Formalization, User Strategy and Interaction Design:
Users’ Behaviour with Discourse Tagging Semantics

Bertrand Sereno*, Simon Buckingham Shum & Enrico Motta
Knowledge Media Institute, The Open University, Milton Keynes, UK

* Now at: Centre for Advanced Learning Technologies, INSEAD, Fontainebleau, France

Acknowledgements:
Overview

- **The problem:** collaboration semantics in contested domains — no consensus assumed; possibly not even possible

- **Previous work:** Scholarly discourse as Collaborative Knowledge Structuring (CKS)
  - Modelling and querying Web argument structures

- **How to help users engage in CKS?**
  - Evaluating the ClaimSpotter tool

- **Summary of evaluation results and design principles**
  - Formalization / User Strategy / Interaction Design
Scholarly discourse as CKS...

From:

Le Journal des Scavans
January 1665

Philosophical Transactions of the Royal Society of London
March 1665

To:

Chaomei Chen, 2006: Citation analysis

Buckingham Shum et al, 2003: lineage analysis
These annotations are freeform summaries of an idea, as one would also find in researchers’ journals, fieldnotes, lit. review notes or blog entries.

Making formal connections between ideas creates a semantic citation network — novel literature navigation, querying and visualization.

“People try to maximise their rate of gaining information.”

“Web User Flow by Information Scent (WUFIS)”

“Information foraging theory”

“Information foraging models”

Paper: “The Scent of a Site: A System for Analyzing and Predicting Information Scent, Usage, and Usability of a Web Site”

Paper: “Information foraging”
Scholarly discourse as CKS...

Connecting freeform tags with naturalistic connections ("dialects") grounded in a formal set of relations (from semiotics and coherence relations)
Cognitive Coherence Relations—towards a relational ontology for discourse

Scholarly discourse as CKS...
Visualising claims and arguments

The link-tracking service allows the user to specify structures
  e.g. show tags one link out from any tag on the left hand of a claim containing the string “CiteSeer”

Scholarly discourse as CKS... Querying on argumentation structures

machine learning

Perspective in contrast agree

Neural network text categorizer Depth 10 Lineage

machine learning Depth 10 Descendants
Scholarly discourse as CKS...
“What papers contrast with this paper?”

1. Extract concepts for this document
2. Trace concepts on which they build
3. Trace concepts challenging this set
4. Show root documents

<table>
<thead>
<tr>
<th>The key issues you are concerned with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>445</td>
</tr>
<tr>
<td>Decision Forest classifier</td>
</tr>
<tr>
<td>446</td>
</tr>
<tr>
<td>Decision Forest classifier improves on C4.5 and kNN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The related issues you may be concerned with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>446</td>
</tr>
<tr>
<td>Decision Forest classifier improves on C4.5 and kNN</td>
</tr>
<tr>
<td>515</td>
</tr>
<tr>
<td>Instance based learning</td>
</tr>
<tr>
<td>511</td>
</tr>
<tr>
<td>Decision tree learning</td>
</tr>
<tr>
<td>277</td>
</tr>
<tr>
<td>decision trees and naive Bayes perform well for text categorization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The following claims disagree ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Optimised rules outperform Naive Bayes and decision trees disagrees with decision trees and naive Bayes perform well for text categorization</td>
</tr>
</tbody>
</table>

The point is... we think these kinds of structures are worth having

*But can users create them?*
How to help scholars engage in CKS?
Pilot study: paper-based literature modelling

How to help scholars engage in CKS? 
From paper prototype to semiformal mapping tool

- The **ClaiMapper** tool

How to help scholars engage in CKS?

Pilot study: paper-based annotation

How to help scholars engage in CKS?

- The **ClaimSpotter** annotation tool: Web 2.0-style tagging with optional community/system tag recommendations
A user-centred approach to the formative evaluation of a CKS tool

- **Research question:**
  - how do annotators approach the task of using a new Web tool to semantically annotate a document with its key contributions?

- **Focus**
  - ..is on how the tool functionality and UI affordances shape tagging behaviour

- **Quantitative and qualitative analysis**
Example *claims* (tag triples) from users

- Domain ontology *is about* A hierarchy of URIs on multiple levels
- Universal physical access *is unlikely to affect* Digital divide
- Hypertext node juxtaposition *is analogous to* Cinematic shot juxtaposition

- [Evidence] In the Bristol trial, the awareness of the presence of other players was correlated with how much our participants enjoyed the game as well as with how engaged they felt *is consistent with* Presence awareness of many other people is capable of causing, feel good factor

- Magpie moves away from hypermedia towards open service-based architectures *is evidence for* [Magpie *improves on* COHSE]
User study: selected results

- See paper for details
- and the PhD for complete account

Tag length similar for novices and experts (64% 1-3 words)

Figure 6.2: Concepts length distribution graph for all the annotators. 164 concepts out of 257 are composed of three words or less.

Tag reuse
most of them used only twice in this study (1 hour)

<table>
<thead>
<tr>
<th>Reused concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>a community-based project that wired four computing centres (hubs) in a lower socio-economic urban area</td>
</tr>
</tbody>
</table>

Table C.3: Concepts reused by the annotators.

Transcript analysis
to derive themes, sub-categories and codes

Figure 6.7: A coding session in TAMS: text chunks (main window, right side) are selected and assigned a code (selected from the bottom-left window, or created from scratch.)
Themes from the user study

- **Formalization**
  - the cognitive task of structured tagging

- **User Strategy**
  - how users approach the semantic annotation task

- **Interaction Design**
  - how behaviour is shaped by the tool’s affordances
Formalization

the cognitive task of structured tagging

- Looking for the right tag type...
Formalization
the cognitive task of structured tagging

- Looking for the right link type...
User Strategy
how users approach the semantic annotation task

- What granularity and type of claims? When to stop?

Magpie: Supporting Browsing and Navigation on the Semantic Web

John Damianopoulou & Mario Corcho

ABSTRACT

The semantic annotation of web pages is a manual process that requires users to annotate web pages by means of relationships that are defined in a specific domain. This paper describes an automatic tool called Magpie that supports this process. Magpie automatically assigns annotations to web pages based on the content of the page and the user’s preferences.

INTRODUCTION

A lot of research has gone into supporting the task of finding web resources by means of relationships that are defined in a specific domain. This paper describes a tool called Magpie that supports this process. Magpie automatically assigns annotations to web pages based on the content of the page and the user’s preferences.

TABLE OF CONTENTS

- Introduction
- Related Work
- Magpie
- Conclusions
- References

REFERENCES

User Strategy

how users approach the semantic annotation task

- No initial use of tagging aids — focus is on own tags
Interaction Design

how behaviour is shaped by the tool’s affordances

- ‘Flip’ left/right tags to match the link type
Interaction Design

how behaviour is shaped by the tool’s affordances

- Skimming highlighted text
Lessons Learnt & Design Principles

- **Untrained users can do it:** in their first hour they created coherent claims. UI design validated to this degree.
  —future work: longitudinal evaluation at scale

- **New users attend to what is highlighted** for them (matching tags; primary doct.), and generally don’t click down a level
  —next version combines visualizations and document-centric features

- **Support incremental formalization**
  —cf. use of *is-about* as a placeholder link; provide an *Other...* category and try to map automatically to the ontology

- **Users’ strategies vary** — don’t assume a strong workflow
  a paper-based pilot study can provide insights into this

- **Web 2.0 UI simplicity:** good design needed to provide high functionality, walk-up-and-use CKS tools
  —we overwhelmed some users with overlaid suggestions for tags
ClaimSpotter:
papers and demos
http://kmi.open.ac.uk/projects/hyperdiscourse/tools/claimspotter

Hypermedia Discourse project:
theories / tools / case studies / user studies: face-face and asynch. interaction
http://kmi.open.ac.uk/projects/hyperdiscourse

2nd International Conference on
The Pragmatic Web
22-23 Oct., Tilburg, NL
collaboration / semantics / usability / community informatics / argumentation
http://www.PragmaticWeb.info
Short/full paper submission deadline: 14 May