coviz User's Manual

`usage: python coviz.py [reconcile output] [source text] [muc annotations]`

Developed by: the University of Utah's NLP Research Group [http://www.cs.utah.edu/nlp/]
Repository hosted at: http://github.com/DrSleep/coviz
Bug reports: clemmer.alexander@gmail.com

Description

`coviz.py` is a visualization tool created specifically for the University of Utah's Reconcile program (http://www.cs.utah.edu/nlp/reconcile/), and designed to help users understand how similar two interpretations of some text document are. The program is designed to show to interpretations of a document side-by-side, allowing users to see both what a noun phrase is coreferent to in the current interpretation of the document, and what coref chains that noun phrase is a part of in an alternate interpretation of the document.

1 What you need to use the program

The main goal of the project was to have *one* script that generates *one* HTML file that contains everything it needs to do the visualization work. This way, users need only a browser to use the visualization tool.

Because of these things, the following is **required**:

- Python (to run the script)

- Some sort of web browser, preferably as current as possible (to view the visualization)

- JavaScript enabled inside this browser

2 Using the program

2.1 Running the program

The **general form** of the command is:

`python coviz.py [reconcile output] [source text] [muc annotations]`

If you'd like to see an **example** using the test data provided, then you can generate the output file by using this command:

`python coviz.py coref_output raw.txt muc_annots`

This outputs a file called `out.html`. **Open this** in any browser to begin interacting with the visualization.

2.2 Making sense of the output

Once you've opened it, you'll see two interpretations of the same document placed directly next to each other. It should look something like this:
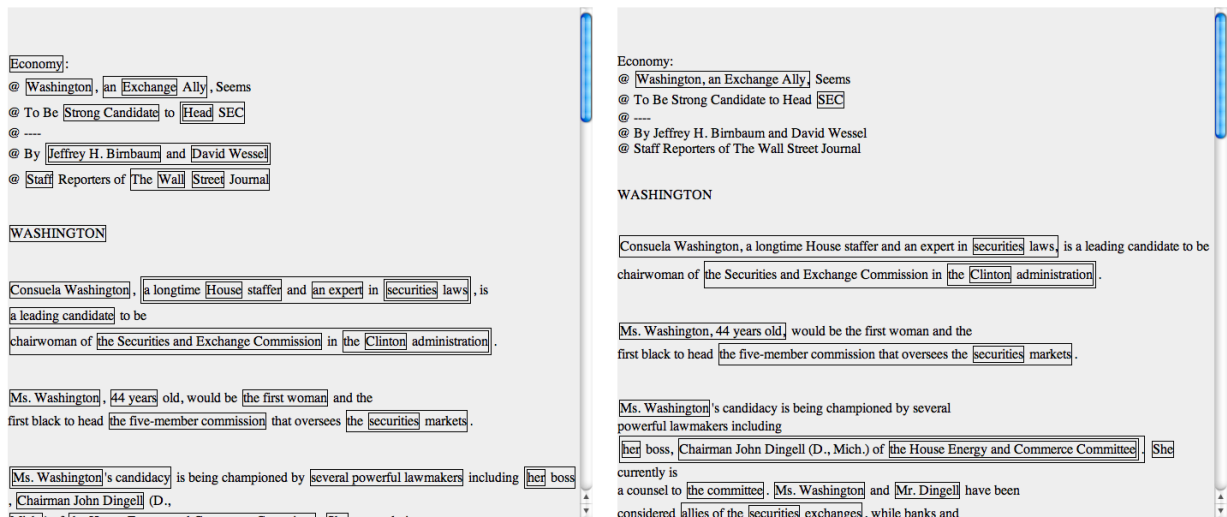
Each of these interpretations has lots of little boxes inside it. Each of these represents one noun phrase (see the next picture). Note that NPs can be nested arbitrarily many times.
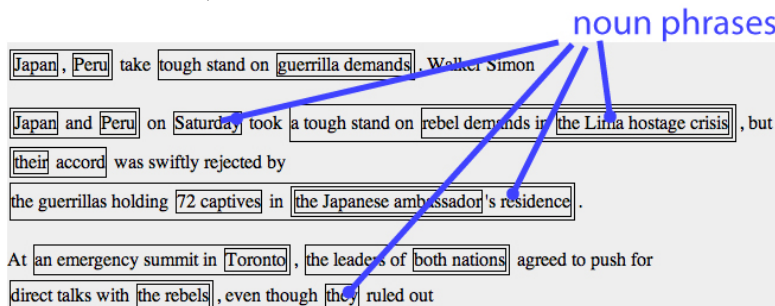
2.3 Clicking on NPs

The basic idea behind this visualization tool is that we want to compare how some NP manifests across two interpretations of some document. So, when users click on a document, obviously all NPs that are coreferent to that particular NP should be highlighted in that interpretation. But we also want to see what coref chains it is a part of in the other interpretation, too! So when users click on any noun phrase, two things should happen:

- All coreferent noun phrases in the interpretation you clicked on will be highlighted. An example follows, although note that the noun phrase on the right has another (separate) noun phrase inside of it.

- Any of the opposing interpretation's coref chains that have any of the words in common with the noun phrase you selected will also be highlighted. Here's an example:

One big noun phrase

Consuela Washington , a longtime House staffer and an expert in securities laws , is ... leading candidate ... be chairwoman of the Securities and Exchange Commission in the Clinton administration .

many NPs that have words in common with this big one

Consuela Washington, a longtime House staffer and an expert in securities laws , is a leading ... car chairwoman of the Securities and Exchange Commission in the Clinton administration .