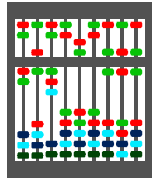


Hypertext News & Information

(HTML, HTMLHelp, & WinHelp)

A Newsletter for Clients and Colleagues of Scribble & Count™ LLC



Issue 2:4 / 3:1

Q4 1998 / Q1 1999

Contact info on p. 8

Contents

From the Partners p. 1

Double Issue (CSS2 & HTML 4)

Analysis pp. 2-7

CSS2 (p. 2)

HTML 4.0 (p. 4)

Calendar p. 7

About Scribble & Count p. 8

Description of company

Contact Information

News: S&C to present at STC conference in Utah; facilities improvements

Distribution

To Receive: Send us your name, title, company, postal AND email addresses, phone, and fax. Use our contact info (p. 8).

To Unsubscribe: We need the first line of your mailing label (print version), or your email address (notification of new posting).

Back Issues:

Paper: request by phone, fax, or email (contact info on p. 8); identify the issue(s).

Diskette: HTML only on cumulating diskette (request as for paper).

Website: All issues are at www.scribble-count.com/Newsltr_main.html(current issue delayed slightly). If you

From the Partners...

This will be another double issue: 2:4 (4Q98) & 3:1 (1Q99). Analyzing HTML 4.0 and its style sheet (CSS2) is too much for a couple pages, even when they're dense. (The fact that Bill was working in New York when 2:4 was due out didn't help, but the size of the project is the real reason it didn't get done on time.)

Issue 2:2 / 2:3 included an overview of hypertext options. Since then, we've had a chance to look more closely at the standards W3C released a year ago. We've distilled what we think are the essences of each, drawn any conclusions we feel comfortable with, and provided resources. The tone is (again) that of an overview, assisting managers and colleagues with their decisions.

Note: A temporary option for some Windows-only folks is to remain with WinHelp 4. Our argument for this hasn't changed since our presentation to the Rocky Mountain Chapter of STC (Nov. 97; syllabus available on our Website: www.scribble-count.com/Move-HTML.html), but time is running out for this alternative.

Because most of the changes in HTML 4.0 are related to accommodating style sheets, we'll cover CSS2 first. (And since CSS will become the main prototype for style sheets in general, there will be some transferability to well.)

The only browsers when upgrades are legacy baggage train. of the newsletter, we'll terms; the online Glossary file, and use of terms that we recognize may be unfamiliar. There will be a link to the Quick-Reference Front Door on our home page when it's ready.

After the TechnoTips, we'll list coming events. We've discontinued the General articles, because they're now on our Website (. . /essays_main.html).

We've overhauled the entire Website, and the newsletters are also on it (. . /newsltr_main.html). If you'd like to be notified by email when each issue is posted instead of or in addition to receiving the paper copy, contact us as listed on page 8. (Posting slightly delayed, depending on our schedules.)

If you have any comments (and we want *very much* to hear them), please email them to Bill Sanders.

Thanks.

Marci Bowman
mbowman@scribble-count.com

Bill Sanders
bsanders@scribble-count.com

TechnoTips (HTML 4.0 & CSS2)...

Cascading Style Sheets 2 and HTML 4.0 have been out for more than a year, so it's possible to examine each and draw some conclusions concerning their strengths, weaknesses, and prospects. Because CSS has driven so many of the changes to HTML 4.0, we'll look at it first.

- Conventions:**
1. To distinguish them, tags are placed in angle brackets (`` or `<style>`); attribute names are capitalized (FONT or STYLE). However, coders should be aware that, although HTML is not case-sensitive, XML is.
 2. Generic content is italicized (*xxx*, or `<div.class>`); substitute whatever is applicable.

Cascading Style Sheets (CSS2)

Status (2/15/99): Final Recommendation released 5/12/98; nothing new scheduled.

Introduction:

The W3C intends to separate formatting from traditional markup dealing with structure. HTML is a mess in this regard: tech writers and desktop publishers have been cluttering the code for years, and it must be cleaned up before it is incorporated into XML as a subset.

Note: HTML was designed to describe simple content (including hypertext links) for transmission by a standard protocol (TCP/IP). Trainers and librarians know that formatting which aids retention and visual recognition is also content – but abuse is easy, and purists' resistance is strong.

This separation of form and content is the real, pragmatic reason behind CSS2; but as soon as we get accustomed to the syntax and the required self-discipline, we'll like the power and flexibility of style sheets. (In any event, the change is inevitable, although not normally imminent.)

Both IE 4 and Netscape 4 were released before HTML 4.0 and CSS2 were final – so browser support is still inconsistent. (Both should be mostly, though incompatibly, compliant in their versions 5, due out this year.) Because support is not consistent, and because there are nearly a billion legacy pages on the Web, the W3C has provided three temporary work-arounds. In decreasing order of preference, they are:

- declarations within a `<style>` tag in the document `<head>`;
- a new STYLE attribute that can be used with most tags;
- the deprecation (but not yet removal) of 3.2 style-related tags (such as `` and `<bgcolor>`), and attributes (such as FONT and most uses of ALIGN).

To whatever extent you utilize these work-arounds, liberal use of `<!--comments-->` in any document with a life expectancy of more than a few months will make later revision much easier.

Main points in CSS2:

The spec is long and complex; we've selected what we think are the most essential aspects.

Precedence. There are four possible levels of style designation in an HTML 4.0-compatible document; later designations override anything earlier with which they conflict. Fortunately, specific styles with the narrowest scope can almost always be placed last. In order of normal appearance (and increasing precedence), these levels are:

External style sheets linked in the `<head>`

External style sheets imported in the `<head>`

Declarations within the `<style>` tag in the `<head>`

Inline styles in the `<body>` (either as a 3.2 tag or attribute, or using the new STYLE attribute).

TechnoTips (cont.)...

The sample code below shows how external sheets are referenced and declarations are made in the `<head>`:

```
<head>
<link rel=stylesheet type="text/css" href="xxxx">
<style>
  @import url (xxxx);
  body {background: xxxx}
  p.this {font-size: xxxx; color: xxxx}
</style>
```

The external sheets are linked and then imported; declarations follow, where they can override in case of conflict. Try to minimize inline styles by making generalizations that you can put in the `<head>`. Use `<!--comments-->` liberally in the style sheet to explain your actions.

Rectangular Structure. CSS2 assigns four nested rectangular parts (margin, border, padding, and core) to every element (a division, an image, a paragraph, a table, a table cell, etc.), as shown in Fig. 1.

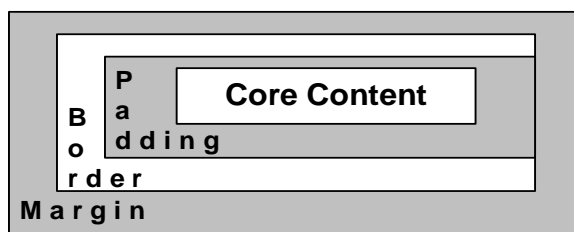


Fig. 1: Rectangular Structure

The widths can be different on all parts of the rectangles, and any of them can be set to zero; all four parts must fit inside whatever container holds it. (For example, if a graph is placed in a table, the image rectangle with its margin/border/padding must fit inside a table cell; the cell rectangle must fit inside the table; the table rectangle must fit inside the `<p>`, `<div>`, or `<body>`.) This isn't as cumbersome as it sounds, and it permits placement to within pixel accuracy.

CLASSES. Assigning classes to various tags will *eventually* be a powerful search tool, but you can use it now for formatting to the extent the browser is CSS-compatible.

For example, eventually you'll be able to search by class (only the abstract, only citations, only quotes, only `<div.class>`, for example). Now you can assign classes to nearly any tag, so you could define different formats in the style sheet. `<p.this></p>` could have different formatting from `<p.that></p>`, if the two classes of `<p>` tags were defined differently in the style sheet (see code sample, above).

Note: Word, WordPerfect, and FrameMaker users will see *conceptual* similarities, but implementation is very different.

Conclusions:

HTML styles have most of the power and flexibility of print-based styles; but they will also do something most print styles don't: combine with events, classes, and IDs to increase the natural advantages of automation.

You'll soon be able to format more with CSS than you ever could with traditional declared and inline styles. HTML purists are happy, because they get their unadorned markup language back.

Moving to CSS. Most HTML books (*e.g.*, Musciano's *HTML: The Definitive Guide*, espec. p. 278) recommend abandoning as quickly as possible inline and declared styles in favor of CSS2. Scribble & Count is slightly more cautious: learn the basics as quickly as possible; then analyze your audience (if you can) and abandon traditional styles only after your ability and your users' browsers are ready. In the meantime, `<!--comment-->` any code you intend to use after 1999.

TechnoTips (cont.)...

Note: The `<!--comments-->` referred to above are for maintenance purposes. If you use style sheets for documents read by old browsers, also enclose the style declarations in comments:

```
<style>
  <!--style declarations-->
</style>
```

Old browsers won't recognize the `<style></style>` tags and will ignore the entire sequence, but IE and Netscape have promised that their next browsers will be savvy enough to know your trick.

Interoperability w/ XML. Because it ignores inline styles, XML can read both CSS2 and HTML 4.0. As soon as the formatting flotsam deposited by us tech writers and desktop publishers is removed (years, not months), it will be subsumed into a later version of XML. At that point, we can still write pure HTML, but as a subset of XML it will have to adhere rigidly to then-current XML standards (including upper and lower case); browsers will not be as tolerant as they are now. And when that day arrives, mastery of style sheets will be an absolute necessity.

Final Warning. When fully implemented, style sheets will do for HTML what they've done for years in word processing and DTP software: allow you to change – for good if you make the right choices; for ill if you don't – the appearance and organization of a document. And since styles are extremely powerful, they can make you very creative and efficient. Or they can help you do incalculable damage. If you haven't caused at least one disaster by misusing print styles, have a long talk with someone who has.

Resources for CSS2:

We've listed the W3C spec, and a text that is at least more readable than the spec. (Each month, introductory articles are published in technical periodicals; when you're ready to tackle CSS2, you might start with one of them.)

Busch, David D.; & J. W. Olsen. *Cascading Style Sheets Complete*. New York: McGraw-Hill, 1998. Pbk., 423pp., glossary, index, CD. ISBN: 0-07-913703-2.

[Covers both CSS1 & CSS2. Perhaps tries to cover too much: begins basic, covers advanced topics such as client/server differences, ends with valuable speculations.]

"Cascading Style Sheets, level 2 CSS2 Specification". www.w3.org/TR/REC-CSS2-19980512 [2/15/99].

[Not written for the learning process; but rigid structure makes it fine for quick-retrieval if you already know CSS.]

HTML 4.0

Status (2/15/99): Ver. 4.0 final specification released 4/24/98; nothing more scheduled.

Introduction

HTML 4.0 is three, apparently contradictory, things:

- Long-term: a huge change;
- Mid-range: a temporary transition vehicle;
- Short-term: hardly a ripple (if you decide to wait a while).

How can that be? And what do "long-term" and "short-term" mean?

Ver. 4.0 is a clear warning of what's ahead: an end in (or at least reduction to) the Hypertext Wars, and a firm commitment to subsume HTML into XML. 4.0 also contains two tracks: fast (strict variant) and slow (transitional variant) – so you have an adjustment period if you want. (**Note:** There's also a variant for extensive use of frames.) And since there are nearly a billion legacy pages out there, the browsers won't jettison support for the old styles any quicker than Gates got rid of DOS. So: enjoy the bleeding edge; move deliberately; or be timid. It's your choice. Total time frame? Years, not months.

Since Netscape 4 and IE 4 still don't implement ver. 3.2 consistently, don't expect the standards set forth so assertively by the W3C in HTML 4.0 (see **Resources**, below) to become rigid any time soon.

TechnoTips (cont.)...

Essences of Ver. 4.0

It seems to us that the major changes pertain to styles, classification techniques, and event handling.

Encouragement To Format w/ Style Sheets and Declarations. Style sheet(s) and declarations are listed in the document <head> (see the **CSS** section above.) To further this goal:

- Most format-related tags and attributes have been "deprecated" (W3C jargon for strongly discouraged, but reluctantly supported for the time being).
- Most tags have a STYLE attribute for inline styles. (**Note:** Probably a temporary device.)
- Use is encouraged of tags that include style aspects:
 - <tt> and <code> imply a monospaced font;
 - <cite> implies italics;
 - <dfn> implies indentation.

The W3C intends to extract many benefits from this separation of form and content:

- If you want specific formatting, you should specify it in a style sheet (where it will also be easier to maintain)
- If you don't so specify, each browser can make its own selections (which increases cross-platform portability)
- Some of the tags (<cite>, <dfn>, and <div.class>, for example) can be used for other purposes, such as searching or compiling bibliographies and glossaries.
- Ultimately, the changes will remove most of the inline formatting and speed integration into the more rigid XML.

All this will not come to pass overnight. Ver. 4.0 has not been implemented completely by anyone (IE 4 and Netscape 4 were released before HTML 4.0 was final), and there are nearly a billion legacy pages. W3C deprecation, therefore, does not mean imminent withdrawal of support. (But the direction is inevitable and irreversible – that's why HTML 4.0 is both a drastic change and barely a ripple.)

Addition of ID and CLASS Attributes. You can put them in almost any tag, and tie them to search techniques or formatting. An ID must be unique (and so can also serve as a bookmark or a substitute for the NAME attribute of the "anchor" <a> tag; the value for CLASS should be common to a group. ID and CLASS can be used like *HTMLHelp* Topic Types; and this ex-reference librarian predicts that writers will come to love the power to categorize that this technique offers.

Addition of EVENT Attributes. Ver. 4.0 identifies 10 EVENT attributes, which are available in most tags:

ON CLICK	ON KEYUP	ON MOUSEUP	ON MOUSEOVER
ON DBLCLICK	ON KEYDOWN	ON MOUSEDOWN	ON MOUSEOUT
	ON KEYPRESS	ON MOUSEMOVE	

If you know how any one of these works, you can intuit the basics of the others, but you'll have to consult a 4.0 reference (two listed in **Resources**) to determine acceptable values.

Note: If you've been following our analysis of the Hypertext Wars (see *Hypertext News & Information*, issue 2:2 / 2:3), you'll know that this is where IE and Netscape differ drastically; but the W3C envisions ECMAScript as a kind of peace treaty. After being properly chastened by the Justice Dept., Microsoft might acquiesce.

Other Aspects

There are dozens of other changes. The ones S&C has found to be the most significant are listed briefly below.

Tag & Attribute Changes. Dozens of tags and attributes have been either deprecated or added/revised.

The deprecations are mostly format-related, reflecting the move toward style sheets. The following, very incomplete, list should provide a feel for the changes.

<u>Tags</u>	<u>Tags</u>	<u>Attributes</u>
<basefont>	<u>	ALIGN
<center>		STYLE (!!both new & deprecated!!)
	<strike>	BORDER

Hypertext News & Information

The tags added or revised deal primarily with content or organization, for example: <abbr>, <acron>, <object>, <q>. Each browser has a default formatting for these, which you can accept or change in a style sheet.

Tables. Tables are vital. In fact, they've been overused as a formatting work-around. Except for the style-sheet-related attributes (such as ALIGN), which are deprecated for reasons described above, all the aspects of ver. 3.2 remain supported, and some previously proprietary extensions are included in the new spec. Some noteworthy changes:

- Partial borders permitted.
- ALIGN=*char* equivalent to the decimal tab in word processors.
Note: The value *char* has been added to an attribute which has been deprecated! This type of inconsistency occurs frequently; there's as much politics as hypertext in HTML 4.0.
- Division of tables into sections (each styled or classed differently with a <tbody> tag).
- Control over columns with the <colgroup> and <col> tags.
- Running headers and footers with <thead> and <tfoot> tags.

Frames. The only major changes in 4.0 are the addition of inline frames (<iframe>), and a variant of the spec for heavy frame users. Inline frames function like inline images, and the attributes they accept are very similar to the <frame> tag. However, the concept was IE's, so it's not clear to what extent Netscape 5 will support it.

Note: Adding the ID attribute to a <frame> tag produces something similar to WinHelp's secondary window.

Conclusions

Moving to HTML 4.0. More than half the features in 4.0 are not yet supported by *any* browser, and none supports more than a third. So even after browser versions 5 come out, you have to decide how quickly you will cut old browsers loose. Scribble & Count has decided that, with Netscape and IE commanding 95+% of the market and both available free, that we will have little patience with laggards — unless a client has other views, of course. A few months after ver. 5 of both are available, we will begin supporting HTML 4.0's new features (normally, the "transitional" variant). But we'll read an analysis or two first to see which features are so differently implemented that they're best avoided. In short, the changes are inevitable and will eventually be total — but they aren't imminent unless you choose to make them so.

Relationship of HTML 4.0 to DHTML and XML. There has been a lot written about XML being the successor to an HTML which has been "overtaken by events" (read: "corrupted by us format-happy authors"). A conference in San Jose in May 98 (less than a month after the latest 4.0 recommendation was released) made this official: once HTML's formatting is tucked away in style sheets, the pure HTML structure and content that remains will be subsumed into XML as a subset. DHTML has a similar future.

Power of Styles. 4.0's style provisions will change the following HTML syntax:

1. A basic style sheet can be identified with default characteristics of font, etc., for each of the formatting tags, and for event behaviors. Maintenance becomes much easier.
2. Documents can be subdivided into large or small divisions (body, glossary, citations, revisions, etc.), each of which can be defined in styles sheets or declarations. Formatting becomes more intricate, searching more focused.
3. A specific type of tag (for example <p>) can be subdivided into classes (<p.this></p> and <p.that></p>), each of which can be defined in style sheets or declarations. Only the structure (<p.this>, <p.that>) is in the HTML code; what the structure means is in the style sheet.
4. Inline formatting can still be done one the fly with the new STYLE attribute. This "halfway house" for recovering format abusers will actually hasten the transition; the syntax for the STYLE attribute is the same as what the style sheets and <style> tag inside the <head> use.

The more generally the style is declared, the more easily it can be modified. We are obviously getting powerful stuff here — with the power to bungle if improperly applied, as any user of Word's or FrameMaker's styles has come to appreciate.

Hypertext News & Information

Transitional Work-arounds. There are two major ones:

1. Until your audience's browsers implement 4.0 inconsistently, you might want to designate classes without defining them in a style sheet or declaration. Add the definitions when you're satisfied your audience is in compliance.
2. Alternatively, you can go ahead declare the styles in the <head>; non-4.0 browsers will ignore the <style> tag and anything inside if it is also <!--commented-->, but compliant browsers should be programmed to act on the style declarations that are inside <style> tags, even if they are packaged as <!--comments-->.

And, of course, <!--comment--> all your techniques throughout the document.

Resources for HTML 4.0:

We've included the W3C spec, a text, two reference works, and even the W3C site for browsing purposes.

Graham, Ian. *HTML 4.0 Sourcebook: A Complete Guide to HTML 4.0 and HTML Extensions.* New York: Wiley, 1998. Pbk., 631pp., index. ISBN: 0-471-25724-9.

[It's all there as a very useful reference – if you know 3.2 well. (If you don't, it's all there as a very ponderous text.)]

"HTML 4.0 Reference". www.htmlhelp.com/reference/html40 [accessed 2/15/99].

[Several excellent reference sections.]

"HTML 4.0 Specification". www.w3.org/TR/REC-html40/ [accessed 2/15/99]

Musciano, Chuck; & Bill Kennedy. *HTML: The Definitive Guide*, 3rd ed. Sebastapol, CA: O'Reilly, 1998. Pbk., 587pp., index, apps. ISBN: 1-56592-492-4.

[Better than Graham as a learning text, though perhaps a bit too optimistic on the speed with which 4.0 will be accepted and 3.2 will fade. Main drawback: does not identify clearly enough which features are new to 4.0.]

World Wide Web Consortium Website. www.w3.org/ [accessed 2/15/99].

[Entire site has been recently restructured; and is worth browsing through.]

Calendar...

Colorado Events

AITP / DPMA. 3rd Th. Normally a dinner-speaker. \$25-30. 303-789-4547.

Boulder Writers' Alliance. 4th Tues. on odd months, 6:30, NIST (325 S. Broadway). www.bwa.org.

Rocky Mountain Internet Users' Group. 2nd Tues., odd months, NIST (325 S. Broadway, Boulder). [Http://plaza/xor.com/rmiug/](http://plaza.xor.com/rmiug/).

STC (Rocky Mtn Chapter). 3rd Th., 5:45 (SIGs & networking), 7:15 (program), venue varies. www.stc-rmc.org/

Windows on the Rockies Users' Group. 3rd Mon., 6:30, usually Best Western Landmark Hotel. (455 S. Colorado). www.wrug.org.

National Events

Apr 20-23 Info Online Conference & Expo. Orlando, FL. 888/333-9088. www.io-conference.com.

Apr 26-30 XML Europe '99. Granada, Spain. 703-519-8160.

May 3 Java Jump Start Conference. Burlingame, CA. 800/838-8999. www.winwriters.com/Java/jhelp99.htm

May 16-19 STC Annual Conference. Cincinnati. www.stc-va.org.

Oct. 14-16 STC Rocky Mountain Conference. Bellevue, WA. www.stc-psc.org/conf.htm.

About Scribble & Count...

Description

Scribble & Count, LLC is a Colorado information services company specializing in hypertext. Its owners, Marci Bowman and Bill Sanders, have English, library, and education degrees; and nearly 20 combined years experience in tech writing, research, records management, and training. Ms. Bowman is past VP of Internet Chamber of Commerce, and Mr. Sanders has been Technical Assistant on CompuServe's Hypertext Forum. The S&C Website has a page devoted to hypertext services, and a section containing essays on hypertext: general intro, news, social impact, etc. These newsletters have been added in a separate section. A lecture/slide presentation with similar content has been given to metro Denver technical, business, and professional groups; to regional professional conferences; and to a national Help/Web developers' convention.

Contact Points

POB 24478, Denver, CO 80224 (POB 1533, Idaho Springs, CO 80452)
303/377-6399(v) / 303/388-2526(f) | www.scribble-count.com | online@scribble-count.com

S&C News

Bill will again address the Intermountain STC regional conference (Salt Lake City, April 17). He will lead a roundtable discussion comparing print and online media. Marci will speak to three local groups; the Website (www.scribble-count.com/S&C_New.html) will have details when they are set.

Work has begun on S&C's new office in the mountains above Idaho Springs; see the Website (.../HDQ_main.html) for pictures and unrealistic completion schedule.

1998 was S&C's best year ever (thanks to our clients who were so instrumental). As a result, we have purchased two more computers, added a third printer, and upgraded the scanner. Marci has networked the three production machines (two on NT one on 95; the stand-alone dual boots 3.1 and 95).

Scribble & Count, LLC
PO Box 24478
Denver, CO 80224

To: