More than Word Cooccurrence: Exploring Support and Opposition in International Climate Negotiations with Semantic Parsing
Pablo Ruiz, Clément Plancq, Thierry Poibeau

To cite this version:

HAL Id: hal-01326692
https://hal.archives-ouvertes.fr/hal-01326692
Submitted on 5 Jun 2016

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.
More than Word Cooccurrence: Exploring Support and Opposition in International Climate Negotiations with Semantic Parsing

Pablo Ruiz Fabo, Clément Plançon, Thierry Poibeau
LATTICE Lab (CNRS, Ecole normale supérieure, U Paris 3,
PSL Research University, Université Sorbonne Paris Cité)
1, rue Maurice Arnoux, 92120 Montrouge, France
E-mail: {pablo.ruiz.fabo,clement.plancon,thierry.poibeau}@ens.fr

Abstract
Text analysis methods widely used in digital humanities often involve word co-occurrence, e.g. concept co-occurrence networks. These methods provide a useful corpus overview, but cannot determine the predicates that relate co-occurring concepts. Our goal was identifying propositions expressing the points supported or opposed by participants in international climate negotiations. Word co-occurrence methods were not sufficient, and an analysis based on open relation extraction had limited coverage for nominal predicates. We present a pipeline which identifies the points that different actors support and oppose, via a domain model with support/opposition predicates, and analysis rules that exploit the output of semantic role labelling, syntactic dependencies and anaphora resolution. Entity linking and keyphrase extraction are also performed on the propositions related to each actor. A user interface allows examining the main concepts in points supported or opposed by each participant, which participants agree or disagree with each other, and about which issues. The system is an example of tools that digital humanities scholars are asking for, to render rich textual information (beyond word co-occurrence) more amenable to quantitative treatment. An evaluation of the tool was satisfactory.

Keywords: climate diplomacy, semantic role labelling, digital humanities

1. Introduction
Text-analysis methods widely used in social or political sciences often involve word co-occurrence. For instance, the concept co-occurrence networks surveyed in Venturini et al. (2012), or the clustering and topic modelling approaches surveyed in Grimmer and Stewart (2013). These methods are useful in order to arrive at an overview of the content of large corpora. However, these techniques do not identify which predicates relate co-occurring elements with each other. If an actor like France is mentioned in the same sentence as a concept, like stricter regulations, which is the verb mediating between both? Is France in favour of, or against stricter regulations?
Several technologies can detect related elements in texts, and the predicate that indicates their relation. A recent approach is Open Relation Extraction (e.g. Mausam et al., 2012), where relations are identified without the need to previously specify a vocabulary of predicates or actors. The corpus we’re working on consists of summaries of international climate negotiations (ENB corpus)¹. A single sentence in this corpus can contain several support and opposition predicates, which can be verbal or nominal (see Figure 1). For this corpus, the results of a workflow based on open relation extraction tools were uneven, particularly with nominal predicates. ² To address these challenges, we developed an application with a domain model and analysis rules which operate on outputs for semantic role labelling, syntactic dependencies and anaphora resolution, provided by a natural language processing (NLP) pipeline.

Our application identifies the points supported and opposed by actors in the negotiations, and aggregates keyphrases and DBpedia concepts³ extracted from those negotiation points. The information is presented on an interface, providing an overview of how different actors’ positions in the negotiation compare to each other. The application helps address a current need identified by digital humanists: tools for the quantitative analysis of textual structures beyond word co-occurrences.

The paper is structured thus: Section 2 discusses related work. Section 3 describes the corpus. Section 4 presents the system. Finally, section 5 provides an evaluation and discussion. Material complementing the paper, and the system itself, are accessible from the project’s website.⁴

2. Related Work
Regarding prior work on the ENB corpus¹, Venturini et al. (2014) analyzed the corpus via co-occurrence networks, using the Cortex² corpus cartography toolkit (Chavalarias and Cointet, 2013). However, the study does not consider which actors were linked to which concepts. Salway et al. (2014) use grammar induction to determine common actor/statement patterns in the corpus, and it could be tested whether these patterns complement our workflow’s outputs.


¹ The Earth Negotiations Bulletin corpus (ENB), http://www.isd.ca/vol12/
² Based on outputs from Open Information Extraction 4.0 (https://github.com/knowitall/openie) and OLLIE (https://github.com/knowitall/ollie, Mausam et al. 2012).
³ wiki.dbpedia.org , Auer et al. (2007). The concepts can sometimes refer to proper nouns or named entities.
⁴ https://sites.google.com/site/climatetlp/
⁵ http://docs.cortext.net
1 - Multiple verbal predicates

The EU, with NEW ZEALAND and opposed by CHINA, MALAYSIA and BHUTAN, supported including the promotion of natural regeneration within the definitions of "afforestation" and "reforestation."

<table>
<thead>
<tr>
<th>Proposnions</th>
<th>Predicate</th>
<th>Negotiation Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 European_Union supported</td>
<td>including the promotion of natural regeneration within the definitions of &quot;afforestation&quot; and &quot;reforestation.&quot;</td>
<td>support</td>
</tr>
<tr>
<td>2 New_Zealand</td>
<td>supported</td>
<td>support</td>
</tr>
<tr>
<td>3 China</td>
<td>~supported</td>
<td>opposition</td>
</tr>
<tr>
<td>4 Malaysia</td>
<td>supported</td>
<td>opposition</td>
</tr>
<tr>
<td>5 Bhutan</td>
<td>supported</td>
<td>opposition</td>
</tr>
</tbody>
</table>

2 - Nominal predicate

Much of the discussion was on a proposal by the G-77/China to include research and development in the transport and energy sectors in the priority areas to be financed by the SCCF.

<table>
<thead>
<tr>
<th>Proposnions</th>
<th>Predicate</th>
<th>Negotiation Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Group_of_77/China proposal</td>
<td>to include research and development in the transport and energy sectors in the priority areas to be financed by the SCCF.</td>
<td>support</td>
</tr>
</tbody>
</table>

Figure 1: Typical corpus sentences. Sentence 1 has predicates supported and opposed, with several actors each. Example 2 shows a nominal predicate (proposal). For Sentence 1, five <actor, predicate, negotiation point> propositions are extracted by the system, and the opposing actors (China, Malaysia, Bhutan) are assigned a proposition which is a negated version (with ~supported as the predicate) of the proposition for the main verb supported.

These studies focus mostly on syntactic dependencies and verbal predicates. We are exploring semantic role labeling as the main source of relation information, and we are addressing nominal predicates besides verbal ones. We also provide an interface to navigate the extraction results.

3. Corpus description

ENB volumes are divided into issues, each of which is a 2000 word summary of the negotiations for one day in a climate summit. The corpus strives for an objective tone. To avoid biases, it uses similar syntactic structures when reporting about all participants’ interventions. Typical sentences, showing how the corpus reports on participants’ support and opposition in the negotiations, are in Figure 1. We analyzed the 255 issues (ca. 35,000 sentences) for the Conference of the Parties (COP) summits. We parsed the HTML into clean text, and added a date for each issue based on ENB’s table of contents.

4. System architecture

The system’s goal is helping researchers analyze patterns of support and opposition between different actors in climate negotiations, as well as examining which issues these actors agree or disagree about. To this end, based on the outputs of an NLP toolkit, and based on a domain model, the system applies rules to extract domain-relevant propositions, formalized as <actor, predicate, negotiation point> tuples (see Figure 1). The information is made navigable, together with the original corpus, on a user interface. In addition, the system extracts keywords and linked entities from the negotiation points, and also displays them on the interface, in response to user queries.

4.1. NLP pipeline

We used the IXA Pipes NLP toolkit\(^7\) (Agerri et al., 2014), and compatible tools. The toolkit's default English modules were used for tokenization, part of speech tagging and constituency parsing.

Anaphora resolution: Some types of pronominal anaphora (see 4.2) were resolved via custom rules based on coreference chains from CorefGraph\(^8\), a Python implementation of Stanford’s dcoref\(^9\) (Lee et al. 2013).

Dependency parsing and semantic role labeling (SRL) were carried out with ixa-pipe-srl\(^9\), which provides a wrapper around the mate-tools library (Björkelund, Bohnet et al., 2010). The dependency and SRL format are the CoNLL ones (Surdeanu et al., 2008). SRL is performed against PropBank (Palmer et al. 2005) and NomBank (Meyers et al. 2004).

Keyphrase Extraction: We used YaTeA (Aubin and Hamon, 2006), which extracts multiple-word and single-word terms in an unsupervised manner, using syntactic and statistical criteria.

Entity Linking (EL) was performed with the ELCO3 tool from our previous work (Ruiz and Poibeau, 2015). This combines EL outputs from several public-domain EL systems, and selects the best outputs via a weighted vote.

Annotation format: IXA Pipes uses NAF, the NLP Annotation Format (Fokkens et al., 2014). This is an

---

\(^6\) The terminology adopted is as follows: <France, proposed, a new approach to carbon sinks> is a proposition, with actor France, predicate proposed and a new approach to carbon sinks as the negotiation point.

\(^7\) http://ixa2.si.ehu.es/ixa-pipes/

\(^8\) https://bitbucket.org/Josu/corefgraph

\(^9\) https://github.com/newsreader/ixa-pipe-srl
XML format composed of layers, each of which represents an analysis step (tokenization, part-of-speech tagging, etc.). The KafNafParserPy library was used to manage NAF annotations.

4.2. Domain model and analysis rules

The domain model contains actors and predicates. Actors represent participants in international climate negotiations, and are formalized as a map between actor variants and their DBpedia URI. The model also contains lemmas for verbal and nominal predicates. Some verbs are neutral reporting verbs (e.g. announce). Other verbs express notions like support or opposition and agreement or disagreement (e.g. criticize). The verbs are contained in PropBank. The nominal predicates (e.g. announcement, objection) express similar notions to the verbs, and were selected from NomBank. The predicate type (i.e. support, oppose or report) is also specified in the model.

Several analysis rules were implemented, that identify propositions based on a predicate’s semantic roles. Most of our domain predicates involve an agent and a message (i.e. a negotiation point) expressed by that agent in a given manner: agreeing with it, objecting to it, or simply mentioning it. In that sense, actor mentions in a predicate’s A0 argument correspond to the actor expressing a message, and the predicate’s A1 argument often corresponds to the negotiation point addressed by the actor. Based on this, the generic rule is in Figure 3.

Rule: Generic proposition

for each predicate p:
· resolve negation (see below)
for each pronoun he, she, it in p’s A0 argument:
· apply anaphora resolution (see below)
for each actor-mention am in p’s A0 argument:
· create a proposition ‹am, p, point›, where point is a concatenation of p’s A1 arguments

Figure 3: Generic proposition rule

In some sentences, the AI role contains actors rather than a negotiation point. This is notably the case in constructions like China, opposed by the EU, preferred... (see Figure 1 for an example). The agent of opposed by is the agent of a proposition that contradicts the main verb’s proposition. The rule to treat such constructions is in Figure 4.

Rule: Proposition for an opposing actor

for each opposed by sequence ob:
· find proposition main for the sentence’s main verb for each actor-mention oam in ob:
· create a proposition ‹oam, ~pmain, pointoam›, where ~pmain is a negated form of main’s predicate, and pointoam is main’s negotiation point

Figure 4: Proposition rule for disagreeing actors
Negation is treated by finding \textit{AM-NEG} roles related to the predicates, as well as negative lexical items (e.g. \textit{not}, \textit{lack}) in a window of two tokens preceding the predicate.

\textbf{Anaphora resolution:} In the corpus, a personal pronoun (\textit{he, she}) can be the anaphor for a country\footnote{The pronoun gender corresponds to the delegate representing the country.}, besides the inanimate pronoun \textit{it}. Two rules were created to deal with this non-standard pronoun use, and anaphora resolution was limited to cases covered by these rules:

- An actor in the subject of a sentence’s main verb (based on dependency parsing) is taken as the antecedent of a sentence-initial \textit{he/she} in the following sentence.
- Antecedents for a pronoun (from CorefGraph’s coreference chains) are only accepted if they are in the same sentence as the pronoun, or in the sentence immediately preceding the pronoun.

Finally, to facilitate searches by date-range, propositions are assigned the date of the document containing them.

\subsection*{4.3. User interface}

The interface (Figure 5) allows researchers to explore countries’ positions in the negotiation, and to compare them based on keyphrases and DBpedia concepts.

A full text search is performed with the \textit{Text} search box. Documents matching the query are displayed on the right panel, and the propositions that have been annotated by the system in those documents are displayed on the left panel. Propositions with a given agent or a given predicate are searched with the \textit{Actors} and \textit{Actions} search boxes respectively. The matching propositions are displayed on the left panel, and their documents on the right.

Tabs for KeyPhrases and DBpedia concepts on the right panel provide an overview of the content retrieved for a query. For a Text search, the keyphrases and concepts have been extracted from full documents. For Actors or Actions searches, the keyphrases and concepts are restricted to propositions matching the query.

The \textit{AgreeDisagree} view (Figure 6) allows selecting two actors or groups of actors, and displays keyphrases and DBpedia concepts from propositions where those actors agree or disagree.

Exporting results or editing the model’s actors and predicates is currently not allowed; this would be useful future work.

Figure 5: \textit{Main view of the interface}. The left panel gives access to the search workflows (\textit{Text, Actors, Actions}). It also shows propositions for a query (e.g. the actor \textit{Canada}), and gives access to the \textit{AgreeDisagree} view. The right panel shows the documents in the \textit{Docs} tab, as well as aggregated keyphrases and DBpedia concepts/entities for the query or for selected propositions, in the other tabs.

Figure 6: The \textit{AgreeDisagree View} displays keyphrases and DBpedia concepts found in sentences where selected actors (here The EU and China) agree or disagree.
5. Evaluation and discussion

Several aspects can be evaluated. First, results from the NLP modules. Second, the outputs of the domain-specific components. Third, the user interface. Finally, for digital humanities applications, it is important whether the system helps researchers gain insights they would not have otherwise achieved (e.g. detect previously unseen generalizations; see Berry, 2012).

The domain model and rules to create domain-relevant propositions were evaluated with two manually annotated test-sets. The first one (ENB-COP) comprises 100 sentences (311 propositions) from the COP climate summit issues in the ENB corpus, that the system was built to analyze. The test-set primarily contains sentences representing the corpus challenges (with negation, multiple actors, multiple predicates, verbal and nominal predicates). The second test-set (ENB-IPCC) also comes from the ENB corpus, but not from the COP issues. Instead, it is based on ENB issues covering scientific report creation discussions by the Intergovernmental Panel on Climate Change. In other words, the ENB-IPCC test-set does not come from the same corpus that the system was developed for, and serves as a way to test the system on texts from a slightly different domain, in terms of syntactic structures and lexical items involved. The second test-set contains 283 sentences (566 propositions). For both test-corpora, a system output was considered correct if all of the proposition components (actor, predicate, negotiation point) match the reference exactly. Based on this notion of a correct output, precision, recall and F1 values are shown on Table 1.

Table 1: Exact-match proposition extraction. Precision, Recall, F1\(^{16}\) on the ENB-COP corpus (on which system development was based), and the ENB-IPCC corpus (covering a somewhat different domain)

<table>
<thead>
<tr>
<th>Corpus</th>
<th>F1</th>
<th>P</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENB-COP</td>
<td>0.69</td>
<td>0.687</td>
<td>0.693</td>
</tr>
<tr>
<td>ENB-IPCC</td>
<td>0.718</td>
<td>0.714</td>
<td>0.722</td>
</tr>
</tbody>
</table>

We consider these results acceptable for corpus exploration. Note that our evaluation is conservative, since propositions partially matching the reference receive no credit. It could have been possible to achieve higher scores by computing F1 over individual proposition elements, or by using the slot error rate metric (Makhoul et al., 1999). Our conservative measure avoids overestimating the system’s value for our users.

The proposition elements for which the system made an error in the ENB-COP corpus are summarized in Table 2. Table 3 contains an error-type analysis in a sample containing the first 45 errors from the ENB-IPCC corpus.

Table 2: Counts and proportion of errors per error-type, for propositions of shape \(\text{actor, predicate, point}\) in the ENB-COP corpus

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Count</th>
<th>% of Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>only predicate wrong</td>
<td>2</td>
<td>2.1%</td>
</tr>
<tr>
<td>only point wrong</td>
<td>63</td>
<td>64.95%</td>
</tr>
<tr>
<td>both predicate &amp; point wrong</td>
<td>32</td>
<td>32.99%</td>
</tr>
</tbody>
</table>

Table 3: Counts and proportion of errors types for propositions of shape \(\text{actor, predicate, point}\) in a sample of errors from the ENB-IPCC corpus

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Count</th>
<th>% of Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>only predicate wrong</td>
<td>5</td>
<td>11.11%</td>
</tr>
<tr>
<td>only point wrong</td>
<td>35</td>
<td>77.78%</td>
</tr>
<tr>
<td>both predicate &amp; point wrong</td>
<td>5</td>
<td>11.11%</td>
</tr>
</tbody>
</table>

Most errors took place identifying the proposition’s negotiation point. It can be challenging to delimit an actor’s negotiation point based on semantic roles, besides the difficulty posed by our evaluation, requiring exact matches. Exploiting dependency information to add syntactically related words to the negotiation point can help (see van Atteveldt, 2015). In both test-corpora, between approx. 25% and 35% of the errors involve a wrongly identified predicate. These errors occur with some types of multi-predicate sentences. Regarding the custom rules for pronominal anaphora, a thorough evaluation against an annotated test-set has not been performed. What can be stated based on informal evaluation is that accuracy was fine for the application’s needs, but given that the rules only consider sentence initial he/she pronouns, coverage may be lacking. In terms of user evaluation, a domain expert as well as two general users have provided comments. They find the application original since the data it outputs (propositions and their associated keyphrases and entities) is not available from other applications they have access to. Our users have pointed out some possible improvements. For instance, there could be more interactivity across the application’s panels, e.g. clicking on a keyphrase or DBpedia entity could highlight the propositions it has been extracted from, or restrict the result set in the proposition pane accordingly. (This function is already available in the AgreeDisagree pane, but not elsewhere in the app). Users have also pointed out that some of the keyphrases extracted are not informative; a better weighting or filtering could be implemented.

\(^{15}\) The results are reproduced on the project’s site: https://sites.google.com/site/climatenlp

\(^{16}\) Both test-sets and related information are on the project’s site

\(^{17}\) The definitions for these metrics were the usual:

\[
F1 = \frac{2 \cdot P \cdot R}{P + R}, \quad P = \frac{\text{nbr. of correct outputs}}{\text{nbr. of system outputs}}, \quad R = \frac{\text{nbr. of correct outputs}}{\text{nbr. of reference outputs}}
\]
6. Outlook
Besides improvement suggestions provided by users, mentioned above, useful future work would be exploiting the propositions extracted by the application to create network graphs representing support and opposition between negotiating parties, and between parties and issues: Opposition vs. support predicates would be represented visually by different types of edges in the network. Other useful features would be annotation export, and an annotation confidence score, which users could employ to prioritize manual result revision.

Acknowledgements
We thank Tommaso Venturini, Audrey Baneyx, Kari de Pryck and Diégo Antolinos-Basso from the Sciences Po médialab in Paris for feedback on the system. Pablo Ruiz is supported through a PhD grant from Région Ile-de-France.

7. References


Chavalarias, D. and Cointet, J.P. (2013). Phylomematic patterns in science evolution—the rise and fall of scientific fields. Plos one, 8(2)


