



W3C

Compact HTML for Small Information Appliances

W3C NOTE 09-Feb-1998

This version:

<http://www.w3.org/TR/1998/NOTE-compactHTML-19980209>

Latest version:

<http://www.w3.org/TR/1998/NOTE-compactHTML-19980209>

Author:

Tomihisa Kamada, [ACCESS Co.,Ltd., tomy@access.co.jp](mailto:tomy@access.co.jp)

Status of this Document

This document is a NOTE made available by the W3 Consortium for discussion only. This indicates no endorsement of its content, nor that the Consortium has had any editorial control in its preparation, nor that the Consortium has, is, or will be allocating any resources to the issues addressed by the NOTE.

This document is a submission to W3C. Please see <http://www.w3.org/Submission/> Acknowledged Submissions to W3C regarding its disposition.

Abstract

The Internet infrastructure has been developed all over the world, and nowadays there are a variety of devices equipped with the Internet-access function, from TV sets to wireless cellular phones. The HyperText Markup Language (HTML) is widely accepted and spread as the standard of the WWW(World Wide Web) document format. The "Compact HTML" proposed here defines a subset of HTML for small information appliances such as smart phones, smart communicators, mobile PDAs, and etc. Such a certain level of HTML is strongly required as a guideline from the manufacturers of small information devices, service providers, carriers, and software developers. Since "Compact HTML" is completely based on the HTML recommendations, we can use millions of HTML-based content resources, various software tools, and public materials (textbooks, magazines, and web information).

Contents

1. [Introduction](#)
2. [Requirements of Small Information Appliances](#)
 1. [Scope of the Products](#)
 2. [Requirements](#)
 3. [Wireless Network](#)
3. [Definition of Compact HTML](#)
 1. [Design Principles](#)

2. [Features of Compact HTML](#)
3. [Detail Definition](#)
4. [Examples and Benefits of Compact HTML](#)
 1. [Examples](#)
 2. [Benefits of Compact HTML](#)
 3. [Another Approach](#)
5. [Concluding Remarks](#)

[References](#)

[Acknowledgments](#)

[Appendix A. Compact HTML Tag List](#)

[Appendix B. Compact HTML DTD](#)

1. Introduction

The Compact HTML is a well-defined subset of HTML 2.0[1], HTML 3.2[2] and HTML 4.0[3] recommendations, which is designed for small information appliances. HTML defines flexible, portable, and practical document format for the documents on the Internet. One direction of HTML is to grow toward richer multimedia document format. A new recommendation HTML 4.0[3] includes new additional features. For example, CSS(Cascading Style Sheets) give a wider range of document styles. On the other hand, there must be another direction for small information appliances. Small information appliances have several hardware restrictions such as small memory, low power CPU, small or no secondary storage, small display, mono-color, single character font, and restricted input method (no keyboard and mouse). The browser for Compact HTML proposed in this document can be implemented in such a restricted environment. Once such a subset of HTML is defined, contents providers and information appliance manufacturers can rely on this common standard. We believe that Compact HTML definitely contributes to the rapid growth of small information appliance market.

2. Requirements of Small Information Appliances

2.1. Scope of the Products

First we describe the scope of target small information appliances. The categories of these devices are often referred to as smart phones, smart communicators, and mobile PDAs. There are some hardware restrictions for these devices. From the hardware point of view, we pick up main characteristics of the target devices.



- Small memory
Typical case: 128-512Kbytes RAM and 512K-1Mbytes ROM
- Low power CPU
Typical case: 1-10 MIPS class CPU for embedded systems
- Small display
Typical case: 50x30 dots, 100x72 dots, and 150x100 dots
- Restricted colors
Typical case: mono-color (black and white)
- Restricted character fonts
Typical case: only single font
- Restricted input method
Typical case: several control buttons and number buttons (0-9)

The picture shows an example of cellular phone which has the HTML browsing function. A wide variety of content services are potentially possible via wireless networks.

2.2. Requirements

To realize the WWW browsing function for such small devices, a suitable subset of HTML is necessary. The requirements are derived from the above hardware restrictions. Also these devices should be easy to use from the standpoint as consumer products. The browser software for a subset of HTML should run within the small memory: e.g., 150-200Kbytes for the working data and also 150-200Kbytes for the program code. The minimum requirement for the CPU power should be 1-2 MIPS, though it may depend on the CPU power required for network communication processing. Easy navigation is also one of the key features for consumer devices. It means that the users can navigate information with a minimum number of operations. A subset of HTML should satisfy this requirement.

2.3. Wireless Network

Compact HTML does not depend on the underlying network protocol. In the typical cases, the

transport protocol for Compact HTML is assumed to be HTTP over TCP/IP. However, current wireless communication networking for cellular phones is low band and low speed. In this area, the transport protocol should be defined as light protocol for better performance on the physical packet layer. It also seems useful to compress HTML contents so that most of HTML data can be stored within one packet data.

3. Definition of Compact HTML

3.1 Design Principles

The Compact HTML is designed to meet the requirements of small information appliances described above. It is designed based on the following four principles.

(1) Completely based on the current HTML W3C recommendations

Compact HTML is defined as a subset of HTML 2.0, HTML 3.2 and HTML 4.0 specifications. This means that Compact HTML inherits the flexibility and portability from the standard HTML.

(2) Lite Specification

Compact HTML has to be implemented with small memory and low power CPU. Frames and tables which require large memory are excluded from Compact HTML.

(3) Can be viewed on a small mono-color display

Compact HTML assumes a small display space of black and white color. However, it does not assume a fixed display space, but it is flexible for the display screen size. Compact HTML also assumes single character font.

(4) Can be easily operated by the users

Compact HTML is defined so that all the basic operations can be done by a combination of four buttons; Cursor forward, Cursor backward, Select, and Back/Stop(Return to the previous page). The functions which require two-dimensional focus pointing like "image map" and "table" are excluded from Compact HTML.

The definition of Compact HTML is derived straightforwardly from the above principles.

3.2 Features of Compact HTML

The Compact HTML is a subset of HTML 2.0, HTML 3.2 and HTML 4.0. We describe the major features which are excluded from Compact HTML, as follows.

- JPEG image
- Table
- Image map
- Multiple character fonts and styles
- Background color and image
- Frame
- Style sheet

We define that Compact HTML includes GIF image support. It should be noted that this subset does not require two-dimensional cursor moving, that is, it can be operated by using only four buttons. We can also expect that well-designed pages for small display fit the screen space and the scrolling is not necessary. Actually the Compact HTML browser can display the pages like "deck of cards" by HDML[4].

Since the memory capacity is the most important issue in implementing the Compact HTML

browser, we recommend the buffer limit for some functions.

- INPUT

The maximum buffer size is 512 bytes.

- SELECT

The maximum buffer size is 4096 bytes.

Though such a limitation belongs to the implementation issues, the common criteria is useful while developing devices.

One recommended implementation for the browser is to support the direct selection of anchors by using number buttons. For example, when five anchors are contained in an HTML page, the third anchor can be selected just by pressing the "3" button. (The HTML 4.0 specification includes a new attribute "accesskey" for the similar purpose of direct key assignment.)

3.3 Detail Definition

The complete list of tags supported by Compact HTML is described in Appendix A. The comparison with HTML 2.0, HTML 3.2 and HTML 4.0 is marked in the table. The document type definition (DTD) for Compact HTML is also described in Appendix B. This gives the intended interpretation of Compact HTML elements. The document type is defined as follows;

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD Compact HTML 1.0 Draft//EN">
```

4. Examples and Benefits of Compact HTML

4.1 Examples

Here we describe the examples of applications by using Compact HTML. The following examples show the compact browser for cellular phones. The screen is the space of 7 text lines and 16 characters wide. The top line is used for displaying the status information.

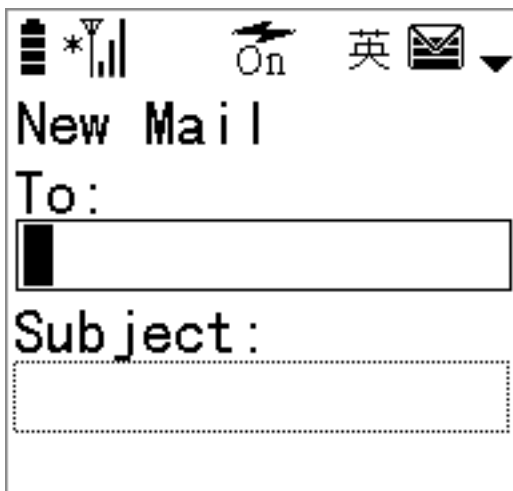
(1) Compact HTML example: Simple Menu

In this example, the cursor focus point is expressed as the reverse text.



(2) Compact HTML example: Mail Send Form

This example shows the mail sending form using INPUT tags. The focused form is expressed as solid surrounding lines, and non-focused forms are expressed as dotted surrounding line. The cursor point for input characters is expressed as a reverse box.



■ * | | | On 英 [Envelope] ▼

New Mail

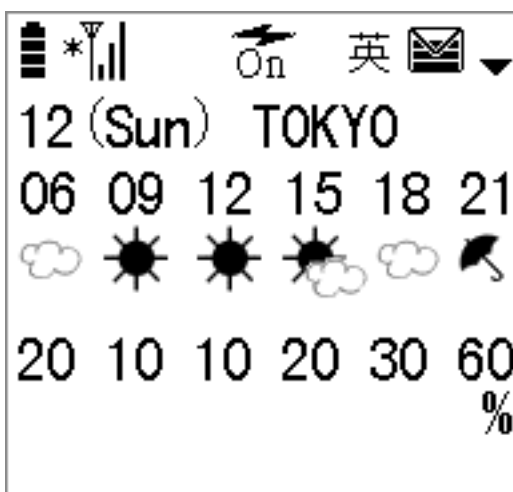
To:

█

Subject:

(3) Compact HTML example: Image Contents

This example shows weather and rain information of the day. It uses mono-color GIF image.



Practical implementations and experiments show that Compact HTML is enough useful for small screen of 5-10 text lines and 10-20 characters wide.

4.2 Benefits of Compact HTML

The Compact HTML, an HTML-based approach, guarantees that small information appliances can connect to the open WWW world. Compact HTML keeps the advantage of HTML features and solves the problems arising from the restrictions of small information appliances.

The Compact HTML specification can be referred to by the tools like HTML authoring systems. In addition, the client-specific web services for such small devices can be realized by using user agent attributes [5]. That is, the server can do the content filter for Compact HTML.

4.3 Another Approach

There may be another approach which is NOT based on HTML standards. The approach of a new language may be accepted in a certain closed service. For example, a language named HDML(Handheld Device Markup Language)[4] is proposed for the mobile handheld devices. The goal of HDML is very similar to the one of Compact HTML. It seems useful for a class of handheld devices. However, the disadvantage of special language approach can be said that everything such as contents, authoring tools, server software, client software, and textbooks have to be prepared. Especially thinking about a product line from high-end PDAs to low-end cellular phones, the consistent HTML-based approach would make sense.

5. Concluding Remarks

We proposed the Compact HTML for small information appliances. This contributes to the WWW community, especially for wireless and mobile small devices. Compact HTML can be referred as a recommended guideline for HTML services and for HTML browsing software in this area.

The Internet world is growing very fast. Several levels will be likely required for Compact HTML in the near future. It is important to evolve the functionality and extend the specification flexibly, according to the requirements from the growing market.

In the early 21st century, the world-wide standard for wireless digital phone networking protocols will be established. The band width of the network is expected to be wide enough for motion picture (video) communications. Compact HTML should include such advanced features in this next generation of wireless network.

References

- [1] Tim Berners-Lee and Dan Connolly, The HyperText Markup Language (HTML) Specification Version 2.0 (RFC 1866), Nov. 1995.
- [2] Dave Raggett, HTML 3.2 Reference Specification, W3C Recommendation. Jan. 1997.
- [3] Dave Raggett, Arnaud Le Hors, and Ian Jacobs, HTML 4.0 Reference Specification, W3C Recommendation, Dec. 1997.
- [4] Tim Hyland, Proposal for a Handheld Device Markup Language (HDML), May. 1997.
- [5] Tomihisa Kamada and Tomohiko Miyazaki, Client-Specific Web Services by Using User Agent Attributes, W3C NOTE, Dec. 1997.

Acknowledgments

We would like to thank to Professors Nobuo Saito and Tatsuya Hagino (KEIO University) and many W3C members for their supports and suggestions. We are also grateful to NTT DoCoMo for their supports to Compact HTML for wireless cellular phones.

Appendix A. Compact HTML Tag List

--Updated 1998/01/27
 --HTML(2.0:HTML2.0, 3.2:HTML3.2, 4.0:HTML4.0)
 --CH : Compact HTML

No	Elements	Attributes	HTML	CH	Comments
1	!-	-	2.0	CH	--
2	!DOCTYPE	-	2.0	CH	--

3	&xxx;	-	2.0	CH	--& ,© ,> ,< ," ,® , ; --� ~
4	A	name= href="URL" rel= rev= title= urn=(deleted from HTML3.2) methods=(deleted from HTML3.2)	2.0	CH CH - - - -	--
5	ABBR	-	4.0	-	--
6	ACRONYM	-	4.0	-	--
7	ADDRESS	-	2.0	-	--Only one font.
8	APPLET	-	3.2	-	--(Deprecated element in HTML4.0)
9	AREA	shape= coords= href="URL" alt= nohref	3.2	-	--
10	B	-	2.0	-	--Only one font.
11	BASE	href="URL"	2.0	CH	--
12	BASEFONT	size=	3.2	-	--Only one font. --(Deprecated element in HTML4.0)
13	BDO	-	4.0	-	--
14	BIG	-	3.2	-	--Only one font.
15	BLOCKQUOTE	-	3.2	CH	--
16	BODY	- bgcolor= background= text= link= vlink= alink=	2.0 3.2 3.2 3.2 3.2 3.2	CH - - - - -	--Non-white colors are drawn as black.
17	BR	- clear=all/left/right	2.0 3.2	CH CH	--
18	BUTTON	-	4.0	-	--
19	CAPTION	-	3.2	-	--
20	CENTER	-	3.2	CH	--(Deprecated element in HTML4.0)
21	CITE	-	2.0	-	--Only one font.
22	CODE	-	2.0	-	--Only one font.
23	COL	-	4.0	-	--
24	COLGROUP	-	4.0	-	--
25	DD	-	2.0	CH	--

26	DEL	-	4.0	-	--
27	DFN	-	3.2	-	--
28	DIR	- compact	2.0	CH -	--(Deprecated element in HTML4.0)
29	DIV	- align=left/center/right	3.2	CH CH	--
30	DL	- compact	2.0	CH -	--
31	DT	-	2.0	CH	--
32	EM	-	2.0	-	--Only one font.
33	FIELDSET	-	4.0	-	--
34	FONT	size=n size=+n/-n color=	3.2	- - -	--Only one font. --(Deprecated element in HTML4.0)
35	FORM	action= method=get/post enctype=	2.0	CH CH CH	--
36	FRAME	-	4.0	-	--(Frameset DTD)
37	FRAMESET	-	4.0	-	--(Frameset DTD)
38	HEAD	-	2.0	CH	--
39	Hn	- align=left/center/right	2.0 3.2	CH CH	--
40	HR	- align=left/center/right size= width= noshade	2.0 3.2 3.2 3.2 3.2	CH CH CH CH CH	--
41	HTML	- version=	2.0 3.2	CH CH	--version="C-HTML 1.0".
42	I	-	2.0	-	--Only one font.
43	IFRAME	-	4.0	-	--(Frameset DTD)
44	IMG	src= align=top/middle/bottom align=left/right width= height= hspace= vspace= alt= border= usemap= ismap=	2.0 2.0 3.2 3.2 3.2 3.2 3.2 2.0 3.2 3.2 2.0	CH CH CH CH CH CH CH CH - -	--Large images compressed automatically.
		type=text name= size= maxlength= value=	2.0	CH CH CH CH	--Max character buffer 512 bytes.

45	INPUT	type=password name= size= maxLength= value=	2.0	CH CH CH CH	--
		type=checkbox name= value= checked	2.0	CH CH CH CH	--
		type=radio name= value= checked	2.0	CH CH CH CH	--
		type=hidden name= value=	2.0	CH CH CH	--
		type=image name= src= align=top/middle/bottom/ left/right	2.0 2.0 2.0 3.2	-	--
		type=submit name= value=	2.0	CH CH CH	--
		type=reset name= value=	2.0	CH CH CH	--
		type=file name= value=	3.2	-	--
46	INS	-	4.0	-	--
47	ISINDEX	prompt=	2.0 3.2	- -	--(Deprecated element in HTML4.0)
48	KBD	-	2.0	-	--Only one font.
49	LABEL	-	4.0	-	--
50	LEGEND	-	4.0	-	--
51	LI	-	2.0	CH	
		type=1/A/a/I/i	3.2	-	--
		type=circle/disc/square	3.2	-	
		value=	3.2	-	
52	LINK	href="URL" rel= rev= urn= methods= title= id=	2.0	-	--
53	LISTING	-	2.0	-	--Only one font. --(Obsolete element in HTML4.0)

54	MAP	name=	3.2	-	--
55	MENU	- compact	2.0	CH -	--(Deprecated element in HTML4.0)
56	META	name= http-equiv= content=	2.0	CH	--http-equiv="refresh" only.
57	NEXTID	n=	2.0	-	--Deleted from HTML3.2.
58	NOFRAMES	-	4.0	-	--(Frameset DTD)
59	NOSCRIPT	-	4.0	-	--
60	OBJECT	-	4.0	-	--
61	OL	- type=1/A/a/I/I/i start= compact	2.0	CH	
			3.2	-	
			3.2	-	
			2.0	-	
62	OPTGROUP	-	4.0	-	--
63	OPTION	- selected value=	2.0	CH	--
				CH	
				-	
64	P	- align=left/center/right	2.0	CH	--
			3.2	CH	
65	PARAM	-	4.0	-	--
66	PLAINTEXT	-	2.0	CH	--(Obsolete element in HTML4.0)
67	PRE	- width=	2.0	CH	--
			3.2	-	
68	Q	-	4.0	-	--
69	S	-	2.0	-	--(Deprecated element in HTML4.0)
70	SAMP	-	2.0	-	--Only one font.
71	SCRIPT	-	3.2	-	--
72	SELECT	name= size= multiple	2.0	CH	--Max character buffer 4 Kbytes.
				CH	
				CH	
73	SMALL	-	3.2	-	--Only one font.
74	SPAN	-	4.0	-	--
75	STRIKE	-	2.0	-	--(Deprecated element in HTML4.0)
76	STRONG	-	2.0	-	--Only one font.
77	STYLE	-	2.0	-	--
78	SUB	-	3.2	-	--
79	SUP	-	3.2	-	--
80	TABLE	- align=left/center/right etc. border= width= cellspacing=	3.2	-	--

		cellpadding=			
81	TBODY	-	4.0	-	--
82	TD	- align=left/center/right valign=top/middle/bottom/ baseline rowspan= colspan= width= height= nowrap	3.2	-	--
83	TEXTAREA	name= rows= cols=	2.0	CH CH CH	--Max character buffer 512 bytes.
84	TFOOT	-	4.0	-	--
85	TH	- align=left/center/right valign=top/middle/bottom/ baseline rowspan= colspan= width= height= nowrap	3.2	-	--
86	THEAD	-	4.0	-	--
87	TITLE	-	2.0	CH	--
88	TR	- align=left/center/right valign=top/middle/bottom/ baseline	3.2	-	--
89	TT	-	2.0	-	--Only one font.
90	U	-	3.2	-	--(Deprecated element in HTML4.0)
91	UL	- type=disk/circle/square compact	2.0 3.2 2.0	CH - -	--
92	VAR	-	2.0	-	--Only one font.
93	XMP	-	2.0	-	--Only one font. --(Obsolete element in HTML4.0)

Appendix B. Compact HTML DTD

```
<!-- Compact HTML Document Type Definition -->
```

```
<!--  
    Date: Tuesday November 25th 1997  
    Author: Tomihisa Kamada <tomy@access.co.jp>  
-->
```

```
<!ENTITY % HTML.Version  
    "-//W3C//DTD Compact HTML 1.0 Draft//EN"  
>
```

```

<!--===== Deprecated Features Switch =====>
<!ENTITY % HTML.Deprecated "INCLUDE">
<!--===== Imported Names =====>
<!ENTITY % Content-Type "CDATA">
<!ENTITY % HTTP-Method "GET | POST">
<!ENTITY % URL "CDATA">
<!-- Parameter Entities -->
<!ENTITY % heading "H1|H2|H3|H4|H5|H6">
<!ENTITY % list "UL | OL | DIR | MENU">
<!ENTITY % preformatted "PRE">
<!--===== Character mnemonic entities =====>
<!ENTITY % ISOlat1 PUBLIC
    "ISO 8879-1986//ENTITIES Added Latin 1//EN//HTML">
%ISOlat1;
<!--===== Entities for special symbols =====>
<!ENTITY amp    CDATA "&"  -- ampersand    -->
<!ENTITY gt     CDATA ">"  -- greater than -->
<!ENTITY lt     CDATA "<"  -- less than   -->
<!--===== Text Markup =====>
<!ENTITY % phrase "DFN">
<!ENTITY % special "A | IMG | BR ">
<!ENTITY % form "INPUT | SELECT | TEXTAREA">
<!ENTITY % text "#PCDATA | %phrase | %special | %form">
<!ELEMENT (%phrase) - - (%text)*>
<!ELEMENT BR      - O EMPTY>
<!ATTLIST BR
    clear (left|all|right|none) none
    >
<!--===== HTML content models =====>
<!ENTITY % block
    "P | %list | %preformatted | DL | DIV | CENTER |
    BLOCKQUOTE | FORM | HR ">
<!ENTITY % flow "(%text | %block)*">
<!--===== Document Body =====>
<!ENTITY % body.content "(%heading | %text | %block )*">
<!ELEMENT BODY 0 0 %body.content>
<!ELEMENT DIV  - - %body.content>
<!ATTLIST DIV
    align (left|center|right) #IMPLIED
    >

```

<!ELEMENT center - - %body.content>

<!--===== The Anchor Element =====-->

<!ELEMENT A - - (%text)* -(A)>

<!ATTLIST A

name CDATA #IMPLIED

href %URL #IMPLIED

>

<!--===== Images =====-->

<!ENTITY % Length "CDATA">

<!ENTITY % Pixels "NUMBER">

<!ENTITY % IAlign "(top|middle|bottom|left|right)">

<!ELEMENT IMG - 0 EMPTY>

<!ATTLIST IMG

src %URL #REQUIRED

align %IAlign #IMPLIED

width %Pixels #IMPLIED

height %Pixels #IMPLIED

hspace %Pixels #IMPLIED

vspace %Pixels #IMPLIED

alt CDATA #IMPLIED

border %Pixels #IMPLIED

>

<!--===== Horizontal Rule =====-->

<!ELEMENT HR - 0 EMPTY>

<!ATTLIST HR

align (left|right|center) #IMPLIED

size %Pixels #IMPLIED

width %Length #IMPLIED

noshade (noshade) #IMPLIED

>

<!--===== Paragraphs=====-->

<!ELEMENT P - 0 (%text)*>

<!ATTLIST P

align (left|center|right) #IMPLIED

>

<!--===== Headings =====-->

<!ELEMENT (%heading) - - (%text;)*>

<!ATTLIST (%heading)

align (left|center|right) #IMPLIED

>

<!--===== Preformatted Text =====-->

<!ENTITY % pre.exclusion "IMG">

<!ELEMENT PRE - - (%text)* -(%pre.exclusion)>

<!--===== Block-like Quotes =====-->

<!ELEMENT BLOCKQUOTE - - %body.content>

<!--===== Lists =====-->

<!ELEMENT DL - - (DT|DD)+>

<!ELEMENT DT - 0 (%text)*>

<!ELEMENT DD - 0 %flow;>

<!ELEMENT (OL|UL) - - (LI)+>

<!ELEMENT (DIR|MENU) - - (LI)+ -(%block)>

<!ELEMENT LI - 0 %flow>

<!--===== Forms =====-->

<!ELEMENT FORM - - %body.content -(FORM)>

<!ATTLIST FORM
action %URL #IMPLIED
method (%HTTP-Method) GET
enctype %Content-Type; "application/x-www-form-urlencoded"
>

<!ENTITY % InputType
"(TEXT | PASSWORD | CHECKBOX | RADIO | HIDDEN
| IMAGE | SUBMIT | RESET)">

<!ELEMENT INPUT - 0 EMPTY>

<!ATTLIST INPUT
type %InputType TEXT
name CDATA #IMPLIED
value CDATA #IMPLIED
checked (checked) #IMPLIED
size CDATA #IMPLIED
maxlength NUMBER #IMPLIED
src %URL #IMPLIED
align %IAAlign #IMPLIED
>

<!ELEMENT SELECT - - (OPTION+)>

<!ATTLIST SELECT
name CDATA #REQUIRED
size NUMBER #IMPLIED
multiple (multiple) #IMPLIED
>

<!ELEMENT OPTION - 0 (#PCDATA)*>

<!ATTLIST OPTION
selected (selected) #IMPLIED
value CDATA #IMPLIED
>

<!ELEMENT TEXTAREA - - (#PCDATA)*>

<!ATTLIST TEXTAREA
name CDATA #REQUIRED
rows NUMBER #REQUIRED
cols NUMBER #REQUIRED
>

<!--===== Document Head =====-->

<!ENTITY % head.content "TITLE & ISINDEX? & BASE?">

<!ELEMENT HEAD 0 0 (%head.content)>

<!ELEMENT TITLE - - (#PCDATA)*>

<!ELEMENT BASE - 0 EMPTY>

<!ATTLIST BASE
href %URL #REQUIRED
>

<!ELEMENT META - 0 EMPTY>

<!ATTLIST META
http-equiv NAME #IMPLIED
>

<!--===== Document Structure =====-->

<!ENTITY % version.attr "VERSION CDATA #FIXED '%HTML.Version;'">

<!ENTITY % html.content "HEAD, BODY">

<!ELEMENT HTML 0 0 (%html.content)>

<!ATTLIST HTML
 %version.attr;
>

<!--===== End of DTD =====-->