Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure

Patrick Saint-Dizier
IRIT-CNRS, 118 route de Narbonne
31062 Toulouse France
stdizier@irit.fr

Abstract

Given a controversial issue, argument mining from texts in natural language is extremely challenging: besides linguistic aspects, domain knowledge is often required together with appropriate forms of inferences to identify arguments. A major challenge is then to organize the arguments which have been mined to generate a synthesis that is relevant and usable. We show that the Generative Lexicon (GL) Qualia structure, enhanced in different manners and associated with inferences and language patterns, allows to capture the typical concepts found in arguments and to organize a relevant synthesis.

1 Introduction to Argument Mining

One of the main goals of argument mining is, given a controversial issue, to identify in a set of texts the arguments for or against that issue. These arguments act as supports or attacks of the issue. Arguments may also attack or support the arguments which support or attack that controversial issue in order to reinforce or cancel out their impact. Arguments are difficult to identify, in particular when they are not adjacent to the controversial issue, possibly not in the same text, because their linguistic, conceptual or referential links to that issue are rarely explicit.

For example, given the controversial issue: *Vaccine against Ebola is necessary*, the link with statements such as *Ebola adjuvant is toxic*, *Ebola vaccine production is costly*, or *7 people died during Ebola vaccine tests* is not straightforward without domain knowledge, including finding the polarity of these statements. For example, a knowledge-based analysis of the third statement shows that it is irrelevant or neutral w.r.t. the issue (Saint-Dizier 2016).

Argument mining is an emerging research area which introduces new challenges in natural language processing and generation. Argument mining research applies to written texts, e.g. (Mochales Palau et ali.., 2009), (Kirschner et ali., 2015), for example for opinion analysis, e.g. (Villalba et al., 2012), mediation analysis (Janier et al. 2015) or transcribed argumentative dialog analysis, e.g. (Budzynska et ali., 2014), (Swanson et ali., 2015). The analysis of the NLP techniques relevant for argument mining from annotated structures is analyzed in e.g. (Peldszus et al. 2016). Annotated corpora are now available, e.g. the AIFDB dialog corpora or (Walker et al., 2012). These corpora are very useful to understand how argumentation is realized in texts, e.g. to identify argumentative discourse units (ADUs), linguistic cues (Nguyen et al., 2015), and argumentation strategies, in a concrete way, possibly in association with abstract argumentation schemes, as shown in e.g. (Feng et al., 2011). In natural language generation, argument generation started as early as (Zuckerman et ali. 2000). Finally, reasoning aspects related to argumentation analysis are developed in e.g. (Fiedler et al., 2007) and (Winterstein, 2012) from a formal semantics perspective. Abstracting over arguments allows to construct summaries and to induce customer preferences or value systems.

In (Saint-Dizier 2016), a corpus analysis identifies the type of knowledge that is required to develop argument mining. It is briefly reported in this paper. Then, the Generative Lexicon (GL) is shown to

be an appropriate model, sufficiently expressive, to characterize the types of knowledge, inferences and lexical data that are required to accurately identify arguments related to an issue. The present contribution focuses on the next stage: given a set of arguments for or against an issue that have been mined in various texts, how to generate a synthesis that is consistent, well-structured and usable?

2 Mining Arguments: the need of knowledge

To explore and characterize the forms of knowledge that are required to develop argument mining in texts, we constructed and annotated four corpora based on four independent controversial issues. The texts considered are extracts from various sources, e.g.: newspaper articles and blogs from associations. Issues deal with (1) Ebola vaccination, (2) women's situation in India, (3) nuclear plants and (4) organic agriculture. The total corpus includes 51 texts, a total of 24500 words for 122 different arguments. From our manual analysis, the following argument polarities are observed: attacks: 51 occurrences, supports: 32, argumentative concessions: 17, argumentative contrasts: 18 and undetermined: 4.

Our analysis shows that for 95 arguments (78%), some form of knowledge is involved to establish an argumentative relation with an issue. An important result is that the number of concepts involved is not very large: 121 concepts for 95 arguments over 4 domains. These concepts are mainly related to purposes, functions, parts, properties, creation and development of the concepts in the issues. These are relatively well defined and implemented in the Qualia structure of the Generative Lexicon, which is the framework adopted in our modeling.

The Generative Lexicon (GL) (Pustejovsky, 1995) is an attempt to structure lexical semantics knowledge in conjunction with domain knowledge. In the GL, the Qualia structure of an entity is a kind of lexical and knowledge repository composed of four fields called roles:
- **the constitutive role** describes the various parts of the entity and its physical properties, it may include subfields such as material, parts, shape, etc.
- **the formal role** describes what distinguishes the entity from other objects,

- **the telic role** describes the entity functions, uses, roles and purposes,
- **the agentive role** describes the origin of the entity, how it was created or produced.

Let us consider the controversial issue (1): *The vaccine against Ebola is necessary*. The Qualia structure of the head term of (1), `vaccine(X)`, is:

$$
\begin{bmatrix}
\text{CONSTITUTIVE:} \begin{bmatrix} \text{ACTIVE\_PRINCIPLE, ADJUVANT} \end{bmatrix}, \\[4pt]
\text{TELIC:} \begin{bmatrix} \text{MAIN: PROTECT\_FROM(X,Y,D),} \\ \text{AVOID(X,DISSEMINATION(D)),} \\ \text{MEANS: INJECT(Z,X,Y)} \end{bmatrix}, \\[4pt]
\text{FORMAL:} \begin{bmatrix} \text{MEDICINE, ARTEFACT} \end{bmatrix}, \\[4pt]
\text{AGENTIVE :} \begin{bmatrix} \text{DEVELOP(T,X), TEST(T,X),} \\ \text{SELL(T,X)} \end{bmatrix}
\end{bmatrix}
$$

The Qualia structure of `Ebola` is:

$$
\begin{bmatrix}
\text{FORMAL:} \begin{bmatrix} \text{VIRUS, DISEASE} \end{bmatrix}, \\[4pt]
\text{TELIC:} \begin{bmatrix} \text{INFECT(E1,EBOLA,P)} \Rightarrow \text{GET\_SICK(E2,P)} \\ \Rightarrow \Diamond\text{DIE(E3,P)} \land \text{E1} \le \text{E2} \le \text{E3} \end{bmatrix}
\end{bmatrix}
$$

Then, via formula expansion, the formal representation of the controversial issue is:

$\Box$ *(protect_from(X,Y, (infect(E1,ebola, Y) $\Rightarrow$ get_sick(E2,Y) $\Rightarrow \Diamond$ die(E3,Y)))
$\land$ avoid(X,dissemination(ebola)).*

3 Generation of an argument synthesis related to a controversial issue

3.1 A Network of Qualias to Characterize the Generative Expansion of Arguments

Our observations show that arguments attack or support (1) specific concepts found in the Qualia of the head terms in the controversial issue (called **root concepts**) or (2) concepts derived from these root concepts, via their Qualia. For example, arguments can attack properties or purposes of the adjuvant or of the protocols used to test the vaccine. Then, a network of Qualias must be defined to develop the argument synthesis. This network is limited to three levels because derived concepts must remain functionally close to the root concepts.

A Qualia Q_i describes major features of a concept such as vaccine(X), it can be formally defined as follows:
$Q_i : [\ R_X : T_j^{i,X}\ ]$, where:

- R_X denotes the four roles: $X \in \{formal, constitutive, agentive, telic\}$ and possibly subroles,
- $T_j^{i,X}$ is a term which is a formula, a predicate or a constant T_j in the role X of Q_i.

A network of Qualias is then defined as follows:
- nodes are of two types: [terminal concept] (no associated Qualia) or [non terminal concept, associated Qualia],
- the root is the semantic representation of the controversial issue and the related Qualias Q_i,
- Step 1: the first level of the network is composed of the nodes which correspond to the terms $T_j^{i,X}$ in the roles of the Qualias Q_i The result of this step is the set T of terminal nodes $\{ T_j^{i,X} \}$ and non terminal nodes $\{ T_j^{i,X}, Q'_{i'}: [R_X : T'^{i',X}_{j'}]\}$.
- Step 2: similarly, the terms $T'^{i',X}_{j'}$ from the $Q'_{i'}$ of step 1 introduce new nodes into the network together with their own Qualia when they are non-terminal concepts. They form the set T', derived from T.
- Step 3: the same operation is carried out on T' to produce T''.
- Final step: production of T'''. The set of concepts involved is: $\{T \cup T' \cup T'' \cup T'''\}$.

This network of Qualias forms the backbone of the argument mining system. This network develops the **argumentative generative expansion of the controversial issue**. This network is the organization principle, expressed in terms of relatedness, that guides the generation of a synthesis where the different facets of the Qualias it contains are the structuring principles.

3.2 Synthesis Generation Input Data

Arguments which have been mined are automatically tagged with the following attributes:
- the argument identifier (an integer),
- the **identifier of what the argument attacks or supports** (issue or another argument),
- the **text span involved** that delimits the argument compound and its kernel,
- the **polarity of the argument** w.r.t. the issue with one of the following values: support, attack, argumentative concession or contrast.
- the **concepts involved**, to identify the argument: list of the main concepts from the Qualias used in the mining process,

- the **strength of the argument**, based on linguistic marks found in the argument,
- the **discourse structures** in the compound, associated with the argument kernel, as processed by our discourse analysis platform TextCoop.

Argument 11 is tagged as follows:
<argument Id= 11, polarity= attack with concession, relationWith=issue, conceptsInvolved= efficiency measure, safety measure, test, evaluation method, strength= moderate >
<concession> *Even if the vaccine seems 100% efficient and without any side effects on the tested population,* < /concession>
<main arg> *it is necessary to wait for more conclusive data before making large vaccination campaigns.* < /main arg>
<elaboration> *The national authority of Guinea has approved the continuation of the tests on targeted populations.* </elaboration> < /argument>.

At this stage no metadata is considered such as date of argument or author status. This notation was defined independently of any ongoing task such as ConLL15.

Argument kernels are expressed in various ways:
- **evaluative expressions:** *Vaccine development is very costly, adjuvant is toxic*,
- **comparatives:** *number of sick people much smaller than for Malaria.*
- **facts related to properties of the main concept(s) of the issue:** *Vaccine is not yet available. There is no risk of dissemination.*
- **facts related to the consequences, functions, purposes, uses or goals** of the issue: *vaccine prevents bio-terrorism, 3 vaccinated people died.*

An indicative evaluation provides an accuracy of about 82% of the mining and tagging processes compared to our manual annotation.

3.3 A Multi-facet Argument Synthesis Generation

The network of Qualias together with the attributes 'relationWith', 'Polarity' and 'ConceptsInvolved' are the most crucial elements to generate a structured argument synthesis that corresponds to the domain conceptual organization. Each node of the Qualia network defines a cluster to which mined arguments are associated. Clusters are organized hierarchically by decreasing relatedness to the issue.

Redundant arguments characterized by lexical duplicates are eliminated.

The generation of a synthesis proceeds informally as follows:

(1) The concepts of the network of Qualias are organized hierarchically and by role. Level 0 of the synthesis contains the concepts $T_j^{i,X}$ of the set T, since they are the most relevant. They are organized by role: constitutive, agentive and telic (which is the main role for argument mining). Each concept in a role is considered separately. In the telic role, the list starts by the 'Main' category and then the other types are considered, such as 'Means'.

(2) Each of these concepts defines a cluster that contains the mined arguments that have this concept in their 'conceptsInvolved' attribute, with the constraint that the argument concerns the issue (relationWith=issue), the other arguments are treated as presented in (6). For example, the concept `sell(T,X)` includes the mined argument with the value *production costs* its its attribute 'conceptsInvolved'.

(3) This process goes on with the lower levels of the network of Qualias: the concepts $T'^{i',X}_{j'}$ associated with the Qualias $Q_{i'}$ are considered and then, the sets T'' and T''' defined in 3.1.

(4) When an argument involves several concepts, it is included into the concept cluster that is the highest in the network of concepts.

(5) For each concept, related arguments are structured by polarity: first supports and then attacks, for each of polarity, arguments are listed, with an indication of their number of occurrences found while mining to give an estimate of their recurrence.

(6) To deal with the generation of a more comprehensive graph of attacks and supports, a further stage of the generation process consists in considering that each argument that has been mined can also be an issue which can be attacked or supported. This is specified in the attribute 'relationWith'. Therefore, the same process as for the controversial issue is applied, while keeping the same network of Qualias.

Let us consider our example on vaccination, then:
(a) the root concepts (Level 0) in the Qualia hierarchy are those of vaccine and ebola, e.g.: constitutive: *active principle, adjuvant*; agentive: *develop, test, sell*; telic: *protect-from, dissemination, infect,* *get-sick, die, inject*.

(b) Level 1 concepts are those associated with the Qualias of the root concepts given in (a), e.g. those in the Qualias of: adjuvant (e.g. *dilute*), tests (e.g. *efficiency measure, evaluation methods*), develop (e.g. *production costs, availability, ethics*), etc.

For example, a synthesis would be composed of the following clusters:

Level 0: e.g.:

Cluster 1: Adjuvant: attack: *adjuvant is toxic* (3 occurrences) ...

Cluster 2: Dissemination: support: *reduces dissemination* (5)..

Cluster 3: Get-sick: concessive support: *limited number of cases and deaths compared to other diseases* (2), ...

Level 1:, e.g.:

Cluster 4: Production costs: attack: *high production and development costs (6) ...*

Cluster 5: Availability: concessive attack: *vaccine not yet available (4)*, etc.

Finally, in order to make the synthesis based on Qualia network structure more clear, each of the concepts is associated with a simple and direct definition, directly generated via language patterns from the Qualia structure network, e.g.:

- *Side-effect* and *toxicity* are related to the use of a medicine,

- *Contamination* entails *disease dissemination*.

- *Population isolation* avoids *disease dissemination*.

- *Production costs* are related to the creation and development of any product,

- *Efficiency* must be measured during the *test* phase.

These definitions are defined for each cluster. The result is a hierarchically organized and well articulated set of clusters that account for the various arguments for or against an issue found in various texts, where each level is made clear to the reader. No evaluation of the relevance of this type of clustering technique has been carried out so far. This evaluation should involve the analysis of the adequacy of this clustering technique by real users, its granularity and its adequacy to the problem to investigate. It may also depend on the type of issue and arguments that have been mined. Deeper forms of argument synthesis could be desirable, but these involve complex conceptual planning issues. To the best of our knowledge, such a task has never been undertaken.

References

K., Budzynska, M., Janier, C., Reed, P. Saint-Dizier, M., Stede, and O. Yakorska. 2014. A model for processing illocutionary structures and argumentation in debates. In proc. LREC, 2014.

V. W., Feng and G, Hirst. 2011. Classifying arguments by scheme. In Proceedings of the 49th ACL: Human Language Technologies, Portland, USA.

A., Fiedler and H., Horacek. 2007. Argumentation within deductive reasoning. International Journal of Intelligent Systems, 22(1):49-70.

M., Janier, C. and Reed, C. 2015. Towards a Theory of Close Analysis for Dispute Mediation Discourse, Journal of Argumentation.

C., Kirschner, J., Eckle-Kohler and I., Gurevych. 2015. Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications. In: Proceedings of the 2nd Workshop on Argumentation Mining, Denver.

R., Mochales Palau and M.F., Moens. 2009. Argumentation mining: the detection, classification and structure of arguments in text. Twelfth international ICAIL'09, Barcelona.

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.

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).

J., Pustejovsky. 1995. The Generative Lexicon, MIT Press.

P. Saint-Dizier. 2016. Argument Mining: the bottleneck of knowledge and lexical ressources, LREC, Portoroz.

R., Swanson, B., Ecker and M., Walker. 2015. Argument Mining: Extracting Arguments from Online Dialogue, in proc. SIGDIAL.

M.G., Villalba and P., Saint-Dizier. 2012. Some Facets of Argument Mining for Opinion Analysis, COMMA, Vienna, IOS Publishing.

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.

G., Winterstein. 2012. What but-sentences argue for: An argumentative analysis of 'but', in Lingua 122.

I., Zuckerman, R., McConachy and K. Korb. 2000. Using Argumentation Strategies in Automatic Argument Generation, INLG.