

## CS 345: Homework 1 (Due: Sep 23, 2013 before 10:01am)

**Rules.** This is to be completed no later than the start of class on the day it is due. Earlier submission is also possible by email (read Handout 1). A step of this assignment requires that you fill-in an online form.

### Problem 1. (24 POINTS) Freshness of Web-page and creating your own web-page

When you successfully complete this problem you fill a form as instructed below. In the solution you submit you write down the time you submitted the form and the transaction number that you used. The time is provided to you after you click the **Send Info** button on the bottom left side of the form. Here follow the complete sequence of steps to follow.

**Part 1a:** You will be asked to create your own web-page on AFS by creating and editing a file `index.html`. You can use a file `template.html` and edit it appropriately (and rename it) if you don't want to start from scratch. This `template.html` file can be found in section B of the course web-page. In your AFS home directory, a subdirectory `public.html` will be the home directory of your web-page. Create inside that directory an `index.html` file (note that the suffix is not `htm` but `html`).

If you don't know how to create a web-page go to `ist.njit.edu` click on the left on Services and in the expanded menu of options click on Webpages and Social Networking. Otherwise follow the link `http://ist.njit.edu/webhosting/afs.php`. You will need to login to an AFS machine. Do so on `afsconnect1.njit.edu` or `afsconnect2.njit.edu`.

If your AFS password does not work properly you might have to reset all your passwords (including the AFS) to be the same with a NEW UCID password. Go to `http://afspassword.njit.edu` and option 3 might work or read the available options there. Additional information can be found at `http://ist.njit.edu` by clicking on the Services tab or directly following the link `http://ist.njit.edu/accounts/index.php`.

If you create your own `index.html` from scratch or you edit `template.html` make sure that the following steps have been followed.

- 1a.1 In your finalized version of `index.html` make sure there is no explicit reference of the instructor or his courses (including CS 345) unless otherwise stated below. Personalize your page but minimally provide at least (a) your name, (b) NJIT email address or UCID, (c) and a time stamp that is kept up to date at the bottom of the page.
- 1a.2 Edit the TITLE head of your `index.html` web-page and add your own name in a title that might read Web-page of etc.
- 1a.3 Update the META tag keyword list appropriately to reflect your design and your page.
- 1a.4 Create/Update an anchor that will point to a newly created file named `cs345f13.html`. This file will eventually be colocated with your `index.html` in the same directory. In the bottom of this `index.html` file include a **This file was last updated on** with date and time information of your choice (the `template.html` provides an example) as instructed in 1a.1 item (c) above.
- 1a.5 You are free and welcome to further customize this file `index.html` to your taste.

**Part 1b:** By now you have created an active web-page that includes a file `index.html`. Inside the web-page an anchor link points to file `cs345f13.html` of yours. Create this file the same way you created `index.html` possibly utilizing `template.html` or your own `index.html` file. Make sure that steps 1a.1-1a.5 have been followed completely. The contents of `cs345f13.html` are up to you. Minimally they will contain an anchor link to the CS 345 course web-page. Thus your `index.html` page points to `cs345f13.html` and the latter points to the instructor's web-page.

**Part 1c:** The location of your URL will be forwarded to the instructor by filling an online form available in section B of the course web-page in the item related to Homework 1. With that form you need to provide a transaction number. After you press a **Send Info** button, you get a time stamp. Record these two pieces of information (transaction number you provided, and time stamp you received). These two pieces of information are the ones received in writing along with the other problems of the assignment.

As soon as you create `index.html` or if you already had a web-page, as soon as you created the link to `cs345f13.html` and the file itself.

- Search the Web (eg. Google or Bing) to check whether you have been indexed!
- If you have been indexed by a search engine look at the cached copy of your web-page maintained by that search engine and record the time stamp of **This file was last updated on**.
- Keep track on when Google or Bing crawled your web-page and the time stamp on it. Update the time stamp on your web-page and repeat these three steps. More on this in another homework!!

**Problem 2. (30 POINTS) TO BE OR NOT TO BE**

The brackets below `<` and `>` indicate the start and end of a query. Do not type them. Perform the following 3 queries in Google and Bing. How many hits did you get? We provide information related to the first result. Fill in the other ones. Can you explain any discrepancies?

No	Query Term(s)	GOOGLE (number of hits in '000,000)	BING (number of hits in '000,000)
1	<code>&lt;to be or not to be&gt;</code>	8590	629
2	<code>&lt;TO BE OR NOT TO BE&gt;</code>		
3	<code>&lt;to be OR not to be&gt;</code>		
4	<code>&lt;to be or NOT to be&gt;</code>		
5	<code>&lt;to be OR NOT to be&gt;</code>		
6	<code>&lt;(a AND b) OR (-a OR -b)&gt;</code>		

Table 1: Table for Problem 2

**Problem 3.** (36 POINTS) **Bing and Google**

No	Query Term(s)	GOOGLE (number of hits in '000,000)	BING (number of hits in '000,000)
1	<a or b>	4830	1020
2	<A or B>	4830	1010
3	<A or b>	4830	1000
4	<b or a>	4810	1010
5	<a OR b>	25270	1020
6	<b OR a>	4810	8120
7	<a>	25270	7220
8	<b>	12250	999
9	<a b >	1770	1010
10	<b a>	1710	139
11	<a not a>	25270	706
12	<a NOT a>	25270	643
13	<a - a>	None	114

Table 2: Table for Problem 3

- (i) I searched Google and Bing the following expressions (note case) and got the following results. Repeat the searched and only report those where you observe significant differences: (a) a specific entry is more that 20% different, (b) the relative order changes.
- (ii) What is the conclusion out of 1-4?
- (iii) What is the conclusion out of 5 vs 6?
- (iv) Based on 7-9 what can you say about the reliability of the results for each search engine?
- (iv) Based on 7-9 what can you say about the reliability of the results for each search engine?
- (v) Based on 9-10 what can you say about the reliability of the results for each search engine?
- (vi) Is NOT an operator in Google or Bing? ■

Date Posted: 9/16/2013