A Two-Level Approach to Generate Synthetic Argumentation Reports

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

ABSTRACT

Given a controversial issue, a major challenge in argument mining is to organize the arguments which have been mined to generate a synthesis that is readable, synthetic enough and relevant for various types of users. Based on the Generative Lexicon (GL) Qualia structure, which is a kind of lexical and knowledge repository, that we have enhanced in different manners and associated with inferences and language patterns, we show how to construct a synthesis that outlines the typical elements found in arguments. We propose a two-level approach: a synthesis of the arguments that have been mined and navigation facilities that allow to access the argument contents in order to get more details.

CCS CONCEPTS

• **Computer systems organization** → **Natural language processing**; *Argument mining*; Knowledge representation;

KEYWORDS

Generative Lexicon, Rhetoric

1 AIMS AND CHALLENGES

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*, identifying the argumentative link with statements such as *Ebola adjuvant is toxic*, *Ebola vaccine production is costly*, or *7 people died during Ebola vaccine tests* cannot be realized solely on the basis of linguistics data, but requires domain knowledge. Furthermore, a knowledge-based analysis of the third statement shows that it is irrelevant or neutral w.r.t. the issue (Saint-Dizier 2016).

1.1 Argument Mining Challenges

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). Finally, reasoning aspects related to argumentation analysis are developed in e.g. (Fiedler et al., 2007) and (Winterstein, 2012) from a formal semantics perspective.

In opinion analysis, the benefits of argument mining are not only to identify the customers satisfaction level, but also to characterize why customers are happy or unhappy. Abstracting over arguments allows to construct summaries and to induce customer preferences or value systems (e.g. low fares are preferred to localization or quality of welcome for some categories of hotel customers).

In (Saint-Dizier 2016a), a corpus analysis identifies the type of knowledge and inferences that are required to develop argument mining. It is briefly reported in this paper. Then, we have shown, on the basis of a set of examples, that the Generative Lexicon (GL) could 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.

1.2 Natural Language Summarization

In natural language generation, the main projects on argument generation was developed as early as (Zuckerman et ali. 2000) and (Fiedler 2007). While there are currently several reseach efforts to develop argument mining, very little has been done recently to produce a synthesis of the mined arguments that is readable, synthetic enough and relevant for various types of users. This includes identifying the main features for or against a controversial issue, but also tasks such as eliminating duplicates, fallacies or ad'hominem statements and identifying arguments which attack or support each other, besides the controversial issue.

18th Workshop on Computational Models of Natural Argument
Floris Bex, Floriana Grasso, Nancy Green (eds)
16th July 2017, London, UK

In (Saint-Dizier 2016b), we show how arguments that have been mined can be organized in hierarchically structured clusters so that readers car navigate over and within sets of arguments according to the conceptual organization proposed by the Generative Lexicon. This approach turned out not to be synthetic enough, since over 100 arguments can be mined for a given issue, making the perception of the main attacks and supports quite difficult. However, this initial approach allows the construction of an argument database useful to readers who wish to access to the exact form of arguments that have been mined.

The present contribution focuses on the next stage, aiming at producing a synthesis that is short and efficient where the concepts present in the GL Qualia structures are used to abstract over arguments while keeping the structure of those clusters of arguments which are accessible via links from the synthesis. The argument cluster system is accessed to get more precise information.

This contribution to natural language argumentation synthesis is not really a summarization task, as e.g. developed in (Mani et al. 1999). In our approach, no text or document is reduced to produce a summary. The synthesis that is proposed is simply a two level re-organization task that involves forms of clustering. From that perspective, it could be viewed as a preliminary step to a summarization procedure. A real summarization task would involve constructing summaries for each cluster of arguments, but this is beyond the present research.

In terms of feature classification and relevance, the concepts used in the Qualia structure of the Generative lexicon are defined a priori, similarly to the features evaluated in most opinion analysis systems. They are used as entry points to the re-organization and to the cluster system. A challenging point is that these concepts must obviously correspond as much as possible to the user perception of the domain to which the issue belongs.

1.3 Paper Structure

In this paper, for the sake of understanding, we first summarize the results elaborated in our previous contributions, we then develop the synthesis production model. This two-level approach, a synthesis of the arguments that have been mined and, associated with navigation facilities that allow to access the argument contents in order to get more details seems to be an efficient approach for readers who want first to get the essentials of the argumentation.

2 MINING ARGUMENTS: THE NEED OF KNOWLEDGE

2.1 Corpus Analysis: 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.

2.2 An introduction to the Generative Lexicon

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 both a 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.

To illustrate this conceptual organization, let us consider the controversial issue (1):
The vaccine against Ebola is necessary.
The main concepts in the Qualia structure of the head term of (1), *vaccine* are organized as follows:

Vaccine(X):

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

The Qualia structure of `Ebola` is:

Ebola:

$$\begin{bmatrix} \text{FORMAL: } \begin{bmatrix}\text{VIRUS, DISEASE}\end{bmatrix}, \\ \text{TELIC: } \begin{bmatrix}\text{INFECT(E1,EBOLA,P)} \Rightarrow \text{GET\_SICK(E2,P)} \\ \Rightarrow \Diamond\text{DIE(E3,P)} \wedge \text{E1}\leq\text{E2} \leq\text{E3}\end{bmatrix} \end{bmatrix}$$

The terms, predicates or constants, found in the different roles of any Qualia are defined on the basis of a domain ontology, when it exists, or via bootstrapping techniques on the web, if it doesn't exist for this domain. Qualia structures can be hierarchically organized, as in any ontology. Vaccine is a kind of medicine, it therefore inherits of the properties, i.e. the predicates present in medicine, unless some blocking is formulated. Similarly, Ebola is a kind of disease, therefore it inherits of the properties of a disease. This rich organization greatly simplifies the description of Qualias. Some Qualia structure resources are available as payware at ELRA, from the SIMPLE EEC project.

Finally, from the two above Qualias and 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)))* $\wedge$ *avoid(X,dissemination(ebola)).*

2.3 Using Qualias for Argument Mining

Originally, the Qualia structure was designed to characterize sense variations around a prototypical one, and the large number of potential combinations of NP arguments with predicates, in particular verbs. This was implemented via a mechanism called type coercion. In (Pustejovsky 1995), the Qualia structure manipulates atomic terms assocated as lists to one of the four qualia roles. This Qualia structure, in our view, is a specific interpretation of a more global typology of object descriptions, realized in various manners from Aristotle.

In our approach, we view the Qualia structure as a means to structure knowledge associated with concepts in a functional way, via telicity (an subtypes of telicity), various types of functional and structural parts, and the way an object was created. This view allows us to have complex structures such as formula, modalities, etc. instead of just the atomic concepts of the original Qualia. Manipulating such structures is more complex, but better corresponds to the reality. Reasoning with these complex forms is addressed in (Saint-Dizier 2016a). In the present paper, we propose an argument synthesis based on atomic concepts, which may be isolated concepts in roles or part of formula.

Other types of resources such as FrameNet, WordNet or VerbNet do not contain the information found in Qualias, which is essential for argument mining. These latter resources are structured around predicative forms and mainly describe the type of arguments and adjuncts predicates can take and how they are combined. VerbNet introduces semantic representations based on primitives which may be of interest for our approach as a way to normalize the complex representations we have implemented and, possibly, the atomic concepts themselves.

In terms of data completeness, it is clear that Qualia descriptions will never be comprehesive knowledge repositories for a given concept, with all its facets. In our approach, due to a lack of existing resources, Qualias are mostly described manually. Even via the use of bootstrapping techniques, it is clear that the Qualia of a concept C (e.g. vaccine) essentially contains the most typical features (encoded via concepts, which themselves can originate Qualias). An incremental automatic acquisition of Qualia features would be crucial and helpful, but this raises complex problems such as consistency or granularity management.

3 A NETWORK OF QUALIAS TO CHARACTERIZE THE GENERATIVE EXPANSION OF ARGUMENTS

Before generating any argument synthesis, it is necessary to organize the set of concepts at stake in these arguments, in particular those which are supported or attacked w.r.t. the controversial issue.

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 directly derived from these root concepts, via their Qualia. In particular, concepts related to various types of parts of the concept, purposes, functions and uses of the concept are frequently found in arguments, whatever their polarity. For example, arguments can attack properties or purposes of the adjuvant, which is a part of a vaccine or the way a vaccine avoids dissemination of a disease. Besides the telic role, the agentive is also a crucial role since, for example, arguments often attack the way a vaccine has been tested, or its purchase cost.

From these observations, a network of Qualias can be defined to organize the concepts and knowledge structures involved in the arguments. This network is, for the time being, limited to three levels because derived concepts must remain functionally close to the root concepts to have a certain argumentative weight. However, some arguments, quite remote from the main concepts of the issue may have a

strong weight because of the hot concepts they include, e.g. *vaccination prevents bio-terrorism.*

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 sub-roles,
- $T_j^{i,X}$ 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 { $T_j^{i,X}$, $Q'_{i'}$: [$R_X : T1_{j'}^{i',X}$]}.

In the case of issue (1), nodes form the set T which corresponds to the terms in the Qualias of vaccine(X) and Ebola, some of which are terminal and others non-terminal.
- Step 2: similarly, the terms $T1_{j'}^{i',X}$ 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 $T1$, derived from T.
- Step 3: the same operation is carried out on $T1$ to produce $T2$.
- Final step: production of $T3$. The set of concepts involved is: $\{T \cup T1 \cup T2 \cup T3\}$.

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 also 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 (Saint-Dizier 2016b). Natural language words or expressions that lexicalize each concepts can be associated with each network nodes.

An important issue is to evaluate if and how this network defines a kind of 'transitive closure' that would characterize the typical and most frequent concepts that appear in arguments that support or attack an issue. Obviously, unexpected arguments may arise with concepts not in this network, probably with a lower frequency and recurrence.

The total number of concepts at stake in arguments for an average size issue, such as issues (1) and (3), is about 40 concepts, with non-homogeneous usages. A rough estimate indicates that about 80% of the arguments related to an issue can be recognized on the basis of these concepts.

The 'transitive closure' induced by this network is obviously not perfect, but quite efficient. The arguments which are not found are rather unexpected, but of much interest. For example, arguments such as: *vaccinations prevents bio-terrorism, vaccination raises ethical and racial problems* are found. These arguments are stored in a specific cluster called 'Other', so that they can be accessed in the synthesis.

Let us now illustrate the construction of this network. For example, from 'vaccine', two nodes are candidates: {active_principle, adjuvant}.

Assuming that, e.g. active_principle is a terminal concept, and adjuvant a non-terminal one, then, active_principle is associated with words such as 'active principle, stem cell'. 'Adjuvant' being non-terminal, its Qualia is included into the network at step 1:

Adjuvant(Y,X1):

$$\begin{bmatrix} \text{FORMAL} : \left[\text{MEDICINE, CHEMICALS}\right], \\ \text{TELIC}: \left[\text{DILUTE(Y,X1), ALLOW(INJECT(X1,P))}\right] \end{bmatrix}$$

The concepts in the formal and telic roles (medicine, chemicals, dilute(Y,X1), inject(X1,P) originate new Qualias, these are considered at step 2. Natural language terms are associated to these concepts, e.g.: *medicine, chemicals, inject, injection, dilute, dilution.*

Similarly, test(T,X), in the agentive role of vaccine(X), applied to vaccines (and medicines more generally), originates a node in T, and additional nodes in T1, T2 from its non-terminal concepts:

Test(T,X):

$$\begin{bmatrix} \text{CONSTITUTIVE}:\left[\text{PARTS OF A TEST: DATA, PROTOCOL}\right] \\ \text{TELIC}: \begin{bmatrix} \text{MAIN: EVALUATE(T,PROTECTION(X,Y, A))}, \\ \text{EVALUATE(T,SIDE-EFFECTS(X,Y, A))} \end{bmatrix} \\ \text{FORMAL} : \left[\text{SCIENTIFIC ACT}\right] \\ \text{AGENTIVE} : \left[\text{ELABORATE(T,X)}\right] \end{bmatrix}$$

Then, arguments may attack or support concepts present in test, such as the evaluation of the protection or the test protocol that has been used.

4 GENERATING AN ARGUMENTATIVE REPORT FROM A CONTROVERSIAL ISSUE

4.1 Main arguments to include in the synthesis

Let us consider the arguments found in issue (1) that must be included in a synthesis. Arguments mainly attack or support salient features of the main concepts of the issue or closely related ones by means of various forms of evaluative expressions. Among 50 non-overlapping arguments, the main arguments associated with issue (1) are, omitting associated discourse structures:

Supports:
vaccine protection is very good;
Ebola is a dangerous disease;
there are high contamination risks;
vaccine has limited side-effects,

there are no medical alternative to vaccine, etc.
Attacks:
there is a limited number of cases and deaths compared to other diseases;
7 vaccined people died in Monrovia,
there are limited risks of contamination,
there is a large ignorance of contamination forms,
competent staff is hard to find and P4 lab is really difficult to develop;
vaccine toxicity has been shown,
vaccine may have high side-effects,
Concessions or Contrasts:
some side-effects;
production and development costs are high;
vaccine is not yet available;
a systematic vaccination raises ethical and freedom problems.

The type of synthesis we propose reduces these expressions to an evaluation of the main concepts, as found in the Qualia structure, as developed in section 3. The number of arguments for and against each concept is given to outline the balance between each tendency. This however remains a tendency because this number of arguments depends on how many texts have been processed and how may arguments have been mined. The comprehensive list of arguments is stored in clusters and are accessible via navigation links from the concepts in the synthesis (section 4.3).

4.2 Synthesis Input Data: annotated arguments

The output of the mining system, which is the starting point of the synthesis construction, includes the following attributes (Saint-Dizier 2016), associated with each argument:

- the argument identifier (an integer), in our first experiment, all arguments attack or support the controversial issue and no other arguments,
- the **text span involved** that delimits the argument compound and its kernel, which ranges from a few words to a paragraph. In the synthesis, only the kernel of the argument is considered,
- the **polarity of the argument** w.r.t. the issue: support or attack. Additional intermediate values (argumentative concessions and contrasts) could be added in the future,
- the **concepts involved**, to identify the argument: list of the main concepts from the Qualias used in the mining process. Only the concepts found in the main argument section are considered. Those in adjoined discourse structures will be considered for higher-level synthesis in a later stage, to identify, e.g. restrictions.
- 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.

The 'ConceptsInvolved' attribute is structured from the root node, as a kind of path, so that the concept that is involved is clearly identified. This attribute may contain an ordered list of paths if several concepts are involved. A typical path is a sequence:
root-concept/(Role/Concept)*,
where the Concept is a predicate or a constant of a Qualia structure found under the role 'Role'.
For example, the concept of 'protocol' is defined as follows:
vaccine(X)/agentive/test(T,X)/constitutive/protocol.
since protocol is a concept associated with the constitutive role of the concept 'test'. Besides a clear identification of the concept 'protocol', this path can be used as (1) a way to structure a synthesis and (2) a way to provide some explanation of why an utterance is an argument by outlining its relation(s) with the root concept.

Argument 11:
Even if the vaccine seems 100% efficient and without any side effects on the tested population, it is necessary to wait for more conclusive tests before making large vaccination campaigns. The national authority of Guinea has approved the continuation of the tests on targeted populations.
is composed of an argument kernel (*it is necessary to wait for more conclusive tests before making large vaccination campaigns*) modified by two discourse structures. This argument is tagged as follows:
<argument Id= 11,
polarity= attack
conceptsInvolved= 'vaccine(X)/agentive/ test(T,X)'
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 tests 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 meta-data is considered such as the date of the argument or the author status. This notation was defined independently of any ongoing task such as ConLL15.

4.3 Example of a argumentation synthesis

Let us now characterize the form of the synthesis. For issue (1) the synthesis of the examples given in 4.1 is organized as follows, starting by the concepts which appear in the issue (root concepts) and then considering those, more remote, which appear in the derived concepts constructed by the network of concepts. Each line of the synthesis is produced via a predefined language pattern. Between parenthesis, the total number of occurrences of arguments mined in texts for that concept is given as an indication. This number is also a link that points to the arguments that have been mined in their original textual form. For each line, the positive facet is presented first, followed by the negative one when they exist, independently of the occurrence frequency, in order to preserve a certain homogeneity in the reading:

5

Vaccine protection is good (3), bad (5).
Vaccine avoids (5), does not avoid (3) dissemination.
Vaccine is difficult (3) to develop.
Vaccine is (4) expensive.
Vaccine is not (1) available.
Ebola is (5) a dangerous disease.
Humans may die (1) from Ebola.
Tests of the vaccine show no (2), high (4) side-effects.
Other arguments (4).

5 ARGUMENT SYNTHESIS GENERATION

Given the input data and the output forms presented in 4.3, let us now develop the grammatical and lexical environment that allows the generation of this synthesis. The ordering of the synthesis is based on the path mentioned in the 'conceptsInvolved' attribute of each argument. Arguments are sorted on the basis of this attribute.

5.1 The lexico-grammatical generation system

The synthesis of arguments is based on abstract linguistic patterns defined as follows:
(1) `[HeadConcept, Be/Predicate, Evaluative, AttributeLexicalization].`
or:
(2) `[HeadConcept, Be/Predicate, AttributeLexicalization. Evaluative].`

The symbol `HeadConcept` is the lexicalization of the rightmost (or leaf) concept in the attribute 'conceptsInvolved'. For example, in:
conceptsInvolved= 'vaccine(X)/agentive/ test(T,X)'
the rightmost concept is 'Test(T,Z)', its lexicxalization, stored in a lexicon, is 'test'. For example:
`word([test], noun, abstract, test(X,Y)).`
The function lexicalization(Word, Predicate) extracts the lexical item from the appropriate lexical entry. Finally, when the path given in 'conceptsInvolved' is long (from 2 concepts), a lexicalization of the whole path is produced, e.g. for the example above: 'test of the vaccine' instead of 'test' alone, using the basic compound NP pattern:
`[A of B]` or
`[A of B of C]`
where A, B and C are concepts of the path:
'C/role/B/role'/A.

The symbol `Be/Predicate` entails the lexicalization of the main predicate of the sentence. It is either the neutral 'be' (*is, are*) or a specific lexicalization if the attribute name that is considered includes a higher-level predicate such as *prevent, evaluate, allow, avoid* as shown in the Qualia structure above. When there are supports and attacks, this verb appears as such and then modified by a negation so that supports and attacks can be differentiated in the synthesis, using the following patterns:
Supports:
`[is/are/Verb, (Stats)]`,
Attacks:
`[is not/ are not/do not Verb/ does not Verb, (Stats)]`
Supports and attacks:
`[is/are/Verb, (Stats1), is not/ are not/do not Verb/ does not Verb, (Stats2)]`

The symbol `(Stats)` simply indicates the number of arguments that have been mined with the 'conceptInvolved' path considered. These statistics are indicative since they depend on the volume of text that has been mined. They also do not account for the strength of the argument. (Stats) is a link to the set of mined arguments that the reader may wish to inspect. These are stored in a cluster (section 4.2) and sorted from the argument(s) that have the highest strength to those that have the lowest one.

The symbol `Evaluative` is an evaluative expression, often a scalar adjective, modified by a negation depending on its polarity, the existence of an antonym and the polarity of the arguments to represent. Adjectives, as well as nouns, have their semantic characteristics stored in their lexical entry. The lexicalization of `Evaluative` is defined as follows:
(1) by default the values good / bad for products and attitudes and easy/difficult for processes. However, these adjectives are not very accurate and specific values are preferred.
(2) by an adjective found in one of the arguments of the cluster. The adjective must be prototypical. For that purpose, we use a resource we defined for opinion analysis, where about 500 of the most standard adjectives are organized in non-branching proportional series (Cruse 1986). Each series corresponds a precise conceptual dimension such as shape, cost, temperature, difficulty, peace, availability, etc. Most series are composed of a few positive and negatively oriented terms and possibly a neutral point, terms are structured with a partial order. Other series correspond to boolean adjectives and are simply composed of two elements. For example, starting from the most negative term:
temperature: *frozen - cold - mild - (warm, hot) - boiling.*
prototypical: cold / warm, neutral: mild.
toxicity : *poisonous - dangerous - neutral - recommended - beneficial.* prototypical: dangerous / beneficial, neutral: neutral.
cost: *expensive - (reasonable, appropriate) - cheap.*
In the cluster of arguments being processed the adjectives used as evaluators are collected and their inclusion in one or more non-branching proportional series is investigated. The series that is the most frequently refereed to is kept, and both the positively and negatively oriented typical adjectives are used in the lexicalization.

The symbol `AttributeLexicalization` is the direct lexical item that corresponds to the concept. In our approach this lexical item is stored in a lexicon where lexical entries include lexical items and their associated logical representations, as described above. When the attribute is propositional, the same strategy is used, in that case, an expression is produced via a pattern instead of a single item. This is the case for example with 'develop(T,X)' which gets the realization *to develop.*

This synthesis generation system is quite simple at the moment. Besides domain lexical entries, related to the concepts in the Qualias (between 50 and 100 lexical entries depending on the issue), the system currently uses 22 patterns that allow to produce the constructions presented above. These patterns are stable for the type of issues we consider, which are simple, with arguments which are direct and concern a single concept of the network. This generation system needs further investigation for more abstract issues or complex situations such as controversial dialogs. In this first stage, only the relations between the controversial issue and arguments have been investigated. In a 'real' argumentation, arguments may attack or support other arguments instead of the issue. It is not clear at the moment whether the same generation procedure can be used. Probably, to keep the synthesis readable, additional devices such as navigation links would be needed to indicate that an argument gets supports or attacks from others.

The system described here is relatively simple to implement. A first implementation has been realized with the logic-based platform <TextCoop> we developed for discourse analysis. This platform can also be used for language generation since it is fully declarative and partly reversible. However, the strategy used in <TextCoop> needs to be revised so that the simplest structure is generated first. The parsing strategy that is used in <TextCoop> is indeed a priori oriented towards language parsing.

5.2 Features of an Evaluation

This approach to argument synthesis generation is relatively simple and straightforward. The two levels: synthesis and links to the exact arguments stored as clusters seems to be a good compromise between proliferation of data and over generalization via a few synthetic lines. Elaborating an appropriate level of synthesis and the way to realize it linguistically needs a lot of experimentations and tunings.

A direct evaluation of this system must be realized along at least the following main features:

- the overall linguistic adequacy of the generation system, based on the patterns presented in the previous section. These patterns produce short sentences that readers can understand, their linguistic adequacy and language overall quality must be evaluated and possibly tuned.

- the types of domains and controversial issues for which this system is adequate, from very concrete to more abstract, and for various amounts of arguments, from just a few to several hundreds, including duplicates. Several experiments show that Qualia structures can quite straightforwardly be specified for concrete areas, this is less easy for areas which manipulate abstract or very general purpose concepts. An evaluation of the limits of the approach and the means to overcome difficulties must be carried out.

- the load in linguistic and conceptual resource description for each domain where arguments are mined. This includes essentially Qualia structures, lexical entries and associated resources such as non-branching proportional series, a few generation patterns. Qualia structures are often related to a domain ontology. These resources are not very large for a domain, but they nevertheless require some manual effort. Another aspect is the management of the coherence of the resources when they are specified at different levels (e.g. lexical, conceptual).

- the adequacy of the conceptual model, here the Qualia structure. It is necessary to show that it is indeed sufficiently accurate in a large number of domains.

A higher level evaluation concerns the adequacy of this synthesis for professionals who want to access to an argument synthesis, where arguments do correspond to their view and analysis of the domain. For example, in the hotel and restaurant domains, the features at stake are well identified in consumer evaluation. For more abstract or less common areas such as the issues developed in this paper, there is a need to make sure that the concepts developed in the Qualias do correspond to the vision of professional users, otherwise such a system will not be of much practical relevance.

We also feel there is no unique form of synthesis: several forms of synthesis could be foreseen that would depend on the reader's interests and profile. A real evaluation of the system presented here requires the development of adequate protocols to measure the relevance of various forms of synthesis (more abstract, more or less concise, using various types of concepts, with appropriate lexicalizations). This can only be done with large and diverse populations of users over several domains.

6 CONCLUSION AND PERSPECTIVES

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 in argument mining is to organize the arguments which have been mined to generate a synthesis that is readable, synthetic enough and relevant for various types of users.

Based on the Generative Lexicon (GL) Qualia structure, which is a kind of lexical and knowledge repository, we have shown how to construct a synthesis that captures the typical elements found in arguments and their polarity. We propose a two-level approach: a synthesis of the arguments that have been mined and, associated with the elements of this synthesis, navigation facilities that allow to access the argument contents in order to get more details.

The work presented in this paper is a first, exploratory experiment. Constructing an argument synthesis from a large diversity of issues, and in various contexts (dialogs, consumer opinion expression, etc.) , where arguments may also attack each other, is a complex task. The type of synthesis that would be really useful to the public and to professionals requires a close cooperation with opinion analysts and related professional. Additional features of arguments such as reliability, strength, validity, persuasion, etc. should also be incorporated at some future stage.

The next steps of our work include:

- the development of other issues and the annotation of related arguments by at least two annotators, this would entail a further validation of our model,
- the development of a larger argument mining system based on knowledge, in particular as structured in the Qualia,
- the development of tools that would contribute to the creation of Qualias from texts, we have some ongoing work on this very crucial dimension,
- the development of adequate and relevant evaluation protocols that would analyze the adequacy of the type of synthesis that is produced, as such and w.r.t. users expectations and implicit model of the domain.

REFERENCES

[1] 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.

[2] A. Cruse, Lexical Semantics, Cambridge university Press, 1986.

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

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

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

[6] 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.

[7] I. Mani, M. Maybury, 1999. Advances in Automatic Text Summarization, MIT Press.

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

[9] 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.

[10] 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).

[11] J., Pustejovsky. 1995. The Generative Lexicon, MIT Press.

[12] P. Saint-Dizier. 2016. Argument Mining: the bottleneck of knowledge and lexical resources, LREC, Portoroz.

[13] P. Saint-Dizier. 2016b. Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure, proceedings of INLG, Edinburgh.

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

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

[16] 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.

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

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