

Cahiers **GUT** *enberg*

∞ PRÉSENTATION DE HTML

¶ Christian ROLLAND

Cahiers GUTenberg, n° 19 (1995), p. 67-84.

<http://cahiers.gutenberg.eu.org/fitem?id=CG_1995__19_67_0>

© Association GUTenberg, 1995, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

implique l'accord avec les conditions générales

d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.

Présentation de HTML*

Christian ROLLAND

E-mail: `Christian.Rolland@univ-brest.fr`

Résumé. Cet article présente le langage de balisage HTML (*HyperText Markup Language*), utilisé pour représenter des documents hypertextes dans le cadre de *Word Wide Web*. Nous présentons les balises indiquant les structures les plus couramment utilisées puis le balisage d'hypertextes.

Abstract. *This article presents a markup language, HTML (HyperText Markup Language), which is used to represent hypertext documents in Word Wide Web. Tags which indicate the most usual structures are shown; then the hypertext tags and other features are exhibited.*

1. Généralités

HTML (*HyperText Markup Language*) a été conçu comme une classe de document SGML (*Standard Generalized Markup Language*, norme Iso 8879, voir l'article de M. GOOSSENS dans ce même numéro) et est assimilable à un langage, basé sur un ensemble de balises permettant de décrire des documents d'une manière structurée. Les différentes structures présentes dans un document HTML sont encadrées par une paire de balises : une balise de début `<NomD'Élément>` et une balise de fin `</NomD'Élément>`. Bien que la présence de certaines balises de fin ne soit pas obligatoire, il est souhaitable qu'à chaque balise de début corresponde une balise de fin. Dans la suite de cet article, les balises de fin seront présentes sauf dans les cas où elles sont structurellement inutiles.

Comme dans tout document SGML, le nom des balises peut être indifféremment écrit en minuscules, en majuscules ou avec une combinaison des deux. De plus, plusieurs espaces successifs interrompus ou non par des changements de ligne sont pris en compte comme un seul espace et les caractères accentués ne doivent pas être utilisés directement.

*. Présentation réalisée à Nanterre, le 19 janvier 1995 lors de la journée *Diffusion des documents électroniques*.

Tous les documents HTML doivent posséder une structure minimale commune divisées en deux parties distinctes :

- un en-tête, où sont donnés des renseignements d'identification ;
- un corps, où se trouve la partie visible (texte) du document.

Ces éléments de structure du document HTML sont marqués par les balises `<HEAD>` et `</HEAD>`, qui encadrent l'en-tête, et `<BODY>` et `</BODY>`, qui encadrent le corps. L'ensemble doit être encadré par les balises `<HTML>` et `</HTML>`. L'en-tête doit contenir au moins un élément, *TITLE*, faisant office de titre « externe¹ » pour le document.

1	<code><HTML></code>
2	<code><HEAD></code>
3	<i>Partie 1 : en-tête</i>
4	<code><TITLE> Titre externe </TITLE></code>
5	<code></HEAD></code>
6	<code><BODY></code>
7	<i>Partie 2 : texte balisé</i>
8	<code></BODY></code>
9	<code></HTML></code>

2. Structures usuelles

Les structures autorisées en HTML sont de plusieurs ordres. Certaines servent à diviser le texte en parties distinctes, d'autres servent à indiquer des préférences de mise en forme de texte et enfin d'autres permettent d'utiliser des fonctions spéciales.

2.1. Division structurelle du document

Dans HTML, le document n'est pas explicitement divisé en parties, chapitres, sections, etc. comme c'est le cas dans d'autres langages. La division se fait par une hiérarchie d'éléments dont chacun est libre de donner la signification qu'il désire (titre de document, section, sous-section, titre de paragraphe, etc.). En effet, le titre de ces éléments de division est précisé selon son niveau dans une hiérarchie : les balises sont nommées *Hn* où *n*

1. Le titre externe n'est pas le titre visible lorsque l'on consulte le document. Il s'agit d'une manière de faire référence à ce document, un peu comme le nom d'un fichier, même si ce titre externe peut être plus explicite. Un document peut donc posséder deux titres à la fois : un titre externe ne faisant pas partie du texte du document et un titre interne inclus dans le texte du document.

est le niveau dans la hiérarchie : H_1, H_2, \dots, H_6 . Les deux exemples suivants montrent chacun une utilisation possible de ces balises.

```

1 <H1> De l'art de cuisiner </H1>
2   <H2> Les entrées </H2>
3   <H2> Le plat principal </H2>
4   <H2> Les desserts </H2>

```

<pre> 1 <H1> Partie 1 : Les entrées </H1> 2 <H2> 1.1 Les crudités </H2> 3 <H2> 1.2 Les surgelés </H2> 4 <H1> Partie 2 : Le plat principal </H1> 5 <H2> 2.1 Viandes </H2> 6 <H2> 2.2 Poissons </H2> 7 <H1> Partie 3 : Les desserts </H1> 8 <H2> 3.1 Fruits </H2> 9 <H2> 3.2 Gâteaux </H2> 10 <H2> 3.3 Glaces </H2> </pre>	<p>Partie 1 : Les entrées</p> <p>1.1 Les crudités</p> <p>1.2 Les surgelés</p> <p>Partie 2 : Le plat principal</p> <p>2.1 Viandes</p> <p>2.2 Poissons</p> <p>...</p>
---	---

Le texte de la division² du document est habituellement une succession de paragraphes et de divisions du niveau inférieur. Les paragraphes sont représentés par l'élément *P*.

<pre> 1 <P> 2 Il fait beau. Mais qu'est-ce que 3 le beau temps s'il n'y a pas eu 4 la pluie auparavant ? 5 [...] 6 </P> 7 <P> 8 Et comment ne le serait-ce 9 pas ? [...] 10 </P> </pre>	<p>Il fait beau. Mais qu'est-ce que le beau temps s'il n'y a pas eu la pluie auparavant? [...]</p> <p>Et comment ne le serait-ce pas? [...]</p>
---	---

2.2. Listes

Cinq types de listes sont disponibles dans HTML. Ces listes sont délimitées par des balises qui indiquent le type de liste représentée. Chaque élément constituant la liste est balisé par les balises `` et ``.

2.2.1. Liste de type répertoire

Une liste de type répertoire est représentée par l'élément *DIR*. Elle doit être composée d'éléments contenant chacun moins de 20 caractères. Selon

2. Partie, chapitre, section, etc. selon la signification qui lui est donnée.

les logiciels utilisés, les listes de ce type pourront être ou non disposées en colonnes.

<pre> 1 <DIR> 2 html.aux 3 html.dvi 4 html.html 5 html.log 6 html.ps 7 html.tex 8 image.eps 9 </DIR> </pre>	<pre> html.aux html.ps html.dvi html.tex html.html image.eps html.log </pre>
--	--

2.2.2. Liste de type menu

Une liste de type menu est représentée par l'élément *MENU*. Elle est composée d'éléments dont la longueur ne dépasse pas une ligne. Chaque objet de la liste est généralement placé au début d'une nouvelle ligne lors de la mise en forme du document pour la visualisation.

<pre> 1 <P> 2 Voici les solutions possibles. 3 </P> 4 5 <MENU> 6 Beau temps 7 Mauvais temps 8 Temps variable 9 </MENU> </pre>	<pre> Voici les solutions possibles. Beau temps Mauvais temps Temps variable </pre>
--	--

2.2.3. Liste ordonnée

Une liste ordonnée est représentée par l'élément *OL*. Il s'agit d'une énumération d'éléments. Le compteur d'énumération est affiché au début d'une nouvelle ligne, suivi de l'objet de liste correspondant.

<pre> 1 2 allumez l'ordinateur ; 3 tapez votre nom de connexion ; 4 tapez votre mot de passe ; 5 lancez le programme que vous 6 désirez utiliser. 7 </pre>	<pre> 1. allumez l'ordinateur ; 2. tapez votre nom de connexion ; 3. tapez votre mot de passe ; 4. lancez le programme que vous désirez utiliser. </pre>
---	---

2.2.4. Liste non ordonnée

Une liste non ordonnée est représentée par l'élément *UL*. Elle est composée d'une série d'éléments non ordonnés (aucun numéro d'ordre ne leur est associé).

<pre> 1 <H1> Contenu du panier </H1> 2 3 4 Oranges 5 Pommes de terre 6 Poireaux 7 Fromage 8 Jambon fumé 9 </pre>	<p>Contenu du panier</p> <ul style="list-style-type: none"> • Oranges • Pommes de terre • Poireaux • Fromage • Jambon fumé
--	--

2.2.5. Liste descriptive

Une liste descriptive est représentée par l'élément *DL*. Contrairement aux listes précédentes, les objets d'une liste descriptive sont composés de deux parties : la partie entrée et la partie description. L'entrée est représentée par l'élément *DT* et la partie description par l'élément *DD*.

L'attribut *COMPACT* peut-être associé à l'élément glossaire en vue d'une disposition de la liste, si possible, plus « compacte » (par exemple, placement de l'entrée en face de la première ligne de la description au lieu de la placer sur une ligne seule).

<pre> 1 <H2> Glossaire </H2> 2 <DL COMPACT> 3 <DT> Glose </DT> 4 <DD>Commentaire d'un mot ou d'un 5 passage écrit entre les lignes 6 ou dans les marges d'un 7 manuscrit.</DD> 8 <DT> Missus </DT> 9 <DD>À l'époque carolingienne, 10 délégué chargé de contrôler 11 les représentants locaux du 12 roi.</DD> 13 </DL> </pre>	<p>Glossaire</p> <p>Glose Commentaire d'un mot ou d'un passage écrit entre les lignes ou dans les marges d'un manuscrit.</p> <p>Missus À l'époque carolingienne, délégué chargé de contrôler les représentants locaux du roi.</p>
---	--

2.3. Élément d'information


Dans les documents HTML, un élément de structure est prévu pour indiquer les coordonnées de l'auteur, le propriétaire du document ou tout autre information permettant l'obtention de compléments sur le sujet traité. Les logiciels de visualisation de documents HTML traitent habituellement ces informations au fer à droite. L'élément *ADDRESS* est généralement utilisé en début ou fin de document.

<pre> 1 <ADDRESS> 2 Christian Rolland
 3 Christian.Rolland@univ-brest.fr
 4 Association GUTenberg 5 <\ADDRESS> </pre>	<p><i>Christian Rolland</i> <i>Christian.Rolland@univ-brest.fr</i> <i>Association GUTenberg</i></p>
---	---

Dans l'exemple précédent, la balise `
` a été utilisée pour forcer le changement de ligne (voir en 4.3).

3. Images

Il est tout à fait possible d'insérer des images dans un document HTML. Ceci peut être fait de deux façons : soit à l'extrémité d'un lien hypertexte, soit directement dans le texte. Pour le premier cas, voir en 5. En ce qui concerne le second cas, c'est l'élément `IMG` qui indique l'image à intégrer dans le formatage final (visuel) du document. Le nom du fichier contenant l'image est précisé par l'intermédiaire de l'attribut `SRC` ; un autre attribut `ALT`, facultatif, permet d'indiquer un texte à afficher au cas où le programme de visualisation du document n'a pas la possibilité de traiter ce type d'image. Si le format de l'image est connu du formateur HTML, celle-ci sera traitée, sinon c'est le message textuel qui sera utilisé.

<pre> 1 Un triangle T peut être contenu dans 2 un cercle C : 3 </pre>	<p>Un triangle T peut être contenu</p>  <p>dans un cercle C :</p>
---	---

Un autre attribut, `ALIGN`, purement dédié à la présentation est utilisable. Il peut prendre les valeurs :

- `TOP` pour aligner le haut de l'image sur la ligne courante ;
- `BOTTOM` pour aligner le bas de l'image sur la ligne courante ;
- `MIDDLE` pour aligner le milieu de l'image sur la ligne courante.

4. Indications de présentation

4.1. Styles de présentation

HTML autorise l'auteur d'un document à ajouter dans le document des indications pour la présentation. Ceci peut se faire de deux manières : soit

en utilisant des styles logiques de présentation, soit en utilisant des styles physiques de présentation. À chacun de ces styles est associé un élément particulier.

4.1.1. Styles logiques

Les styles logiques de présentation sont les plus généraux puisqu'ils n'imposent en aucun cas la taille des caractères, la police, etc. Les styles prévus sont au nombre de huit et permettent de définir des caractéristiques générales (citation, emphase,...) ou informatiques (nom de variable, programme,...) concernant une partie d'un texte. Le logiciel de visualisation a carte blanche pour la mise en forme. Les éléments correspondants sont les suivants :

<i>CITE</i>	pour une citation ;
<i>CODE</i>	pour une partie de programme ;
<i>DFN</i>	pour la définition d'un terme ;
<i>EM</i>	pour un texte mis en emphase simple ;
<i>KBD</i>	pour une séquence tapée au clavier par l'utilisateur ;
<i>SAMP</i>	pour une série simple de caractères ;
<i>STRONG</i>	pour un texte mis en emphase forte ;
<i>VAR</i>	pour le nom d'une variable.

<pre> 1 Ce cher auteur a écrit 2 dans son livre <CITE>Micro, typo, 3 dodo</CITE> que cette opération 4 nécessitait la saisie de la 5 commande <KBD>traduc<KBD> 6 <VAR>NomDuFichier</VAR>.</pre>	<p>Ce <i>cher</i> auteur a écrit dans son livre <i>Micro, typo, dodo</i> que cette opération nécessitait la saisie de la commande <code>traduc</code> <i>NomDuFichier</i>.</p>
--	--

L'emphase simple *pourrait* se traduire, lors du formatage, par l'utilisation de l'italique (dans un texte en romain). L'emphase dans l'emphase n'est pas explicitement prévue dans la spécification de HTML, pas plus que la combinaison des différents styles.

4.1.2. Styles physiques

Les styles physiques prévus dans HTML imposent le choix de la police de caractères, ce qui rend le document dépendant des possibilités de l'application chargée de l'interpréter. Si celle-ci n'est pas capable d'utiliser du

gras, alors la mise en forme particulière du texte qui demande du gras ne sera pas restituée. Les éléments correspondant à ces préférences sont :

- B* pour mettre le texte en caractères gras ;
- I* pour mettre le texte en italiques ;
- TT* pour mettre le texte en caractères de type machine à écrire ;
- U* pour mettre le texte en caractères soulignés.

<pre>1 Comme disait <U>ma</U> 2 grand-mère, <I>la culture, 3 il n'y a que ça</I>.</pre>	<pre>Comme disait <u>ma</u> grand-mère, <i>la</i> <i>culture</i>, <i>il n'y a que ça</i>.</pre>
--	---

4.2. Styles de paragraphes

En plus du paragraphe simple délimité par les balises `<P>` et `</P>`, deux autres éléments assimilables aux paragraphes sont utilisables.

4.2.1. Citation de texte

Alors que l'élément *CITE* sert principalement à citer des noms d'ouvrages, l'élément *BLOCKQUOTE* est utilisé pour des citations de texte.

<pre>1 Comme le disait Harpagon dans 2 <CITE>L'Avare</CITE> de 3 Molière : 4 <BLOCKQUOTE> 5 Il me semble que j'entends un 6 chien qui aboie. N'est-ce point 7 qu'on en voudrait à mon argent ? 8 </BLOCKQUOTE></pre>	<pre>Comme le disait Harpagon dans L'Avare de Molière : Il me semble que j'en- tends un chien qui aboie. N'est-ce point qu'on en voudrait à mon argent?</pre>
--	--

4.2.2. Texte préformaté

Une disposition de texte prédéfinie peut être réalisée par l'intermédiaire de l'élément *PRE*, permettant ainsi la représentation de textes destinés à une mise en forme précise. Le texte que l'élément *PRE* contient est dit préformaté car les passages à la ligne du texte source HTML sont conservés, ainsi que les espaces (si deux espaces se suivent, deux espaces seront imprimés et non un seul comme normalement). D'autre part, c'est une police de caractères non proportionnels qui est employée dans ce cas par l'interpréteur HTML. Il est possible de préciser un attribut de présentation qui

indique le nombre de colonnes sur lesquelles les caractères du texte seront placés. Généralement, c'est le nombre 80 qui est pris par défaut.

<pre> 1 <PRE WIDTH=40> 2 <KBD> ls -laF *.tex </KBD> 3 4 +--- nom des fichiers 5 +----- options 6 +----- commande 7 </PRE></pre>	<pre> ls -laF *.tex ^ ^ ^ +--- nom des fichiers +----- options +----- commande</pre>
---	--

4.3. Éléments de présentation

Il existe, dans la liste des balises de HTML, deux balises qui n'ont pas de signification structurelle. Ces balises ne s'utilisent pas en couple, seule la balise de la forme « balise de début » a une signification.

 indique un retour à la ligne forcé ;

<HR> indique un filet horizontal de la largeur de la « page ».

5. Parlons hypertexte

Le propre de HTML est de permettre de réaliser des documents hypertextes. Un document hypertexte possède des zones actives pointant vers des parties de lui-même ou bien vers des parties ou l'intégralité d'autres documents. Dans le cas de WWW/HTML, ces documents sont bien souvent disséminés sur tout le globe et accessibles *via* Internet. Lorsque ces zones spéciales sont activées, un lien hypertexte est suivi et le document pointé est visualisé.

HTML possède la notion d'*ancree*. Elle correspond à une zone du document qui est soit un point d'arrivée soit un point de départ de lien hypertexte. L'ancree de départ est une zone active alors que l'ancree d'arrivée ne l'est pas.

Une ancree — active ou non — est représentée par l'élément *A* dont les propriétés sont précisées par des attributs. Dans le cas où l'élément indique une ancree de départ, c'est une adresse URL (voir l'article de L. OTTAVI dans ce même numéro) qui est précisée en destination ou bien une étiquette si cette destination est dans le même document (lien externe ou interne). Si c'est une arrivée explicite de lien, l'élément est obligatoirement nommé par une étiquette. Seule l'une de ces deux indications est obligatoire pour qu'un

lien soit utile. Les attributs autorisés pour indiquer les caractéristiques de l'ancre sont les suivants.

HREF Cet attribut précise que l'ancre est active et donne l'URL de la destination du lien. En plus de la syntaxe de type URL, une forme du type *#étiquette* est aussi autorisée pour compatibilité avec l'attribut *NAME*. Cette forme en *#étiquette* peut être collée à la suite d'une adresse URL pour préciser un endroit précis d'un document.

```
<A HREF="http://www.demo.fr/html/intro.html">introduction</A>
```

NAME Cet attribut précise que l'ancre peut être la destination d'un lien et donne un identificateur (\simeq étiquette) à celle-ci.

```
<A NAME="conclusion">conclusion</A>
```

REL Cet attribut précise une relation entre le document actif et un autre document. Il ne doit pas être employé séparément de l'attribut *HREF* dans un lien. Son utilisation correspond souvent à la spécification d'une relation de parenté, de succession, etc. permettant ainsi de diviser un gros document en plusieurs parties et de diriger la consultation.

```
<LINK REL="Next" HREF="partie2.html">Suite</A>
```

REV Cet attribut précise une relation entre le document actif et un autre document et correspond à la relation inverse de *REL*. Son emploi correspond aux mêmes conditions que celui-ci.

```
<LINK REV="Previous" HREF="partie1.html">Revenir</A>
```

URN Cet attribut précise le *numéro de ressource uniforme* du document. Ce serait un moyen unique de référencer la plus proche copie d'un document par rapport à l'endroit (réseau) où la demande de consultation a été faite. Encore à l'étude par divers groupes de travail.

TITLE Cet attribut précise le nom du document pointé par l'attribut *HREF*. Il est utilisable à titre d'information ou pour nommer des documents non-HTML.

```
<A HREF="ccl.html" TITLE="Conclusion">conclusion</A>
```

METHODS Cet attribut précise les *méthodes* HTTP acceptées par l'objet lors d'une utilisation publique.

```
<A METHODS="GET,HEAD,DELETE" HREF="zx45.html">suite</A>
```

```

1 Pour la définition de
2 <A HREF="#glose">glose</A>, voir le
3 <A HREF="#glossaire">glossaire</A>
4 ou le
5 <A HREF="http://www.biblio.fr/dico.html">
6 dictionnaire </A>.
7 <A NAME="glossaire"></A>
8 <H2>Glossaire</H2>
9 <DL>
10 <DT><A NAME="glose">Glose</A></DT>
11 <DD>Commentaire d'un mot ou d'un
12 passage écrit entre les lignes ou
13 dans les marges d'un manuscrit.</DD>
14 </DL>

```

Pour la définition de glose, voir le glossaire ou le dictionnaire.

Glossaire

Glose

Commentaire d'un mot ou d'un passage écrit entre les lignes ou dans les marges d'un manuscrit.

Un lien, dans un document HTML, peut pointer sur n'importe quel type de document : autre document HTML, image graphique (Gif, Jpeg, etc.), fichier PostScript accessible par le protocole ftp, répertoire d'un serveur ftp, etc. Chaque document est traité selon le protocole de transmission choisi et selon son type. Si le logiciel de visualisation n'est pas capable de traiter tel ou tel format de fichier, il fait généralement appel à un autre programme qui en est capable.

6. En-tête

La partie en-tête d'un document HTML n'est pas visualisable car les renseignements qu'elle contient n'ont de signification que pour le logiciel de visualisation du document.

L'élément « titre externe » (*TITLE*) est le seul des éléments inclus dans l'en-tête qui soit obligatoire. Il est aussi possible de préciser un identificateur unique pour le document qui pourra ainsi être référencé par lui (il est généralement fixé arbitrairement par des éditeurs hypertexte) par l'élément *NEXTID* qui admet l'attribut *N* pour cela ; le nom du fichier contenant le document HTML actif sous la forme d'une adresse URL par l'élément *BASE* qui admet l'attribut *HREF* pour indiquer l'URL ; si le document peut faire l'objet d'une recherche par mot clé par l'élément *ISINDEX* qui est normalement généré par le serveur de documents HTML ; ainsi qu'une série de relations entre le document et d'autres documents (précédent, suivants, etc.) par l'élément *LINK* qui admet les mêmes attributs qu'une ancre hypertexte *A*. Hormis dans le cas de l'élément *TITLE*, seule la balise de début d'élément est utilisée.

```

1 <HEAD>
2 <BASE HREF="/usr/local/www/index.html">
3 <TITLE> Association GUTenberg </TITLE>
4 <ISINDEX>
5 <NEXTID N="GUTenberg034">
6 <LINK REL="Parent" HREF="menu.html">
7 <LINK REV="Previous" HREF="bienvenue.html">
8 </HEAD>

```

7. Commentaires

L'utilisation de commentaires est possible dans HTML sous la forme de balises spéciales : `<!--` pour le début et `-->` pour la fin du commentaire.

```

1 <H1> Partie 3 </H1>
2 <!-- Partie encore à écrire -->
3
4 <!-- <H1> Partie 4 </H1> -->

```

Partie 3

8. Caractères accentués et symboles

Le texte courant d'un document HTML est écrit en code Ascii (7 bits). Ceci est tout à fait adapté à la langue anglaise pour la saisie, mais pas aux langues latines par exemple. Pour remplacer les caractères accentués et d'autres caractères spéciaux, il faut faire appel à des *entités* SGML. Ces entités sont composées du symbole `&`, suivi d'un mot clé et d'un point-virgule (;). Une fois lues par le logiciel de visualisation, elles sont interprétées et affichées sous la forme du caractère correspondant dans le codage du système où l'affichage se fait. Dans ces entités, sont aussi définis des symboles comme `<`, `>`, `&`, `"`, etc. qui permettent de représenter des balises et des entités SGML/HTML.

```

1 Si la balise &lt;HEAD&gt; n'a pas
2 &eacute;t&eacute; utilisée,
3 il faut le signaler &agrave;
4 l'auteur.

```

Si la balise `<HEAD>` n'a pas été utilisée, il faut le signaler à l'auteur.

Certains auteurs de documents HTML saisissent leur texte en 8 bits, leurs documents ne sont dès lors plus portables car les caractères accentués ne sont pas codés de la même manière sur un ordinateur de type PC ou un Macintosh ou sur des stations de travail Unix. Une solution au problème consiste à saisir les documents HTML avec les caractères accentués de la plateforme locale puis d'utiliser un programme dédié au transcodage comme

recode³ qui, dans sa version 3,5, propose une bonne traduction depuis et vers n'importe quel codage de caractères y compris le format des entités SGML et le format (L^A)T_EX (voir aussi l'article de M. GOOSSENS sur la traduction des documents L^AT_EX, dans ce même numéro).

9. Classement des éléments HTML

HTML est un langage qui évolue. Peu à peu, des éléments y sont ajoutés de manière à offrir plus de facilités et des fonctions nouvelles. Les différents éléments de HTML sont répartis sur plusieurs niveaux (actuellement de 0 à 3). Les éléments les plus courants sont au niveau 0 (listes, liens hypertextes, structures principales, entités, etc.) ; il s'agit du niveau de correspondance minimale avec les intentions de l'auteur. Au niveau 1 sont groupés les éléments du niveau 0 plus les images et les styles de présentation ; là peuvent commencer à se manifester des différences par rapport aux intentions de l'auteur selon les logiciels de visualisation utilisés (images non affichées, polices non disponibles, etc). Ensuite, le niveau 2 contient le niveau 1 plus des possibilités introduites récemment ; dans l'état d'avancement des travaux sur l'évolution de HTML, le niveau 2 ne contient actuellement que les formulaires. Enfin, des propositions concernant des nouveaux éléments (tableaux, mathématiques, etc.) sont contenues dans le niveau 3 qui correspond à ce que l'on appelle HTML+.

Dans cet article, nous nous sommes limités aux éléments des niveaux 0 et 1 de HTML 2.0. Voir l'article à suivre de F. Dagorn pour HTML niveau 2.

Bibliographie

- [BERNERS-LEE *et al.*, 1994] T. BERNERS-LEE, D. CONNOLLY et K. MULDRON. – Hypertext Markup Language (HTML) version 2.0. – *Request for Comments* Internet (<ftp://www.ics.uci.edu/pub/ietf/html/htmlspec2.ps.gz>), octobre 1994.
- [RAGGETT, 1993] D. RAGGETT. – HTML+ (*hypertext markup format*). – *Draft* Internet (<ftp://curia.ucc.ie/pub/sgml/htmlplus.ps>), octobre 1993. Document de discussion.
- [RAGGETT, 1994] D. RAGGETT. – A review of the HTML+ document format. *First International Word-Wide Web Conference*. – mai 1994.

3. Logiciel réalisé par F. Pinard et distribué dans le domaine public parmi les logiciels GNU. Une version peut être transférée par ftp depuis l'adresse [prep.ai.mit.edu](ftp://prep.ai.mit.edu), répertoire `/pub/gnu`.

```
<HEAD>
<TITLE> Association GUTenberg </TITLE>
</HEAD>

<BODY>
<IMG SRC="logogut.gif" ALT="Association GUTenberg"><BR><HR>
<H2> Des distributions de <IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"></H2>
<P>GUTenberg distribue des versions multilingues de
<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"> pour les
environnements suivants :</P>
<UL>
<LI> Unix, version bas&eacute;e sur le
    M1<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"> de M. FERGUSON ;
    ex&eacute;cutables disponibles pour Sun, AIX, HP, SGI,
    DEC Alpha, etc. ;
<LI> PC : version bas&eacute;e sur le moteur
    Em<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"> de E. MATTES,
    avec de nombreux outils graphiques et PostScript ;
<LI> Macintosh : Euro-Oz<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm">,
    version multilingue du logiciel
    Oz<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"> de A. TREVORROW.
</UL>
<P>Ces distributions sont disponibles soit sur support physique
(cartouche ou disquettes), soit &agrave; travers le r&eacute;seau
sur le serveur de Rennes (voir plus loin).</P>
<!-- Texte coupé -->
<H2> Un serveur de fichiers</H2>
<P>Physiquement install&eacute; sur une machine du CRI-CICB de Rennes,
le serveur de fichiers de GUTenberg est accessible par la commande
FTP, &agrave; l'adresse :
<A HREF="ftp://ftp.univ-rennes1.fr/pub/GUTenberg/">
ftp.univ-rennes1.fr</A>,
login <KBD>anonymous</KBD>
sous le r&eacute;pertoire : <KBD>/pub/GUTenberg</KBD>.</P>
<P>Les distributions GUTenberg sont disponibles sur ce serveur,
ainsi que les fichiers d'archive de la messagerie GUT, le style
<KBD>french.sty</KBD>, etc.</P>
<P>Une version source de la famille des produits
<IMG ALIGN=BOTTOM ALT="" SRC="tex.xbm"> est en cours
d'&eacute;laboration, pour une installation simplifi&eacute;e
par les outils du type <KBD>configure</KBD>.</P>
<!-- Texte coupé -->
<ADDRESS>
Pour GUTenberg : <BR>
<EM>Christian.Rolland@irisa.fr <BR>
Derni&egrave;re mise &agrave; jour : vendredi 18 mars 1994</EM>
</ADDRESS>
</BODY>
```

FIGURE 1 – Exemple de document HTML

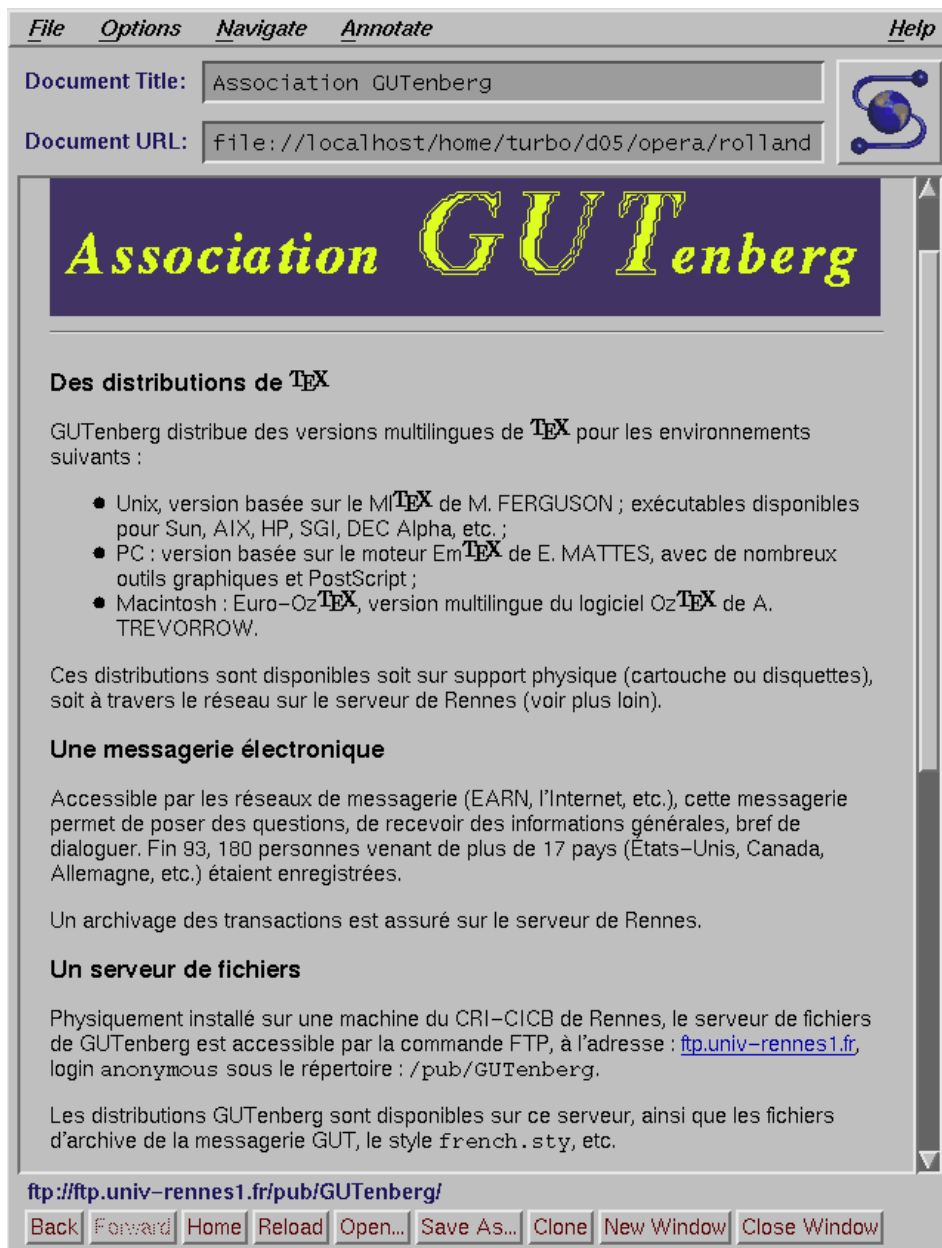


FIGURE 2 – Visualisation par Mosaic du document de la figure 1

Quelques pointeurs supplémentaires

<http://web.urec.fr/docs/WWW/WWW.html>

World-Wide Web, par F. DAGORN et C. GROSS

<http://www.hal.com/%7Econnolly/html-spec/>

Informations sur la spécification de HTML par D. CONOLLY.

<http://info.cern.ch/hypertext/WWW/MarkUp/MarkUp.html>

Informations diverses sur HTML, collectées par T. BERNERS-LEE.

<http://www.ics.uci.edu/pub/ietf/html/index.html>

Groupe de travail IETF *HyperText Markup Language* (HTML), par R. FIELDING.

<http://www.willamette.edu/html-composition/strict-html.html>

Comment bien écrire des documents HTML, par J. E. TILTON.

<http://www.ncsa.uiuc.edu/demoweb/html-primer.html>

Guide du débutant en HTML, par M. ANDREESSEN.

<http://kcgl1.eng.ohio-state.edu/www/doc/html/doc.html>

Comment écrire des fichiers HTML, par P. FLYNN.

<http://www.utirc.utoronto.ca/HTMLdocs/NewHTML/htmlindex.html>

Documentation HTML, par I. GRAHAM.

Annexe : Table des entités HTML

Caractère	Entité	Description
<i>Symboles</i>		
<	<	signe inférieur à
>	>	signe supérieur à
&	&	et commercial
"	"	apostrophe double
␣	 	espace insécable
.	­	tiret invisible
Suite du tableau page suivante...		

Caractère	Entité	Description
<i>Caractères accentués, étrangers et ligatures</i>		
Æ	Æ	AE majuscule, diphtonge (ligature)
Á	Á	A majuscule, accent aigu
Â	Â	A majuscule, accent circonflexe
À	À	A majuscule, accent grave
Å	Å	A majuscule, rond
Ã	Ã	A majuscule, tilde
Ä	Ä	A majuscule, tréma ou <i>umlaut</i>
Ç	Ç	C majuscule, cédille
Ð	Ð	ETH majuscule, islandais
É	É	E majuscule, accent aigu
Ê	Ê	E majuscule, accent circonflexe
È	È	E majuscule, accent grave
Ë	Ë	E majuscule, tréma ou <i>umlaut</i>
Í	Í	I majuscule, accent aigu
Î	Î	I majuscule, accent circonflexe
Ì	Ì	I majuscule, accent grave
Ï	Ï	I majuscule, tréma ou <i>umlaut</i>
Ñ	Ñ	N majuscule, tilde
Ó	Ó	O majuscule, accent aigu
Ô	Ô	O majuscule, accent circonflexe
Ò	Ò	O majuscule, accent grave
Ø	Ø	O majuscule, barré
Ö	Õ	O majuscule, tilde
Ö	Ö	O majuscule, tréma ou <i>umlaut</i>
Ð	Þ	THORN majuscule, islandais
Ú	Ú	U majuscule, accent aigu
Û	Û	U majuscule, accent circonflexe
Ù	Ù	U majuscule, accent grave
Û	Ü	U majuscule, tréma ou <i>umlaut</i>
Ý	Ý	Y majuscule, accent aigu
á	á	a minuscule, accent aigu
â	â	a minuscule, accent circonflexe
æ	æ	minuscule ae, diphtonge (ligature)
à	à	a minuscule, accent grave
å	å	a minuscule, rond
Suite du tableau page suivante...		

Caractère	Entité	Description
ã	ã	a minuscule, tilde
ä	ä	a minuscule, tréma ou <i>umlaut</i>
ç	ç	c minuscule, cédille
é	é	e minuscule, accent aigu
ê	ê	e minuscule, accent circonflexe
è	è	e minuscule, accent grave
ð	ð	eth minuscule, islandais
ë	ë	e minuscule, tréma ou <i>umlaut</i>
í	í	i minuscule, accent aigu
î	î	i minuscule, accent circonflexe
ì	ì	i minuscule, accent grave
ï	ï	i minuscule, tréma ou <i>umlaut</i>
ñ	ñ	n minuscule, tilde
ó	ó	o minuscule, accent aigu
ô	ô	o minuscule, accent circonflexe
ò	ò	o minuscule, accent grave
ø	ø	o minuscule, barré
õ	õ	o minuscule, tilde
ö	ö	o minuscule, tréma ou <i>umlaut</i>
ß	ß	ss minuscules, allemand (ligature sz)
þ	þ	thorn minuscule, islandais
ú	ú	u minuscule, accent aigu
û	û	u minuscule, accent circonflexe
ù	ù	u minuscule, accent grave
ü	ü	u minuscule, tréma ou <i>umlaut</i>
ý	ý	y minuscule, accent aigu
ÿ	ÿ	y minuscule, tréma ou <i>umlaut</i>