Chapter 4
A Hypertext Markup Language Primer
XHTML Mark Up with Tags

• Extensible Hypertext Markup Language

• Format
  – Word/abbreviation in < >
  – <title> PAIR </title>
  – Singleton (not surround text) />
    <br />
    <hr />
  – XHTML lowercase tags
  – Text between tags is case sensitive
• <!-- comments -->
Figure 4.2 Standard form for an HTML Web page; every page must have these tags nested this way: (a) HTML code as seen in Notepad++ and (b) HTML code as seen in Text Wrangler.
Paradoxes

Russell's Paradox

The Twentieth Century logician Bertrand Russell introduced a curious paradox: *This statement is false.* The statement can't be true, because it claims the converse. However, if it is not true, then it's false, just as it says. That makes it true. Paradoxically, it seems to be neither true nor false, or perhaps both true and false.

Magritte's Paradox

The famous Belgian artist René Magritte rendered the idea of Russell's Paradox visually in his famous painting *Ceci n'est pas une pipe.* The title translates from French, *This Is Not A Pipe.* The painting shows a pipe with the text *Ceci n'est pas une pipe* below it. Superficially, the painting looks like a true statement, since it is a picture of the pipe, not an actual pipe. However, the assertion is also part of the picture, which seems to make it false, because it is clearly a painting of a pipe. Paradoxically, the truth seems to depend on whether the statement is an assertion about the painting or a part of it. But, it's both.
<body>
    <img src="beach.jpg" alt="tropical beach" />
</body>
</html>

**Figure 4.4** Diagram of the compose and check process; with both the browser and text editor open alternate between composing HTML and saving (^S), and refreshing the Firefox display of the HTML and checking.

**Figure 4.5** The W3C Markup Validation Service page; browse for your HTML file and then click Check.
Bold, Italic, Underline

- **Bold:** `<b>` `</b>`
- **Italic:** `<i>` `</i>`
- **Underline:** `<u>` `</u>`
  - Tag pair surrounds text to format
- **Nesting**

`<b><i> Veni, Vidi, Vici! </i></b>` produces: **Veni, Vidi, Vici!**
<html>
<head>
<title>Formatting Examples</title>
</head>
<body>
<p>
This is <b>bold</b> text.
This is <big>big</big> text.
This is <em>emphasized</em> text.
This is <i>italic</i> text.
This is <tt>small</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
This is <tt>smaller</tt> text.
</p>
</body>
</html>

---

**Figure 4.1** HTML source code for formatting tags and the result (inset).
Headings

• 6 levels
• Display on a new line
• **Indent** HTML tags (source code)

   `<h1>Pope</h1>`

   `<h2>Cardinal</h2>`

   `<h3>Archbishop</h3>`

Pope
Cardinal
Archbishop
<table>
<thead>
<tr>
<th>Start Tag</th>
<th>End Tag</th>
<th>Meaning</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>&lt;doctype html&gt;</code></td>
<td></td>
<td>First tag in an HTML5 file</td>
<td>✓</td>
</tr>
<tr>
<td><code>&lt;html&gt;</code></td>
<td><code>&lt;/html&gt;</code></td>
<td>Tag enclosing all HTML text</td>
<td>✓</td>
</tr>
<tr>
<td><code>&lt;title&gt;</code></td>
<td><code>&lt;/title&gt;</code></td>
<td>Title bar text; describes page</td>
<td>✓</td>
</tr>
<tr>
<td><code>&lt;head&gt;</code></td>
<td><code>&lt;/head&gt;</code></td>
<td>Preliminary material; e.g., title at start of page</td>
<td>✓</td>
</tr>
<tr>
<td><code>&lt;body&gt;</code></td>
<td><code>&lt;/body&gt;</code></td>
<td>The main, content part of the page</td>
<td>✓</td>
</tr>
<tr>
<td><code>&lt;p&gt;</code></td>
<td><code>&lt;/p&gt;</code></td>
<td>Paragraph</td>
<td></td>
</tr>
<tr>
<td><code>&lt;hr/&gt;</code></td>
<td></td>
<td>Line (horizontal rule)</td>
<td></td>
</tr>
<tr>
<td><code>&lt;h1&gt;...&lt;h6&gt;</code></td>
<td><code>&lt;/h1&gt;...&lt;/h6&gt;</code></td>
<td>Headings, six levels</td>
<td></td>
</tr>
<tr>
<td><code>&lt;b&gt;</code></td>
<td><code>&lt;/b&gt;</code></td>
<td>Bold</td>
<td></td>
</tr>
<tr>
<td><code>&lt;i&gt;</code></td>
<td><code>&lt;/i&gt;</code></td>
<td>Italic</td>
<td></td>
</tr>
<tr>
<td><code>&lt;pre&gt;</code></td>
<td><code>&lt;/pre&gt;</code></td>
<td>Preformatted text in which white space matters</td>
<td></td>
</tr>
<tr>
<td><code>&lt;a href=&quot;fn&quot;&gt;</code></td>
<td><code>&lt;/a&gt;</code></td>
<td>Anchor reference, <code>fn</code> must be a pathname to an HTML file</td>
<td></td>
</tr>
<tr>
<td><code>&lt;img src=&quot;fn&quot;/&gt;</code></td>
<td></td>
<td>Image source reference, <code>fn</code> must be a pathname to a .jpg, png, or .gif file</td>
<td></td>
</tr>
<tr>
<td><code>&lt;br/&gt;</code></td>
<td></td>
<td>Break, continue text on the next line</td>
<td></td>
</tr>
</tbody>
</table>
• Ordered, Unordered
• Sublist
• Different bullet symbol

<ol>
  <li>Lab 1</li>
  <li>Lab 2</li>
  <li>Lab 3</li>
</ol>
In his *Journal* of October 27, 1853 Thoreau wrote that he was obligated to buy back from the printer the remaining copies of his *A Week On the Concord and Merrimack Rivers*. Of the 1000 books printed he had to buy 706, which he still owed money on and had to carry up two flights of stairs. "I have now a library of nearly 900 volumes," he wrote, "over 700 of which I wrote myself."

Steinbeck traveled to Russia several times, but never mastered the language. Traveling with photographer Robert Capa in 1947 he wrote, "...I admit our Russian is limited, but we can say hello, come in, you are beautiful, oh no you don't, and one which charms us but seems to have an application rarely needed, 'The thumb is second cousin to the left foot.' We don't use that one much."

**Figure 4.5.** The display of the two pages from Figures 4.3 and 4.4 in a small window showing that the table keeps the links in a single row (bottom) rather than wrapping them (top).
Steinbeck traveled to Russia several times, but never mastered the language. Traveling with photographer Robert Capa in 1947 he wrote, "...I admit our Russian is limited, but we can say hello, come in, you are beautiful, oh no you don't, and one which charms us but seems to have an application rarely needed, 'The thumb is second cousin to the left foot.' We don't use that one much."
**Tables**

- A list of lists
- rows, columns, cells

```html
<table border="1">
  <caption>Bilingual Countries</caption>
  <tr><th>Country</th><th colspan="2">Languages</th></tr>
  <tr><td>Belgium</td><td>Dutch</td><td>French</td></tr>
  <tr><td>Canada</td><td>English</td><td>French</td></tr>
  <tr><td>Cyprus</td><td>Greek</td><td>Turkish</td></tr>
  <tr><td>Philippines</td><td>English</td><td>Filipino</td></tr>
</table>
```
Figure 4.14 The Bilingual Countries table, styled to enhance the caption and make the cells easier to read.
```html
td, th {
  border-style:solid;
  border-width:1px;
  padding:5px;
  text-align:center;
  background-color:cornsilk;
  color:saddlebrown;
}
tr.alt td {
  background-color:blanchedalmond;
}

<table>
  <caption>Bilingual Countries</caption>
  <tr>
    <th>Country</th> <th colspan="2">Languages</th>
  </tr>
  <tr>
    <td>Belgium</td> <td>Dutch</td> <td>French</td>
  </tr>
  <tr class="alt">
    <td>Canada</td> <td>English</td> <td>French</td>
  </tr>
  <tr>
    <td>Cyprus</td> <td>Greek</td> <td>Turkish</td>
  </tr>
  <tr class="alt">
    <td>Philippines</td> <td>English</td> <td>Filipino</td>
  </tr>
</table>

**Figure 4.15** Table with rows of alternating color, and the styling that achieves it.
White space and Attributes

• White space (spaces, tabs, new lines)
  - converted into a single space
  
  <pre> (pre - format) </pre>

• Attributes – inside angle brackets, after tag word
  
  <p align = "center"> (default left)

• Horizontal rule
  
  <hr width="50%" size="3" />
Anchor Tags

• **Absolute pathnames:**
  
  Page at other web sites using complete URLs
  
  http://server/directory_path/filename

  `<a href="http://www.aw.com/snyder/index.html">FIT</a>`

• **Relative pathnames:**
  
  Page stored in same directory (filename)
  
  – More flexible — move web files around as a group
  
  – Specify path deeper or higher in directory structure
Thoreau

In his *Journal* of October 27, 1853 Thoreau wrote that he was obligated to buy back from the printer the remaining copies of his *A Week On the Concord and Merrimack Rivers*. Of the 1000 books printed he had to buy 706, which he still owed money on and had to carry up two flights of stairs. "I have now a library of nearly 900 volumes," he wrote, "over 700 of which I wrote myself."

*Figure 4.3. A page and its HTML for a simple listing of links.*
Image Tags

```html
<img src="pathname" />
```

- Using anchor tag, pictures as links

```html
<a href="http://www.ucsd.edu"> <img src="ucsdlogo.gif" /> </a>
```

**GIF:** Graphics Interchange Format (.gif)
- 8 bits (256 colors or levels)

**JPEG:** Joint Photographic Experts Group (.jpg .jpeg)
- 24 bits (millions of colors)

**PNG:** Portable Network Graphics (.png)
Image Positioning

- Images inserted where tag is in HTML, aligns text with bottom of image (default)
- Attribute `align`
- Values: top line of text, middle, or bottom left, center, or right side of image
- Image on separate line, enclose in `<p>` tags
- Attributes `height` and `width`:
  (size in pixels to display image)
Color

- Hexadecimal color numbers (0-9 A-F)
  `<body bgcolor="#FF00FF"> (fuchsia)`

- Color terms
  `<body bgcolor="magenta">`

- `color` with body text, link, or font tags
  `<font color="red">text</font>`

<table>
<thead>
<tr>
<th>Table 4.3</th>
<th>Original HTML colors</th>
</tr>
</thead>
<tbody>
<tr>
<td>black</td>
<td>silver</td>
</tr>
<tr>
<td>red</td>
<td>fuchsia</td>
</tr>
<tr>
<td>blue</td>
<td>navy</td>
</tr>
<tr>
<td>lime</td>
<td>green</td>
</tr>
</tbody>
</table>
### Table 4.4. Hexadecimal Digit Equivalents

<table>
<thead>
<tr>
<th>Hex</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td>61</td>
<td>62</td>
<td>63</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>90</td>
<td>91</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
</tr>
<tr>
<td>6</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>103</td>
<td>104</td>
<td>105</td>
<td>106</td>
<td>107</td>
<td>108</td>
<td>109</td>
<td>110</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>112</td>
<td>113</td>
<td>114</td>
<td>115</td>
<td>116</td>
<td>117</td>
<td>118</td>
<td>119</td>
<td>120</td>
<td>121</td>
<td>122</td>
<td>123</td>
<td>124</td>
<td>125</td>
<td>126</td>
<td>127</td>
</tr>
<tr>
<td>8</td>
<td>128</td>
<td>129</td>
<td>130</td>
<td>131</td>
<td>132</td>
<td>133</td>
<td>134</td>
<td>135</td>
<td>136</td>
<td>137</td>
<td>138</td>
<td>139</td>
<td>140</td>
<td>141</td>
<td>142</td>
<td>143</td>
</tr>
<tr>
<td>9</td>
<td>144</td>
<td>145</td>
<td>146</td>
<td>147</td>
<td>148</td>
<td>149</td>
<td>150</td>
<td>151</td>
<td>152</td>
<td>153</td>
<td>154</td>
<td>155</td>
<td>156</td>
<td>157</td>
<td>158</td>
<td>159</td>
</tr>
<tr>
<td>A</td>
<td>160</td>
<td>161</td>
<td>162</td>
<td>163</td>
<td>164</td>
<td>165</td>
<td>166</td>
<td>167</td>
<td>168</td>
<td>169</td>
<td>170</td>
<td>171</td>
<td>172</td>
<td>173</td>
<td>174</td>
<td>175</td>
</tr>
<tr>
<td>B</td>
<td>176</td>
<td>177</td>
<td>178</td>
<td>179</td>
<td>180</td>
<td>181</td>
<td>182</td>
<td>183</td>
<td>184</td>
<td>185</td>
<td>186</td>
<td>187</td>
<td>188</td>
<td>189</td>
<td>190</td>
<td>191</td>
</tr>
<tr>
<td>C</td>
<td>192</td>
<td>193</td>
<td>194</td>
<td>195</td>
<td>196</td>
<td>197</td>
<td>198</td>
<td>199</td>
<td>200</td>
<td>201</td>
<td>202</td>
<td>203</td>
<td>204</td>
<td>205</td>
<td>206</td>
<td>207</td>
</tr>
<tr>
<td>D</td>
<td>208</td>
<td>209</td>
<td>210</td>
<td>211</td>
<td>212</td>
<td>213</td>
<td>214</td>
<td>215</td>
<td>216</td>
<td>217</td>
<td>218</td>
<td>219</td>
<td>220</td>
<td>221</td>
<td>222</td>
<td>223</td>
</tr>
<tr>
<td>E</td>
<td>224</td>
<td>225</td>
<td>226</td>
<td>227</td>
<td>228</td>
<td>229</td>
<td>230</td>
<td>231</td>
<td>232</td>
<td>233</td>
<td>234</td>
<td>235</td>
<td>236</td>
<td>237</td>
<td>238</td>
<td>239</td>
</tr>
<tr>
<td>F</td>
<td>240</td>
<td>241</td>
<td>242</td>
<td>243</td>
<td>244</td>
<td>245</td>
<td>246</td>
<td>247</td>
<td>248</td>
<td>249</td>
<td>250</td>
<td>251</td>
<td>252</td>
<td>253</td>
<td>254</td>
<td>255</td>
</tr>
</tbody>
</table>

**Note:** Find the decimal number in the table and then combine the entries in the left column and the top row symbols to form the hexadecimal equivalent. Thus decimal 180 is hexadecimal B4.
<html>
<head>
    <title>Twentieth Century Paradoxes</title>
</head>
<body bgcolor="#000000" text="#DDDDDD" link="#FFCC66">
    <h1 align="center"><font COLOR="yellow">Paradoxes</font></h1>
    <h2><font color="#FF8E2A">Russell's Paradox</font></h2>
    <p>The Twentieth Century logician Bertrand Russell introduced a curious paradox: <b>This statement is false.</b> The statement can't be true, because it claims the converse. However, if it is not true, then it's false, just as it says. That makes it true. Paradoxically, it seems to be neither true nor false, or perhaps both true and false.</p>

    <hr width="75%">

    <h2><font color="#FF8E2A">Magritte's Paradox</font></h2>
    <p><img src="pipe.jpg" height="130" width="192" align="right"> The famous French artist rendered the idea of Russell's Paradox visually in his famous painting <i>Ceci n'est pas une pipe</i>. The title translates from French, This Is Not A Pipe. The painting shows a pipe with the text <i>Ceci n'est pas une pipe</i> below it. Superficially, the painting looks like a true statement, since it is a picture of the pipe, not an actual pipe. However, the assertion is also part of the picture, which seems to make it false, because it is clearly a painting of a pipe. Paradoxically, the truth seems to depend on whether the statement is an assertion about the painting or a part of it. But, it's both.</p>
</body>
</html>

<figure>
    <img src="paradoxes.jpg" alt="Completed Web page and the HTML source."><br>
    <figcaption><strong>Figure 4.2</strong>. Completed Web page and the HTML source (continues next page).</figcaption>
</figure>

<figure>
    <img src="paradoxes.jpg" alt="Completed Web page and the HTML source."><br>
    <figcaption><strong>Figure 4.2 (continued)</strong>. Completed Web page and the HTML source.</figcaption>
</figure>
CSS - Cascading Style Sheets

<style>
  body { background-color: darkslategray }

  p { color: lightyellow }

  h1 { color: gold; text-align: center }

  h2 { color: darkorange }

</style>
Chapter 5
Locating Information on the WWW
Figure 5.4 NPR hierarchies: (a) Navigation links after clicking on Music.
**Figure 5.4** Google’s Advanced Search window. Notice that text panes are provided for AND-words, quote phrases, OR-words, and NOT-words; the combined query is in the text window surrounded by blue.
Figure 5.1  Crawling over the Green Eye Cat page: The crawler adds the page’s URL to the lists for each word in its title; for words in the anchor text, the link URL is added to their lists.
Crawlers

- **Crawler** visits every website it can find:
  - Identifies all links to other Web pages
  - Checks records if visited recently
  - If not, adds to list of URLs to be crawled
  - Records index of keywords used on page (`<title>`, `<meta>`, `<a...>` `<img alt=.../>`)
Figure 5.1  Sample index entry lists for tokens around “cat” produced by a Web crawler. Some lists are tiny (caszzzzzzzz has one entry) and others are very long—there are more than 2.05 million URLs following cat.
How a Web Search Engine Finds Links

1. A Web site that includes a search engine, such as Google, Yahoo!, or Excite, regularly runs programs called Web crawlers, or spiders, to gather information about what's available on the Internet. The crawlers travel across the World Wide Web by following hypertext links they encounter. Some Web crawlers follow every link they find. Others follow only certain types of links, ignoring those that lead to graphics, sound, and animation files.

2. When a crawler encounters a document, it sends the address of the document—its universal resource locator, or URL—along with the text from the document, back to the search engine's indexing software.

3. The indexing software extracts information from the documents, organizes it into a database, and, based on the frequency of different words found in the document, indexes the information. Like the index in the back of this book, a Web index makes it quicker to locate a specific piece of information. The type of information indexed depends on the search engine. It could include every word in each document; the most frequently used words; words in its title, headings, and subheadings; and the document's size and date of creation. Some search engines employ humans to view and rate sites found by crawlers.
Query Processors

• Gets keywords (search terms / tokens) from user and looks in its index
• Returns a *hit* list (URL, links)
• Answer queries quickly (1/5 second)
• Multi word searches (AND) all words
• Intersects lists (alphabetized)
<table>
<thead>
<tr>
<th>token1</th>
<th>token2</th>
<th>token3</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.ab.com">www.ab.com</a></td>
<td><a href="http://www.aa.com">www.aa.com</a></td>
<td><a href="http://www.rs.org">www.rs.org</a></td>
</tr>
<tr>
<td><a href="http://www.rs.org">www.rs.org</a></td>
<td><a href="http://www.ab.com">www.ab.com</a></td>
<td><a href="http://www.rs.org">www.rs.org</a></td>
</tr>
<tr>
<td><a href="http://www.ru.com">www.ru.com</a></td>
<td><a href="http://www.m.edu">www.m.edu</a></td>
<td><a href="http://www.zz.edu">www.zz.edu</a></td>
</tr>
</tbody>
</table>
Figure 5.2 Illustrating the Intersecting Alphabetized Lists rules: In each step (row of boxes) one or more arrows advance; notice Step 3 where the arrow has advanced in two lists in the same step because the earliest URL is on both lists.
Page Ranking

• Google's idea: PageRank
  – Orders links by relevance to user
  – Relevance: counting links to page
    (more pages link to a page, more relevant page must be)
      • “vote” - each page links to another page
      • Google considers whether "voting page" is highly ranked
What??

• Search engines consider each word *separately*
• Exact phrase, use " "
• Research
  • *Select* search *terms* in query
    • Advanced search, general topic, descriptive terms, refine,
  • Anatomy of a *hit* (how to use)
  • Using hit list (*skim* quickly)
  • Find the page – *locate* desired data
Logical Operators

• **AND, OR, NOT**
  
  – **AND**: Pages containing **both** terms  
    Thai AND restaurants  
  
  – **OR**: Pages containing **either** term (also both terms)  

  – **NOT**: **Excludes** pages with given term  

  • AND and OR go **between** terms (infix operator)  
  • NOT **precedes** term to exclude (prefix operator) “ – “  
  • ( Group search terms )
Filtered Searches

• Pinpoint specific pages

Then narrow your results by...

- language: any language
- region: any region
- last update: anytime
- site or domain: .edu

• Limit search hits on advanced search
Figure 5.4  A diagram illustrating how a Web search’s hits are narrowed by adding additional terms; each circle represents a Web site, the four boxes represent words descriptive of the Web page; the four searches based on the red words, burgundy words, fuchsia words, and purple words collect the pages into the area of that color.
**Figure 5.6** A schematic diagram of how sources relate to an original information source.
Sources

• Open Source – Wikipedia
• Respected Source (authoritative – experts)
• Primary Source – direct knowledge
• Secondary Source – interview primary
• Tertiary Source – watch, read secondary