

Knowledge-driven argument mining based on the qualia structure

Patrick Saint-Dizier

IRIT-CNRS, 118 route de Narbonne, 31062 Toulouse Cedex, France

Abstract. Given a controversial issue, argument mining from texts in natural language is extremely challenging: besides linguistic aspects, domain knowledge is often required together with appropriate forms of inferences to identify arguments. Via the analysis of various corpora, this contribution explores the types of knowledge that are required to develop an efficient argument mining system. We show that the Qualia structure of the Generative Lexicon with some extensions and a specific interpretation has some expressive capabilities which are appropriate for this task.

Keywords: Argument mining, natural language processing, knowledge representation, Generative Lexicon qualia structure

1. Aims and challenges of argument mining

One of the main goals of argument mining is, given a controversial issue, to identify in a set of texts, the arguments for or against that issue. 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 and referential links to that issue are rarely direct and explicit. Arguments are often evaluative natural language statements which become arguments because of the specific relations they have with another evaluative statement considered as a controversial issue. Arguments can be supports or attacks, with various degrees of strength. Arguments can also attack other arguments.

Let us illustrate the need of knowledge in argument mining by two examples taken from two different situations:

Issue: *The situation of women has improved in India,*

Statement found in another text:

- (a) *early morning, we now see long lines of happy young girls with school bags walking along the roads.*

(a) is a support of the issue, but it requires knowledge and inferencing to make explicit, and possibly explain, the relationships between women's conditions and young girls carrying school bags, here education.

Let us now consider:

- (b) *School buses must be provided so that schoolchildren do not reach the school totally exhausted after a long early morning walk.*

(b) is an attack of (a), indeed: *these young girls may not be so happy of walking early morning*, but it is not an attack of the issue: the facet that is concerned in the relation between (b) and (a) does not concern women's conditions in particular.

Let us now consider another example taken from the BBC Moral Maze debates (with non-adjacent arguments, data from the Dundee IAT group):

Speaker 1: *Let's be straight, the situation in Kenya was a bloody dirty war.*

is related to and supported by an utterance from another speaker in the debate:

Speaker 2: *I think there was horror at Mau Mau.*

To identify this support relation, knowledge about Mau Mau populations is needed, in particular that they live in Kenya and had conflicts with the British.

Except in specific contexts, and for certain types of arguments (e.g. warnings, threats, advice, requirements), most statements do not have any specific linguistic mark that allow to directly identify (1) them as supports or attacks of a controversial issue, and (2) the facet of the issue they support or attack and how. Therefore, evaluating the role of knowledge issues in argument mining is necessary. Argument mining has to deal with two major problems: (1) **relatedness**: the argument must be thematically related to the issue at stake, and (2) **focus and relevance**: what facets of the issue are involved, how and how much.

This article explores this problem that has never been investigated before in the context of argument mining, to the best of our knowledge. For that purpose, an empirical analysis of the phenomena, largely manual, has been carried out. The aim of this exploratory task is to construct a relevant corpus and then to identify and categorize the forms of knowledge that are required to perform efficient argument mining. A preliminary conclusion to this exploration tends to show that a specific interpretation and use of the Qualia structure of the Generative Lexicon with some extensions has expressive capabilities which are appropriate for this task.

This paper develops preliminary results on a knowledge-driven argument mining approach and outlines the difficulties, the research questions and also the expected gains in accuracy and explicative power. It attempts to develop a working method to go beyond standard argument annotation so that the role and the importance of knowledge can be characterized. From these results, a more elaborated and principle-based analysis must be realized to characterize the relatedness aspects between an issue and an argument, or between an argument that supports or attacks another argument in an argumentation graph. This paper contributes to answering the following questions:

- How to construct a valid corpus to investigate the need of knowledge in argument mining? Our approach is to define issues a priori and then to manually search for arguments found in various types of texts (Sections 3.1 and 3.2),
- How to annotate arguments in order to explore how they are related to a controversial issue (Section 3)? What kind of new tags should be introduced, what level of granularity should be targeted and how to gradually go into more details on knowledge issues?
- How to categorize the needs in terms of knowledge and reasoning (Sections 3.2 and 4)? What are the main forms of knowledge which are involved (Sections 4 and 5)?
- How to categorize the knowledge involved (Section 5.2)?
- How to account for the diversity of arguments w.r.t. an issue (Section 5.3).

2. Challenges of argument mining

Argument mining is an emerging research area which introduces new challenges both in natural language processing (NLP) and in Artificial Intelligence. So far, most experiments and projects focused on

NLP techniques, based on corpus annotation, to mine arguments in various types of contexts. A synthesis of the techniques, annotated corpora and language resources which are used in the main current projects is given in <http://argumentationmining.disi.unibo.it/resources.html>. The analysis of the NLP techniques relevant for argument mining from annotated structures is analyzed in e.g. [1]. The analysis of the various forms of reasoning involved in argumentation is emerging; e.g. the articles in [2] which mainly focus on formal aspects of reasoning in argumentation.

Argument mining research applies to written texts, e.g. [3,4], for example for opinion analysis, e.g. [5], mediation analysis [6] or transcribed dialogue analysis, e.g. [7,8]. Annotated corpora are now available, e.g. the AIFDB dialog corpora (<http://corpora.aifdb.org/>). These corpora are very useful to understand how argumentation is realized in texts, e.g. to identify argumentative discourse units (ADUs), linguistic clues [9], and argumentation strategies, in a concrete way, possibly in association with abstract argumentation schemes, as shown in e.g. [10]. Finally, reasoning aspects related to argumentation analysis are developed in e.g. [11] and [12] from a formal semantics perspective with a study of connectors of contrast and concession.

In opinion analysis, the benefits of argument mining are not only to identify a customer 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). Categorizing arguments allows the construction of various forms of synthesis, organized e.g. in clusters by topic and subtopic to enhance the ease of access for readers when there are large amounts of arguments [13].

The analysis provided below shows that for about 75% of the situations some knowledge is required to accurately identify arguments w.r.t. a controversial issue. The analysis provided below is based on a simple and direct logical analysis of arguments, it does not include any pragmatic considerations which, however, play a prominent role in argumentation analysis. Pragmatic factors also involve knowledge, probably of a more generic nature, e.g. [14]. This is not developed in the remainder of this paper, but let us outline a few dimensions of these factors that will need to be taken into account at some stage of argument analysis. First the author of arguments frequently adapts his arguments to the potential readers, taking into account the *Doxa* (commonly admitted facts) and the main opinions shared by his audience. This *Doxa* may convey information that is incorrect, including potentially incorrect argumentative schemes in some contexts (e.g. quality must be preferred to quantity, what exists is preferable to what is only possible), e.g. [15]. Next, the author of arguments cares about the positive image he constructs of himself (*discursive Ethos*) in the organization of his arguments, to give a higher impact to his arguments, which may get a higher strength [16]. Finally, most, if not all, arguments rely on implicit structures: enthymemes, analogies, presuppositions, but also logical and pragmatic entailments. The reader therefore must participate to reconstruct the implicit data and has therefore a higher personal commitment w.r.t. the arguments. These pragmatic factors will need to be investigated in future developments of argument mining in order to elaborate accurate argumentation graphs.

3. Corpus construction and analysis

3.1. Corpus characteristics and construction

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. These

corpora are relatively small, they are designed to explore the needs in terms of knowledge, knowledge organization and reasoning schemas. The goal is to elaborate the main features of a larger empirical analysis and functional model.

For this first experiment, we considered the four following issues, which involve very different types of arguments, forms of knowledge (concrete or relatively abstract) and language realizations. These issues are:

- (1) *Ebola vaccination is necessary,*
- (2) *Women's conditions have improved in India,*
- (3) *The development of nuclear plants is necessary,*
- (4) *Organic agriculture is the future.*

Texts have been collected on the web, considering the issues as web queries. The text fragments which have been selected are extracts from various sources where these issues are discussed, in particular: newspaper articles and blogs from associations. These are documents accessible to a large public, with no professional consideration, they can therefore be understood by almost every reader. The knowledge that is involved to understand them is relatively simple and is shared by most of the readers. Language is French (glosses are given in the paper) or English. Sources are, for example:

- issue 1: The Lancet (UK journal), a French journal that develops results from the Howard Hughes Medical Institute, a web site: Ebola vaccines, therapies, and diagnostics Questions and Answers
- issue 2: http://saarthakindia.org/womens_situation_India.html, <http://www.importantindia.com/20816/women-in-india-role-and-status-of-women-in-india/>,
- issue 3: Pour/Contre le nucléaire (les centrales) web site in French, etc.

A large number of texts have been collected for each issue. For each text, the next task was to manually identify arguments related to the issue at stake. The unit considered for arguments is the sentence. These sentences are then tagged ⟨argument⟩. Since texts may be quite verbose, may include factual data of no argumentative interest, have pictures or data that cannot be taken into account in this investigation, some cleaning is necessary. In this investigation, only the paragraphs that contain a sentence identified as an argument are kept.

A preliminary task aimed at identifying all the arguments related to each issue. In a second stage, arguments judged by the annotator to be similar or redundant are eliminated and a single utterance, judged to be the most representative in terms of content and structure, is kept. The arguments eliminated on the redundancy criteria are kept separately for further analysis or for tests. So far, due to a lack of human resources, these tasks have been realized only by the author. It is clear that decisions made are somewhat subjective and that at least two annotators would have been necessary. In spite of these limitations, we consider that this work is a valid preliminary investigation whose main aim is (1) to gradually develop relevant annotation guidelines, as it is the case for most argument annotation projects, (2) to evaluate the impact of knowledge in argument mining processes and (3) to suggest new investigation directions. A more elaborated analysis must follow with the involvement of several trained annotators, with the progressive development of clear and stable annotation guidelines.

For each of these issues, the corpus characteristics and the different arguments found are summarized in Table 1. In this table, the number of words in column 2 represents the total size of the text portions that have been considered for this task, i.e. those paragraphs that contain at least one argument.

This corpus shows that the argument diversity per issue is not very large. A relatively high overlap rate has been observed: while there are original arguments, authors tend to borrow quite a lot of material

Table 1
Corpus characteristics

Issue	Corpus size	Nb of different arguments	Overlap rate (nb of similar arguments)
(1)	16 texts, 8300 words	50	4.7
(2)	10 texts, 4800 words	27	4.5
(3)	7 texts, 5800 words	31	3.3
(4)	23 texts, 6200 words	22	3.8
Total	56 texts, 25100 words	130	4.07

from each other. For example, for issue (1) an average redundancy rate of 4.7 has been observed, i.e. the same argument is found 4.7 times on average in different texts. This overlap rate is somewhat subjective since it depends on the annotator analysis and the corpus size. In spite of this subjectivity, this rate gives an interesting rough redundancy level. It is probable that with a larger corpus, this overlap rate would increase, while the number of new arguments would gradually decrease. A more detailed analysis of those repetitions would be of much interest from a rhetorical and sociological perspective.

A last step in the corpus analysis consists in tagging the discourse structures found in those sentences identified as arguments. For that purpose, the TextCoop platform we developed [17] is used with an accuracy of about 90%, since sentences are relatively simple. Discourse structures which are identified are those usually found associated with arguments: conditions, circumstances, causes, goal and purpose expressions, contrasts and concessions. The goal is to identify the kernel of the argument (tagged *<main arg>*), in general the main proposition of the sentence, and its sentential modifiers. In addition, the discourse structures may give useful indications on the argumentation strategy that is used. The result of the tagging, at this stage, is e.g.:

*<argument> <concession> Even if the vaccine seems 100% efficient and without any side effects on the tested population, </concession>
<main arg> it is necessary to wait for more conclusive data before making large vaccination campaigns </main arg>
<elaboration> The national authority of Guinea has approved the continuation of the tests on targeted populations.</elaboration> </argument>.*

3.2. Annotation parameters

The next step is to define a set of preliminary tags appropriate for analyzing the impact and the types of knowledge involved in argument mining. In this annotation task, our approach is twofold:

- (1) the standard tags are kept, these include the text span and its associated discourse structure, the polarity and the strength of the argument, and
- (2) new tags are introduced, whose goal is to clarify the need of knowledge in argument mining.

This second set of tags is preliminary and relatively global and must be revised in a later stage, when the main types of knowledge have been identified. There are therefore no guidelines yet a priori for these latter tags, since their perimeter, definition and implementation is ongoing. This is the usual methodology, essentially empirical and exploratory, that is used when dealing with a new dimension of argument mining: tags and guidelines progressively emerge from initial annotation tasks. In this paper, tags are relatively global and open so that as many details and facets as possible can be specified. Once the facets

encoded in those tags are made more clear, and the value they may take are identified, more detailed tags can be defined together with their scope and constraints.

An argument and its context (the discourse structures it is associated with, such as an illustration or a condition) are tagged between XML `<argument>` tags. The argument kernel is tagged `<main arg>`. Attributes, associated with the main tag `<argument>`, informally characterize the knowledge that is required, they include:

- the **text span involved**, which is a sentence. Arguments are numbered for referencing needs,
- the **discourse relations** associated with the argument, these are annotated using the tags defined in the TextCoop platform as described above,
- the **polarity of the argument** w.r.t. the issue which has one of the following values: support, concession (argumentative concession is a weak support), contrast (a weak attack), and attack. The concession and contrast categories have been introduced to account for cases where the support or the attack are weak. These however need to be defined more precisely (formally and in annotation guidelines). Concessions and contrasts are both discourse structures and criteria to evaluate argument polarity and strength.
- the **strength of the argument**, which is, in our view an a priori value, that is related to linguistic clues. It must be contrasted with persuasion effects that depend on the context and on the listener.
- the **conceptual relation to the issue**: informal specification why it is an attack or a support, based on the terms used in the argument or in the ‘knowledge involved’ attribute. The expressions that are used must have the form of a simple evaluative expression. Paraphrases are allowed. These expressions will be analyzed in a later stage to characterize the semantic relations involved in the conceptual relation to the issue (Section 4),
- the **knowledge involved**: list of the main concepts used. These come preferably from a predefined domain ontology, or from the annotator intuitions, if no ontology is available. This list may be quite informal. This list nevertheless contributes to characterize the nature of the knowledge involved to identify arguments.

A argument from issue (1) is tagged as follows:

```
<argument nb= 11, polarity= attack, strength= moderate,
relationToIssue= limited proofs of efficiency, limited safety of vaccination,
conceptsInvolved= efficiency measure, safety measures, evaluation methods>
<concession> Even if the vaccine seems 100% efficient and without any side effects on the tested
population, </concession>
<main arg> it is necessary to wait for more conclusive data before making large vaccination campaigns
</main arg>
<elaboration> The national authority of Guinea has approved the continuation of the tests on targeted
populations.</elaboration> </argument>.
```

This example is a rather weak attack of issue (1) since it says that more conclusive data is necessary before making large vaccination campaigns. From that point of view, this argument may be analyzed and labeled as a *concession: vaccination is necessary, however it is necessary to wait for more conclusive data.*

The annotation task is realized in several steps. The specification and the language expressions used in the two attributes ‘Conceptual relation to the issue’ and ‘Concepts involved’ are proof-read and polished several times in order to have homogeneous specifications for each issue and over the four issues

which are developed in this investigation. Concepts used are kept as minimal as possible, while the most accurate level of granularity is preserved from our intuitions. After this stage, a conceptual analysis of how arguments are related to an issue is carried out. This is developed in the next section.

4. Analysis of the types of knowledge involved in argument identification

Issues (1) to (4) involve different types of analysis which show the different facets of the knowledge needs. While issues (1) and (4) involve relatively concrete and simple concepts, issue (2) is much more abstract. It involves abstract concepts related to education, the family and human rights. Finally issue (3) involves arguments which are more fuzzy, and which are essentially comparisons with other sources of energy. This panel of issues, even if it is far from comprehensive, provides a first analysis of the types of knowledge used to identify arguments.

Dealing with knowledge remains a very vague issue in general. Knowledge can cover a large diversity of forms, from linguistic knowledge (e.g. semantic types assigned to concepts, roles played by predicate arguments), to forms involving inferences (presuppositions, implicit data), via domain and general purpose knowledge, contextual knowledge, etc. Each of these aspects of knowledge require different representation formalisms, associated inferential patterns and often involve complex acquisition procedures. This section first explores the types of knowledge involved in argument identification with respect to an issue and how it is organized. Then, a synthesis is proposed to characterize more formally the conceptual relations between the issue and the arguments. From this exploration, a relatively minimal and simple knowledge representation model emerges. The notion of concept that is used in this analysis corresponds to the notion of concept in a domain ontology, where these concepts can be either terminal (basic notions) or relational.

4.1. Main concepts in arguments related to issues (1) and (4)

In issues (1) and (4), 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. Samples of arguments associated with issue (1) are:

Supports: *vaccine protection is very good; Ebola is a dangerous disease; high contamination risks; vaccine has limited side-effects; no medical alternative to vaccine, etc.*

Attacks: *limited number of cases and deaths compared to other diseases; limited risks of contamination, ignorance of contamination forms; competent staff and P4 lab difficult to develop; vaccine toxicity and high side-effects,*

Concessions or Contrasts: *some side-effects; high production and development costs; vaccine not yet available; ethical and freedom problems.*

For issue (1), the term *vaccine* is the central concept of the issue. The concepts used to express arguments for or against that issue can be structured as follows, from this central concept *vaccine*:

- (1) concepts which are **parts** of a vaccine: the *adjuvant* and the active principle. For example, a part of the vaccine, the adjuvant, is said to be toxic for humans,
- (2) concepts associated with vaccine **super types**: a vaccine is a type of or an instance of the concept medicine; it therefore inherits the properties of ‘medicine’ (e.g. an argument says the vaccine has high side effects) unless inheritance is blocked, and

- (3) concepts that describe the **purposes, functions, goals and consequences** of a vaccine, and how it is **created**, i.e. developed, tested and sold. These concepts are the most frequently advocated in arguments for or against issue (1).

For example, the concepts of *side-effect* and *toxicity* are consequences of vaccines and more generally of a medicine, which, for a large number of them may have side-effects, besides curing a given disease. The concept of *contamination* is related to one of the purposes of a vaccine, namely to avoid that other people get the disease via contamination, and therefore on a larger scale, the purpose of a vaccine is to prevent *disease dissemination*. Finally, *production costs* are related to the creation and development of any product, including medicines and vaccines.

Without knowing that a vaccine protects humans from getting a disease, it is not possible e.g. to say that *prevents high contamination risks* is a support for issue (1) and to explain why. Similarly, without knowing that the active principle of a vaccine is diluted into an adjuvant that is also injected, it is not possible to analyze *the adjuvant is toxic* as an attack. Without this knowledge, this statement could be e.g. purely neutral or irrelevant to the issue.

The conceptual categories used in this short analysis: purpose, functions, goals, properties, creation and development, etc. are foundational aspects of the structure of a concept. They allow an accurate identification of arguments and what facet they exactly attack or support in the issue and how. This is also useful to construct various types of argument synthesis.

4.2. Main concepts in arguments related to issue (2)

This issue is much more complex and abstract than the issues (1) and (4); it is useful to evaluate the boundaries of our approach and its feasibility. The arguments related to issue (2) mainly involve comparisons with men's living conditions or refer to general and quite diverse principles of human welfare. Some arguments are justified by means of figures while others remain vague, possibly not up-to-date or relative to specific situations. Samples of arguments found for issue (2) are:

Supports: *increased percentage of literacy among women; women are allowed to enter into new professional fields; at the upper primary level, the enrollment increased from 0.5 million girls to 22.7 million girls.*

Attacks: *practices of female infanticide, poor health conditions and lack of education still persisting; home is women's real domain; they are suffering the violence afflicted on them by their own family members; malnutrition is still endemic.*

No concessions or contrasts have been observed: argument polarities are very clearcut, with a very positive or negative tonality, proper to highly controversial and overheated debates.

The concepts used in arguments related to issue (2) concentrate on facets of humans in the society and in the family and evaluates them for women. The identification of these concepts is crucial to characterize argumentative relations between issue (2) and their polarity. For example, improving literacy means higher education, better jobs and therefore more independence and social recognition, which are typical of a living condition improvement.

Roughly, the concepts used in arguments supporting or attacking issue (2) can be classified into two categories:

- those related to the services provided by the society to individuals: education, safety, health, nutrition, human rights, etc. The arguments evaluate the quality of these services for women.

- those related to the roles or functions humans can play in the society: job and economy development, family development, cultural and social involvement, etc. Arguments evaluate if and how women play these roles and functions.

Each of these concepts needs to be structured, as above for the notion of vaccine, to allow the identification of arguments and what facet of women's living conditions they exactly support or attack. The number of concepts involved may be very large, however, most arguments concentrate on a few prototypical ones which are the most striking. This is developed in Section 5.2.

4.3. Main concepts in arguments related to issue (3)

In the case of issue (3) (nuclear plants are necessary), supports or attacks mainly involve comparisons between various sources of energy. Samples of arguments are:

Supports: *allows energy independence; creates high technology jobs; risks are over-estimated; wastes are well-managed and controlled by AIEA; preserves the other natural resources,*

Attacks: *there are alternative solutions with less pollution; alternatives create more jobs than nuclear; there are risks of military uses: more dangerous than claimed; nuclear plants have high maintenance costs, etc.*

Concessions or Contrasts: *nuclear plants use dangerous products, but we know how to manage them; difficult to manage nuclear plants, but we have competent persons.*

Arguments are quite developed and illustrated; they require knowledge about energy, energy management, pollution, health, and various related dangers.

In terms of concept and concept organization, nuclear plants are complex entities. A first set of concepts concern their **creation**, i.e. construction, development and maintenance, a second set concerns their **purposes, functions, and consequences** which can be further decomposed into facets. **Parts** are also addressed, for example in the case of nuclear wastes. The same structures must be developed for other sources of energy (coal, water, wind, etc.) to analyze those arguments which include comparisons between nuclear energy and other sources of energy.

Most arguments lack precise comparative data, e.g. wastes are all said to be a nuisance, but the degree of nuisance is not given. These arguments rather play on the pathos. These are **underspecified arguments**. Finally, some arguments are fallacious, for example, the 'energetic independence' advocated as a support for nuclear plants is not acceptable since uranium is bought abroad: the use of knowledge should contribute to detect them.

4.4. Observation synthesis

From the above manual analysis, the following argument polarities are observed:

- attacks: 53 occurrences,
- supports: 33,
- argumentative concessions: 21,
- argumentative contrasts: 19 and
- undetermined: 4.

Table 2
Evidence for Knowledge

Issue	Need of knowledge nb of cases (rate)	Total number of concepts involved (rough estimate)
(1)	44 (88%)	54
(2)	21 (77%)	24
(3)	18 (58%)	19
(4)	17 (77%)	27
Total	100 (77%)	124

The corpus shows a tendency to argue against an issue: while attacks and contrasts reach 55%, supports and concessions reach only 41%. Probably unhappy persons tend to express themselves more than if they were happy.

Arguments where the need of knowledge between an issue and an argument has been characterized is summarized in Table 2. For example, for issue (1), 44 arguments for a total of 50 (88%) require knowledge to be identified as arguments. For these 44 arguments, 54 different concepts are required to establish that these statements are arguments for or against issue (1).

From the annotation schemes (Section 3) and the analysis provided in the above Sections 4.1 to 4.3, it is possible to identify the main types of knowledge which are involved. From our corpus observations, it turns out that the types of knowledge involved in relating an argument to an issue are based on the existence of lexical semantics relations between the concepts of the main terms of the issue and the argument. These relations typically include:

- **Concepts related via paradigmatic relations:** this includes in particular the concepts derived from the head terms of the issue which are either parts or more generic or more specific concepts (hyponyms or hyperonyms) of these head terms. For example, the concept of *adjuvant* is part of a vaccine, and the notion of *side effect* is a part of the generic concept *medicine*. Iteratively, parts of more generic concepts or subparts of parts are also candidates. The notion of part covers various types of parts, from physical to functional ([18]). Transitivity is only valid between parts of the same type. [18] also addresses a number of methodological problems when defining parts, hyponyms or hyperonyms which must be observed for subsequent steps of our analysis. To a lesser extent, antonyms have been observed for issues (2) and (3) (e.g. *literacy/lack of education*). These allow to develop attacks between arguments.
- **Concepts related via generic functional relations:** these relations mainly include two classes: (1) those related to the functions, uses, purposes and goals of the concepts in the issue, and (2) those related to the development, construction or creation of the entities represented by the concepts in the issue. The concepts involved in these relations are relatively straightforward for issues which are concrete, such as vaccination, but they are much more complex to specify for more abstract issues, such as issue (2).
- **Concepts related via a combination of paradigmatic and functional relations:** the relation between an issue and an argument can be for example the purpose of one of the parts or of a more generic concept of the issue head concept. For example, the notion of *industrial independence* associated with issue (3) involves in a more generic concept (a plant, an industrial sector) the functional relation of goal (independence w.r.t. e.g. to other countries). This type of combination is frequent, but does not involve too many conceptual levels to preserve a clear relatedness.

Paradigmatic and the specific functional relations presented above cover a large majority of the relations between issues and their related arguments estimated to about 80% of the total. There are obvi-

ously other types of relations, more difficult to characterize, e.g. *vaccination prevents bio-terrorism*, but of much interest for argument mining.

The figures in Table 2 show that for about 77% of the statements identified as arguments, some form of knowledge is involved to establish an argumentative relation with a controversial issue. An important result is that the number of concepts involved is not very large: 124 concepts for 100 arguments over 4 domains. Even if the notion of concept remains vague, these results are nevertheless interesting to develop large argument mining systems.

Argument behavior can be summarized as follows:

- Most arguments found directly attack or support the issue (about 75%),
- Some authors anticipate arguments from other parties by attacking the most frequent ones (about 25%),
- Groups of arguments may show forms of **incoherence** in their premises as well as in their conclusions. This may affect the strength of the arguments. For example: *with an annual increase rate of literacy of 2%, women will be shortly equal to men* versus *with an annual increase rate of literacy of 9%, women will be shortly equal to men* shows some form of incoherence in the premises for a similar conclusion.
- About 52% of the arguments involve concepts which are directly related to the head terms of the issue (e.g. adjuvant or tests for the vaccine), while about 39% require two levels of concepts, only 9% of them involve longer conceptual chains. These figures however depend on the way concepts are structured.

The above analysis shows that the introduction of knowledge in argument mining is a crucial feature which allows to improve the identification of:

- (1) the relatedness of an evaluative statement w.r.t. a controversial issue,
- (2) which aspect(s) of the issue it attacks or supports,
- (3) the argument polarity: attack, support, concession, contrast, underspecified (when knowledge is not comprehensive enough to decide on the polarity of an argument) and fallacious.

4.5. Language realizations

In terms of language realizations, the following three main types of argument realizations are observed:

- (1) use of **evaluative expressions** (46% of cases), in attribute-value form, where the attribute is proper to the concepts of the controversial issue or of closely derived concepts: *Vaccine development is very expensive, adjuvant is toxic.*
- (2) use of **comparatives** applied to related concepts (21% of the cases), e.g. *electricity produced by nuclear energy is cheaper than electricity produced from coal or wind, number of sick people much smaller than for Malaria, nuclear wastes are more dangerous than coal wastes.*
- (3) use of **facts related to the uses, consequences or purposes** of the main concept of the issue (33% of the cases), e.g.: *women are allowed to enter into new professional fields.*

Based on the notion of evaluative expression in (1) above, the following constructions are frequently observed:

- (1) if the concept is a verb with a clear polarity (e.g. *protects, prevents, pollutes*) or a formula, then language realizations include: (1a) the negation of the VP; (1b) the use of adverbs of frequency, completion, etc. possibly combined with a negation: *never, almost never, seldom, rarely, not frequently, very frequently, fully, systematically*, or (1c) the use of modals expressing doubt or uncertainty: *seem, could, should*,
- (2) if the concept is a noun, then language realizations involve attribute structures with one or more adjectives that evaluate the concept: *toxic, useless, expensive*, etc., which can be modified by intensifiers such as: *100%, totally*. Those adjectives must have a clear polarity in the context at stake.

These constructs allow to evaluate arguments related to purposes and goals such as: *the vaccine has not been fully tested; it seems to protect people from Ebola*; or related to the quality of parts of the objects in the issue: *the adjuvant is toxic*. Specific language patterns can be developed to identify these constructions as shown in [19].

5. A suitable knowledge representation model: The generative lexicon

5.1. Perspective and position

The conceptual organization described in the above analysis tends to suggest that the type of conceptual categorization offered by the Qualia structure, in particular in its last development in Generative Lexicon (GL) [20], with some extensions to the formalism, is an adequate representation framework to deal with knowledge based argument mining. The remainder of this article illustrates this proposal that needs to be evaluated in more depth when the annotation phase described above becomes more accurate and confirmed by several annotators.

The main other lexical semantics approaches such as FrameNet or VerbNet mainly concentrate on predicate argument structure for verbs and their adjuncts: they characterize the roles that these elements (NPs, PPs and S) play in the verb and proposition meaning. According to our observations, these are not central features for knowledge-based argument mining although they may be useful to develop lexical resources and argument mining templates as shown in [19].

The GL is a model that organizes concept descriptions via structures called roles. Roles describe the purposes and goals of an object or an action, its origin, and its uses. These are the main features which are supported or attacked in arguments. These, to the best of our knowledge, are specific features of the GL Qualia structure, in particular the telic and the agentive roles. However, a main limitation is however that the GL has very little resources available.

The notion of Qualia structure emerged from Aristotle categories. It has been much debated over the centuries in epistemology and philosophy. Contradictory views and definitions abound which will not be addressed in this article. Our perspective is that some features of Qualias developed in the GL seems to be appropriate to represent the type of knowledge that is crucial in knowledge-based argument mining. A number of features and uses of the Qualia (e.g. as supports of type coercion to represent logical metonymy) are not relevant for our investigations. The Qualia is therefore a convenient formalism and a set of tags that does correspond to argument mining needs.

In the remainder of this section, after a short presentation of the Qualia structure of the GL, we develop a few motivational examples, that make more concrete the proposals developed in Section 4. To conclude, some insights on the potential diversity of arguments associated to an issue are developed with an outline of a model that partly characterizes how Qualias can be used to account for the **generative**

expansion of arguments. Our experience is that while there are a few unexpected arguments, most of them can be relatively well circumvented on the basis of the concepts developed in Qualias.

5.2. An introduction to the Generative Lexicon

The Generative Lexicon (GL) [20] emerged from a series of research efforts inspired by Aristotle's notion of modes of explanation. The GL is an attempt to structure lexical semantics knowledge in conjunction with domain knowledge. It allows to explain a number of language phenomena such as various types of metonymies via a decompositional view of lexical meaning. Various forms of so-called 'generative aspects of lexical combinations' have been characterized via the operation of type shifting, where the original type that is expected has been coerced to another type, allowing metaphors such as 'to devour books' or various forms of sense variations, which are frequent in language. The GL introduces some original forms of representations such as dotted types, that account for the different facets of an entity (e.g. the physical and contents facets of a book). The GL also develops a specific argument structure with semantic types, lexical paradigms, an event structure and the Qualia structure, which is the structure that is considered in our investigations.

The Qualia structure of an entity is a kind of lexical and knowledge repository composed of four fields called roles:

- **the constitutive role** describes the various parts of the entity and its physical properties, it may include subfields such as material, parts, shape, and components,
- **the formal role** describes what distinguishes the entity from other objects, i.e. the entity in its environment, in particular the entities which are more generic.
- **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.

Roles are composed of constants or predicates a priori defined from a domain or from a general purpose ontology. A well-known example is the case of *novel(X)*:

$$\text{Novel}(X): \left[\begin{array}{l} \text{CONSTITUTIVE: } \left[\begin{array}{l} \text{PHYSICAL OBJECT: PAGES, COVER,} \\ \text{CONTENTS: CHAPTERS, STORY, CHARACTERS, ETC.} \end{array} \right], \\ \text{TELIC: } \left[\text{CONTENTS: READ}(Y,X), \text{INFORM}(Y,X), \text{ENTERTAIN}(Y,X), \text{ETC.} \right], \\ \text{FORMAL: } \left[\text{PUBLICATION, BOOK} \right], \\ \text{AGENTIVE: } \left[\begin{array}{l} \text{CONTENTS: WRITE}(T,X), \\ \text{PHYSICAL OBJECT: PRINT}(P,X), \text{PUBLISH}(P,X), \text{ETC.} \end{array} \right] \end{array} \right]$$

In this Qualia structure, X is the novel, Y, a reader, T, the author, and P, the publisher. Concepts are represented by means of predicates or constants. In the constitutive and agentive roles, two 'facets' of novel are encoded: the physical object and the contents. In the telic role, the prototypical predicate is *read*; the other predicates such as *inform*, *entertain*, develop additional purposes. In the agentive role, the creation of a book is characterized in particular by its writing (content facet), printing and publishing (physical object facet). Predicates such as 'sell' may belong to the two facets.

The formal role is close to the isa (hyponymy) relation frequently found in ontologies and in WordNet. The constitutive role has a rich informational structure since the different types of parts or constituents can be specified in dedicated sub-roles. The argument structure of the GL and its semantic typing is different from frame semantics descriptions such as FrameNet which basically develops semantic roles, not present in the GL. We show below that these two roles structure most of the knowledge that is required for a knowledge-driven approach to argument mining.

5.3. A Qualia representation of the issue head concepts

This section shows how the concepts informally presented in Sections 4.1 to 4.3 can be represented in Qualias and what extensions are required. The goal is to show how concepts can be specified in a concrete manner. This section introduces (1) the extensions to the GL that are required for argument mining and (2) the elements of a method to describe the type of data that is needed for knowledge driven argument mining. The Qualias are a relatively open representations where roles may be more or less developed and constrained: a method to describe Qualia roles is crucial to develop homogeneous and coherent sets of Qualias. Some role sub-typing have been introduced in the past in the EEC Parole-Simple project (http://catalog.elra.info/product_info.php?products_id=881&language=en).

The Qualia structure of the head term of issue (1), vaccine, describes the structure and the main properties and functions of a vaccine:

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

where X is the variable that represents the vaccine, Y is the person that is vaccinated, T is the biologist or company that develops the vaccine, Z is the doctor that makes the injection, and D is the disease associated with the vaccine (Ebola in our example). The agentive role develops the way the vaccine is created while the telic role develops its functions and purposes. An example of typing is introduced in this role: main functions (protect, avoid dissemination) and the means: how these functions are realized. This Qualia representation can be extended to develop causal and temporal chains:

$$\text{develop}(E1, T, X) \wedge \text{test}(E2, T, X) \wedge \text{sell}(E3, T, X) \wedge E1 < E2 < E3,$$

where $E1$, $E2$ and $E3$ are event-denoting variables.

This Qualia is written manually and the concepts it contains are derived from the arguments found in texts. Additional purposes or functions can be found using e.g. bootstrapping methods on the web via patterns such as ‘a vaccine allows to’. If quite a lot of data can be acquired automatically, we feel that some parts of the overall organization needs to be done manually, as in most knowledge representation systems, to guarantee coherent and well-structured representations.

Next, the Qualia structure of Ebola (and more generally, of a virus) is:

$$\text{Ebola: } \left[\begin{array}{l} \text{FORMAL: } [\text{VIRUS, DISEASE}], \\ \text{TELIC: } [\text{INFECT}(E1, \text{EBOLA}, P) \Rightarrow \text{GET_SICK}(E2, P) \Rightarrow \Diamond \text{DIE}(E3, P) \wedge E1 \leq E2 \leq E3] \end{array} \right]$$

P represents the patient that gets the disease. The purpose of Ebola is to infect people (P) who get sick and may die. There is no agentive role since there is no volition in the Ebola virus. Ebola is a constant in this representation. To characterize the sequence of events, a sequence of causal chains from ‘infect’ to ‘may die’ is introduced in the telic role.

The Qualia associated with the head concept in issue (2) is more complex and has several facets. The Qualia represents the facets of humans in the society and at home. The Qualia represents the functions, roles and purposes of humans in the society and at home. Women’s conditions is evaluated w.r.t. these features. This Qualia includes the main roles of humans as found in ontologies or in texts via mining. They are organized via several informal types such as: socio-economic, society, family, etc.:

$$\text{human}(X): \left[\begin{array}{l} \text{AGENTIVE: } \left[\begin{array}{l} \text{GET}(X, \text{FULL EDUCATION}), \text{GROW}(\text{PARENTS}, X), \\ \text{EDUCATE}(\text{PARENTS}, X), \text{GET}(X, \text{JOB}), \\ \text{GET}(X, \text{SECURITY}), \text{GET}(X, \text{RIGHTS}, \text{HEALTH}), \text{ETC.} \end{array} \right], \\ \text{TELIC: } \left[\begin{array}{l} \text{SOCIO-ECONOMIC: } \text{ENTER}(X, \text{WORK CAREER}), \\ \text{DEVELOP}(X, \text{INNOVATIONS PRODUCTS}), \text{EARN}(X, \text{MONEY}), \\ \text{CULTURAL, SCIENCE, EDUCATION: } \text{DEVELOP}(X, \text{CULTURE}) \\ \text{FAMILY: } \text{HAVE}(X, \text{FAMILY}), \text{EDUCATE}(X, \text{CHILD}), \text{CHOOSE}(X, \text{HUSBAND/WIFE}) \\ \text{SOCIETY: } \text{CONTRIBUTE_TO}(X, \text{EVOLUTION OF IDEAS}) \\ \text{PARTICIPATE}(X, \text{PUBLIC LIFE}), \text{MAINTAIN}(X, \text{SAFETY OF POPULATION}), \text{ETC.} \end{array} \right] \end{array} \right]$$

The agentive role develops the emergence of the citizen and the parent, e.g. via education. The telic describes the main roles humans play in these environments. This example, abstract and very generic, shows the limits of Qualias and also probably of most knowledge representation systems (Cyc, Sumo, etc.). However the level of granularity shown here seems to be sufficient for the purpose of argument mining.

For issue (3) the head term is the compound ‘nuclear plant’. Its Qualia can be summarized as follows:

$$\text{nuclear-plant}(X): \left[\begin{array}{l} \text{CONSTITUTIVE: } [\text{BUILDINGS, REACTOR, ETC.}, \text{MATERIALUSED: URANIUM}], \\ \text{FORMAL: } [\text{PLANT, ARTIFACT}], \\ \text{AGENTIVE: } \left[\begin{array}{l} \text{MAIN: } \text{CONSTRUCT}(E1, Y, X), \text{DEVELOP}(E2, Y, X), \text{TEST}(E3, Y, X), \\ \text{WRITE}(E4, Y, \text{DOCS}), \text{MAINTAIN}(E5, Y, X), \text{CONTROL}(E6, X, \text{AIEA}), \\ \text{CONSEQUENCE: } \text{MANAGE}(E9, Y, \text{NUCLEAR WASTES}) \end{array} \right], \\ \text{TELIC: } \left[\begin{array}{l} \text{MAIN: } \text{PRODUCE}(E7, X, \text{ELECTRICITY}), \\ \text{SIDE-EFFECTS: } \text{PRODUCE}(E8, X, \text{NUCLEAR WASTES}) \end{array} \right] \end{array} \right]$$

Table 3
Estimate number of Qualia structures to identify the arguments

Issue	Need of knowledge nb of cases (rate)	Total number of concepts involved (rough estimate)	nb of Qualias (estimate)
(1)	44 (88%)	54	9
(2)	21 (77%)	24	11
(3)	18 (58%)	19	8
(4)	17 (77%)	27	10
Total	100 (77%)	124	38

The events E_i are subject to partial temporal constraints, in particular: E1, E2, E3, E4 < E5, E9, E7, E8. There is a temporal overlap between the agentive and the telic. For plants producing electricity from coal, wind or water, the Qualia is almost identical, with the difference that uranium must be replaced by either coal, wind or water. The control organizations are also different, probably government institutions. Wastes or nuisances must also be specified. Given these conceptual similarities, the comparisons found in arguments can be related to the controversial issue.

Describing Qualias raises several feasibility problems since there is little data available. The problem of an automatic or semi-automatic acquisition of Qualias is addressed in e.g. [21], and to a lesser extent in [19]. While it is possible to encode Qualias manually, we feel that a bootstrapping method from corpora should enhance the process by providing a lot of data, which, however, needs to be organized by hand in a second stage. An interesting point is that, for a given issue, the knowledge that is required is not so large, as shown in Table 2. It seems that for an issue with a large scope, a maximum of 20 Qualias are needed to deal with most of the arguments. For the issues presented in this article, the number of required Qualias is given in Table 3, for the arguments that have been identified.

The number of Qualias for an issue in the corpus presented in Section 3 is on average 10. Some Qualias may be quite large while others are relatively sketchy. Some Qualias may be close to or inherit from others, e.g. vaccine and medicine. The number of Qualias may slightly change depending on how knowledge is structured, but the variation is quite limited and the number of concepts to include in the Qualias remains approximately constant. Table 3 shows that an average of 4 concepts are used per Qualia in our corpus.

5.4. A network of Qualias to characterize the generative expansion of arguments

Our observations show that arguments attack or support (1) specific concepts found in the Qualia of the head terms in the controversial issue (called **root concepts**) or (2) concepts derived from these root concepts, via their Qualia. For example, arguments can attack properties or purposes of the adjuvant or of the protocols used to test the vaccine. Then, a network of Qualias can be defined to develop the knowledge structures involved in the recognition of potential arguments [19]. This network is limited to three levels because derived concepts must remain conceptually and functionally close to the root concepts. The objective is to evaluate 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, probably with a lower frequency and recurrence.

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 [13]. Natural language words or expressions that lexicalize each concepts can be associated with each network nodes.

6. Conclusion

In this contribution, we have presented a preliminary analysis of the different forms of knowledge which are frequently required to determine whether an utterance is an argument for or against a given controversial issue. The corpus analysis shows that more than 75% of the arguments that have been manually identified require some form of knowledge and inference to be accurately related to a controversial issue. A conceptual analysis shows that the Qualia structure of the GL is a useful knowledge and lexical representation approach that provides a conceptual organization which makes explicit the main topics used in arguments.

This analysis is exploratory and must be expanded in various directions before any development of a knowledge-driven argument mining system. A more detailed empirical analysis is necessary, with a larger corpus, several annotators, and annotation guidelines before concluding that the Qualia structure is the main resource that is needed and finalizing the extensions which are required. Similarly, the reasoning schemes which are proper to argumentation must be explored and modeled so that the knowledge described in the Qualias and the reasoning schemes can accurately be used to mine arguments.

References

- [1] 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](https://doi.org/10.4018/jcini.2013010101).
- [2] J. Redmond, *Logic, Argumentation and Reasoning, Interdisciplinary Perspectives from the Humanities and Social Sciences*, Springer, 2017.
- [3] R.M. Palau and M.-F. Moens, Argumentation mining: The detection, classification and structure of arguments in text, in: *Proceedings of 12 ICAIL*, 2009.
- [4] C. Kirschner, J. Eckle-Kohler and I. Gurevych, Linking the thoughts: Analysis of argumentation structures in scientific publications, in: *Proceedings of 2nd Workshop on Argumentation Mining, Denver*, 2015.
- [5] M.G. Vilalba and P. Saint-Dizier, Some facets of argument mining for opinion analysis, in: *Proceedings of COMMA12, IOS Publishing*, 2012.
- [6] M. Janier and C. Reed, Towards a Theory of Close Analysis for Dispute Mediation Discourse, *Journal of Argumentation* 22 (2015).
- [7] 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.
- [8] R. Swanson, B. Ecker and M. Walker, Argument mining: Extracting arguments from online dialogue, in: *Proceedings of SIGDIAL15*, 2015.
- [9] 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.
- [10] V.W. Feng and G. Hirst, Classifying arguments by scheme, in: *Proceedings of 49th ACL: Human Language Technologies*, 2011.
- [11] A. Fiedler and H. Horacek, Argumentation within deductive reasoning, *Journal of Intelligent Systems* 22(1) (2007), 49–70.
- [12] G. Winterstein, What but-sentences argue for: An argumentative analysis of ‘but’, *Lingua* 122 (2012).
- [13] P.S. Dizier, Challenges of argument mining: Generating an argument synthesis based on the qualia structure, in: *Proceedings of INLG16*, 2016.
- [14] C. Perelman and L.O. Tyteca, *The New Rhetoric: A Treatise on Argumentation*, University of Notre Dame Press, 1977.
- [15] D. Walton, *Goal-Based Reasoning for Argumentation*, Cambridge University Press, 2015. doi:[10.1017/CBO9781316340554](https://doi.org/10.1017/CBO9781316340554).
- [16] R. Amossy, *L’argumentation dans Le Discours*, Armand Colin, 2010.

- [17] P. Saint-Dizier, Processing natural language arguments with the TextCoop platform, *Journal of Argumentation and Computation* **3**(1) (2012), 49–82.
- [18] A. Cruse, *Lexical Semantics*, Cambridge University Press, 1986.
- [19] P.S. Dizier, The bottleneck of knowledge and language resources, in: *Proceedings of LREC16*, 2016.
- [20] J. Pustejovsky, *The Generative Lexicon*, MIT Press, 1986.
- [21] V. Claveau and P. Sebillot, Automatic acquisition of GL resources, using an explanatory, symbolic technique, in: *Proceedings of Advances in Generative Lexicon Theory*, 2013.