

Richiami sugli elementi del linguaggio HTML

Un documento in formato Web può essere aperto con un browser, attraverso un collegamento a Internet oppure caricandolo dal disco del proprio computer senza connettersi alla rete. Sullo schermo viene visualizzata una pagina in formato grafico: la visualizzazione è il risultato di un'elaborazione del browser che interpreta i codici contenuti nel file e li trasforma in comandi per la costruzione della pagina.

Le pagine di Internet, in particolare tutti i documenti scritti per la tecnologia **WWW** (*World Wide Web*), sono redatti con un linguaggio studiato per la gestione degli ipertesti: non si tratta di un linguaggio di programmazione, ma piuttosto di un **linguaggio di formazione della pagina**.

Un **ipertesto** è un insieme di documenti che ha la possibilità di essere consultato in modo non sequenziale, passando da un documento all'altro attraverso collegamenti (**link**) realizzati mediante parole o immagini, sui quali occorre fare clic con il mouse.

Esso può contenere elementi non solo testuali, ma anche immagini, suoni, filmati, e qualunque tipo di oggetto trattabile da un sistema informatico. In ambiente Internet, in particolare, utilizzando i riferimenti **URL** (*Uniform Resource Locator*), cioè gli indirizzi dei siti, si può accedere da un documento all'altro registrati su

computer ubicati in qualsiasi parte del mondo, realizzando così un'attività che viene comunemente definita **navigazione** nella rete (*netsurfing*).

Il linguaggio **HTML** (*HyperText Markup Language*, cioè linguaggio che utilizza dei contrassegni per la formazione di ipertesti) consente di sfruttare i vantaggi e le prestazioni derivanti dall'organizzazione ipertestuale e dall'uso degli oggetti multimediali: questo linguaggio è diventato lo standard nell'architettura WWW per creare e riconoscere i documenti ipermediali.

I documenti WWW, tipicamente scritti in HTML, sono dei normali testi di caratteri, e di conseguenza sono visibili e modificabili con qualunque programma di trattamento dei testi, dal *word processor* al semplice *editor*.

I testi scritti in HTML si distinguono dai normali file con estensione *.txt* o comunque da file di testo non HTML tramite l'estensione *.html* nel nome del file; seguendo l'abitudine derivante dalle regole del sistema operativo DOS, per il quale i suffissi dei nomi di file non possono eccedere i tre caratteri, spesso i testi HTML hanno come suffisso *.htm*. Il suffisso non è un elemento obbligatorio, ma è utilizzato normalmente per identificare il tipo di file. In Windows i documenti in HTML vengono rappresentati con un'icona che raffigura il logo del browser utilizzato.

Tali documenti, quindi, non sono altro che dei normali file di caratteri; in più vengono aggiunti i codici di formazione della pagina che prendono il nome di **tag** (cioè contrassegno, da cui deriva il termine *MarkUp* della sigla di HTML) consistenti in sequenze di caratteri proprie del linguaggio HTML che non fanno parte del testo normale, e che consentono di realizzare gli elementi caratteristici dell'ipertesto: i *link* (salti ipertestuali) e gli oggetti grafici e multimediali. Altri codici servono per semplificare l'impaginazione dei documenti (stili del testo, titolo dei documenti, paragrafi, liste), così da rendere il linguaggio utile anche per creare documenti complessi.

Per rappresentare all'interno dei documenti i link a oggetti multimediali e i link alle pagine della rete, il linguaggio HTML utilizza i riferimenti di tipo URL, con i quali è praticamente possibile rappresentare ogni file o servizio in Internet.

I *tag* del linguaggio HTML sono costituiti da stringhe di caratteri comprese tra i simboli di minore (<) e maggiore (>):

```
<CODICE parametro1 parametro2 ... parametroN>
```

In HTML un codice è, nella maggior parte dei casi, abbinato a un altro codice corrispondente, che segna la fine dell'azione prevista da quel codice. Questo secondo tag ha una barra (*slash*) all'inizio:

</CODICE >

Per esempio:

il codice **** attiva il neretto nella scrittura dei caratteri; per specificare la fine del neretto si usa il codice che indica la fine dell'azione: ****.

Tutti i caratteri del testo compresi tra **** e **** verranno visualizzati in neretto dal browser.

Tutti i codici possono essere indifferentemente specificati in caratteri maiuscoli o minuscoli, però è utile scrivere i tag sempre in maiuscolo per distinguerli dalle frasi del testo.

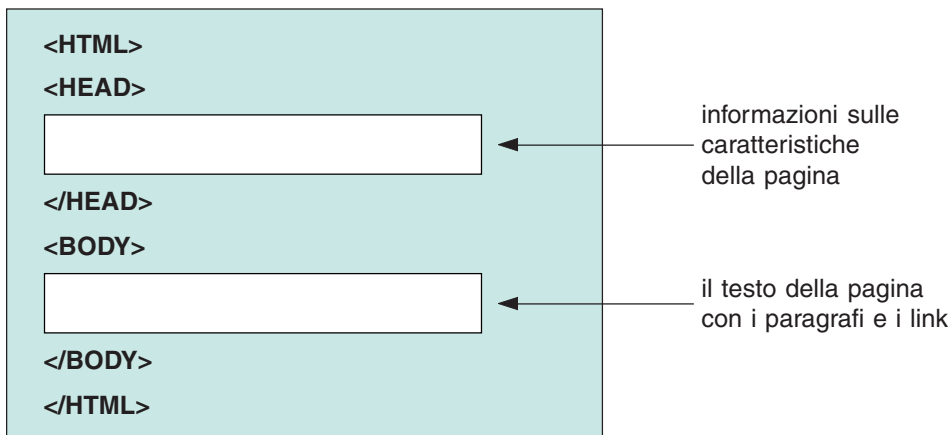
Di norma un testo HTML viene suddiviso in almeno due **sezioni**:

- la sezione *intestazione* (**HEAD**) contenente informazioni sul titolo e sulle altre caratteristiche (sfondo, colore dei caratteri) della finestra nella quale il documento viene visualizzato;
- la sezione *corpo* (**BODY**), contenente il documento vero e proprio con tutti gli elementi caratteristici di un ipertesto.

La codifica di queste due sezioni avviene attraverso le coppie di tag **<HEAD> </HEAD>** per l'intestazione e **<BODY> </BODY>** per il corpo.

Inoltre un documento HTML è di norma delimitato dai due tag **<HTML> </HTML>**, perché alcuni browser devono essere informati su quale porzione di testo è necessario applicare l'interpretazione del linguaggio HTML.

Quindi un documento HTML ha la seguente struttura di base:



Di seguito vengono illustrati i tag più importanti che servono per definire un testo in linguaggio HTML.

Solitamente è necessario differenziare le **dimensioni dei caratteri**, anche solo per distinguere il titolo dal resto del testo, o comunque per definire dei sottotitoli o elementi testuali evidenziati. Per questo è prevista una serie di sei codici differenti per altrettante dimensioni dei caratteri. I valori di queste dimensioni sono

preceduti nel tag dalla lettera H:

<Hn>...</Hn>, dove *n* indica il codice numerico corrispondente alla dimensione del carattere.

I tag, in ordine di grandezza dal più grande al più piccolo, sono i seguenti:

```
<H1> testo </H1>
<H2> testo </H2>
<H3> testo </H3>
<H4> testo </H4>
<H5> testo </H5>
<H6> testo </H6>
```

Il numero fra 1 e 6 specifica in effetti l'importanza della parte di testo (1 risulta essere il più importante).

Alcune parti possono essere evidenziate usando uno **stile** diverso dal testo normale, con il corsivo, il neretto o il sottolineato. La parte di testo deve essere racchiusa tra una delle seguenti coppie di tag:

```
<I>...</I> per ottenere il testo in corsivo (italic)
<B>...</B> per il neretto o grassetto (bold)
<U>...</U> per il sottolineato (underline).
```

Un **paragrafo** è una porzione di testo separata dalle altre da una riga di spazio o semplicemente da un ritorno a capo.

Per specificare l'inizio di un nuovo paragrafo è sufficiente usare il tag **<P>**; l'inserimento di questo codice, oltre a creare una riga di separazione tra paragrafi, viene utilizzata per assegnare le impostazioni tipografiche al paragrafo stesso.

Il codice **<P>** è anche utilizzato per imporre un salto di paragrafo quando serve.

Per inserire invece un ritorno a capo si usa il tag **
** per l'interruzione di riga, che si può scrivere ovunque serva.

Talvolta può essere utile separare un paragrafo da un altro attraverso una linea orizzontale: questa riga viene inserita con il tag **<HR>**.

A ogni pagina HTML può essere associato un **titolo** che ne descrive il contenuto e che viene riportato nella barra del titolo della finestra del browser di Internet.

Il titolo è usato principalmente per identificare il documento (per esempio quando si cerca un documento tra tanti attraverso un motore di ricerca oppure quando una pagina viene inserita in un bookmark di siti preferiti): è opportuno quindi che ogni titolo sia diverso dagli altri all'interno dell'insieme di pagine che formano un sito e che sia formato da una frase significativa per ricordarne il contenuto.

Il titolo viene racchiuso tra la coppia di tag **<TITLE> </TITLE>**.

Per rappresentare i **colori** dei caratteri o dello sfondo di una pagina, il linguaggio HTML utilizza una combinazione dei tre colori fondamentali RGB (*red, green, blue*) attraverso tre numeri, compresi ciascuno tra 0 e 255, che specificano la gamma cromatica del rosso, del verde e del blu. Questi valori sono scritti in formato esadecimale (base 16) e quindi variano da 00 a FF, essendo FF la rappresentazione in base 16 del numero 255.

Per esempio:

```
nero      00 00 00
bianco    FF FF FF
rosso     FF 00 00
blu       00 00 FF
verde     00 FF 00
```

I colori intermedi corrispondono a valori compresi tra 00 e FF per i tre colori.

Per esempio:
giallo FF E3 AC.

Per assegnare un colore di sfondo a una pagina si usa il tag **BGCOLOR** all'interno della sezione <BODY>. Per esempio il tag <BODY BGCOLOR="FFFFFF"> consente di visualizzare una pagina con lo sfondo bianco, per ottenere uno sfondo giallo chiaro, si può usare il tag <BODY BGCOLOR="FFFFE0">.

Come già detto, la principale caratteristica di un ipertesto è la possibilità di consultare il documento in modo non sequenziale attraverso l'uso di collegamenti ad altri oggetti o pagine. Con i **link** si effettuano riferimenti a file remoti oppure ad altre parti del documento stesso (consentendo per esempio di spostarsi velocemente da un indice alla voce richiesta). Il link appare sulla pagina come un'immagine, un'animazione o semplicemente una parola (o una breve frase) di norma sottolineata e di colore blu: passando su di esso il puntatore del mouse si trasforma diventando una mano; questo simbolo indica che si può attivare il collegamento premendo il tasto del mouse. Un link già visitato, di solito, assume il colore rosso porpora. La struttura del tag che definisce un link ipertestuale è la seguente:

** parola **

Il *nomefile* indica la destinazione del link, e deve essere indicato con la sintassi degli indirizzi URL. *Parola* indica la parte di testo che compare sottolineata e in colore per rappresentare il link. La lettera A maiuscola, all'inizio e alla fine del tag, sta per **àncora** (*anchor*), cioè punto di attracco, destinazione del salto ipertestuale; quindi questo link specifica il riferimento ipertestuale (*HREF*) di un'ancora.

Per esempio la riga:

**Prova **

produce sul video la parola *Prova* sottolineata e in colore; se con il mouse si fa clic su di essa, il browser richiama il file *prova.htm*: abbiamo quindi realizzato un link ipertestuale ad un'altra pagina.

Con i link definiti dal tag <A HREF> ... , è possibile anche produrre un salto ipertestuale a un altro sito Internet, specificandone l'indirizzo URL:

**Visitate il sito Prova **

Sul video la frase *Visitate il sito Prova* appare sottolineata e in colore; passando sopra di essa il puntatore del mouse diventa una mano e, facendo clic su di essa, il browser richiama la pagina di apertura del sito richiesto.

I browser per Internet sono in grado di visualizzare **immagini** che sono state memorizzate in uno dei diversi formati grafici. In particolare i formati grafici più usati sono:

- **GIF** (Graphics Interchange Format)
- **JPEG** (Joint Photographic Experts Group).

Le immagini possono essere inserite all'interno di un documento HTML.

Per includere un'immagine nel documento si deve usare il seguente tag HTML:

dove *nomeimmagine* è l'indirizzo URL del file contenente l'immagine.

Per realizzare l'interattività con l'utente si possono inserire all'interno delle pagine Web i moduli con gli elementi standard dell'interfaccia grafica.

I **moduli** (*form*) in HTML sono racchiusi tra la coppia di tag <FORM> ... </FORM>.

All'interno del *form* si possono inserire oggetti per acquisire dati dall'utente (*INPUT*), che possono essere:

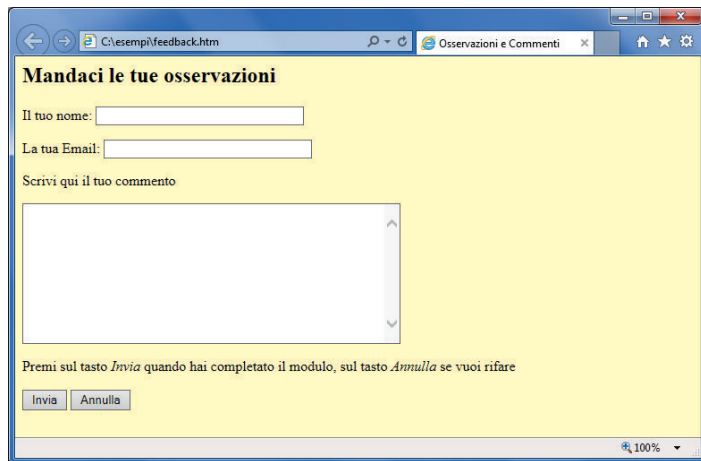
- caselle di testo (*TYPE = "text"*) di cui specifichiamo la lunghezza (*SIZE*),
- aree di testo con barre di scorrimento (*TEXTAREA*), indicando anche il numero di righe (*ROWS*) e di colonne (*COLS*) che vengono visualizzate,
- bottoni per inviare i dati (*TYPE = "submit"*) oppure per annullare quanto inserito (*TYPE = "reset"*); *VALUE* indica la parola che deve comparire sul bottone.

Il seguente esempio presenta una pagina Web che contiene molti dei tag HTML descritti in precedenza.

ESEMPIO

Predisporre una pagina Web per raccogliere suggerimenti o commenti dai visitatori, formata da moduli sul video da riempire con il nome del mittente, l'indirizzo di posta elettronica e il testo del commento.

Una pagina di questo tipo si chiama *modulo di feedback*. Il modulo completato dall'utente viene inviato tramite posta elettronica facendo clic su un bottone di comando inserito nella pagina.



```
<HTML>
<HEAD>
<TITLE>Osservazioni e Commenti</TITLE>
</HEAD>
<BODY BGCOLOR="FFFFCC">
<H2><B>Mandaci le tue osservazioni</B></H2>
<P>
<FORM METHOD="post" ACTION="mailto:rossi@prova.it">
<P>Il tuo nome: <INPUT TYPE="text" NAME="mittente" SIZE="35">
<P>La tua Email: <INPUT TYPE="text" NAME="email" SIZE="35">
<P>Scrivi qui il tuo commento
<P><TEXTAREA NAME="domanda" ROWS="10" COLS="50"></TEXTAREA>
<P>Premi sul tasto <I>Invia</I> quando hai completato il modulo, sul tasto
<I>Annulla</I> se vuoi rifare
<P><INPUT TYPE="submit" VALUE="Invia" NAME="B1">
<INPUT TYPE="reset" VALUE="Annulla" NAME="B2">
</FORM>
</BODY>
</HTML>
```