Visualizing links between sentences in asynchronous conversations:
a new feature for ConVis

Jordon Johnson
M.Sc. Student
UBC Computer Science
jordon@cs.ubc.ca

“How do you kill that which has no life?” – South Park
Background: Asynchronous Conversations

• Participants are not constrained to a common timeframe; **can insert a new reply at almost any point** in the existing discussion

• Examples:
  – email threads
  – blogs with comments and replies to comments
  – **reddit** posts with comment trees
Background: ConVis @ UBC

http://www.cs.ubc.ca/cs-research/lici/research-groups/natural-language-processing/image/convis/3.jpg
Background: Links Between Sentences

“Sentence B refers to sentence A”

Link Attributes:
- Linked sentences
- Agreement value
- Sentiment value

GOAL: extend ConVis so users can visualize/edit links
Data

• Online articles with comments, replies to comments, etc.
  – Sentences are indexed
  – XML tags provide a “reply-to” tree structure

• **Links**: annotations by humans or algorithms
  – Currently using sets of “gold standard” links, annotated by humans

Image source: http://www.zastavki.com/
Task 1: Evaluate/Revise Links

- Some of the gold standard links have **issues**:
  - Incorrect sentiment/argument values
  - Same pairs of sentences, different sentiment/argument values
  - One sentence is a single word ("you?") or entirely punctuation ("?"")
  - Unlikely given positions in conversation
  - Pointed in the wrong direction

Image source: http://www.wow247.co.uk/
Task 1: Evaluate/Revise Links

• We’d like to be able to:
  – Improve the existing gold standard through revision
  – Add to the gold standard by evaluating/revising the output of candidate linking algorithms
Task 2: Evaluate Article Effectiveness

• Which parts of the article sparked the most commentary?

Task 3: Explore Link Chains

• Can we learn something useful from chains of links?
Idiom: Arc Diagram

I employ a modified version by using the ConVis glyphs as the spine
Murphy’s Law Slides: Task 1

ha ha ah aha - do you really believe the people you vote for are in control?

But I'm still never going to vote LibDem.

I agree.
He's admitted his mistakes and taken a lot of vitriol by coming
Not many disgraced MPs would have the bottle to that.
If any.
But I'm still never going to vote LibDem.
Cabinet was told nothing about GCHQ spying programmes, says Chris Huhne. Cabinet ministers and members of the national security council were told nothing of the existence and scale of the vast data-gathering programmes run by British American intelligence agencies, a former member of the government has revealed.

Chris Huhne, who was in the cabinet for two years until 2012, said ministers were in "utter ignorance" of the two biggest covert operations, Prism and Tempora. The former Liberal Democrat MP admitted he was shocked and mystified by the surveillance capabilities disclosed by the Guardian from files leaked by the whistleblower Edward Snowden.

"The revelations put a giant question mark into the middle of our surveillance," he said.

"The state should not feel itself entitled to know, see and memorise every private citizen communicates. The state is our servant."

"Writing in Monday’s Guardian, Huhne also questioned whether the Home Office deliberately misled parliament about the need for the communications data. GCHQ, the government's eavesdropping headquarters, already had remote extensive snooping capabilities.

He said this lack of information and accountability showed "the supervisory arrangements for our intelligence services need as much updating as their techniques."

Over the past three months the Guardian has made a series of disclosures about the activities of GCHQ and its much bigger American counterpart, the National Security Agency.

Two of the most significant programmes uncovered in the Snowden files were run by the NSA, and Tempora, which was set up by GCHQ. Between them, they allow the agencies to harvest, store and analyse data from millions of phone calls, emails and search engine queries.

As a cabinet minister and member of the national security council (NSC), he would have expected to be told about these operations, particularly as the
Murphy’s Law Slides: Task 3

Cabinet was told nothing about GCHQ spying programmes, says Chris Huhne.

If Obama isn’t really in control of the US and Cameron isn’t really in control, then who the f**k is?

Whoever it is, they certainly weren’t democratically elected to those positions.

If I don’t know what they propose to do, then my MP cannot raise any objection or any other constituent may have, which seems very undemocratic.

There seems to be an arrogance about this - it’s as if they know what’s best, nobody could possibly point out any flaws in their thinking, or highlight any flaws with their proposal.

I think this snooping will be found to be illegal by the ECHR.

Have Labour pledged to stop the mass snooping?

No, didn’t think so.

How can a Cabinet & NSC member not have known about Prism & Tempora?

Who’s governing us?

Why would the security services inform the Cabinet?

When we have seen a Minister for Defence travelling around the world with security vettet ‘friends’ and observe that unteen MPs are members of ‘friends’ wherever, the last people you’d tip off are the subjects of the surveillance.

Because democracy depends on an informed public, that happens to include members.

Well I’d like to move on from Huhne’s past. And course everyone has a right to keep throwing the past at the man, but if you will lose the opportunity to see if the person has changed.

MI5/6, FBI/CIA, The MOSSAD, and most secret services, are all totally out of remit of governments.

I discovered the next day that it had happened to other people involved with me specifically at NUS (sea-men) offices.

These people had a firing squad in their offices and GCHQ etc.
Murphy’s Law Slides: Task 3

Cabinet was told nothing about GCHQ spying programmes, says Chris Huhne.

Why would the security services inform the Cabinet?

Huhne would have been vetted as a matter of course.

Well I’d like to move on from Huhne’s past. And Mr Huhne’s revelations should prove vital in helping to move this unhealthy surrounding GCHQ further out into the open.
<table>
<thead>
<tr>
<th>Idiom</th>
<th>Modified Arc Diagram (using ConVis glyphs as a spine)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What:</strong> Data</td>
<td>Indexed sentences, grouped into article/comments <em>(tree)</em>&lt;br&gt;Links with attributes:&lt;br&gt;- Indices of linked sentences <em>(ordered)</em>&lt;br&gt;- Sentiment <em>(ordered, diverging)</em>&lt;br&gt;- Argument <em>(ordered, diverging)</em></td>
</tr>
<tr>
<td><strong>What:</strong> Derived</td>
<td>Relative positions of sentences in article/comments <em>(ordered)</em></td>
</tr>
<tr>
<td><strong>How:</strong> Encode</td>
<td>Curved line marks connecting relative positions of linked sentences in their respective article/comments&lt;br&gt;- Relative positions in article/comments: vertical positions on corresponding ConVis glyph&lt;br&gt;- Sentiment: line color&lt;br&gt;- Argument: line stroke pattern</td>
</tr>
<tr>
<td><strong>How:</strong> Facet</td>
<td>Use ConVis comment view to display sentences related to selected links</td>
</tr>
<tr>
<td><strong>How:</strong> Manipulate</td>
<td>Highlight selections; use animated transitions between different highlighted states</td>
</tr>
<tr>
<td><strong>How:</strong> Reduce</td>
<td>Filter links displayed based on selected article/comment or sentence(s)</td>
</tr>
<tr>
<td><strong>Why:</strong> Tasks</td>
<td>Annotate: Revise or delete existing links&lt;br&gt;Browse: View number and types of links from article/comment/sentence of interest&lt;br&gt;Explore: Find and examine chains of links&lt;br&gt;Discover Features: Examine attribute patterns in chains of links</td>
</tr>
<tr>
<td>Scale</td>
<td>Hundreds of sentences in the article and comments; hundreds of links between sentences</td>
</tr>
</tbody>
</table>
The **Good**, the **Bad** and the **Ugly**

- Allows users to accomplish all three tasks
- Seems to work well within ConVis constraints
- Current interactions are limited
- Current interactions are inconsistent; makes the vis less intuitive
- No aggregation or filtering based on link type
- Doesn’t currently fit on a laptop screen
- Links between nearby sentences are hidden in some cases
- “Hairy” clutter not always well handled yet
- Text view looks like a 6-year-old learning HTML
// TODO Future Work

- REFACTOR/IMPROVE (optimization, quality-of-life improvements)
  - Make interactions more consistent for the users
  - Make links between nearby sentences less likely to pass through comment glyphs
  - Improve contrast between selected link and other links in some cases
  - Attempt scaling algorithms to fit entire vis on standard laptop screen
Future Work

• Further integration with other ConVis features to allow additional tasks
• Aggregation/filtering by link type
• Zoom capabilities
• Additional ways of filtering links to facilitate studying link chains
• Implement saving link revisions to file
• Add sliders to text view for amount of context to be shown
• USER STUDIES
# Acknowledgements Matrix

<table>
<thead>
<tr>
<th></th>
<th>Design feedback</th>
<th>Debugging assistance</th>
<th>Language reference (D3, JavaScript, etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giuseppe Carenini</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enamul Hoque</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamara Munzner</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaden Masrani</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thank you! Any questions?