

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

张宏 张亦弛 编著
飞思科技产品研发中心 监制

HTML 代码优化 原理与技巧

电子工业出版社

House of Electronics Industry

北京·BEIJING

95

内 容 简 介

本书介绍了 70 多种有关优化网页 HTML 代码的原理与技巧, 内容涉及 HTML、图像、多媒体、表格、报头、网址、JavaScript、CSS、网站等十个方面的优化问题。相信读者能够利用这些技巧, 减少网页的字节数, 提高网页的显示速度, 以达到“网页小、速度快”的目的。

本书适合于中高级网页制作者和网络管理者。

未经许可, 不得以任何方式复制或抄袭本书的部分或全部内容。

版权所有, 侵权必究。

图书在版编目 (CIP) 数据

HTML 代码优化原理与技巧 / 张宏, 张亦弛编著. —北京: 电子工业出版社, 2003.9
ISBN 7-5053-9019-8

I. H... II. ①张... ②张... III. 超文本标记语言, XHTML—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2003) 第 071354 号

责任编辑: 陆舒敏

印 刷: 北京中科印刷有限公司

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京海淀区万寿路 173 信箱 邮编: 100036

经 销: 各地新华书店

开 本: 880×1230 1/32 印张: 6.625 字数: 169.6 千字

版 次: 2003 年 9 月第 1 版 2003 年 9 月第 1 次印刷

印 数: 6 000 册 定价: 15.00 元

凡购买电子工业出版社的图书, 如有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系调换。联系电话: (010) 68279077

前 言

在我国，人们在制作网页时，通常使用一个 WYSIWYG（所见即所得）的网页制作软件，例如 FrontPage 和 Dreamweaver。从技术角度看，用这些软件制作的网页有一个共同的缺点，那就是 HTML 代码过于庞大。在本书第 1 章第 1 节中我们举了一个例子：有一个网页，上面只有“你好”两个汉字，用手工方法制作这个网页，HTML 代码只有 8 行，大小为 76 字节。而用 Word(XP 简体中文版)制作这个网页，HTML 代码竟然有 127 行，是手工方法的 16 倍，大小为 3961 字节，是手工方法的 52 倍。

HTML 代码过于庞大，必然造成网页显示速度下降。假设网络传输速度是每秒 1KB，那么第一个“你好”网页最多需要 1 秒钟即可显示出来，而第二个“你好”网页最少需要 3 秒钟才能显示出来。在我国，绝大多数上网者仍然使用调制解调器，以电话拨号方式上网，上网速度是比较慢的。在这种情况下，网页显示速度下降，不仅浪费他们的时间，而且浪费他们的金钱。

针对这个问题，我们编写了本书。书中介绍了 70 多种有关优化网页 HTML 代码的原理与技巧，内容涉及 HTML、图像、多媒体、表格、报头、网址、JavaScript、CSS、网站等十个方面的优化问题。我们希望读者能够利用这些技巧，减少网页的字节数，提高网页的显示速度，以达到“网页小、速度快”的目的。

本书由飞思科技产品研发中心策划并组织编写。本书作者长期从事网络技术工作，书中内容是作者多年经验的总结，文字部分由张宏、张亦弛编写，代码、图像部分由程翼、合欢编写。本书在编写过程中，编者虽然未敢稍有疏虞，但纰缪和不尽人意之处仍在所难免，恳请本书的读者提出宝贵的意见或建议，以便修订并使之更加完善。

我们的联系方式如下：

咨询电话：(010) 68134545 68131648

答疑邮件：support@fecit.com.cn

网 址：<http://www.fecit.com.cn> <http://www.fecit.net>

答疑网址：<http://www.fecit.com.cn/>“问题解答”专区

通用网址：计算机图书、FECIT、飞思教育、飞思科技、飞思

飞思科技产品研发中心

目 录

第 1 章 概述.....	1
1.1 为什么要优化网页的 HTML 代码.....	1
1.2 如何优化网页的 HTML 代码.....	6
1.3 关于本书内容的几点说明.....	9
第 2 章 基本优化方法.....	11
2.1 删除 HTML 代码中可有可无的内容.....	11
2.2 删除网页 HTML 代码中的空语句.....	14
2.3 删除 HTML 语句的默认属性.....	16
2.4 删除 HTML 语句的 lang 属性.....	19
2.5 将简短的注释放入 HTML 语句里面.....	20
2.6 最大限度优化<option>语句.....	22
2.7 用一个<basefont>语句代替一组语句.....	25
2.8 用一个 语句简化一组 语句.....	27
2.9 用 语句使小字号文字的行距恢复正常.....	29
2.10 用短语句替换网页效果完全相同的长语句.....	30
2.11 color 属性最好使用英文单词.....	31
2.12 优先使用特殊字符的第二种编码.....	33
第 3 章 优化图像和多媒体.....	37
3.1 降低图像的分辨率, 减少图像的字节数.....	37
3.2 优先使用三种格式中字节数最少的图像.....	39
3.3 利用图像处理软件缩小图像的尺寸.....	40
3.4 为大幅图像制作高分辨率、低分辨率两个版本.....	43
3.5 将大尺寸画面制作成 GIF 格式 Interlaced 模式的图像.....	46
3.6 将一个大图像分割成多个小图像.....	48
3.7 大、中型网站的常用图像应当重复利用.....	50
3.8 用 alt 属性为图像指定简短的注释.....	51
3.9 用 width 和 height 属性指定图像的宽度和高度.....	53
3.10 用表格或映像替代字节数过大的图像菜单.....	55

3.11	提高图像与主题的关联性	59
3.12	利用 Flash 制作质量高、效果好的动画	60
3.13	优先使用 RealMedia 格式的音频、视频文件	62
3.14	避免使用<bgsound>语句播放网页背景音乐	68
第 4 章	优化表格	71
4.1	将一个大型表格拆分成多个小型表格	71
4.2	四种常用的嵌套表格替代方法	75
4.2.1	用段落代替嵌套的表格	77
4.2.2	用图像代替嵌套的表格	78
4.2.3	用映像代替嵌套的表格	80
4.2.4	用层代替嵌套的表格	83
4.3	用<pre>语句代替表格, 将一组文本框对齐	85
4.4	优先使用 100%宽度的表格	88
4.5	将 align=center/right 属性放入<tr>语句中	90
4.6	将 valign=bottom/top 属性放入<tr>语句中	93
4.7	用<colgroup>语句简化<td>语句的属性	97
4.8	省略</tr>和</td>语句	100
第 5 章	优化网页报头	103
5.1	删除网页报头中的非必需代码	103
5.2	减少网页报头中的注释语句	108
5.3	将网页报头中的公共代码存放外部文件中	110
5.4	将网页报头中的 JavaScript 代码转移到网页的其他部分	116
5.5	在网页报头中加入四条与 spider 有关的 HTML 语句	118
5.6	用<meta>语句指定网页的语言	122
5.7	利用 Prefetching 技术提高网页显示速度	124
5.8	动态内容的网页最好不缓存	125
5.9	提高网站在搜索引擎的排名位置	127
第 6 章	优化网址	131
6.1	用<base>语句将绝对网址转换为相对网址	131
6.2	常用网站的<a>语句可以采用 IP 地址	132
6.3	网址中最后一个目录一定要添加斜线	134
6.4	缩短 Apache 服务器上的过长网址	136

6.5	利用网址重新定向技术使老网站发挥作用	139
第 7 章	优化 JavaScript 和 CSS	143
7.1	在网页中加入检查表单输入错误的功能	143
7.2	不提倡将表单内容发送给一个邮箱	147
7.3	使用 JavaScript 技术来完成站内搜索	151
7.4	用 history.back()语句加快网页后退的速度	154
7.5	告诉访问者如何启用浏览器的 JavaScript 和 Cookie 功能	155
7.6	不要让访问者知道 JavaScript 代码发生错误	157
7.7	尽可能消除 JavaScript 代码中的空白区域	158
7.8	重复利用 CSS 代码和 JavaScript 代码	160
7.9	用 CSS 语句将汉字的大小固定	162
7.10	用一条<style>语句代替一组内容相同的 style 属性	164
7.11	将一组内容相同的 CSS 语句合并为一条语句	166
7.12	优先采用 CSS 语句的简化写法	167
7.13	用 id 属性代替 class 属性	169
第 8 章	优化网站	171
8.1	控制网站主页的大小	171
8.2	查找网站上访问量最高的一组网页	174
8.3	简化网站的任务流程	177
8.4	不要在自己的网站上大量链接其他网站的资源	180
8.5	网站开通之前的测试	181
第 9 章	其他优化方法	185
9.1	大型网页中不适合使用可展开菜单	185
9.2	利用两个帧来模拟 Prefetching 技术	192
9.3	指定网页的背景、前景颜色以避免网页文字不可见	194
9.4	将网页 HTML 代码中的邮箱地址转换成十六进制编码	196
9.5	充分利用 HTML 4.0 的新功能	199

第 1 章 概 述

1.1 为什么要优化网页的 HTML 代码

制作一个网页，通常有两种方法，一种是手工方法，也就是手工录入网页的文字，手工编写网页的 HTML 代码，所有工作都是手工完成的，专业制作者通常采用这种方法；另外一种软件方法，制作者只需要录入网页的文字，网页的 HTML 代码则由一个网页制作软件（例如 FrontPage 或 Dreamweaver）自动生成，业余制作者通常采用这种方法。手工方法的优点是灵活，缺点是制作者必须熟练掌握 HTML 语言；软件方法的优点是简单，缺点是网页制作软件自动生成的 HTML 代码过于庞大。下面我们举一个比较典型的例子来说明制作网页的两种方法。

假设有一个网页，上面只有“你好”这两个汉字。我们首先用手工方法编写这个网页，该网页的 HTML 代码只有 8 行，网页大小为 76 字节，代码如下：

```
<html>
<head>
<title>你好</title>
</head>
<body>
你好
</body>
</html>
```

我们再用 Word 生成这个网页。在 Word (XP 简体中文版) 中输入“你好”这两个汉字，然后以 Web 网页格式保存到磁盘上。该网页的 HTML 代码竟然有 127 行，是手工方法的 16 倍，网页大小为 3961 字节，是手工方法的 52 倍，代码如下：

```
<html xmlns:o="urn:schemas-microsoft-com:office:office"
xmlns:w="urn:schemas-microsoft-com:office:word">
```

```
xmlns="http://www.w3.org/TR/REC-html40">
<head>
<meta http-equiv=Content-Type content="text/html; charset=gb2312">
<meta name=ProgId content="Word.Document">
<meta name=Generator content="Microsoft Word 10">
<meta name=Originator content="Microsoft Word 10">
<link rel=File-List href="你好.files/filelist.xml">
<title>你好</title>
<!--[if gte mso 9]>
<xml>
  <o:DocumentProperties>
    <o:Author>Zhang Hong</o:Author>
    <o:Template>Normal</o:Template>
    <o:LastAuthor>Zhang Hong</o:LastAuthor>
    <o:Revision>1</o:Revision>
    <o:TotalTime>1</o:TotalTime>
    <o:Created>2003-05-22T11:32:00Z</o:Created>
    <o:LastSaved>2003-05-22T11:33:00Z</o:LastSaved>
    <o:Pages>1</o:Pages>
    <o:Characters>2</o:Characters>
    <o:Lines>1</o:Lines>
    <o:Paragraphs>1</o:Paragraphs>
    <o:CharactersWithSpaces>2</o:CharactersWithSpaces>
    <o:Version>10.2625</o:Version>
  </o:DocumentProperties>
</xml>
<![endif]-->
<!--[if gte mso 9]>
<xml>
  <w:WordDocument>
    <w:GrammarState>Clean</w:GrammarState>
    <w:PunctuationKerning/>
    <w:DrawingGridHorizontalSpacing>5.25 磅
    </w:DrawingGridHorizontalSpacing>
    <w:DrawingGridVerticalSpacing>7.8 磅
    </w:DrawingGridVerticalSpacing>
    <w:DisplayHorizontalDrawingGridEvery>0
    </w:DisplayHorizontalDrawingGridEvery>
    <w:DisplayVerticalDrawingGridEvery>2
```

```

</w:DisplayVerticalDrawingGridEvery>
<w:Compatibility>
  <w:SpaceForUL/>
  <w:BalanceSingleByteDoubleByteWidth/>
  <w:DoNotLeaveBackslashAlone/>
  <w:ULTrailSpace/>
  <w:DoNotExpandShiftReturn/>
  <w:AdjustLineHeightInTable/>
  <w:BreakWrappedTables/>
  <w:SnapToGridInCell/>
  <w:WrapTextWithPunct/>
  <w:UseAsianBreakRules/>
  <w:UseFELayout/>
</w:Compatibility>
<w:BrowserLevel>MicrosoftInternetExplorer4</w:BrowserLevel>
</w:WordDocument>
</xml>
<![endif]-->
<style>
  <!--
  /* Font Definitions */
  @font-face
  { font-family:宋体;
    panose-1:2 1 6 0 3 1 1 1 1 1;
    mso-font-alt:SimSun;
    mso-font-charset:134;
    mso-generic-font-family:auto;
    mso-font-pitch:variable;
    mso-font-signature:3 135135232 16 0 262145 0;}
  @font-face
  { font-family:"\@宋体";
    panose-1:2 1 6 0 3 1 1 1 1 1;
    mso-font-charset:134;
    mso-generic-font-family:auto;
    mso-font-pitch:variable;
    mso-font-signature:3 135135232 16 0 262145 0;}
  /* Style Definitions */
  p.MsoNormal, li.MsoNormal, div.MsoNormal
  { mso-style-parent:"";

```

```
margin:0cm;
margin-bottom:.0001pt;
text-align:justify;
text-justify:inter-ideograph;
mso-pagination:none;
font-size:10.5pt;
mso-bidi-font-size:12.0pt;
font-family:"Times New Roman";
mso-fareast-font-family:宋体;
mso-font-kerning:1.0pt;}
/* Page Definitions */
@page
{mso-page-border-surround-header:no;
mso-page-border-surround-footer:no;}
@page Section1
{size:595.3pt 841.9pt;
margin:2.0cm 2.0cm 2.0cm 2.0cm;
mso-header-margin:42.55pt;
mso-footer-margin:49.6pt;
mso-paper-source:0;
layout-grid:15.6pt;}
div.Section1
{page:Section1;}
-->
</style>
<!--[if gte mso 10]>
<style>
/* Style Definitions */
table.MsoNormalTable
{mso-style-name:普通表格;
mso-tstyle-rowband-size:0;
mso-tstyle-colband-size:0;
mso-style-noshow:yes;
mso-style-parent:"";
mso-padding-alt:0cm 5.4pt 0cm 5.4pt;
mso-para-margin:0cm;
mso-para-margin-bottom:.0001pt;
mso-pagination:widow-orphan;
font-size:10.0pt;
```

```
font-family:"Times New Roman";}  
</style>  
<![endif]-->  
</head>  
<body lang=ZH-CN style='tab-interval:21.0pt; text-justify-trim:punctuation'>  
<div class=Section1 style='layout-grid:15.6pt'>  
<p class=MsoNormal>  
<span style='font-size:12.0pt; font-family:宋体'>你好</span>  
</p>  
</div>  
</body>  
</html>
```

上面这段 HTML 代码之所以这样庞大,主要原因是 Word 自动加入了大量的 XML 语句和 CSS 语句,前者用来指定网页的属性,后者用来指定网页的式样。但在这个例子中,无论这些语句是否存在,浏览器都将以默认的字体(通常是宋体)和默认的字号(通常是 3 号),显示“你好”这两个汉字(如图 1-1 所示),因此这些语句是可有可无的。

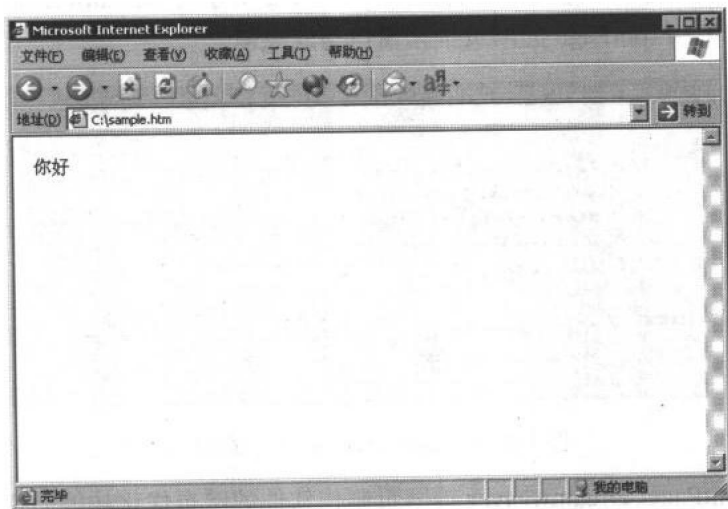


图 1-1 只有“你好”两个汉字的网页

从这个例子中可以看出, Word 在生成网页时,自动加入大量的可有可无的 HTML 代码,其他网页制作软件(例如 FrontPage 和 Dreamweaver)在生成网页时,也会自动加入一些可有可无的 HTML 代码,但情况没有

Word 这样严重。这种情况的最大坏处，就是增加了网页的字节数，降低了网页的显示速度，这就是我们主张优化网页 HTML 代码的根本原因。本书介绍了 70 多种优化方法，我们希望通过这些方法，能够帮助大家减少网页的字节数，提高网页的显示速度，以达到“网页小、速度快”的最终目的。

1.2 如何优化网页的 HTML 代码

大多数网页制作软件（例如 FrontPage 和 Dreamweaver）都具有网页优化功能，制作者利用这些软件制作一个网页之后，可以从菜单中选择网页优化窗口，对网页中的 HTML 代码进行各种优化处理。在 FrontPage XP 简体中文版中，选择【工具】→【网页优化】菜单命令，可以打开网页优化窗口。在 Dreamweaver MX 英文版中，选择【Command】→【Clean Up HTML】菜单命令，可以打开网页优化窗口。具体方法很简单，只要单击一个执行按钮即可完成网页优化操作。图 1-2 所示为 Dreamweaver 的网页优化窗口，FrontPage 的网页优化窗口与此类似。

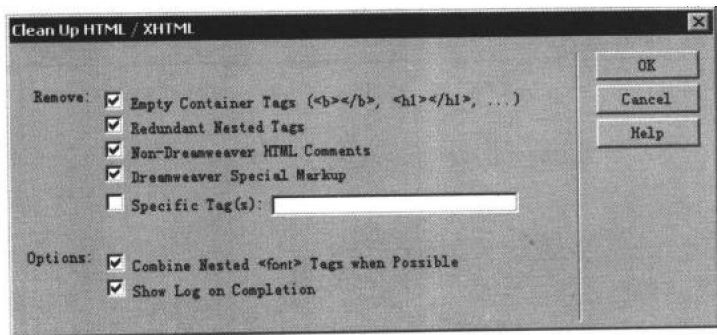


图 1-2 Dreamweaver 的网页优化窗口

FrontPage、Dreamweaver 等网页制作软件的网页优化功能通常是删除 HTML 代码中可有可无的内容，一般包括六个方面内容：

(1) 删除多余的引号。例如将下列第一条语句修改为第二条语句（下同）。

```
<font color="red">你好</font>  
<font color=red>你好</font>
```

(2) 删除多余的空格。

```
<p>你好</p>
```

```
<p>你好</p>
```

(3) 删除注释。

```
<blockquote>你好</blockquote> <!-- 这里是用户的留言 -->
```

```
<blockquote>你好</blockquote>
```

(4) 合并嵌套的 HTML 语句。

```
<p><p>你好</p></p>
```

```
<p>你好</p>
```

(5) 删除没有内容的 HTML 语句。

```
<p>你好<span></span></p>
```

```
<p>你好</p>
```

(6) 删除 HTML 语句的默认属性。

```
<td align="left">你好</td>
```

```
<td>你好</td>
```

除此之外, 还有一些专门的网页优化软件, 也能够完成上面的操作。图 1-3 所示是一个常用的网页优化软件 Advanced HTML Optimizer, 它可以将一个网页的大小从 2575 字节减少到 1843 字节。

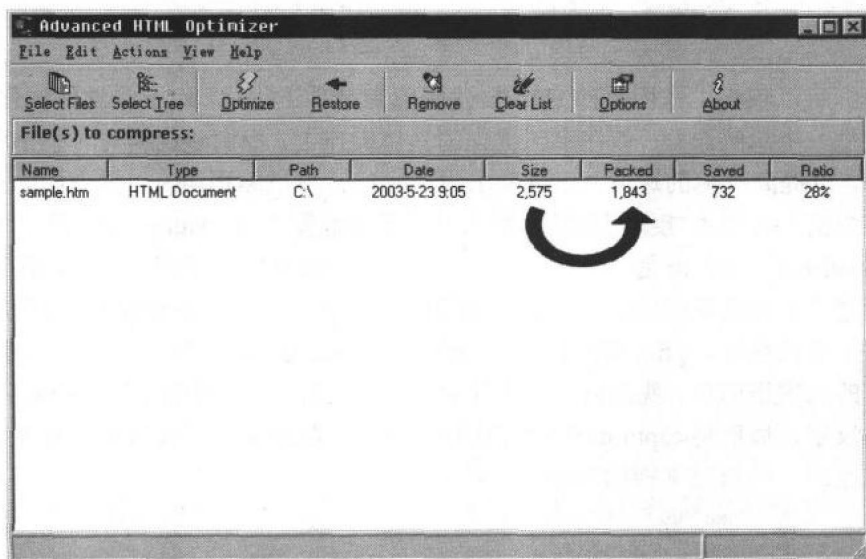


图 1-3 常用的网页优化软件 Advanced HTML Optimizer

这些软件的网页优化功能确实有一定的作用，但坦率地讲，它们的优化效果远远不能满足我们的要求。以下面的一组 HTML 语句为例，如果用这些软件优化，软件会认为没有问题，不需要优化，而如果用手工方法优化，所有 value 属性和所有 </option> 语句都可以删除，HTML 代码可以减少一半以上。

(1) 未优化的 HTML 代码。

```
<option value=星期一>星期一</option>
<option value=星期二>星期二</option>
<option value=星期三>星期三</option>
<option value=星期四>星期四</option>
<option value=星期五>星期五</option>
<option value=星期六>星期六</option>
<option value=星期日>星期日</option>
```

(2) 已优化的 HTML 代码。

```
<option>星期一
<option>星期二
<option>星期三
<option>星期四
<option>星期五
<option>星期六
<option>星期日
```

所以我们主张网页的 HTML 代码应当是手工优化，而不是软件优化，这是一个总的原则。但在手工优化的具体操作上，我们建议编写一些小程序，以提高优化的速度，这就要求掌握一门计算机编程语言。以上面的代码为例，如果不能编写程序，那么只能手工删除 7 个 value 属性和 7 个 </option> 语句。但是，一些大、中型网站，由成百上千个网页组成，手工删除就显得很很不现实了。为此，必须编写一个小程序自动删除有关内容。以上面的这组 <option> 语句为例，我们用 Visual Basic 编写了一个小程序，它的原理很简单，就是从一行语句中找出“星期一”、“星期二”、“星期三”等汉字，最后将 <option> 语句与这些汉字合并起来即可。读者可以参考它的原理，用自己掌握的其他语言编写。

```
Private Sub Form_Load()
    Dim code As String
    Dim position As Integer
```

```
Open "C:\old.htm" For Input As #1
Open "C:\new.htm" For Output As #2
Do While Not EOF(1)
    Line Input #1, code

    position = InStr(code, ">")
    code = Mid(code, position + 1)

    position = InStr(code, "<")
    code = "<option>" & Left(code, position - 1)

    Print #2, code
Loop
Close #1
Close #2
End
End Sub
```

1.3 关于本书内容的几点说明

本书大部分内容是介绍如何优化 HTML 代码，其余部分内容是介绍如何优化 JavaScript 和 CSS 代码的，因此我们要求读者熟悉 HTML 语言，了解 JavaScript 和 CSS 语言，并且有一定的手工制作网页的经验。目前许多设计者使用 WYSIWYG（所见即所得）的网页制作软件，例如 FrontPage 和 Dreamweaver，他们基本上不懂这三种语言，因此本书适合中、高级的网站制作者和网站管理者，不适合一般的初学者。

本书内容是在 Windows XP 简体中文版上编写的，所有优化方法均在两种常用浏览器（Internet Explorer 6 中文版和 Netscape Navigator 6 英文版）上测试通过，但由于条件限制，没有在其他操作系统（例如 Linux）和其他浏览器（例如 Opera）上测试，这一点需要说明。不过，考虑到 HTML 是一种跨平台的脚本语言，我们认为这些优化方法对 Linux 和 Opera 也同样适用。

本书介绍的优化方法，有些符合 HTML 语法，有些不符合 HTML 语法。以上节介绍的<option>、</option>语句为例，按照 HTML 语法，</option>语句是不能省略的，但根据我们长期从事网页制作的经验，省略