



CMNA X

The 10th workshop on
Computational Models of
Natural Argument

16 August 2010, Lisbon, Portugal

Introduction

Welcome to CMNA X!!

The series of workshops on Computational Models of Natural Argument has reached its 10th anniversary, with the current edition hosted in 2010 with ECAI in the beautiful Lisbon.

The longest standing event on Argument and Computation, CMNA has always been characterised by its aiming at the broader interdisciplinary audience, interested in natural, that is “real” argumentation. Naturalness may involve the use of means which are more immediate than language to illustrate a point, such as graphics or multimedia. Naturalness can also relate to the preference for one particular style of reasoning as opposed to another to structure complex arguments. Or to the use of more sophisticated rhetorical devices, interacting at various layers of abstraction. Or the exploitation of “extra-rational” characteristics of the audience, taking into account emotions and affective factors.

The current edition is no different in presenting a broad showcase of uses and models of natural argumentation, from the linguistic perspective to the interface issues, from the more formal approaches to the rhetorical and persuasive aspects, and with applications ranging from the medical, to the legal, the educational, the political and the security one.

CMNA will not be possible without the unvalued contribution of our programme committee. They have worked very hard in providing top quality, detailed reviews: on busy times like these, we are profoundly grateful for their tremendous support.

We hope you will enjoy CMNA X variegated programme, and we look forward, as always, to creative and stimulating discussions.

Chris Reed
Floriana Grasso
Nancy Green

Table of Contents

Argumentation and explanation in the context of dialogue	1
<i>Floris Bex, Katarzyna Budzyńska</i>	
Towards the Model of Central and Peripheral Arguments	5
<i>Katarzyna Budzyńska</i>	
Applying Argumentation to Enhance Dialogues in Social Networks	10
<i>Stella Heras, Katie Atkinson, Vicent Botti, Floriana Grasso, Vicente Julian, Peter McBurney</i>	
About the Ways of Criticism	20
<i>Erik C. W. Krabbe, Jan Albert van Laar</i>	
Generating Abstractive Summaries of Conversations through Automatic Argumentative Analysis	29
<i>Vincenzo Pallotta, Rodolfo Delmonte</i>	
Handling explanation in operational contexts	34
<i>Patrick Saint-Dizier</i>	
<i>Obligationes</i> as Formal Dialogue Systems	38
<i>Sara L. Uckelman</i>	
Working on the Argument Pipeline: Through Flow Issues between Natural Language Argument, Instantiated Arguments, and Argumentation Frameworks	47
<i>Adam Wyner, Tom van Engers, Anthony Hunter</i>	
The <TextCoop> Platform. Analyzing arguments in procedural texts . . .	57
<i>Thomas de Filippo, Laureline Marsal, Patrick Saint-Dizier</i>	

Argumentation and explanation in the context of dialogue

Floris Bex¹ and Katarzyna Budzyska²

Abstract. Whilst computational argumentation and computational explanation have both been studied intensively in AI, models that incorporate both types of reasoning are only just starting to emerge. The two forms of reasoning need to be clearly distinguished, as they may influence dialogue protocol and strategy. We show that this distinction can be made by considering the speech acts used to put forth the reasoning structures. Using the language of the Argument Interchange Format, ideas from speech act theory are integrated into a conceptual model that allows us to perform both argumentative and explanatory reasoning.

1 INTRODUCTION

Reasoning can be characterized as the process of moving from certain starting statements, assumptions or premises, to other statements, conclusions [17]. At the same time, reasoning is also the outcome of this process (i.e. the product), a static structure. Reasoning is typically used in the context of argumentation, where premises are offered as proof of a conclusion or a claim often in order to persuade someone or settle an issue. However, reasoning is also used in the context of explanation, where the *explananda* (facts to be explained) are explained by a coherent set of *explanans* (facts that explain). The usual purpose of explanation is not necessarily to convince someone but rather to help someone understand why the explananda are the case. In this paper, we aim to explore the similarities and differences between argumentation and explanation and make a first step towards an integrated computational model of the two.

Both argumentation and explanation are well-presented in their respective sub-fields of AI. A number of computational models of argumentation have emerged and matured in the past twenty-or-so years [10] and the computational aspects of the dialectics of argument (cf. [6]) and of the structure of argument [9] are well understood. Computational models for explanation are mainly based on the technique of abductive (model-based) reasoning, which has been studied in the context of medical and system diagnosis (e.g. [3, 8]); other examples of computational explanation are [4], who models explanatory dialogues, and [14], who uses explanations for natural language understanding. Despite the important role explanations can play in argumentative dialogue, there have not been many attempts to combine argumentation and explanation into one formal model. Perhaps the most thorough work thus far is [2], who combines structured arguments (cf. [10]) with abductive-causal reasoning into one model of inference to the best explanation. Other formal work that mentions both explanation and argumentation is [8].

Argumentation and explanation are often used in concert when performing complex reasoning: explanations can themselves be the subject of argumentation or they may be used in an argumentative way. Hence, we need a model that integrates argumentation and explanation. In such a model, the two types of reasoning should be clearly distinguished, because argumentation and explanation have different properties and in a dialogical setting the difference can influence protocol and strategies of dialogue. This distinction is not always easy to make because of the overlap between argumentation and explanation and the shifting between them in complex dialogue. It is even more complicated in the case of “reasoning-as-product”, that is, distinguishing between static arguments and explanations, which often have a similar (logical) structure.

In our opinion, the only way to distinguish between argumentation and explanation is by looking at the context in which the reasoning was originally performed. In this paper, we concentrate on the contextual property of the intention of the speaker. We are interested in how to represent the connection between the intentions and the static reasoning structure under consideration. In this paper, we show that this connection can be made by using ideas from speech act theory [15]. More specifically, not the propositional contents of a speech act (i.e. that which forms the static reasoning structure) but rather the illocutionary force of the speech act in a dialogue determines whether reasoning is argumentation or explanation. We will use the conceptual model of the Argument Interchange Format [5, 12] so as to provide a model that is not tied to any specific dialogue or argument formalism.

The rest of this paper is organized as follows. In section 2 we elaborate on the (structural and contextual) similarities and differences between argumentation and explanation and we give some intuitive examples of both types of reasoning. Section 3 discusses our ideas for a conceptual framework for argumentation and explanation. Finally, section 4 discusses some preliminary conclusions and ideas for future research.

2 SIMILARITIES AND DIFFERENCES

Argumentation is a type of reasoning used in a specific *probative function*, to prove a claim [17]. By its very nature, it involves some sort of opposition between parties³ and reasons are not just given to support for a conclusion but also to remove an opponent’s doubts about this conclusion. For example, a reasoning $\alpha \vdash \beta$ is argumentation when β is questioned (dubious) and a proponent of this argument uses α not only to support β , but also to remove an opponent’s doubts about β . Explanation, on the other hand, has not as its main goal to prove but rather to explicate why something is the case. Explanation

¹ School of Computing, University of Dundee, Dundee, UK, DD1 4HN, email: florisbex@computing.dundee.ac.uk

² Institute of Philosophy, Cardinal Stefan Wyszyński University in Warsaw, Poland and School of Computing, University of Dundee, Dundee, UK, DD1 4HN, email: k.budzyska@uksw.edu.pl

³ Hence the use of the term “calculus of opposition” for argumentation-theoretic semantics that allow one to calculate the acceptability of arguments.

in its purest form is not inherently dialectical and an explanation is given to help the other party, not to convince them. Consider the following example. Say I arrive at work at ten in the morning and my boss asks why I am late. I can either explain to him that the bridge was open and that I had to wait or I can argue that I am not “late”, because my contract does not specify the exact hours I have to be at the office. In first case, I am answering my boss’ question by explaining to him what caused my being late. In the latter case, I am arguing against my boss claim that I am late.

Argumentation and explanation are often used in conjunction. Explanations can themselves be the subject of argumentation, as one may argue in support or in opposition of a particular explanation or parts of it. For example, if my boss questions my explanation by arguing that I never cross a bridge on my way to work, I can argue (e.g. by providing evidence) that I do. Furthermore, explanations may be used in an argumentative way, as having someone agree to a particular explanation of a phenomenon might help us to persuade them. For example, if my boss accepts my explanation for being late I might convince him not to fire me.

Because argumentation and explanation are often intertwined in complex reasoning, they can sometimes be hard to distinguish from one another. However, it is important that we do distinguish the two types of reasoning. Apart from providing a measure of conceptual neatness, there are also more concrete reasons for not confusing the two types of reasoning. One of them is that circular arguments are usually considered fallacious while circular explanations are not. Take [18]’s recession example. An economist is asked why the economy is in recession in a certain state at present, and she replies: “Right now a lot of people are leaving the state, because taxes are too high”. But when asked why taxes are so high, she responds: “Well, a lot of people are unemployed, because of the recession”. The economist has not committed the fallacy of arguing in a circle, because he was explaining human behavior which has inherent feedback loops. The second reason for correctly distinguishing between argument and explanation is that the type of reasoning used might influence the allowed and desired moves in a dialogue. The ways in which to correctly respond to an explanation are different from the ways in which one should respond to argumentation; for example, it does often not make sense for the other party to deny the explananda whilst it does make sense to deny the conclusion of an argument. Similarly, a request for information is often better met by explaining something than by arguing that something is the case.

One possible way of distinguishing between argumentation and explanation might be to look at the product of reasoning, that is, the argument or the explanation put forth, and the structure and type of this product. At first sight, it often seems an explanation is *abductive* and *causal* whilst an argument is *modus-ponens* style, non-causal reasoning. The basic idea of abductive inference is that if we have a general rule $\alpha \rightarrow \beta$, meaning α *causes* β , and we observe β , we are allowed to infer α as a possible explanation of β . In contrast, argumentation is often seen as reasoning from a premise α to a conclusion β through an inference rule $\alpha \rightarrow \beta$, where this rule need not necessarily be causal. However, as it turns out it is also possible to give abductive or causal *arguments* (see e.g. [19]’s argument from evidence to hypothesis and causal argument). Similarly, one may perform explanatory reasoning by taking a rule $\beta \rightarrow \alpha$, meaning β is *evidence for* α (see [2] for a discussion on evidential and causal reasoning).

In our opinion, the distinction between argumentation and explanation is not one that is inherent to the product of reasoning, the static structure. Rather, the distinction follows from the dialogical context

in which the reasoning was originally performed. In order to determine this context, we need not just look at the original intention of the speaker but also at the broader dialogical context, such as the utterance that was replied to by the speaker and the intentions of the other participants. In other words, the context is largely determined by the *speech acts* that were performed (see e.g. [16] for the very basic concepts of the speech act theory needed in this paper).

According to the pragmatic theory of speech act, argumentation and explanation can be treated as different speech acts. A speech act $F\alpha$, such as: *claim* α , *why* α , consists of an illocutionary force F and a propositional content α . An illocutionary force is an intention of uttering a propositional content. That is, the performer of a speech act may utter α with an intention of asserting, asking, promising and so on.

Originally, Searle & Vanderveken recognized argumentation as an instance of content’s property, i.e. *argue* α is a speech act consisting of an assertive illocutionary force of uttering α which is a conclusion intended to be supported by premises that the performer of the speech act provides: “When one argues that P one asserts that P and gives reasons which support the proposition that P , normally with the perlocutionary intention of convincing the hearer that P ” [16, p. 184]. Observe that in such an account premises of a reasoning are hidden. For example, the argumentation “I am not late, because my contract does not specify the exact hours I have to be at the office” has to be formalized as *argue* p , where p represents the conclusion “I am not late”, while the premise “My contract does not specify the exact hours I have to be at the office” remains unexpressed in the formalization.

We use a different approach proposed in [1], where argumentation and explanation are both instances of illocutionary acts that represent a relation between premises and conclusions: *argue*(α, β) and *explain*(α, β), where α denotes a conclusion and β denotes premises. The distinction between argumentation and explanation cannot just be made by looking at the original speech act; one also needs to consider the broader dialogical context. In the next section, we show how this can be represented in the AIF⁺.

3 SPECIFICATION IN THE AIF⁺

In this section, we present how in general the AIF⁺ describes argument and its context (Section 3.1). Then, we propose how to model argumentation and explanation in the AIF⁺ (Section 3.2) and finally we show how they can be represented as a context of reasoning (Section 3.3).

3.1 Architecture of the AIF⁺

The AIF⁺ [13] is a dialogical extensions to the Argument Interchange Format, AIF (see e.g. [5]). The AIF is an attempt to bring together a wide variety of argumentation technologies so that they can work together. The AIF⁺ extends this approach by allowing to explicitly handle the context of dialogue in which a reasoning is put forth. It enables to connect the locutions uttered during a dialogue (argument₂) and the underlying arguments expressed by the content of those locutions (argument₁).

In the AIF⁺, the argument₁ is represented by two kinds of nodes:

- information (I-) nodes, which refer to data, and
- scheme (S-) nodes, which refer to the passage between information nodes, which are classified into three groups:
 - rule application (RA-) nodes which correspond to inference or support,

- conflict application (CA-) nodes which correspond to conflict or refutation,
- preference application (PA-) nodes which correspond to value judgments or preference orderings.

The argument₂ is also described by two types of nodes:

- locution nodes (L-), which refer to utterances and constitute a subclass of information nodes, and
- transition application (TA-) nodes, which refer to the passage between locutions and constitute a subclass of rule application nodes.

The TA-nodes are governed by the protocol of a dialogue system, recording e.g. that a given assertion has been made in response to an earlier question. [12] shows two examples of protocols that can be represented in the AIF+: the Two Party Immediate Response (TPI) [7] and Argument Scheme Dialogue (ASD) [11].

The interaction between argument₁ and argument₂ is captured by means of two types of illocutionary application (YA-) nodes [12]:

- the YA-nodes between I-nodes and L-nodes, and
- the YA-nodes between RA-nodes and TA-nodes.

For example, an YA-node may represent the relation between an assertion claim α with its propositional content α . The YA-link is determined and warranted (authorized) by the constitutive rules for speech acts [15]. These rules determine what constitutes a successful speech act. For example, an assertion may be unsuccessful and attacked, if its performer did not have enough evidence for the statement or he declared what he actually disbelieves.

3.2 YA-nodes

In this section, we propose the specification of argumentation and explanation in the AIF⁺. We will illustrate it on the example adapted from [18].

Allen The Evanston City Council should make it illegal to tear down the city's old warehouses.

Beth What's the justification for preserving them?

Allen The warehouses are valuable architecturally.

Beth Why are they so valuable?

Allen The older buildings lend the town its distinctive character.

Walton points out that Allen's first response is argumentation, while a second one is explanation. We follow this assumption without further considerations.

In the dialogue between Allen and Beth (see Fig. 1), the argument₂ consists of five speech acts represented by L-nodes (we use abbreviation L_i to denote subsequent locution nodes). The argument₁ consists of three propositions represented by I-nodes (I_i means subsequent information nodes). The interaction between the argument₂ and the argument₁ is described by means of the YA-nodes. The speech acts L₁, L₃ and L₅ have assertive illocutionary force connecting them with propositional contents I₁, I₂ and I₃, respectively. The passage between L₁ (resp. L₃, L₅) and I₁ (resp. I₂, I₃) is represented by YA₁ (resp. YA₄, YA₇). The illocutionary node YA₂ (resp. YA₅) links the directive L₂ (resp. L₄) and its propositional content I₁ (resp. I₂): not all YA-nodes are assertive schemes.

The most interesting is the complex type of illocutionary force which could be treated as intention of arguing and explaining. In the

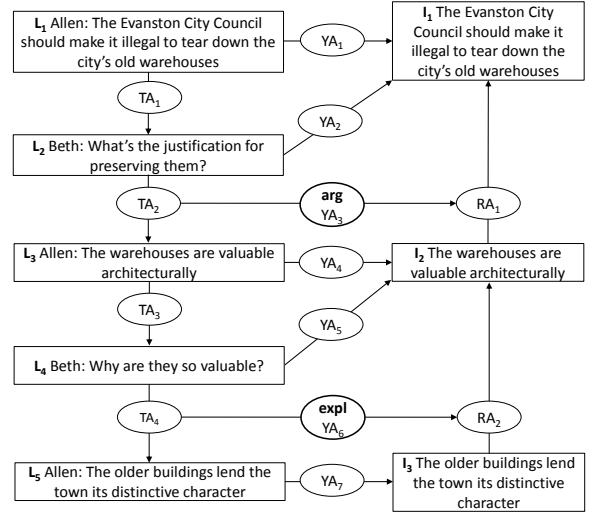


Figure 1. The AIF⁺ description of the example from [18]

AIF⁺, the complex illocution is represented by the YA-nodes between RA-nodes and TA-nodes [12]. In Fig. 1, there are two such nodes: YA₃ and YA₆. According to the assumption made above, YA₃ corresponds to argumentation and YA₆ to explanation. The illocution YA₃ links Allen's response to Beth's challenge (i.e. TA₂) with the argument "The warehouses are valuable architecturally" for the claim "The Evanston City Council should make it illegal to tear down the city's old warehouses" (i.e. RA₁). This captures the intuition that Allen's argumentation is invoked by Beth's challenge. On the other hand, the illocution YA₆ links Allen's response to Beth's request for information (i.e. TA₄) with the explanation "The older buildings lend the town its distinctive character" for the claim "The warehouses are valuable architecturally" (i.e. RA₂). This captures the intuition that Allen's explanation is invoked by Beth's request for information.

Observe that we could represent argumentation and explanation as YA₄ and YA₇, respectively. However, in such a representation they are indistinguishable from simple assertion. Assigning argumentation and explanation to the TA- and RA-nodes captures the intuition that they are social processes that emerge from the interaction between agents such that one agent responds to interlocutor's request for justification or explanation.

3.3 Context of reasoning

In Section 2, we emphasized that argumentation and explanation can be distinguished not by the structural properties of an underlying reasoning, but by its contextual properties. In this section, we propose how to model argumentation and explanation as a context of reasoning.

In the AIF⁺, the context of reasoning structure is represented by argument₂, in which a reasoning was performed, as well as by the interaction between argument₁ and argument₂. The structural indistinguishability of argumentation and explanation means that the same structure may be either argumentation or explanation depending on the intentions of the speaker who performed a given speech act. Consider the following dialogue between John and Ann.

John The warehouses are valuable architecturally.
Ann Why? I don't think so!
John The older buildings lend the town its distinctive character.

Observe that in this dialogue we have the same reasoning structure (argument₁) as in the previous dialogue: “The warehouses are valuable architecturally, since the older buildings lend the town its distinctive character”, however, the intentions of providing the premise are different in the case of John and in the case of Allen. That is, John utters the premise “The older buildings lend the town its distinctive character” to argue (prove) the conclusion “The warehouses are valuable architecturally”, while Allen utters the same premise to explain (make understandable) the conclusion.

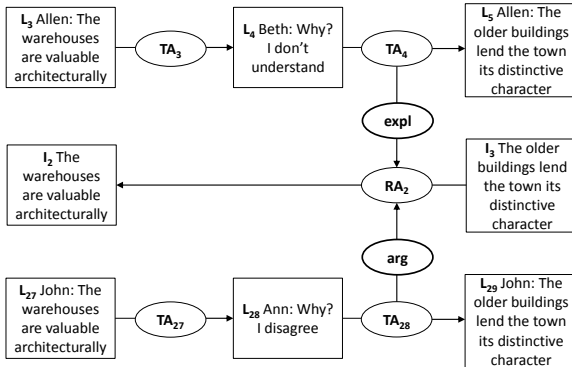


Figure 2. The same reasoning structure in different contexts

Fig. 2 shows how the distinction between reasoning contexts can be captured. Say that the moves in the dialogue between John and Ann are represented by L₂₇, L₂₈ and L₂₉. The passage between those moves are modeled by TA₂₇ and TA₂₈. The fragment of the previous dialogue (represented by L₃, L₄, L₅ and TA₃, TA₄) generates the reasoning structure modeled by I₂, I₃ and RA₂. Since the John-Ann dialogue refers to the same reasoning structure, the nodes L₂₇, L₂₈, L₂₉, TA₂₇ and TA₂₈ should be related to the nodes I₂, I₃ and RA₂. In Fig. 2, we focus on the relation that corresponds to the intentions of argumentation and explanation. If Allen’s response to Beth’s question was intended to explain Beth why the older buildings lend the town its distinctive character, then the node TA₄ should be linked with the node RA₂ by means of an explanatory type of YA-node (denoted as expl in Fig. 2). On the other hand, if John’s response to Ann’s question was intended to prove Ann why the older buildings lend the town its distinctive character, then the node TA₂₈ should be linked with the node RA₂ by means of an argumentative type of YA-node (denoted as arg in Fig. 2).

4 CONCLUSIONS AND FUTURE WORK

In the paper, we propose the basic framework for representing argumentation and explanation as the context of reasoning with the use of the AIF⁺ language. We propose to model them as the YA-nodes that link together the TA-nodes and the RA-nodes. The TA-nodes correspond to the response of an interlocutor in a dialogue that consists in giving the justification (in the case of argumentation) or explanation. The RA-nodes correspond to the reasoning structure which

originates from the interaction between interlocutors in a dialogue. Thus, we show how illocutionary force can be used to distinguish between the two types of reasoning. This presents a good basis for further research on the subject. In particular, it would be interesting to see how the semi-formal AIF⁺ specification relates to a more formal framework for argumentation and explanation, such as [2].

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of EPSRC under grant EP/G060347/1 for Floris Bex, and the support from Polish Ministry of Science and Higher Education under grant N N101 009338 for Katarzyna Budzynska.

REFERENCES

- [1] N. Asher and A. Lascaride, *Logics of Conversation*, Cambridge University Press, Cambridge, 2003.
- [2] F.J. Bex, *Evidence for a Good Story. A Hybrid Theory of Arguments, Stories and Criminal Evidence*, Doctoral dissertation University of Groningen, 2009.
- [3] T. Bylander, D. Allemang, M. C. Tanner, and J. R. Josephson, ‘The computational complexity of abduction’, *Artificial Intelligence*, **49**, 25–60, (1991).
- [4] A. Cawsey, *Explanation and Interaction: The Computer Generation of Explanatory Dialogues*, MIT Press, Cambridge, MA, 1992.
- [5] C. Chesnevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk, and S. Willmott, ‘Towards an argument interchange format’, *The Knowledge Engineering Review*, (21(4)), 293–316, (2006).
- [6] P.M. Dung, ‘On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and *n*-person games’, *Artificial Intelligence*, **77**, 321–357, (1995).
- [7] P.E. Dunne and T.J.M. Bench-Capon, ‘Two party immediate response disputes: Properties and efficiency’, *Artificial Intelligence*, **149**(2), 221–250, (2003).
- [8] D. Poole, ‘Learning, bayesian probability, graphical models, and abduction’, in *Abduction and Induction: essays on their relation and integration*, eds., P. Flach and A. Kakas, Kluwer, (1998).
- [9] H. Prakken, ‘An abstract framework for argumentation with structured arguments’, *Argument and Computation*, **1**, (2010). to appear.
- [10] H. Prakken and G.A.W. Vreeswijk, ‘Logics for defeasible argumentation’, in *Handbook of Philosophical Logic*, eds., D. Gabbay and F. Günthner, volume 4, 219–318, Kluwer Academic Publishers, Dordrecht/Boston/London, second edn., (2002).
- [11] C. Reed and D. Walton, ‘Argumentation schemes in dialogue’, in *Dissensus and the Search for Common Ground (Proceedings of OSSA 2007)*, (2007).
- [12] C. Reed, S. Wells, K. Budzynska, and J. Devereux, ‘Building arguments with argumentation: the role of illocutionary force in computational models of argument’, *Submitted to COMMA2010*, (2010).
- [13] C. Reed, S. Wells, J. Devereux, and G. Rowe, ‘Aif+: Dialogue in the argument interchange format’, in *Computational Models of Argument: Proceedings of COMMA-2008*, p. 311323, (2008).
- [14] R. Schank, *Explanations Patterns: Understanding Mechanically and Creatively*, Lawrence Erlbaum, Hillsdale, NJ, 1986.
- [15] J. Searle, *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, New York, 1969.
- [16] J. Searle and D. Vanderveken, *Foundations of Illocutionary Logic*, Cambridge University Press, Cambridge, 1985.
- [17] D. Walton, ‘What is reasoning? what is an argument?’, *The Journal of Philosophy*, ((87)8), 399–419, (1990).
- [18] D. Walton, ‘Epistemic and dialectical models of begging the question’, *Synthese*, (152), 237–284, (2006).
- [19] D.N. Walton, C. Reed, and F. Macagno, *Argumentation Schemes*, Cambridge University Press, Cambridge, 2008.

Towards the Model of Central and Peripheral Arguments

Katarzyna Budzynska¹

Abstract. The paper lays foundations for the computational model of persuasive argumentation. It is inspired by the elaboration likelihood model (ELM) proposed by Petty and Cacioppo. The ELM is one of the most influential contemporary theories of persuasion in social psychology. It assumes that there are two routes to persuasion and two types of persuasive means (arguments): central and peripheral. The central route to persuasion (CR) is related to content-based arguments, while the peripheral route (PR) is related to unsubstantial kind of impact, such as credibility or attractiveness of the source of information.

1 INTRODUCTION

In most of the cases, natural arguments have a persuasive function, i.e. they are put forth to make others believe or do something. Therefore, in real-life practice argumentation combines aspects related to the logical structures of reasoning with the rhetorical (persuasive) aim of its performance. While in theory of argumentation the structure of arguments is well studied (see e.g. [17, 12, 20, 11]), their persuasive nature is less explored. Moreover, the analysis of communication in natural contexts requires models which determine not only how to *correctly* build argumentation (such as models provided by logic-oriented theories), but also models which allow to reconstruct messages adequately to how people *actually* build arguments (such as models based on psychological theories). Such analysis shall start with a psychology-based type of model, which would enable to reconstruct components that people actually identify in a communication. Next step would be to use a logical model which allows to evaluate this message. In particular, we could analyze if arguments support a claim in the satisfactory degree, if sources of information can be trusted, etc.

As a result, there is a need to combine logical approaches with theories of persuasion. Theory of argumentation exploits some elements of rhetoric. Computational models of arguments (such as e.g. ARGUMED [19], Araucaria [15], the AIF [4], ArgDF [14] or Avicenna [13]) find particularly important the concept of argumentation schemes which corresponds to the classical idea of *topoi* [1]. However, there have not been many attempts to use theories of persuasion, which has been developed more recently. The rich source of such theories is social psychology (see [6] for an overview). To the best of the author's knowledge, the only model that directly links computational approach with social psychology is the model inspired by cognitive coherence theory proposed by Pasquier *et al* [7]. Cognitive coherence is used to represent the cognitive aspects of communication through the concept of attitude change which captures the persuasive aspect inherent to each types of communication including argumentation.

This paper considers the possibility of exploiting another theory from social psychology in the computational models of natural argument. The elaboration likelihood model (ELM) proposed by Petty and Cacioppo [8] is one of the most influential contemporary theory of persuasion. The ELM proved its importance as a theory which allowed to explain the inconsistency in the results of experiments. Other psychological models assumed one route to persuasion and one type of persuasive means (arguments). The ELM suggested that there are two routes and two types of means: central and peripheral. The two-route theory turned out to predict the results of experiments more adequately. It showed why one factor was important in some experiments, while in others it played insignificant role. In fact, the significance of a factor depends on the route which is activated in the particular experiment.

The central route to persuasion (CR) is related to content-based arguments, while the peripheral route (PR) is related to unsubstantial kind of impact, such as credibility or attractiveness of the source of information or the influence on emotions of the receivers (audience). In the CR, the outcome of persuasive effort is the result of the receiver's thoughtful consideration of issue-relevant material of the message. When the CR-processes are engaged, the quality (strength) of arguments' content is likely to influence persuasive success. On the other hand, in the PR the outcome of persuasion is the result of less thoughtful processes, such as some heuristic (a simplifying decision rule). The persuasive success is influenced by cues other than an argument's content, e.g., the receiver might reach a conclusion based on the communicator's credibility or likeability. In consequence, the PR-arguments, e.g. built upon someone's credibility, were an important factor in the experiments, in which the peripheral route was activated, while it was insignificant in those cases, in which the central route was dominant [10].

The likelihood that a receiver will engage in elaboration (issue-relevant thinking) depends on the elaboration ability (such as e.g. prior background knowledge or the presence of distraction in the communication setting) and the elaboration motivation (such as e.g. the receiver involvement, i.e., the personal relevance of the issue). For example, when an issue is not involving and there is some distraction present, the receiver may rely on PR-arguments such as the communicator's expertise. In such a case, high-credibility communicators will be more successful than low-credibility communicators. On the other hand, when the issue is personally relevant, the quality of argument content will become more important.

The usage of the PR-arguments has both strengths and weaknesses. First, they allow societies to function efficiently. For example, the PR-arguments based on the mechanism of authority enables to create the hierarchy in a society, which in turn enables to develop complex systems of production, defence, etc. [5]. The peripheral route is also beneficial from the point of view of an individual agent. It allows him to save time, energy and his "mental resources" in pro-

¹ Institute of Philosophy, Cardinal Stefan Wyszyński University in Warsaw, Poland and School of Computing, University of Dundee, Dundee, UK, DD1 4HN, email: k.budzynska@uksw.edu.pl

cessing information he receives. Nowadays, peripheral arguments are especially (and will become even more) important, since the pace of life is so fast and the amount of messages received is so high that it makes processing all of messages in the central route impossible. However, the peripheral route has also some disadvantages. It is particularly susceptible to manipulation. The receivers do not fully control peripheral cues, since they are automatic (“subconsciously driven”) social means of influence. Thus, it is especially important to support the skills of identifying and evaluating the elements related to the peripheral route to persuasion.

The main tasks of the paper are: (1) to bridge the gap in modeling persuasive argumentation bringing together traditions of informal logic, philosophy and social psychology, (2) to extend the model of argumentation with the components specific for persuasive functions of communication.

The paper is organized as follows. Section 2 presents a basic framework for persuasive argumentation. Section 3 proposes the specification in the AIF language. And Section 4 discusses the characteristic properties of central and peripheral arguments.

2 BASIC FRAMEWORK

One of the tools used to represent and analyze argumentation is the graph-theoretic method of argument diagramming [17]. In this section, I show how this method can be applied to represent persuasive argumentation inspired by the ELM. I specify the model only for the relation of support between arguments. In the future work, I want to enrich this account with the relation of attack.

Let D be a digraph (directed graph) (V, E) representing the *network of persuasive arguments*, where:

- V is a set of vertices (nodes), what in argument diagramming corresponds to *messages* (formulas that are premises or conclusions),
- E is a set of arrows (directed edges) which are 2-element ordered pairs of V , what in argument diagramming corresponds to the relation of *support* (not the relation of attack).

For a digraph, two maps are defined $\text{init}: E \rightarrow V$, and $\text{ter}: E \rightarrow V$, which assign to every edge e an *initial vertex* $\text{init}(e)$ and a *terminal vertex* $\text{ter}(e)$. If $(\text{init}(e), \text{ter}(e)) = (x, y)$, then we will write $e = xy$.

In the standard approach there is one type of vertices in a digraph. A digraph for the model of persuasive argumentation needs to represent two types of vertices representing CR-arguments (in a diagram marked with a solid line) and PR-arguments (marked with a dotted line). The CR-arguments are central arguments that are content-based and involves issue-relevant thinking, while the PR-arguments are peripheral arguments that includes means such as credibility or attractiveness of communicator, expert opinion, testimony, popular opinion or audience’s emotions. Thus, we distinguish two types of vertices $V(D)$ by defining the mapping that allows to color them with two colors CR and PR corresponding to two types of routes, i.e. central and peripheral:

Let D be a digraph (V, E) . A *vertex coloring* for the model of persuasive argumentation is a map $c: V \rightarrow \{CR, PR\}$.

In tools that uses the method of argument visualization (such as e.g. Araucaria), there may be a need to diagram not only the whole argumentative text, but also these of its parts which uses the peripheral route. For example, if we analyze and visually represent a message, we may want to have a general view where in this message PR-

arguments were involved. It is especially important, when the analyst want to find those parts of the message which are potentially susceptible to manipulation. As a result, it can support the skills of identifying and evaluating the elements related to the peripheral route to persuasion.

The simplest way to “extract” PR-arguments is to subtract the set of vertices colored with CR from the set of all vertices. We will call it a maximal separation and a PR-digraph obtained, written D_{PR}^{max} , will represent only those arguments which are peripheral. The CR-digraph for maximal separation can be obtained in a similar manner.

Let $D = (V, E)$ be a network of persuasive arguments. *Peripheral* directed subgraph for *maximal separation* $D_{PR}^{max} \subseteq D$ is a pair $(V_{PR}^{max}, E_{PR}^{max})$ such that:

- $V_{PR}^{max} = \{x \in V \mid c(x) = PR\}$
- $E_{PR}^{max} = \{xy \in E \mid x \in V_{PR}^{max} \wedge y \in V_{PR}^{max}\}$

Consider an example given in Fig. 1. Say that a case 1.1 shows a network of persuasive arguments D . Then, a case 1.2 represents the result of maximal separation. The set V_{PR}^{max} consists of the vertices from V which are colored with PR , i.e. b, d and e . The edges E_{PR}^{max} , that remained after maximal subtraction, have to start and end in peripherally colored vertices, i.e. ed .

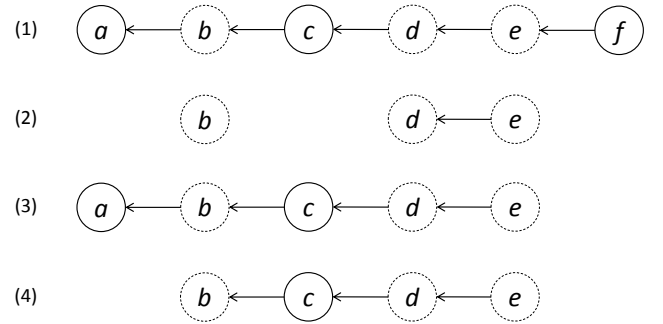


Figure 1. An example of a digraph for persuasive argumentation (1), and separation between CR- and PR-arguments: (2) maximal, (3) partial, and (4) minimal

Such an operation allows to “quickly” separate those fragments of a text that uses the peripheral route from those based on the central route. This could be useful in some applications. However, it may also be not easy to analyze a diagram obtained. In more complex diagrams, the visual representation may become “inconsistent” or “discontinuous”, what can make it difficult to see the general picture of interrelation among the particular fragments of the text (compare Fig. 1.2 with 1.1). Perhaps, the more useful would be to “remove” those CR-arguments which do not “disconnect” a network of persuasive arguments. We will call it a partial separation and a PR-digraph, written D_{PR}^{part} , is obtained by subtracting those CR-arguments which are not supported by any peripheral arguments. In this manner, the “continuity” of a diagram is prevented and a main conclusion remains at the diagram even if it is central (see Fig. 1.3). The CR-digraph for partial separation can be obtained in a similar manner.

Let $D = (V, E)$ be a network of persuasive arguments. *Peripheral* directed subgraph for *partial separation* $D_{PR}^{part} \subseteq D$ is a pair

$(V_{PR}^{part}, E_{PR}^{part})$ such that:

- $V_{PR}^{part} = \{x \in V \mid c(x) = PR \vee (c(x) = CR \wedge \exists x_1 \dots x_n x \subseteq D(c(x_i) = PR \text{ for } i \in \{1, \dots, n\}))\}$
- $E_{PR}^{part} = \{xy \in E \mid x \in V_{PR}^{part} \wedge y \in V_{PR}^{part}\}$

The set V_{PR}^{part} contains those elements of V , which are: either peripheral (e.g. the vertices b, d, e in Fig. 1.3), or central (e.g. a in the case 1.3) but there need to be a path in D ending at this central element (e.g. the path $edcba$) such that one of path's element is peripheral (e.g. b). The set E_{PR}^{part} , that remained after partial subtraction, has to start and end in vertices belonging to V_{PR}^{part} , e.g. cb or ba .

We may also want to have “continuous” subgraph which not only does not start, but also does not end in the CR-arguments. We will call it a minimal separation and a PR-digraph, written D_{PR}^{min} , is obtained by subtracting those CR-arguments which are not supported by any PR-arguments or not supporting any PR-arguments. In this manner, the “continuity” of a diagram is prevented, however, a main conclusion may be removed from the diagram, if it is central. The CR-digraph for minimal separation can be obtained in a similar manner.

Let $D = (V, E)$ be a network of persuasive arguments. Peripheral directed subgraph for *minimal separation* $D_{PR}^{min} \subseteq D$ is a pair $(V_{PR}^{min}, E_{PR}^{min})$ such that:

- $V_{PR}^{min} = \{x \in V \mid c(x) = PR \vee (c(x) = CR \wedge \exists x_1 x_2 \dots x_m x_{m+1} \dots x_n x \subseteq D(c(x_i) = PR \text{ for } i \in \{1, \dots, m\} \wedge c(x_j) = PR \text{ for } j \in \{m+1, \dots, n\}))\}$
- $E_{PR}^{min} = \{xy \in E \mid x \in V_{PR}^{min} \wedge y \in V_{PR}^{min}\}$

The set V_{PR}^{min} contains those elements of V , which are: either peripheral (e.g. b, d, e in Fig. 1.4), or central (e.g. c in the case 1.4), but there need to be a path in D (e.g. the path dcb) such that this central element has to follow a PR-argument (in this case, the element d) and has to be followed by a PR-argument (in this case, the element b). Observe that a is removed from a diagram D , since no peripheral argument follows a . The set E_{PR}^{min} , that remained after minimal subtraction, have to start and end in vertices belonging to V_{PR}^{min} , e.g. cb , but not ba (since $a \notin V_{PR}^{min}$).

3 SPECIFICATION IN THE AIF

The Argument Interchange Format, AIF (see e.g. [4]), is an attempt to bring together a wide variety of argumentation technologies so that they can work together. At the most abstract layer, the AIF provides an ontology of concepts which can be used to represent argument structure. In the upper ontology, the argument is described by two kinds of nodes: (1) information (I-) nodes, which refer to data, and (2) scheme (S-) nodes, which refer to the passage between information nodes. The S-nodes are classified into three groups: (2.1) rule application (RA-) nodes which correspond to inference or support, (2.2) conflict application (CA-) nodes which correspond to conflict or refutation, and (2.3) preference application (PA-) nodes which correspond to value judgments or preference orderings.

At a more concrete layer, the AIF provides a set of specific argumentation schemes, i.e. forms of argument or patterns of reasoning. They can have a deductive, inductive or presumptive form. If an argument is presumptive and its premises are true, then the conclusion may presumably be taken to be true [16]. In the form ontology, argumentation schemes (such as argument from expert opinion or argument from consequences) are defined as types of RA-nodes with

a conclusion description (such as “ A may plausibly be taken to be true”) and one or more premise descriptions (such as “ E is an expert in domain D ”).

In order to represent persuasive argumentation in the AIF, we could distinguish two types of I-nodes corresponding to the central and peripheral routes. However, since the AIF allows to represent argumentation schemes, we may model persuasive arguments with the use of two types of RA-nodes:

- RA^{CR} -nodes representing the patterns of reasoning which involves a thoughtful considerations of arguments’ content, and
- RA^{PR} -nodes representing the cognitive processes which are based on shortcuts such as communicator’s credibility.

Consider two arguments: “We should lower taxes, since it will make people happy”, and “We should lower taxes, since Mary said so”. The first one (an instance of argument from consequences) is central, since to accept a conclusion “We should lower taxes” the receiver needs to consider the content of a premise “It will make people happy”. Thus, in the AIF it is represented by the RA^{CR} -nodes (see Fig. 2). The second one is PR-argument, since in the absence of other arguments for lowering taxes the receiver has to rely on Mary’s credibility. Thus, in the AIF it is represented by the RA^{PR} -nodes (Fig. 2). In this way, it is possible to query, e.g. ArgDB, not only for arguments supporting war in Iraq, but also for the PR-arguments for war in Iraq, i.e. arguments which are potentially used with an intention of manipulating the public opinion.

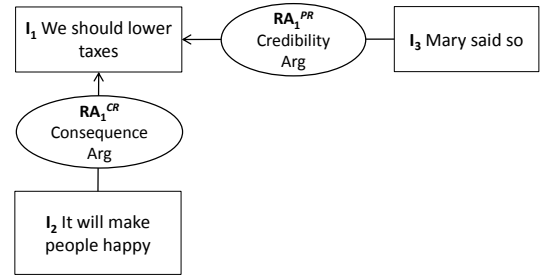


Figure 2. An example of CR-argument and PR-argument represented in the AIF

The set of RA^{PR} -nodes should include schemes of source indicators reasoning (see e.g. [22]) or ethotic [2] and pathotic arguments [3]. The ELM suggests that in the peripheral route the receivers rely on the source credibility, attractiveness or likeability. Thus, the following RA^{PR} -type schemes could be formulated:

- x says α , x is credible, therefore (plausibly) α ,
- x says α , x is attractive, therefore (plausibly) α ,
- x says α , I like x , therefore (plausibly) α .

In the future work, I plan to elaborate the set of PR-schemes in a more systematic manner.

In computational approaches, the most influential taxonomy of argumentation schemes is a model proposed by Walton [23]. In this model, CR- and PR-argumentation schemes can be distinguished.

Among the Walton's set of schemes available in Araucaria, the following RA^{CR} -type schemes can be identified: argument from sign, argument from example, argument from verbal classification, argument from evidence to a hypothesis, argument from falsification of a hypothesis, argument from correlation to cause, argument from cause and effect, argument from consequences, argument from analogy, argument from waste, argument from established rule, argument from exceptional case, argument from precedent, argument from gradualism, slippery slope arguments, argument from vagueness of verbal classification, argument from arbitrariness of verbal classification and arguments from ignorance. In Araucaria, there are also the RA^{PR} -type schemes in Walton's set: argument from position to know, appeal to expert opinion, argument from popular opinion, argument from popular practice, ethotic argument, argument from bias and argument against the person.

It may seem that argument from expert opinion is a good quality argument that should belong to the RA^{CR} -type of arguments. Even though it may be used this way, it has also a potential to be used in a manipulatory manner: "[arguments from expert opinion] can be deceptive and misleading - they are sometimes powerfully persuasive even when they turn out to be groundless and flawed" [21, p.30].

The PR-arguments may be "rationalized" or "strengthen", if we testify them with the use of critical question. Thus, even if the proponent performs a "weak" argument from expert opinion trying to use its peripheral potential (based on the social mechanism of authority [5]), then the receiver may check its quality by asking critical questions associated with this type of arguments. On the other hand, if the receiver elaborates the message in the peripheral route (since an issue is not involving or there is some distraction present), then even a "weak" appeal to expert opinion may be successful. The problem of the type of route activated in a particular communicative situation is inexpressible in the pure AIF, since it does not allow to represent the context of argument performance. Partly, it can be expressed in the AIF^+ [18], if the argument was performed in the dialogue, but perhaps its more satisfactory representation will be available in $AIF2.0$.

4 PROPERTIES OF CR- AND PR-ARGUMENTS

According to the ELM, in the peripheral route the receiver accepts a claim not on the basis of cognitively expensive reasoning, but on the basis of some shortcuts (such as credibility or attractiveness of communicator). That is, the PR-arguments are used by the receiver to decide whether to believe the claim "without engaging in any extensive cognitive work relevant to the issue under consideration" [9, p. 256]. Those arguments do not activate mechanisms of processing and comparing message's contents, but mechanisms of social influence (such as e.g. the mechanisms of authority or liking [5]) which are based on people's built-in (automatic) reactions.

In other words, the arguments that are effective in this type of route requires execution of different cognitive processes. In the central route, a receiver processes not only contents of arguments and a claim, but also the relation between them. In case of a message "We should lower taxes, since it will make people happy", a receiver processes the content of both reasoning's components as well as their relevance. If a hearer received a message such as "We should lower taxes, since $2+2=4$ ", he would reject this piece of reasoning at once, not because of the false argument, but because there is no relation (relevance) between the argument and the claim.

In the peripheral route, a hearer does not (or at least "is not intended to") process this relation. As mentioned above, he does not have to engage in an extensive cognitive work *relevant* to the issue

under consideration. It could be understood in this way that a hearer uses peripheral argument as a cue to believe a claim, but not necessary has to think if and what they have in common. In case of a message: "We should lower taxes, since Mary said so", Mary's credibility is used as a shortcut for believing that we should lower taxes. Observe that when the grounds for trusting Mary's words is her credibility, then the hearer does not have to, but still may consider the relation between her credibility and the issue of lowering taxes (i.e. if she is credible in the domain of taxes). However, when the grounds for trusting Mary's words is her attractiveness, then, clearly, there is no relation which could be extensively, thoughtfully considered and testified with the use of critical questions.

Generally, the arguments have different properties when executed in the different routes such as: they are activated in the different circumstances (involvement vs. lack of involvement, etc.), they are executed in the different way (higher vs. lower cognitive cost, etc.), they have different outcomes (more vs. less resistant to counterarguments, etc.), and finally - they have different structure basis (content-based vs. heuristic-based).

5 CONCLUSIONS AND FUTURE WORK

In legal argumentation, the most important function of message is the interaction among arguments. Thus, analysis of this type of communication may concentrate on how reasons support e.g. a judge's decision. However, when we are interested more about the communication typical for politics or advertisement, we can discover that substantial (content-based) arguments play the secondary role, and the majority of the message is supposed to influence the receiver with the use of different means. For example, when a beautiful girl stands by a car and the advertisement says "Prices start from 8.000 pounds", the advertiser hopes that the presence of the attractive girl will influence potential buyers. Otherwise, he would use only the CR-argument referring to the good price of his product.

The paper proposes the basic formal framework of the persuasive arguments inspired by the elaboration likelihood model (ELM). ELM suggests that there are two different types of routes to persuasion and two different types of persuasive arguments: central and peripheral. Three directions of future research are particularly attractive. First, I plan to consider a taxonomy of peripheral schemes of argumentation. Next, other theories of persuasion as well as formal rhetoric should be explored and compared with the account of persuasive function of communication proposed by ELM. Finally, the relation of attack in the network of persuasive argumentation should be examined. One way to accomplish it would be to assume that in case of a CR-argument attacking a PR-argument the CR-argument wins because of its "quality". However, the ELM could be also interpreted in this way that depending on the route activated we need different semantics to compute the result of an attack. For example we could assume that a CR-argument will win in the central route and lose in the peripheral route.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the support from Polish Ministry of Science and Higher Education under grant N N101 009338.

REFERENCES

- [1] Aristotle, *Rhetoric*, Dover Publications, 2004.
- [2] A. Brinton, 'Ethotic argument', *History of Philosophy Quarterly*, **3**, 245-257, (1986).

- [3] A. Brinton, 'Appeal to the angry emotions', *Informal Logic*, 77–87, (1988).
- [4] C. Chesnevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk, and S. Willmott, 'Towards an argument interchange format', *The Knowledge Engineering Review*, (21(4)), 293–316, (2006).
- [5] R. B. Cialdini, *Influence: Science and practice (4th ed.)*, Allyn and Bacon, Boston, 2001.
- [6] D. J. O'Keefe, *Persuasion: Theory and research*, Thousand Oaks, CA: Sage, 2002.
- [7] P. Pasquier, I. Rahwan, F. Dignum, and L. Sonenberg, 'Argumentation and persuasion in the cognitive coherence theory', in *First International Conference on Computational Models of Argumentation (COMMA)*, pp. 223–234, (2006).
- [8] R. E. Petty and J. T. Cacioppo, *Communication and persuasion: Central and peripheral routes to attitude change*, New York: Springer-Verlag, 1986.
- [9] R. E. Petty and J. T. Cacioppo, *Attitudes and Persuasion: Classic and Contemporary Approaches*, Westview Press, 1996.
- [10] R. E. Petty, J. T. Cacioppo, and R. Goldman, 'Personal involvement as a determinant of argument-based persuasion', *Journal of Personality and Social Psychology*, (41), 847–855, (1981).
- [11] J. Pollock, *Cognitive Carpentry. A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA., 1995.
- [12] H. Prakken, 'An abstract framework for argumentation with structured arguments', *Argument and Computation*, 1, (2010). to appear.
- [13] I. Rahwan, B. Banihashemi, C. Reed, D. Walton, and S. Abdallah, 'Representing and classifying arguments on the semantic web', *The Knowledge Engineering Review*, (2010). to appear.
- [14] I. Rahwan, F. Zablith, and C. Reed, 'Laying the foundations for a world-wide argument web', *Artificial Intelligence*, (171(10-15)), 897–921, (2007).
- [15] C. Reed and G. Rowe, 'Araucaria: Software for argument analysis, diagramming and representation', *International Journal of AI Tools*, 14(3-4), 961–980, (2004).
- [16] C. Reed and D. Walton, 'Towards a formal and implemented model of argumentation schemes in agent communication', *Autonomous Agents and Multi-Agent Systems*, (11(2)), 173–188, (2005).
- [17] C. Reed, D. Walton, and F. Macagno, 'Argument diagramming in logic, law and artificial intelligence', *The Knowledge Engineering Review*, ((22)1), 87–109, (2007).
- [18] C. Reed, S. Wells, K. Budzynska, and J. Devereux, 'Building arguments with argumentation: the role of illocutionary force in computational models of argument', *Proc. of COMMA2010*, (2010). to appear.
- [19] B. Verheij, 'Automated argument assistance for lawyers', *Proc. of Conference on Artificial Intelligence and Law*, 43–52, (1999).
- [20] B. Verheij, 'The toulmin argument model in artificial intelligence. or: how semi-formal, defeasible argumentation schemes creep into logic', *Argumentation in Artificial Intelligence*, 219–238, (2009).
- [21] D. Walton, *Appeal to Expert Opinion : Arguments from Authority*, University Park, Pa., Penn State Press, 1997.
- [22] D. Walton, *Ad Hominem Arguments*, University of Alabama Press, Tuscaloosa, 1998.
- [23] D. N. Walton, *Argumentation Schemes for Presumptive Reasoning*, Erlbaum, Mahwah NJ, USA, 1996.

Applying Argumentation to Enhance Dialogues in Social Networks

Stella Heras¹ and Katie Atkinson² and Vicente Botti
and Floriana Grasso and Vicente Julián and Peter McBurney

Abstract. Nowadays, many websites allow social networking between their users in an explicit or implicit way. In this work, we show how argumentation schemes theory can provide a valuable help to formalize and structure on-line discussions and user opinions in decision support and business oriented websites that held social networks among their users. Two real case studies are presented and analysed. Then, guidelines to enhance social decision support and recommendations with argumentation are provided.

Introduction

Currently, the Web has evolved from read-only HTML documents and personal websites to social sites where people interact, share and constantly update huge amounts of decentralised information. This new conception of the Web as a platform for computing and collaborative interaction has been supported by the development of so-called Web 2.0 technologies and standards like AJAX, SVG, FOAF, PHP, Ruby, XHTML, P2P or XSL. The result has been the fast proliferation of web-based communities, on-line social networks, web applications and webservices.

Nowadays, many websites allow social networking between their users in an explicit or implicit way. While some are declared leisure oriented social networking sites, others are more decision support or business oriented while still allowing their users to interact, share their preferences and profiles, form communities with other users and give advice, recommendations and feedback about their experiences. Some examples of the former type are Facebook¹, Flickr², MySpace³, Orkut⁴ and Twitter⁵, whose general purpose is to keep users in contact by providing them with a smart interface to show their profile, manage their acquaintances list, join groups, share media content and chat. The latter is the case of on-line shopping companies, such as Amazon⁶ or eBay⁷ and consumer review sites, like Tripadvisor⁸ or Epinions⁹, which provide their users with a means of networking and discussing matters related to their business or reviewing topic. In addition, there is a third type of on-line social networking site whose orientation lies between leisure, business and decision support. Some examples of this further type are websites such as Ethicaleconomy¹⁰ or Kiva¹¹, which support networking and debate to promote ethical

values and social help. Regardless of the purpose of the social networking, in all of these communities discussions arise from the difference of opinion between users, and individual views are mixed in the tangle of user-generated content posted in discussion boards, wikis and blogs. Therefore, there is an obvious need for mechanisms to structure this information and to elicit as much useful knowledge as possible from it.

In this work, we show how Argumentation Theory (concretely Argumentation Schemes Theory [19]) can provide a valuable help to formalize and structure on-line discussions and user opinions. Argumentation schemes are stereotyped patterns of human reasoning that can improve the user's understanding about discussions and provide a means to evaluate what users have stated and why. When opinions are product recommendations to other users, they are usually justified because they match the user profile (i.e. fit the content of user's declared preferences and likes), the profile of similar users (i.e. collaborative filtering) or both (i.e. hybrid recommendations). Usually, there is not an explanation about the reasoning process that has been followed to come up with specific recommendations. In fact, these recommendations tend to come directly from the recommendation algorithm that runs the website and not from the acquaintances that a user has in his social network. However, this does not follow future trends on the Web, where discovering is becoming social (as reported by Joe Kraus, Google's director of product management in a talk at the Supernova conference 2008) and consequently, recommendations could be expected to come directly from acquaintances in a decentralised way. Moreover, people trust recommendations more when the engine can explain why it made them [13] and what is understood as a good recommendation is changing from the one that minimises some error evaluation measure about the output of content, collaborative filtering or hybrid recommendation methods to the one that really makes people happier.

On the other hand, when user opinions are conveyed in reviews and guides that users write to provide pieces of advice to other users, the reasons that the author has put forward his ideas may be implicit in the text. However, for non-expert users identifying and developing a deep understanding about all such implications can be difficult. Moreover, many reviews and guides are written collaboratively between several users, starting with an initial review followed up with many comments and replies. Thus, each individual opinion can be blurred as the number of posts grows.

Regarding evaluation, user opinions are commonly assessed using some measures of trust and reputation (e.g. usefulness degrees, reviewer ranks, seller ratings and customer feedbacks) in decision support or business oriented websites. These values are internally computed in the website by providing users with rating tools to score

¹ Departamento de Sistemas Informáticos y Computación. Universidad Politécnica de Valencia, Spain, email: sheras@dsic.upv.es

² Department of Computer Science, University of Liverpool, UK

¹www.facebook.com ²www.flickr.com ³www.myspace.com

⁴www.orkut.com ⁵twitter.com ⁶www.amazon.com

⁷www.ebay.com ⁸www.tripadvisor.com

⁹www.epinions.com ¹⁰www.ethicaleconomy.com

¹¹www.kiva.org

the posts of other users or to leave feedback about their experiences. However, they are not usual on leisure oriented social networking sites, where the truthfulness of the user opinion has a lesser importance since its final consequences do not usually give rise to a wrong decision or an unsuccessful commercial transaction. Nevertheless, even when user opinions are attached with trust and reputation values, these measures do not provide an objective way of assessing them, for instance, by looking at the reasoning patterns that they follow to come up with specific conclusions. Thus, user opinions can be misunderstood and rated low, decreasing unfairly the trust and reputation values of their authors.

In this work, we focus in particular on business oriented websites that allow a social interaction among their users. We leave out of our analysis for the time being decision support, leisure or ethics oriented social networking sites, as the scope and target of on-line business websites makes them more suitable to define tools to analyse opinions and elicit knowledge from their users. In this paper, we start with a brief overview of related work which, though not necessarily explicitly related to social networks, applies argumentation theory to classical social network activities. Then we formalise our notion of social network, and we show how this definition fits to Amazon and eBay, probably best epitomize how implicit social networks emerged from business websites. Finally, we discuss how argumentation schemes could be best utilised to improve social networking features of these sites.

1 Related Work

Although not directly applied to social networks, the work most relevant to our purposes is perhaps the work on recommender systems [1]. Recently, research has investigated the impact of the social network dynamics on the recommendation, typically based on the notion of "social trust" [7, 16, 17]. A relevant work [9] applies argumentation to manage the interaction that emerges from the recommendation dialogues between a social network of agents, by means of a dialogue game that controls the recommendation process. It uses Argumentation Schemes [19] to define a set of potential attacks to the recommendations provided by agents. Here, the concepts of argumentation, recommender systems and social networks are studied together for the first time. The recognition that current recommendation systems are unable to make use of the large amount of quantitative data to empower recommendation has led to the consideration that a more sophisticated approach based on argumentation could be the key [5]: preferences formalised according to Defeasible Logic Programming [6] inform an inference tool able to analyse and decide conflicts among the set of equally viable recommendations. An important additional contribution of [5] is the identification of a number of research questions, such as *exposing underlying assumptions* behind recommendations, *approaching trust and trustworthiness* from the perspective of backing arguments and *providing rationally compelling arguments* for recommendations. The work in the present paper contributes to these areas but from the different perspective of structuring and clarifying the reasoning process followed by users to provide pieces of advice and recommendations to other users of their social network.

Other relevant research, though not directly applied to social networking, applies argumentation to support trust and reputation. In [11], a framework for evaluating trust-related arguments in on-line stores is proposed, evaluated empirically, and extended by applying the Toulmin model of argument to provide guidelines for the implementation of well-structured trust-assuring arguments and to in-

vestigate if the provision of these arguments actually increases the customer trust in Internet stores. A novel quantitative trust model for argumentation-based negotiating agents is proposed in [2], suitable to emerging applications like e-business, based on a negotiation dialogue game, and presenting a model for securing agent oriented systems in which agents can utilise an argumentation system to reason over the reputation of other agents. Our approach of applying argumentation to social networks does not intend to provide new measures of trust and reputation for the users of the network, but support the existing measures with arguments that justify them and clarify the reasoning process behind them.

Finally, preliminary work on applying argumentation to prediction markets was proposed in [14]. Prediction, or information, markets, are a special type of social network whose purpose is to aggregate the information held by their users to make predictions about specific events or parameters. The work analyses the influence of the social relationships on the predictions made by group judgement, where a group of agents linked via a social network argue on the final outcome of a prediction.

2 Definition of a Social Network Model

We studied a number of social networks, focusing in particular to the argumentation activities that, either implicitly or explicitly, users would engage in. In particular, following the typology of argumentative dialogue in [18], we assessed how different social networks compare with this feature. This comparison is shown in Table 1. From this analysis, we extrapolated a general abstraction of social network.

For our purposes, we consider a social network as an abstraction to represent social structures that link individuals or organisations. Links can stand for different types of interdependency, such as friendship, trade, shared knowledge, common hobbies, etc. We distinguish between **explicit social networks**, which openly represent users and links among them, so that users can, for instance, search their *contact list* to interact with other users, and **implicit social networks**, which may or may not store information about social relationships among users, but which usually do not make this information accessible to users, who cannot access their contact lists to retrieve previous partners or do not have an easy way of searching reports about previous exchanges. For both types, we identify the following features that define a social network in our model:

- **Overall purpose** of the network: e.g. friendship, business, shared hobbies.
- **Permitted tasks**: e.g. recommend, provide opinions, evaluate others' opinions.
- **Nodes** representing individuals or organisations.
- **Roles** that individuals or organisations can play in the social network.
- **Knowledge databases**: individual or shared knowledge databases associated with each node and representing information about the issues related with each role.
- **Ties**, or links, between nodes, which can be of different sorts, depending on the overall purpose of the network (e.g. values, visions, ideas, financial exchange, friendship, personal relationships, kinship, dislikes, conflict, trade). Ties can be directed or not. Undirected links represent social relations that are present in the network, but whose related information is not stored nor explicitly supported.
- **Social network analysis measures**, used to evaluate the relations that a tie represents. Values of trust and reputation are common examples of these measures.

- Types of **argumentative dialogues** that can be held in the network.

In what follows we concentrate on two case studies, analyzing arguably the two most popular business oriented websites that allow social interaction among users, despite this not being their primary purpose: Amazon and eBay. For each, we analyse the features that make them considered *de facto* social networks, and we represent them in the light of the model we defined above.

2.1 Amazon

Amazon (www.amazon.com) is possibly the largest on-line retailer offering, either directly or via "marketplace" associated sellers, a very wide range of products, from books, to groceries, from furniture to clothes and shoes, and so on. Social networking features allow users to interact in different ways. Amazon's users can:

- write reviews about products, whether purchased or not. As part of their review, users can rate products. Reviews can be annotated with the nickname of the reviewer or his popularity as reviewer ("reviewer rank"), based on both positive and negative votes received, as well as the time when the review was published. In addition, other users can write comments on reviews, rate them as useful/unuseful, and report them to the company if they consider them offensive or inappropriate.
- leave feedback about "marketplace" sellers after a purchase, with a comment. Seller ratings are computed using the votes received over the transactions performed in a specific period of time. Sellers have the opportunity to respond to the comment/rating and rate the transaction, but they cannot rate buyers (only feedback submitted by buyers is considered to compute a seller rating).
- join customer communities: users can create a profile and share it with other users, join different communities, participate in forum, create *Listmania* lists with the Amazon products they like or recommend and *Wish* lists with the products they are interested in, suggest products to their communities by adding a tag, write *So You'd Like to...* guides to directly recommend products. Posts can be replied to, rated and reported, but these ratings are not used to compute customer ranks.

On top of this, Amazon website runs a powerful recommendation algorithm that matches each customer's purchased and rated item to similar items, and outputs a personalised recommendation list [12]. This algorithm follows an *item-to-item collaborative filtering* approach that scales to massive data, producing recommendations in real time with a brief explanation (e.g. "N% customers buying X also bought Y").

Depending on the activities that customers perform in Amazon, different customer **roles** can be identified. Figure 1 shows a use case diagram with the roles that take part in the Amazon social networking features and the activities that customers can carry out when they are playing each role. Note that customers can play different roles when they are carrying out different activities on the website. In addition to them, Amazon customers can play other roles (e.g. administrators, developers, associates, etc.) and perform other operations using other features (e.g. Amazon Web Services, Amazon Marketplace Payments, etc.), but they are out of the scope of this paper.

The most frequent type of customer is a potential *buyer* that checks the website searching for a specific product. For clarity purposes, we consider this role as the default role that any customer that registers on the website plays. Furthermore, specific users can also play

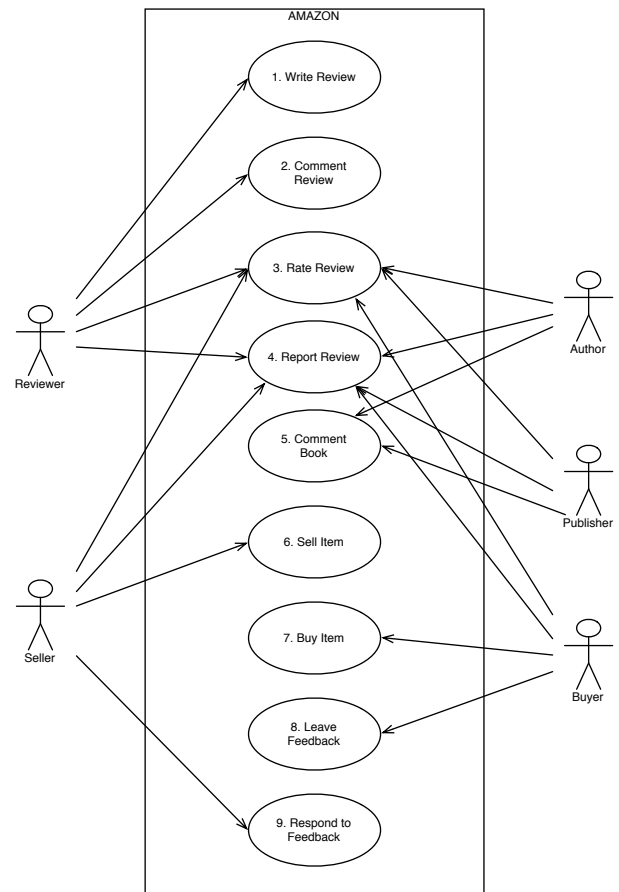


Figure 1. Amazon Use Case Diagram

the *seller*, *author* and *publisher* roles, depending on their business or their relation with the product. As explained above, after a purchase, the buyer has the opportunity of leaving feedback about his experience with the seller, reviewing the transaction. Apart from selling their products, *sellers* can also answer to this feedback if they want to add any comment or reply. In addition, Amazon allows *authors* and *publishers* to write comments on their works.

Other important roles from the social networking perspective is the *reviewer* role, which is played by any Amazon customer as soon as he writes a review on any product or a comment on a review, sharing his knowledge and opinions about the product and making recommendations to other Amazon customers. Note that we consider that customers play the reviewer role both when they start a new review and also when they make comments on a review written by other customers. Finally, any Amazon customer is able to rate a review as useful/unuseful or report a review to the website if it is considered inappropriate.

2.1.1 Amazon Social Network Models

Explicit social networks are formed by the users joining communities, while implicit social networks emerge from the activity of writing reviews and from sales and their subsequent feedback. In the spirit of our analysis, we focus here on the latter, and we analyse the social networks emerging from reviews and sales according to

Activity	Persuasion	Negotiation	Inquiry	Deliberation	Information Seeking	Eristics	Examples
Blogs	X				X		Pingback, Slashdot, LiveJournal, BlogSpot
Collaborative RT Editors					X		SubEthraEdit, SynchroEdit, ACE, Moonedit, Google Docs & Spreadsheets and Zoho
Commercial Sites	X		X	X	X	X	Amazon and eBay
Commercial Social Networks		X			X		Dell IdeaStorm
Consumer Review Sites with Social Networking features	X		X	X	X	X	TripAdvisor and ePinions
Deliberative Social Networks	X			X		X	Webs of discussion and debate for decision-making purposes between individuals and government
Ethical Sites	X		X	X	X	X	Kiva and Ethicaleconomy
Forums	X			X	X	X	Yahoo! Groups and Google Groups
Instant Messaging					X		Gtalk, Skype, ICQ, Yahoo! Messenger, MSN, Pidgin AOL and Jabber
Social Cataloguing			X		X		CiteULike, Connotea, BibSonomy and rebase
Social Guides					X		WikiTravel
Social Libraries			X		X		discogs.com, imdb.com and LibraryThing
Social Network Search Engines					X		Newstrove
Social Network Sites	X			X	X	X	Facebook, Flickr, MySpace, Orkut and Twitter
Social On-line Storage					X		Using servers or P2P technology
Text Chat					X		IRC and other technologies
Virtual Worlds	X	X	X	X	X	X	Dotsoul, SecondLife, Active Worlds, the Sims on-line, There, Planeshift, Croquet project, VOS, Solipsis, Everquest and World of Warcraft
Wikis			X		X		Wikipedia, Wikisource

Table 1. Argumentation dialogues enabled by social networking activities.

our model.

Social Network of Reviews

Reviews that Amazon users write give rise to social relations from which emerges an *Amazon Social Network of Reviews*, with the **purpose** of sharing information on the products, and with **nodes** representing buyers, sellers, reviewers, authors (of books for example) and publishers/manufacturers. The **tasks** permitted are: writing reviews on Amazon products, writing comments on reviews, rating reviews or reporting reviews. The main type of **dialogue** enabled by this social networking activity is information seeking and sharing, but persuasion is also enabled by means of free comments and responses to them. Figure 2 is an example of a network of reviews with six Amazon customers playing each role that is involved in the activity of writing reviews. Arrows stand for **social ties** implicitly created from users' activity. In the example, arrows from User 1 to Users 3, 4, 5 and 6 stand for social ties meaning that User 1 has written a review about a product related with those other users (because they are sellers, authors, publishers or buyers of the product). Arrows from Users 3, 4, 5 and 6 to User 1 stand for a social tie that represents that those users rated or reported the review of User 1. Finally, arrows between Users 1 and 2 represent social ties that can be created when reviewers comment, rate or report the review of other reviewers. Also, comments on reviews can also be commented, rated and reported. **Knowledge databases** are attached to each role, denoting the information that the website stores for each role, e.g. the

reviews made by a reviewer and how they were rated, though not always the information is accessible to a user (e.g. there is no obvious way for users to check their list of reviews). **Analysis measures** evaluate users' performance in the social network, e.g. the reviewer's rank can be used as a reputation measure to evaluate the importance of a reviewer. Measures can also label ties, e.g. individual reviewers' trust measures could be computed from the usefulness ratings assigned over a certain period of time.

Social Network of Sales

The overall **purpose** of the *Amazon Social Network of Sales* is to run commercial transactions between its members and to inform about them. Therefore, the **tasks** permitted on the network are sell and buy products and leave feedback about these commercial transactions, while **nodes** are simply buyers and sellers (as feedback cannot be left unless a transaction has occurred). The main type of **dialogue** that these tasks enable is that of information seeking and sharing, by leaving feedback, and persuasion, by supporting this feedback and responding to it. Figure 3 shows an example of an Amazon social network of sales with one seller (User 1) and two buyers (Users 2 and 3) and their respective **knowledge databases**. As before, arrows show **social ties**: the arrow from User 1 to Users 2 and 3 means that this customer has sold a product to Users 2 and 3 and thus, a social tie between them has been created. The arrow from User 2 to User 1 represents that User 2 has provided feedback about his sale with User 1 and hence, this has generated a new social tie. This arrow can

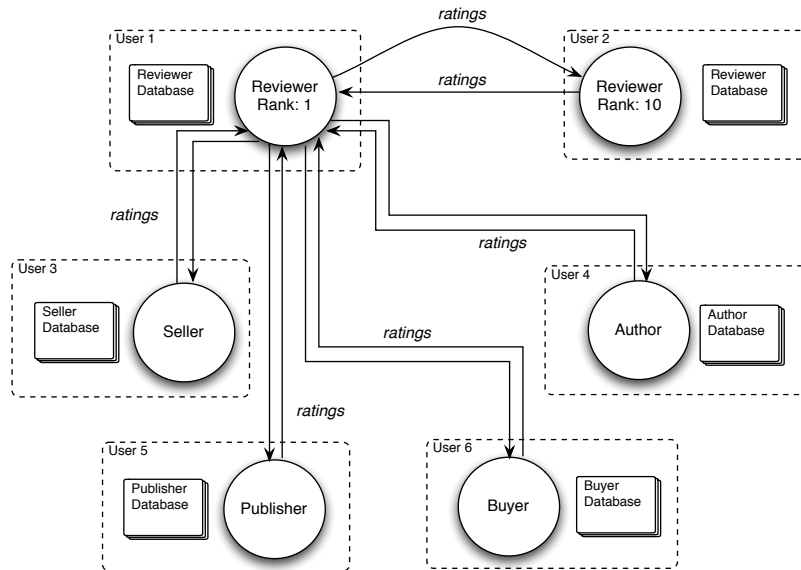


Figure 2. Amazon Social Network of Reviews

be labelled with the numeric feedback (ranging from 1 to 5 stars) that the buyer has left about the transaction, which we consider as a social network analysis measure. Amazon aggregates these numbers to compute a seller rating, which can also be used as a reputation measure to label nodes that represent customers playing the role of sellers.

In addition, the seller can also leave comments about the transaction and the feedback received (although not about the buyer himself and hence, this assessment does not count towards any rating about him). This is represented in the figure by the arrow from User 1 to User 2. Thus, this arrow cannot be labelled with a feedback score.

2.2 eBay

eBay (www.ebay.com) is the world's largest on-line marketplace, an open trading platform that allows sellers and buyers to get in touch. Unlike Amazon, eBay does not sell products directly and does not centralise transactions and payment, but it charges a fee to sellers for publishing their adverts. After the purchase, buyers directly pay sellers, either on-line (though they are encouraged to do this via the system, e.g. with PayPal) or in person. The core business model of eBay are auctions, though products can also be sold directly via a "Buy It Now" option. The main social network features in eBay are:

- reviews and guides. Reviews are labelled with reviewer information, which consists, among other things, of the reviewer's profile with the number of reviews, number of times starting the reviews for an item, number of guides written, helpfulness votes and *reviewer rank* (i.e. number of positive votes on the total amount of the reviewer's reviews). Reviews can be rated as helpful/unhelpful, but, unlike Amazon, they cannot be commented.
- feedback about completed transaction. Unlike Amazon, in eBay both feedback on sellers and buyers counts towards customer ratings. Detailed comments can also be left: buyers can review sellers on the details of their purchase (e.g. if the item really was as described in the ad, if the delivery was quick enough, etc.) and

viceversa (the buyer paid promptly, etc.). These detailed comments are anonymous and do not count towards the overall feedback score. Users' feedback profiles can be made private, so that comments on the purchase are hidden to other users, though the positive or negative ratings received cannot.

- Community services: such as *Neighbourhoods*, communities of people who share common interests; *My World*, allowing the creation of a profile; *Discussion boards*; *Groups*, which are member led discussion areas that allow members to share ideas, opinions and information quickly and easily; *Blogs*, and *Chat rooms*.

As pointed out before, all these activities allow eBay customers to interact and form social networks. The main **roles** that eBay customers play when they are performing them are shown in figure 4. For comparison with Amazon and clarity purposes, we have only included in the figure the roles that eBay customers play when they write reviews or perform sales, since the activities enabled by the eBay community services are mostly equal to them. Moreover, posts in other community forums can be commented upon, but not rated, so users cannot leave a numeric value that stands for their opinion about the post, which makes it more difficult for other users and even for the website administrators to have an overall view of the real opinion of the user that has written the post. However, any post in eBay can be reported to the site for misuse.

The most common roles that eBay customers play are the *seller* and the *buyer* roles, selling and buying items on the website respectively. By default, any user that registers on eBay plays the role of a potential *buyer*. In addition, both users can leave feedback about their counterpart and their commercial experience with him. The figure also shows an important difference between Amazon and eBay users when they play the reviewer role, since in eBay reviews cannot be commented by other users which would also play in that case the role of reviewers. Thus, reviews can be only written by one user and discussions are not permitted. However, any eBay customer can rate reviews, leaving his opinion about them. Also, reviews can be reported to the website if they are considered as inappropriate by any user.

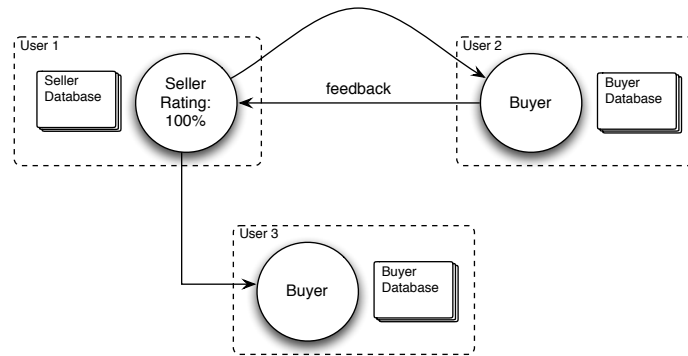


Figure 3. Amazon Social Network of Sales

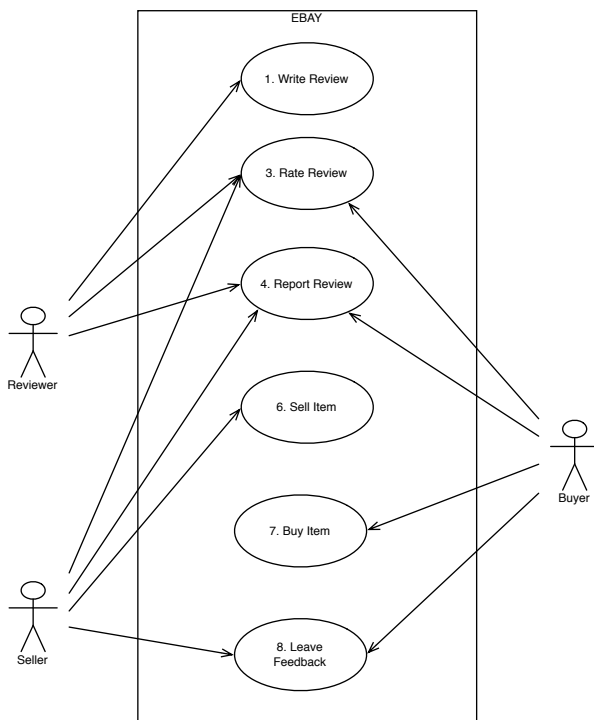


Figure 4. eBay Use Case Diagram

2.2.1 eBay Social Network Models

As with Amazon, eBay enables explicit social networks, via community services, and implicit social networks, emerging from the reviews and sales, with similar characteristics. As before, we concentrate on the implicit, business driven networks.

Social Network of Reviews

The reviewing facility that eBay offers to its customers enables the creation of an implicit social network that we have named the *eBay Social Network of Reviews*, whose main **purpose** is to share knowledge and experience about the items sold in eBay. Therefore, the **tasks** that eBay customers perform on this network are writing re-

views and evaluate them (by rating or reporting them). Unlike Amazon, reviews on eBay cannot be commented upon, so the main type of **dialogue** enabled in this network is information seeking and sharing.

Figure 5 shows an example of the network. As shown in the figure, **nodes** represent individuals playing the possible roles that take part in the social network of reviews with their associated **knowledge databases** and arrows stand for **social ties**. These are customers that have written reviews (i.e. reviewers) and other customers that have rated or reported them (i.e. sellers and buyers). Therefore, arrows from User 1 to Users 2, 3 and 4 imply that User 1 has written a review about an item sold or bought by the other users. Also, arrows from Users 2, 3 and 4 to User 1 mean that those users have rated or reported to the website the review written by User 1.

Figure 5 also shows some values computed by the website that can be considered as **analysis measures** to evaluate the performance of each user on the social network. For instance, nodes are labelled with the reviewer's rank computed by the website. Also, arrows from Users 2, 3 and 4 to User 1 are labelled with the ratings that those users have provided about the review of User 1, which are also used in the reviewer's rank calculation. In addition, some other complex analysis measures could be considered. For instance, the number of ratings or reports provided by a user could be used to evaluate his level of participation in the activities developed on the network.

Social Network of Sales

The core activity held on the eBay website has the **purpose** of buying and selling of products. Underlying to this activity, eBay allow both buyers and sellers to provide ratings on their partners, sharing with the website and other users their opinions about their commercial transactions. These activities give rise to social relations between the eBay customers, which we have represented by means of an *eBay Social Network of Sales* abstraction. Therefore, as in Amazon social network of sales, the **tasks** permitted on the network are buy and sell products and leave feedback about these transactions. Regarding types of **dialogues**, both information seeking and sharing, and persuasion are enabled, since customers are not only able to leave their opinions on sales, but also about their partners on these sales (i.e. opposite to Amazon, eBay allows both sellers and buyers to evaluate the commercial transaction and the user that they have done business with). Figure 6 shows an example of the eBay social network of sales.

In this network, customers interact playing either the seller or the

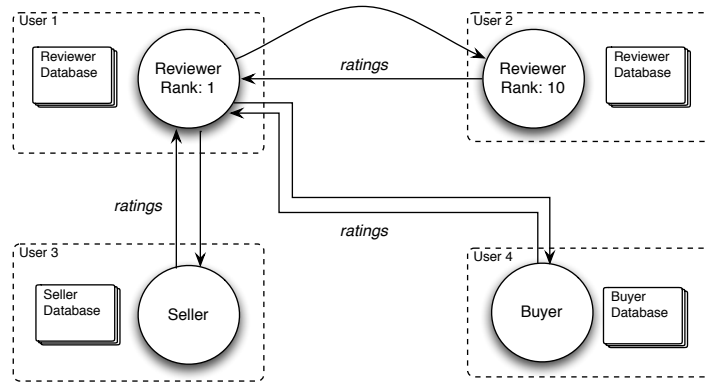


Figure 5. eBay Social Network of Reviews

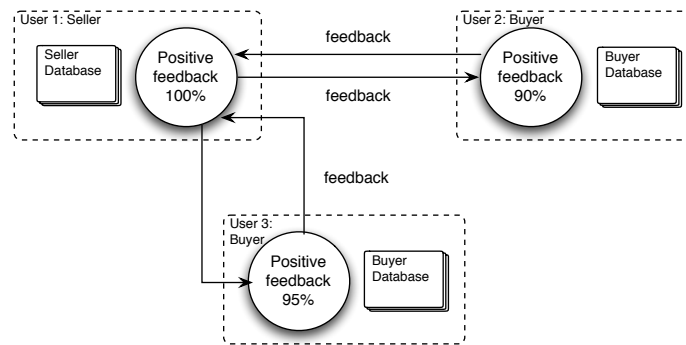


Figure 6. eBay Social Network of Sales

buyer role, which are represented on the network by **nodes** and their associated **knowledge databases**. The arrow from User 1 to User 3 means that User 1 has sold a product to User 3 and hence, a **social tie** has been generated. As pointed out before, sales can be evaluated by buyers, who leave feedback scores about it. An example of the associated social tie is represented in the figure from the arrows from Users 2 and 3 to User 1. Optionally, sellers can also evaluate buyers (although they are only allowed to leave positive feedback on buyers). This type of tie is shown in the figure by means of the arrow between User 1 and User 2. Feedback ratings can be used as social network **analysis measures**. eBay computes several scores for feedback (see eBay website for more information¹²). The figure shows, for instance, how the percentage of positive feedback could be used to label nodes with a reputation measure for each user playing a specific role on the network.

3 Argumentation Schemes to Support On-line Dialogues in Social Networks

So far, we analysed the reviews and sales features of Amazon and eBay, characterising the underlying interactions among their users by means of social network abstractions, which, we believe, are general enough to be transferable to other similar websites. In this section, we concentrate on how argumentation could enhance the performance of the emergent activities carried out by the users of a social

network, with a preliminary step towards the application of argumentation schemes to formalise the underlying reasoning shown in the dialogues held among the users of the networks in our cases of study. Argumentation schemes [19] are characterised by a set of premises and their underlying conclusion, and are associated with a set of critical questions (CQs) that stand for potential attacks that could refute the conclusion drawn from the scheme. This feature is very useful to guide argumentation dialogues. Thus, if a proponent of a position uses a pattern of reasoning that matches with an argumentation scheme, an opponent can try to pose one of its critical questions to attack that position. We analysed a number of typical dialogues held in the situation described in the two cases studies above, and we identify the following advantages of applying argumentation schemes to social networked business:

- *To provide a formal structure to opinions and recommendations*, allowing for explanations and justifications that clarify the position of the reviewer.
- *To provide a way of evaluating user opinions and recommendations*, by looking at their associated reasoning patterns, with critical questions as a way to show weaknesses and possible attacks.
- *To provide a formal structure to the dialogue as a whole*, clarifying the dynamics of each individual contribution in terms of the overall argument.

To illustrate these advantages, consider for example the conversation extract (inspired by real posts on Amazon) shown in figure 7 review-

¹²<http://pages.ebay.com/help/feedback/>

Customer Review

255 of 282 people found the following review helpful:

★★★★★ A must in your Argumentation bibliography, September 18, 2009

By User1
New Reviewer Rank: 2,525,408

...

....this book is an excellent reading. It's the third book that I've read from this author and it's as good or better than the last two. Any student or researcher on Argumentation in AI will enjoy the reading, which starts with some introductory chapters in the area and nicely flows to more specific topics. As a scholar in AI, I strongly recommend it...

...

[Permalink](#) | Was this review helpful to you? ([Report this](#))

[Add a comment](#)

Comments

[Track comments by e-mail](#)

[Showing 1-2 of 2 posts in this discussion](#)

User2 says:
New Reviewer Rank: 1,326,523

...so I'm still not sure about the quality of the book, since I read the 2nd of this series and I found it quite difficult to follow. What confuses me the most are what you (*i.e. User1*) said on your review of this 2nd book, where you wrote a hard criticism and strongly discourage the reading. Up to my knowledge, this could be a hard reading...

...

[Reply to this post](#)

[Permalink](#) | [Report abuse](#) | [Ignore this customer](#)

10 of 17 people think this post adds to the discussion. Do you?

User3 says:
New Reviewer Rank: 15,782

...I totally agree with User1. I haven't read other books on the series, but looking to this one, I guess they are also good. Moreover, although User1 discourage the reading of the 2nd book of the series for the non-scholars, he does so because its contents assume previous expertise on the area. This does not necessarily mean that the 2nd book is a bad reading...

...

[Reply to this post](#)

[Permalink](#) | [Report abuse](#) | [Ignore this customer](#)

5 of 6 people think this post adds to the discussion. Do you?

Figure 7. An example on Amazon reviews

ing book *B*. The argumentation can be summarised as:

- *User1* provides an argument in favour of the book:

A1: I am a scholar in the area of AI; I strongly recommend the reading of the book; THEREFORE this is a good reading

- *User2* replies with two arguments: an opinion about the topic and an attack to A1:

A2: I have read the 2nd book of the series of *B*; This wasn't a good reading; THEREFORE book *B* couldn't be a good reading either.

A3: *User1* says that book *B* and its series are good; *User1* posted a hard criticism and discouraged the reading of the book *B* in a previous review; THEREFORE the review of *User1* is inconsistent with what he said previously

- Finally, *User3* replies to *User2* with an argument that supports the argument of *User1*:

A4: *User1* is a scholar in the area of AI; *User1* discourages the reading for non-scholars of the 2nd book of the series of *B*; THEREFORE the 2nd book of the series of *B* isn't a good reading for non-scholars

Following [19], these arguments could be translated into argumentation schemes as:

A1: Argument From Expert Opinion

Major Premise: Source *User1* is an expert in subject domain *AI* containing proposition *book B is a good reading*

Minor Premise: *User1* asserts that *book B is a good reading* is true

Conclusion: *book B is a good reading* is true

CQ1: How credible is *User1* as an expert source?

CQ2: Is *User1* an expert in the field *AI* for which *book B is a good reading*?

CQ3: What did *User1* assert that implies that *book B is a good reading*?

CQ4: Is *User1* personally reliable as a source?

CQ5: Is the proposition *book B is a good reading* consistent with other experts assert?

CQ6: Is *User1*'s assertion based on evidence?

A2: Argument From Position to Know

Major Premise: Source *User2* is in position to know about things in a certain subject domain *books on B series* containing proposition *book B is a good reading*

Minor Premise: *User2* asserts that *book B is a good reading* is false

Conclusion: *book B is a good reading* is false

CQ1: Is *User2* in position to know whether *book B is a good reading* is true or false?

CQ2: Is *User2* an honest source?

CQ3: Did *User2* assert that the *book B is a good reading* is true or false?

A3: Argument From Inconsistent Commitment

Initial Commitment Premise: *User1* has claimed that he is committed to proposition *book B and its series are a good reading*

Opposed Commitment Premise: Other evidence shows that *User1* is not committed to proposition *book B and its series are a good reading* since he discouraged the reading of the book *B* in a previous review

Conclusion: *User1*'s commitments are inconsistent

CQ1: What is the evidence supposedly showing that *User1* is committed to proposition *book B and its series are a good reading*?

CQ2: What further evidence in the case is alleged to show that *User1* is not committed to proposition *book B and its series are a good reading*?

CQ3: How does the evidence from premise 1 and premise 2 prove that there is a conflict of commitments?

A4: Argument From Expert Opinion

Major Premise: Source *User1* is an expert in subject domain *AI* containing proposition *book B is a good reading*

Minor Premise: *User1* asserts that *the 2nd book of the series isn't a good reading for non-scholars* is true

Conclusion: *the 2nd book of the series isn't a good reading for non-scholars* is true

(CQs as in A1)

By associating a scheme to each argument, opinions are given, obviously enough, a formal structure, which makes the pattern of reasoning explicit. Users could be asked to explain their arguments by using the critical questions of a schema. For instance, in the example above, A3 attacks A1 in fact by instantiating its CQ4. Or, A4 attacks

A3 instantiating its CQ2. Moreover, users could be encouraged to clarify their position better: we have often found negative ratings of a product where the free text reveals that the bad experience was in fact related to the transaction (e.g. late shipment, item broken, etc.).

Of course for this situation to be realistic, users need to find the use of argumentation natural enough not to be discouraged to use it. Currently, there are several argumentation tools that offer support for on-line debates, with varying degrees of structure given to the arguments. Some examples are Debategraph (<http://debategraph.org/>), an evolution of DebateMapper that includes it as a view to comment, build, edit and rate debates; Debatepedia (<http://wiki.idebate.org/>), a new free wiki encyclopedia of arguments and debates and the tool for evaluate debate TruthMapping (<http://truthmapping.com/>). Recent developments have introduced Web 2.0 standards to support on-line debate. Some contributions of this type are Cope.it! [10], which encourages collaboration by sharing opinions and resources; the semantic web-based argumentation system ArgDF [15]; Cohere [3], a web tool for social bookmarking, idea-linking and argument visualisation; the Argument Blogging project [20], which intends to harvest textual resources from the Web and organise them into distributed argumentative dialogues and the On-line Visualisation of Argument (OVA at ARG:dundee: www.arg.dundee.ac.uk) tools, which facilitate argument analysis and manipulation in on-line environments. Some examples of tools that are of a more formal and structured nature include the Parmenides system [4] and the Carneades system [8]. Despite the proliferation of these tools, their uptake by business oriented websites like Amazon or eBay is questionable, as their main interest is not to alienate users from their site by providing a seamless and natural interaction.

4 Conclusions: Desiderata for Argumentation enhanced Social Networks

In this work we showed how argumentation theory can provide valuable insights in formalising and structuring on-line discussions and user opinions in business oriented websites. We gave a model of social network, and we provided two case studies of commercial websites, Amazon and eBay, fitting this model. Finally, we demonstrated how typical interactions in these environments could be seen as argumentation dialogues, and could in fact be enhanced by such features. Several conditions need to be verified before a more widespread uptake of argumentation techniques could be possible, however.

First, sites like Amazon or eBay should make each underlying social network explicit, so that users could exploit all information resources available in the website, in turn enhancing trust and reputation by providing public and transparent measures. Secondly, sites should provide easy-to-use tools for the quick and seamless identification of argumentation schemes in the line of reasoning that a user is following in a post. Although this aspect is more related to the advancement of the state of the art on argumentation and computation research, websites which decide for the uptake of a particular tool could grant some reward (e.g. positive feedback) to the users of these tools. Third, sites should provide tools to represent the dynamics of dialogues among users, so that attack and defense statements can be easily identified. Again, this comes at a considerable cost to the users (who would not necessarily be prepared to engage in a dialogue each time they want to leave a review for a product), so reward mechanisms should be used. Finally, sites should provide tools for summarising and analysing the information gathered from the schemes and attacks identification. A "summary" showing statistics and a graphical representation of debate on a product would repre-

sent a concrete added value for users, and an effective motivation to engage in argumentative activities. This it would allow, for instance, users to understand at a glance which is the most prominent view of a particular product they want to purchase, without having to read all reviews.

We believe that argumentation *can* make business driven social networking more rewarding, and we see this as one of the most promising application areas for research in argument and computation.

REFERENCES

- [1] G. Adomavicius and A. Tuzhilin: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, No. 6, pp. 734-749, 2005.
- [2] J. Bentahar, J.-J. Ch. Meyer and B. Moulin: Securing Agent-Oriented Systems: An Argumentation and Reputation-based Approach. *Proceedings of the 4th International Conference on Information Technology: New Generations (ITNG 2007)*, IEEE Computer Society, pp. 507-515, 2007.
- [3] S. Buckingham Shum: Cohere: Towards Web 2.0 Argumentation. *Proceedings of the 2nd Int. Conf. on Computational Models of Argument, COMMA*, pp. 28-30, 2008.
- [4] D. Cartwright and K. Atkinson. Political engagement through tools for argumentation. In *Proceedings of the 2nd Int. Conf. on Computational Models of Argument, COMMA*, pp. 116-127, 2008.
- [5] C.I. Chesñevar, A.G. Maguitman and M.P. González: Empowering Recommendation Technologies Through Argumentation. *Argumentation in Artificial Intelligence*, pp. 403-422, Springer, 2009.
- [6] A.J. García, J. Dix and G.R. Simari: Argument-based Logic Programming. *Argumentation in Artificial Intelligence*, pp. 153-171, Springer, 2009.
- [7] J. Golbeck: Generating Predictive Movie Recommendations from Trust in Social Networks. *Proceedings of the Fourth International Conference on Trust Management, LNCS*, Vol. 3986, pp. 93-104, 2006.
- [8] T. Gordon, H. Prakken, and D. Walton. The Carneades model of argument and burden of proof. *Artificial Intelligence*, Vol. 171, No. 10-15, pp. 875-896, 2007.
- [9] S. Heras, M. Navarro, Vicente Botti and V. Julián: Applying Dialogue Games to Manage Recommendation in Social Networks. *Proc. of the 6th Int. Workshop on Argumentation in MAS, ArgMAS*, 2009.
- [10] N. Karacapilidis and M. Tzagarakis: Web-based collaboration and decision making support: A multi-disciplinary approach. *Web-Based Learning and Teaching Technologies*, Vol. 2, No. 4, pp. 12-23, 2007.
- [11] D. Kim and I. Benbasat: The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation. *Information Systems Research*, Vol. 17, No. 3, pp. 286-300, 2006.
- [12] G. Linden, B. Smith and J. York: Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, Vol. 7, No. 1, pp. 76-80, 2003.
- [13] G. Linden, J. Hong, M. Stonebraker and M. Guzdial: Recommendation Algorithms, Online Privacy and More. *Communications of the ACM*, Vol. 52, No. 5, 2009.
- [14] S. Ontañón and E. Plaza: Argumentation-based Information Exchange in Prediction Markets. *Proceedings of the 5th International Workshop on Argumentation in Multi-Agent Systems, ArgMAS*, 2008.
- [15] I. Rahwan, F. Zablith and C. Reed: Laying the Foundations for a World Wide Argument Web. *Artificial Intelligence*, Vol. 171, No. 10-15, pp. 897-921, 2007.
- [16] J. Sabater and C. Sierra: Reputation and Social Network Analysis in Multi-Agent Systems. *Proceedings of AAMAS 2002*, No. 1, pp. 475-482, 2002.
- [17] F.E. Walter, S. Battiston and F. Schweitzer: A model of a trust-based recommendation system on a social network. *Journal of Autonomous Agents and Multi-Agent Systems*, Vol. 16, No. 1, pp. 57-74, 2007.
- [18] D. Walton and E. Krabbe: *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, 1995.
- [19] D. Walton, C. Reed and F. Macagno: *Argumentation Schemes*. Cambridge University Press, 2008.
- [20] S. Wells, C. Gourlay and C. Reed: Argument Blogging. *Proceedings of Computational Models of Natural Argument, CMNA 2009*.

About the Ways of Criticism

Erik C. W. Krabbe and Jan Albert van Laar¹

Abstract. This paper attempts to systematically characterize critical reactions in argumentative discourse, such as ‘objections’, ‘critical questions’, ‘rebuttals’, ‘refutations’, ‘counterarguments’, and ‘fallacy charges’, in order to contribute to the dialogical approach to argumentation. We shall make use of four parameters to characterize distinct types of critical reaction. First, a critical reaction has a *focus*, for example on the standpoint, or on another part of an argument. Second, critical reactions appeal to some kind of *norm*, argumentative or other. Third, they each have a particular illocutionary *force*. Fourth, a critical reaction occurs a particular *level* of dialogue (the ground level or some meta-level). The concepts here developed shall be applied to discussions of critical reactions by Finocchiaro, Pollock and Snoeck Henkemans.

1. INTRODUCTION

1.1 The notion of a critical reaction

The notions of criticism and of argument are very much related, both at a practical and at a theoretical level. In practice, a critical attitude is often manifested by ‘being argumentative’ in one’s comments and appreciations, whereas arguments are associated with a critical stance sooner than with a constructive one. In daily parlance, both ‘criticism’ and ‘argument’ even share some negative connotations, such as meddlesomeness and quarrelsomeness. In the theory of argumentation, there are no such connotations, but the theoretical concepts of criticism and of argument are all the same closely related. Argumentation can be either critical (opposing someone else’s point of view) or constructive (defending one’s own point of view) or both. Moreover, some sort of critical stance is often seen as essential for all argumentation, including the constructive kind, since argumentation is conceived as an instrument to overcome doubt, and doubt seems to imply a critical stance. In pragma-dialectics, the normative model for argumentation proposed is that of a critical discussion in which standpoints are critically tested. Also at the intersection of argumentation studies and artificial intelligence, dialogue protocols and models for persuasion dialogue have been developed that start from the assumption that argumentation and criticism are closely interwoven [1, 2]. Thus criticism seems not only to lie at the origin of argument, but also to pervade the whole argumentative procedure.

But then, there is not just one kind of criticism. Merely expressing critical doubt is certainly different from expressing an opposite point of view, and expressing such a point of view is again different from arguing for that point of view. All three are different from raising specific objections against a point of view, or against an argument, or against parts of an argument, or against the arguer, or against the circumstances in which the argument has

been presented. This paper purports to contribute to a systematic classification of these and other kinds of critical reaction and thus to contribute to the dialectical approach to argumentation. In this, others, such as Aristotle [3], Finocchiaro [4], Pollock [5], Govier [6] and Walton [7] have preceded us, and we have ourselves each attempted to contribute to this enterprise as well [8, 9].

In this paper, we deal with the term ‘criticism’ in the sense in which the term pertains to negative evaluations, rather than in a sense that also pertains to positive evaluations. Nevertheless, such criticism can itself be called *constructive* when making valuable contributions to a discussion. We aspire to discuss negative critical reactions in a wide sense, encompassing such criticisms as pertain to (expressions of) propositions, arguments, parts of arguments, and (the applications of) argument schemes, as well as to arguers and institutional circumstances, and relating to such issues as understandability, admissibility, validity, appropriateness, reasonableness, consistency, timeliness, or offensiveness. But we shall not discuss such aspects of critical reactions as fail to contribute to the contents of an argumentative exchange. Thus one could ‘critically react’ to an opponent by grabbing his shoulders and shaking him gently. Would this add content to the exchange? Of course, it might. If in some culture or in some special circumstances, this would be the way to express that one disagrees with the opponent’s point of view, it would as such add some content and be among the critical reactions we intend to cover; however, the circumstance that the expression of disagreement is performed by grabbing and shaking, rather than by a speech act, will not be part of our concerns. And then, the grabbing and shaking may also fail to express anything that must be taken into account as a part of the argumentative exchange, and thus fail to be part of our concerns altogether. From now on, we shall use the term ‘critical reaction’ exclusively for those (aspects of) reactions that do contribute to an argumentative exchange (dialogue).

It should be mentioned that not all reactions in dialogue are critical. Reactions of agreement or acceptance, or requests to grant a concession would not count as such. The same holds for elucidations and explanations of earlier contributions, and indeed for arguments offered in response to criticism.² What is missing in

² There are more moves that are not critical. Consider for example the move of calling into question in the confrontation stage of a critical discussion, as specified by van Eemeren and Grootendorst [10]. Calling into question may sound as if it implies some criticism, but actually the move by which a party in the confrontation stage of a critical discussion *calls into question* the standpoint expressed by the interlocutor is merely aimed at expressing a neutral stance towards the standpoint and not directly at getting the other side to make repairs [10, p. 136]. Therefore, calling into question in the confrontation stage is not yet really critical. However, the move in the opening stage by which the antagonist *challenges* the protagonist to defend his standpoint [10, p. 137] and the move in the argumentation stage by which she *attacks* a standpoint of the protagonist in the sense of *posing a request for an argument* [10, p. 143–4], are clearly critical reactions, as we understand them.

¹ University of Groningen, The Netherlands, email: e.c.w.krabbe@rug.nl and j.a.van.laar@rug.nl

these reactions is a negative evaluation of the move they react upon or at least a suggestion that such a negative evaluation may be forthcoming. One might stretch the concept of critical reaction to the extent that an elucidation of one's earlier contribution would count as criticism of a request for elucidation, and that arguments would count as criticisms of doubts or requests for arguments. One might also claim that acceptance of a statement is a criticism of that statement as being superfluous, since one agrees. Taking this line, all reactions in dialogue could be said to be critical in some sense. We shall not go that far, but exempt from the realm of critical reactions those reactions that merely comply with the requests (to accept, to elucidate or to argue) contained in the move one reacts upon. We do so because of the lack of obviousness of the negative evaluation content of such reactions, if any.³

1.2 The approach in this paper

Rather than straightforwardly heading towards a general classification of types of critical reaction – based upon a division of genera into species – we shall attempt to characterize critical reactions in terms of four parameters or factors (Section 2): the *focus* of a critical reaction (Section 2.1), the *norm* appealed to in a critical reaction (Section 2.2), the illocutionary *force* of a critical reaction (Section 2.3), and the *level* at which a critical reaction is put forward (Section 2.4). Each parameter can take several values, which are characteristic features of critical reactions of certain types.

Thus, the parameter of *focus* allows a critical reaction to be characterized by its focus on (a part of) a move or contribution of a particular kind by the interlocutor; for instance, on the conclusion of an (elementary) argument, or on one of its premises, or on its connection premise. In fact any contribution or part of a contribution in an argumentative exchange can be at the focus of some critical reaction; critical reactions themselves not excluded.

The parameter of *norm* allows a critical reaction to be characterized by a norm appealed to in the criticism, for instance a rule of critical discussion that the critic claims to have been violated. But violation of norms is not the only ground for criticism, nor is a charge of norm violation the only way norms are appealed to in critical reactions. The norm may also be appealed to merely because the criticism puts one's interlocutor under some kind of obligation, as for instance when a critic expresses critical doubt vis-à-vis a standpoint taken by his opponent.

The parameter of *force* allows a critical reaction to be characterized by the illocutionary force of the speech act used. Thus, critical reactions may come forth as directives, for instance as recommendations, requests, challenges, or commands, but also

as assertives, for instance as accusations, and of course as arguments.

The parameter of *level* allows a critical reaction to be characterized by the level at which it is put forward. A critical reaction can aim at eliciting a response from the proponent that contributes directly to the construction of the proponent's case, and thus constitute a ground level move. Alternatively, a critical reaction can belong to a dialogue about the ground level dialogue and thus aim at influencing the course of the latter dialogue, and only in that indirect manner contribute to the construction of the proponent's argumentation, thus constituting a meta-level move.

By examining these parameters, we attempt to contribute to a systematic conceptual analysis of the various ways of criticism. An inventory of the distinct kinds of critical reactions is important, for example when trying to understand an argumentative discourse. But also when developing models or protocols for reasonable persuasion dialogue we are in need of a theoretically motivated classification of critical reactions. After having expounded and illustrated in Section 2 the main features of our system of characterizing critical reactions, based upon Van Laar [9], we shall perform a first test of the system by applying it, in Section 3, to three contemporary approaches to critical reactions.

2. FOUR PARAMETERS

As explained above, we expect that each fully developed and articulated critical reaction can be characterized in terms of four parameters or factors: focus, norm, force, and level. In the case of a particular critical reaction, each parameter will take on a specific value (or, equivalently, each factor will be specified by a specific feature of the critical reaction). We shall deal with these parameters in turn.

2.1 Focus

Each critical reaction has a *focus*, which functions as a precondition for a critical reaction of a particular type [cf. 12]. This may be a focus on a move of a particular type, or on a special part of a move, or on a sequence or combination of moves, put forward by the interlocutor, possibly reconstructed by the critic. Because one can take a critical stance towards any kind of contribution, each type of speech act in an argumentative exchange can be at the focus of a critical reaction. What is more, an argumentative move can be seen as having four aspects: it expresses a particular *proposition*, by employing a particular *locution* put forward with a particular illocutionary force, by a particular *person*, within a particular *situation*. So, the focus of a critical reaction, besides being aimed at a particular kind of speech act, can be *propositional*, *locutional*, *personal* or *situational* in character. We shall first list the most prominent kinds of focus and then discuss the aspects.

First, a critical reaction can focus on (parts of) an elementary argument as reconstructed by the critic. An elementary argument is an illative core of a (possibly more complex) argument, having just one justificatory step. It contains a standpoint (or conclusion) and a set of premises (reasons) containing exactly one connection premise [cf. 13, p. 128]. The connection premise is a conditional statement, having the conjunction of the other premises as its

³ Our notion of 'critical reaction' is both wider and narrower than what Van Eemeren et al defined as the speech act of criticism, the essential condition of which is: 'The speaker, S, says or does something that counts as a negative evaluation of the actions or attributes of the target, T' [11, p. 109]. Unlike a critical reaction, as we understand that term, this speech act of criticism need not be a reaction to a dialogue move. Moreover, this speech act seems to exclude requests for arguments or clarifications, which do not by themselves count as a negative evaluation but merely allude to the possibility that such a negative evaluation may result.

antecedent and the standpoint as its consequent, which – within an argumentative context – expresses the commitment to accept the standpoint as soon as one has accepted the reasons in the antecedent. Often, the connection premise remains implicit, and in such cases the procedure for making it explicit is straightforward.⁴

One of the parts of an elementary argument a critical reaction can focus on is the standpoint advanced by the protagonist. This may happen before the elementary argument has been advanced – and in fact elicit the argument. Such a critical reaction may be focused on an expression of an opinion by the interlocutor, whether this expression has been marked as a standpoint or not (if not, the criticism will turn the expression of opinion into a standpoint). Of course, critical reactions can also focus on other parts of an elementary argument, or on a combination of parts. Where critical reactions on individual parts of an elementary argument are concerned, a threefold distinction can be upheld: such a critical reaction focuses on a standpoint or on a reason in support of a standpoint (turning that reason itself into a substandpoint), or on a connection premise (on the three ways hypothesis, cf. [7]).

It can be useful to characterize a critical reaction on an elementary argument in more detail as being focused on a special type of reason belonging to a specific kind of argumentation. For instance, a reaction could focus on the ‘normality premise,’ belonging to defeasible arguments, which expresses that circumstances are not exceptional, or it could focus on the ‘desirability premise,’ belonging to the pragmatic argument scheme (which is a kind of practical reasoning), which expresses the desirability of a particular goal.

Second, a critical reaction can focus on a more complex argument, such as a basic argument that is built up from several elementary arguments [cf. 13, p. 129]. This happens when it is pointed out that there occurs a shift in the meaning of a particular term in the course of a chain of arguments, or when it is alleged that a chain of arguments is question begging due to circularity, or when it is shown that various parts of the complex argument are mutually inconsistent. The critic can also charge the arguer of having made mistakes in suppositional arguments: for instance, when the arguer has derived an absurdity after having introduced a supposition to be refuted, but then subsequently misidentifies the responsible premise (see Aristotle [15] in *Sophistical Refutations* 5 on the fallacy of *non causa*, 167b21-36).

Third, the focus of a critical reaction can be on a kind of argumentative move that is different from the presentation of an argument. A challenge, to take an example, can be the focus of a critical reaction when it is alleged that the critic’s challenge is inappropriate due to the critic’s having conceded the proposition at issue at an earlier stage. In a similar vein, one can critically react towards requests for clarification, for example because any further clarification would be superfluous. In such cases, a request can be pictured as a delaying tactic. More in general, a critical reaction can be focused on any kind of critical reaction. But there are also other moves that can be critically reacted to, for instance proposals. When one party, defending a standpoint, proposes a premise that is to function as a shared point of departure, a possible critical reaction by the other party could be that accepting that premise as a starting point would come down to accepting the standpoint. The

critical reaction, in such a case, is aimed at preventing an arguer from begging the question.

Fourth, a critical reaction can focus at a combination of argumentative moves (which could all be different from moves needed for constructing an elementary or complex argument). For example, it could be pointed out that one’s opponent refuses to concede a proposition that is immediately implied by a proposition granted earlier. In that case the criticism focuses on the combination of the present move of refusal and the earlier move of concession.

When focusing on such (parts or combinations of) moves of the interlocutor, the emphasis can be towards one or other of the four aspects of a move. Consider first *propositional critical reactions*. If such a reaction focuses directly on the content of a standpoint or of a reason, it can be called a *tenability criticism*, ‘Why *P*?’ [16, p. 161]; if it focuses on the content of a connection premise, it can be called a *connection criticism*, ‘Why if *P* then *Q*?’ [16, p. 160].

A *locutional critical reaction* focuses on the formulation of a standpoint, reason or connection premise, or of some other contribution. It may either be concerned with unclarity about the propositional content or with unclarity about the illocutionary force of the contribution. In the first case, it aims at getting the speaker to indicate into more detail what proposition he tries to express, ‘What do you mean by *P*?’; or it aims at pressing him to adapt his formulation on some other ground, for example because the terminology is biased, or distasteful. A locutional criticism concerned with propositional content can also focus on a complex argument when pointing out a fallacy of equivocation, or when pointing out the lack of terminological coherence in the antagonist’s set of commitments. In the second case, when the illocutionary force is unclear, a locutional criticism aims at getting clearer about the kind of speech act performed by the other side: is he offering an argument or an explanation? Is this multiple argumentation or coordinative argumentation? Is this a mere concession or a stronger, assertive kind of commitment? Or it can be told to the interlocutor that he has performed an inappropriate kind of speech act: he ought not himself have made a *concession* for he is the protagonist in a unmixed interchange who is not asked to make concessions to defend his standpoint, but to take advantage of concessions by the antagonist to do so.

A *personal critical reaction* ‘attacks’ the person who brought forward an argumentative contribution, for example by saying something like ‘you’re not in a position to argue in favor of (or: against) *P* in a credible way due to a general flaw in your character (or a specific bias, etc.)’.

A *situational critical reaction* can point out that the external circumstances are such that the other side’s contribution is inappropriate: ‘defending this very standpoint in the current societal circumstances enhances violence’, or ‘challenging *P* is impolite and therefore not allowed in this family’ or ‘You shouldn’t argue about Burma; you have never been there.’

2.2 Norm

Each critical reaction appeals to a particular kind of argumentative norm. One can relate to a norm in various ways. One merely *follows* a norm, without appealing to it, when one fulfills the obligations prescribed by the norm.; for example, if, when one is

⁴ The procedure is not complicated by the goal of finding a pragmatic optimum [cf. 10, pp. 117-118], or an appropriate general warrant [cf. 14, p. 214].

supposed to provide an argument if asked to so, and is indeed asked to do so, one provides an argument. One merely *utilizes* a norm, again without appealing to it, when one makes use of a right provided by the norm. For example, one utilizes the norm according to which the parties can take turns, simply by performing one's move when the interlocutor has finished speaking. However, one *appeals* to a norm by putting forward a critical reaction (of a kind that is sanctioned by the norms) in order to put some pressure on the interlocutor to respond in a certain way. So, by challenging a standpoint, the critic is utilizing the freedom rule [10, p. 190] which allows him to challenge, but also, although implicitly, appealing to the obligation-to-defend rule [10, p. 191] in order to press the arguer to present an argument. One appeals to a norm, in the special sense of *emphasizing* it, in case the critic not only appeals to the norm, but is also rubbing it in, meaning that she is more or less clearly conveying the message that her critical reaction is pertinent because of the fact that this norm is operative. So, when the critic puts forward a challenge, and in addition stresses that the arguer is under the obligation to provide an argument, she is quite explicitly emphasizing a burden of proof rule. These two ways of appealing to norms (implicitly and explicitly by emphasizing the norms) are examined in some more detail in the next subsection about the force of a critical reaction. In the remainder of this subsection, however, we shall concentrate on the distinction between three *kinds* of norms, rather than on ways to refer or appeal to them.

First, there are the so-called *rules for critical discussion* (a normative model for persuasion dialogue). These rules mark the distinction between argumentatively reasonable and unreasonable dialogue moves (*fallacies*). A critic may charge an arguer with having violated one of these rules. Such a charge would amount to an appeal to the rule in the sense of emphasizing. Of course the charge may be ill-founded. When a critic appeals to a norm that she considers to be part of the constitution of genuine critical discussion but we do not, her critical reaction must be seen by us as an incorrect appeal to a rule for critical discussion.

Second, there are *norms of optimality*, which mark the distinction between argumentative moves that are really good and those that, though not fallacies, are unsatisfactory in some argumentative respect (*lapses* or *blunders*). For instance, if a proponent can choose between a stronger and a weaker argument, the stronger argument is to be preferred (cf. Krabbe [17] on the discussion rule 'Try to win'). Since lapses or blunders are usually 'advantageous' for one's interlocutor, the latter may leave them unnoticed. But she may also point out that the argument, though not fallacious, is flawed and therefore unconvincing. External observers of an argumentative discussion, however, often appeal to optimality norms to criticize the participants.

Third, there are the so-called *institutional norms*. Argumentative norms that are institutional can be seen as marking the distinction between dialogue moves that are appropriate within the institutional setting, and those which are inappropriate within the setting. In the latter case we may speak of *faults*. In contradistinction to the rules for critical discussion, these norms are not part of the general explication of argumentative reasonableness. However, they do apply in particular types of context, where the participants use argumentation for special purposes that supplement the goal of resolution of a difference of opinion. Van Eemeren and Houtlosser discuss these institutional settings as 'argumentative activities' [18, p. 76-7]. For example, when engaged in legal proceedings, additional rules apply to the

argumentative moves put forward by the participants, for in order for the difference of opinion to be resolved in a manner that is not merely dialectically reasonable but also legally admissible, various additional constraints must have been taken into account. These additional constraints can be emphasized as norms in critical reactions. For example, in a legal context, one could think of a critic charging the other side with proposing a proposition as a shared starting point that has been obtained by legally inadmissible methods. We take the idea of an institution in a broad sense, including rather mundane activities such as having a colloquial conversation, or discussing current affairs, in addition to more formalized activities such as being engaged in a parliamentary discussion, a public debate or a debating contest. Norms to the effect that particular topics are, within certain circumstances, not up for debate, or to the effect that certain character traits or personal circumstances can disqualify a person as a serious participant can be regarded as special norms that characterize some (and not all) argumentative activities.

2.3 Force

A third parameter to be used for characterizing the ways of criticism is that of the illocutionary force of a critical reaction. Conspicuous here are reactions in the form of directives or assertives.

First, a critical reaction, whatever the norm appealed to and whatever the focus, can be put forward as a directive in the form of a request; either for argument or for clarification. Requests for argument (or: challenges) have a propositional focus, 'Why *P*?', whereas requests for clarification have a locutional focus, 'What do you mean by formulation *P*?' In both cases, the request aims at an extension of the argument as constructed at some stage of the dialogue. Requests utilize the rules for critical discussion, and appeal to them in an implicit manner. By filing a request for an argument or a clarification, the critic is capable of pressing the arguer to provide the requested argument or clarification on the basis of certain rules for critical discussion. The implicit, normative appeal of a request for an argument would, if made explicit, yield something like: 'in order for you to fulfill your burden of proof, as laid down in Rule 3 for critical discussion, or Commandment 2 of the code of conduct [10, pp. 139 and 191], you must provide an argument as requested.' The urgency of a request for clarification becomes clear from a similar message, which could be made explicit to yield: 'in order for you to adequately express yourself, as required in Rule 15 for critical discussion or Commandment 10 of the code of conduct [10, pp. 157 and 195], you must provide a clarification as requested.' Normally, the reference to the applied rules remains fully implicit in such requests; however, the normative appeal can be made explicit along the lines indicated, in which case the norms are emphasized, rather than merely appealed to implicitly.

Second, instead of merely requesting an argument or a further explication, a critic can reconstruct and negatively evaluate (a part of) a contribution by the other side, by making an assertion to the effect that there is a flaw of some kind in the interlocutor's contribution. Critical reactions such as these have been dealt with by Finocchiaro as 'active evaluations' [4, p. 339]. When pointing out a flaw, the critic is actively taking part in the discussion about the matters at issue in the criticized contribution by putting forward

a negative evaluation in which she appeals to one or more norms: the flaw needs repair. The critic can do so but nonetheless refrain from alleging that her interlocutor has been unreasonable on the ground of having violated some rule for critical discussion (a norm of the first kind) or inapt on the ground of having violated some institutional norm (a norm of the third kind). In that case the norm appealed to must be of the second kind or of yet some other kind. As will be shown in the next subsection, in such cases any further discussion about whether the evaluation is correct or not, can best be seen as taking place at the ground level of dialogue. In other cases, if the alleged flaw is taken to be a transgression of a norm of the first or third kind, the critical reaction can best be seen as initiating a meta-level dialogue [cf. Krabbe 19].

One prominent way of pointing out a flaw is to deny a proposition that has been expressed or employed by the interlocutor or to assert a proposition that implies the denial. Such denials come in two kinds, depending upon the messages conveyed to the other participant. If party A denies a proposition *P* that has been used by party B, saying ‘not *P*’, this denial can convey the relatively weak message that B will not be able to defend his standpoint that *P* vis-à-vis party A. This so-called *weak denial* is not itself a kind of standpoint that requires a defense when challenged. Instead, it expresses an expectation to the effect that, according to A’s assessment, party B will not be capable of constructing a case for his main standpoint that will turn out to be convincing for A. If requested to *defend* ‘not *P*’, party A can justifiably answer ‘It is not my opinion that *P* is not the case, and therefore I am not willing to present an argument in favour of ‘not *P*’; instead I am evaluating negatively your strategic chances of finding an argument that will convince me.’ A weak denial does, however, come with an obligation for the critic to be open about her considerations that brought her to this assessment: what makes her think that B lacks the means for persuading her? So, there is, instead of a burden of proof, a kind of burden of giving some explanation, be it that this burden will have to be rather limited considering that the critic herself may not have full access to the grounds of her assessment. In short, a weak denial will always be a purely critical move, rather than a constructive one.⁵

A second kind of denial is the *strong denial*. With a strong denial, ‘not *P*’, party A conveys the message that A will be able to defend this denial against B’s critical testing. Such a counterstandpoint does carry a burden of proof, when challenged. So, besides being critical, such a move is constructive, generating a mixed dispute.

If the focus of a weak or strong denial is on the propositional content of the connection premise, the critic is pointing out a justificatory flaw. Such flaws can also be pointed out in ways other than by denials, for example by presenting a counterexample. Methods, other than denials, for pointing out flaws can also be found in critical reactions in which it is alleged that a formulation used by the other side contains biased terms or harmful ambiguities. Or when the evidence is pictured as legally inadmissible, or when it is held that the interlocutor has exceeded the time limit.

Third, when raising a challenge or when pointing out a flaw, party A can choose to accompany this critical reaction by some of

the counter-considerations that party B must take into account when making further decisions as to whether and, if so, how to proceed in his attempts to persuade A of B’s standpoint *P*. These counter-considerations function as A’s *strategic advice* to B. We will provide a few examples. First, a challenge can be accompanied by a consideration that explains to B why A is critically disposed to *P*. The message to B then is that B must adapt his persuasive strategy in such a way that this motive for a critical stance will be defused. Second, a similar analysis can be provided when A gives an argument in favour of taking a neutral position towards *P*, rather than accepting *P*. Third, it has been stated above that weak denials should generally be accompanied by considerations that explain why party B will turn out to be unable to persuade A. But such considerations would of course be overruled if B were to defuse them in some way or other. Hence they provide strategic advice for B. Fourth, strong denials can be accompanied by argumentation. Such argumentation can fulfill two functions: a constructive persuasive function (persuading B of not-*P*), but we refrain from discussing this function since we are concerned with critical, rather than with constructive moves. In the present context it is more to the point to stress the other function, that of providing party B with considerations that must be refuted before party A will retract her critical doubt towards *P*.⁶

2.4 Level

A fourth parameter is that of level. Although it is difficult to draw a borderline, the distinction we have in mind has to do with the directness with which a dialogue move contributes to the argumentation in favour of one of the standpoints adopted in the discussion. Quite direct contributions will be located at the ground level dialogue, while more indirect contributions – moves that are about the dialogue rather than about the issue at hand – are to be located at a meta-level dialogue or at dialogues even higher up in the hierarchy.

Clearly, a move in which a proponent puts forward an argument in favour of a challenged proposition contributes directly to the complex case he is constructing for his conclusion, and so this move is a ground level move. The same applies to the clarification of a part of the argument, for example by explaining what was meant by this or that expression. Requests for further arguments or for clarification of an argument will be seen as quite directly contributing to the argumentation in that the response aimed for is

⁵ This notion of a weak denial resembles what Rescher discusses as ‘cautious denial’ [20, p. 9], which expresses that *P* is not the case ‘for all that you (the adversary) have shown’ [p. 6].

⁶ In Rescher [20] two notions occur that resemble such counter-considerations, namely ‘provisoed denial’ [pp. 6, 9] and ‘weak distinction’ [p.12]. In a provisoed denial of *P* the opponent expresses something of the form: ‘*Q* is the case for all that you’ve shown and *Q* constitutes *prima facie* evidence for not *P*’. Not *P* is a weak denial and we may interpret *Q* as a part of the explanation why the opponent thinks *P* is not (yet) to be accepted. The advice is to defuse *Q* as a counter consideration to *P*. A weak distinction is a way for the opponent to focus on the connection premise of an argument ‘*Q* so *P*’. A weak distinction is expressed by something of the form: ‘For all you’ve shown, both *Q* and *R*, while *Q* and *R* taken together provide *prima facie* evidence for not-*P*’. Again, ‘not-*P*’ is a weak denial and we may interpret *R* to be a part of the explanation why the connection between *Q* and *P* does not apply in the current circumstances. The advice is to defuse *R* as a counter consideration to the connection between *Q* and *P*.

an argument or a clarification. So, these moves are considered to be ground level moves as well.

However, if a party's move deals with the strategy adopted by himself or by the other side, the contribution can be seen as dealing with the standpoints at issue, but only indirectly so. The primary topic is a strategy that has been, can be or should be adopted. So, what we have called weak denials are to be seen as initiating a meta-level dialogue. Similarly, moves offering strategic advice are meta-level moves.

When a party asserts that the other side has transgressed a rule for critical discussion or an applicable institutional norm of some kind, the moves must be seen as being primarily about the legitimacy or appropriateness of part of the preceding dialogue, and thus as initiating and contributing to a meta-level dialogue. So, when the critic puts forward a negative evaluation by charging her interlocutor with having breached a norm, strongly emphasizing the norm, her evaluation will also count as a request for some kind of repair, as is generally the case with pointing out flaws. But in addition, the interlocutor is accused of having put forward a move that hinders or even blocks either the resolution-goal of their discussion (a fallacy) or one of the goals inherent in the institutional activity (a fault). All such charges take place at a meta-level of dialogue.

Charges of faults in the present sense occur for instance when party A points out to party B that defending a certain proposition will have unacceptable social consequences (the charge may of course be unjustified). One may think of the self-fulfilling prophecy that ensues when a prime minister too much stresses its country's economical troubles, or of cases where it is said that our adversaries will profit if anyone would take a critical stance towards a standpoint. Also personal attacks can be seen as charges that the interlocutor has violated an institutional norm, in that case a norm to the effect that for instance the arguer's financial involvement, lack of expertise or insincerity is inappropriate for the kind of discussion at hand. Those personal attacks that are dialectically illegitimate constitute *ad hominem* fallacies.

3. APPROACHES TO CRITICAL REACTIONS

3.1 Finocchiaro on Active Involvement

In Chapter 15 of his *Galileo and the Art of Reasoning*, Finocchiaro discusses fallacies and the understanding and evaluation of reasoning or arguments. Unfavorable (negative) evaluation – which according to Finocchiaro is often more interesting than favorable (positive) evaluation – is called ‘criticism’ [4, p. 332]. On the basis of his study of Galileo's critiques of various Aristotelian arguments (in the *Dialogue Concerning the Two Chief World Systems*) Finocchiaro concludes ‘that it is both possible and effective to evaluate arguments ‘actively’ in the sense that the inferential interrelationships among the propositions involved are tested by reasoning on the level of, and largely in terms of, the object argument and by checking whether what follows from asserted premises is the conclusions drawn in the object argument or other propositions.’ [4, p. 339]. In dialogical terms: the opponent may join the proponent as an arguer as they discuss the ins and outs of an argument put forward by the proponent, and this may be done on what we called the ground level of dialogue.

The focus of the criticisms Finocchiaro discusses is always the argument itself, and more specifically the connection premise, which is the premise that specifies ‘the inferential interrelationships’ between the other premises and the conclusion. The criticisms can take various shapes – spelled out in Chapter 17 [4] as pointing to different kinds of ‘invalidity’, which later were described as six types of fallaciousness [21] – but, to put it bluntly, they all claim that the conclusion does not follow from the premises [21, p. 267]. The differences lie in the arguments offered to substantiate that claim. Here the main categories are formal, explanatory, presuppositional, positive, semantical, and persuasive fallaciousness, which each correspond to a particular way of arguing by the critic.⁷

The dialectical norm appealed to in these criticisms is evidently that the arguments offered by a proponent should be improved until they are such that the conclusion follows from the premises.⁸ Finocchiaro [21] does not explain what ‘following from’ means. In our view ‘following from’ need not be understood in the sense of deductive validity: rather it is a contextual notion referring to there being in a specific context no further objections to an inference step [cf. 4, p. 422-3]. Thus, with respect to the parameter of force, we may notice that the criticisms (when used in dialogue) contain not only a claim and arguments, but also a request to the proponent to improve his argument taking into account the counter-considerations brought forward.

As to the level, we saw that these criticisms could take place at the ground level, but it is not of course excluded that the participants take more distance from the object argument and get involved in a meta-dialogue about ways of arguing. Finocchiaro's use of the term ‘fallaciousness’ may even suggest this, since a discussion about fallacies typically takes place at a meta-level. However, Finocchiaro is quite explicit about preferring the weaker term ‘fallaciousness’ to the term ‘fallacy’ [21, p. 266-7]. So ‘fallaciousness’ need not refer exclusively to a situation where someone is charged with a fallacy (infringement of a rule for critical discussion) and a shift to a meta-dialogue takes place.

3.2 Pollock on rebutting defeaters and undercutting defeaters

We will provide another test of the worth of our four parameters by trying to characterize two notions that Pollock employs [5], and that have proven useful in the area of argument and computation. These two notions, to wit ‘rebutting defeater’ and ‘undercutting defeater’, can be seen as specific kinds of critical reactions. In order to show that to be the case, by applying the four parameters, we first need to transform these notions from Pollock's epistemic perspective to our dialectical perspective.

Given his interest in ‘the construction of a person’ who is capable of reasoning and reasoned action, Pollock chooses to focus on reasoning as the process by which beliefs are inferred from perceptual input and from previously held beliefs. An argument is considered as ‘a record of the state transitions involved in the

⁷ We cannot here discuss these in detail, but these ways of arguing and their use by Galileo form the most intriguing part of Finocchiaro's paper [21].

⁸ This is a dialectification of the simpler norm that in arguments the conclusion should follow from the premises.

agent's reasoning' [5, p. 39]. According to Pollock, most reasoning is defeasible insofar as it allows for rebutting defeaters or undercutting defeaters. Pollock defines a rebutting defeater thus: 'If $\langle \Gamma, p \rangle$ is a *prima facie* [= non deductive] reason [= elementary argument with premises Γ and conclusion p], $\langle \Delta, q \rangle$ is a rebutting defeater for $\langle \Gamma, p \rangle$ iff $\langle \Delta, q \rangle$ is a reason and $q = \neg p$ ', where ' $\neg \phi$ ' is the denial of ' ϕ ' [5, p. 85]. His definition of undercutting defeater is: 'If $\langle \Gamma, p \rangle$ is a *prima facie* reason, $\langle \Delta, q \rangle$ is an *undercutting defeater* for $\langle \Gamma, p \rangle$ iff $\langle \Delta, q \rangle$ is a reason and $q = \neg(\Pi\Gamma \gg p)$ '. In this definition, ' $\neg(\Pi\Gamma \gg p)$ ' is the denial (' $\neg\phi$ ' symbolizes the negation of ϕ) of a conditional proposition (' $\phi \gg \psi$ ' stands for: ϕ would not be true unless ψ were true) with the conjunction of the premises of Γ (symbolized as ' $\Pi\Gamma$ ') as its antecedent and p as its consequent, saying: 'It is not the case that $\Pi\Gamma$ wouldn't be true unless p were true' [5, p. 86]. Suppose one reasons from 'The table looks red to me' to 'the table is red.' Then an undercutting defeater that defeats this argument reasons from 'A red light illuminates this table' to 'It's not the case that the table wouldn't look red to me unless the table were red'.

Given that these two kinds of defeaters are defined at an abstract level, we surmise that they can best be seen as propositional or linguistic entities that can be, but need not have been, employed by a person with the aim of defeating, or criticizing, an argument. A dialectical way of reading the definition of a rebutting defeater would be: 'Where the proponent puts forward " P so Q " as an elementary argument, " R so S " is a *rebutting defeater* if and only if " R so S " is an elementary argument that can be put forward by the opponent, where S is the denial of Q '. Similarly, the idea of an undercutting defeater allows of the following interpretation: 'Where the proponent puts forward " P so Q " as an elementary argument, " R so S " is an *undercutting defeater* if and only if " R so S " is an elementary argument that can be put forward by the opponent, where S is the denial of the connection premise $P \rightarrow Q$ '. Note that in a dialectical context, the critic's use of a rebutting defeater against an arguer's argument ' P so Q ' normally conveys the message that the critic does not consider herself committed to both the proposition ' P ' and the proposition ' $P \rightarrow Q$ ', though she may not dispose of any defeaters for these propositions.

What would be a plausible way of characterizing these critical reactions in terms of the four parameters? Both kinds of defeaters can be seen as critical reactions having at least the force of assertives, and more specifically as argued strong denials by which a critic points out the flaw on the proponent's side of having used a proposition that happens to be false. In both kinds, the strong denial is accompanied by counter-argumentation, which, as we discussed, also functions as strategic advice by giving the considerations to be defused by the proponent. Further, both kinds of defeater are, when put to use, contributions to the ground level dialogue. What distinguishes the two defeaters is first of all that rebutting defeaters are focused on the propositional content of a (intermediate) conclusion or (sub-)standpoint put forward by the proponent, while undercutting defeaters are focused on the propositional content of a connection premise taken for granted by the proponent. As to norms it may be observed that a rebutting defeater and an undercutting defeater appeal to different norms for argumentative dialogue.

A rebutting defeater is a somewhat unspecific kind of critical reaction, in the sense that it implies a negative evaluation of the elementary argument presented by the interlocutor, but without specifying whether the connection premise or one of the other premises is taken to be unacceptable. As we have seen, the

strategic advice to the proponent is to defend his position in a way that meets the considerations of the critic by defusing, that is defeating, the defeater. According to Rule 7 for critical discussion [10, pp. 145-148], the protagonist has developed a successful defense only if his argumentation is ultimately based on propositions that may not be called into question in the discussion, something that can be settled by the parties in the opening stage. But this rule determines the meaning of 'successful defence', rather than stipulating a norm. The implicit norm, however, is clear: try to develop a successful defence (a norm of optimality, see Section 2.2). A rebutting defeater can be seen as appealing to this norm. An undercutting defeater appeals to rules for persuasion dialogue according to which the arguer must employ appropriate and correctly applied argument schemes [e.g. Commandment 8 in 10, p. 194] or accepted logical schemes [e.g. Commandment 7 in 10, p. 193].

3.3 Snoeck Henkemans on complex argumentation in critical discussion

In order to explain how complex argumentation comes about and what functions the various kinds of elements of an argumentation fulfill, Snoeck Henkemans [22] discusses the critical reactions that can occur in a critical discussion (as well as the various ways of responding to these critical reactions). She points out that in the ideal model of a critical discussion, an antagonist can criticize an argument by indicating the lack of acceptability of 'the propositional content of the argumentation' [22, p. 408], or by pointing out that 'the argument does not provide sufficient support for the standpoint' [22, p. 408], or by making it clear that 'he regards the argument as irrelevant to the standpoint' [22, p. 409].

The force of a critical reaction by which the antagonist makes it clear that a premise is unacceptable can, when the reaction has been put forward in a questioning mode, be seen as that of a request for argumentation in defense of a premise of an elementary argument. But then, the antagonist may also offer a counter-argument to the effect that the premise is unacceptable. In that case the critical reaction can be seen as having the force of a strong denial and as being accompanied by argumentation in favor of that denial. Snoeck Henkemans makes it clear that these critical reactions can give rise to a subordinative (or: serial) argumentation on the protagonist's part. So, in both cases, the implicit strategic advice for the proponent is to provide such subordinative argumentation. Further (in both cases), the critical reaction takes place on the ground level of dialogue. As in the case of Pollock's rebutting defeaters, these two kinds of critical reaction appeal to a norm according to which the proponent must make a serious effort to succeed in the discussion.

Snoeck Henkemans distinguishes between two ways of indicating that, as far as the antagonist is concerned, the adduced reasons do not suffice to yield the conclusion. The antagonist may either call their sufficiency into question by asking for more reasons or by raising an objection against the argument's sufficiency. In the former case, the critical reaction can be characterized as being focused on the connection premise and as having the force of a request. In the latter case, the antagonist 'mentions a specific objection that can be seen as an argument in favor of his claim that the amount of support is insufficient' [22, p.

409]. In our terms, this critical reaction can be seen as having the assertive force of a strong denial of the connection premise and as being accompanied by an argument in favor of that denial. Snoeck Henkemans makes it clear that both situations prompt the protagonist to offer a coordinative (or: a kind of convergent) argumentation. In our view, the protagonist could also try to argue that the argumentation given is in fact sufficient, i.e., he could offer subordinative argumentation supporting the connection premise. To put these matters in our terms, the implicit strategic advice for the proponent is to provide either coordinative argumentation or subordinative argumentation. In the case of a request, the advice is to provide cumulative argumentation by adding new evidence, unless the protagonist sees an opportunity to convince the antagonist, by the use of subordinative argumentation, that the argumentation given is in fact sufficient. In case the critic has mentioned an objection, the advice is to provide complementary argumentation by adding a reason that defuses the objection [23, pp. 96-97]. In both cases, the critical reaction takes place on the ground level of dialogue. As in the case of undercutting defeaters, the norm appealed to is that the connection premise must correspond to an appropriate and correctly applied argumentation scheme [e.g. Commandment 8 in 10, p. 194] or accepted logical scheme [e.g. Commandment 7 in 10, p. 193].

Finally, the antagonist may indicate that ‘the argument is irrelevant to the standpoint’ [22, p. 409]. Such a critical reaction focuses, again, on the connection premise. Its force can be merely that of a request for further argumentation to support the relevance. However, the antagonist may also offer a counter-argument against the relevance of a premise or premise set. In that case, the critical reaction has the force of a strong denial and is accompanied by argumentation. Snoeck Henkemans makes it clear that charges of irrelevance occasion the protagonist to offer a new argument in favor of a premise (made explicit) that states the connection between the allegedly irrelevant premise (set) and the main standpoint; the result is subordinative argumentation. In our terms, the implicit strategic advice for the proponent is to provide such subordinative argumentation. In the special case where the antagonist has supported the objection of irrelevance by argument, this argument must be defused by the argument of the proponent. Whether or not it contains an argument, the critical reaction takes place on the ground level of dialogue. The norm appealed to is that the connection premise must correspond to an appropriate and correctly applied argument scheme [e.g. Commandment 8 in 10, p. 194] or accepted logical scheme [e.g. Commandment 7 in 10, p. 193]. A difference between attacking sufficiency and attacking relevance that Snoeck Henkemans alludes to is to be found in the strategic advices they offer. In the case of attacking the insufficiency, the strategic advice is (either to give up or) to repair an elementary argument by adding one or more premises, such that the result is an elementary argument that has an acceptable connection premise or, in our view, to provide a new elementary argument that supports the connection premise of the original argument. In the case of attacking irrelevance, the only options offered by the strategic advice are either to give up or to support the connection premise by argument.

4. CONCLUSION

As has become evident from our discussion of the four parameters and of the notions of criticism used by Finocchiaro, Pollock and Snoeck Henkemans, there exists an enormous variety of critical reactions. These must be taken into account within argumentation studies aimed at the development of norms for argumentation and of practical guidelines for those who wish to engage in argumentative activities, displaying rationality as well as persuasiveness. In order to proceed in these areas we think it to be important to compare notions of critical reactions as they exist within approaches such as formal dialectic, pragma-dialectic and computational approaches in a more or less systematic manner, so as to facilitate the development of a clear and useful inventory of critical reactions. In this paper we took some steps in that direction.

What became clear to us, is that criticisms often constitute subtle argumentative instruments that do not only carry negative messages for the interlocutor, but are often helpful in that they provide various kinds of strategic advice.

REFERENCES

1. Henry Prakken, ‘Coherence and Flexibility in Dialogue Games for Argumentation’, *Journal of Logic and Computation*, **15**, 1009 – 1040, (2005).
2. Simon Parsons, Michael Wooldridge and Leila Amgoud, ‘Properties and Complexity of Some Formal Inter-agent Dialogues’, *Journal of Logic and Computation*, **13**, 347 – 376, (2003).
3. Aristotle, *Topica*, In: Aristotle, *Posterior Analytics, Topica* (pp. 263-739), Harvard University Press, Cambridge, MA and William Heinemann, London, 1976. Translated by E. S. Forster. First ed. 1960. (*The Loeb Classical Library* 391.)
4. Maurice A. Finocchiaro, *Galileo and the Art of Reasoning: Rhetorical Foundations of Logic and Scientific Method*, Reidel, Dordrecht, 1980. (*Boston Studies in the Philosophy of Science* 61).
5. John Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA, 1995.
6. Trudy Govier, *The Philosophy of Argument*, Ed. by John Hoaglund, with a preface by J. Anthony Blair, Vale Press, Newport News, VA, 1999.
7. Douglas N. Walton, Objections, Rebuttals and Refutations. In: Juho Ritola (Ed.), *Argument Cultures* (CD-ROM), Ontario Society for the Study of Argumentation, Windsor, ON, 2010. (Proceedings from the 8th Conference of the Ontario Society for the Study of Argumentation, June 3-6, 2009, University of Windsor.)
8. Erik C.W. Krabbe, Nothing but Objections! In: Hans V. Hansen and Robert C. Pinto (Eds.), *Reason Reclaimed: Essays in Honor of J. Anthony Blair & Ralph H. Johnson* (pp.51-63, references pp. 283-299), Vale Press, Newport News, VA, 2007.
9. Jan Albert van Laar, Commentary on: D. Walton’s “Objections, Rebuttals and Refutations,” In: Juho Ritola (Ed.), *Argument Cultures* (CD-ROM), Ontario Society for the Study of Argumentation, Windsor, ON, 2010. (Proceedings from the 8th Conference of the Ontario Society for the Study of Argumentation, June 3-6, 2009, University of Windsor.)

10. Frans H. van Eemeren and Rob Grootendorst, *A Systematic Theory of Argumentation*, Cambridge University Press, Cambridge, 2004.
11. Frans H. van Eemeren, Rob Grootendorst, Sally Jackson and Scott Jacobs, *Reconstructing Argumentative Discourse*, The University of Alabama Press, Tuscaloosa, AL, 1993.
12. Simon Wells and Chris Reed, Formal Dialectic Specification. In: Iyad Rahwan, Pavlos Moraitis and Chris Reed (Eds.), *Argumentation in Multi-Agent Systems* (pp. 31-43), Springer, New York, 2005. (Proceedings of ArgMAS 2004.)
13. Douglas N. Walton and Erik C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State University of New York Press, Albany, NY, 1995.
14. David Hitchcock, Good Reasoning on the Toulmin Model. In: David Hitchcock and Bart Verheij (Eds.), *Arguing on the Toulmin Model: New Essays in Argument Analysis* (pp. 203-218), Springer, Dordrecht, 2006.
15. Aristotle, *On Sophistical Refutations*. In: Aristotle, *On Sophistical Refutations, On Coming-To-Be and Passing-Away, On the Cosmos*. Harvard University Press, Cambridge, MA, and William Heinemann, London, 1965. Translated by E. S. Forster. First ed. 1955 (*The Loeb Classical Library 400/442*).
16. Erik C.W. Krabbe, Profiles of Dialogue as a Dialectical Tool. In: Frans H. van Eemeren (Ed.), *Advances in Pragma-Dialectics* (pp. 153-167), Sic Sat, Amsterdam and Vale Press, Newport News, VI, 2002.
17. Erik C.W. Krabbe, Strategies in Dialectic and Rhetoric. In: Hans V. Hansen, Christopher W. Tindale, J. Anthony Blair, Ralph H. Johnson and Robert C. Pinto (Eds.), *Argumentation and its Applications* (CD-ROM), Ontario Society for the Study of Argumentation, Windsor, ON, 2001. (Proceedings from the Conference of the Ontario Society for the Study of Argumentation, May 17-19, 2001, University of Windsor, Windsor, Ontario).
18. Frans H. van Eemeren and Peter Houtlosser. Theoretical Construction and Argumentative Reality: An Analytical Model of Critical Discussion and Conventionalised Types of Argumentative Activity. In: David Hitchcock (Ed.), *The Uses of Argument: Proceedings of a Conference at McMaster University* (pp. 75-84), Ontario Society for the Study of Argumentation, Hamilton, 2005.
19. Erik C.W. Krabbe, Metadialogues. In: Frans H. van Eemeren, J. Anthony Blair, Charles A. Willard and A. Francisca Snoeck Henkemans (Eds.), *Anyone Who has a View: Theoretical Contributions to the Study of Argumentation* (83-90), Kluwer, Dordrecht, 2003.
20. Nicholas Rescher, *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*, State University of New York Press, Albany, 1977.
21. Maurice A. Finocchiaro, 'Types of Fallaciousness: Toward a Realistic Theory of Logical Criticism', *Argumentation*, **1**, 263-82, 1987. Reprinted in Maurice A. Finocchiaro, *Arguments About Arguments: Systematic, Critical and Historical Essays in Logical Theory* (pp.128-47), Cambridge University Press, Cambridge, etc., 2005.
22. A. Francisca Snoeck Henkemans, 'Complex Argumentation in a Critical Discussion', *Argumentation*, **17**, 405-419, 2003.
23. A. Francisca Snoeck Henkemans, *Analysing Complex Argumentation: The Reconstruction of Multiple and Coordinatively Compound Argumentation in a Critical Discussion*, Sic Sat, Amsterdam, 1992.

Generating Abstractive Summaries of Conversations through Automatic Argumentative Analysis

Vincenzo Pallotta¹ and Rodolfo Delmonte²

Abstract. Abstractive summarization of conversations is a very challenging task that requires full understanding of the utterance/s contained in a dialog turn, their roles and relationships in the conversation. We present an efficient system, derived from a full-fledged text analysis system, which performs the necessary linguistic analysis of turns in conversations and provides useful argumentative labels to build synthetic abstractive summaries.

1 Introduction

The generation of text summaries from transcribed spoken dialogs occurring between two or more people is considered an open problem in Natural Language Processing (or Computational Linguistics). Voice communications can happen either during face-to-face encounters or when parties are remotely connected and using communication services such as VoIP, teleconferencing or just telephones. In this work, we start from the assumption that voice communication can be transcribed accurately and that it is segmented into turns assigned to each speaking participant. While this assumption might seem restrictive, most corpora of voice communications are of this kind. Moreover, the cost of transcribing speech is far lower than the cost of summarization.

Automatic summarization of conversations would be a very useful tool that allows enterprises in extracting knowledge from many sources (e.g. web forums, social media, meeting records) and integrating this knowledge into corporate knowledge bases for future content-based access. For instance, one could easily find answers to questions like “why this decision was made” or “who rejected the proposal made by X?” or “How the decision of doing X impacted the progress of project Y?”. It is apparent that answering this kind of questions requires a deep understanding of the conversational situation such as its dynamics, the rules of order adopted, the specific language used, and even culture-specific rituals.

1.1 Related Work

The problem of automatic summarization of conversations was initially investigated in the 80’ in the context of several DARPA projects [20] focusing on meeting data. Automatic summarization of meetings has been typically approached in a so-called “extractive” fashion that is by extracting excerpts of the dialogs and by assembling them into a hopefully coherent text [12]. This method has severe limitations due to the intrinsic characteristics of the source data: dialogs are not as coherent as an ordinary narrative text (such as news or scientific articles) and obtaining a coherent text from dialog turns is

practically impossible using the extractive approach. Moreover, the proposed solutions for extractive summarization of meetings have already reached their qualitative upper bounds [17].

Abstractive summarization of dialogs has been recently only partially addressed by some research project among which the European FP6 AMI, and the DARPA’s CALO project. On this side, some advances have been achieved such as extracting (very basic) “dialog acts” [1], detecting “agreement and disagreement” [7, 5], segmenting the meeting into “thematic zones” [8], and, only recently, detecting “decision points” [9, 10], and “action items” [13]. At the University of Twente, also in the framework of the AMI project, a group was investigating on the automatic argumentative structuring of meeting dialogs [18].

Unfortunately, although very related, none of these projects directly address the problem of abstractive summarization of conversations and all the advances on the detection of useful content information from meeting data are mainly exploited to improve extractive summarization, which, as said before, can be only improved up to a given (unsatisfactory) upper bound.

The problem we address here, instead, is that of abstractive summarization of conversations. This problem is apparently much harder to solve than the extractive one since it requires almost the full understanding of the source data [6].

1.2 Argumentative Analysis of Dialogs

We start from transcribed turns in conversations. Our solution to the problem of abstractive summarization of conversations would be learning a general semantic/pragmatic structure for dialogs data that could be specialized for different conversations types (e.g. meetings, debates, focus groups, forums and interviews).

We were looking at pervasive pragmatic phenomena in dialogs and we realized that most of encountered dialogs include argumentative processes. Basically, exchange of opinions, questions-answers, negotiations and collaborative decision-making are at the heart of voice communications. Unfortunately, these phenomena are not sufficiently lexically marked in the dialogs to justify extractive summarization. For instance, a disagreement to a proposal about an issue cannot be summarized as just the turn “no” occurring in the dialog. This action is the result of a complex, sometimes long exchange between two or more participants.

By simply extracting some turns from the dialog is often insufficient to fully render the situation. In contrast, it is necessary to provide a paraphrase where some argumentative terminology is used such as, for instance, “Mr. X disagreed to the proposal made by Mr. Y on the issue raised by Mr. Z”. Our system is capable to recognize a small, but highly relevant set of general argumentative action that

¹ Webster University Geneva, Switzerland, pallotta@webster.ch

² University of “Ca Foscari” - Venice, Italy, delmonte@unive.it

can be used to produce such paraphrases thus creating a meaningful abstractive summary of the analyzed conversations. While the generation of the summaries is quite trivial given the extracted information, we will focus in this paper on the recognition of argumentative events as the core of the technology required to implement an abstractive summarizer of voice communications.

The paper is organized with a presentation of the system in section 2; then some data and an evaluation in section 3. A final section will be devoted to conclusions.

2 Automatic Argumentative Annotation

Computing semantic representations for argumentative annotation of conversations is a particularly hard task which – when compared to written text processing – requires the following additional information to be made available:

- adequate treatment of fragments;
- adequate treatment of short turns, in particular one/two-words turns;
- adequate treatment of first person singular and plural pronominal expressions;
- adequate treatment of disfluency, thus including cases of turns made up of just such expressions, or cases when they are found inside the utterance;
- adequate treatment of overlaps;
- adequate treatment of speaker identity for pronominal coreference;

In addition, every dialog turn must receive one polarity label, indicating negativity or positivity, and this is computed by looking into a dictionary of polarity items. This can subsequently be used to decide on argumentative automatic classification.

2.1 Argumentative structure – issues and theories

As shown by [19], tracking argumentative information from meeting discussions is of central importance for building summaries of project memories since, in addition to the "strictly factual, technical information", these memories must also store relevant information about decision-making processes [14].

We consider the adoption of a deeper structured representation based on argumentation theory. The argumentative structure defines the different patterns of argumentation used by participants in the dialog, as well as their organization and synchronization in the discussion. A conversation is decomposed into several stages such as *issues*, *proposals*, and *positions*, each stage being possibly related to specific aggregations of elementary dialog acts. Moreover, argumentative interactions may be viewed as specific parts of the discussion where several dialog acts are combined to build such an interaction; as for instance, a disagreement could be seen as an aggregation of several acts of reject and accept of the same proposal. From this perspective, we adopted an argumentative coding scheme, the Meeting Description Schema (MDS) from [15]. In MDS, the argumentative structure of a meeting is composed of a set of topic discussion *episodes* (a discussion about a specific topic). In each discussing topic, there exists a set of issue discussion episodes. An issue is generally a local problem in a larger topic to be discussed and solved. Participants propose alternatives, solutions, opinions, ideas, etc. in order to achieve a satisfactory decision. Meanwhile, participants either express their positions and standpoints through acts of accepting or rejecting proposals, or by asking questions related to the current

proposals. Hence, for each issue, there is a corresponding set of proposals episodes (solutions, alternatives, ideas, etc.) that are linked to a certain number of related positions episodes (for example a rejection to a proposed alternative in a discussing issue) or questions and answers.

3 Computing Argumentative Annotations

The core of our solution is based on adapting and extending GETARUNS [2, 3], a system for text understanding developed at the University of Venice. GETARUNS is organized as a pipeline which includes two versions of the system: the Partial and the Deep GETARUNS. The Deep version, used in this work, is equipped with three main modules:

- a lower module for parsing, where sentence strategies are implemented;
- a middle module for semantic interpretation and Discourse Model (DM) construction which is cast into Situation Semantics;
- a higher module where reasoning and generation takes place.

GETARUNS, has a highly sophisticated linguistically based semantic module which is used to build up the DM. Semantic processing is strongly modularized and distributed amongst a number of different submodules which take care of Spatio-Temporal Reasoning, Discourse Level Anaphora Resolution, and other subsidiary processes like Topic Hierarchy which cooperate to find the most probable antecedent of coreferencing and cospecifying referential expressions when creating semantic individuals. These are then asserted in the DM, which is then the sole knowledge representation used to solve nominal coreference. The system has been adapted to cope with some peculiar phenomena in dialogs such as overlap and short turns or fragments. Overlaps are an important component of all spoken dialogue analysis. In all dialogue transcription, overlaps are treated as a separate turn from the one in which they occur, which usually follows it. This is clearly wrong from a computational point of view. For this reason, when computing overlaps we set as our first goal that of recovering the temporal order. This is done because:

- overlaps may introduce linguistic elements which influence the local context;
- eventually, they may determine the interpretation of the current utterance;

For these reasons, they cannot be moved to a separate turn because they must be semantically interpreted where they temporally belong. In addition, overlaps are very frequent. The algorithm we built looks at time stamps, and everytime the following turn begins at a time preceding the ending time of current turn it enters a special recursive procedure. It looks for internal interruption in the current turn and splits the utterance where the interruption occurs. Then it parses it split initial portion of current utterance and continues with the overlapping turn. This may be reiterated in case another overlap follows which again begins before the end of current utterance. Eventually, it returns to the analysis of the current turn with the remaining portion of current utterance.

Fragments and short turns are filtered by a lexical lookup procedure that searches for specific linguistic elements which are part of a list of backchannels, acknowledgements expressions and other similar speech acts. If this procedure succeeds, no further computation takes place. However, this only applies to utterances shorter than 5 words, and should be made up only of such special words. No

other linguistic element should be present apart from nonwords, that is words which are only partially produced and have been transcribed with a dash at the end. The two remaining cases are the following:

- graceful failure procedures for ungrammatical sentences, which might be fullfledged utterances but semantically uninterpretable due to the presence of repetitions, false starts and similar disfluency phenomena;
- failure procedures for utterances that are constituted just by disfluency items and no linguistically interpretable words.

We also implemented a principled treatment of elliptical utterances and contribute one specific speech act or communicative act. They may express agreement/disagreement, acknowledgements, assessments, continuers etc. All these items are computed as being complements of abstract verb SAY which is introduced in the analysis, and has as subject, the name of current speaker.

Finally, Automatic Argumentative Annotation is carried out by a special module of the GETARUNS system activated at the very end of the computation of the each dialog. This module takes as input the complete semantic representation produced by the system. The elements of semantic representation we use are the following ones:

- all facts in Situation Semantics contained in the Discourse Model, which include individuals, sets, classes, cardinality, properties related to entities by means of their semantic indices;
- facts related to spatio-temporal locations of events with logical operators and semantic indices;
- vectors of informational structure containing semantic information at propositional level, computed for each clause;
- vectors of discourse structure with discourse relations computed for each clause from informational structure and previous discourse state;
- dialog acts labels associated to each utterance or turn following ICSI classification;
- overlaps information computed at utterance level;
- topic labels associated to semantic indices of each entity marked as topic of discourse;
- all utterances with their indices as they have been automatically split by the system.

To produce Argumentative annotation, the system uses the following 21 *Discourse Relations labels* computed by GETARUNS in the higher module:

statement, narration, adverse, result, cause, motivation, explanation, question, hypothesis, elaboration, permission, inception, circumstance, obligation, evaluation, agreement, contrast, evidence, hypoth, setting, prohibition.

These are then mapped onto five general argumentative labels:

ACCEPT, REJECT/DISAGREE, PROPOSE/SUGGEST, EXPLAIN/JUSTIFY, REQUEST.

In addition, we use the label DISFLUENCY for all those turns that contain fragments which are non-sentences and are not semantically interpretable. Full details about the argumentative annotation algorithm are provided in [4]. The argumentative annotation algorithm is outlined as follows:

1. It recovers Dialog Acts for each dialog turn as they have been assigned by the system;
2. It recovers Overlaps as they have been marked during the analysis;

3. It produces an Opinion label which we call Polarity, which can take one of two values: Positive or Negative;
4. It produces a list of Hot Spots (i.e. a set of turns in sequence where the interlocutors overlap each other frequently) and builds up Episodes (i.e. a set of turns in which a single speaker "arguments" his/her topics without interruption). Episodes may occasionally be interrupted by overlaps or by short continuers, backchannel or other similar phenomena by other speakers without however grabbing the floor;
5. It assigns a set of argumentative labels, one for each clause. The system then chooses the label to associate to the turn utterance from a hierarchy of argumentative labels graded for Pragmatic Relevance which establishes that, for instance, Question is more relevant than Negation, which is more relevant than Raising an Issue, etc.

We are also able to evaluate the degree of collaboration vs. competitiveness of each participant in the conversation and make a general statements like the ones we include here below taken from the summaries produced for a dialog of our test corpus.

3.1 Abstractive Summary Generation

We present here the structure and the core content of the abstractive summary (Memo) that is generated automatically by the system. The Memo is made of five sections of paraphrased summary, touching upon different conversational features that are relevant to the general understanding of the conversation.

The first section reports general information regarding the *participants and their level of involvement* in terms of number of turns and actual presence in the conversation. In this meeting, for instance, two of the participants joined late and contributed less.

GENERAL INFORMATION ON PARTICIPANTS

The participants to the meeting are 6. Participants less actively involved are Adam and Andreas who only intervened respectively for 9 and 72 turns.

The second section reports about *levels of interactivity* and is based on the use of time stamps and the computation of overlaps. We compute number of turns effectively held by each speaker, then the number of overlaps done by each interlocutor. These numbers allow us to define levels of competitiveness in the conversation.

LEVEL OF INTERACTIVITY IN THE DISCUSSION

The speaker that has held the majority of turns is Don with a total of 512 turns, followed by Morgan with a total of 456 turns, followed by Jane with a total of 263.

The speaker that has undergone the majority of overlaps is Morgan followed by Jane. The speaker that has done the majority of overlaps is Morgan followed by Jane. Morgan is the participant that has been most competitive. This participant: Andreas, only intervened after turn no. 1091

The third section reports lexical chains that highlight the *main topics* of the conversation.

DISCUSSION TOPICS AND LEXICAL CHAINS

The discussion was centered around the following topics: format, file and tool.

The main topics have been introduced by the second most important speaker of the meeting.

The participant who introduced the main topics in the meeting is: Morgan.

The most frequent entities in the whole dialogue partly coincide with the best topics, and are the following: *format*, *utterance*, *representation*, *annotations*, *information*, *p_file*, *equals*, *ATLAS*, *prosodic*, *database*, *data*, *diff*, *a_p_i_*, *tool*, *sequence*, *files*, *external*, *boundaries*, *whole*, *versions*, *timeline*, *text*, *speaker*, *rising*, *file_format*, 'Adam', *transcripts*, *prosody*, *phrase*, *perl*, *lattice*, *idea*, *channels*.

The fourth section highlights *speakers' attitudes* towards topics and they are computed from pragmatic information collected from idiomatic expression and sentiment analysis. It uses five values: Positive, Negative, None, Acceptance, Suspension, where Suspension indicates an expression of doubt or some modality and None refers to simple declaratives without opinion specified.

SPEAKERS ATTITUDES TOWARDS THE TOPIC "format"
We report here below the lexical chains associated to the most frequent entities and topics discussed in the dialogue. The lexical chain headed by the word "format", is present in the following turns, headed by a given predicate and introduced by a given speaker, with a given attitude.

TNo. 45, Pred. develop, Spk. Don, Att. Positive
TNo. 205, Pred. have, Spk. Don, Att. Positive
TNo. 230, Pred. like, Spk. Jane, Att. Positive
TNo. 279, Pred. see, Spk. Morgan, Att. Negative
TNo. 300, Pred. describe, Spk. Don, Att. None
TNo. 345, Pred. have, Spk. Jane, Att. Positive
TNo. 490, Pred. have, Spk. Don, Att. Negative
TNo. 523, Pred. be, Spk. Don, Att. Positive
TNo. 569, Pred. find, Spk. Morgan, Att. Suspension
TNo. 583, Pred. be, Spk. Don, Att. Suspension
TNo. 605, Pred. do, Spk. Jane, Att. Positive
TNo. 773, Pred. add, Spk. Morgan, Att. Suspension
TNo. 775, Pred. be, Spk. Don, Att. Suspension
....
TNo. 1207, Pred. be, Spk. Jane, Att. Acceptance

The last fifth section highlights the *argumentative content* of the conversation, is made of two parts. The first is an overview of the argumentative activity of participants; the second is an outline of the argumentative structure detected for each thematic episode. Each turn is paraphrased according to its recognized argumentative function and linked through a discourse relationship to previous turns.

ARGUMENTATIVE CONTENT

The following participants: *Andreas*, *Dave*, *Don*, *Jane* and *Morgan*, expressed their dissent 58 times; however *Dave*, *Andreas* and *Morgan* expressed dissent in a consistently smaller percentage.

The following participants: *Adam*, *Andreas*, *Dave*, *Don*, *Jane* and *Morgan*, asked questions 55 times.

The remaining 1204 turns expressed positive content by proposing, explaining or raising issues. However *Adam*, *Dave* and *Andreas* suggested and raised new issues in a consistently smaller percentage.

The following participants: *Adam*, *Andreas*, *Dave*, *Don*, *Jane* and *Morgan*, expressed acceptance 217 times.

EPISODE ISSUE No. 1 In this episode we have the following argumentative exchanges between *Don*, *Jane* and *Morgan*:
Don raises the following issue: [um so - so it definitely had that as a concept.]

then he, overlapped by *Morgan*, continues: [so th ... it has a single timeline.]

Morgan accepts the previous explanation: [uhhuh,.]

Then *Don* elaborates the following explanation: [and then you can have lots of different sections.]

EPISODE ISSUE No. 2 In this episode we have the following argumentative exchanges between *Don*, *Jane* and *Morgan*:

Jane, overlapped by *Dave*, suggests the following explanation [actually we - we use a generalization of the - the sphere format.] [yeah so there is something like that.] [but it's um probably not as sophist...]

Don asks the following question [well what does H.T.K. do for features?]

then he, overlapped by *Jane*, continues [or does it even have a concept of features?]

Jane provides the following answer [they h... it has its own ...]

then she, overlapped by *Morgan*, continues [I mean Entropic has their own feature format that's called like s... s.d. or some s... s.f. or something like that.]

Morgan accepts the previous explanation [yeah.]

EPISODE ISSUE No. 3 In this episode we have the following argumentative exchanges between *Don*, *Jane* and *Morgan*:

Don, overlapped by *Morgan*, provides the following explanation [you know we'll do something where we - some kind of data reduction where the prosodic features are sort ... uh either at the word level or at the segment level.]

Morgan agrees [right,.]

then he continues [or - or something like that.] [they're not going to be at the phone level and they're n- ... not going to be at the frame level when we get done with sort of giving them simpler shapes and things.] [and so; the main thing is just being able ...] [well I guess the two goals.]

Don disagrees with the previous explanation [um one that *Chuck* mentioned is starting out with something that we don't have to start over that we don't have to throw away if other people want to extend it for other kinds of questions.]

Morgan agrees [right,.]

It's important to notice the use of pronouns in the paraphrase and the insertion of discourse markers to indicate the continuation by the same speaker. Formulating adequate paraphrases may sound awkward sometimes and this is due to the inherent difficulty of the task.

4 Experimental Results

We started using the system to parse ICSI corpus of meetings [11] and we realized that the semantic representation and the output of the parser were both inadequate. By looking at errors, we became aware of the peculiarities of spoken dialog texts such as the ones made available in ICSI corpus, and to the way to implement solutions in such a complex system. These dialogs are characterized by the need to argument in an exhaustive manner the topics to be debated which are the theme of each multiparty dialog. The following are some statistics that characterize the corpus:

- percent of turns made of one single word: 30%
- percent of turns made of up to three words: 40%
- average number of words x turn overall: 7
- average number of words x turn minus short utterances: 11

These values correspond to those found for PennTreebank corpus where we can count up to 94K sentences for 1M words – again 11 words per sentence. In analyzing ICSI, we found turns with as much as 54 words depending on the topic under discussion and on the people on the floor. The system has been used to parse the first 10 dialogs of the ICSI corpus for a total number of 98523 words and 13803 turns. We had to tune all the modules and procedures carefully. In particular, the module for argumentative automatic classification was incrementally improved in order to cover all conventional ways to express agreement.

We had one skilled linguist to provide a turn level annotation for argumentative labels: we don't have any agreement measure in this case, even though we expect the annotation to be in line with current experiments on the same subject [16]. In Table 1 we report data related to the experiment of automatic annotation of argumentative categories.

	Correct	Incorrect	Total Found
Accept	662	16	678
Reject	64	18	82
Propose	321	74	395
Request	180	1	181
Explain	580	312	892
Disfluency	19		19
Total	1826	421	2247

Table 1. Accuracy of automatic argumentative annotation

On a total of 2304 turns, 2247 have received an argumentative automatic classification, with a Recall of 97.53%. We computed Precision as the ratio between Correct Argumentative Labels/Found Argumentative Labels, which corresponds to 81.26%. The F-score is 88.65%.

5 Conclusions

We have presented work carried out to extend and adapt a language analysis system designed for text understanding in order to make it fit for argumentative analysis as the core technology for abstractive summarization of conversations. Summaries are generated by aggregating the statistics of the analysis, by highlighting the topic discussed and by segmentating the content of conversation into argumentative episodes. The turns of the episodes are then paraphrased according to their argumentative functions highlighting their discourse relationship (e.g. replies and elaborations).

For this purpose, we implemented a module that computes Argumentative automatic classification labels out of a small set, on top of discourse relations and other semantic markers determined by the semantic component of the system. The system has been evaluated for the argumentative classification module and results are very encouraging.

5.1 Future work

Of course we intend to complete the analysis of all dialogues contained in the ICSI corpus and refine our algorithms. In particular we still need to work at the level of DECISION labelling, and to improve the discrimination of really argumentative from pragmatically irrelevant utterance, a choice that in some cases is hard to make on an automatic basis. We would like to formally assess the robustness

of the system by applying the algorithm to the output of automatic transcription systems and evaluate its degree of degradation.

References

- [1] J. Ang, Y. Liu, and E. Shriberg, 'Automatic dialog act segmentation and classification in multiparty meetings', in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (2005).
- [2] R. Delmonte, *Computational Linguistic Text Processing Logical Form, Semantic Interpretation, Discourse Relations and Question Answering*, Nova Science Publishers, New York, 2007.
- [3] R. Delmonte, *Computational Linguistic Text Processing Lexicon, Grammar, Parsing and Anaphora Resolution*, Nova Science Publishers, New York, 2009.
- [4] R. Delmonte, A. Bristot, and V. Pallotta, 'Towards automatic dialogue understanding', in *Proceedings of the International Conference on Spoken Communication (GSCP09)*, Naples, Italy, (2009).
- [5] M. Galley, K. McKeown, J. Hirschberg, and E. Shriberg, 'Identifying agreement and disagreement in conversational speech: Use of bayesian networks to model pragmatic dependencies', in *Proceedings of 42nd Meeting of the Association of Computational Linguistics (ACL)*, (2004).
- [6] U. Hahn and I. Mani, 'The challenges of automatic summarization', *IEEE Computer*, **33(11)**, 29–36, (2000).
- [7] D. Hillard, M. Ostendorf, and E. Shriberg, 'Detection of agreement vs. disagreement in meetings: Training with unlabeled data', in *Proceedings of HLT-NAACL*, (2003).
- [8] P. Hsueh and J.D. Moore, 'Automatic topic segmentation and labeling in multiparty dialogs', in *Proceedings of IEEE Spoken Language Technology Workshop*, pp. 98–101, (December 2006).
- [9] P. Hsueh and J.D. Moore, 'Automatic decision detection in meeting speech', in *Proceedings of the workshop on Machine Learning for Multimodal Interfaces (MLMI)*, pp. 168–179, (2007).
- [10] P. Hsueh and J.D. Moore, 'What decisions have you made: Automatic decision detection in conversational speech', in *Proceedings of HLT-NAACL*, (2009).
- [11] A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Piskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters, 'The icsi meeting corpus', in *Proceedings of IEEE/ICASSP*, pp. 364–367, (2003).
- [12] G. Murray, S. Renals, and J. Carletta, 'Extractive summarization of meeting recordings.', in *Proceedings of the 9th European Conference on Speech Communication and Technology*, pp. 593–596, (2005).
- [13] J. Niekrasz, M. Purver, J. Dowding, and S. Peters, 'Ontology-based discourse understanding for a persistent meeting assistant', in *Proceedings of the AAAI Spring Symposium Persistent Assistants: Living and Working with AI.*, (2005).
- [14] V. Pallotta, H. Ghorbel, A. Ballim, A. Lisowska, and S. Marchand-Maillet, 'Towards meeting information systems: Meeting knowledge management', in *Proceedings of ICEIS*, pp. 464–469, (2004).
- [15] V. Pallotta, H. Ghorbel, G. Coray, and P. Ruch, 'An argumentative annotation schema for meeting discussions', in *Proceedings of the LREC 2004 international conference*, (2004).
- [16] V. Pallotta, V. Seretan, and M. Ailomaa, 'User requirements analysis for meeting information retrieval based on query elicitation', in *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL 2007)*, pp. 1008–1015, (2007).
- [17] K. Riedhammer, D. Gillick, B. Favre, and D. Hakkani-Tür, 'Packing the meeting summarization knapsack', in *Proceedings of the 9th International Conference of the ISCA (Interspeech 2008)*, pp. 2434–2437, (2008).
- [18] R.J. Rienks, D. Heylen, and E. van der Weijden, 'Argument diagramming of meeting conversations', in *Proceedings of the Workshop on Multimodal Multiparty Meeting Processing at the 7th International Conference on Multimodal Interfaces (ICMI)*, pp. 85–92, (2005).
- [19] D. Rosemberg and J.A.A. Silince, 'Common ground in computer-supported collaborative argumentation', in *Proceeding of the Workshop on Computer Supported Collaborative Argumentation for Learning Communities (CSCL99)*, (1999).
- [20] K. Zechner, 'Automatic summarization of open-domain multiparty dialogues in diverse genres', *Computational Linguistics*, **28(4)**, (2002).

Handling explanation in operational contexts

Patrick Saint-Dizier¹

Abstract. In this document, we present some preliminary facets of a model for explanations and their realizations in language. We introduce the notion of explanation function realized by means of explanation schemes. We show how these schemes are realized, what purpose they serve and on what linguistic basis they can be recognized.

1 Introduction and Context

Explanation and its relations to Language is a vast area of investigation. Explanation analysis is quite well developed via writing recommendations in didactics [3]. In artificial intelligence, it is often centered around the notion of argumentation [4], [26], [32], but argumentation is just one (major) facet related to explanation. Let us also note studies on causality, and an emerging field around negotiation [4] and explanation in multi-agent systems. Two decades ago, explanations were used to produce natural language outputs for experts systems, often from predefined templates [29]. Quite interesting principles have emerged from this research. In ergonomics and cognitive science, the ability for humans to integrate explanations about a task (possibly via a guidance system) when they perform that task is measured as such and in relation with the document properties (typography, pictures) [18], [20], [7].

In linguistics, a lot of efforts have been devoted to the definition and the recognition of discourse frames [34], [23] [28] and rhetorical relations [21], [19], which are, for some of them, central to explanation. However, we now observe a proliferation of rhetorical relations, which, in general, turn out to be quite difficult to recognize from language marks since they involve some pragmatic interpretation. Finally, explanation is a field which is investigated in pragmatics (e.g. cooperative principles, dialogue principles) [25] and in philosophy (e.g. rationality and explanation, phenomenology of explanation, causality, etc.) [16], [36] and [10].

Explanations are in general structured with the aim of reaching a goal [1], [6]. Explanations are often associated with a kind of instructional style, which ranges from injunctive to advice-like forms. Procedures of various kinds (social recommendations, as well as DIY, maintenance procedures, health, didactic texts) form an excellent source of corpus to observe how explanations are constructed, linguistically realized and what goals they target. Indeed, in procedures, style is often quite straightforward [12], procedures being essentially oriented towards action: there must be little space for inferences that may lead to misinterpretations, hence the need of explanations. Explanation occurs also in goal-driven but non procedural contexts, for example, as a means to justify a decision as in legal reasoning, or as a way to explain the reasons of an accident in accident reports. Explanation may also be associated with various pragmatic effects (irony,

emphasis, dramatisation, etc.) for example in political discourse. In each case, explanation has a goal-oriented structure [9], [30].

Explanation analysis and production is essential in opinion analysis to make more explicit how a certain opinion is supported [17], it is also essential in question answering when the response which is produced is not the direct response: the user must then understand why the response provided is appropriate. Finally, it is central in a number of types of dialogues, clarification situations, persuasion strategies, etc.

Our main goal is to identify a number of prototypical explanation schemes (See [32] for argumentation forms), their linguistic basis (e.g. prototypical language marks or constructs), and to categorize their communicative goals. We aim at identifying the language and pragmatic means which are used to support, motivate and convince the reader. We will basically concentrate here on explanations associated with procedures. We consider a large variety of procedures, from large public ones (cooking, DIY), to professional ones (maintenance, health), with a large diversity of target readers and application domains. Our aim is to have a closer analysis of the forms and organizations explanations may take in operational contexts (installation, production, maintenance procedures) compared to general studies on discourse. Besides the theoretical dimension, the main aim is to have a better and concrete analysis of the difficulties, the gaps, the risks that one may encounter while realizing a procedure. This is a major issue in e.g. maintenance or production to prevent risks.

In this short paper, we first introduce the notion of **explanation function** that specifies communicative goal of explanations. Explanation functions are abstract constructs which are realized in language via **explanation schemes**. An important feature of explanation is that it is transcategorical: it includes syntactic and lexical semantics factors, as well as typographic and pragmatic factors. The boundaries of explanations are fuzzy: e.g. a whole procedure can indeed be viewed as an explanation, or just e.g. its illustrations, advice and warnings.

2 Corpus analysis and annotation

Our initial corpus is composed of 2000 procedural texts in French, from 24 very diverse domains. The backbone of procedures is based on the 'goal (titles) - instructions' structure. However, this structure is associated with a large diversity of explanations meant to guide, motivate, evaluate, warn users. The explanation structure is therefore crucial for the understanding of a procedure and its correct execution.

In order to identify explanation functions and general explanation principles and goals, we have first carried out an 'informal' annotation of the corpus. This was realized by 3 Master students in linguistics on 62 different texts (about 85 pages), with the same training and annotation instructions. This is a difficult task: identifying and categorizing rhetorical relation is almost never straightfor-

¹ IRIT-CNRS Toulouse France, email: sstdizier@irit.fr

ward. To help annotators, the task was realized on the outputs of our <TextCoop> system (see demo paper in this volume), that tags instructions, goals, various discourse markers, arguments, and within an instruction: some manners, instruments, conditional expressions and simple causal forms. The kappa test indicates an agreement level of 80 types of elements (rhetorical is more difficult than arguments for example). We then had a discussion and constructed an annotated corpus which is the result of this discussion, where disagreement have been partly or totally resolved. This helped us refine our definitions for the different objects to tag.

The tags to use were first given a priori, from our experience of the domain (<TextCoop> started 5 years ago) and then gradually enriched when tagging the corpus. After an adaptation period, annotators could do an homogeneous task. The principles are to tag first the most crucial and less ambiguous structures (e.g. reformulation) which are most of the time indicated via clear linguistic marks. Once these settled, we moved on to more subtle ones, proceeding by contrast with the previously defined ones.

We basically have the following preliminary set of tags corresponding to frequently encountered structures (called hereafter Elementary Explanation Structures: EES):

- rhetorical: elaboration, illustration, reformulation, result, contrast, analogy, evidence, encouragement, hint, evaluation.
- arguments: warning, advice (very few threats or rewards, where the author is involved), these often have the form of a causal expression, formally: do X 'because/otherwise' Z, with quite a large variety of causal marks (Fontan et al. 2008),
- conditions: involve at least 2 structures: condition and then - else, and the assumption structure, which is an hypothetical statement.
- cause: which involve a statement and a structure: cause or consequence. We limit annotations to trans-sentential causal expressions, i.e. those operating over instructions,
- concession: statement followed by a concession statement,
- goal expression: purpose,
- frame: circumstance, and some propositional attitudes such as: commitment, authority.

Annotations can be embedded, leading to complex structures. In our corpus, annotations are realized in XML with some attributes, such as e.g. the force (weight) of arguments, and meta-annotations such as the certainty of the annotator. A given structure may receive several annotations in case of ambiguity, overlap or multiple functions. The / below is an 'or'.

An example, in readable form, from didactics, is the following (most marks produced by our <TextCoop> system are omitted, only EES are given, instructions appear on new lines):

[*procedure* [*purpose* Writing a paper: [*elaboration* Read light sources, then thorough]]

[*assumption/circumstance* Assuming you've been given a topic,]
[*circumstance* When you conduct research], move from light to thorough resources [*purpose* to make sure you're moving in the right direction].

Begin by doing searches on the Internet about your topic [*purpose* to familiarize yourself with the basic issues;]

[*temporal-sequence* then] move to more thorough research on the Academic Databases;

[*temporal-sequence* finally], probe the depths of the issue by burying yourself in the library.

[*warning* Make sure that despite beginning on the Internet, you don't simply end there.

[*elaboration* A research paper using only Internet sources is a weak

paper, [*consequence* which puts you at a disadvantage...]]

While the Internet should never be your only source of information, [*contrast* it would be ridiculous not to utilize its vast sources of information. [*advice* You should use the Internet to acquaint yourself with the topic more before you dig into more academic texts.]]

Besides its illustrative purpose, this example shows the difficulty of the task and its subjectivity. Nevertheless, by going through a reasonable amount of pages of annotated texts, we can discover regularities related to (1) the communicative goals and (2) the ways these goals are realized. The analysis is here entirely manual. Annotated corpus is also used to test the implementation.

3 Explanation Functions

From our corpus investigations, we can propose a first, global classification of the functions realized by explanations. Our view on explanation is quite broad, with the introduction of quite a large number of functions that play a central role in explanation, even if their associated language realization may play several roles, besides explanation (e.g. guidance or expected result below). To carry out this task, our strategy was to identify the underlying communicative aims and to categorize them. Again, this is somewhat intuitive, but it was realized with the contribution of annotators, identifying, for each annotated fragment, what is its communicative goal(s). Within an operational context, explanation functions can schematically be subdivided into two fields: 'Why do action A?' and 'How-to do A?'. To avoid any confusion with existing terms, our explanation functions are prefixed with E-.

The first subset, 'why do A?' functions, is composed of information providers: **E-structure** and **E-information**. The function 'structure' enhances the structure and the coherence of the text. It is in a large part composed of low-level goal and function expressions (*push to open the box*), indicating motivations and expected results. The aim of 'information', which operates at the ideational level, is to enhance or contradict the beliefs of the reader by providing precise information on some aspects of the action at stake. This very general function may be subdivided into more precise functions such as e.g. E-clarification or E-precision, still to be investigated. The second subset of this group operates at the inter-personal level, and aims at motivating the user to realize the action, via some precautions or recommendations. This subset is formed of the various types of arguments, as usually found in argumentation classifications: **E-warnings** and **E-advice** when there is no implication on the author's part, and **E-threats** and **E-rewards** otherwise. These are designed to justify the importance of an action and the necessity of doing it as required (warnings) or to indicate the optional character of an action (advice) and the benefits of doing it. Besides the recognition of arguments, evaluating their illocutionary or persuasion force is of much interest. This is realized in general via a series of marks, essentially adverbial.

The second subset, 'How-to-do A?' functions, contains several families. The first one deals with functions related to controls on the user related to his actions and interpretations. In the first set, control on user actions, a first group of functions is associated with the notion of **E-guidance**. This function has quite fuzzy boundaries, it can just include temporal marks (Muller et al, 2004), but also possibly manners and instruments which offer a number of details and ways to realize and coordinate actions, up to instructions themselves. It is deeply related to the argument structure of the action verb of the instruction (mainly adjuncts and temporal marks). The next function is **E-framing**, which indicates via a statement the range of application

of an instruction or of a group of instructions (*for X25-01 pumps:....*); it often has the form of a low level title or a condition. **E-Expected result** describes the target result, it is a means for the user to evaluate his performance and to make sure he is on the right way. Finally, **E-elaboration** explains in more depth how to realize an action.

The second family is related to the control of the interpretations made by the user. The goal is to make sure he correctly understands the instructions. In this class fall relatively well know functions directly associated with rhetorical relations: **E-definition**, **E-reformulation**, **E-illustration**, **E-elaboration**. Their goal is to ensure that the user correctly understands the terms of the procedure. The third family is composed of two functions which provide basic help to the user: **E-encouragements**, **E-evaluation**.

4 Explanation schemes

Explanation schemes allow us to introduce a grammar that quite globally describes the semantic and pragmatic structure of explanations. This grammar is based on EES and constrained statements, which form its terminal elements, explanation functions being the non-terminal symbols. We can associate with every explanation function a set of explanation schemes. There is no unique assignment: an explanation scheme may be shared by several explanation functions (Walton et al. 2008).

As an illustration, we give below a few schemes associated with explanation functions (the * indicates multiple occurrences):

E-warning: $[[warning]]$,

$[[warning]] [[illustration]]* [[elaboration]]*$, etc.

E-Definition: $[[definition]] [[illustration]]*$

E-Expected-Result: $[[circumstance]] [[statement]]$
 $expr(+modal, +probability)]]$

this latter example requires a statement with a modal expression such as 'should' introducing a probability.

E-illustration: $[[circumstance]] [[illustration]]*$

Explanation functions can be complex compounds and may include other explanation functions; e.g.:

E-warning:

$[[warning]] E-illustration* E-elaboration*$.

The identification of each EES is based on patterns. These patterns are mainly based on marks and lexical constraints. The marks identified are basically relational: they bind two statements (basically, kernel and satellite). We developed subcategorization frames for these marks [28], [23]. The simplest example is probably *illustration* which has the following patterns in English:

S for example S/NP (NP can be plural) ; (this means any S followed by the string 'for example' followed either by another S or an NP)

S for instance S/NP ;

S such as NP, S e.g. NP ;

S like NP

Elaboration includes marks such as: *actually*, *that is (to say)*, *namely*, *as a matter of fact*, *more precisely*, with similar subcats.

For the tag 'Illustration', tested over a set of 32 procedures has the following recognition rate, we get the following results where precision has been favored over recall:

recall	precision
85%	98%

The recall rate is now so high because in a number of contexts, the notion of illustration (not counting pictures) is not so clear, in particular it may overlap with elaboration, for which we have the

following results:

recall	precision
79%	91%

However, these results are really good considering information retrieval systems for example. Elaboration is a complex relation, where there may be a large gap between the kernel and its satellite, making the recognition by means of patterns quite challenging.

Some EES, such as warnings and advice are composite structures: they include a conclusion and a support for that conclusion. Warnings are basically organized around a unique structure composed of an 'avoid expression' combined with a proposition. The variations in the 'avoid expressions' capture the illocutionary force of the argument via several devices, ordered here by increasing force:

(1) 'prevention verbs like avoid' NP / to VP (*avoid hot water*)

(2) do not / never / ... VP(infinitive) ... (*never put this cloth in the sun*)

(3) it is essential, vital, ... to never VP(infinitive).

Supports are identified from various marks:

(1) via connectors such as: *otherwise*, *under the risk of*, or via verbs expressing consequence,

(2) via negative expressions of the form: *in order not to*, *in order to avoid*, etc.

(3) via specific verbs such as risk verbs introducing an event (*you risk to break*). In general the embedded verb has a negative polarity.

(4) via the presence of very negative terms, such as: nouns: *death*, *disease*, etc., adjectives, and some verbs and adverbs. We defined a lexicon of about 200 negative terms found in our corpora.

Some supports may be empty, because they can easily be inferred by the reader. In that case, the argument is said to be truncated. Warning patterns are implemented on the <TextCoop> platform (see demo paper in this volume). We carried out an indicative evaluation (e.g. to get improvement directions) on a corpus of 36 procedural texts from professional domains, containing 262 warnings. We get the following recognition rates, compared to our manually annotated reference corpus:

conclusion recognition	support recognition	(3)	(4)
88%	91%	95%	95%

(3) conclusions well delimited (4) supports well delimited, w.r.t. warnings correctly identified.

5 Ongoing work

This work remains preliminary and exploratory. So far, we have fully implemented on the <TextCoop> platform (version V2, in Java) the recognition forms within procedures for advice, warnings, purpose, concession, condition, result, elaboration and illustration. Cause is ongoing. EES such as circumstance, analogy and contrast are more difficult to characterize. The result is an enriched semi-structured document with explicit explanation marks, together with procedural marks. This allows us to have a more accurate analysis of explanation, yet to be improved. In particular it is of much interest to consider speech act verbs as a 'support' for explanation, considering e.g. the very detailed descriptions given in [35].

Out <TextCoop> platform is designed for text semantics. It offers a formalism to describe text semantics structures (discourse relations or simply semantic patterns) based on recognition rules.

<TextCoop> V.2 is implemented in Java, and it is in particular based on JFLEX and JCUP technology. Input-output structures tend to be close to UIMA recommendations to allow for a certain interoperability.

Explanation schemes need further work to characterize in more detail their structure, and functions. We will also investigate structural regularities that would reveal forms of know-how in explanation organisation.

From an applied perspective, our aim is to evaluate the impact of explanations on users performances (how they understand them, use them, memorize them). In parallel, we are developing an analysis of risk expressions, since risk prevention is a major challenge in procedures, in particular in developing countries. From a theoretical perspective, we aim at integrating this work into Action theory to be able to formally define the complexity of a procedure.

REFERENCES

- [1] Adam, J.M., *Types de Textes ou genres de Discours? Comment Classer les Textes qui Disent De et Comment Faire*, Langages, 141, pp. 10-27, 2001.
- [2] Adam, J.M., *Types de Sequences Textuelles Elementaires*, Pratiques n56, Metz, 1987.
- [3] Ashley, K.D., Desai, R., Levine, J.M., Teaching case-based argumentation concepts using dialectic arguments vs. didactic explanations, ITS 2002, Lecture notes in computer science, 2002.
- [4] Amgoud, L., Bonnefon, J.F., Prade, H., *An Argumentation-based Approach to Multiple Criteria Decision*, in 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'2005, Barcelona, 2005.
- [5] Amgoud, L., Parsons, S., Maudet, N., *Arguments, Dialogue, and Negotiation*, in: 14th European Conference on Artificial Intelligence, Berlin, 2001.
- [6] Aouladomar, F., Towards Answering Procedural Questions. Dans : KRAQ'05 - IJCAI workshop, Edinburgh, F. Benamara, P. Saint-Dizier (Eds.), p. 24-36, juillet 2005.
- [7] Bieger, G.R., Glock, M.D., *The Information Content of Picture-text Instructions*, Journal of Experimental Education, 53, 68-76, 1984-85.
- [8] Bouffier, A., From texts to structured documents: The case of Health Practice Guidelines, International Semantic Web Conference, Doctoral Consortium, Pusan, Novembre 2007.
- [9] Carberry, S., Plan Recognition in natural language dialogue, Cambridge university Press, MIT Press, 1990.
- [10] Davidson, D., *Essays on Actions and Events*, Oxford: Oxford University Press, 1980.
- [11] Delin, J., Hartley, A., Paris, C., Scott, D., Vander Linden, K., *Expressing Procedural Relationships in Multilingual Instructions*, Proceedings of the Seventh International Workshop on Natural Language Generation, pp. 61-70, Maine, USA, 1994.
- [12] Di Eugenio, B. and Webber, B.L., Pragmatic Overloading in Natural Language Instructions, International Journal of Expert Systems, 1996.
- [13] Gardent, C., Discourse tree adjoining grammars, report nb. 89, Univ. Saarlandes, Saarbrücken, 1997.
- [14] Greimas, A., *La Soupe au Pistou ou la Conservation d'un Objet de Valeur*, in Du sens II, Seuil, Paris, 1983.
- [15] Grosz, B., Sidner, C., Attention, intention and the structure of discourse, Computational Linguistics 12(3), 1986.
- [16] Keil, F.C., Wilson, R.A., *Explanation and Cognition*, Bradford Book, 2000.
- [17] Kosseim, L., Lapalme, G., *Choosing Rhetorical Structures to Plan Instructional Texts*, Computational Intelligence, Blackwell, Boston, 2000.
- [18] Lemarie, J., Lorch, R. F., Eyrolle, H., Virbel, J., SARA: A text-based and reader-based theory of text signaling. Educational Psychologist, 43, 27, 2008.
- [19] Longacre, R., *Discourse Typology in Relation to Language Typology*, Sture Allen éd., Text Processing, Proceeding of Nobel Symposium 51, Stockholm, Almqvist and Wiksell, 457-486, 1982.
- [20] Luc, C., Mojahid, M., Virbel, J., Garcia-Deban, C., Pery-Woodley, M-P., *A Linguistic Approach to Some Parameters of Layout: A study of enumerations*, In R. Power and D. Scott (Eds.), Using Layout for the Generation, Understanding or Retrieval of Documents, AAAI 1999 Fall Symposium, pp. 20-29, 1999.
- [21] Mann, W., Thompson, S., *Rhetorical Structure Theory: Towards a Functional Theory of Text Organisation*, TEXT 8 (3) pp. 243-281, 1988.
- [22] Mann, W., Thompson, S.A. (eds), *Discourse Description: diverse linguistic analyses of a fund raising text*, John Benjamins, 1992.
- [23] Miltasaki, E., Prasad, R., Joshi, A., Webber, B., *Annotating Discourse Connectives and Their Arguments*, new frontiers in NLP, 2004.
- [24] Muller, P., Tannier, X., *Annotating and Measuring Temporal Relations in Texts*, In Proceedings of Coling 2004, volume I, pages 50-56, Geneva, 2004.
- [25] Pollock, J.L., *Knowledge and Justification*, Princeton university Press, 1974.
- [26] Reed, C., *Generating Arguments in Natural Language*, PhD dissertation, University College, London, 1998.
- [27] Rosner, D., Stede, M., *Customizing RST for the Automatic Production of Technical Manuals*, in R. Dale, E. Hovy, D. Rosner and O. Stock eds., Aspects of Automated Natural Language Generation, Lecture Notes in Artificial Intelligence, pp. 199-214, Springer-Verlag, 1992.
- [28] Saito, M., Yamamoto, K., Sekine, S., *Using Phrasal Patterns to Identify Discourse Relations*, ACL, 2006.
- [29] Schank, R., *Explanation Patterns: Understanding Mechanically and Creatively*, Laurence Erlbaum, 1986.
- [30] Takechi, M., Tokunaga, T., Matsumoto, Y., Tanaka, H., *Feature Selection in Categorizing Procedural Expressions*, The Sixth International Workshop on Information Retrieval with Asian Languages (IRAL2003), pp.49-56, 2003.
- [31] Talmy, L., *Towards a Cognitive Semantics*, vol. 1 and 2, MIT Press, 2001. Vander Linden, K., *Speaking of Actions Choosing Rhetorical Status and Grammatical Form in Instructional Text Generation* Thesis, University of Colorado, 1993.
- [32] Walton, D., Reed, C., Macagno, F., *Argumentation Schemes*, Cambridge University Press, 2008.
- [33] Webber, B.L. and Di Eugenio, B. *Free Adjuncts in Natural Language Instructions*, Proceedings COLING-90, Helsinki, Finland, August 1990
- [34] Webber, B., D-LTAG: extending lexicalized TAGs to Discourse, Cognitive Science 28, pp. 751-779, Elsevier, 2004.
- [35] Wierzbicka, A., *English Speech Act Verbs*, Academic Press, 1987.
- [36] Wright, von G.H., *Explanation and understanding*, Cornell university Press, 2004.

Obligationes as Formal Dialogue Systems¹

Sara L. Uckelman²

Abstract. Formal Dialogue Systems (FDSs) model rule-based interaction between agents and as such have multiple applications in multi-agent systems and AI more generally. Their conceptual roots are in formal theories of natural argumentation, of which Hamblin’s formal systems of argumentation in [15, 16] are some of the earliest examples. Hamblin cites the medieval theory of *obligationes* as inspiration for his development of formal argumentation. In an *obligatio*, two agents, the Opponent and the Respondent, engage in an alternating-move dialogue, where the Respondent’s actions are governed by certain rules, and the goal of the dialogue is establishing the consistency of a proposition. We implement *obligationes* in the formal dialogue system framework of [31] using Dynamic Epistemic Logic [38]. The result is a new type of inter-agent dialogue, for *consistency-checking*, and analyzing *obligationes* in this way also sheds light on interpretational and historical questions concerning their use and purpose in medieval academia.

1 Introduction

Rule-based interactions such as dialogues or arguments are ubiquitous and diverse; they are the basic method of communication between agents. Such interactions are modeled in artificial intelligence and computer science by formal dialogue systems (FDSs) and dialogue games [18, 22, 23, 24, 31]. These systems give formal, and hence potentially implementable, methods for modeling real-life dialogue situations, such as complex reasoning in legal domains. More generally, dialogue systems and games are used in multi-agent systems to model distributed cognition and interaction between intelligent agents, and they can also be used in the specification of complex software systems and programs [24].

The conceptual roots of formal dialogues, however, come not from within AI but from without. They are located in the sphere of natural argumentation, that is, philosophical logic, argumentation theory, and rhetoric. One of the earliest attempts to provide a theory of formal dialogues is Hamblin’s [16, 15]. In [16], Hamblin locates part of the motivation for his development of formal argumentation in historical formal dialogue systems, that is, dialogical or disputational settings where explicit rules are given governing the actions of the participants. One such system of natural argumentation that he considers in particular is the medieval theory of *obligationes*, developed in the 13th and 14th centuries. In an *obligatio*, two agents, the Opponent and the Respondent, engage in an alternate-move dialogue, where the Respondent’s actions are governed by certain rules, and the goal of the dialogue is, in the most basic case, to establish the consistency of a proposition. We argue that *obligationes* are best modeled by FDSs because of their intrinsic dialogical nature. Furthermore,

taking *obligationes* from the realm of natural argumentation to the realm of formal dialogue systems results in the determination of a new type of dialogue system different from those generally discussed in the AI literature, and thus they make a novel contribution to formal modeling of dialogue and interaction in multi-agent systems.

The plan of the paper is as follows. In §2 we present the medieval theory of *obligationes*, focusing specifically on the works of one author, Walter Burley, and give examples. In §3 we briefly survey previous work on *obligationes*, both formal and philosophical, and motivate modeling *obligationes* as FDSs by showing how they can make sense of King’s interpretation of *obligationes* as a meta-disputational framework. In §4 we introduce formal dialogue systems, following the presentation in [31], and show generically how *obligationes* can be viewed as a formal dialogue system. Before we give a precise specification of *obligationes* as FDSs in §6, we first outline the logic used in the argumentation, a type of Dynamic Epistemic Logic, in §5. We discuss the properties of the protocol we introduce, compare our results with standard types of FDSs, and define a new type of inter-agent dialogue, for *consistency-* or *feasibility-checking*, in §7. In §8 we conclude and point towards future work.

2 Medieval theories of *obligationes*

An *obligatio* is a dialogue between two agents, the Opponent and the Respondent, where the Opponent puts forward a sequence of propositions, and the Respondent is obligated (hence the name) to follow certain rules in his responses to the Opponent’s propositions. More precisely, the Opponent puts forward an initial statement, called the *positum*, which the Respondent can either accept or refuse to accept. If he accepts, the *obligatio* begins. If he does not, no *obligatio* begins. If the *obligatio* begins, the Opponent puts forward propositions and the Respondent has three ways that he can respond: He can grant or concede the proposition, he can deny the proposition, or he can doubt it, where ‘doubt’ should be understood as ‘remain agnostic about’; doubting φ does not entail any commitment to $\neg\varphi$. (Some authors, such as William of Ockham [26] and the anonymous author of the *Obligationes Parisienses* [11], mention a fourth option, which is to ‘draw distinctions’, that is, to clarify an ambiguity on the part of the Opponent.) The *obligatio* continues until the Opponent calls “*Cedat tempus*” (“Time’s up”), whereupon the responses of the Respondent are analysed with respect to the Respondent’s obligations, to determine whether he has responded well or badly.

The earliest texts on *obligationes* date from the beginning of the 13th century [10, 11, 12], and many of the leading logicians from that century and the next wrote treatises on the subject. While the roots of obligational disputations are clearly grounded in Aristotle’s discussion of dialectical exchanges in the *Topics* VIII, 4 (159a15–24) and in the *Prior Analytics* I, 13 (32a18–20) (cf. [41, §II.A]), the systematic development of the theory of *obligationes* over the course of the 13th

¹ A short version of this paper will appear in the proceedings of STAIRS’10.

² Institute for Logic, Language, and Computation, University of Amsterdam, The Netherlands, email: S.L.Uckelman@uva.nl

and 14th centuries tends to show little adherence to the Aristotelian definitions. While the specific details vary from author to author, a number of distinct types of *obligationes* discussed by multiple authors can be identified. The six most common are *positio*, *depositio*, *dubitatio*, *sit verum* or *rei veritatis*, *institutio*, and *petitio*. Of these six, *positio* is universally the most widely studied, both by medieval and modern authors; as a result, it is the focus of the current paper. For further information on *obligationes*, including a discussion of their purpose and their role in medieval philosophy, see [41].

To make the above more precise, we look at the theory of *obligationes* of a specific writer, Walter Burley. Burley's treatise *De obligationibus*, written around 1302, gives a standard treatment of *positio*. The text of this treatise is edited in [7] and a partial translation of the text, including the section on *positio* in its entirety, is found in [6]. (In the interests of conserving space, we quote all medieval texts in translation only, rather than in the original Latin.) Burley defines the general goal of an *obligatio* as follows:

The opponent's job is to use language in a way that makes the respondent grant impossible things that he need not grant because of the *positum*. The respondent's job, on the other hand, is to maintain the *positum* in such a way that any impossibility seems to follow not because of him but rather because of the *positum* [6, p. 370].

Thus, it is clear that in an *obligatio*, the goal is consistency, not logical truth or validity. In *positio*, the primary obligation of the Respondent is to grant, that is, to hold as true, the *positum*. If the Respondent accepts the *positum* and the *obligatio* begins, he is obliged to follow the following rules:

Rule 1 Everything that is posited and put forward in the form of the *positum* during the time of the *positio* must be granted [6, p. 379].

Rule 2 Everything that follows from the *positum* must be granted. Everything that follows from the *positum* either together with an already granted proposition (or propositions), or together with the opposite of a proposition (or the opposites of propositions) already correctly denied and known to be such, must be granted [6, p. 381].

Rule 3 Everything incompatible with the *positum* must be denied. Likewise, everything incompatible with the *positum* together with an already granted proposition (or propositions), or together with the opposite of a proposition (or the opposites of propositions) already correctly denied and known to be such, must be denied [6, p. 381].

In Rule 1, 'in the same form as' should be understood syntactically; if the *positum* is 'Marcus is Roman', then the Respondent doesn't have an obligation to accept 'Tullius is Roman' unless it is explicit (either through common knowledge or through previous concessions) that Marcus is Tullius.

Burley also defines a notion of relevance of propositions which applies to all types of *obligatio*. A proposition is *irrelevant* or *impertinent* if neither it nor its negation follows from the set of propositions which have already been conceded (which includes the negations of propositions which have been denied).

Rule for Irrelevant Propositions One must reply to what is irrelevant in accordance with its own quality [6, p. 375].

I.e., the Respondent should reply by conceding the proposition if it is true, by denying if it is false, and by doubting if he does not know which is the case. The notion of 'relevance' in *obligationes* parallels

	Opponent	Respondent	
1	φ .	I admit it.	$\Phi_0 := \{\varphi\}$.
2	$\neg\varphi \vee \psi$.	I grant it.	$\Phi_1 := \{\varphi, \neg\varphi \vee \psi\}$.
3	ψ	I grant it.	

Figure 1. An example *obligatio*.

	Opponent	Respondent
1	φ or φ must be granted.	I admit it.
2	φ must be granted.	I deny it.
3	φ follows from the <i>positum</i> and the opposite of something correctly denied	I grant it.
4	φ must be granted.	??

Figure 2. A more interesting example.

the concept of 'support' in formal argumentation theory where an argument is defined as a pair $A = (S, p)$ where S is a consistent subset of the knowledge base such that $S \models p$ (cf., e.g., [9, p. 155]).

A simple example illustrating Burley's rules for *positio* is given in Figure 1. Suppose φ does not imply $\neg\psi$ and φ is known to be contingently false. In the first round, the Opponent puts forward a contingent (but false) proposition; the Respondent grants it in accord with Rule 1. In the second round, either φ implies ψ , then the sentence is relevant and follows from Φ_0 (the set of propositions conceded so far along with the negations of propositions denied to this point); or it doesn't, in which case it is irrelevant and true (since φ is false). In both cases, the Respondent is required to concede; the first case falls under Rule 2, and the second under the Rule for Irrelevant Propositions. In the third round, the Respondent likewise must concede because ψ follows from Φ_1 . This example *obligatio* shows how, given a *positum* which is false but not contradictory, the Opponent can force the Respondent to concede any other consistent proposition.

More interesting examples, such as the example in Figure 2, involve statements about the obligational rules themselves. Let φ be the proposition 'you are in Rome' (spoken by the Opponent to the Respondent). The *positum* is a disjunction between a simple proposition and the assertion that that proposition must be granted. Because the disjunction is not a logical contradiction (in particular the first disjunct is possible, though it is in fact false), the Respondent is correct in accepting the *positum*. The second disjunct is irrelevant, as it is not a logical consequence of the *positum*, and furthermore it is false: Since φ is false, and φ is also irrelevant, the Respondent is not under any obligation to accept φ . Thus it is false that φ must be granted, so he correctly denied the second proposition. The third proposition expresses a logical necessity, about the validity of disjunctive syllogism, and so is accepted. But now it is unclear how the Respondent should respond to the re-assertion that φ must be granted. On the one hand, this proposition has been put forward before, and was denied, and so it should continue to be denied. On the other hand, once the third proposition has been granted, by Rule 2, φ must be granted. So superficially it appears that the Respondent is obliged to both accept and deny this final statement. Burley's resolution to the problem is to argue that (3) is not only not necessary, but it is repugnant, since it is inconsistent with the opposite of (2). Since it is repugnant, the Respondent should have in fact denied, and thus (4) can also be denied without contradiction [41, pp. 152–155].

3 Previous work on *obligationes*

Green’s Ph.D. dissertation [7], containing an edition of and commentary on two treatises on *obligationes*, now generally ascribed to William of Sherwood and Walter Burley, marks the beginning of modern research on *obligationes*. Since then, many philosophers and historians have devoted themselves to the question of the goal or purpose of obligational disputations and the role they played in medieval academic life, while somewhat fewer have focused on the logical properties of *obligationes*. Despite this, the purpose of *obligationes* and their role in medieval academic life remains stubbornly unclear [34]. Two recent sources which discuss the various views, which range from the view that *obligationes* were mere academic exercises, that they were used for counter-factual reasoning, that they were a tool for evaluating *sophismata* and *insolubilia* (paradoxical and problematic sentences), that they are thought experiments, and that they provide a theory of belief revision, are [41] and [42].

Hamblin is the first modern author to attempt to formalize *obligationes* [16, pp. 260–263]. Given his interest in formalizing argumentation generally, he focuses on the dialogical aspects of *obligationes*. His formalization is rudimentary and models only one variant, that given by William of Sherwood³, but it marks the beginning of modern scholarship on the formal properties of *obligationes*. Recent scholarship has focused on the game-like nature of *obligationes*, e.g., [11, 13, 41]. In particular, there is an immediate apparent similarity between *obligationes* and Lorenzen’s dialogical logic [20]. It may therefore seem natural to look to game-based structures in logic to provide a general framework for modeling different types of *obligationes*. However, there are a number of aspects which do not immediately lend themselves nicely to a game-like interpretation (e.g., the notion of a winning strategy for an *obligatio*, for example, is difficult to define⁴), and despite the strongly logical component of *obligationes*, to date relatively little work has been done on the formal properties of the logic involved and few attempts have been made to provide an explicit specification of the game(s) involved.

The most extensive attempt is in [13]. In this book, Dutilh Novaes analyses the obligational theories of three authors, Walter Burley, Richard Swyneshed (c.1330), and Ralph Strode (second half of the 14th C), giving separate formalizations for each one. Her formalizations are based on models $\mathfrak{M} = \langle K_c, \Phi, \Gamma, R(\varphi) \rangle$ where K_c is the set of common knowledge among the participants of the disputation (expressed as a set of propositions); Φ is a sequence of propositions, which keeps track of the assertions of the Opponent; Γ is a sequence of propositions, which keeps track of the responses of the Respondent; and $R(\varphi)$ is a function from φ to 1 (standing for ‘concede’), 0 (‘deny’), and ? (‘doubt’). These formalizations are not very satisfying for a number of reasons. Each obligational theory studied is provided with a different framework, which means that it is difficult to make cross-theory comparisons. Further, only *positio* is studied; the other types of *obligationes* are not discussed. Additionally, the frameworks all presuppose a significant amount of background information which is taken for granted and never specified: the semantic model(s) in which truth of propositions (particularly the *positum*

and irrelevant propositions) and the Respondent’s knowledge of both individual propositions as well as how the consequence relations are to be evaluated, and the syntactic rules governing $\vdash$ (which is used to generate Γ). For example, the set of common knowledge K_C is not defined in any explicit fashion, and there is nothing which grounds the knowledge of the participants. Finally, since the nature of the proof system being used in the definition of $R(\varphi)$ is never specified, her frameworks are essentially incomplete; it is impossible to implement the logical model without making the proof-system explicit [13, p. 169].

When looking for an alternative to a game-theoretic approach to modeling *obligationes*, one fruitful suggestion that presents itself can be rooted in a particularly interesting interpretation of *obligationes* given by King [19]. King takes his starting point from Spade, who, in [35], looked to the textual evidence for actual uses of *obligationes* to understand how they were used by the medievals. While to date there is no historical record for actual obligational disputations, we have many examples of philosophers using obligational techniques as part of their argumentation [19, p. 1]. King explains the apparent “content-freeness” of obligational disputations by pointing out that “they operate at a higher level of logical generality than that at which substantive debate occurs. If this is correct, then actual obligational moves—perhaps even recognized as such—are the vehicle whereby real argument takes place” [19, p. 6], and thus *obligationes* provide a “meta-methodology” for reasoning [19, p. 7].

We use this suggestion as the motivation for our approach to modeling *obligationes*. An *obligatio* is essentially a dialogue; and any dialogue can be seen as a game played according to the rules specified by an FDS [22]. We believe that viewing *obligationes* as FDSs, which require that we explicitly specify the logic of argumentation/inference and the models against which the dialogue is to be evaluated, provide a more fruitful approach to modeling *obligationes*. On this view, Hamblin’s modeling approach has the advantage over others proposed in recent literature because it takes the dialogical nature of the disputation seriously. By varying the rules governing the disputation, radically different types of *obligationes* arise, which result in radically different types of dialogues/disputations. Despite the wide range of difference that can be found, the basic structure of an *obligatio* remains the same, making the general framework of FDSs an appropriate modeling choice. Specifying *obligationes* from within the context of FDSs allows us to situate them formally in current research on formal dialogues, which in turn can help to clarify the interpretational question, by helping us understand the possible purposes to which *obligationes* could be disposed. In particular, we argue that the naturalness of modeling *obligationes* as dialogue systems supports King’s suggestion that *obligationes* provide agents with a meta-methodology for argumentation. That is, *obligationes* give frameworks within which dialectical argumentation—dialogue—can take place.

4 Formal dialogue systems

The standard taxonomy of formal dialogue systems is based on the argumentation-based typology given by Walton and Krabbe [39], who identify six different basic types of dialogues: information seeking, inquiry, persuasion, negotiation, deliberation, and eristic. The division is made on the basis of the preconditions and postconditions that must hold for successful dialogue to take place [9, 24]. Three of these classes, information seeking, inquiry, and persuasion, deal with beliefs and knowledge, and thus are of particular interest. Walton and Krabbe make no claim as to the comprehensiveness of their classi-

³ Hamblin routinely questions the attribution to Sherwood of the text he is considering; however, more recent scholarship is agreed that the text was almost certainly written by Sherwood, sometime in the middle of the 13th century [4].

⁴ Yrjönsuuri mentions the possibility of modeling *obligationes* as games, but he says that “defining the results of the game in any manner appropriate to modern game-theory seem utterly problematic” though despite this “[i]n the following I will keep to the English word *game*, assuming that the problems pointed out above can just be left unsolved” [41, pp. 9–10].

fication, and others [2, 8, 9, 14, 32, 33] have researched types not covered in the Walton and Krabbe typology. In particular, Cogan et al. argue that “there remain several situations in which it seems natural to engage in dialogues, but to which the basic Walton and Krabbe dialogue types do not apply” [9, p. 161]. In [8, 9], they take a systematic approach to defining dialogue types on the basis of pre- and post-conditions, and extend Walton and Krabbe’s list of belief-based dialogues with four new types, verification and three types of queries. In §7, we discuss how *obligationes* fit into these different types. First, we make precise what we mean by a formal dialogue system.

In this section, we follow the presentation of formal dialogue systems given in [31], an overview paper which discusses different formal argumentation systems that have been proposed for the analysis of persuasion dialogues and provides a unified approach within which each of these different systems can be modeled. While Prakken focuses on persuasion dialogues, his framework is in fact general enough to handle other types as well [31, pp. 170, 173]. Thus, it is appropriate to use it to consider *obligationes*.

The specification of a *formal dialogue system* contains the following elements [31, p. 166]:

- A *topic language* $\mathcal{L}_t$, closed under classical negation.
- A *communication language* $\mathcal{L}_c$. We denote the set of *dialogues*, that is, the set sequences of $\mathcal{L}_c$, by $M^{\leq \infty}$, and the set of finite sequences of $\mathcal{L}_c$ by $M^{< \infty}$. For a dialogue $d = m_0, \dots, m_n, \dots$, the subsequence $m_0, \dots, m_i$ is denoted d_i .
- A *dialogue purpose* or *goal*.
- A set $\mathcal{A}$ of *agents* (participants) and a set $\mathcal{R}$ of *roles* that the participants can occupy. Each participant a has a (possibly empty) *belief base* $\Sigma_a \subseteq \mathcal{L}_t$ and a (possibly empty) *commitment set* $C_a(d_n) \subseteq \mathcal{L}_t$. The belief base may or may not change during the dialogue; the commitment set usually does.
- A *context* $K \subseteq \mathcal{L}_t$, representing the (shared, consistent, and unchanging) knowledge of the agents specified at the outset.
- A *logic* L for $\mathcal{L}_t$.
- A set E of *effect rules* $C_a(d_n) : M^{< \infty} \rightarrow \mathcal{P}(\mathcal{L}_t)$ for $\mathcal{L}_c$, specifying how utterances $\varphi \in \mathcal{L}_c$ in the dialogue affect the commitment stores of the agents. The effect rules are such that if $d = d'$ then $C_a(d, m) = C_a(d', m)$, that is, the changes in commitments are determined solely by the most recent move in the dialogue along with the commitments at that step.
- A *protocol* P for $\mathcal{L}_c$, specifying the legal moves of the dialogue, which is a function from the context and a non-empty $D \subseteq M^{< \infty}$ to $\mathcal{P}(\mathcal{L}_c)$, satisfying the requirement that if $d \in D$ and $m \in P(d)$, then $d, m \in D$. The elements of D are called *legal finite dialogues*, and $P(d)$ is the set of moves allowed after move d . At any stage, if $P(d) = \emptyset$, then the dialogue has *terminated*. A protocol will often be accompanied by a *turn-taking* function $T : D \rightarrow \mathcal{P}(\mathcal{A})$, which takes a finite dialogue d_n and specifies who governs move m_{n+1} , and *termination* conditions, which specify when $P(d) = \emptyset$.
- A set of *outcome rules* O .

Dialogue systems can be explicitly connected with the games they specify by defining the agents’ strategies in the standard game-theoretical way. Formally, a *strategy* s_a for agent a is a function $D_a \rightarrow \mathcal{L}_c$, where $D_a \subseteq D$ is the set of all finite legal dialogues d_n in which $T(d_n) = a$. A strategy s_a is called *winning* if in every dialogue where a follows this strategy, he realizes his dialogue goal.

We can identify a number of properties of protocols [31, p. 170]:

- A protocol has *public semantics* iff the set of legal moves is always independent from the agents’ belief bases.

- A protocol is *context-independent* iff the set of legal moves and the outcome is always independent of the context, that is, $P(K, d) = P(\emptyset, d)$.
- A protocol is *fully deterministic* iff P always returns a singleton or the empty set.
- A protocol is *unique-move* iff the turn shifts after each move; it is *multiple-move* otherwise.

Protocols which are not fully deterministic are *permissive*, that is, they specify what moves are *legal* or *allowed* for the agent, rather than specify what moves are *required*. Thus, *obligationes* are a type of DS where the protocol for the Respondent is fully deterministic; for each proposition the Opponent puts forward, there will be exactly one correct move that the Respondent can make.

We now show how generically *obligationes* can be viewed as DSs; we give precise examples in §6. In *obligationes*, there are two designated roles Opp (Opponent) and Res (Respondent) that members of $\mathcal{A}$ can have; those members of $\mathcal{A}$ which do not fill either role are irrelevant for modeling the disputation. We explain below how Σ_{Opp} , Σ_{Res} , C_{Opp} , C_{Res} , and the context K are generated. In Burley-style *positio*, the dialogue purpose is consistency: If we take Res’s commitment set to be the set of formulas he has conceded along with the negation of those that he’s denied over the course of a *positio*, then the goal for Res is to maintain the consistency of his commitment set, and the goal for Opp is to force Res into contradiction.

In general, the topic language $\mathcal{L}_t$ and the communication language $\mathcal{L}_c$ are the same. This allows, among other things, the participants in an *obligatio* to dispute about the allowed moves of the other players. (For example, Opp may ask Res to respond to the claim “You deny φ ”.) The turn-taking protocol in an *obligatio* is unique-move: $T(\emptyset) = \text{Opp}$, $T(d_n) = \text{Opp}$ if n is odd, and $T(d_n) = \text{Res}$ if n is even. (Throughout we assume that we label the steps in the sequence from 0, so in an *obligatio* it is always Opp that goes first.) The protocol P will be such that the moves of Opp are not constrained in any way, but Res’s moves must be made in reaction to the move of Opp at the previous stage. The same will be true for the effect rules E ; in a disputation, Opp makes a series of claims or assertions, but these actions have no effect on his commitment store. On the other hand, Res is constrained to be reactive only: He can only concede statements claimed by Opp, concede their negations, or remain ambivalent. Res never asserts any statement of his own devising, he only ever responds to propositions put forward by Opp. Thus, *obligationes* are essentially asymmetric, in that the rules governing the behavior of the Opp and Res are disjoint⁵, and so are their actions.

The outcome rules for *obligationes* are simple: If Res realizes the goal, then he wins. If Opp realizes the goal, then he wins. There is nothing further that hinges upon winning or losing an obligational disputation (except, of course, the individual prestige or embarrassment of the participants!).

Above we noted that in an arbitrary DS, the commitment set of an agent will generally change during the course of the dialogue. It can either strictly grow, so that the agents are only adding new propositions to their commitment-base at each turn, or they can also revise their commitments by rejecting previous commitments in favor of new ones. This latter case arises in ordinary circumstances when agents utilize a form of default reasoning, which is defeasible and non-monotonic, in that an agent can be forced to accept information which contradicts his previous commitments, requiring that

⁵ In fact, in most texts, no rules for Opp are given. One exception is the early text *Tractatus Emmeranus* [10], which gives some rules (better thought of as guidelines, or strategic advice) to the Opponent.

his commitments be revised in order to maintain consistency (cf. [3, 5]). In AI contexts, the ability to simulate non-monotonic reasoning is of great importance; monotonic dialogues and discussions are more commonly found in philosophical contexts. One of the benefits of Prakken's approach to DS is that it can handle both approaches, merely by the specification of the underlying logic [31, p. 173].

5 The underlying logic

By specifying the logic L and its underlying models, we are able to explicitly generate Σ_{Opp} , Σ_{Res} , C_{Opp} , C_{Res} , and K satisfying desired properties. In our approach to modeling *obligationes* as DSs, the underlying logic is multi-agent Dynamic Epistemic Logic (DEL, [38]). This logic is monotonic and not argument based. An epistemic logic is an extension of propositional logic with a family of modal operators K_a for $a \in \mathcal{A}$. We are interested in a particular extension of standard epistemic logic, namely, *epistemic logic with common knowledge*, which has a further family of operators C_G , for $G \subseteq \mathcal{A}$. For a set Φ_0 of propositional letters and set $\mathcal{A}$ of agents, the set Φ_{EL}^A of wffs of EL is defined by:

$$\varphi := p \in \Phi_0 \mid \neg\varphi \mid \varphi \vee \varphi \mid K_a\varphi : a \in \mathcal{A} \mid C_G\varphi : G \subseteq \mathcal{A}$$

$K_a\varphi$ is read ‘agent a knows that φ ’. $C_G\varphi$ is read ‘it is common knowledge amongst the group of agents G that φ ’. We will use C_G to represent explicitly the knowledge of the two agents at the beginning of the disputation.

Epistemic logic is interpreted on Kripke frames. A structure $\mathfrak{M} = \langle W, w^*, \{\sim_a : a \in \mathcal{A}\}, V \rangle$ is an *epistemic model* if

- W is a set (of possible worlds), with $w^* \in W$ a designated point (representing the actual world).
- $\{\sim_a : a \in \mathcal{A}\}$ is a family of equivalence relations on W , one for each member of $\mathcal{A}$. The relation $w \sim_a w'$ is interpreted as ‘ w and w' are epistemically equivalent for agent a ’. $\sim_G : G \subseteq \mathcal{A}$ is defined as the reflexive and transitive closure of $\bigcup_{a \in G} \sim_a$.
- $V : \Phi_0 \rightarrow 2^W$ is a valuation function associating atomic propositions with subsets of W . For $p \in \Phi_0$, if $w \in V(p)$, we say that ‘ p is true at w ’.

The semantics for the propositional connectives are as expected. We give just the semantics for the epistemic operators.

$$\begin{aligned} \mathfrak{M}, w \models K_a\varphi & \quad \text{iff} \quad \forall w' (\text{if } \langle w, w' \rangle \in \sim_a \text{ then } \mathfrak{M}^E, w' \models \varphi) \\ \mathfrak{M}, w \models C_G\varphi & \quad \text{iff} \quad \forall w' (\text{if } \langle w, w' \rangle \in \sim_G \text{ then } \mathfrak{M}^E, w' \models \varphi) \end{aligned}$$

EL models cover the knowledge of the agents; to model their actions, we add dynamics, via Propositional Dynamic Logic (PDL, [17]). PDL is an extension of propositional logic by a family of modal operators $[\alpha]$ for $\alpha \in \Pi$, a set of programmes (or more generally, a set of actions or events). The language of PDL is two-sorted, with a set Φ_0 of atoms and a set Π_0 of atomic actions. We do not need the full expressivity of PDL to model *obligationes*, so we introduce only the fragment we require. We let $\Pi_0 = \emptyset$, and the sets Φ_{Ob} and Π_{Ob} of complex well-formed formulas and programmes are defined by mutual induction:

$$\begin{aligned} \varphi &:= \varphi \in \Phi_{\text{EL}}^A \mid [\alpha]\varphi : \alpha \in \Pi_{\text{Ob}} \\ \alpha &:= \varphi? : \varphi \in \Phi_{\text{EL}}^A \end{aligned}$$

The programme $\varphi?$ is to be interpreted as a test operator, which tests for the truth of φ . Note that the only programmes that we allow are testing of formulas which do not themselves contain any

programmes. The semantics for the new $[\varphi?]$ operator are given in terms of model reduction. Let $\mathfrak{M} \upharpoonright \varphi := \langle W^{\mathfrak{M}, \varphi}, \{\sim_a^{\mathfrak{M}, \varphi} : a \in \mathcal{A}\}, V^{\mathfrak{M}, \varphi} \rangle$, where $W^{\mathfrak{M}, \varphi} := \{w \in W : \mathfrak{M}, w \models \varphi\}$, and the relations and valuation functions are just restrictions of the originals. For a set of ordered propositions Γ_n , let $\mathfrak{M} \upharpoonright \Gamma_n = \mathfrak{M} \upharpoonright \gamma_0 \upharpoonright \dots \upharpoonright \gamma_n$, that is, $\mathfrak{M} \upharpoonright \Gamma_n$ is the result of the sequential restriction of $\mathfrak{M}$ by the elements of Γ_n . Then:

$$\mathfrak{M}, w \models [\varphi?]\psi \quad \text{iff} \quad \forall v \in \mathfrak{M} \upharpoonright \varphi, v \models \psi$$

One advantage of using an epistemic logic for our disputation logic is that it allows us to model the epistemic bases of the agents, and the context of the disputation, explicitly (for a fuller discussion of the advantages, which includes the ability to use this framework to model different types of *obligationes* beyond just the one considered here, see [36]). While above we defined the concept of a ‘belief base’ in a dialogue, in the context of *obligationes* it is the agent’s knowledge, not their beliefs that is important. Given an epistemic model $\mathfrak{M}$, the knowledge bases of Opp and Res are defined as follows:

$$\begin{aligned} \Sigma_{\text{Opp}}^{\mathfrak{M}} &:= \{\varphi : \mathfrak{M}, w^* \models K_{\text{Opp}}\varphi\} \\ \Sigma_{\text{Res}}^{\mathfrak{M}} &:= \{\varphi : \mathfrak{M}, w^* \models K_{\text{Res}}\varphi\} \end{aligned}$$

In an arbitrary model $\mathfrak{M}$, the set of propositions which are common knowledge amongst a group of agents is not explicitly specified. In an *obligatio*, the set of common knowledge, against which the truth of irrelevant propositions is evaluated, is likewise often left implicit. In some cases, before the *obligatio* begins, a *casus* is introduced. A *casus* is a hypothesis about how the world is, or extra information about how the *positum* should be analyzed [40]. A common example of a *positum* introduced with a *casus* (understood in the first sense) is ‘In truth Socrates is black. It is posited that Socrates is white.’ The first sentence is the *casus*; it tells the participants not only that Socrates is black, but also that he exists, and is colored, all of which facts the Respondent must take into account when responding to the Opponent’s *posita*. Thus, in the first sense, the *casus* can be understood as a set of literals expressing the *explicit common knowledge* at the start of the dialogue, so the *casus* can be implemented by a restriction on V .

Definition 5.1. (*Casus*). Let Lit_{Φ_0} be the set of literals formed from Φ_0 , and $K \subseteq \text{Lit}_{\Phi_0}$ be the *casus*. Then $\mathfrak{M}$ models the *casus* if there is a $P_c \subseteq P$ of W with $w^* \in P_c$, such that if $w \sim_{\text{Res}} w^*$, then $w \in P_c$, if $v \sim_{\text{Opp}} w^*$, then $v \in P_c$, and for all $w, v \in P_c$, $w \sim_{\text{Res}} v$ and $w \sim_{\text{Opp}} v$; and for every positive literal $p \in K$ and every $w \in P_c$, $w \in V(p)$, and for every negative literal $\neg q \in K$ and every $w \in P_c$, $w \notin V(q)$.

Unlike contexts in dialogue systems, it is not assumed that the *casus* of an *obligatio* is consistent, but if it is not, then Res should not accept the *positum*, since Opp could easily force him into conceding a contradiction. However, if the *casus* is consistent, we can easily show that if $\mathfrak{M}$ models a *casus* K , then for every $\varphi \in K$, $\mathfrak{M} \models C_{\{\text{Opp}, \text{Res}\}}\varphi$, and so $K \subseteq \Sigma_{\text{Opp}}^{\mathfrak{M}}$ and $K \subseteq \Sigma_{\text{Res}}^{\mathfrak{M}}$.

6 Protocols, effect rules, and outcomes

Different types of *obligationes* can be modeled by changing the protocols, effect rules, and outcome conditions. First, we specify the general properties shared by all *obligationes*. We identify our set of agents with their roles, i.e., our set of agents is $\mathcal{A} = \{\text{Opp}, \text{Res}\}$, and our topic language and commitment language is the language of dynamic epistemic logic $\mathcal{L}_{\text{DEL}}$ introduced in the previous section. Let

α be a designated formula representing “*cedat tempus*”. We can identify two types of protocols used in *obligationes*. The first type of protocol is uniform throughout all different systems; the second varies from author to author and type to type. The uniform protocol P_u is invariant over all contexts and is defined for a finite dialogue d_n :

$$\begin{aligned} & \text{if } m_n = \alpha & P_u(\emptyset) &= \mathcal{L}_c \\ & \text{otherwise, if } n \text{ is odd,} & P_u(d_n) &= \emptyset \\ & \text{and if } n \text{ is even,} & P_u(d_n) &= \mathcal{L}_c \\ & & P_u(d_n) &= \{[m_n?] \top, [\neg m_n?] \top, [\top?] \top\} \end{aligned}$$

That is, if it is Opp’s turn, he is allowed to assert any statement in the communication language (we allow repetitions). If it is Res’s turn, he must either concede, deny, or doubt Opp’s statement from the previous round. And if “*cedat tempus*” has been called, the dialogue ends and there are no more legal moves available. Since m_n , the move of Opp, will always be a statement in the communication language $\mathcal{L}_c$, and the communication language allows for the embeddings of the test programme, this protocol is well-defined. This protocol has public semantics and is context-independent, but it is not fully deterministic, since whenever it is Res’s turn, he has a choice of actions. For ease of future reference, we introduce meta-names for the actions of Res: **concede**: $\varphi := [\varphi_n?] \top$, **deny**: $\varphi := [\neg \varphi_n?] \top$, and **doubt**: $\varphi := [\top?] \top$. The last clause is equivalent to saying “I don’t know”; $[\top?] \top$ will always be valid, in any model.

The rules governing the commitment sets C_{Opp} and C_{Res} are defined as follows:

$$\begin{aligned} & \text{for all } n & C_{\text{Opp}}(d_n) &= \emptyset \\ & \text{if } n \text{ is even} & C_{\text{Res}}(d_n) &= C_{\text{Res}}(d_{n-1}) \\ & \text{if } n \text{ is odd} & C_{\text{Res}}(d_n) &= C_{\text{Res}}(d_{n-1}) \cup \{m_n\} \end{aligned}$$

That is, Opp has no commitments, Opp’s moves do not change Res’s commitments, and Res’s commitment store strictly grows on the basis of his actions, and thus obligational dialogues are monotonic. As above, since $\mathcal{L}_c$ and $\mathcal{L}_t$ coincide, the final clause of the definition is well-defined. Note that in general, C_{Res} and Σ_{Res} will be disjoint, and similarly for C_{Res} and K (contra, e.g., [27, §3], where “an agent’s commitment store is just a subset of its knowledge base”).

The general protocol defined above specifies what the possible moves of Res are. In an *obligatio*, however, we want to say more than what moves are *allowed*, we also want to specify a set of possible moves which are in fact *required*, since in an obligational disputation Res is under obligation to respond to Opp in certain ways. This is done by specifying a more refined protocol. Such a protocol, because it makes reference to the agents’ knowledge bases, will always be defined with respect to a particular DEL model $\mathfrak{M}$. We give as an example Burley’s protocol P^{Bur} for *positio*, introduced in §2. Let Γ_n be the sequence of Res’s move in a dialogue d_n . For a DEL model $\mathfrak{M}$ and context K , $P^{\text{Bur}}(K, \emptyset) = P_u(\emptyset)$ and if n is odd, $P^{\text{Bur}}(K, d_n) = P_u(d_n)$. For n even,

- For $d_0 = m_0 = \text{the } \textit{positum}$,

$$P^{\text{Bur}}(K, d_0) = \begin{cases} \text{concede:}m_0 & \text{iff } \exists w \in W, \mathfrak{M}, w \models m_0 \\ \text{deny:}m_0 & \text{iff } \forall w \in W, \mathfrak{M}, w \not\models m_0 \end{cases}$$

- For $d_n, n > 0$:

$$\begin{aligned} & \text{If } \mathfrak{M} \upharpoonright \Gamma_n \models m_n: & P^{\text{Bur}}(K, d_n) &= \text{concede:}m_n \\ & \text{If } \mathfrak{M} \upharpoonright \Gamma_n \models \neg m_n: & P^{\text{Bur}}(K, d_n) &= \text{deny:}m_n \\ & \text{Otherwise:} & & \\ & \text{If } \mathfrak{M}, w^* \models K_{\text{Res}}m_n: & P^{\text{Bur}}(K, d_n) &= \text{concede:}m_n \\ & \text{If } \mathfrak{M}, w^* \models K_{\text{Res}}\neg m_n: & P^{\text{Bur}}(K, d_n) &= \text{deny:}m_n \\ & \text{If } \mathfrak{M}, w^* \models \neg(K_{\text{Res}}m_n \vee K_{\text{Res}}\neg m_n): & P^{\text{Bur}}(K, d_n) &= \text{doubt:}m_n \end{aligned}$$

Finally, we define two outcome rules for Burley-style *positio*, governing who wins. Generally speaking, Opp wins if he can force Res into inconsistency, and Res wins otherwise. Since any individual *obligatio* = d_n for some finite n , we can define a weak notion of “local” winning: If $m_n = \alpha$, then Opp wins if $\mathfrak{M} \upharpoonright \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M}, \Gamma_n} : a \in A\}, V^{\mathfrak{M}, \Gamma_n} \rangle$ and Res wins otherwise. But even though individual *obligationes* are finite, they are all potentially infinite. This view gives rise to a “global” winning condition: Opp wins if there is some n such that $\mathfrak{M} \upharpoonright \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M}, \Gamma_n} : a \in A\}, V^{\mathfrak{M}, \Gamma_n} \rangle$. Res wins otherwise. In both cases, the only time W will be empty is when $C_{\text{Res}} \models \varphi \wedge \neg \varphi$, that is, over the course of the disputation Res has conceded an inconsistent set, and has thus “responded badly”. Thus, protocol P^{Bur} ensures the dialogical consistency of Res (cf. [31, p. 171] and [13, ch. 3]).

There are also two ways that “responded badly” can be explicated, a broad-grained way and a fine-grained way. On the broad-grained view, we are only interested in whether Opp or Res has locally won, that is, whether Opp has been able to force Res to concede a contradiction, or whether Res has remained consistent in his answers. This is the view generally considered by medieval authors.

Before we discuss some interesting properties of protocols like the ones introduced here, we close this section by formalizing the example in Figure 1. Fix a model $\mathfrak{M}$ such that $\exists w \in W, \mathfrak{M}, w \models \varphi \wedge \psi$, and let $K = \{\neg \varphi\}$. The first move is Opp’s, and he asserts the *positum* $\varphi := d_0$. According to $P^{\text{Bur}}(d_0)$, Res should **concede**: $\varphi := d_1$. The protocol now allows Opp to make any assertion he likes, so he asserts $\neg \varphi \vee \psi := d_2$. Now, by the argument given above, either $\mathfrak{M} \upharpoonright d_1 \models d_2$, or, if not, then $\mathfrak{M}, w^* \models K_{\text{Res}}d_2$. By the protocol, in both cases, Res should concede and $d_3 = \text{concede:}d_2$. Then Opp asserts $d_4 = \psi$, and since $\mathfrak{M} \upharpoonright d_1, d_3 \models \psi$, Res is again required to respond with **concede**: $d_4 := d_5$. Then, Opp calls “*cedat tempus*”, and $d_6 = \alpha$, and, by the general protocol P_u there are no more legal moves and the dialogue ends.

7 Discussion

7.1 The protocol

The protocol P^{Bur} defined above is semi-public, as it depends on Res’s knowledge, but does not depend on Opp’s; context-dependent; and fully deterministic. It also meets all but four of the 13 desiderata for agent argumentation protocols given in [25]. There McBurney et al. consider dialogue protocols from the point of view of design, and identify 13 *desiderata* that a designer might want to incorporate. These are:

- stated dialogue purpose** A system should have one or more publicly stated purpose, and its structure should facilitate its achievement.
- diversity of individual purposes** Participants should be allowed to have their own (distinct) purposes, consistent with the stated dialogue purpose.
- inclusiveness** Any qualified participant who wants to participate can.
- transparency** Participants should know the rules and structure of the system prior to its commencement.
- fairness** Either all participants should be treated equally, or asymmetries in treatment should be explicit.
- clarity of argumentation theory** The dialectical system should conform to a stated theory of argumentation.
- separation of syntax and semantics** The syntax and the semantics of the system should be defined separately.

rule-consistency The rules and locutions should not lead to deadlocks or to infinite cycles of repeated locutions.

encouragement of resolution Resolution should not be precluded by the rules.

discouragement of disruption The rules should discourage disruptive behavior, such as repeating the same locution repeatedly.

enablement of self-transformation Participants should be able to change their preferences, knowledge, degrees of belief, etc.

system simplicity The locutions and rules of the system should be as simple as possible, consistent with the other desiderata.

computational simplicity The system should minimize computational demands on the participants.

In the FDS constructed from Burley’s obligational theory, the dialogue purpose is stated explicitly (cf. §2). The purposes of Res and Opp are distinct, and anyone who is qualified may participate in an *obligatio* as either Res or Opp may. The rules are agreed upon in advance, and the asymmetry between Res and Opp are explicit. The system conforms to a stated theory of argumentation, namely Burley’s rules for *positio*. The rules do not preclude resolution, and it is also quite simple, storable in a case-based structure with minimal cases to consider. Further, to the extent that the protocol is rule-consistent (which we discuss below), it discourages disruption in that Res is never able to continuously repeat the same locution, at least, not without responding badly.

The issue of self-transformation enablement requires further comment. McBurney et al. further specify that self-transformation requires that “participants should have the right to retract commitments made earlier in the same dialogue” for otherwise “in such circumstances, there would be no point for the agents to engage in dialogue” [25, p. 403]. We disagree with this assessment; the *obligationes* framework provides a counterexample, a type of dialogue which is worthwhile engaging in but in which the agents’ commitments are monotonic. *Obligationes* satisfy the less restrictive view of self-transformation, given that the commitment sets and responses of Res are flexible and can change over the course of the disputation.

For the remaining three desiderata, because the topic language and the communication language coincide, it is not clear to what extent *obligationes* satisfy the requirement of the separation of syntax and semantics. It does not satisfy rule-consistency of discouragement of disruption, because Opp can continually put forward the same proposition; however, while the protocol allows Opp to act in such a fashion, there are pragmatic reasons why he will generally not do so. In Res responds well (that is, follows the protocol), the only time he will change his response to a proposition φ is to move from doubt to either concede or deny. If he has responded badly at some point, then Opp may by repeating a locution be able to force Res into losing, in which case there will be no point in Opp repeating that proposition again, as he should instead call *Cedat tempus*. Whether the final criterion is satisfied is uncertain: As we note in the final section, the computational complexity of certain decision problems that can be extracted from this protocol is not yet known. Regardless, the protocol scores quite highly—as well, or better, than the protocols analysed in [25]. It should also be clear that this high score is not specific to Burley’s protocol: Any other FDS protocol developed from a different medieval theory would score similarly high.

7.2 Comparison with other dialogue systems

We discussed Walton & Krabbe’s taxonomy of dialogue types, and Cogan et al.’s extensions, in §4 before introducing Prakken’s frame-

work. Prakken himself distinguishes two types of persuasion dialogues, those that are for *pure persuasion* and those for *conflict resolution* [31, p. 169]. In his terminology, a DS S is for *pure persuasion* iff for any terminated dialogue d , agent a is a winner for topic t ($a \in w_t(d)$) iff either a is a proponent of t and $t \in C_{a'}(d)$ for all a' who are either proponents or opponents of t , or a is an opponent of t and $t \notin C_{a'}(d)$ for all a' who are either proponents or opponents of t . A DS S is for *conflict resolution* otherwise.

Where do *obligationes* fit in these schemes? Prakken says that frameworks for persuasion dialogues can be found in the Middle Ages [31, p. 163, citing [1]]. Angelelli’s discussion of medieval disputations in [1] focuses on *obligationes*, but nowhere does he make any mention of persuasion. If *obligationes* are a type of persuasion dialogue, then because Opp’s commitment store is always empty in an *obligatio*, *obligationes* would be classified, on Prakken’s distinction, as conflict resolution dialogues instead of pure persuasion dialogues. However, given the stated goal of *obligationes* and the fact that the players agree on the truth value of the proposition in question at the start of the dialogue, this classification does not seem appropriate.

Despite its breadth, the decempartite division of [9], incorporating Walton and Krabbe’s typology, also does not accommodate *obligationes*. Because they are about the consistency of a formula, *obligationes* are not negotiation or deliberation dialogues. Because the truth value of the proposition in dispute is known to both, and the Opponent is not trying to persuade the Respondent of anything, they are not information-seeking, inquiry, or pure persuasion dialogues. Since they are not pugilistic in nature, they are not eristic dialogues. Nor are they any of the four new kinds introduced in [9], since those types require as well that at least one party not know the truth-value of the proposition.

Obligationes are somewhat similar to the ‘elicit-inform’ dialogue game of [21, 32]. These dialogues were developed by Ravenscroft and Matheson in the context of collaborative e-learning, where the dialogues are between a tutor system and a student. The asymmetry between the players is similar to the asymmetry between the Opponent and the Respondent. In an elicit-inform dialogue, the student is questioned by the tutorial system, and “after reasoning about the learner’s contributions, the tutor system either *sanctions* their explanations by informing them they were correct, or points out that they were ‘incorrect’ and so *informs* them of a consistent, or ‘correct’ answer” [32, p. 96]. This is very similar to the behavior of the Opponent when he calls *Cedat tempus* and evaluates the actions of the Respondent to determine whether he has responded well or badly. However, as elicit-inform dialogues have as their goal the persuasion of the student to adopt a certain belief, they are not a perfect match for *obligationes*, since persuasion is not at issue in obligational dialogues.

Instead of trying to shoehorn *obligationes* into a type of dialogue system already introduced in the literature, we think it is more interesting, and highlights the unique nature of obligational reasoning better, to consider them as introducing a new type of dialogue into the typology. Thus, one of the contributions of the current paper is the introduction of a new type of inter-agent dialogue, for checking the feasibility of a set of propositions. Proving that a proposition or a set of propositions is feasible—can be consistently maintained—has various applications in constraint and allocation satisfaction, where a consistent solution meeting certain requirements has to be found. Thus, the new dialogue type that we have introduced may turn out to have useful applications beyond the context of *obligationes*, though we do not investigate this potential extension further here.

7.3 The role of obligations

One final point of interest, in comparing protocols based on *obligationes* to other dialogue protocols, is the role of obligations (in the ordinary sense of the term) in the dialogues. Generally, in an FDS, “[s]trictly speaking the only dialectical obligation that a participant has is making an allowed move when it is one’s turn” [31, p. 170]. In an *obligatio*, however, Res has a two-tiered obligation: He is required to follow both the uniform protocol P_u , and the appropriate specific protocol for the type of *obligatio* he is in.⁶ It is possible for Res to violate his obligation to follow the second protocol (in which case he loses), but if he does not follow the uniform protocol than no disputation even takes place. This two-tiered nature of the obligation of the Res helps us understand King’s analysis of *obligationes* as a meta-methodology of argumentation. The specific protocol is the methodology—it tells Res how to respond within a particular disputation—while the general protocol constrains the types of specific protocols that are allowed, and hence can be understood as a meta-methodology (a higher order method).

Additionally, in an *obligatio* there is no connection between an agent’s commitment store and his assertions; Opp has no commitments, even though all he makes are assertions, and Res makes no assertions, but his commitments are generated from his concessions and denials of Opp’s assertions. This is a significant difference from standard commitment rules such as the one discussed by Prakken in his Paul and Olga example [31, p. 169]. He says “As for commitment rules, the following ones seem to be uncontroversial and can be found throughout the literature:

- If $s(m) = \text{claim}(\varphi)$ then $C_s(d, m) = C_s(d) \cup \{\varphi\} \dots$ ”

That is, an agent’s discursive commitments are generated from his assertions. A similar position is advocated when he says elsewhere that “Commitments are typically incurred by making or conceding claims and stating” [30, p. 1017].

8 Conclusion

We have now seen how at least one type of medieval obligational theory can be interpreted as giving rise to a formal dialogue system; it is straightforward to extend this analysis not just to theories of *positio* outlined by other medieval authors but also other types of *obligationes*, such as *dubitatio*, by varying the underlying logic.⁷ The result of such an analysis shows that, just as a particular dialogue can be viewed as a game played according to a set of rules specified by an FDS (cf. §3), so too *obligationes* can be naturally understood as giving the participants a methodology of argumentation or reasoning to follow. By specifying the protocols and rules of an FDS, an particular obligational theory gives participants a framework within which to do philosophical analysis. This provides formal support for King’s interpretation of *obligationes* as functioning at the meta-level, rather than at the content level. We close our paper by pointing towards questions that we hope to answer in future work.

In §6 we distinguished two ways “responds badly” can be explained. Both of these give rise to decision problems whose complexity we intend to investigate in future work (cf. §7). In the first, the local winning outcome conditions corresponds to the decision problem

⁶ The obligation to follow the uniform protocol is what Walton and Krabbe term an ‘action commitment’: an obligation to execute a particular course of action, given the action(s) of the other player(s) [39, ch. 1].

⁷ The case of *dubitatio* is considered in [37].

RESPONDS-WELL($d_n, \mathfrak{M}$) defined as follows: Given a finite *obligatio* d_n and model $\mathfrak{M}$, check whether $\mathfrak{M} \models \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M}, \Gamma_n} : a \in A\}, V^{\mathfrak{M}, \Gamma_n} \rangle$, that is, whether Opp has locally won.

Question. What is the computational complexity of RESPONDS-WELL($d_n, \mathfrak{M}$)?

Note: The answer is not necessarily a straightforward adaptation of complexity results for model-checking in Public Announcement Logic (PAL, [38, ch. 4]), since in PAL, attention is restricted to announcements which are *truthful*, and the introduction of announcements which may be false, but which are nonetheless effective (in that they reduce the model), is not generally considered.

The fine-grained view asks whether, at each step n , Res has followed the specific protocol he was obligated to follow. It is possible for Res to have responded correctly in the sense of winning locally on the broad-grained view, but to still have not followed the rules correctly, by choosing the wrong response for irrelevant propositions. In this case, the decision problem RULE-FOLLOWING($d_n, P, \mathfrak{M}$) is defined as follows: Given a finite *obligatio* d_n , protocol P , and model $\mathfrak{M}$ check whether the construction of Γ_n satisfies the conditions of P .

Question. What is the computational complexity of RULE-FOLLOWING($d_n, P, \mathfrak{M}$)?

Note: It may be possible to extend the complexity results of [27, 28, 29], though it is not *prima facie* clear how this would be done. In particular, the results in [29, §7] are for argumentation-based logics, which DEL is not, and the results in [27, §5–6] are for protocols where repetition is not allowed.

ACKNOWLEDGEMENTS

This research was funded by the NWO project “Dialogical Foundations of Semantics” (DiFoS) in the ESF EuroCoRes programme Log-ICCC (LogICCC-FP004; DN 231-80-002; CN 2008/08314/GW). The author would like to thank the anonymous referees for their meticulous and useful comments on an earlier version of this paper.

REFERENCES

- [1] I. Angelelli, ‘Techniques of disputation in the history of logic’, *Journal of Philosophy*, **67**(20), 800–815, (1970).
- [2] M. Baker, ‘The role and models in artificial intelligence and educational research: A prospective view’, *International Journal of Artificial Intelligence in Education*, **11**, 122–143, (2000).
- [3] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni, ‘An abstract, argumentation-theoretic approach to default reasoning’, *Artificial Intelligence*, **93**, 63–101, (1997).
- [4] H.A.G. Braakhuis, ‘Obligations in early 13th century Paris: the *Obligations* of Nicholas of Paris (?)’, *Vivarium*, **36**(2), 152–233, (1998).
- [5] G. Brewka, ‘Dynamic argument systems: a formal model of argumentation processes based on situation calculus’, *Journal of Logic and Computation*, **11**(2), 257–282, (2001).
- [6] W. Burley, ‘Obligations (selections)’, in *Cambridge Translations of Medieval Philosophical Texts*, eds., N. Kretzmann and E. Stump, volume 1, 369–412, Cambridge University Press, (1988).
- [7] W. Burley and R. Green, O.F.M., *An Introduction to the Logical Treatise ‘de Obligationibus’*, volume 2, Université Catholique Louvain, 1963.
- [8] E. Cogan, S. Parsons, and P. McBurney, ‘What kind of argument are we going to have today?’, in *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, (2005).
- [9] E. Cogan, S. Parsons, and P. McBurney, ‘New types of inter-agent dialogues’, in *Argumentation in Multi-Agent Systems*, eds., S. Parsons, N. Maudet, P. Moraitis, and I. Rahwan, LNAI 4049, 154–168, Springer, (2006).
- [10] L. M. de Rijk, ‘Some thirteenth century tracts on the game of obligation’, *Vivarium*, **12**(2), 94–123, (1974).

- [11] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation II', *Vivarium*, **13**(1), 22–54, (1975).
- [12] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation III', *Vivarium*, **14**(1), 26–49, (1976).
- [13] C. Dutilh Novaes, *Formalizing Medieval Logical Theories: Suppositio, Consequentiae and Obligationes*, volume 7 of *Logic, Epistemology, and the Unity of Science*, Springer, 2007.
- [14] R. A. Gire, 'Commands in dialogue logic', in *Practical Reasoning: Proceedings of the First International Conference on Formal and Applied Practical Reasoning (FAPR 1996)*, Bonn, Germany, eds., D. M. Gabbay and H. J. Ohlbach, volume 1085 of *LNAI*, pp. 246–260. Springer, (1996).
- [15] C. Hamblin, 'Mathematical models of dialogue', *Theoria*, **37**, 130–155, (1971).
- [16] C. L. Hamblin, *Fallacies*, Methuen, London, [1970].
- [17] D. Harel, D. Kozen, and J. Tiuryn, 'Dynamic logic', in *Handbook of Philosophical Logic*, eds., D. M. Gabbay and F. Guenther, volume 4, Kluwer, 2 edn., (2002).
- [18] N. C. Karunatillake, N. R. Jennings, I. Rahwan, and P. McBurney, 'Dialogue games that agents play within a society', *Artificial Intelligence*, **173**, 935–981, (2009).
- [19] P. King, 'Opposing and responding: comments on Paul Spade'. preprint, http://individual.utoronto.ca/pking/presentations/Spade_Comments.pdf, 2004.
- [20] P. Lorenzen and K. Lorenz, *Dialogische Logik*, Wissenschaftliche Buchgesellschaft, 1978.
- [21] M. Matheson and A. Ravenscroft, 'Evaluating and investigating learning through collaborative argumentation: An empirical study', Technical Report DDRG-01-01, The Open University, Milton Keynes, (2001).
- [22] N. Maudet, 'Negotiating dialogue games', *Autonomous Agents and Multi-Agent Systems*, **7**, 229–233, (2003).
- [23] P. McBurney and S. Parsons, 'Games that agents play: A formal framework for dialogues between autonomous agents', *Journal of Logic, Language and Information*, **11**, 315–334, (2002).
- [24] P. McBurney and S. Parsons, 'Dialogue games for agent argumentation', in *Argumentation in Artificial Intelligence*, eds., I. Rahwan and G. Simari, 261–280, Springer, (2009).
- [25] P. McBurney, S. Parsons, and M. Wooldridge, 'Desiderata for agent argumentation protocols', in *Proceedings of AAMAS'02, July 15–19, 2002, Bologna, Italy*, pp. 402–409, (2002).
- [26] W. of Ockham, *Opera Philosophica*, volume I, Franciscan Institute, St. Bonaventure, 1974.
- [27] S. Parsons, M. Wooldridge, and L. Amgoud, 'An analysis of formal inter-agent dialogues', in *Proceedings of AAMAS'02, July 15–19, 2002, Bologna, Italy*, pp. 394–401, (2002).
- [28] S. Parsons, M. Wooldridge, and L. Amgoud, 'On the outcomes of formal inter-agent dialogues', in *Proceedings of the 2nd International Conference on Autonomous Agents and Multi-Agent Systems, Melbourne*, (2003).
- [29] S. Parsons, M. Wooldridge, and L. Amgoud, 'Properties and complexity of some formal inter-agent dialogues', *Journal of Logic and Computation*, **13**(3), 347–376, (2003).
- [30] H. Prakken, 'Coherence and flexibility in dialogue games for argumentation', *Journal of Logic and Computation*, **15**(6), 1009–1040, (2005).
- [31] H. Prakken, 'Formal systems for persuasion dialogue', *Knowledge Engineering Review*, **21**(2), 163–188, (2006).
- [32] A. Ravenscroft and M. P. Matheson, 'Developing and evaluating dialogue games for collaborative e-learning', *Journal of Computer Assisted Learning*, **18**(1), 93–101, (2002).
- [33] E. Sklar and S. Parsons, 'Toward the application of argumentation-based dialogues for education', in *Proceedings of the 3rd International Conference on Autonomous Agents and Multi-Agent Systems*, eds., C. Sierra and E. Sonenberg, (2004).
- [34] P. V. Spade, 'Why don't medieval logicians ever tell us what they're doing? or, what is this, a conspiracy?'. preprint, <http://pvspade.com/Logic/docs/Conspiracy.pdf>, 2000.
- [35] P. V. Spade. Unpublished talk on *obligationes* presented at the Midwestern Division of the American Philosophical Association, April 1993.
- [36] S. L. Uckelman. 'A unified dynamic framework for modeling *obligationes*', 2010. preprint, <http://staff.science.uva.nl/~suckelma/latex/jolli-ob/jolli-ob.pdf>, in submission.
- [37] S. L. Uckelman, 'Deceit and nondefeasible knowledge: The case of *dubitatio*'. preprint, <http://staff.science.uva.nl/~suckelma/latex/loft2010/loft2010.pdf>, 2010.
- [38] H. van Ditmarsch, W. van der Hoek, and B. Kooi, *Dynamic Epistemic Logic*, volume 337 of *Synthese Library Series*, Springer, 2007.
- [39] D. N. Walton and E. C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State University of New York Press, 1995.
- [40] M. Yrjönsuuri, 'The role of casus in some fourteenth century treatises on sophismata and obligations', in *Argumentationstheorie. Scholastische Forschungen zu den logischen und semantischen Regeln korrekten Folgerns*, ed., K. Jacobi, 301–321, Brill, (1993).
- [41] M. Yrjönsuuri, *Obligationes: 14th Century Logic of Disputational Duties*, volume 55 of *Acta Philosophica Fennica*, Societatis Philosophia Fennica, 1994.
- [42] M. Yrjönsuuri, ed., *Medieval Formal Logic*, Kluwer, 2001.

Working on the Argument Pipeline: Through Flow Issues between Natural Language Argument, Instantiated Arguments, and Argumentation Frameworks

Adam Wyner¹ and Tom van Engers² and Anthony Hunter³

Abstract. In many domains of public discourse such as arguments about public policy, there is an abundance of knowledge to store, query, and reason with. To use this knowledge, we must address two key general problems: first, the problem of the knowledge acquisition bottleneck between forms in which the knowledge is usually expressed, e.g. natural language, and forms which can be automatically processed; second, reasoning with the uncertainties and inconsistencies of the knowledge. Given such complexities, it is labour and knowledge intensive to conduct policy consultations, where participants contribute statements to the policy discourse. Yet, from such a consultation, we want to derive policy positions, where each position is a set of consistent statements, but where positions may be mutually inconsistent. To address these problems and support policy-making consultations, we consider recent automated techniques in natural language processing, instantiating arguments, and reasoning with the arguments in argumentation frameworks. We discuss application and “bridge” issues between these techniques, outlining a pipeline of technologies whereby: expressions in a controlled natural language are parsed and translated into a logic (a literals and rules knowledge base), from which we generate instantiated arguments and their relationships using a logic-based formalism (an argument knowledge base), which is then input to an implemented argumentation framework that calculates extensions of arguments (an argument extensions knowledge base), and finally, we extract consistent sets of expressions (policy positions). The paper presents a progress report towards the implementation of dynamic, web-based, distributed, collaborative knowledge base construction in natural language which supports incomplete and inconsistent knowledge.

1 INTRODUCTION

In many domains of public discourse, participants have the opportunity to present their views. In public policy-making discussions, members of the public and politicians gather information and consult about issues of public import such as copyright, banking, immigration and others. Such consultations promote democratic institutions, support regulatory compliance, and make government more efficient. With the advent of the Internet, the opportunities to broaden

such public participation have increased dramatically as has the volume of information. While consultation exercises are necessary and beneficial, the volume and complexity of the information introduces problems to store, analyse, represent, and reason. Moreover, the participants may enter uncertain, incomplete, or inconsistent statements; the knowledge may be non-monotonic, where statements are introduced or removed, context changes, and inferences change relatedly. These create substantive problems in reasoning.

Standard, current tools which are used to gather knowledge of policy are face-to-face meetings and rounds of consultation/commenting on written position reports. On line forums are also used to broaden participation and make policy more efficiently [23]. Broadly speaking, the meetings and on line forums serve to record the uncertain and potentially inconsistent points of view of the participants. Analysts subsequently try to identify and reason with the inconsistency in order to formulate and clarify the policy, a task which is done “manually”. Moreover, the tools do not support participants to submit their contribution in a way that facilitates further processing.

None of the current tools address the knowledge acquisition bottleneck, nor knowledge representation and reasoning (KR) as understood in the subfield of artificial intelligence. The knowledge acquisition bottleneck is the problem of translating between the form in which participants know or express their knowledge and the formal representation of knowledge which a machine can process; the bottleneck has limited the advance of artificial intelligence technologies [18]. KR is understood as a structured, logical representation of the knowledge for automated reasoning, querying, and processing. Currently, knowledge engineers (or legislative assistants) manually structure and formalise unstructured linguistic information into a knowledge base (or legislative act) which supports information extraction, identification of relationships among statements, inference, query, development of ontologies, redundancy, identification of contradictions, reasoning with uncertain and inconsistent knowledge, or visualisations. Yet, in doing so, the knowledge engineers rely on their own fine-grained, unstructured, implicit linguistic knowledge, which itself is problematic.

In order to support policy-making, we address the knowledge acquisition bottleneck and inconsistency of statements with automated techniques in natural language processing, instantiating arguments, and reasoning with the arguments in argumentation frameworks. We discuss application and *bridge* issues between these techniques, outlining a pipeline of technologies whereby: expressions in a controlled natural language are automatically parsed and translated into a logic

¹ Corresponding author. University of Leeds, United Kingdom, email: adam@wyner.info. 2010 ©Adam Wyner, Tom van Engers, and Anthony Hunter.

² Leibniz Center for Law, The Netherlands, email: vanEngers@uva.nl

³ University College London, United Kingdom, email: a.hunter@cs.ucl.ac.uk

(a literals and rules knowledge base – LRKB) using *Attempto Controlled English* (ACE)⁴; from this knowledge base, we generate instantiated arguments and their relationships using a logic-based formalism (an argument knowledge base – AKB) with a system based on [5]; we input this knowledge base into an implemented argumentation framework which calculates extensions of arguments (an argument extensions knowledge base AEKB) with [17] that is grounded in [14]. At the end, we extract consistent sets of expressions (policy positions).⁵

While each of these components is well developed independently, little consideration has been focused on the *integration* and *through flow* between them. The important novelty of this paper is an exploration of these bridge issues which must be addressed to construct an on line system which allows users to enter expressions in natural language that feed into an implemented logic-based formalism and argumentation frameworks, creating an *argument processing pipeline*. The objective of the pipeline is support and automate what is now an unsupported and manual task: input a (possibly inconsistent, dynamic) knowledge base of literals and rules expressed in natural language and output *policy positions*.

Many questions arise. Where we rely on input from public participants, who are not logicians or knowledge engineers with training in building well-formed rules, ill-formed arguments could be entered. What prompts can be introduced to make KB construction systematic and meaningful? Is Propositional or First-order logic appropriate as a semantic formalism for natural language. What limitations are imposed by using a controlled language? How can context and dynamics be addressed? In the face of these questions, our approach is pragmatic and incremental, using available technologies while acknowledging limitations. For example, we do not address the problem of ill-formed arguments; First-order logic is known to have a range of limitations; controlled languages are inherently less expressive than 'full' natural languages, and we consider only *static* knowledge bases relative to a *fixed context*. These are abstractions which helps us focus on other central aspects of the system. Given further theoretical and technological advances, our approach is to be extended and refined towards the goal of an argumentation system which supports dynamic, web-based, distributed, collaborative knowledge base construction using natural language and reasoning with incomplete and inconsistent knowledge. This paper represents an intermediate report towards that goal.

In Section 2, we provide a sample policy-making discussion, a graphical representation of the relationships among the statements, and our translation methodology using the Attempto Controlled Language System. In this section, we develop a literals and rules knowledge base (LRKB), which represents the domain knowledge. In Section 3, we take LRKB and instantiate arguments for and against statements, forming an argument knowledge base (AKB). In Section 4, we take the arguments and relations defined in AKB as input to an argumentation framework AF and calculate extensions of arguments, which are sets of mutually compatible arguments, which forms an argument extensions knowledge base (AEKB). From the extensions, we extract the literals and rules for each extension to provide *policy positions* as shown in Section 5. The paper concludes with a discussion section.

2 Example, Translation Methodology, and Considerations

In this section, we discuss our source material and working example, the methodology of using the Attempto Controlled Language system to translate the sentences to parse and semantically translate the sentences, and considerations which arise in developing a set of sentences for input to an argumentation system.

2.1 Sources and Examples

[40] present a policy-making discussion adapted and abstracted from a BBC *Have Your Say* discussion list. The BBC's *Have Your Say* introduces a question for discussion, provides brief background points, then allows users to write in comments in a monitored, unthreaded discussion list.⁶ For the particular study, [40] considered comments in response to the question *Should people be paid to recycle?*, where 697 comments were published (and 35 were rejected).⁷

The comments as they are present a knowledge engineer with a wealth of unstructured textual data since the discussion list is unthreaded and has very few substantive constraints on what participants can contribute or how. Among the many issues such a discussion raises for the analyst, we see the following. Participants may use any vocabulary or syntactic structure they wish (so long as it does not lead to the rejection of their comment); they may use novel words or expressions, metaphors, misspellings, and ungrammatical sentences. There is no check that the meaning as interpreted is the meaning as intended. Implicit or pragmatic information in the comments must be inferred by the reader. There are no over indications of inter-comment relationship (unless the commenter explicitly refers to one or more previous comments). The comments can represent uncertain or inconsistent information. The comments can be vague and ambiguous. The interpretation of comments is left to the analyst to infer.

Beyond these empirical observations, it is unclear what current computational linguistic system could accurately parse the large majority of sentences, yield the acceptable correlated semantic interpretation, and be used to support the acquisition of the knowledge. Text-mining can extract some useful information (e.g. sentiment analysis [25], some aspects of rules [39], automatic implicatures [13], and contradictions [32]), but is still limited, error prone, and does not capture the complex meaning of each comment, much less the complex web of meanings comprised of the semantic relationships among the comments. Uncertainties and inconsistencies (other than explicit contradictions) are difficult to identify. It would be a significant task to take such a list of comments and transform them into a knowledge base that is suitable for a significant KR task.

[40] scope the issues, which is a common practice in semantic studies in Linguistics. By scoping the issues, the data is filtered, normalised, and reduced in scale so as to define a problem space that is on the one hand close to or derived from the source data and is on the other hand amenable to yet challenging for the theoretical and implemented techniques. It allows us to address and solve particular problems, abstract issues, and make progress which otherwise is prohibited by an ill-defined scope. While the resulting example is relatively far from the source material, it is still a useful starting point. The intention is that by addressing relatively artificial examples, we

⁴ <http://attempto.ifi.uzh.ch/site/description/>

⁵ The term policy position is perhaps related to *normative positions* of deontic logic [29].

⁶ http://news.bbc.co.uk/2/hi/talking_point/

⁷ See <http://newsforums.bbc.co.uk/nol/thread.jspa?forumID=7269&edition=2&ttl=20100218141845>.

can highlight issues, then extend the approach incrementally to solve additional and less artificial problems and examples.

With these programmatic points in mind, we adopt and adapt the approach and sentences in [40], where there are 16 sentences are automatically translated into First-order logic, manually indicating where there are rules and contradictions: users input sentences using the controlled language editor, then indicate *rule relationships* between sentences such as *premise*, *claim*, or *contradiction*. The result is a (possibly inconsistent) knowledge base.

In the current paper, we have 19 example sentences; in Section 2.3, we justify our revisions and discuss the sentences. For the purposes of this paper, we only work with Propositional Logic in this paper because the theory of argument formation for expressions in First-order logic is more complex than we need here and because the current implemented system for generating arguments and relations only works for Propositional Logic, though a system for First-order Logic is in development. A representation in First-order Logic is crucial in order to *identify* and *reason with* contradiction, redundancy, implication, ontological argument, fine-grained information extraction, and query; in future work, we anticipate working with First-order Logic Statements. We have, therefore, assigned each sentence a lower case propositional variable. Each rule has a lower case rule name.⁸

Example Sentences

- p1: Every household should pay some tax for the households garbage.
- p2: No household should pay some tax for the households garbage.
- p3: Every household which pays some tax for the households garbage increases an amount of the households garbage which the household recycles.
- p4: If a household increases an amount of the households garbage which the household recycles then the household benefits the households society.
- p5: If a household pays a tax for the households garbage then the tax is unfair to the household.
- p6: Every household should pay an equal portion of the sum of the tax for the households garbage.
- p7: No household which receives a benefit which is paid by a council recycles the households garbage.
- p8: Every household which does not receive a benefit which is paid by a council supports a household which receives a benefit which is paid by a council.
- p9: Tom says that every household which recycles the households garbage reduces a need of a new dump which is for the garbage.
- p10: Every household which reduces a need of a new dump benefits the households society.
- p11: Tom is not an objective expert about recycling.
- p12: Tom owns a company that recycles some garbage.
- p13: Every person who owns a company that recycles some garbage earns some money from the garbage which is recycled.
- p14: Every supermarket creates some garbage.
- p15: Every supermarket should pay a tax for the garbage that the supermarket creates.
- p16: Every tax which is for some garbage which the supermarket creates is passed by the supermarket onto a household.
- p17: No supermarket should pay a tax for the garbage that the supermarket creates.
- p18: Tom is an objective expert about recycling.

⁸ One general issue that arises about the pipeline concerns the representations which suit each application. This is discussed further below.

p19: If an objective expert says every household which recycles the households garbage reduces a need of a new dump which is for the garbage, then every household which recycles the households garbage reduces a need of a new dump which is for the garbage.

As in [40], where a rule has more than one premise, the assumption is that the premises must conjunctively hold for the claim to hold. The rules are *definite clauses* - Horn clauses with exactly one positive literal. We assume that users construct rules from the propositions.

Rule Set

- r1: $[p4 \wedge p3] \rightarrow p1$
- r2: $[p9 \wedge p18 \wedge p19] \rightarrow p4$
- r3: $p10 \rightarrow p3$
- r4: $[p12 \wedge p13] \rightarrow p11$
- r5: $[p5 \wedge p15] \rightarrow p2$
- r6: $[p6 \wedge p7 \wedge p8] \rightarrow p5$
- r7: $p14 \rightarrow p15$
- r8: $p16 \rightarrow p17$

We can represent the structure of the literals and rules into a graphic where solid lines indicate a premise relation and a dashed line indicates a contradiction.

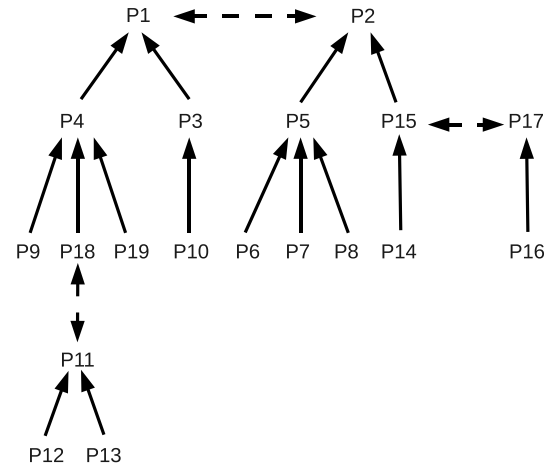


Figure 1. Relationships between Statements – Premises, Claims, Contradictions

In the next section, we briefly outline issues bearing on the syntactic parsing and semantic interpretation provided by ACE.

2.2 Parsing and Semantic Representation with Attempto Controlled English

[40] provide a detailed discussion of the application of Attempto Controlled English (ACE) to our candidate sentences. In the context of this paper, where the emphasis is on providing the relationship between the input statements, rules, and an argumentation system which works only on propositions, the importance of using ACE is reduced. Yet, it is relevant here to give a brief overview in that ACE

controls the language of the policy discussion and also provides expressions in First-order Logic that can be used in future implemented argument instantiation systems ([4] and [16]).

To facilitate the processing of sentences, we use a well-developed controlled natural language system – ACE (for an overview of the properties and prospects of controlled natural languages, see [37]). A controlled language has a specified vocabulary and a restricted range of grammatical constructions which are a subset of a natural language (e.g. English). Sentences written and read in the controlled language appear as normal sentences, but can be automatically translated into a formal representation.

ACE supports a large lexicon, a range of grammatical constructions, discourse anaphora, and correlated semantic interpretations: negation on nouns or verbs, conjunction, disjunction, conditionals, quantifiers, adjectives, relative clauses, discourse anaphora, modals (e.g. *necessity* and *possibility*), possessives, prepositional phrases, verbs with three arguments, and verbs with subordinate clauses. ACE checks that the sentences input to the system satisfy the constraints of the syntax and semantics of the language, thus the user is only able to input grammatically acceptable and semantically interpretable sentences in building the knowledge base. ACE has an associated reasoning engine to test for consistency and draw inferences.

As discussed in [40], the syntactic form must be crafted to generate the intended semantic interpretation, but this is relatively straightforward to do. Our 19 sentences are grammatically well-formed in ACE and have the intended interpretations.⁹

2.3 Considerations

In developing the sentences for input to a generator of instantiate arguments, three issues arose – making contradictoriness explicit, the construction of well-formed rules in the LRKB, and the introduction of implicit premises. We discuss these in turn.

To generate instances of arguments and their relations, we require explicit indicators of contradictoriness. To take a simple example, suppose that one participant makes a statement such as *Bill exercises daily* and another claims *Bill does not exercise daily*; the lexical items and syntactic form are such that it is relatively easy to identify these as contradictory. Systematically translating from natural language expressions with negation to logical forms with negation is a complex issue itself, though we do not discuss it here. However, rather than stating *Bill does not exercise daily*, suppose a participant claims *Bill is a couch potato*, which is also contradictory with *Bill exercises daily*. Yet, the lexical items and syntactic form are significantly different; it requires lexical semantic and syntactic knowledge to know that these expressions are contradictory. Indeed, there are a broad range of linguistic forms for expressing contradictoriness [11]; more broadly, contradictoriness is a rich, complex, poorly understood domain ([11], [9], [12, 32]). While ACE easily identifies the inconsistency between *Bill exercises daily* and claims *Bill does not exercise daily*, it misses the inconsistency between *Bill exercises daily*, and *Bill is a couch potato*.

There are two approaches to addressing such contradictions. One strategy is for the user to introduce an explicit *constraint* which simply states which two sentences are contradictory. If *Bill exercises*

daily is represented as the propositional variable P, *Bill does not exercise daily* is $\neg P$, and *Bill is a couch potato* is R, $\neg[P \wedge R]$ is a constraint – P and R cannot both be true in a model. Yet, unlike the contradiction $[P \wedge \neg P]$ of Propositional Logic, constraints must be independently given in the knowledge base. Besides having to be specified, an additional problem of constraints is that there is no explanation of the contradiction; that is, they are *facts* of the knowledge base and follow from or relate to no other elements of the knowledge base.

To overcome this, an alternative strategy is to express the *enthymemes* or rules (if known) which give rise to the inconsistency; enthymemes are implicit premises that are required to allow an inference to go through and which must be made explicit in the KB. For instance, *Bill is a couch potato* implies that *Bill does not exercise daily*. Thus, in stating *Bill is a couch potato* one ought to introduce as well the rule that it implies that *Bill does not exercise daily*; then, where it is asserted that *Bill is a couch potato* and *Bill exercises daily*, we would find, by inference, that an inconsistency arises. Such an approach relies on some pre-existing and accepted terminology by the users, that is, that *exercise daily* is in the lexicon already such that the contradiction of it is meaningful.

A second issue relates to the construction of well-formed rules. Given a set sentences, participants can, in principle, introduce any combination of literals to produce a rule. Yet, some of these appear to be well-formed and meaningful, while others do not. For example, the following argument is formally and syntactically well-formed as well as semantically coherent: *Socrates is a man* and *Every man is mortal* therefore *Socrates is mortal*. In contrast, the following is formally and syntactically well-formed, but semantically incoherent, we would not want to accept it as an argument: *Bill is rich* and *Jill sails* therefore *Socrates is mortal*. Clearly in the latter example, we are missing some semantically relevant relationships between the premises themselves and between the premises taken together and the claim; in this regard, the former example is more coherent. While the former example may have a relatively simple explanation for this coherence (terms are shared), the coherence of more complex examples such as argumentation schemes which represent normative patterns of presumptive reasoning are more difficult to explain [34]. On this topic, we have some of the following general questions. What are the conditions on semantic well-formedness for the rules provided to the knowledge base? What makes a set of statements semantically cohere as premises of a claim? What makes a set of statements semantically imply a conclusion? While the mathematical or syntactical conceptions of a rule may be clear, and logicians or knowledge engineers may implicitly know how to write well-formed rules, the principles which underlie the linguistic semantic instantiations are not clear. If we allow non-logicians or knowledge engineers to build rules, what guidance can be provided to build meaningful rules? That is, how can we make explicit what logicians and knowledge engineers implicitly know?

These topics touch directly on complex issues relating to the interface of human reasoning, language, and formal reasoning ([15] and [31]); indeed, in addition to the issues that arise about controlled languages, it is problematic to presume that participants can, without training, contribute to a normative *rational* discussion in a way which supports automated reasoning; a similar point applies to the reconstruction of arguments from texts by participants. In other words, a degree of literacy in operating the system and interpreting the results is going to be required.

A related point bears on enthymemes, which are implicit premises that nonetheless bear on the inference ([7] and [35]). There are two

⁹ ACE translates sentences such as *Every supermarket creates some garbage*. as generally done in First-order Logic, $\forall x[\text{supermarket}(x) \rightarrow \exists y[\text{garbage}(y) \wedge \text{create}(x,y)]]$. However, generalised quantifiers are the more desirable translation of quantifiers for natural language [2]. ACE also translates conditional statements as conditionals, though we have not included conditional statements among our rules in order to simplify. This does not impact on the key points in the paper.

sides to the issue – how enthymemes are used in reasoning and how they are incorporated into the knowledge base. The former presumes a solution to the latter since enthymemes are *implicit* in reasoning, which implies that they are *explicitly* represented in the knowledge base in some way. We do not address the former since in developing a knowledge acquisition device, we are concerned with soliciting implicit knowledge, making it explicit, and formally incorporating it into the knowledge base.

While all three of these issues are substantial in and of themselves, they are partially addressed in our approach to collaborative, incremental, distributive knowledge base construction. For example, suppose there are two sentences in the knowledge base which are mutually incompatible (e.g. our *Bill exercises daily* and *Bill is a couch potato*) but were not overtly recognised or marked as such by users; this could arise simply because of what the users attend to or the size of the knowledge base. Some subsequent users could (on separate occasions and without being aware of the input of other participants) introduce the relevant implications to yield the representation of contradiction. Similarly, rules which are constructed by some parties could subsequently be rendered inapplicable by asserting that a premise does not hold. Finally, as implicit knowledge is solicited, it is explicitly introduced into the knowledge base which can then support further explicit reasoning. In this industrialised vision of knowledge engineering, individual participants may be logically imperfect and partially ignorant, yet contribute to the construction of globally perfect knowledge bases which arise after the calculation of argument extensions. In this regard, the system does not so much rule out erroneous statements or rules as to allow them to be modified by the input of other participants and, more importantly, reasoned with to determine consequences which would highlight the problems.

Finally, we should touch on one representational issue. At different stages of the pipeline, different forms of representation are useful to support distinct reasoning processes. There is a balance to be struck at every step between abstraction and expressiveness. For instance, while we have used lower case letters and numbers, we could have used the sentence strings themselves as propositional variables. While we have left resolution of these issues to future work, it is not particularly problematic so long as we are clear about and consistent about the conversions at each state.

3 INSTANTIATING ARGUMENTS AND RELATIONSHIPS FROM A KNOWLEDGE BASE

In this section, we discuss instantiating arguments and their relationships from a *literals and rules knowledge base* (LRKB) as given above; the result is an *argument knowledge base* (AKB). The AKB is then input to an abstract argumentation framework to calculate extensions as discussed in Section 4; the result is an *argument extensions knowledge base* (AEKB), from which we can extract consistent sets of propositions which represent policy positions.

We work with a *logic-based* approach of [5] and [6], which represents arguments in terms of classical logic (for related work see ([28], [19], [21], [1], [5], and [27])). While we could discuss other approaches to instantiating arguments and relationships which use defeasible rules (e.g. [27] or [20]), we keep to a logic-based approach for several reasons: it is founded in a well-known and widely used logic (classical propositional logic), it has an extension to First-order Logic, it is broadly compatible with the logical translations provided by *ACE* (which has no defeasible rules), an implementation is available, and issues about generating and structuring arguments in rela-

tions are well-developed (e.g. minimal arguments, redundancy, and argument tree pruning among others). However, as we are primarily interested to explore an implemented example, we do not examine these issues further.

[5] and [6] provide numerous examples of knowledge bases and instantiated arguments primarily in Propositional Logic. The novelty of our example is twofold: our example is an abstraction from input sentences which have been automatically translated to First-order Logic; we have related the arguments and relations to an argumentation framework, where, as we see, additional issues arise.

Instantiating arguments and their relationships is necessary for calculating extensions of arguments, which are arguments which are mutually consistent. Abstract argumentation frameworks (AFs), which we briefly discuss in Section 4, define how to determine such extensions. Though [6] discuss AFs, in this paper, we take the primary focus of logic-based approach to be on instantiating arguments and their relationships. We review key components of the logic-based approach.

In a logic-based approach, statements are expressed as atoms (lower case roman letters), while formulae (greek letters) are constructed using the logical connectives of conjunction, disjunction, negation, and implication. The classical consequence relation is denoted by $\vdash$. Given a knowledge base Δ comprised of formulae and a formula α , $\Delta \vdash \alpha$ denotes that Δ entails α . Δ can be inconsistent and comprised of a range of declarative statements. We assume a set of formulae Δ from which arguments are constructed. Where $\perp$ denotes inconsistency, $\Delta \vdash \perp$ denotes that Δ is inconsistent. An argument is an ordered pair $\langle \phi, \alpha \rangle$, where $\phi \subseteq \Delta$, ϕ is a minimal set of formulae such that $\phi \vdash \alpha$, and $\phi \not\vdash \perp$. ϕ is said to support the claim α . For example, where p and q are atoms, and where the KB is comprised of p and $p \rightarrow q$, then $\langle \{p, p \rightarrow q\}, q \rangle$ is an argument, where $p, p \rightarrow q$ is the support for the claim q .

The knowledge base Δ may be inconsistent, which here arises where Δ contains contradictory propositions (and not necessarily just constraints); this bears on issues discussed in Section 2.3. With contradictory propositions, we can construct arguments in relations, where the propositional claim of an argument is contradictory to the propositional claim of another argument or is contradictory to some proposition in the support of another argument. These are *attack* relations between arguments $\langle \Psi, \beta \rangle$ and $\langle \Phi, \alpha \rangle$ such as *undercutter* and *rebuttal*; attacking arguments are referred to as *counter-arguments*. $\langle \Psi, \beta \rangle$ is an undercutter for $\langle \Phi, \alpha \rangle$ where β is $\neg(\phi_1 \wedge \dots \wedge \phi_n)$ and $\{\phi_1 \dots \phi_n\} \subseteq \Phi$; in essence, the claim of one argument is the negation of a set of formulae in the support of another argument.¹⁰ $\langle \Psi, \beta \rangle$ is a rebuttal for $\langle \Phi, \alpha \rangle$ if and only if $\beta \leftrightarrow \neg\alpha$ is a tautology; the claims of the arguments are inconsistent. For example, supposing the following LRKB (from [6]): $p, p \rightarrow \neg q, r, r \rightarrow \neg p, \neg p \rightarrow q$. From this LRKB, we can construct an argument to support the claim $\neg q$: $\langle \{p, p \rightarrow \neg q\}, \neg q \rangle$. With respect to this argument, we have an undercutter $\langle \{r, r \rightarrow \neg p\}, \neg p \rangle$ and a rebuttal $\langle \{r, r \rightarrow \neg p, \neg p \rightarrow q\}, \neg p \rangle$.

Given a large and complex LRKB, arguments will have structural relationships such as subsumption of supports, where one support is a subset of another support, and implication between claims, where one claim entails another. Moreover, there may be more than one argument which undercuts or rebuts another argument. [5] and [6] define and discuss a range of these relationships among arguments; however, additional definitions are not directly relevant to our key

¹⁰ There is an additional notion of *canonical undercut*, where the atoms are ordered; it is useful for efficiency. For the presentation here, we presume it.

points in this paper. For our purposes, given an LRKB, we can generate not only the arguments, but also the counterarguments, the counterarguments to these arguments (counter-counterarguments), and so on recursively; such a structure is an *argument tree*, a graph where arguments are nodes and attack relations are (undifferentiated) arcs. From a given LRKB, [5] generate all possible arguments and counterarguments. To illustrate this with an example, we discuss *JArgue*, which implements the logic-based approach.

3.1 JArgue

JArgue is an implementation in Java of logic-based argumentation [5]; given an LRKB, *JArgue* generates arguments, attacks between arguments, and an argument tree.¹¹ One provides *JArgue* with an LRKB as a text file of propositional clauses (*conjunctive normal form*); implicational rules such as $p \rightarrow q$ are rewritten as $!p|q$, where $!$ is negation and $|$ is disjunction. Where the antecedent of a rule is a conjunction of literals as in $[r \wedge s] \rightarrow q$, we rewrite the negation of the antecedent as a disjunct of negated literals – $!r|!s|q$. Starting *JArgue* and inputting an LRKB, one can request the generation of arguments where a particular literal is the claim; given the arguments, one can then select among the arguments and request the generation of the argument tree. We discuss *JArgue* and illustrate it with our example.

Our example in Section 2.1 has what we are treating as atoms $p1$ – $p19$ and rules $r1$ – $r8$. Several preliminary comments are needed before presenting the input LRKB. First, since incompatible literals is a significant issue, we did not presume to represent statements which were incompatible as logical contradictions, e.g. $p1$ and $p2$, $p11$ and $p18$, and $p15$ and $p17$, even though ACE translates these pairs as contradictory. In other words, we presumed the problematic examples of contradiction rather than the unproblematic.

- $p1$: Every household should pay some tax for the households garbage.
- $p2$: No household should pay some tax for the households garbage.
- $p11$: Tom is not an objective expert about recycling.
- $p18$: Tom is an objective expert about recycling.
- $p15$: Every supermarket should pay a tax for the garbage that the supermarket creates.
- $p17$: No supermarket should pay a tax for the garbage that the supermarket creates.

However, to determine attack relations, the logic-based approach requires the explicit expression of contradiction as in classical logic. We therefore make the following substitutions: $p2 \equiv \neg p1$, $p11 \equiv \neg p18$, and $p17 \equiv \neg p15$. This is an issue relevant to the illustration, not to the functionality of either the logic-based approach or *JArgue* since expressions of contradiction can be added to the knowledge base as it is built. However, the choice here is to represent such incompatibilities as logical contradiction rather than as constraints of the form $\neg[p1 \wedge p2]$ (or in conjunctive normal form for *JArgue* $!p1|!p2$). We explore this choice in future research.

Given these equivalences, the related rules are adjusted by substituting in equivalent expressions:

- $r1$: $[p4 \wedge p3] \rightarrow p1$
- $r2$: $[p9 \wedge p18 \wedge p19] \rightarrow p4$

¹¹ Requests for *JArgue* should be addressed to Tony Hunter a.hunter@cs.ucl.ac.uk.

- $r3$: $p10 \rightarrow p3$
- $r4$: $[p12 \wedge p13] \rightarrow \neg p18$
- $r5$: $[p5 \wedge p15] \rightarrow \neg p1$
- $r6$: $[p6 \wedge p7 \wedge p8] \rightarrow p5$
- $r7$: $p14 \rightarrow p15$
- $r8$: $p16 \rightarrow \neg p15$

Our LRKB for *JArgue* will be comprised of literals and (conjunctive normalised) rules. However, there is a somewhat tangential issue of whether we need include *all* literals since some serve only as *intermediaries* between rules (see the discussion of *intermediate concepts* in [36]); that is, if we have a language comprised of $p, q, r, p \rightarrow q, q \rightarrow r$, there is no reason to include in the KB q (or r) itself since it is enough to include all the rules and only those literals which are terminal leaves in a graph of *all the rules* as in Figure 1; in other words, the KB need only represent what is *asserted* along with the rules, not what is simply expressed in the language. In a dynamically constructed KB, what is asserted and what is inferred may change over time. We presume this simplification and leave further discussion to future research. Thus, we have rules $r1$ – $r8$ as revised above and only $p6, p7, p8, p9, p10, p12, p13, p14, p16, p18$, and $p19$. Other literals are inferred.¹²

With these points in mind, our input LRKB for *JArgue* is as follows:

- $p6, p7, p8, p9, p10, p12, p13, p14, p16, p18, p19, !p10|p3, !p9|!p18|!p19|p4, !p12|!p13|!p18, !p3|!p4|p1, !p5|!p15|!p1, !p6|!p7|!p8|p5, !p14|p15, !p16|!p15$

To run *JArgue*, we add this LRKB, then query the KB to generate an argument. Querying the KB can only be done for one literal at a time. For our example, if we query for $p1$, we get the following argument, which we label $a1$:¹³

- $a1: < \{p10, p9, p18, p19, !p10|p3, !p9|!p18|!p19|p4, !p3|!p4|p1\}, p1 >$

It happens to be the case that there is only one argument $a1$ for $p1$ generated from our KB, but with another KB, there could have been more.

JArgue then requests the user to select from an argument with respect to which it should generate an argument tree. Selecting $a1$, *JArgue* generates an argument tree with two counterarguments $a2$ and $a3$ along with one counterargument $a4$ to $a3$; for the sake of discussion, we discuss each node and their relationship rather than present the graph. The first counterargument $a2$ is an undercutter of $a1$ since the claim of $a2$, $!p18$, is the negation of a literal which appears in the support of $a1$:¹⁴

- $a2: < \{p12, p13, !p12|!p13|!p18\}, !p18 >$

The second counterargument $a3$ is a rebuttal of $a1$ since the claim of $a3$, $!p1$, is the negation of the claim of $a1$:

- $a3: < \{p6, p7, p8, p14, !p6|!p7|!p8|p5, !p14|p15, !p5|!p15|!p1\}, !p1 >$

¹² This may be a more substantive issue since what arguments and relations are generated and used for input to the abstract framework may depend on what is explicitly available in the knowledge base rather than the explicit and inferred statements. We discuss this further below, and it may need further examination.

¹³ Labelling the arguments is not a functionality of Logic-based argumentation or *JArgue*. We do so to provide elements for the AEKB.

¹⁴ Logic-based argumentation and *JArgue* represent the claim of a canonical undercutter with $*$; for the purposes of our illustration, we make it explicit.

Finally, we have a counterargument to a_3 , where one of the implied literals of a_3 , p_{15} , is contradicted by the claim of a_4 , $\neg p_{15}$:

$$a_4: < \{p_{16}, \neg p_{16} | p_{15}\}, p_{15} >$$

If we query the LRKB for $\neg p_1$ and keep the argument labeling the same, we have a_3 . The argument tree for a_3 gives us a_1 as a counterargument; this shows that where we have contradictory *claims*, the arguments attack one another. Otherwise, a_2 is a counterargument of a_1 , and a_4 is a counterargument for a_3 as before.

Querying just the argument trees for p_1 and $\neg p_1$, we can create an AKB, which is a knowledge base of generated arguments in attack relations (indicated with *att*), which is given in the argument tree. Such an AKB is suitable for input to an abstract argumentation framework to generate an AEKB, which allows us to reason abstractly over arguments. This is AKB1:

Arguments: a_1, a_2, a_3, a_4
 Attack Relations: $\text{att}(a_1, a_3), \text{att}(a_3, a_1), \text{att}(a_2, a_1), \text{att}(a_4, a_3)$

3.2 Considerations

Where we choose to argue for and against literals at the root of an argument tree such as about p_1 and $\neg p_1$, the arguments and relations among arguments appear relatively straightforward. It is important to note that we have used JArgue to generate argument trees with respect to a query on a proposition. However, we are using it as well to creating AKBs for input to abstract argumentation frameworks, which is the reason to investigate the argument trees for a proposition *and* its negation; however, using JArgue to generate arguments and relations to support abstract argumentation frameworks is somewhat beyond the brief, as we see in a moment.

A relevant general question arises – how many arguments and attack relations among them can be generated from the underlying LRKB? Given that we have classical logic, *contrapositive* reasoning applies; that is, given $\neg q$ and $p \rightarrow q$, we can infer $\neg p$. Thus, if we query p_3 , which is implied by p_{10} , which is itself asserted, we can generate an argument and an argument tree.

$$a': < \{p_{10}, \neg p_{10} | p_3\}, p_3 >$$

In the argument tree, we have a rebuttal argument which is generated by contraposition – where $\neg p_1$ and p_4 hold, $\neg p_3$ must hold:

$$a'': < \{p_6, p_7, p_8, p_9, p_{14}, p_{18}, p_{19}, \neg p_6 | \neg p_7 | \neg p_8 | p_5, \neg p_{14} | p_{15} | p_9 | \neg p_{18} | \neg p_{19} | p_4, \neg p_5 | p_{15} | \neg p_1, \neg p_3 | p_4 | p_1\}, \neg p_3 >$$

This argument itself has two counterarguments. And in particular, a' and a'' rebut one another. More generally, every literal in the language (appearing in a rule or as an assertion) has an argument and a counterargument (each of which may also have counterarguments). Given this, should all such arguments in their relations be in the AEKB for a given AKB?

For our purposes, we only include those arguments and relations relative to a debate under investigation, here p_1 and $\neg p_1$; these are the arguments and relations as generated by JArgue. Other arguments and counterarguments such as for p_3 are not clearly relevant. Indeed, perhaps all the other arguments and relations arise via contraposition are problematic; indeed, it is questionable whether contraposition is a *natural* reasoning rule (see [15] and [31]).

The hedge, however, bears on how we use AKB as input to the abstract argumentation framework discussed in Section 4, where we find the AKB with four arguments and four attacks is not sufficient

to get a correct result. This is so because while a_2 attacks a_1 , and a_4 attacks a_3 , nothing attacks a_2 or a_4 ; thus, these arguments attack and defeat a_1 and a_3 . The extensions leave just a_2 and a_4 , which is contrary to our desired goal, which is that a_1 and a_4 could hold together, while a_3 and a_2 could hold together. We need, then, to generate an attack on a_2 where we query for the claim of a_1 , and to generate another attack on a_4 when we query for the claim of a_3 .

This would seem to require that where we query for a proposition and its negation, we want to generate at least one additional argument to counterattack attackers, if we can. In other words, we want to generate attacks (either rebuttals or undercutters) to a_2 and a_4 . For example, we could do so by finding a rebuttal which uses (in a sense to be explained) the claim of the queried argument. In other words, we want to find one rebuttal a_5 where the claim of a_5 is the negation of the claim of a_2 and where the claim of a_1 , namely p_1 , is used is found among the support for the claim of a_5 ; similarly, we want to find another rebuttal a_6 where the claim of a_6 is the negation of the claim of a_4 and where the claim of a_3 , namely $\neg p_2$, is used. Such arguments may be generated from the underlying LRKB or they may be generated by contraposition, which then provides some intuitive motivation to support such a rule in the context of argumentation. In other words, we want to find instantiated arguments to realise AKB2:

Arguments: $a_1, a_2, a_3, a_4, a_5, a_6$
 Attack Relations: $\text{att}(a_1, a_3), \text{att}(a_3, a_1), \text{att}(a_2, a_1), \text{att}(a_4, a_3), \text{att}(a_5, a_2), \text{att}(a_2, a_5), \text{att}(a_6, a_4), \text{att}(a_4, a_6)$

The result is a graph as in:

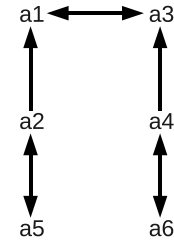


Figure 2. Arguments and Relationships

Note that arguments a_1 - a_4 are clearly related to our earlier representation of the relationships among statements in Figure 1, but that a_5 and a_6 are not.

However, to find a_5 and a_6 , we have some additional issues to consider, in particular the argument trees of each of these arguments. We consider the issues by example. a_2 has as claim $\neg p_{18}$, thus the rebuttal a_5 must have p_{18} as claim. In fact, we have an argument by assertion for this claim: $<\{p_{18}\}, p_{18}>$. However, this has a counterargument (derived by contraposition) as well as a counter-counterarguments:

$$< \{p_6, p_7, p_8, p_{14}, p_{10}, \neg p_6 | \neg p_7 | \neg p_8 | p_5, \neg p_{14} | p_{15}, \neg p_{10} | p_3, \neg p_5 | p_{15} | \neg p_1, p_9, p_{19}, \neg p_3 | p_4 | p_1, \neg p_9 | \neg p_{18} | \neg p_{19} | p_4\}, \neg p_{18} >$$

Must we include, then, the whole argument tree for a_5 ? However, we observe that p_1 and $\neg p_1$ are used in this argument. If p_1 is in fact asserted (given that is the argument we query), then this counterargument to a_5 (and the counter-counterargument) can no longer be

generated, leaving $a5 <\{p18\}, p18>$. By the same token, we need an argument to defend $p15$ from attack by argument $a4$; this is $a6 <\{p14, !14|15\}, p15>$

While these are remarks about what are required to have a AKB suitable for reasoning with AFS, we leave to future research how this is theoretically developed and implemented. In any case, adopting these additional arguments gives more satisfactory results in terms of AFS. The important point is that we must carefully consider what results we might want with respect to AFS in order to create the appropriate arguments and relations, then build into our argument generator the mechanisms to generate the arguments and relations, otherwise the desired goal from the AF will not be achieved.

4 FROM INSTANTIATED ARGUMENTS TO AN AF

Having generated an AKB of arguments and their relations, we can input it to an abstract argumentation framework AF to calculate extensions. In this section, we review basic components of AFS, outline results with respect to an implementation, and then discuss the results.

4.1 AFS

For our purposes, we consider the AF of [14], where there is one set of undifferentiated objects, *arguments*, which are nodes in a graph as well as one undifferentiated relationship between the nodes, the *attack* relation, which can be represented as a graph in which attacks are arcs between nodes representing the arguments (for related work, see [8] [3], [10], among others).

Definition An argumentation framework AF is a pair $\langle \mathcal{X}, \mathcal{R} \rangle$, where $\mathcal{X}$ is a set of *objects*, $\{a_1, a_2, \dots, a_n\}$ and $\mathcal{R}$ is an *attack* relation between objects. For $\langle a_i, a_j \rangle \in \mathcal{R}$ we say the object a_1 attacks object a_2 . We assume that no object attacks itself.

Some of the relevant auxiliary definitions are as follows, where S is a subset of $\mathcal{X}$. There are a range of *semantics*, which are definitions of extensions (sets) of arguments; we give only admissible and preferred semantics:

Definition We say that $p \in \mathcal{X}$ is *acceptable with respect to S* if for every $q \in \mathcal{X}$ that attacks p there is some $r \in S$ that attacks q . A subset, S , is *conflict-free* if no argument in S is attacked by any other argument in S . A conflict-free set S is *admissible* if every $p \in S$ is acceptable to S . A *preferred extension* is a maximal (w.r.t. $\subseteq$) admissible set.

In the next section, we use the ASPARTIX system to generate argument extensions.

4.2 CALCULATING EXTENSIONS WITH ASPARTIX

The ASPARTIX program computes extensions of argumentation frameworks [17].¹⁵ It is, then, very useful for our purposes, where we have created arguments in relations and then wish to compute

extensions. ASPARTIX is implemented in the answer set programming language *Disjunctive Datalog System* (DLV).¹⁶ Using ASPARTIX, one can compute any of the standard extensions for the classical AF (following Dung) including admissible and preferred extensions. Extensions for preference-based, value-based, and bipolar argumentation frameworks can also be generated.

To run ASPARTIX, we input to the program a text file for the AKB, containing the arguments and relations, as well as the desired sort of extension. We used our previously defined AKB, which has the arguments and attack relations:

```
arg(a1).          att(a3,a1).
arg(a2).          att(a2,a1).
arg(a3).          att(a4,a3).
arg(a4).          att(a5,a2).
arg(a5).          att(a2,a5).
arg(a6).          att(a6,a4).
att(a1,a3).       att(a4,a6).
```

ASPARTIX generates the following preferred extensions, where *in* indicates that the argument is in the extension:

```
{in(a1), in(a5), in(a6)}
{in(a2), in(a3), in(a6)}
{in(a3), in(a5), in(a6)}
{in(a1), in(a4), in(a5)}
{in(a2), in(a4)}
```

Such extensions which are generated by the AF we refer to as the *Argument Extensions Knowledge Base* (AEKB).

This seems to be an attractive result. As we are only interested in extensions with $p1$ and $\neg p1$, we will not look at $\{in(a2), in(a4)\}$. As a remark, preferred extensions seem to be the argumentation correlate to the notion *normative positions* of deontic logic [29], which are maximal sets of consistent deontic expressions. In light of this, we will call them *policy positions*.

5 Extracting Policy Positions

Our sets of arguments are an abstract representation of the contents of the arguments themselves, which are comprised of literals and rules. Thus, given the sets of arguments, we can extract the literals and rules that are consistent and which represent the correlated linguistic expressions of the policy positions. While we can reconstruct the content of arguments $a1$, $a2$, $a3$, and $a4$, the additional arguments for $a5$ and $a6$ are less clear. We will suppose that these are, as we suggested, generated as rebuttals.

Recall that our arguments are:

```
a1: < {p10, p9, p18, p19, !p10|p3, !p9|!p18|!p19|p4,
      !p3|!p4|p1}, p1 >
a2: < {p12, p13, !p12|!p13|!p18}, !p18 >
a3: < {p6, p7, p8, p14, !p6|!p7|!p8|p5, !p14|p15,
      !p5|!p15|!p1}, !p1 >
a4: < {p16, !p16|!p15}, !p15 >
a5: < {p18}, p18 >
a6: < {p14, !14|15}, p15 >
```

Let us informally define a policy position (PP) as supports and claims of all the arguments in the extension (we could do without the claim, but in a sense, the objective is making explicit what is

¹⁵ We thank the ASPARTIX team, and particularly Sarah Gaggl, for making the ASPARTIX system available to us. See <http://www.dbai.tuwien.ac.at/staff/gaggl/systempage/index.html>.

¹⁶ See <http://www.dbai.tuwien.ac.at/proj/dlv/>.

otherwise implicit and inferred). Formally, we have two functions: F1 is a function on the extension such that it outputs the union of the result of the application of F2 on every argument of the extension; F2 is a function on arguments which outputs a set comprised of the union of the support with the claim. We suppose these functions. Furthermore, since we have sets, any redundancy is removed.

PP1: $\{p1, p4, p9, p10, p14, p15, p18, p19, !p10|p3, !p9|!p18|!p19|p4, !p3|!p4|p1, !p14|p15\}$
 PP2: $\{!p1, p5, p6, p7, p8, p12, p13, p14, p15, !p18, !p12|!p13|!p18, !p6|!p7|!p8|p5, !p14|p15, !p5|!p15|!p1\}$
 PP3: $\{!p1, p5, p6, p7, p8, p14, p15, p18, !p6|!p7|!p8|p5, !p14|p15, !p5|!p15|!p1\}$
 PP4: $\{p1, p3, p4, p9, p10, !p15, p16, p18, p19, !p10|p3, !p9|!p18|!p19|p4, !p3|!p4|p1, !p16|!p15\}$

The propositions and rules here correlate directly with the propositions and rules introduced in Section 2.1, so we can substitute them here. We will put in the sentences and indicate the rules for PP2:

PP2 Sentences and Rules

!p1: No household should pay some tax for the household's garbage.
 p5: If a household pays a tax for the households garbage then the tax is unfair to the household.
 p6: Every household should pay an equal portion of the sum of the tax for the households garbage.
 p7: No household which receives a benefit which is paid by a council recycles the households garbage.
 p8: Every household which does not receive a benefit which is paid by a council supports a household which receives a benefit which is paid by a council.
 p12: Tom owns a company that recycles some garbage.
 p13: Every person who owns a company that recycles some garbage earns some money from the garbage which is recycled.
 p14: Every supermarket creates some garbage.
 p15: Every supermarket should pay a tax for the garbage that the supermarket creates.
 !p18: Tom is not an objective expert about recycling.
 r4: $[p12 \wedge p13] \rightarrow !p18$
 r5: $[p5 \wedge p15] \rightarrow !p1$
 r6: $[p6 \wedge p7 \wedge p8] \rightarrow p5$
 r7: $p14 \rightarrow p15$

This seems an accurate representation of this policy position from the onset. While PP2 and PP4 seem acceptable representations of policies.

5.1 CONSIDERATIONS

The policy positions of PP1 and PP3 seem to have some additional superfluous information. For instance, PP1 contains the propositions p14 and p15, which are propositions in the support of !p1, which itself is the rebuttal of the main point of PP1. Since other portions of the support for !p1 do not hold, !p1 does not hold. Yet, it would seem, in an intuitive argument, odd to have as part of one's position statements in support of the opposite point, even if these statements are not inconsistent with one's position. We see, then, that AFs *overgenerate* policy positions. In addition, the output sentences of the policy positions are without discourse markers, so the relations among them are unclear. We leave for future research just how to refine the results

to generate all and only the policy positions that correlate with an intuitive notion as well as how to use natural language generation technology to provide well-formed, easy-to-understand output paragraphs of the policy positions.

6 DISCUSSION

In this paper, we have shown how statements expressed in natural language can be represented as a formal knowledge base of literals and rules (LRKB), that the LRKB can be fed into a framework for generating arguments and relations to create a knowledge base of arguments (AKB), and then in turn, the AKB can be provided to an AF to generate a set of extensions which is the knowledge base of extensions of arguments (AEKB). The extensions we related to policy positions, which are consistent sets of statements and rules; we represented the policy positions in terms of the underlying source statements. At each point of the transformation, we discussed relevant issues that arose.

The main novelty of our approach is that formal, implemented approaches to the syntax and semantics of natural language have been used to develop the LRKB, that we have used implemented systems for generating AKB and AEKB, and that we have identified some important novel issues and problems for future research in the development of the argument pipeline. In these respects, our approach is distinct from similar work such as [5], [6], [26], and [38], where disputes expressed in natural language are translated manually into AKB and AEKB. While manual analysis can work, it does not address the problem of the knowledge acquisition bottleneck, nor does it make explicit, formal, or implemented the knowledge engineer's knowledge in making the analysis.

In addition to observations made in the *considerations* sections, the work reported here does not discuss a range of other interesting and relevant issues which we seek to address in the future. For instance, we have left aside any issues relating to knowledge base or context dynamics. Rather, as with classical logics, we have presumed static knowledge bases and fixed contexts. This is entirely a strategic move in order to pin down interesting issues about the static domain per se which are also relevant to dynamic system and context. The approach also emphasises dialectical argumentation such as found in legal disputes, where there is one specific statement being argued for and against and where we generate policy positions. There are, as we know, a range of other dialogical modes and goals, e.g. deliberation, information gathering, and consensus building (see, for example, [33]). Nonetheless, it all of these dialogues, it is fundamental to identify and work with *difference* and *inconsistency*, whether to resolve it, eliminate it, or to adapt to it. Different dialogical modes and goals may be thought of as different strategies to work with difference and inconsistency. Finally, the approach we have taken here makes use of *rhetorical relationships*, e.g. premise, claim, and contradiction, between statements that suit a literals and rules knowledge base that is compatible with Propositional or First-order Logic. However, there are other rhetorical relationships that are relevant to dialogue and argumentation which ought to be incorporated into future analysis and implementations (see [24], [22], and [30]).

ACKNOWLEDGEMENTS

The first two authors thank the European Commission for supporting this research in the IMPACT Project (Integrated Method for Policy making using Argument modeling and Computer assisted Text analysis) FP7 Grant Agreement No. 247228.

REFERENCES

- [1] Leila Amgoud, Martin Caminada, Claudette Cayrol, Marie-Christine Lagasque, and Henry Prakken, 'Towards a consensual formal model: inference part', Technical report, ASPIC project, (2004). Deliverable D2.2: Draft Formal Semantics for Inference and Decision-Making.
- [2] Jon Barwise and Robin Cooper, 'Generalized quantifiers and natural language', *Linguistics and Philosophy*, **4**, 159–219, (1981).
- [3] Trevor J. M. Bench-Capon, 'Persuasion in practical argument using value-based argumentation frameworks', *Journal of Logic and Computation*, **13**(3), 429–448, (2003).
- [4] Philippe Besnard and Anthony Hunter, 'Practical first-order argumentation', in AAAI, eds., Manuela M. Veloso and Subbarao Kambhampati, pp. 590–595. AAAI Press / The MIT Press, (2005).
- [5] Philippe Besnard and Anthony Hunter, *Elements of Argumentation*, MIT Press, 2008.
- [6] Phillip Besnard and Anthony Hunter, 'Argumentation based on classical logic', in *Argumentation in Artificial Intelligence*, eds., Iyad Rahwan and Guillermo Simari, 133–152, Springer, (2009).
- [7] Elizabeth Black and Anthony Hunter, 'Using enthymemes in an inquiry dialogue system', in *Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'08)*, pp. 437–444. ACM Press, (2008).
- [8] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni, 'An abstract, argumentation-theoretic approach to default reasoning', *Artificial Intelligence*, **93**, 63–101, (1997).
- [9] Emma Bryne and Anthony Hunter, 'Man bites dog: Looking for interesting inconsistencies in structured news reports', *Data and Knowledge Engineering*, **48**(3), 265–295, (2004).
- [10] Martin Caminada and Leila Amgoud, 'On the evaluation of argumentation formalisms', *Artificial Intelligence*, **171**(5–6), 286–310, (2007).
- [11] Alan Cruse, *Lexical Semantics*, Cambridge University Press, 1986.
- [12] Marie-Catherine de Marneffe, Anna N. Rafferty, and Christopher D. Manning, 'Finding contradictions in text', in *Proceedings of ACL-08: HLT*, pp. 1039–1047, Columbus, Ohio, (June 2008). Association for Computational Linguistics.
- [13] Valeria de Paiva, Daniel G. Bobrow, Cleo Condoravdi, Dick Crouch, Tracy H. King, Lauri Karttunen, Rowan Nairn, and Annie Zaenen, 'Textual inference logic: Take two', in *C&O:RR*, (2007).
- [14] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial Intelligence*, **77**(2), 321–358, (1995).
- [15] Dorothy Edgington, 'Conditionals', in *Stanford Encyclopedia of Philosophy*, ed., Edward Zalta, Stanford University, (2006). <http://plato.stanford.edu/entries/conditionals/>.
- [16] Vasiliki Efstathiou and Anthony Hunter, 'An algorithm for generating arguments in classical predicate logic', in *ECSQARU*, pp. 119–130, (2009).
- [17] Uwe Egly, Sarah Alice Gaggl, and Stefan Woltran, 'Answer-set programming encodings for argumentation frameworks', *Argument and Computation*, to appear, to appear, (2008).
- [18] Diana E. Forsythe and Bruce G. Buchanan, 'Knowledge acquisition for expert systems: some pitfalls and suggestions', in *Readings in knowledge acquisition and learning: automating the construction and improvement of expert systems*, 117–124, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, (1993).
- [19] Alejandro Javier García and Guillermo Ricardo Simari, 'Defeasible logic programming: An argumentative approach', *Theory and Practice of Logic Programming*, **4**(1), 95–137, (2004).
- [20] Thomas Gordon, Henry Prakken, and Douglas Walton, 'The carneades model of argument and burden of proof', *Artificial Intelligence*, **171**, 875–896, (2007).
- [21] Guido Governatori, Michael J. Maher, Grigoris Antoniu, and David Billington, 'Argumentation semantics for defeasible logic', *Journal of Logic and Computation*, **14**(5), 675–702, (2004).
- [22] Floriana Grasso, 'Towards a framework for rhetorical argumentation', in *Proceedings of the 6th Workshop on the Semantics and Pragmatics of Dialogue (EDIALOG-2002)*, eds., Johan Bos, M. E. Foster, and C. Matheson, pp. 53–60, Edinburgh, UK, (2002).
- [23] Ann Macintosh, 'Moving toward "intelligent" policy development', *IEEE Intelligent Systems*, **24**(5), 79–82, (2009).
- [24] William C. Mann and Sandra A. Thompson, 'Rhetorical structure theory: A theory of text organization', Technical Report ISI/RS-87-190, University of Southern California, Information Sciences Institute (ISI), (June 1987).
- [25] Rudy Prabowo and Michael Thelwall, 'Sentiment analysis: A combined approach', *Journal of Informetrics*, **3**(1), 143–157, (2009).
- [26] Henry Prakken, 'Formalising ordinary legal disputes: a case study', *Artificial Intelligence and Law*, **16**, 333–359, (2008).
- [27] Henry Prakken, 'An abstract framework for argumentation with structure arguments', *Argument and Computation*, **?**, ?–?, (2010). To appear.
- [28] Henry Prakken and Giovanni Sartor, 'Argument-based extended logic programming with defeasible priorities', *Journal of Applied Non-Classical Logics*, **7**(1), (1997).
- [29] Marek Sergot, 'Normative positions', in *Norms, Logics, and Information Systems*, eds., McNamara and Prakken, 289–308, IOS Press, (1998).
- [30] Caroline Sporleder and Alex Lascarides, 'Using automatically labelled examples to classify rhetorical relations: An assessment', *Natural Language Engineering*, **14**(3), 369–416, (2006).
- [31] Keith Stenning and Michiel van Lambalgen, 'A little logic goes a long way: basing experiment on semantic theory in the cognitive science of conditional reasoning', *Cognitive Science*, **28**(4), 481–529, (2004).
- [32] Ellen M. Voorhees, 'Contradictions and justifications: Extensions to the textual entailment task', in *Proceedings of ACL-08: HLT*, pp. 63–71, Columbus, Ohio, (June 2008). Association for Computational Linguistics.
- [33] Doug Walton, Katie Atkinson, Trevor Bench-Capon, Adam Wyner, and Doug Cartwright, 'Argumentation in the framework of the deliberation dialogue', in *Arguing Global Governance: Agency, Lifeworlds and Shared Reasoning*, eds., Corneliu Bjola and Markus Kornprobst, Routledge, (2010). Forthcoming.
- [34] Douglas Walton, *Argumentation Schemes for Presumptive Reasoning*, Erlbaum, Mahwah, N.J., 1996.
- [35] Douglas Walton, 'The three bases for the enthymeme: A dialogical theory', *Journal of Applied Logic*, **6**(3), 361–379, (2008).
- [36] Adam Wyner, 'An ontology in OWL for legal case-based reasoning', *Artificial Intelligence and Law*, **16**(4), 361–387, (December 2008).
- [37] Adam Wyner, Krasimir Angelov, Guntis Barzdins, Danica Damjanovic, Brian Davis, Norbert Fuch, Stefan Hoefler, Ken Jones, Kaarel Kaljurand, Tobias Kuhn, Martin Luts, Jonathan Pool, Mike Rosner, Rolf Schwitter, and John Sowa, 'On controlled natural languages: Properties and prospects', in *Workshop on Controlled Natural Languages, CNL 2009*, ed., Norbert E. Fuchs, number 5972 in LNCS/LNAI. Springer, (2010). To appear.
- [38] Adam Wyner, Trevor Bench-Capon, and Katie Atkinson, 'Arguments, values and baseballs: Representation of popov v. hayashi', in *Legal Knowledge and Information Systems. JURIX 2007*, eds., Arno R. Lodder and Laurens Mommers, pp. 151–160, Amsterdam, (2007). IOS Press.
- [39] Adam Wyner, Raquel Mochales-Palau, Marie-Francine Moens, and David Milward, 'Approaches to text mining arguments from legal cases', in *Semantic Processing of Legal Texts*, eds., Simonetta Montemagni, Daniela Tiscornia, and Enrico Francesconi, Lecture Notes in Computer Science, Springer, (2009). To appear.
- [40] Adam Wyner, Tom van Engers, and Kiavash Bahreini, 'From policy-making statements to first-order logic', in *Proceedings of International Conference on Electronic Government and the Information Systems Perspective (EGOVIS 2010)*, Lecture Notes in Computer Science. Springer Verlag, (2010). To appear.

The <TextCoop> Platform: Analyzing arguments in procedural texts

Thomas De Filippo, Laureline Marsal, Patrick Saint-Dizier¹

1 Introduction and Aims

<TextCoop> is an environment for text semantics analysis. In a first experimental stage, it is dedicated to procedure processing for testing and evaluation. Processing procedures has obviously major applications. Besides instructions, procedures also contain a number of forms of explanation, in particular warnings and advice. Therefore, our platform can also be used as a tool for argument mining.

The goal of this document is to show the current possibilities of the software, version V2, in Java, after a proof of concept realized two years ago in Perl, based on scripts (Delpech et al. 2008).

Procedural texts are organized sets of instructions, they may also be sets of advice, as in social behavior texts. In our perspective, procedural texts range from apparently simple cooking recipes to large maintenance manuals. They also include documents as diverse as teaching texts, medical notices, social behavior recommendations, directions for use, assembly notices, do-it-yourself notices, itinerary guides, advice texts, savoir-faire guides etc. Even if procedural texts adhere more or less to a number of structural criteria, which may depend on the author's writing abilities and on traditions associated with a given domain, we observed a very large variety of realisations, which makes parsing such texts quite challenging.

Procedural texts explain how to realize a certain goal by means of actions which may be temporally organized. Procedural texts can indeed be a simple, ordered list of instructions to reach a goal, but they can also be less linear, outlining different ways to realize something, with warnings, advice, conditions, hypothesis, preferences. They also often contain a number of recommendations and comments of various sorts. The organization of a procedural text is in general made visible by means of linguistic and typographic marks.

Research on procedural texts was initiated by works in psychology, cognitive ergonomics, and didactics, (Mortara et al. 1988) (Adam 1987), (Kosseim 2000) to cite just a few. Several facets, such as temporal and argumentative structures have then been subject to general purpose investigations in linguistics, but they need to be customized to this type of text. There is however very little work done in Computational Linguistics circles. The present work is based on a preliminary experiment we carried out (Delpech et al., 07), (Aouladomar, 05) where a preliminary structure was proposed, from corpus analysis.

Let us now give an illustrative example (translated from French), extracted from the 'Do-It-Yourself Home' domain, to show how texts are processed: *In the bedroom, it is necessary to clean curtains. These are cleaned first with a vacuum-cleaner to remove dust, then, if they are in cotton, they can be washed in the washing machine at 60 degrees; if they are white, it is even recommended to add some bleach*

so that they look whiter. With some starch, they can be easily ironed.

In this text, the sequence: *In the bedroom, it is necessary to clean curtains* is analyzed as a justification of the actions to undertake. The next portion: *These are cleaned first with a vacuum-cleaner to remove dust, then, if they are in cotton, they can be washed in the washing machine at 60 degrees.* is the instruction kernel, where the last instruction is associated with a condition. Finally, *If they are white, it is even recommended to add some bleach so that they look whiter. With some starch, they can be easily ironed.* are two subordinated clauses, analyzed as being in the rhetorical relation advice to the kernel instructions.

Another example in parenthetical format (French gloss) is the following:

[*instruction* The first step consists in opening the computer box, then to place it on a large, clean surface,

[*advice* where you have ample space to work comfortably,] and then to withdraw all the caches at the PC front.]

A more complex case is:

[[*Goal* To clean leathers,]

[*instruction* choose specialized products dedicated to furniture,

[*advice* [*instruction* and prefer them colorless],

[*support* they will play a protection role, add beauty, and repair some small damages.]]]]

In the framework of argumentation, <TextCoop> can be used as a simple but efficient and flexible argument mining system. This is realized via the description of text patterns that describe the linguistic structure of arguments.

2 The software and the demo

<TextCoop> analyses procedures (dedicated to installation, maintenance, production, etc. over various technical and large public domains) in French and in English. Texts come either from the Web or from textual databases. <TextCoop> identifies structures and tags them in XML, in particular:

- Titles (goals of the text) and their structure, it also indexes titles (similarly to a search engine),
- Instructions and instructional compounds, temporal structure and conditions,
- Prerequisites (materials, tools, prior knowledge, etc.),
- Explanation structures: advice, warnings, illustrations, hints, evaluations, etc.

<TextCoop> is based on linguistic analysis, knowledge representation and language processing. The environment is modular and flexible so that it can adapt to users' needs and specificities via tailored interfaces. The linguistic interface allows for the specification of grammar rules or patterns (supporting gaps representing finite sequences

¹ IRIT-CNRS Toulouse France, email: sstdizier@irit.fr

of symbols), and lexical and ontological data. It also includes morphological analysis, and utilities such as web access, web page cleaning, and querying based on XML tags on the output structures. It is being made compliant with current data representation norms, via its planned embedding into the UIMA framework.

The rules and patterns include terminal (triggers as well as grammatical words like connectors) and non-terminal symbols (referring in general to 'local grammars', e.g. temporal expressions, instrument expressions, etc.). They also include gaps, standing for finite strings of words of no present interest, XML tags (treated as terminal elements), functions (for handling controls and for computing attribute values for tags), and insertion point specifications, designed to insert annotations at precise places. Symbols may also be associated with attribute value structures to encode morphology, syntax as well as semantic elements. An extraction rule for an advice is, for example:

begin-mark, Pronoun, to be/modal, adv,
verb/advice-exp, gap, end-mark.

as in: *It's better to install a grounding electrode.*
We advise you to use a waterproof box.

Similarly, for warnings, we have:

begin-mark, it, is, adverb,
important/necessary/compulsory/essential/
vital/fundamental, to, verb, gap, end-mark.

as in: *It is very important to be aware of electrical grounding safety.*

It is important to wash your hands before entering the room.

<TextCoop> recently got a much more stable implementation using Java, and in particular we developed an engine based on JFLEX and JCUP. This engine is based on an LALR(1) automaton that we have adapted so that it can handle non-determinism and gaps. Interfaces (user and administrator) have been defined so that it is quite easy, via some training, to describe patterns and lexical data for various semantic recognition treatments. At the functional level, <TextCoop> V.2 allows to (via customization):

- Index large volumes of written procedures and query your database,
- Enrich procedures, and make some controls on their contents (e.g. via business requirements),
- Add useful information : list of tools, materials, etc.,
- Develop advice and precautions to take for a given task, via document synthesis,
- Contribute to various types of applications, in particular risk analysis and prevention as detected from instructions.

A typical XML output of <TextCoop> is given below in Fig. 1. <TextCoop> can be viewed as the kernel of the system, to which dedicated add-ons or plug-ins can be added to realize specific, application-oriented tasks. These confer a stronger value to this system.

In terms of performances, we have at the moment the following results. On a standard PC, windows XP, an average of 300Mo of input texts is processed per hour, with a lexicon and ontology of a few thousand terms and about 25 tagging patterns or rules. We give below the performances for the major structures found in procedures.

As reported in (Fontan et al. 2008), we carried out an indicative evaluation on a corpus of 66 texts over various domains, containing 262 arguments. Those texts were manually annotated by a trained linguist, and the results were then compared with the system output. We get the following results for warnings:

conclusion recognition	support recognition	(3)	(4)
88%	91%	95%	95%

(3) conclusions well delimited (4) supports well delimited, with respect to warnings correctly identified.

We also carried out an indicative evaluation on the same corpus of 66 texts containing 240 manually identified advice. We get the following results for advice:

conclusion recognition	support recognition	(3)	(4)	(5)
79%	84%	92%	91%	91%

The software demonstration will include:

- Introduction to the software and its aims and functions,
- Automatic extraction of arguments (essentially warnings and advice) from procedural documents (technical and social) in French and English, attendees can propose texts, discussion of results (and errors),
- Introduction of new rules or patterns in the system to recognize new semantic structures,
- Demonstration of the adequacy of interfaces, of the introduction of domain knowledge (via an ontology), feedback from the audience,
- Automatic construction of a list of advice/warnings for a given topic.

An example of output is given below in Fig. 2.

REFERENCES

- [1] Adam, J.M., *Types de Textes ou genres de Discours? Comment Classer les Textes qui Disent De et Comment Faire*, Langages, 141, pp. 10-27, 2001.
- [2] Aouladomar, F., Saint-Dizier, P., *An Exploration of the Diversity of Natural Argumentation in Instructional Texts*, 5th International Workshop on Computational Models of Natural Argument, IJCAI, Edinburgh, 2005.
- [3] Delpech, E., Murguia, E., Saint-Dizier, P., *A Two-Level Strategy for Parsing Procedural Texts*, VSST07, Marrakech, October 2007.
- [4] Fontan, L., Saint-Dizier, P., *Analyzing the explanation structure of procedural texts: dealing with advice and Warnings*, STEP conference, Venice, August 2008.
- [5] Kosseim, L., Lapalme, G., *Choosing Rhetorical Structures to Plan Instructional Texts*, Computational Intelligence, Blackwell, Boston, 2000.
- [6] Mortara Garavelli, B., *Tipologia dei Testi*, in G. Hodus et al.: lexicon der romanistischen Linguistik, vol. IV, Tübingen, Niemeyer, 1988.
- [7] Rosner, D., Stede, M., *Customizing RST for the Automatic Production of Technical Manuals*, in R. Dale, E. Hovy, D. Rosner and O. Stock eds., *Aspects of Automated Natural Language Generation*, Lecture Notes in Artificial Intelligence, pp. 199-214, Springer-Verlag, 1992.

```

<instruction> <verb sem="action"> Energize </verb> the Aircraft Electrical Circuits <tool> from the APU </tool>
</instruction> <warning strength="normal"> Make sure that circuit breakers are closed </warning>. <instruction>
<temp-exp type="durative"> During the test </temp-exp>, <verb type="epistemic"> obey </verb> the instructions
shown <loc> on the terminal. </loc> </instruction> <instruction> <verb sem="action"> Start </verb> the mod-
ule test.</instruction> </explanation type="evaluation"> The test is in progress as <verb type="communication">
shown </verb> on the screen. </explanation> <instruction type="check"> Indication that test is OK comes into view.
</instruction> <instruction> <verb sem="action"> Close </verb> the OMS session. </instruction>

```

Figure 1. Extract of an annotated procedure

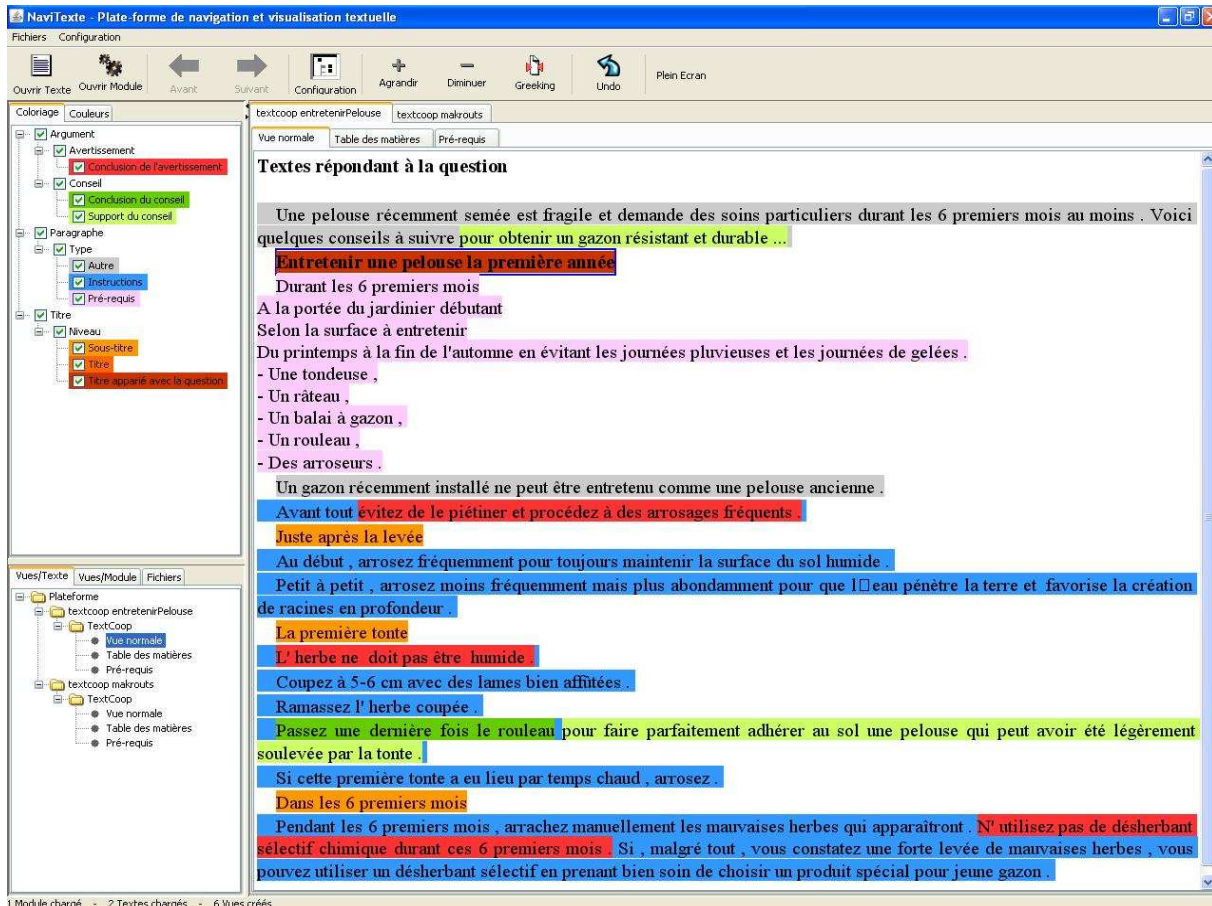


Fig. 2 Navitexte Output

Author Index

Atkinson, Katie	10
Bex, Floris	1
Botti, Vicent	10
Budzynska, Katarzyna	1, 5
de Filippo, Thomas	57
Delmonte, Rodolfo	29
Grasso, Floriana	10
Heras, Stella	10
Hunter, Anthony	47
Julian, Vicente	10
Krabbe, Erik C. W.	20
Marsal, Laureline	57
McBurney, Peter	10
Pallotta, Vincenzo	29
Saint Dizier, Patrick	34, 57
Uckelman, Sara L.	38
van Engers, Tom	47
van Laar, Jan Albert	20
Wyner, Adam	47

Argument & Computation

CALL FOR PAPERS

Co-Editors:

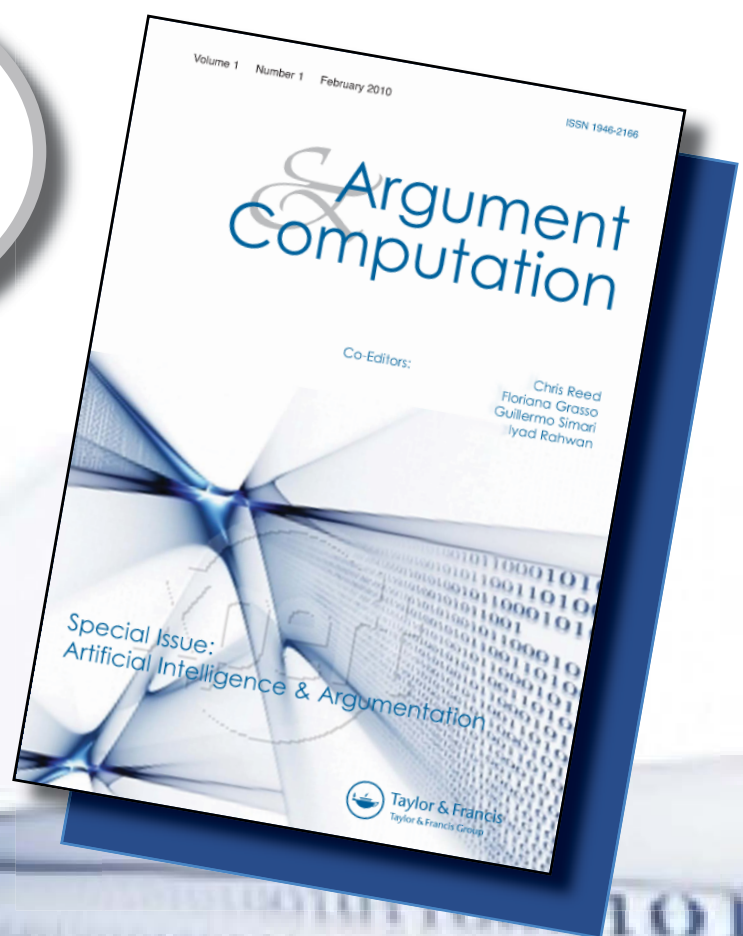
Chris Reed, University of Dundee, UK

Floriana Grasso, University of Liverpool, UK

Guillermo Simari, University Nacional del Sur, Argentina

Iyad Rahwan, British University of Dubai, UAE & University of Edinburgh, UK

**New for
2010!**



Taylor & Francis
Taylor & Francis Group

New for 2010!

Argument & Computation

Argument and Computation aims to promote the interaction and cross-fertilisation between the fields of argumentation theory and computer science. It will be of interest to researchers in the fields of artificial intelligence, multi-agent systems, computer science, logic, philosophy, argumentation theory, psychology, cognitive science, game theory and economics. Topics of interest include, but are not limited to:

- Argumentation and cognitive architectures
- Argumentation and computational game theory
- Argumentation and defeasible reasoning
- Argumentation and nonmonotonic logics
- Argumentation and Decision Theory
- Argumentation and Logic Programming
- Argumentation and game semantics
- Software for teaching argumentation skills
- Argumentation-based interaction protocols
- Argumentation-based semantics of programs
- Argumentation in natural language processing
- Argumentation in human computer interaction
- Argumentation in multi-agent systems
- Computational models of natural argument
- Dialogue games and conversation policies
- Dispute resolution and mediation systems
- Electronic democracy and public deliberation
- Legal and medical applications
- Models of bargaining and economic interaction
- Reasoning about action through argumentation
- Computational tools for argumentation support

In order to distinguish itself from the competition, the new journal will not solicit papers that are wholly within the theory of argumentation without application (practical or theoretical) with in artificial intelligence or computer science, nor will it solicit computational work that fails to employ argumentation as a core focus.

TAYLOR & FRANCIS ONLINE SERVICES

Alerting Services

To sign up for table of contents alerts for **Argument and Computation** visit the Journal's homepage at:

www.tandf.co.uk/journals/tarc

To sign up for other table of contents, new publication and citation alerting services from informaworld visit:

www.informaworld.com/alerting

Online Access

Online access is included with a print institutional subscription to the journal or alternatively is available as an online only option. For further information connect to:

www.tandf.co.uk/journals/online.asp

Librarians' Area

Visit our web resource for Librarians and Information Professionals at:

www.tandf.co.uk/libsite

eUpdates

Register your email address at:

www.tandf.co.uk/journals/eupdates/asp to receive information on books, journals and other news within your areas of interest.

To the Librarian

Please include this journal in your next serials review meeting with my recommendation to subscribe:

From _____

Dept _____

Date _____

Please note this Taylor & Francis journal offers either print & online or online only subscriptions for institutions.

To find out more about our journals visit www.tandf.co.uk/journals

Submissions

Taylor & Francis would like to invite you to submit your paper to **Argument and Computation**.

The journal will accept full articles, describing novel theoretical or applied research in any of the areas of interest; reviews, condensing and critiquing an appropriate subfield of research; system descriptions, focusing on implementations (typically offering online access or downloadable code) and letters, providing pithy polemic on burning issues.

All submissions should be made online at the Argument and Computation Manuscript Central site, <http://mc.manuscriptcentral.com/tarc>.

For more information please visit www.tandf.co.uk/journals/tarc or contact chris.reed@computing.dundee.ac.uk



Taylor & Francis
Taylor & Francis Group