HAL
open science

A Two-Level Approach to Generate Synthetic Argumentation Reports

Patrick Saint Dizier

To cite this version:

Patrick Saint Dizier. A Two-Level Approach to Generate Synthetic Argumentation Reports. Argument and Computation, 2018, 18 years of Computational Models of Natural Argument, 9 (2), pp.137-154. 10.3233/AAC-180035 . hal-02191823

HAL Id: hal-02191823

https://hal.science/hal-02191823v1

Submitted on 23 Jul 2019

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.

Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in:
http://oatao.univ-toulouse.fr/22553

Official URL

DOI : https://doi.org/10.3233/AAC-180035

To cite this version: Saint-Dizier, Patrick *A Two-Level Approach to Generate Synthetic Argumentation Reports.* (2018) Argument and Computation Journal, 9 (2). 137-154. ISSN 1946-2166

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

DOI 10.3233/AAC-180035

A two-level approach to generate synthetic argumentation reports

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

Abstract. Given an issue and related propositions for or against it, a major challenge to make these propositions accessible to readers is to produce a synthesis that is readable, synthetic enough, and relevant for various types of needs and points of view. 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, via a number of preliminary experiments, how to construct a synthesis that outlines the typical elements found in propositions for or against a controversial issue. We propose a two-level approach: a synthesis of the propositions that have been mined, and navigation facilities that allow users to access arguments, structured in clusters, in order to get more details. This approach contributes to characterize why and how propositions support or attack an issue or some facets of that issue.

Keywords: Argument mining, natural language processing, knowledge representation, Generative Lexicon Qualia Structure

1. Aims and challenges

One of the main goals of argument mining is, given an issue, to identify in a set of texts the propositions for or against that issue. These propositions are supports (or justifications) or attacks of the issue or of some facets of the issue when the concepts in this issue are complex.These propositions may also attack or support other propositions which support or attack that issue, in order to reinforce or cancel out their impact. These propositions are difficult to identify, in particular when they are not adjacent to the issue, possibly not in the same text, because their linguistic, conceptual or referential links to that issue are rarely explicit but require knowledge and inferences. Within the framework of mining arguments, our aim is to identify as precisely as possible how and why these propositions become part of arguments for or against a given issue, in particular which facet(s) of the issue is concerned and how it is supported and attacked.

For example, given the issue: *Vaccine against Ebola is necessary*, identifying whether the following propositions participate in an argument and, if so, their argumentative orientation: *Ebola adjuvant is toxic*, *Ebola vaccine production is costly*, and *7 people died during Ebola vaccine tests*, cannot be realized solely on the basis of linguistics data, but requires domain knowledge. As an illustration, a knowledge-based analysis of the third statement shows that it is irrelevant or neutral w.r.t. the issue, as shown in [18], whereas a 'naive' reading would rather interpret it as an attack. We show in [20] that argument mining cannot be realized solely on surface form indicators such as linguistic clues but requires knowledge and various forms on inferences.

Constructing a synthesis of the propositions associated with a controversial issue is a necessary component of argument mining: it is frequent to mine in various types of texts and media, more than 100

This article is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0).

reasons related to an issue, where a number of them may largely overlap. It is then difficult and inefficient for users to read long lists of unstructured propositions where some essential points may be missed. Looking at how users proceed, it turns out than even simple forms of synthesis, via e.g. clustering techniques, are essential to make argumentation accessible. So far very few efforts have been devoted to this aspect of argument mining. Designing an argument synthesis is a delicate task that requires input from users, e.g. they may be required to specify their profile and needs and the level of granularity they expect. A synthesis must be simple: complex graphs must be avoided since they are rarely readable by users.

To carry out a well-organized synthesis, our claim is that such a task requires the same type of knowledge and conceptual organization as for mining arguments. The remainder of this article develops this perspective for the generation of an argument synthesis. The results presented in this article are exploratory and based on simple experimentations. They remain relatively preliminary and need further investigations. They however induce research directions in this area that have not been explored so far in conjunction with argument mining. A set of preliminary investigations towards the construction of an efficient argument synthesis are discussed.

This article is organized as follows. Section 1 presents the challenges and the state of the art in argument mining and some relevant features of text summarization. Section 2 develops the features of our knowledge-driven argument mining system [19]. This section introduces knowledge representation aspects which will be used in the remainder of this article. These are inspired by and enhance the Qualia structure of the Generative Lexicon (Section 2.2). Section 3 shows how a network of concepts, relevant for a synthesis, can be induced from a given controversial issue. Sections 4 and 5 develop our synthesis approach and the set of language patterns which are necessary for that purpose.

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. [10,13], for example for opinion analysis, e.g. [22], mediation analysis [9], scientific text analysis [8], advertising [2], or transcribed argumentative dialog analysis, e.g. [3,21]. The analysis of the NLP techniques relevant for argument mining from annotated structures is reviewed and analyzed in e.g. [14]. Goal-based reasoning for argumentation analysis is developed in [24]. Annotated corpora of various types are now available, e.g. the AIFDB dialog corpora or works reported in [23], among many others. These corpora are very useful to understand how argumentation is realized in texts and dialogs, e.g. to identify Argumentative Discourse Units (ADUs), linguistic cues at stake [12], and argumentation strategies, in a concrete way, possibly in association with abstract argumentation schemes, as shown in e.g. [6]. Finally, reasoning aspects related to argumentation analysis are developed in e.g. [7,17] and [25] from a formal semantics perspective with e.g. a study of concessive and contrastive connectors.

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

In [20], a corpus analysis identifies the types of knowledge and inferences that are required to develop argument mining within a symbolic perspective. The structure of this corpus is briefly reported in this paper. Then, we show, on the basis of that corpus, that the Generative Lexicon (GL) [16] could be an appropriate linguistic and conceptual model and organization, sufficiently expressive, to characterize the types of knowledge and lexical data that are required to accurately identify propositions related to

an issue. This model develops the crucial conceptual aspects that relate propositions to an issue. The conceptual dimensions attacked or supported in arguments are indeed related to the roles, functions, and purposes of the head entities in the issue at stake, or how these entities came into being, were created or produced. The GL Qualia structure is therefore considered here for the conceptual and linguistic structure it offers, independently of the initial uses developed by its initiator, in particular to model language variations and metonymy construction [16].

1.2. Natural language synthesis vs. summarization

In natural language generation, the main projects on argumentation generation were developed as early as 2000 [26] and [7]. Focus was on the production and the organization of arguments from a semantic representation, via some forms of macro-planning, an essential component of natural language generation of texts. A synthesis, and in particular an argument synthesis is a much simpler operation which consists in grouping already existing propositions related to an issue on the basis of their similarity in order to enhance readability. It is therefore more superficial that the generation of arguments.

While there are currently several research 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 redundancies and various types of fallacies which are less relevant in a synthesis that is aimed at being as neutral as possible. This synthesis may have the form of a short text, possibly including navigation links, or it may be represented by means of a network or by various types of graphical representations that organize the reasons or justifications for or against an issue. The major problems are readability, granularity and conceptual adequacy with respect to users.

In [19], we show how propositions that have been mined can be organized in hierarchically structured clusters, according to the conceptual organization proposed by the Generative Lexicon, so that readers can navigate over and within sets of related arguments. This approach turned out not to be synthetic enough, since over 100 reasons can be mined for a given issue, making the perception of the main attacks and supports quite difficult. Nevertheless, this initial approach allows the construction of an argument database with respect to an issue which is useful to readers who wish to access the exact form of arguments that have been mined. This is illustrated in Fig. 1, where the relation of the concept to the issue (part of, purpose, etc.) is specified to enhance the structure of the description. This figure simply shows a form of synthesis readers would like to have, independently of any knowledge representation consideration.

The present contribution goes one step further: it aims at producing a synthesis that is short and efficient where the concepts present in the GL Qualia structures are used to abstract over sets of propositions

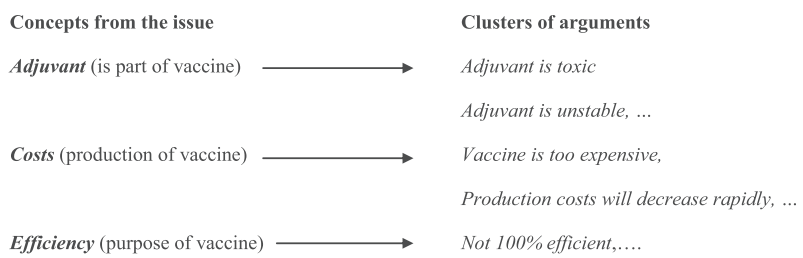

Fig. 1. Illustration of a synthesis.

that support or attack an issue at stake. The structure of clusters of propositions is kept and made accessible via links from this synthesis.

This contribution to natural language argumentation synthesis is not really a summarization task, as e.g. developed in [11]. 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 of the reasons mined for or against an issue, that involves forms of clustering.

From that perspective, this work could be viewed as a preliminary step to a summarization procedure. A real summarization task would involve e.g. identifying prototypical arguments and themes, and then constructing summaries for each group of related arguments, but this is beyond the present research.

In terms of feature classification and relevance, we show below that the concepts used in the Qualia structure of the Generative Lexicon are appropriate for this synthesis task. The concepts in the Qualia are defined from a domain ontology, similarly to the features evaluated in most opinion analysis systems. We show that these features can be used as entry points to the cluster system and as a measure of the relevance of a proposition with respect to the issue at stake. A challenging point is that these concepts and their internal organization must correspond as much as possible to the user perception of the domain to which the issue belongs, otherwise the synthesis may not be so useful.

2. Features of a knowledge-based argument mining approach

2.1. Corpus analysis: Identification of the required knowledge

To explore and characterize the forms of knowledge that are required to develop argument mining in texts, we assembled and annotated four corpora based on four independent controversial issues as reported in [18]. The main results are summarized below. These corpora are relatively small. They are designed to explore the problem, and to elaborate a model, not to design a comprehensive argument mining system. The texts considered are extracts from various sources, e.g.: newspaper articles and blogs from associations. Issues are:

(1) vaccination against Ebola is necessary,
(2) women's situation in India has improved,
(3) nuclear plants are not so dangerous and
(4) organic agriculture is the future.

The total corpus includes 51 texts, a total of 24500 words for 122 different propositions that support or attack one of the above issues. From our manual analysis, the following polarities are observed: attacks: 51 occurrences, supports: 32, argumentative concessions: 17, argumentative contrasts: 18 and undetermined: 4.

Our analysis shows that for 95 propositions (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 propositions over 4 domains. Even if the notion of concept remains somewhat vague, these results are nevertheless interesting to develop large argument mining systems. 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) [16] is an attempt to structure lexical semantics knowledge in conjunction with domain knowledge. Its main structure is the Qualia structure. The notions used in the GL Qualia structure are not new: the roles postulated in the Qualia structure are derived from the Aristotelian class of causes: constitutive for material, formal for formal, telic for final and agentive for efficient. These roles do not overlap with the notions of grammatical roles or semantic roles. They are much more powerful from a semantic and knowledge representation point of view.

In the GL, the Qualia structure of an entity is both a lexical and knowledge repository composed of four fields called qualia 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, i.e. the entity in its environment. It may be structured into several sub-roles such as shape, dimensions, position, etc.
- **the telic role** describes the entity functions, uses, roles and purposes
- **the agentive role** describes the origin of the entity, how it came into being, was created or produced.

To illustrate this conceptual organization, let us consider the controversial statement (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):

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

The main concepts to be found in the Qualia structure of *Ebola* are:

Ebola:

$$\begin{bmatrix} \text{FORMAL: [VIRUS, DISEASE]}, \\ \text{TELIC: } \begin{bmatrix} \text{INFECT}(E_1, \text{EBOLA, P}) \Rightarrow \text{GET_SICK}(E_2, \text{P}) \\ \Rightarrow \diamond\text{DIE}(E_3, \text{P}) \wedge E_1 \leqslant E_2 \leqslant E_3 \end{bmatrix} \end{bmatrix}$$

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 such an ontology does not exist for this domain. Besides these predicates and constants, we have included formula as attribute values. In the above example of Ebola, event variables E_i have been added to deal with complex events. In the Qualia of vaccine, subtypes (e.g. MAIN, MEANS) have been added to provide a more accurate structure to the main four roles. More examples are available in [18].

Qualia structures can be hierarchically organized, as in any ontology. *Vaccine* is a kind of *medicine*, it therefore inherits of the properties of a *medicine*, i.e. the predicates present in *medicine* are inherited,

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-PAROLE EEC project.

Finally, from the two above Qualias and via formula expansion, the formal representation of the controversial issue is (using the modal symbols of necessity and possibility):

$$\Box\ (protect_from(X, Y, (infect(E1, ebola, Y) \Rightarrow get_sick(E2, Y) \Rightarrow$$

$$\Diamond\ die(E3, Y))) \wedge avoid(X, dissemination(ebola)).$$

2.3. Using knowledge for argument mining

Originally, the Qualia structure was designed to characterize sense variations around a prototypical (or domain-stabilized) 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 [16]. In our view, this is a specific interpretation of a more global typology of object descriptions.

In our approach, we view the Qualia structure as a means to represent 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 introduce more flexibility and higher-order representations in Qualias, in particular complex structures such as formula, modalities, etc. instead of just atomic concepts. Manipulating such structures is more complex, but corresponds to the needs in terms of representation. Reasoning with these complex forms is addressed in [19].

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 comprehensive 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. These are encoded via concepts, which themselves can originate Qualias. An incremental automatic acquisition of Qualia features would be crucial and helpful. This raises complex problems such as consistency or granularity management which are outside the scope of this article. Some elements for automatically acquiring Qualias are developed in [4]. In this article, focus is on the model provided by the GL Qualia, which seems to be particularly relevant for knowledge-based argument mining, without excluding the use of other paradigms to specify e.g. the contents of constitutive roles (from WordNet) or verb arguments as implemented in VerbNet via primitives.

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 in the controversial issue that is considered with respect to their possibility to be attacked or supported. Our observations show that propositions found in texts attack or support (1) specific concepts found in the Qualia

of the head terms in the controversial issue (called **root concepts**) or (2) concepts found in one of the Qualias associated with these root concepts (called **derived concepts**). In particular, concepts associated with or derived from 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 (described in the adjuvant Qualia), or the way a vaccine avoids dissemination of a disease (described in the dissemination Qualia). Besides the telic role, the agentive is also a crucial role since, for example, arguments often attack the way a vaccine has been produced or 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 topics they include, e.g. *vaccination prevents bio-terrorism*.

3.1. Network construction

A Qualia Q_i describes a number of functional and structural features of a concept. It can be formally defined as follows:

$$Q_i : \left[R_X : T_j^{i,X}\right], \quad \text{where:}$$

- R_X denotes the four roles: $X \in \{formal, constitutive, agentive, telic\}$, possibly sub-roles may be added as in the illustrations above,
- $T_j^{i,X}$ is a term which is a formula, a predicate or a constant T_j in the role X of Q_i.
- Nodes are of two types: [terminal concept], with no associated Qualia or [non terminal concept], with an associated Qualia

A network of Qualias, for a given controversial issue, is then defined as follows:

- the root is the semantic representation of the controversial issue and the Qualias Q_i of its head concepts (in our four examples: vaccine and Ebola, women's conditions, nuclear plants and organic agriculture),
- Step 1: the first level of the network is composed of the nodes which correspond to the concepts $T_j^{i,X}$ 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 $T2$.
- Step 4 (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 organi-

zation 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 [19]. Natural language words or expressions that lexicalize each concept are associated with each network nodes.

An important concern 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 propositions which are not found are rather unexpected, but of much interest. These may indeed be related to specific local social or historical contexts. For example, propositions such as: *vaccination prevents bio-terrorism, vaccination raises ethical and racial problems* are found. These propositions are stored in a specific cluster called 'Other', so that they can be accessed from the synthesis.

The total number of concepts at stake in arguments for an average size issue, such as issues (1) and (3), is about 40 to 50 concepts, depending on the granularity implemented in Qualia. Quite different occurrence frequencies are observed. A rough estimate indicates that about 85% of the propositions related to an issue can be recognized on the basis of these concepts. The 'transitive closure' induced by this network is obviously not complete, but quite relevant.

3.2. An illustration

Let us now illustrate the construction of this network. For example, from 'vaccine', two nodes of the constitutive role 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* or *stem cell* which may be found in a terminology or a lexicon. The distinction between terminal and non-terminal concepts is defined a priori in domain ontologies or may be related to knowledge engineering decisions, not develop in this article. 'Adjuvant' being non-terminal, its Qualia is included into the network at step 1:

Adjuvant(Y, X1):

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

The concepts in the formal and telic roles are stored in T1:

{medicine, chemicals, dilute(Y, X1), allow(M, inject(X1, P))}.

Those concepts which are non-terminal originate new Qualias. These are considered at step 2 to construct T2. The higher-order predicate 'allow' is not considered as a concept in this example, therefore inject(X1, P) is the concept that is considered. 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 from its non-terminal concepts:

Test(T, X):

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

From these two examples, T (root concepts) and $T1$ (immediate daughter concepts) include the following elements, assuming that variables are bound within each set:

$$T = \big\{\text{adjuvant(Y, X), test(T, X)}\big\}.$$

$$T1 = \{\text{medicine, chemicals, dilute(Y, X), allow (M, inject(X, P)), data, protocol,}$$
$$\text{evaluate(T, protection(X, Y, A)), evaluate (T, side-effects(X, Y, A)), elaborate(T, X)}\}.$$

These concepts are associated with lexical terms or expressions, such as:

$$\textit{lexicalization}(T) = \{\textit{adjuvant}, \textit{test}, \textit{try out}, \textit{experiment}, \textit{etc.}\}.$$

$$\textit{lexicalization}(T1) = \{\textit{medicine}, \textit{chemicals}, \textit{inject}, \textit{injection}, \textit{dilute}, \textit{dilution}, \textit{evaluate protection},$$
$$\textit{test protection}, \textit{evaluate side effects}, \textit{side effects}, \textit{test protocol}, \textit{etc.}\}.$$

Besides other elements, such as evaluative expressions, these lexicalizations are used to detect the utterance relatedness to the controversial issue and to possibly qualify it as an argument. The linguistic structure of propositions that may attack or support an issue is described in [20].

Then, propositions may attack or support concepts present in T or T1, such as the evaluation of the protection offered by the vaccine or the test protocol that has been used.

4. Generating an argumentative report from a controversial issue

4.1. Main arguments found in texts

Let us consider the arguments found in texts for 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. Contrasts in our terminology include propositions which are weak attacks, whereas concessions are weak supports. 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 alternatives to vaccine, etc.

Attacks:

there is a limited number of cases and deaths compared to other diseases;
7 vaccinated 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.

4.2. Structure of a argumentation synthesis

Let us now characterize the form of the synthesis, as it can be derived from the examples given in Section 4.1. A synthesis of the examples given in Section 4.1 can be organized as follows, starting by the concepts which appear in the issue (root concepts in T) and then considering those, more remote, which appear in the derived concepts constructed by the network of concepts, e.g. those in $T1$.

To make the synthesis readable and informative, between parentheses, the total number of occurrences of arguments mined in texts for that concept is given as an indication. The goal is to outline strengths, recurrences and tendencies. This however remains informal because occurrence frequencies are very much dependent on how many texts have been processed and how many arguments have been mined. 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:

Vaccine protection is good (3), bad (5).
Vaccine avoids (5), does not avoid (3) dissemination.
Vaccine is difficult (3) to develop.
Vaccine is expensive (4).
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).

Each line of the synthesis is produced via a predefined language pattern. This is developed in the next sections. The comprehensive list of arguments is stored in clusters (Section 4.4) and are accessible via navigation links, represented in this synthesis by means of underlined numbers.

An alternative synthesis approach would have been to select for each concept an argument that is prototypical paired with a link to the other arguments related to the same concept. This approach sounds less artificial than the generation of abstract sentences, but (1) it is difficult to characterize what a prototypical argument is and (2) the formulation we propose is a priori shorter, more neutral and standardized. However, we feel that exploring other forms of synthesis, in close cooperation with users, should be investigated. For example, forms of rhetoric could be incorporated for non-neutral synthesis. Introductions to these aspects can be found, among many others, in [15], in [1] specific cases can be found in a number of articles of this journal.

4.3. Input data: Annotated arguments

The output of the mining system, which is the starting point of the synthesis construction, includes the following attributes [18], associated with each argument:

- the **argument identifier** (an integer): in this 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. This text span is found between the different tags of the annotation, there is no specific attribute to represent it
- the **polarity of the argument** with respect to 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, so far limited a priori to a few values: strong, average, moderate, weak. Investigations on this topic are ongoing: they show the high impact of context
- 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' in $T1$. Besides a non-ambiguous 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.

In terms of annotations, for example, Proposition 11 attacks the claim on vaccination:

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 kernel (*it is necessary to wait for more conclusive tests before making large vaccination campaigns*) modified by two discourse structures. This proposition 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>.

Discourse structures are those defined within the RST framework (http://www.sfu.ca/rst/). At this stage no meta-data is considered such as the date of the statement or the author status. This notation was defined independently of any ongoing task such as ConLL15.

4.4. Clustering arguments

The last step before the production of a synthesis is to cluster arguments according to their relatedness with respect to the controversial issue. The network of Qualias together with the attributes 'relationWith', 'Polarity' and 'ConceptsInvolved' are the most crucial elements in the annotation system to generate clusters that correspond to the domain conceptual organization. Each concept of the Qualia network potentially defines a cluster.

The generation of clusters proceeds as follows:

(1) Each concept 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 (4)

(2) When an argument involves several concepts, to avoid redundancy and large clusters, it is included into the concept cluster that is the highest in the network of concepts. In this experiment, clusters are defined around a single concept. However, there are cases where several concepts operate at the same level of relevance, and therefore clusters based on the conjunction of clusters must be constructed. This is the case for example for issue (3) where arguments are often based on comparisons: *nuclear wastes induce higher maintenance costs that coal wastes*. Concepts involved are: nuclear wastes, coal wastes and maintenance costs. These should define a specific cluster. In the

present experiment, only the concept 'nuclear waste' is considered, and the argument is tagged 'attack'

(3) For each concept, related arguments are structured by polarity: first supports and then attacks,

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

Let us consider the example on vaccination, then a set of clusters would be composed of the following arguments:

Clusters at level 0 of the concept graph: e.g.:

Cluster 1: Adjuvant: support: *there are different types of adjuvants and most of them are fine*, attack: *adjuvant is toxic, adjuvant is unstable* ...
Cluster 2: Dissemination: support: *vaccine reduces disease dissemination* ...
Cluster 3: Get-sick: concession: *limited number of cases and deaths compared to other diseases, vaccination allows the definition of a protection belt around infected zones*, ...

Clusters at level 1:, e.g.:

Cluster 4: Production costs: attack: *high production and development costs, need highly paid staff*, ...
Cluster 5: Availability: support: *vaccine can be produced in poor countries*, contrast: *vaccine not yet available, vaccine must be kept at a low temperature*, etc.

5. Argument synthesis generation

Given the input data as specified in Section 4.3, the type of clusters of arguments that have been produced Section 4.4 and general form of a synthesis presented in Section 4.2, let us now develop the grammatical and lexical environment that allows the generation of this synthesis. The ordering of the items of the synthesis is based on the path mentioned in the 'conceptsInvolved' attribute of each argument cluster. Arguments are sorted on the basis of this attribute, starting by the level 0 of the network, i.e. by those concepts which appear in the issue, and then considering derived concepts: those in levels 1, 2 and 3, corresponding to $T1$, $T2$ and $T3$ respectively.

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].`

Let us now describe informally how the different symbols of these patterns generate natural language expressions.

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 prototypical lexicalization, stored in a lexicon, is 'test'. For example:

```
word([test], noun, abstract, test(X, Y),
  [[try, out], [analysis]]).
```

The function lexicalization(Word, Predicate) extracts the lexical item from the appropriate lexical entry. In this lexical entry, the last argument contains closely related lexical terms. These are used in particular to detect relatedness.

When complex forms are given as attributes, such as 'avoid(T, dissemination(X))', then the content that is lexicalized corresponds to the whole formula: 'avoids the dissemination'. In our current approach such formula are generated on the basis of predefined patterns, for example:

```
Verb(X, Predicate(Y))  ⇒
  lexicalization(Verb(X)), lexicalization(Predicate(Y)),
```

In our example, *avoids dissemination* would be generated. Some morphological features need to be taken into account in the generation rules.

Finally, when the path given in 'conceptsInvolved' is long (from 2 concepts), in order to avoid any ambiguity, 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.

Next, the symbol `Be/Predicate` in patterns (1) and (2) above 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 value that is considered includes a higher-level predicate such as *prevent, evaluate, allow, avoid*. In that case, `Be/Predicate` is the lexicalization of the higher-level predicate found in the attribute. For example, the attribute: 'evaluate(T, protection(X,Y,A))' is lexicalized by: *evaluate the protection of*. The variables X, Y and A are lexicalized by other parameters in the patterns (1) or (2).

When there are supports and attacks, this predicative form appears as such and is then modified by a negation so that supports and attacks can be differentiated in the synthesis, using the following patterns:

Supports:

```
[is/are/Predicate, (Stats)],
```

Attacks:

`[is not/ are not/do not Predicate/ does not Predicate, (Stats)]`

In these patterns, `Predicate` represents the whole attribute. For example, an attack of the function of a vaccine `avoid(X, dissemination(D))` produces the expression:
does not avoid the dissemination.

Next, the symbol `(Stats)` indicates the number of arguments that have been mined with the 'conceptInvolved' path considered, i.e. the number of arguments in the corresponding cluster (Section 4.4). 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 corresponding cluster of arguments that the reader may wish to inspect.

Next, the symbol `Evaluative` in the patterns (1) and (2) is an evaluative expression, often a scalar adjective, possibly modified by a negation depending on its polarity, the existence of an antonym (which could be prefered to avoid negation) or the polarity of the argument. Adjectives, as well as nouns, have these 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 a scalar adjective found in one of the arguments of the corresponding 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 [5]. Each series corresponds to a precise conceptual dimension, concrete or abstract, such as height, weight, danger, cost, temperature, difficulty, peace, availability, etc. Most series are composed of a few positively and negatively oriented terms and possibly a neutral term. 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) - fine- cheap*, prototypical: *expensive/cheap*, neutral: *reasonable*.

In the cluster of arguments, the adjectives used as evaluators are collected and their inclusion in one or more non-branching proportional series is investigated. The terms that are the most frequently refereed to is kept, and both the positively and negatively oriented typical adjectives are used in the lexicalization.

Finally, the symbol `AttributeLexicalization` is the 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 and sub-patterns that allow to produce the constructions presented in Section 4.2. 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.

An evaluation of the current system, in order to identify improvement, 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 quantities of arguments, from just a few to several hundreds, including duplicates. Our experience shows 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, such as issue (3) of this experiment. 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). A strategy for automatically acquiring Qualias is developed in [4].
- 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, or to identify additional paradigms that must be integrated into the system.

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 own view and to the analysis they have

of the domain. For example, in the hotel and restaurant domains, the features at stake are predefined by professionals in consumer evaluations. For more abstract or less common areas such as the issues given in Section 2.1, 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, e.g. 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. This article is both an exploration of this latter issue and a simple preliminary proposal.

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 types of communication 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 professional users. Additional features of arguments such as reliability, in-depth analysis of strength, validity, persuasion, etc. should also be incorporated at some future stage.

The next steps of our work include:

- the development of other types of controversial issues and the annotation of their 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 procedures to facilitate language and Qualia structure acquisition
- the development of tools that would contribute to the creation of Qualias from texts: some work is ongoing 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 with respect to users expectations and implicit model of the domain.

Acknowledgements

The author is grateful to an anonymous reviewer for her/his insightful comments that contributed to improve this text. I also thank my institution, the CNRS.

References

[1] R. Amossy, *L'argumentation dans le Discours*, Armand Colin, 2010.
[2] A. Baker, Arguments in advertising, in: *Proceedings of CMNA17*, 2017.
[3] K. Budzinska, M. Janier, C. Reed, P. Saint-Dizier, M. Stede and O. Yakorska, A model for processing illocutionary structures and argumentation in debates, in: *Proceedings of LREC14*, 2014.
[4] V. Claveau and P. Sebillot, Automatic acquisition of generative lexicon resources, using an explanatory, symbolic technique, in: *Proceedings of Advances in Generative Lexicon Theory*, 2013.
[5] A. Cruse, *Lexical Semantics*, Cambridge University Press, 1986.
[6] V.W. Feng and G. Hirst, Classifying arguments by scheme, in: *Proceedings of 49th ACL: Human Language Technologies*, 2011.
[7] A. Fiedler and H. Horacek, Argumentation within deductive reasoning, *Journal of Intelligent Systems* **22**(1) (2007).
[8] N. Green, Argumentation mining in scientific discourse, in: *Proceedings of CMNA17*, 2017.
[9] M. Janier and C. Reed, Towards a theory of close analysis for dispute mediation discourse, *Journal of Argumentation* **22** (2015).
[10] C. Kirschner, J. Eckle-Kohkler and I. Gurevych, Linking the thoughts: Analysis of argumentation structures in scientific publications, in: *Proceedings of 2nd Workshop on Argumentation Mining*, Denver, 2015.
[11] I. Mani, *The Generative Lexicon*, John Benjamins Publishing, 2001.
[12] H. Nguyen and D. Litman, Extracting argument and domain words for identifying argument components in texts, in: *Proceedings of 2nd Workshop on Argumentation Mining*, Denver, 2015.
[13] R.M. Palau and M.-F. Moens, Argumentation mining: The detection, classification and structure of arguments in text, in: *Proceedings of 12 ICAIL*, 2009.
[14] A. Peldszus and M. Stede, From argument diagrams to argumentation mining in texts: A survey, *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)* **7**(1) (2013), 1–31. doi:10.4018/jcini.2013010101.
[15] C. Perelman and L.O. Tyteca, *The New Rhetoric: A Treatise on Argumentation*, University of Notre Dame Press, 1977.
[16] J. Pustejovsky, *The Generative Lexicon*, MIT Press, 1986.
[17] J. Redmond, *Logic, Argumentation and Reasoning, Interdisciplinary Perspectives from the Humanities and Social Sciences*, Springer, 2017.
[18] P. Saint Dizier, The bottleneck of knowledge and language resources, in: *Proceedings of LREC16*, 2016.
[19] P. Saint Dizier, Challenges of argument mining: Generating an argument synthesis based on the qualia structure, in: *Proceedings of INLG16*, 2016.
[20] P. Saint Dizier, Knowledge-driven argument mining based on the qualia structure, *Journal of Argumentation and Computation* **8**(2) (2017).
[21] R. Swanson, B. Ecker and M. Walker, Argument mining: Extracting arguments from online dialogue, in: *Proceedings of SIGDIAL15*, 2015.
[22] M.G. Villalba and P. Saint-Dizier, Some facets of argument mining for opinion analysis, in: *Proceedings of COMMA12*, IOS Publishing, 2012.
[23] M. Walker, P. Anand and R. Abbot, A corpus for research on deliberation and debate, in: *Proceedings of LREC 2012*, 2012.
[24] D. Walton, *Goal-Based Reasoning for Argumentation*, Cambridge University Press, 2015.
[25] G. Winterstein, What but-sentences argue for: An argumentative analysis of 'but', *Lingua* **122** (2012). doi:10.1016/j.lingua.2012.09.014.
[26] I. Zukerman, M. Roger and K. Korb, Using argumentation strategies in automatic argument generation, in: *Proceedings of INLG 2000*, 2000.