Identifying Consumers' Arguments in Text

Jodi Schneider[1] and Adam Wyner[2]

[1] Digital Enterprise Research Institute, National University of Ireland
jodi.schneider@deri.org
[2] Department of Computer Science, University of Liverpool, Liverpool, UK
adam@wyner.info

Abstract. Product reviews are a corpus of textual data on consumer opinions. While reviews can be sorted by rating, there is limited support to search in the corpus for statements about particular topics, e.g. properties of a product. Moreover, where opinions are justified or criticised, statements in the corpus indicate arguments and counterarguments. Explicitly structuring these statements into arguments could help better understand customers' disposition towards a product. We present a semi-automated, rule-based information extraction tool to support the identification of statements and arguments in a corpus, using: argumentation schemes; user, domain, and sentiment terminology; and discourse indicators.

Keywords: argumentation schemes, information extraction, product reviews

1 Introduction

Product reviews such as found on Amazon or eBay represent a source of data on consumer opinions about products. Current online tools allow reviews to be sorted by star rating and comment threads. Yet, there is no support to search through the data for statements about particular topics, e.g. properties of the product. Such statements are distributed throughout the corpus, making it difficult to gain a coherent view. Moreover, reviewers justify their opinions as well as support or criticise the opinions of others; that is, reviewers provide arguments and counterarguments. Extracting data about particular topics and structuring it into arguments would be informative: it could help producers better understand consumers' disposition to their products; and, it could help consumers make sense of the product options, as reported in the reviews, and so then decide what to buy. We present an information extraction tool to support the extraction of arguments from reviews, using: user, domain, and sentiment terminology; discourse indicators; and argumentation schemes.

To set the context, consider the information in the reviews from the point of view of a consumer or manufacturer. We use product reviews from the Amazon consumer web site about buying a camera as a use case. From the consumer side, suppose a photo enthusiast wants to buy a new camera that gives quality indoor pictures. The enthusiast consults a shopping website and reads the reviews of camera models. The information in product reviews about this topic is dispersed through a number of reviews, using different terminology, and expressing opinions on different sides of the situation. So, currently, the enthusiast must read through the reviews, keeping in mind the relevant

Jodi Schneider and Adam Wyner. "Identifying Consumers' Arguments in Text" In SWAIE 2012: Semantic Web and Information Extraction. In conjunction with EKAW 2012. Galway, Ireland. October 9, 2012.

statements, organising and relating them; this is a difficult task. Instead, the enthusiast would like all statements bearing on the camera's indoor picture quality to be reported and sorted according to whether the statement supports the claim that the camera gives quality indoor pictures or supports the claim that it does not. Moreover, it is not sufficient for the enthusiast to be provided with one 'layer' of the argument, since those statements which support or criticise the claim may themselves be subject to support or criticism. From the manufacturer's side, there is a related problem since she wishes to sell a product to a consumer. Looking at the reviews, the manufacturer must also extract information about specific topics from the corpus and structure the information into a web of claims and counterclaims. With this information, the manufacturer could have feedback about the features that the consumer does or doesn't like, the problems that the consumer experiences, as well as the proposed solutions.

There are a variety of complex issues to address. For instance, to overcome the linearity of the corpus and terminological variation, we want a tool that searches and extracts information from across the corpus using semantic annotations, allowing us to find statements about the same semantic topic; searches for strings do not suffice since the same semantic notion might be expressed with different strings. Sentiment identifiers, which signal approval or disapproval, are relevant. Discourse markers indicate relationships between statements, e.g. premise or claim. In addition, users argue from a *point of view:* different user classes, e.g. amateurs and professionals, argue differently about the same object.

While a fully automated system to reliably extract and structure all such information is yet in the future, we propose a semi-automated, rule-based text analytic support tool. We first manually analyse the corpus, identifying the sorts of semantic information to be annotated. We develop reasoning patterns, *argumentation schemes*, and identify slots in these schemes to be filled. The schemes represent different aspects of how users reason about a decision to buy a product. We structure the schemes into a decision tree, hypothesising a main scheme which is used to argue for buying the product. This main scheme is supported by subsidiary schemes that argue for premises of the main scheme. In turn, the subsidiary schemes are grounded in textual information related to the user and the representation of the product. In effect, we reverse engineer an argumentative expert system which takes as input material from the corpus. Thus, the schemes give us *targets* for information extraction in the corpus, namely, those components that can be used to instantiate the argumentation schemes. The information extraction tool supports the identification of relevant information to instantiate the argumentation schemes. As a result of the analysis and instantiation, we gain a rich view on the arguments for or against a particular decision. The novelty is that the tool systematically draws the analyst's attention to relevant terminological elements in the text that can be used to ground defeasible argumentation schemes.

The outline of the paper is as follows. In Section 2, we discuss our use case and materials. Several components of the analysis are presented in Section 3: user, domain, and sentiment terminology; and discourse indicators. The argumentation schemes that we propose to use are given in Section 4. The tool is outlined in Section 5, followed by sample results in Section 6. Related work is discussed in Section 7, and we conclude in Section 8 with some general observations and future work.

2 Use Case and Materials

As a use case, we take reviews about buying the (arbitrarily chosen) Canon PowerShot SX220 HS Digital Camera from the Amazon UK e-commerce website[3], where a very typical question is: *Which camera should I buy?* There are 99 reviews in our corpus, distributed as shown in Table 1.

Table 1. Review distribution by star rating

5-star	54
4-star	27
3-star	9
2-star	8
1-star	1

In these product reviews, many topics are discussed. By careful reading and analysis, we find comments about cameras such as their features and functions. Further, accessories, such as memory cards, batteries, and cases are also discussed – both with regard to their necessity or utility, and their suitability. The brand reputation and warranty are discussed. Users also give conditions of use – recommendations for who the camera would or would not suit, and warnings and advice about how to get the best results. These incorporate the purpose or context in which the camera is or could be used (e.g. "traveling") or values that the camera fulfils (e.g."portable"). Users also give clues to their own experience and values, by talking about how they evaluated the camera, their experience with photography, or personal characteristics (e.g. "ditzy blonde").

Point of view is key to making sense of the overall discussion. For subjective aspects, the impact of a statement may depend on the extent to which consumers share values and viewpoints. Such qualitative aspects of the reviews are not captured by quantitative measures of the discussion since the most popular comment may not advance the analysis with respect to that user or may only sway individuals who are susceptible to popular opinion. Given this, we focus on representing justifications and disagreements with respect to classes of users.

In the course of our manual examination of the corpus, we identified five "components" of an analysis: several consumer argumentation schemes; a set of discourse indicators, and user, domain, and sentiment terminology. The user and domain terminology are used to instantiate the schemes, while the discourse indicators and sentiment terminology structure the interrelationships between the statements within a scheme (e.g. premises, claim, and exception) and between schemes (e.g. disagreement). We begin by discussing the last four components, then turn to argumentation schemes in Section 4.

3 Components of Analysis

The objective of information extraction in our context is to extract statements about a topic (e.g. a camera takes good pictures indoors) and structure them into arguments for (e.g. justifications for this claim) or against it (counterclaims and their justifications).

[3] Accessed 2012-07-22 `https://www.amazon.co.uk/product-reviews/B004M8S152/re`

In the following, we briefly outline the components of our analysis, which are implemented in the tool discussed in Section 5. In our approach, we identify a *terminological pool* that helps us investigate the source text material for relevant passages; thus, we presume that we can search throughout the corpus to instantiate an argumentation scheme using the designated terminology.

In our approach to analysis of the source material, we have presumed that in the context of product reviews, contributors are trying to be as helpful, informative, and straightforward as possible, so the interpretation of language is at *face value*. In other contexts, problematic, interpretive aspects of subjectivity may arise, e.g. irony or sarcasm, which require significant auxiliary, extra-textual knowledge to accurately understand. For our purposes, we do not see irony or sarcasm as a significant problem as we can rely on the normative reading of the text that is shared amongst all readers.

Camera Domain We have terminology from the camera domain that specifies the objects and properties that are relevant to the users. Analysing the corpus, consumer report magazines (e.g. *Which?*), and a camera ontology[4], we identified some of the prominent terminology. These refer both to parts of the camera (e.g. lens, li-ion battery) as well as its properties (e.g. shutter speed). While users may dispute particular factual matters about a camera, these remain objective aspects about the camera under discussion.

User Domain Users discuss topics relative to their point of view, knowledge, and experience. This introduces a *subjective aspect* to the comments. For instance, whether an amateur finds that that a particular model of camera takes *very poor* pictures indoors may not agree with an expert who finds that the same model takes *good* pictures indoors; each is evaluating the quality of the resulting pictures relative to their own parameter of quality and experience with camera settings. To allow such user-relative judgements, we introduce user terminology bearing on a user's attributes (e.g. age), context of use (e.g. travel), desired camera features (e.g. weight), quality expectations (e.g. information density), and social values (e.g. prestige).

Discourse Indicators Discourse indicators express discourse relations within or between statements [1] and help to organise statements into larger scale textual units such as an argument. The analysis of discourse indicators and relations is complex: there many classes of indicators, multiple senses for instances of indicators depending on context, and implicit discourse relations. However, in this study, we keep to a closed class of explicit indicators that signal potentially relevant passages; it remains for the analyst to resolve ambiguities in context.

Sentiment Terminology We use *sentiment* terminology that signals lexical semantic contrast: *The flash worked poorly* is the semantic negation of *The flash worked flawlessly*, where *poorly* is a negative sentiment and *flawlessly* is a positive sentiment. An extensive list of terms is classified according to a sentiment scale from highly negative to highly positive [2]. Text analytic approaches to sentiment analysis are well-developed, but for our purposes we take this relatively simple model to integrate with other components.

In the following, we provide argumentation schemes that use the camera and user terminology. The discourse indicators and sentiment terminology are only used in the tool to identify relevant passages to instantiate the schemes.

[4] http://www.co-ode.org/ontologies/photography/

4 Argumentation schemes

Argumentation schemes represent prototypical patterns of defeasible reasoning [3]. They are like logical syllogisms in that they have premises, an implicational rule (e.g. *If....Then....*), and a conclusion that follows from the premises and rule. Moreover, they can be linked as in proof trees. Yet, unlike classical syllogisms, the conclusion only defeasibly follows since the rule or the conclusion may not hold. Argumentation schemes have been formalised [4] and can be used for abstract argumentation [5]. Example schemes include *practical reasoning*, *expert opinion*, and *analogy*. However, schemes are not widely used to support text analysis, are not tied to user terminology, and not usually tied to some particular domain. This paper makes progress in addressing these issues. In this section we develop a number of argument schemes found in customer reviews, based on manual review of the corpus. Our approach is to remain grounded in the source, and to choose example schemes based on their relevance to arguing for or against purchase of the product. In this way, the schemes give us *targets* for information extraction in the corpus: in particular, the targets are those textual passages that can be used to instantiate the argumentation schemes.

4.1 Argumentation Schemes - Abstract

We present the schemes propositions with variables such as aP_1; the list of premises is understood to hold conjunctively and the conclusion follows; the rule is left implicit.

User Classification With this scheme, we reason from various attributions to a user to the class of the user. This scheme is tied to the particular data under consideration, but could be generalised. We have a variety of users such as amateur or professional.

User Classification Argumentation Scheme (AS1)

1. *Premise*: Agent *x* has user's attributes $aP_1, aP_2, ...$
2. *Premise*: Agent *x* has user's context of use $aU_1, aU_2, ...$
3. *Premise*: Agent *x* has user's desirable camera features $aF_1, aF_2, ...$
4. *Premise*: Agent *x* has user's quality expectations $aQ_1, aQ_2, ...$
5. *Premise*: Agent *x* has user's values $aV_1, aV_2, ...$
6. *Premise*: User's desirable camera features $aF_1, aF_2, ...$ promote/demote user's values $aV_1, aV_2, ...$
   Conclusion: Agent *x* is in class *X*.

Camera Classification We have a scheme for classifying the camera. Note that we have distinguished a user's context of use from a camera's context of use (and similarly for other aspects); in a subsequent scheme (AS3), these are correlated.

Camera Classification Argumentation Scheme (AS2)

1. *Premise*: Camera *y* has camera's context of use $cU_1, cU_2, ...$
2. *Premise*: Camera *y* has camera's available features $cF_1, cF_2, ...$
3. *Premise*: Camera *y* has camera's quality expectations $cQ_1, cQ_2, ...$
   Conclusion: Camera *y* in class *Y*.

Combining Schemes for Camera Evaluation To reason about the camera and the course of action, we use some ontological reasoning, i.e. the class of the camera and of the user, plus argumentation. Given that a user is in class X with certain requirements and a camera is in class Y with certain features, and the features meet the requirements, then that camera is appropriate. The argument that conjoins the user and camera schemes works as a filter on the space of possible cameras that are relevant to the user. We realise this as follows.

Appropriateness Argumentation Scheme (AS3)

1. *Premise*: Agent *x* is in user class *X*.
2. *Premise*: Camera *y* is in camera class *Y*.
3. *Premise*: The camera's contexts of use satisfy the user's context of use.
4. *Premise*: The camera's available features satisfy the user's desireable features.
5. *Premise*: The camera's quality expectations satisfy the user's quality expectations.

Conclusion: Cameras of class *Y* are appropriate for agents of class *X*.

Premises (1) and (2) of the appropriateness scheme (AS3) are the conclusions of the user (AS1) and consumer (AS2) classification schemes, respectively. The other premises (3)-(5) have to be determined by subsidiary arguments which nonetheless ground variables in the same way (in Logic Programming terms, the variables are *unified*). Each of these subsidiary schemes have a similar form, where premises correlate elements from AS1 and AS2 and conclude with one of the premises of (3)-(5). The redundancy ensures that the variables match across schemes. We leave such *intermediary schemes* as an exercise for the reader.

Practical Reasoning The objective of reasoning in this case is for the user to decide what camera to buy. The reasoning is based on the user and the camera. This information is then tied to the decision to buy the camera. Since reasoning about the camera relative to the user is addressed elsewhere in the reasoning process, our scheme (AS4) is a simplification of [6].

Consumer Relativised Argumentation Scheme (AS4)

1. *Premise*: Cameras of class *Y* are appropriate for agents of class *X*.
2. *Premise*: Camera y is of class *Y*.
3. *Premise*: Agent x is of class *X*.

Conclusion: Agent x should buy camera y.

The important point is that if the class of the camera and user do not *align*, or if there are counterarguments to any of the premises or conclusions, then the conclusion from AS4 would not hold.

5 Components of the Tool

To build an analytic tool to explore and extract arguments, we operationalise the components needed to recognise in the text some of the relevant elements identified in Section 4. In this section, we briefly describe the relevant aspects of the General Architecture for Text Engineering (GATE) framework [7] and samples of how we operationalise the components. In Section 6, we show results of sample queries.

5.1 GATE

GATE is a framework for language engineering applications, which supports efficient and robust text processing [7]; it is an open source desktop application written in Java that provides a user interface for professional linguists and text engineers to bring together a wide variety of natural language processing tools and apply them to a set of documents. The tools are formed into a pipeline (a sequence of processes) such as a sentence splitter, tokeniser, part-of-speech tagger, morphological analyser, gazetteers, Java Annotation Patterns Engine (JAPE) rules, among other processing components. For our purposes, the important elements of the tool to emphasise are the gazetteers and JAPE rules: a gazetteer is a list of words that are associated with a central concept; JAPE rules are transductions that take annotations and regular expressions as input and produce annotations as output. Our methodology in using GATE is described elsewhere [8], and in this paper, we focus just on the key relevant elements - the gazetteers and JAPE rules.

Once a GATE pipeline has been applied to a corpus, we can either view the annotations of a text by using the ANNIC (ANNotations In Context) corpus indexing and querying tool [9] or view them *in situ* in a whole text. We illustrate both.

5.2 Gazetteers and JAPE Rules

In section 4, we presented terminology for discourse indicators and the camera domain. The terminology is input to text files such as *cameraFeatures.lst* for terms relating to the camera domain and *conclusion.lst* for terms that may indicate conclusions. The lists are used by a *gazetteer* that associates the terms with a *majorType* such as *cameraproperty* or *conclusion*. JAPE rules convert these to annotations that can be visualised and used in search. For example, suppose a text has a token term “lens” and GATE has a gazetteer list with “lens” on it; GATE finds the string on the list, then annotates the token with majorType as *cameraproperty*; we convert this into an annotation that can be visualised or searched for such as *CameraProperty*. A range of terms that may indicate conclusions are all annotated with *Conclusion*. We can also create annotations for complex concepts out of lower level annotations. In this way, the gazetteer provides a *cover concept* for related terms that can be queried or used by subsequent annotation processes.

In the implementation, we have gazetteer lists for camera domain terminology and for user domain terminology, one list each for conclusions, premises, and contrast, and a range of sentiment terminology lists. Samples of the lists (with number of items) are:

- conclusion.lst (26): be clear, consequent, consequently, deduce, deduction, ...
- cameraFeatures.lst (130): 14X Optical Zoom, action shots, AF tracking, ...
- posThree.lst (172): astound, best, excellent, splendid, ...
- userContextOfUse (32): adventure, ambient light indoors, astronomy photos, ...

In the next section, we show sample results.

6 Sample Results

To identify passages that can be used to instantiate the argumentation schemes, we use ANNIC searches to investigate the entire corpus. Figure 1 shows a result of a search for negative sentiment, followed by up to 5 tokens, followed by a user context; the search returns six different strings that match the annotation pattern.

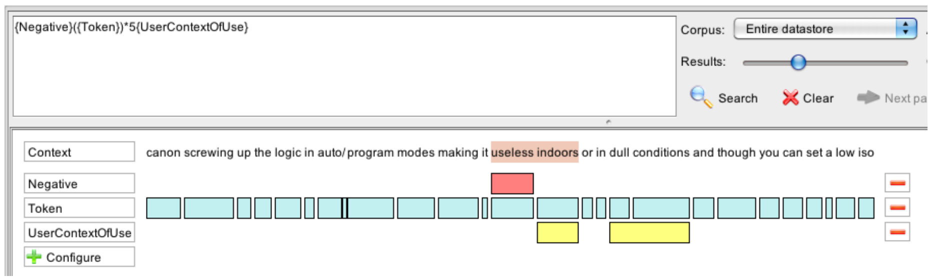

Fig. 1. Sample output from an ANNIC search.

We can also look at annotations *in situ* in a text. Figure 2 shows one review document, with a variety of annotation types, where different highlights indicate different annotation types (differentiated with colour in the original); from this review we extract instantiations for the user and camera schemes. This passage makes the argument that the camera is not appropriate since the user's context of use – baby pictures – does not match the camera context of use. In other words, we use the annotations to instantiate the two schemes below.

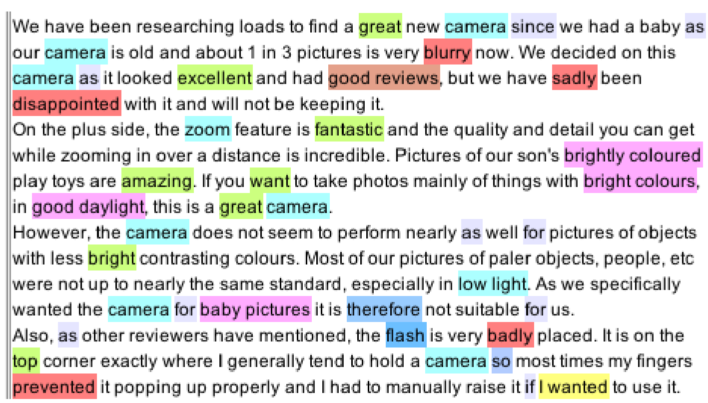
We have been researching loads to find a great new camera since we had a baby as our camera is old and about 1 in 3 pictures is very blurry now. We decided on this camera as it looked excellent and had good reviews, but we have sadly been disappointed with it and will not be keeping it.
On the plus side, the zoom feature is fantastic and the quality and detail you can get while zooming in over a distance is incredible. Pictures of our son's brightly coloured play toys are amazing. If you want to take photos mainly of things with bright colours, in good daylight, this is a great camera.
However, the camera does not seem to perform nearly as well for pictures of objects with less bright contrasting colours. Most of our pictures of paler objects, people, etc were not up to nearly the same standard, especially in low light. As we specifically wanted the camera for baby pictures it is therefore not suitable for us.
Also, as other reviewers have mentioned, the flash is very badly placed. It is on the top corner exactly where I generally tend to hold a camera so most times my fingers prevented it popping up properly and I had to manually raise it if I wanted to use it.

Fig. 2. An annotated review.

User Classification Argumentation Scheme - Baby Picture Reviewer

1. *Premise*: Agent *x* has user attributes: *little experience*.
2. *Premise*: Agent *x* has constraints: *single camera*.
3. *Premise*: Agent *x* has context of use *portrait*.
4. *Premise*: Agent *x* has user's desirable camera features *easy to hold, flash doesn't require user attention, zoom*.
5. *Premise*: Agent *x* has quality expectations *good pictures of pale objects, good pictures of objects that don't have contrast*.
6. *Premise*: Agent *x* has values *good reviews, photo quality, photo detail*.
7. *Conclusion*: Agent *x* is in class *Novice*.

Camera Classification Argumentation Scheme (AS2) - Baby Picture Reviewer

1. *Premise*: Camera *y* has camera's context of use *daylight*.
2. *Premise*: Camera *y* has camera's available features *zoom, flash*.
3. *Premise*: Camera *y* has camera's quality expectations *annoying flash, amazing for bright colours, poor when colours do not contrast (people, pale objects), good quality with zoom, good detail with zoom*.

   Conclusion: Camera *Canon PowerShot SX220* in class *daylight, contrast-oriented, zoom camera*.

One argument against the above camera classification is given by another reviewer: "This camera takes amazing low light photos...". Based on the full text of that review, we can instantiate the camera classification argumentation scheme differently, as follows:

Camera Classification Argumentation Scheme (AS2) - Great low light

1. *Premise*: Camera *y* has camera's context of use *video, photos*.
2. *Premise*: Camera *y* has camera's available features *HD video recording, screen, zoom, flash, colour settings*.
3. *Premise*: Camera *y* has camera's quality expectations *lens shadow, awkward flash location, vibrant colours*.

   Conclusion: Camera *Canon PowerShot SX220* in class *video, general photo camera*.

This shows some advantages of argumentation schemes. First of all, they can help an analyst make explicit the points of contention between reviews. The reviews disagree on the camera's quality expectations: this particular disagreement could not easily be discovered statistically from the text. Second, we can separate out different levels of subjective information to be found in the reviews. The user classification scheme separates the purely subjective information that cannot be attacked from the camera classification scheme, which can be fruitfully attacked. Further, by classifying cameras and users, an entire line of reasoning follows: we only need to instantiate those two schemes.

Some issues do arise, and will need to be considered in future work. First, we cannot always instantiate some premises. For example, users may not indicate user attributes or constraints in a review. In that situation, presumptive values could be used, or found elsewhere in the corpus. Second, there are other arguments and counterarguments that are made. For instance, some reviews suggest ways of dealing with the popup flash so that it's not annoying, making the camera more comfortable to use indoors. To handle more types of arguments and counterarguments, we will want to develop further argumentation schemes. Some negative implications depend on a deeper analysis of the camera domain, for instance: "You need to learn all functions in order to shoot really good photos." or "People look either washed out or with a flat looking red/orange complexion." Other arguments, such as arguments from expertise, are common, and should be analysed further to provide support for information extraction.

7 Related Work and Discussion

In this section, we outline related work, which includes opinion and review mining, user preferences, and ontological approaches, and use of argumentation. What makes our proposal novel and unique is the combination of *rule-based* text analytics, user models, and defeasible argumentation schemes, which together highly structure the representation of information from the source materials. In previous work we have introduced argumentation schemes for understanding evaluative statements in reviews as arguments from a point of view [18]. Our earlier, preliminary implementation, used a single argumentation scheme [6]; this paper extends that work by implementing user terminology and increasing the specification of camera terminology, and by using a cascade of argumentation schemes, where the conclusions of two schemes are the premises of the appropriateness scheme.

Opinion and Review Mining Existing work includes review mining – information extraction using sentiment terminology [10] – and feature extraction of pros and cons [19]. Matching customers to the most appropriate product based on the heterogeneity of customer reviews, rather than just statistical summaries, is an important problem; Zhang et al. develop sentiment divergence metrics, finding that the central tendency or polarity of reviews is insufficient [20]. Our goals, in matching customers to products by distinguishing views based on a customer profile, are similar; unlike that study, we focus on textual analysis, rather than statistical summarization of the text.

User Preferences Case-based reasoning has been used to incorporate critique-based feedback and preference-based feedback into recommendation systems. [21]. To predict ratings in Chinese-language restaurant reviews, Liu et al. model how frequently users comment on features ('concern') and how frequently they rate features lower than average ('requirement') in order to predict ratings [22]. Rather than inferring user preferences from multiple reviews written by a user, we extract user information from a single review; although some personal information (such as the user demographics) is consistent across items in different departments (such as books, movies, consumer electronics, clothing, etc.), the key information about the user is that related to the product, which depends on the category, and in some cases on the item being purchased. For instance, preferences about an item having a flash or a viewfinder are not universal amongst consumer electronics, but apply mainly to cameras.

Ontology-related approaches Yu et al. automatically construct a hierarchical organization for aspects from product reviews and domain knowledge [23]. This approach could be used to further enhance our extraction systems, and there are available tools in GATE to support this: OwlExporter is a GATE plugin for populating OWL ontologies via NLP pipelines [24]; KIM uses an ontology and knowledge base to add semantic annotations based on information extraction [25].

Argumentation Argumentation schemes have been used as a theoretical framework for reviews [26]. Another closely related problem is argumentation mining – using natural language processing to detect disagreement [11,12,13] or stance [14,15].

8 Conclusions and Future Work

We have presented an information extraction tool that supports the identification of relevant information to instantiate argumentation schemes, by annotating discourse indicators as well as user, domain, and sentiment terminology. Textual fragments are associated with annotation types, highlighting the role the text may play in instantiating an argumentation scheme. As we can identify positive and negative sentiment, we can find statements that contribute to arguments for or against other statements. The novelty of our proposal is the combination of *rule-based* text analytics, terminology for various particular components of the analysis, and defeasible argumentation schemes, which together highly structure the representation of information from the source materials. As a result of the analysis and instantiation, we can provide a rich, articulated analysis of the arguments for or against a particular decision.

In future work, we plan to further instantiate the schemes using the tool, noting where they work as intended and where they stand to be improved. Along with this, conceptual issues will be addressed, for instance to clarify distinctions between the camera's quality expectations and features as well as to support matches between a user's values and camera properties. We will develop additional schemes bearing on, for example, expertise, comparison, or particular features (e.g. warranties). An evaluation exercise will be carried out using a web-based annotation editor and evaluation tool, GATE Teamware, to measure the extent of interannotator agreement on the annotation types. Important *logical* developments would be an ontology for users and cameras that would support text extraction and import of scheme instances into an argumentation inference engine to test inferences.

Acknowledgements

The first author's work was supported by Science Foundation Ireland under both Grant No. SFI/09/CE/I1380 (Líon2) and a Short-term Travel Fellowship. The second author gratefully acknowledges support by the FP7-ICT-2009-4 Programme, IMPACT Project, Grant Agreement Number 247228. The views expressed are those of the authors.

References

1. Webber, B., Egg, M., Kordoni, V.: Discourse structure and language technology. Natural Language Engineering (December 2011) Online first.
2. Nielsen, F.Å.: A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. Making Sense of Microposts at ESWC 2011 (2011)
3. Walton, D., Reed, C., Macagno, F.: Argumentation Schemes. Cambridge Univ. Press (2008)
4. Wyner, A., Atkinson, K., Bench-Capon, T.: A functional perspective on argumentation schemes. In Proceedings of the 9th International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2012). (2012) 203–222
5. Prakken, H.: An abstract framework for argumentation with structured arguments. Argument and Computation **1**(2) (2010) 93–124
6. Wyner, A., Schneider, J., Atkinson, K., Bench-Capon, T.: Semi-automated argumentative analysis of online product reviews. In: Proceedings of the 4th International Conference on Computational Models of Argument (COMMA 2012). (2012)

7. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V.: GATE: A framework and graphical development environment for robust NLP tools and applications. In: Proceedings of the Association for Computational Linguistics (ACL'02). (2002) 168–175
8. Wyner, A., Peters, W.: On rule extraction from regulations. In Legal Knowledge and Information Systems - JURIX 2011, IOS Press (2011) 113–122
9. Aswani, N., Tablan, V., Bontcheva, K., Cunningham, H.: Indexing and querying linguistic metadata and document content. In: Proceedings of 5th International Conference on Recent Advances in Natural Language Processing, Borovets, Bulgaria (2005)
10. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1-2) (January 2008) 1–135
11. Albert, C., Amgoud, L., de Saint-Cyr, F.D., Saint-Dizier, P., Costedoat, C.: Introducing argumention in opinion analysis: Language and reasoning challenges. In: Sentiment Analysis where AI meets Psychology (SAAIP at IJCNLP'11). (2011)
12. Saint-Dizier, P.: Processing natural language arguments with the <TextCoop>platform. Argument & Computation **3**(1) (2012) 49–82
13. Wyner, A., Mochales-Palau, R., Moens, M.F., Milward, D.: Approaches to text mining arguments from legal cases. In Semantic Processing of Legal Texts. Volume 6036 of Lecture Notes in Computer Science. Springer (2010) 60–79
14. Abbott, R., Walker, M., Anand, P., Tree, J.E.F., Bowmani, R., King, J.: How can you say such things?!?: Recognizing disagreement in informal political argument. In: Proceedings of the NAACL HLT 2011 (2011)
15. Walker, M.A., Anand, P., Abbott, R., Tree, J.E.F., Martell, C., King, J.: That's your evidence?: Classifying stance in online political debate. Decision Support Systems (2012)
16. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. Text-Interdisciplinary Journal for the Study of Discourse **8**(3) (1988) 243–281
17. Somasundaran, S., Wiebe, J.: Recognizing stances in ideological on-line debates. In: Proceedings of the NAACL HLT 2010, (2010) 116–124
18. Wyner, A., Schneider, J.: Arguing from a point of view. In: First International Conference on Agreement Technologies. AT '12 (2012)
19. Liu, B. "Opinion Mining and Sentiment Analysis." In: Web Data Mining. Springer, (2011) 459–526
20. Zhang, Z., Li, X., Chen, Y.: Deciphering word-of-mouth in social media: Text-based metrics of consumer reviews. ACM Trans. Manage. Inf. Syst. **3**(1) (April 2012) 5:1–5:23
21. Smyth, B.: Case-based recommendation. The Adaptive Web. Volume 4321 of Lecture Notes in Computer Science. Springer (2007) 342–376
22. Liu, H., He, J., Wang, T., Song, W., Du, X.: Combining user preferences and user opinions for accurate recommendation. Electronic Commerce Research and Applications (2012)
23. Yu, J., Zha, Z.J.J., Wang, M., Wang, K., Chua, T.S.S.: Domain-assisted product aspect hierarchy generation: towards hierarchical organization of unstructured consumer reviews. In: Proceedings of EMNLP '11 (2011) 140–150
24. Witte, R., Khamis, N., Rilling, J.: Flexible ontology population from text: The OWL exporter. In: LREC 2010. (2010) 3845–3850
25. Popov, B., Kiryakov, A., Kirilov, A., Manov, D., Ognyanoff, D., Goranov, M.: KIM semantic annotation platform. In The Semantic Web - ISWC 2003. Volume 2870 of Lecture Notes in Computer Science. Springer (2003) 834–849
26. Heras, S., Atkinson, K., Botti, V., Grasso, F., Julián, V., McBurney, P.: Applying argumentation to enhance dialogues in social networks. In: CMNA 2010. (2010)
27. Sporleder, C., Lascarides, A.: Using automatically labelled examples to classify rhetorical relations: A critical assessment. Natural Language Engineering **14**(3) (2008) 369–416