



Announcements

Project 1 is assigned today
 Assignment 2 assigned today
 * Monday is a holiday, so next week's
 'first lab' (M/Tu) is canceled



Quiz #1

Get out a sheet of paper and pencil /
 pen; put everything else away.



Quiz #1

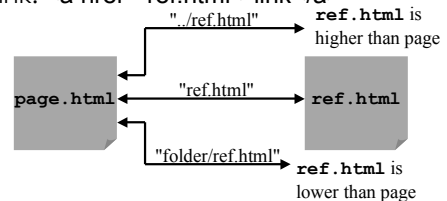
Answer all of the following questions.

1. What is the name of Google's system that decides how important a page is?
2. Which has more hits:
sailboards AND oregon
sailboards AND oregon AND rental
3. Give a query to find pages about Agatha Christie's mystery *Three Blind Mice*



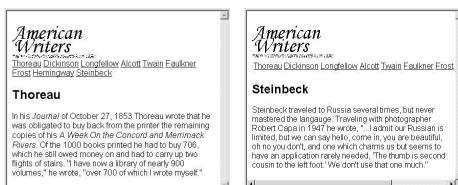
Placing Relative Links

A link in file **page.html** refers to a referenced file **ref.html** by relative link: `link`



Tables Control Position

Tables keep information in a specific relationships to other content



Tables

```

<table>
  <tr>
    <td><a href="hdt.html">Thoreau</a></td>
    <td><a href="ed.html">Dickinson</a></td>
    <td><a href="hw1.html">Longfellow</a></td>
    <td><a href="lma.html">Twain</a></td>
    <td><a href="sc.html">Frost</a></td>
    <td><a href="wf.html">Alcott</a></td>
    <td><a href="rf.html">Emerson</a></td>
    <td><a href="eh.html">Hawthorne</a></td>
    <td><a href="js.html">Stevenson</a></td>
  </tr>
</table>

```



Searching the WWW

Locating the right information on the WWW requires effort



Looking In the Right Place

The WWW is not the first place to look

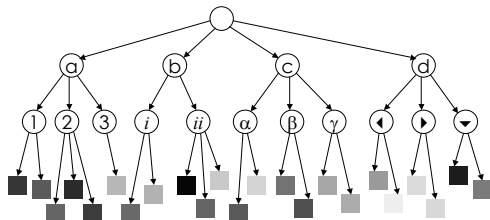
- Go directly to a site -- www.irs.gov
- Guessing a site's URL is often very easy, making it a fast way to find information
- Go to right sight -- dictionary.cambridge.org
- Go to the library -- www.lib.washington.edu
- Go for the kind of information you want -- www.npr.org

Ask, "What site provides this information?"



Hierarchical Organization

Information is organized in hierarchies



Search Engines

No one controls what's published on the WWW ... it is totally decentralized
To find out, *search engines crawl* Web

- * Two parts
 - Crawler visits Web pages building an *index* of the content
 - Query processor checks user requests against the index, reports on known pages

Only a fraction of the Web's content is crawled



Google Advanced



Queries

Searching strategies ...

- * Limit by top level domains or format
- * Find terms maximally specific to topic
- * Look elsewhere for key words, e.g. bio
- * Use exact phrase only when universal
- * If too many hits, requery
- * "Search within results" using "--"
- * Once found, ask if site is best source



Assignment 2

The question of color arose in class:
Why is yellow "full red, full green, no blue" rather than a "primary" color?

Assignment 2

Using Google Advanced Search find the answer. Send email to your TA with the query you used to find the page with the answer, the page's URL, and a short 1 paragraph explanation in your words of why a monitor mixes colors to get yellow

*Due: Before class on Wednesday 1/22



Truth on the Web

- Much Web information is wrong
- Using the Web effectively means recognizing quality information
 - Information from reliable organizations is usually preferred -- check out ownership
 - Look for accuracy, currency, ...
 - Follow links to verify that the content supports the original page

Quote independent verification



A Bogus Site

The Burmese Mountain Dog



Burmese Mountain Dog Guarding
Gandawpala Temple

The Burmese Mountain Dog is a medium sized, muscular dog originally bred in Burma (Myanmar) to guard Buddhist temples. It was bred to guard the temples, and keep the temples free of rodents and beggars. It is also known as the Burmese Temple Dog. In 1954, a group of Burmese Open Lords set up a standard for the Burmese Mountain Dog which has remained virtually unchanged ever since. The Burmese Mountain Dog Club of America was established in 1985 to foster the breed in the United States and the world.

So you want to own a Burmese Mountain Dog?

- The Burmese Mountain Dog is a breed of dog able to guard, ferret small game, and protect property.
- The Burmese Mountain Dogs are remarkably clean dogs. They are easy to keep as they are rarely noisy or

www.mnsu.edu/akcj3/bmd.html



True Site, Bogus Implication

DHMO.org
Dihydrogen Monoxide Research Division

WELCOME

Welcome to the web site for the Dihydrogen Monoxide Research Division (DHMO), currently located in Newark, Delaware. The controversy surrounding dihydrogen monoxide has never been more widely debated, and the goal of this site is to provide an unbiased data clearinghouse and a forum for public discussion.

Explore our many Special Reports, including the **SCIENTIFIC PAPER**, a definitive primer on the subject, plus reports on the **environmental**, **current research**, and a number

www.dhmo.org



Project 1

Create a bogus (fictitious) Web page
To appreciate how easy it is fake quality info you will build a bogus Web page

- Modify photograph, changing its meaning
- Write misleading text
- Add "authenticity" links, fake credentials ...

* Your page should look as legitimate as possible, but contain false information



Project 1 Details on Web

You will ...

- * Write your Web page in HTML
- * Find a photo (or use one of your own)
- * Modify the photo using Photoshop
- * Find links to serve as "authenticity"
- * Milestones
 - Turn In January 24th: Web page
 - Turn In February 3rd: Web page + photo



Here's How It Works

Follow these steps:

- * Create your page locally
- * When finished, publish it on the server
- * Do not touch after the deadline (2:00AM)
- * Print off a copy and turn it in in class
- * Leave your page unchanged until grading is complete

We will check the timestamps