



An approach to bridge the gulf from free-text annotations to ontology-supported interpretative claims.

Bertrand Sereno, Simon Buckingham Shum, Enrico Motta

► To cite this version:

Bertrand Sereno, Simon Buckingham Shum, Enrico Motta. An approach to bridge the gulf from free-text annotations to ontology-supported interpretative claims.. Jun 2004. sic_00001232

HAL Id: sic_00001232

https://archivesic.ccsd.cnrs.fr/sic_00001232

Submitted on 7 Dec 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An approach to bridge the gulf from free-text annotations to ontology-supported interpretative claims.

Bertrand Sereno, Simon Buckingham Shum and Enrico Motta
Knowledge Media Institute
The Open University
Milton Keynes MK7 6AA
United Kingdom

E-mail: {b.seren, s.buckingham.shum, e.motta}@open.ac.uk

Abstract

The work described here looks at the difficulties hidden behind the transformation of typical scholarly document annotations (expressing their contributions or their connections) into ontology-supported knowledge triples, and presents our approach to assist this process, through the association of different text analysis techniques in a supportive interface, ClaimSpotter. Finally, we also introduce why typical annotation evaluation measures cannot be applied here and motivate the need for a different approach.

1. Introduction

The Scholarly Ontologies project [BSMD00] (or ScholOnto) proposes an alternative way to organize and browse scholarly documents, by developing a system where authors' and annotators' arguments are represented explicitly, to form an interpretational layer sitting on top of these documents. This layer results from a document interpretation (and annotation) process, and its semi-formal nature (more about this below) is expected to help answering questions such as “Where does this particular idea come from?” or “Who has taken a position against that particular argument?”, questions whose answers are at the heart of any research process, but for which there is only limited support offered at the moment.

These interpretations may for instance summarize the core contributions of an article and/or its connections to related work, which are deemed relevant in the eyes of an annotator. They are formalized as triples (or claims) $\langle \text{node}, \text{relation}, \text{node} \rangle$, where the nodes can be chunks of text or (typed) concepts (like a theory, a methodology or an approach), and the relation is an instance of a class defined

in a formal ontology of discourse, which organizes the way interpretations can be articulated; figure 1 gives some examples of the relations which can be drawn between nodes.

For a single document, different elements can be retained and reused by different annotators, and connections can be drawn between nodes contained in different documents or within a unique document. Annotators are encouraged to make links to concepts backed by other documents and to reuse concepts. They may extend the models built by other contributors, adding further claims, or take issue with them if they feel the original interpretation is flawed.

However, we expect that moving from utterances expressed in natural language (*I believe the document [li02claimaker] describes a mechanism to enhance documents with machine understandable information, which supports the notion of Semantic Web, as introduced in [bernerslee01semantic]*) to a set of ScholOnto claims (cf. figure 1) is not going to be straightforward, as annotators will have to translate their opinions in a claim-compatible form, which implies translating them, making them ‘fit’ in the canvas of relations.

The first problem users might witness lies in the decision of what to actually use as nodes and relations, and the level of detail these elements should carry, as noted by Shipman and McCall: “Users are hesitant about formalization because of a fear of prematurely committing to a specific perspective on their tasks; this may be especially true in a collaborative setting, where people must agree on an appropriate formalism.” [SM94]. The added level of formality also means that opinions will appear more clear-cut, which could mean that many annotators might be ‘forced’ to make stronger statements than they wanted to.

Our task is therefore not really to solve these problems, which are inherent to the problem being addressed, but instead to look at ways to potentially reduce them, by providing assistance to the annotators. In other words, by helping

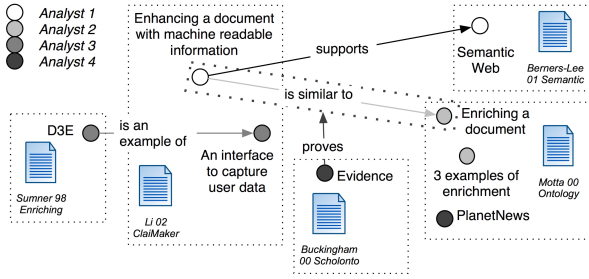


Figure 1. Multiple interpretations encoded as sets of ScholOnto claims, expressing contributions and connections to and from documents.

them bridge the formalization gulf between the freedom of expression of an annotation and the rigidity imposed by the canvas of relations.

2. Active Recommendations

The assistance we are hoping to bring is composed of text analysis techniques wrapped in an interface initiating a dialogue between the annotator and the repository of claims. A crucial aspect to mention is that we are not interested in actually summarizing a scholarly document and ‘reducing’ it to a number of claim-worthy elements that will be suitable for every user. Indeed, interpretations being of course personal, any particular aspect of it might be of interest to at least one annotator. We are instead trying to identify components [Bis98] which might be of more interest, but we are also keeping the whole document for consideration. We would like to coin the term ‘recommendation’ as a way to described these components (a similar view can be found for instance in [LMR⁺03]).

These recommendations support the first step of a dual-annotation process, composed of an annotation with ‘simple’ claims, for which machine tools can help by spotting potentially relevant claim elements or valuable areas of the document; and in a second step, an annotation with ‘complex’ claims, which result from a human sense-making process.

The recommendations are based on the analysis of an observation study we carried out to get some insight on the different needs of annotators, when faced with the task of inputting their interpretations [SSM04]. Following this analysis, we proposed an architecture to plug in different text analysis parsers (cf. figure 2) and we have implemented preliminary versions of some of these:

- We are looking first of all for existing concepts and relations, as defined by all the annotators of the doc-

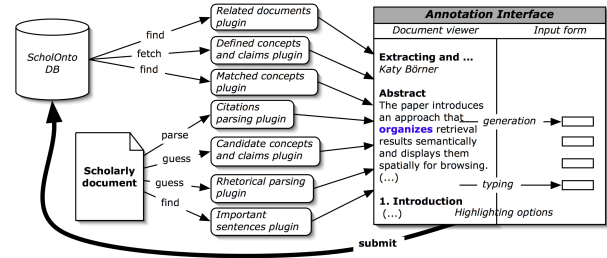


Figure 2. The recommendation system architecture.

ument under consideration; we are also considering the concepts defined in any document, which we can match in the text (using regular expressions).

- Secondly, we are looking at citations, and try to get some insight on the author’s intention by extracting a window around the citation signal (matched with a regular expression) and looking at some predefined keywords to guess their motivation. Once these documents are identified, we can get their claims and concepts and propose them. We are also fetching the same information for what we called ‘related documents’, which are the documents connected through a claim (but not necessarily through a “cites/cited by” relation).
- We are identifying important areas in the document, by looking at passages (sentences or sets of sentences) containing words from the title, from the headers or from the abstract of the document.
- Basing ourselves on the assumption that authors have to defend their position and their contributions, and relate them (through praise or criticism) to the positions of their colleagues [Swa90], we would like to get as much insight as possible from the document authors’ intentions, and from what they wanted to express. The observation study has shown that the ability to guess the role played by a sentence in this position-defending process was providing a valuable resource in the task of interpretation: interpreting a document also means positioning oneself (by agreeing or disagreeing) with the research carried out and presented in the document, positioning oneself with respect to the arguments being proposed by the authors to defend themselves, and positioning oneself with the citations being made and their underlying motivation [Wei71]. We have implemented an initial rhetorical parser which allows us to identify sentences describing the authors’ own work, sentences related to work attributed to external authors, and sentences describing background information.

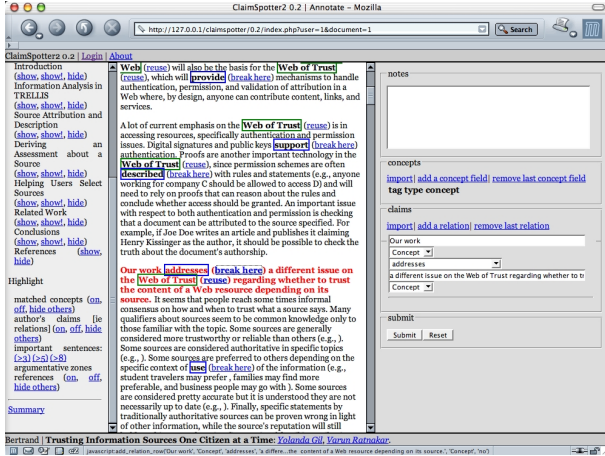


Figure 3. Our annotation interface, ClaimSpotter, allows users to interact with the interface through the selectors, which transform the representation of the text and provide recommendations.

Implementing state of the art components was not the primary goal of our work. We were interested rather in a modular architecture and in the possibility to plug in more robust elements. Indeed, some work has started to replace our rhetorical parser with a more comprehensive version [TM02].

These resources can be wrapped in an environment to support the claim formulation process, as in our ClaiMapper knowledge map interface [USBSL03]. We have also developed an additional interface, ClaimSpotter, displaying side by side the document under consideration and a form to input claims. Using this interface, an annotator can manipulate the text the way she wants, by highlighting the results of the recommendations described above: the different rhetorical zones can be colored and tentative concepts underlined for instance. These elements can directly be dragged and dropped in a claim. Figure 3 shows the current version of this document-centric interface.

3. Evaluation

Recommending components is fine in theory. Assessing their usefulness, however, is a difficult prospect. Traditional annotation reliability measures are of limited use to us [Car96]. In effect, we are not looking to “formulate in a different way” or “formalize” information which does already exist in a document, like in the CREAM approach for instance [HS02]. We are instead interested in knowledge (i) which might not exist *per se* in a document, (ii) which uses this document as a basis point, (iii) and which will also derive from a number of factors (like personal research in-

terests) over which obviously we have no control.

We started our evaluation with the idea of assessing the actual ‘impact’ the recommendations were making on the process. The questions we had in mind were similar to “do users make more claims with the recommendations than without them ?” or “is the nature of the claims submitted richer ?”. This point being based on the idea that a claim making use of a richer discourse relations, say ‘is consistent with’, was of more value than one using ‘is about’ for instance, as it bore more commitment for its author. Our initial experiments were based on the following settings: 4 subjects (2 being familiar with the task and the project, 2 being newcomers) were given a paper and asked to create claims for an amount of time (which was freely set, to match ‘real-life’ conditions). After this time, they were provided the recommendations identified above, and asked whether they would like to add some claims to complement their initial interpretation. By doing so, we were hoping to be able to assess the added value provided by the suggestions.

To summarize the results, we could say that there have been as many experiences as annotators. Some made more claims. But some did not make any use of the suggestions. As far as the ‘quality’ of the claims was concerned, it proved to be very useful for one of our subjects, as it allowed her to submit many more ‘addresses’ claims (*id est*, about the problems being tackled in the test document). A way to explain this could be that presenting authors’ argument in a more direct form (by lifting up these zones [TM02]) gives annotators some additional aspects to react against, which might have been overlooked by the annotator in the first place.

However, it is not because more claims are made, or because the nature of these claims is different than an interpretation is better than another, as there is no ‘right’ or ‘wrong’ interpretation. Therefore, the points we have expressed before have to be moderated and new dimensions of evaluation must be sought. For instance, we could find out how the recommendations and their integration in an interface help users getting their job done. Aspects like the ones we eschewed before about the number and the intrinsic quality of the claims would move to the background, leaving the place to higher-level ones, such as the usability of the tool, the ease of access of the different recommendations and the overall user-friendliness of this multi-windowed environment. Table 1 lists a number of these aspects that we will study in the forthcoming months. The recording of users’ interactions with the system, coupled with a questionnaire based on the evaluation dimensions we are proposing is our next task.

IMPACT OF THE RECOMMENDATIONS	
a	Do they help understand what the document is about ?
b	Do they help hold an internal model of the document and of its connections to the literature ?
c	Do they help the breaking of one's interpretation into 'acceptable' chunks of text and relations and model it as a structured network of nodes and relations ? [SM94]
d	Do the recommendations give additional ideas about the document (which she might not have thought about), additional claims to express, or to counter-argue about a particular point (claim) made by someone else ?
PRESENTATION	
e	Is the ability to browse the text, hide some of its components (sections), and to show the recommendations in situation (through highlighting) helping ?
f	Are the recommendations easily available and accessible ? Does their presentation make sense ?

Table 1. Evaluation dimensions for our ClaimSpotter interface

4. Conclusion

We have presented in this document an approach to assist the translation of an interpretation of a scholarly document into the semi-formal representation scheme provided by the Scholarly Ontologies project. This approach is based on the identification of recommendations, which are proposed for consideration in the hope that they will lighten the burden imposed on an annotator, and their integration in a dialogue interface. Although initial experiments have showed these elements to be interesting, at least to some of our participants, we have also motivated a need for a higher-level evaluation process, in particular based on the way the system eases or obstructs the claim formulation process. The discovery of an acceptable evaluation measure will be our the next priority.

References

- [Bis98] Ann Peterson Bishop. Digital Libraries and Knowledge Disaggregation: the Use of Journal Article Components. In *Proceedings of the 3rd ACM International Conference on Digital Libraries*, 1998.
- [BSMD00] S. Buckingham Shum, E. Motta, and J. Domingue. ScholOnto : an Ontology-based Digital Library Server for Research Documents and Discourse. *International Journal on Digital Libraries*, 3:237–248, September 2000.
- [Car96] Jean Carletta. Assessing Agreement on Classification Tasks: the Kappa Statistic. *Computational Linguistics*, 22(2):249–254, 1996.
- [HS02] Siegfried Handschuh and Steffen Staab. Authoring and Annotation of Web Pages in CREAM. In *Proceedings of the 11th International World Wide Web Conference (WWW2002)*, 2002.
- [LMR⁺03] David B. Leake, Ana Maguitman, Thomas Reichherzer, Alberto Cañas, Marco Carvalho, Marco Arguedas, Sofia Brenes, and Tom Eschridge. Aiding Knowledge Capture by Searching for Extensions of Knowledge Models. In *Proceedings of the International Conference On Knowledge Capture (KCAP), FL, USA*, October 2003.
- [SM94] Frank M. Shipman and Raymond McCall. Supporting Knowledge Base Evolution with Incremental Formalization. In *Proceedings of Human Factors in Computing Systems conference*, pages 285–291, April 1994.
- [SSM04] Bertrand Sereno, Simon Buckingham Shum, and Enrico Motta. Semantic annotation support in the absence of consensus. In *Proceedings of the 1st European Semantic Web Symposium, ESWS 2004, Heraklion, Crete*, May 2004.
- [Swa90] John M. Swales. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press, 1990.
- [TM02] Simone Teufel and Marc Moens. Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status. *Computational Linguistics*, 28(4):409–445, December 2002.
- [USBSL03] V. Uren, B. Sereno, S. Buckingham Shum, and G. Li. Interfaces for Capturing Interpretations of Research Literature. In *Proceedings of the Distributed and Collective Knowledge Capture Workshop, held with the Knowledge Capture Conference (KCAP), FL, USA*, October 2003.
- [Wei71] Melvin Weinstock. Citation Indexes. In *Encyclopedia of Library and Information Science*, volume 5, pages 16–40. 1971.