H1>Intestazione</H1>
</BODY>

</HTML>
```

## menu.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Menu di navigazione </TITLE>
</HEAD>

<BODY>
<H3>Menu di navigazione</H3>
</BODY>

</HTML>
```

## contenuto.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Contenuto della pagina </TITLE>
</HEAD>

<BODY>
<H3>Contenuto della pagina</H3>
</BODY>

</HTML>
```

## I frame

- Non tutti i browser supportano in modo adeguato i frame (vi possono essere dei problemi soprattutto con versioni datate di alcuni browser).
- Quindi gli sviluppatori del linguaggio HTML hanno pensato di introdurre un marcatore apposito per consentire di visualizzare comunque qualche cosa anche nel caso in cui l'utente non disponga di un browser che supporti i frame.
- E' sufficiente inserire, ad esempio, dopo la chiusura del tag FRAMESET più esterno quanto segue:

```
<NOFRAMES>
<BODY>
    Il browser non supporta i frame.
</BODY>
</NOFRAMES>
```

In questo modo l'utente potrà leggere l'avviso che lo informa che il suo browser non supporta i frame (invece di vedere una pagina vuota).

## Attributi del tag FRAMESET

- **ROWS**: serve a specificare da quante righe (con le relative dimensioni) è formato un insieme di frame. Ad esempio ROWS="20%,\*" indica che vi saranno due righe (la prima alta il 20% della dimensione della finestra, la seconda alta ciò che rimane, i.e., l'80% a causa della presenza dell'asterisco).
- **COLS**: serve a specificare da quante colonne (con le relative dimensioni) è formato un insieme di frame. Ad esempio COLS="200,500,1200" indica che vi saranno tre colonne (la prima larga 200 pixel, la seconda 500 pixel e la terza 1200 pixel).
- **BORDER**: indica lo spessore del bordo che circonda i vari frame; ad esempio BORDER="10" indica che il bordo sarà spesso 10 pixel.
- **FRAMEBORDER**: FRAMEBORDER="NO" indica che i frame non devono essere circondati da un bordo.
- **FRAMESPACING**: specifica la distanza in pixel tra un frame e quelli adiacenti; ad esempio FRAMESPACING="20" specifica una distanza di 20 pixel.

## Attributi del tag FRAME

- **SRC**: specifica il percorso della pagina HTML che verrà caricata nel frame.
- **NAME**: specifica il nome del frame (in modo che ci si possa riferire a quest'ultimo nei collegamenti tramite l'attributo TARGET del tag A).
- **SCROLLING**: controlla se sia possibile scorrere (SCROLLING="YES") oppure no (SCROLLING="NO") il contenuto del frame nel caso che le dimensioni di quest'ultimo siano insufficienti a visualizzare tutto in una sola schermata.
- **NORESIZE**: se presente impedisce che il frame sia ridimensionato dall'utente trascinando i bordi.

## Frame ricorsivi

- Consideriamo la struttura definita dal codice seguente:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Frame ricorsivi </TITLE>
</HEAD>

<FRAMESET ROWS="25%,50%,25%">
  <FRAME SRC="top.html" NAME="top">
  <FRAMESET COLS="*,2*">
    <FRAME SRC="midleft.html" NAME="midleft">
    <FRAME SRC="midright.html" NAME="midright">
  </FRAMESET>
  <FRAME SRC="bottom.html" NAME="bottom">
</FRAMESET>

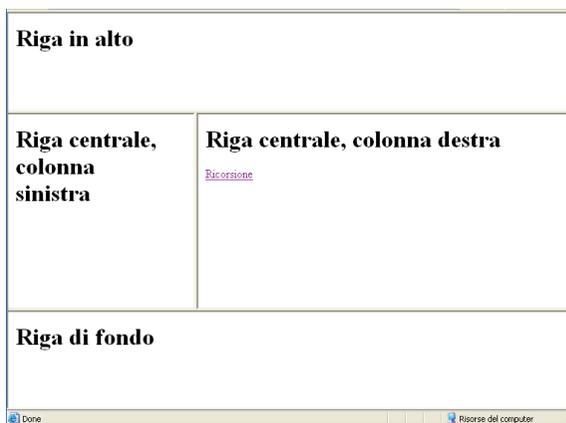
</HTML>
```

## Frame ricorsivi

- La pagina HTML caricata nel frame a destra della seconda riga contiene un link che, se attivato, produce un effetto ricorsivo, ovvero, ricarica l'intero insieme di quadri al suo posto (index.html è il nome del file del lucido precedente).
- ```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Riga centrale, colonna destra </TITLE>
</HEAD>

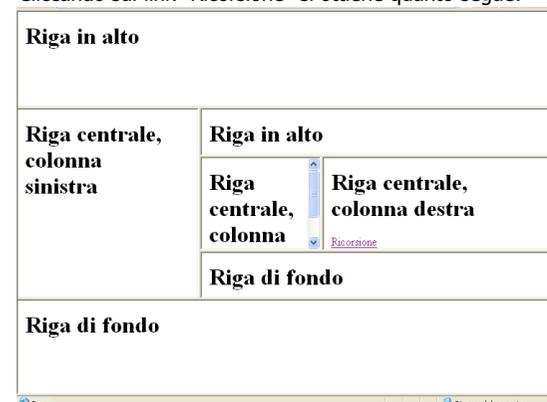
<BODY>
<H1>Riga centrale, colonna destra</H1>
<A HREF="index.html">Ricorsione</A>
</BODY>
</HTML>
```

## Frame ricorsivi



## Frame ricorsivi

Cliccando sul link "Ricorsione" si ottiene quanto segue:



82

## Vantaggi dei Frame

- I frame sono utili in quanto consentono di risparmiare tempo quando si vuole modificare una parte della struttura che rimane fissa durante la navigazione (ad esempio l'intestazione o il menu di navigazione).
- Infatti è sufficiente modificare soltanto una pagina HTML per vedere le modifiche riprodotte automaticamente in qualunque sezione del documento ipertestuale/sito.
- Invece, con un'impostazione tradizionale senza frame, bisognerebbe andare a modificare ogni singola pagina HTML in cui è contenuta la parte "fissa" da cambiare.
- Con i frame bisogna prestare attenzione ai collegamenti che vanno predisposti con cura per evitare di caricare una pagina nel quadro sbagliato oppure ottenere effetti di ricorsione.
- Per caricare un documento in un frame diverso da quello in cui compare il link, bisogna utilizzare l'attributo TARGET del tag A, specificando il nome del frame in cui si vuole caricare il documento collegato.  
Esempio: <A HREF="pagina.html" TARGET="contenuto">Link</A>

## Fogli di Stile

- I fogli di stile a cascata (**Cascading Style Sheets – CSS**) sono stati introdotti per consentire la separazione fra contenuto logico di un documento ipertestuale e aspetto esteriore.
- In effetti un uso accorto dei fogli di stile a cascata consente di "ripulire" il codice HTML da tutti i tag e gli attributi che influenzano soltanto l'aspetto estetico.
- Ciò rende i sorgenti HTML molto più semplici da mantenere e da capire, in quanto i tag e gli attributi rimanenti descrivono soltanto la struttura logica del documento.
- Esistono due livelli di fogli di stile a cascata, descritti in dettaglio ai seguenti indirizzi:
  - [www.w3.org/pub/WWW/TR/REC-CSS1](http://www.w3.org/pub/WWW/TR/REC-CSS1)
  - [www.w3.org/pub/WWW/TR/REC-CSS2](http://www.w3.org/pub/WWW/TR/REC-CSS2)

## Fogli di stile

- Un foglio di stile è un insieme di regole ognuna delle quali è composta dai seguenti elementi:
  - selettore (ciò che viene influenzato);
  - dichiarazione racchiusa tra parentesi graffe (effetto).
- Esempi:  
`H2 {color:blue}`  
In questo modo tutte le intestazioni di livello 2 (contraddistinte dal tag H2) saranno di colore blu.  
La regola seguente invece fa in modo che tutti i paragrafi (contraddistinti dal tag P) rientrino di mezzo pollice per quanto riguarda la prima riga:  
`P {text-indent:0.5in}`

## CSS nel documento HTML

E' possibile usare i fogli di stile inserendo le regole direttamente Nel documento HTML tramite il tag STYLE (nell'intestazione):

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
<TITLE> Fogli di stile a cascata (CSS) </TITLE>

<STYLE TYPE="text/css">
H2 {color:blue}
P {text-indent:0.5in}
H2, H3, H4, H5, H6 {font-family:"fantasy, Arial"; font-weight:Bold; font-size:18pt}
EM {color:red}
BODY {font-style: italic; font-family:"Verdana"; font-size:20pt}
TABLE {font-style: normal; font-family:"Verdana"; font-size:12pt}
TD {padding: 30}
</STYLE>
...
```

## Fogli di stile

- Oltre ai marcatori, è possibile specificare come selettore una classe, ovvero, un nome assegnato dall'utente (il nome deve iniziare con un punto) che servirà a richiamare una regola di stile in punti diversi del file (anche in associazione a marcatori distinti).
- Ad esempio le regole  
`.paragrafo {text-align: justify}`  
`.intestazione {text-align: right}`  
possono essere richiamate nel codice HTML nel modo seguente:  
`<H1 CLASS="intestazione">Titolo</H1>`  
e  
`<DIV CLASS="paragrafo">Testo</DIV>`

## Fogli di stile in file separati

- Per limitare le modifiche riguardanti gli aspetti esteriori ad un singolo file, è preferibile mantenere le regole in un file con estensione css ed includerlo nelle pagine in cui lo si vuole utilizzare con il marcatore seguente (nell'intestazione):  
`<LINK REL="stylesheet" TYPE="text/css" HREF="stile.css">`
- In questo caso le regole appaiono soltanto nel file stile.css:  
`H2 {color:blue}`  
`P {text-indent:0.5in}`  
`H2, H3, H4, H5, H6 {font-family:"fantasy, Arial"; font-weight:Bold; font-size:18pt}`  
`EM {color:red}`  
`BODY {font-style: italic; font-family:"Verdana"; font-size:20pt}`  
`TABLE {font-style: normal; font-family:"Verdana"; font-size:12pt}`  
`TD {padding: 30}`

## Il protocollo HTTP

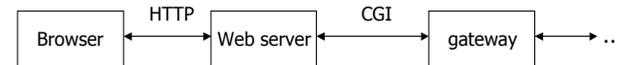
- Il World Wide Web funziona per mezzo degli HTTP (HTML Transfer Protocol) server (detti anche Web Server) che forniscono le pagine richieste ai client (browser).
- Il protocollo HTTP specifica quattro fasi di interazione fra client e server:
  - Apertura della connessione: il browser richiede al server (specificato nell'URL) l'apertura di una connessione TCP (Transport Control Protocol) per verificare la disponibilità del servizio.
  - Connessione: il server (se possibile) accetta la connessione ed invia una conferma al browser.
  - Richiesta: il browser invia al server la richiesta di una risorsa (e.g., pagina HTML).
  - Risposta (reply): il server comunica al browser se la risorsa richiesta è disponibile ed in caso affermativo la invia, terminando nel contempo la connessione.
- Appare evidente quindi che il protocollo HTTP è **stateless**, i.e., non ha modo di tener traccia delle operazioni portate a termine per uno stesso client.

InfoDoc 04-05 - Modulo II

89

## Gateway (programma CGI)

- Un gateway è un programma richiamato dal Web server che permette di stabilire una connessione fra Web ed un ambiente esterno (e.g., un DBMS).
- La comunicazione fra Web server ed i gateway è regolata da un meccanismo noto come CGI (Common Gateway Interface):
  - L'utente richiede l'esecuzione di un gateway cliccando su un apposito link o inviando i dati tramite un form.
  - Il Web server lancia il gateway passandogli i parametri inviati dall'utente.
  - Il gateway termina la sua esecuzione (nel corso della quale può interagire con altri elementi di un sistema informativo, e.g., un database) e restituisce i risultati al Web server.
  - Il Web server trasmette i risultati ricevuti dal gateway al client (browser).



InfoDoc 04-05 - Modulo II

90