Memoplex (Browser)++: A Semantic Document Browser

Clarence Chan
CPSC 533
Nov. 12, 2006

Memoplex Browser++

- ***Text document browser***
  - *Uses semantic links, spatial arrangement*
  - *Color-based clustering for easier identification*
- ***Can view local information more easily, but no advantage over a flat list of results***
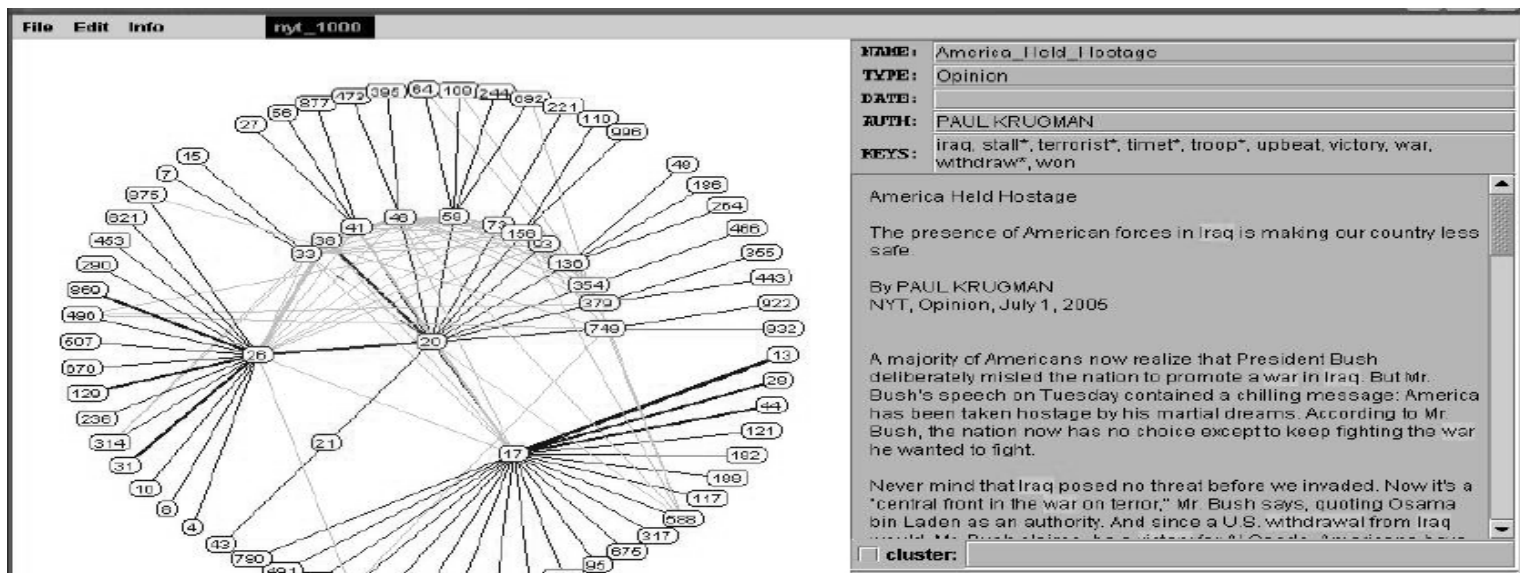
The problem

- Directed search: find docs easily if
  - we have the document title
  - we know exactly what keywords are involved
  - ***We know what we're looking for***

- ***What if we don't exactly know?***
  - ***look through a list of titles***
  - ***"feel around", browse***

Navigation, browsing

- Analogy:
  - Directed search: Look up exact item in library call number system
  - Browsing: Go to a shelf and look around at various books
- What if we combine these?
  - Go to a bookshelf in a general call number range
  - Browse around various titles and books that are related

Memoplex Browser, v.1

- Originally a 533 project from a previous year

  - Built on top of Mike Huggett's Memoplex server

  - Backend: Semantic network of documents

  - Frontend: node-link graph of network

  - Nodes are individual documents
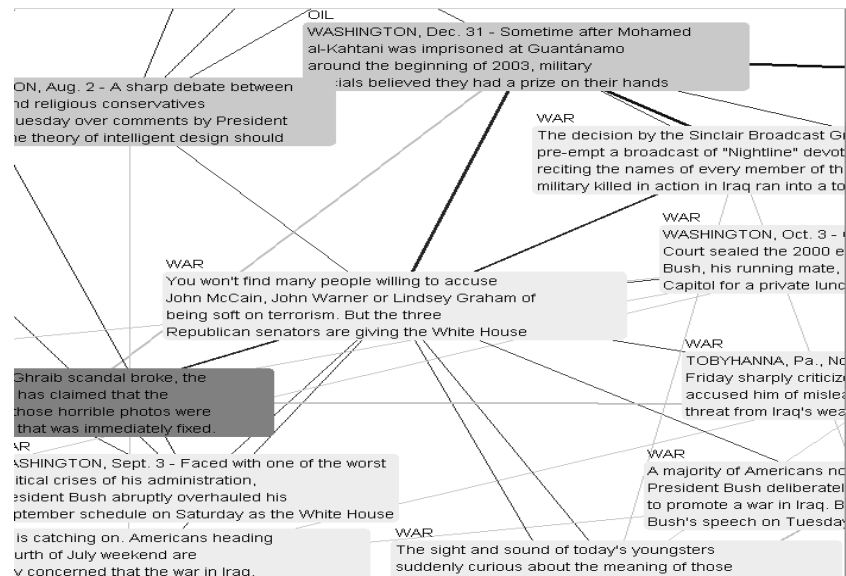
My contribution

- Make the system useful beyond visualizing structure

- Graph structure (browsing) augmented to support the search functionality
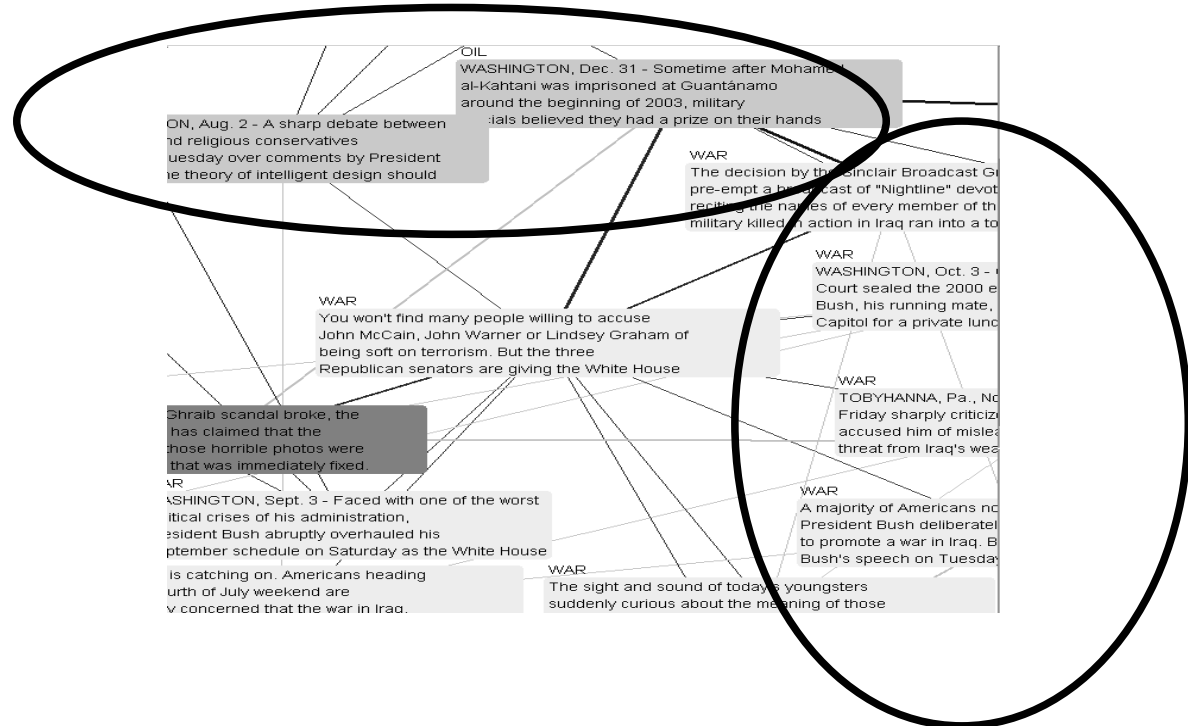
- Built in *prefuse*

Memoplex Browser, v.2

- Attempts to make browsing of ASN more useful / relevant
    - View actual document snippets, only a few at a time
    - Idea: following local links / highly connected links makes for simpler exploratory search

Memoplex Browser, v.2

- Takes up to three keywords in search
    - Finds a set of documents in search pane
    - Doubleclick: opens graph view
    - Elements are spatially organized / coloured into "clusters"

OIL

WASHINGTON, Dec. 31 - Sometime after Moham
al-Kahtani was imprisoned at Guantánamo
around the beginning of 2003, military
...ials believed they had a prize on their hands

ON, Aug. 2 - A sharp debate between
nd religious conservatives
uesday over comments by President
e theory of intelligent design should

WAR
The decision by the Sinclair Broadcast Gr
pre-empt a broadcast of "Nightline" devot
reciting the names of every member of th
military killed in action in Iraq ran into a to

WAR
WASHINGTON, Oct. 3 -
Court sealed the 2000 e
Bush, his running mate,
Capitol for a private lunc

WAR
You won't find many people willing to accuse
John McCain, John Warner or Lindsey Graham of
being soft on terrorism. But the three
Republican senators are giving the White House

WAR
TOBYHANNA, Pa., No
Friday sharply criticize
accused him of misled
threat from Iraq's wea

Ghraib scandal broke, the
has claimed that the
hose horrible photos were
that was immediately fixed.

AR
ASHINGTON, Sept. 3 - Faced with one of the worst
itical crises of his administration,
esident Bush abruptly overhauled his
ptember schedule on Saturday as the White House

WAR
A majority of Americans no
President Bush deliberate
to promote a war in Iraq. B
Bush's speech on Tuesday

is catching on. Americans heading
urth of July weekend are
v concerned that the war in Iraq.

WAR
The sight and sound of todays youngsters
suddenly curious about the meaning of those

Memoplex Browser, v.2

- Interaction

    – Double-click brings item into focus, re-orients nodes

    – Also displays it in content pane

    – Simplified panning controls?

Compared to Google Desktop ...

- *Flat list simpler to navigate - familiar*

- *Spatial layout – very easy to get lost*

- *Clustering / search results were skewed*

- *Good for orienting to local (1 hop away)*

- *Easy to find heavily referenced articles*

Conclusion: Recap

- *A text document browser that uses semantic links and spatial clustering to aid in browsing tasks*

- *Very tricky to do spatial layout of text*

Conclusion

- *Exploiting intricacies of network may not be necessarily useful*

- *List form is easier to navigate*

- *However, some kind of indication of "location" and relative referencing is useful*

- *May have better applications in other domains*