

Memento de la syntaxe html

Internet et Outils (IO2)

Généralités

Un document HTML commence toujours par une ligne contenant `<!DOCTYPE html>`.

Le document est structuré à l'aide de *balises* qui indiquent la nature de ce qu'elles encadrent et qu'on appelle *élément*. Dans la suite on dira que les balises `<p>` et `</p>` encadrent l'élément `<p>`. Les balises sont écrites entre chevrons (les symboles `<` et `>`). Il en existe deux types :

- Les balises en paires : l'une ouvre et l'autre ferme le champ (et commence donc par une barre oblique `/`).

Exemple : `<p> Ceci est un paragraphe. </p>`

- Les balises orphelines : il n'y a qu'une seule balise qui représente un élément ponctuel.

Exemple : `<meta>`

Chacune de ces balises peut avoir des *attributs* qui apparaissent entre les chevrons après le nom de la balise. Par exemple, la balise `<meta>` permet entre autres de préciser l'encodage des caractères. Aujourd'hui, il vaut mieux utiliser l'`utf-8` qui permet de tenir compte de toutes les langues : `<meta charset="utf-8">`

Structure du fichier

- `<!--` et `-->` : encadrent les commentaires.
- `<!DOCTYPE html>` : pour que le document soit reconnu en tant qu'HTML.
- `<html>` : élément qui contient tous les autres ; ses fils sont `head` et `body`.
- `<head>` : en-tête du document (informations diverses).
- `<body>` : contenu du document.

Blocs

- `<h1>` à `<h6>` : titres du plus important (`h1`) au moins important (`h6`).
- `<p>` : paragraphe.
- `<div>` : bloc quelconque (utile pour le CSS).

A l'intérieur d'un paragraphe

- `` : mettre en évidence (*emphasize*), ce qui correspond à l'italique par défaut.
- `` : mettre fortement en évidence, ce qui correspond au gras par défaut.
- `<kdb>` : pour représenter du texte machine.
- `` : permet de délimiter un passage (pour en faire quelque chose en CSS ou en Javascript).
- `<a>` : lien hypertexte.
- `` : image.

A éviter

Les éléments suivants existent mais peuvent avantageusement être remplacés à l'aide du CSS (qui permettra de les modifier de façon uniforme), et doivent donc être évités :

- `<i>` : permet de mettre le texte en italique sans connotation d'insistance (par exemple, pour un moteur de recherche).
- `` : permet de mettre le texte en gras sans connotation d'insistance (par exemple, pour un moteur de recherche).
- `
` : saut de ligne.
- `<hr>` : barre horizontale.

Listes

- `` : liste à puces (*unordered list*).
- `` : liste numérotée (*ordered list*).
- `` : élément de l'une des listes précédentes (*list item*).
- `<dl>` : liste de définitions (*definition list*).
- `<dt>` : terme défini (*definition term*) dans une `dl`.
- `<dd>` : définition d'un terme (après une `dt` dans une `dl`).

Il est possible de mettre une liste dans une liste, etc.

Tableaux

- `<table>` : définit un tableau.
- `<tr>` : définit une ligne (*table row*).
- `<th>` : définit une case d'entête (*table head*).
- `<td>` : définit une case simple (*table data*).

On pourra aussi utiliser les éléments suivants pour construire donner la sémantique (utilisés par le CSS) :

- `<thead>` : en-tête.
- `<tfoot>` : bas du tableau.
- `<tbody>` : corps du tableau.

Structuration sémantique (HTML 5)

- `<nav>` : liste de liens (menu).
- `<article>` : délimite un texte, qui peut inclure les éléments qui suivent.
- `<header>` : contient des éléments d'introduction (titre, sommaire...) à l'élément qui le contient (la page principale, un article...).
- `<footer>` : contient des éléments annexes (à mettre en pied-de-page) à l'élément qui le contient (auteur, date, références...).
- `<section>` : délimite une partie de l'élément qui le contient.
- `<aside>` : contient des éléments hors du texte (une liste de liens, une image, etc.).

Caractères réservés

En HTML comme dans tous les langages, certains caractères sont réservés (par exemple, `<` signale le début d'une balise). On peut accéder à ces caractères par leur nom :

- ` ` : espace insécable.
- `<` : inférieur (`<`).
- `>` : supérieur (`>`).
- `≤` : inférieur ou égal (`≤`).
- `≥` : supérieur ou égal (`≥`).
- `&` : éperluette (`&`).
- `€` : euro (€).
- `é` : e accent aigu (é).
- `É` : E accent aigu (É).

- `ê` : e accent circonflexe (ê).
- `Ê` : e accent circonflexe (Ê).

Les accents peuvent être obtenus sur n'importe quelle lettre majuscule ou minuscule (il suffit d'indiquer la lettre en majuscule ou en minuscule et le nom de l'accent comme ci-dessus).

Avec le codage `utf-8`, on peut aussi écrire directement les caractères accentués lorsque le clavier le permet.

À peu près tous les caractères peuvent s'obtenir de cette manière. Pour plus d'information, faire une recherche (par exemple avec les mots clés « caractères spéciaux HTML »).