Nile Grice

I intend to make a browser extension that allows the user to search over the full contents of their browsing history. This would be useful to me, personally, because instead of making bookmarks, I rely on the browser's history function to revisit web pages I've found interesting or useful. Sometimes, though, the browser's built-in title-only history search isn't sufficient for me to find the page I'm looking for. I believe that, if history search considered the full content of pages browsed in the past, that this issue would be significantly lessened.

This project, being a search engine, touches on a lot of what was discussed in the first half of this course. The task breakdown would be something like:

- Create a program external to the browser that takes <URL, CONTENT> input, and constructs an index (Unsure if it's best for this to be internal or external to the extension. If I make it external, it will probably leverage lucene and be in python/java)
- On page load, send the full text contents of the page to the program for indexing (extension in javascript)
- Choose some retrieval function and use it to get URLs back from the index based on queries
- Create either an extension dialog or web page from which the user can enter their query and receive results

As I am working alone, however, most of the heavy lifting (index construction, document ranking) is likely to be left to a preexisting, yet to be chosen, library. Even then, I have never worked with browser extensions or any search libraries, so I don't doubt this will take over 20 hours.

To demonstrate that the extension works, I can start from a new browser profile, and load a few pages. I can then attempt to use history search to try to find a page using text not present in the title of the page, but relevant to/present in the content of the page. It will fail with the inbuilt history search, but should return the desired document with the extension search.