



HAL
open science

Investigating Argument Relatedness by Means of Frames

Sarah Bourse, Marie Garnier, Patrick Saint Dizier

► **To cite this version:**

Sarah Bourse, Marie Garnier, Patrick Saint Dizier. Investigating Argument Relatedness by Means of Frames. 16th ARGDIAP (Argumentation, Dialogue, Perssuasion) conference: “Argumentation and Corpus Linguistics” - ARGDIAP 2018, ArgDiaP association, a Polish nationwide initiative, Sep 2018, Varsovie, Poland. pp.(electronic medium). hal-03033967

HAL Id: hal-03033967

<https://hal.science/hal-03033967>

Submitted on 7 Dec 2020

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.

Investigating Argument Relatedness by Means of Frames

Sarah Bourse, UT2J Toulouse, France

Marie Garnier, UT2J Toulouse, France

Patrick Saint-Dizier, CNRS IRIT Toulouse

Aims and Motivations

One of the main challenges of argument mining is to identify relevant statements for or against a given claim in a sample of text (for example news articles, blogs, consumer reviews). The problem of relevance is called argument relatedness. It is a central point in information retrieval and is also essential in argument mining (Mochales et al. 2009), (Peldszus et al. 2016), (Swanson et al. 2015). The aim is to mine statements which develop the same topic as the given claim and have an argumentative orientation. Broadly speaking, relatedness is a measure of the semantic and topic proximity of two text spans. They may differ lexically (via the use of synonyms or more generic terms) or syntactically (using e.g. alternations).

In (Saint-Dizier 2016), we show that establishing relatedness between an argument and a statement requires knowledge in 80% of the situations. Since supports and attacks of a claim mainly address the purposes, goals, functions or structure of the main concepts of the claim, we show that an adequate knowledge representation system is the Qualia structure of the Generative Lexicon (Pustejovsky 1995). This approach pairs domain knowledge with lexical descriptions in an efficient and principled way. However, this contribution also shows that Qualia structures are

somewhat difficult to develop and must be defined for each topic. This makes knowledge-based argument mining an approach that, although effective, is difficult to re-use over different domains.

This contribution examines and evaluates the possibility of establishing relatedness solely on the basis of linguistic factors. The development of general purpose linguistic processes and resources that characterize relatedness would make the implementation of its identification much simpler and much more re-usable over domains. This contribution explores this hypothesis and the linguistic resources which are required.

Analysis Protocol

Our analysis is based on two considerations: (1) mining for arguments is driven by the topical content of the claim and (2) the analysis is not based on standard text annotations but on the use of frames encoded in XML. The use of an XML-Frame approach is motivated by the fact that the elements found in statements and that are decisive for the analysis of relatedness may not be adjacent: this makes text annotation, which is linear, almost intractable.

XML-Frames are filled in manually by annotators. Each statement found to be related to the claim and with an argumentative orientation originates an instance of the frame. After the analysis of a number of texts, the result is a set of frames which can be organized as a tree, where the root is the frame representing the claim and the children are those statements found in texts and that introduce additional constraints on the topic. Their relations with the claim are described in each frame instance. In this contribution, we explore the linguistic nature of these relations.

Our corpus is based on texts about social issues, addressing topics such as affirmative action or the gender pay gap. To illustrate it, let us consider the following claim: *affirmative action in education is good for the economy*. This claim is composed of a topic: *affirmative action in education* and an evaluative expression: *is good for the economy*. The goal is then to mine statements which are related to

this claim in various texts. These statements must have a topic that is subsumed by the claim topic and an argumentative orientation which may support or attack the claim depending on the content of the statement. Those statements are also frequently associated with discourse structures which further develop them.

The frame template we have defined for the study of relatedness is as follows:

```

<statement> <topic> <main markers= , link-to-claim= , concepts= ,
  restrictions = , annotator-confidence= >,
  <field markers= , link-to-claim= , concepts= ,
  restrictions = , annotator-confidence= > <\topic>
<evaluative> <evmain markers= , polarity= , strength= ,
  restrictions = , annotator-confidence= >,
  <field markers= , link-to-claim= , concepts= ,
  restrictions = , annotator-confidence= > <\evaluative>
<discourse text= , type= >, % several occurrences possible
<argument-scheme type = , annotator-confidence= >
<\statement>

```

To say it briefly, this frame allows the description of most features that characterize relatedness. The topic field is composed of two parts: the main part, e.g. for the claim (*affirmative action*) and its area(s) of application that restrict it (*in education*). Statements develop subtypes of these elements. The ‘link to claim’ and ‘concepts’ attributes respectively specify the linguistic link (exact words, derivation, synonymy, etc.) and the conceptual link (function, purpose) between the statement topic and the claim topic. The same description is realized for the evaluative part with, in addition, the orientation and strength of the evaluation. The discourse tag describes adjuncts such as elaborations, illustrations, comparisons, conditions or circumstances. Finally, the annotator is invited to specify the kind of argument scheme(s) that has been used, from a standard list of arguments (Walton et al. 2008) (Feng et al 2011).

Toward a Linguistic Categorization of Relatedness

One of the goals of this investigation is to elaborate the feature ‘link to claim’ which develops, from the ‘main markers’ attribute (a list of words found in the statement that establish the relatedness) the nature of that link. To describe the linguistic and conceptual links with the claim, the annotators can use predefined categories or natural language. Then a categorization of the main linguistic operations can be carried out, and the associated resources can be developed.

The aim of this categorization is to characterize the linguistic operations behind relatedness and to evaluate its efficiency and scope, i.e. how much of relatedness analysis can be resolved via linguistic processes. The parameters which are under investigation, categorization and evaluation are as follows:

- the paradigmatic lexico-semantic transformations developed from the topic claim and its restrictions, in particular: forms of synonymy, reformulations, paraphrases, restrictions, negation, forms of inchoativity, etc., for example: *gender parity* → *gender gap*.
- the functional transformations which are related to the nature of the topic, and may induce some domain dependent lexical data,
- the local syntactic transformations on the claim topic, in particular for complex NPs,
- forms of discourse transformations such as: summarization (when the topic is long), illustration or instantiation, expression of consequence,
- the lexical data which is necessary, its structure according to lexical semantics principles (Cruse 1986), and its availability.

We argue that an accurate categorization of these elements and an evaluation of their effectiveness should contribute to overcoming the challenge of relatedness and more generally to argument mining for those systems which are based on linguistic factors.

References

- [1] A. Cruse, *Lexical Semantics*, Cambridge university Press, 1986.
- [2] V. W., Feng and G, Hirst. 2011. Classifying arguments by scheme. In *Proceedings of the 49th ACL: Human Language Technologies*, Portland, USA.
- [3] A., Fiedler and H., Horacek. 2007. Argumentation within deductive reasoning. *International Journal of Intelligent Systems*, 22(1):49-70.
- [4] R., Mochales Palau and M.F., Moens. 2009. Argumentation mining: the detection, classification and structure of arguments in text. *Twelfth international ICAIL'09*, Barcelona.
- [5] H., Nguyen and D. Litman. 2015. Extracting Argument and Domain Words for Identifying Argument Components in Texts. In: *Proc of the 2nd Workshop on Argumentation Mining*, Denver.
- [6] A., Peldszus and M., Stede. 2016. From argument diagrams to argumentation mining in texts: a survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*.
- [7] J., Pustejovsky. 1995. *The Generative Lexicon*, MIT Press.
- [8] P. Saint-Dizier. 2016. *Argument Mining: the bottleneck of knowledge and lexical resources*, LREC, Portoroz.
- [9] R., Swanson, B., Ecker and M., Walker. 2015. *Argument Mining: Extracting Arguments from Online Dialogue*, in proc. SIGDIAL.
- [10] M.G., Villalba and P., Saint-Dizier. 2012. *Some Facets of Argument Mining for Opinion Analysis*, COMMA, Vienna, IOS Publishing.
- [11] M., Walker, P., Anand, J.E., Fox Tree, R., Abbott and J., King. 2012. *A Corpus for Research on Deliberation and Debate*. Proc. of LREC, Istanbul.

- [12] Walton, D., Reed, C., Macagno, F., *Argumentation Schemes*, Cambridge University Press, 2008.