

# Argumentation and explanation in the context of dialogue

Floris Bex<sup>1</sup> and Katarzyna Budzynska<sup>2</sup> and Douglas Walton<sup>3</sup>

**Abstract.** Whilst computational argumentation and explanation have both been studied intensively in AI, models that incorporate both types of reasoning are few and far between. The two forms of reasoning need to be clearly distinguished, as they may influence dialogue protocol and strategy. Using the language of the Argument Interchange Format, we show that the distinction can be made by considering the speech acts used in a dialogue, and explore some of the implications of the combination of argument and explanation.

## 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 to convince someone but rather to help someone understand why the explananda are the case, that is, to help the explainee understand something that she claims she does not now understand, or does not completely understand [19]. 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.

Argumentation and explanation are well-presented in their respective sub-fields of AI. Computational models of argumentation have emerged and matured in the past twenty-or-so years [11]. Computational models for explanation are mainly based on the technique of abductive (model-based) reasoning, and have been studied in the context of medical and system diagnosis and natural language understanding (e.g. [4, 9, 5, 14]). 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 by Bex et al. [3], who combine structured arguments with abductive-causal reasoning into one model of inference to the best explanation. Other examples of work in which argumentation and explanation are combined are [9, 16].

Argumentation and explanation are often used in concert when performing complex reasoning: the explanations can 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 which the two types of reasoning are clearly distinguished; argumentation and explanation have different properties and the reasoning with arguments and explanations adheres to different patterns.

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, we argue that it is the illocutionary force of the speech act in a dialogue that determines whether reasoning is argumentation or explanation. We will use the conceptual model of the Argument Interchange Format [6, 13] 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 examples of both. Section 3 discusses our ideas for a framework for argumentation and explanation. Section 4 briefly explores the ramifications of combining argumentation and explanation in one comprehensive model of dialogical reasoning, and section 5 concludes the paper.

## 2 ARGUMENTATION AND EXPLANATION

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 parties<sup>4</sup> and reasons are not just given to support a conclusion but also to remove an opponent's doubts about this conclusion. For example, a reasoning  $\alpha \vdash \beta$  is argumentation when  $\beta$  is questioned (dubious) and a proponent of this argument uses  $\alpha$  not only to support  $\beta$ , but also to remove an opponent's doubts about  $\beta$ . Explanation, on the other hand, has not as its main goal to prove but rather to explicate why something is the case. Explanation in its purest form is not inherently dialectical: 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. Thus, explanations are usually offered for propositions that both parties agree on (i.e. we agree I am "late" and

<sup>1</sup> School of Computing, University of Dundee, Dundee, UK, DDI 4HN, email: florisbex@computing.dundee.ac.uk

<sup>2</sup> Department of Logic and Cognitive Science, Institute of Philosophy and Sociology Polish Academy of Sciences, Poland email: budzynska.argdiap@gmail.com

<sup>3</sup> Centre for Research in Reasoning, Argumentation and Rhetoric (CRRAR), University of Windsor, Canada email: dwalton@uwindsor.ca

<sup>4</sup> Hence the use of the term "calculus of opposition" for argumentation-theoretic semantics that allow one to calculate the acceptability of arguments

I explain why) whilst arguments are offered for propositions that are not immediately agreed upon (i.e. I contest the fact 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. The most thorough work on combining argumentation and explanation thus far is [3], who combines tree-structured arguments in the ASPIC<sup>+</sup> framework [10] with abductive-causal reasoning based on standard models of explanation [9] into a *hybrid theory* of inference to the best explanation. The basic idea of this hybrid approach is as follows. A logical model of abductive-causal reasoning takes as input a causal theory (a set of causal rules) and a set of observations that has to be explained, the explananda, and produces as output a set of hypotheses that explain the explananda in terms of the causal theory. Arguments can be used to support and attack stories, and these arguments can themselves be attacked and defeated. Thus, it is possible to reason about, for example, the extent to which an explanation conforms to the evidence. This is important when comparing explanations: the explanation that is best supported and least falsified by arguments is, *ceteris paribus*, the best explanation.

## 2.1 Distinguishing Argumentation and Explanation

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 Walton’s [18] 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  $\alpha$  causes  $\beta$ , and we observe  $\beta$ ,

we are allowed to infer  $\alpha$  as a possible explanation of  $\beta$ . In contrast, argumentation is often seen as reasoning from a premise  $\alpha$  to a conclusion  $\beta$  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. Walton’s [21] argument from evidence to hypothesis and causal argument). Similarly, one may perform explanatory reasoning by taking a rule  $\beta \rightarrow \alpha$ , meaning  $\beta$  is evidence for  $\alpha$  (see Bex et al. [3] 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. According to the pragmatic theory of speech act [15] argumentation and explanation are different speech acts. A speech act  $F\alpha$ , such as: *claim* $\alpha$ , *why* $\alpha$ , consists of an illocutionary force  $F$  and a propositional content  $\alpha$ . An illocutionary force is an intention of uttering a propositional content. That is, the performer of a speech act may utter  $\alpha$  with an intention of asserting, asking, promising and so on. Thus, argumentation and explanation are both instances of illocutionary acts that represent a relation between premises and conclusions: *argue*( $\alpha, \beta$ ) and *explain*( $\alpha, \beta$ ), where  $\alpha$  denotes a conclusion and  $\beta$  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 core ontology.

## 3 ARGUMENTATION AND EXPLANATION IN THE ARGUMENT INTERCHANGE FORMAT

The Argument Interchange Format (AIF) is a communal project which aims to consolidate some of the defining work on computational argumentation [6]. Its aim is to facilitate a common vision and consensus on the concepts and technologies in the field so as to promote the research and development of new argumentation tools and techniques. In addition to practical aspirations, such as developing a way of interchanging data between tools for argument manipulation and visualization [7], a common *core ontology* for expressing argumentative information and relations is also developed. Thus, the AIF ontology aims to provide a bridge between linguistic, logical and formal models of argument and reasoning.

The AIF core ontology is first and foremost an abstract, high-level specification of information and the various argumentative relations (e.g. inference, conflict) between this information.<sup>5</sup> The core ontology is intended as a conceptual model of arguments and the schemes or patterns arguments generally follow. It defines arguments and their mutual relations as typed graphs [6, 12], which is an intuitive way of representing argument in a structured and systematic way without the formal constraints of a logic [6]. This section briefly describes how in general the AIF describes argument and its dialogical context (Section 3.1). Then, we propose how to model argumentation and explanation in the language of the AIF (Section 3.2).

<sup>5</sup> The name Argument Interchange *Format* is in this respect somewhat misleading, as it seems to imply that AIF is a file format, whereas the AIF ontology can be implemented in a number of specific formats (XML, DOT, SQL). However, the name is retained for historical reasons.



valuable architecturally” (i.e.  $RA_2$ ). 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  $YA_4$  and  $YA_7$ , 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.

## 4 SCHEMES FOR ARGUMENTATION AND EXPLANATION

Using the machinery of the AIF core ontology described in section 3, we can distinguish between argumentation and explanation according to the dialogical context in which they are used. This allows us to combine argumentation and explanation in a principled way and paves the way for complex reasoning where, for example, arguments are used to justify explanations (cf. [3]) or explanation is used to clarify parts of an argument (as is the case in the example in Fig. 1, where the premise of the argument  $RA_2$  is explained).

The introduction of explanation into the AIF core ontology has a profound effect on the patterns of reasoning that are included in the ontology. Recall that the AIF core ontology is not only intended as a conceptual model of (object-level) arguments and explanations like the ones presented in section 3, but also as a repository of the schemes or patterns arguments generally follow. To this end, the core ontology also includes a so-called Forms Ontology, which contains these schemes. So in the ontology, relations like inference, conflict, transition and so on are treated as genera of a more abstract class of schematic relationships, which allows the three types of relationship to be treated in more or less the same way, which in turn greatly simplifies the ontological machinery required for handling them. Thus, inference schemes, conflict schemes and transition schemes in the Forms Ontology embody the general principles expressing how it is that  $q$  is inferable from  $p$ ,  $p$  is in conflict with  $q$ , and  $p$  is answerable with  $q$ , respectively. The individual RA-, CA- and PA-nodes that fulfil these schemes then capture the passage or the process of actually inferring  $q$  from  $p$ , conflicting  $p$  with  $q$  and answering  $p$  with  $q$ , respectively.

Inference schemes in the AIF ontology are similar to the rules of inference in a logic, in that they express the general principles that form the basis for actual inference. They can be deductive (e.g. the inference rules of propositional logic) or defeasible (e.g. argumentation schemes). Take, for example, the inference scheme for Argument from Expert Opinion [21]:

- *premises*:  $E$  is an expert in domain  $D$ ,  $E$  asserts that  $P$  is true,  $P$  is within  $D$ ;
- *conclusion*:  $P$  is true;
- *presumptions*:  $E$  is a credible expert,  $P$  is based on evidence;

Now, AIF arguments fulfil these schemes in a similar way to how inferences in logic instantiate inference schemes. For example, the argument *Peter says that the buildings are valuable architecturally and Peter is an expert on architecture*  $\rightarrow RA_3 \rightarrow$  *the buildings are valuable architecturally* would fulfil the scheme for argument from expert opinion. Note that the presumption, that Peter is credible and that his assertion is based on evidence, is not explicitly needed in the argument that fulfils the scheme: the idea of presumptions is that they can be assumed to hold unless proven otherwise. Thus, specific

(but still generalizable) knowledge can be modelled in the AIF in a principled way using argumentation schemes, for which we can assume, for example, a raft of implicit assumptions which may be taken to hold and exceptions which may be taken not to hold. These argumentation schemes then tell us how we can build valid and coherent arguments.

### 4.1 Transition Schemes

An argumentative dialogue (i.e.  $argument_2$ ) has an (often implicit) reply structure that contains the connections between the locutions in a dialogue. In the language of the AIF core ontology, these connections are explicitly rendered as transitions or TA-nodes (section 3). These transitions form the “glue” that keeps the locutions together and makes a dialogue coherent. This is analogous to non-dialogical argument, where logical (inference) connections (in the form of RA-nodes) form the glue between the individual propositions. The exact principles that make a dialogue coherent have been formulated and studied in the literature on formal dialogue systems [8]. At the heart of these systems are the dialogue protocols that describe a dialogue games permitted locutions, how and when the dialogue starts and ends and, perhaps most importantly, how locutions may be combined into exchanges.

In [2], the authors discuss *transition schemes* (following earlier work by Reed et al. [13]), schematic representations of a single transition in a dialogue. These transition schemes which can be instantiated to form transitions (i.e. a step in a dialogue), and these transitions can then be chained to form a dialogue. Note that the ontological machinery at work here is (intentionally) very similar to that of argumentation schemes, schematic representations of inference that can be instantiated to form inferences, which can be chained to form arguments. As an example of a transition scheme, consider the transition  $TA_3$  in Fig. 1, which is a particular instantiation of the following general scheme.

- *start locution*: Assert  $P$ ;
- *end locution*: Request Explanation  $P$

This scheme stands for the fact that assertions may be responded to by requesting an explanation of the information that is asserted. Another scheme is the one that is fulfilled by  $TA_4$ , which says that an explanation can be given if the other party requests it.

- *start locution*: Request Explanation  $P$ ;
- *end locution*: Explain  $P$

Thus transition schemes can be used to enforce, for example, that (as in the above scheme) an explanation may only be given when the other party asks for it. As Bex and Reed [2] show, it is also possible to define presumptions for transition schemes. For example, we might say that in order for someone to request an explanation after an assertion, the requesting party must somehow not understand the assertion completely (recall that explanations are often aimed at improving understanding). This can be incorporated into the above *assert – request explanation* scheme as a presumption, which means that the fact that the requester does not understand the assertion is implicitly assumed. That is, the requester does not have to explicitly say “I don’t understand” unless his understanding is actively challenged (i.e. “Why are you asking for an explanation, I think you understand perfectly!”).

Exactly which transition schemes are important and which conditions on these schemes (in the form of presumptions) we need has

been discussed in [19], where pre- and postconditions for the use of explanation are proposed. It remains to be investigated which types of conditions would be appropriate for a combination of argumentation and explanation. For example, one would only request an argument for some claim if there is doubt about this claim, and one would only request an explanation about a claim if there is a lack of understanding. Exactly how doubt or understanding should be defined remains as of yet an open question.

## 4.2 Explanation Schemes

In addition to argumentation schemes, there has also been work on what we call explanation schemes or scripts [14]. An explanation scheme is a generic scenario, an abstract rendering of a sequence of actions or events of a kind that is familiar to both the explainer and the explainee based on their common knowledge of how things can be normally expected to go in situations they are both familiar with. For example, the restaurant-script [14] contains information about the standard sequence(s) of events that take place when somebody goes to dine in a restaurant. Similar to argumentation and transition schemes, general explanation schemes can be instantiated by particular explanations and the scheme in a sense provides the conditions for the explanation's coherence (just as the argumentation scheme tells us what a coherent argument is and a combination of transition schemes tells us what a coherent dialogue is).

Take, for example, a man who enters a restaurant, orders some soup and gets his soup from the waiter. A natural continuation of this script would be that the man proceeds to eat his soup. If, for example, the man would instead remove his pants and offer them to the waiter, the story would be less coherent, because it does not seem to adhere to the typical restaurant scheme. But if this story fits another explanation scheme it can still be coherent. Suppose information is added to the script that the waiter spilled the hot soup on the man's legs. This new information would fill out the story in such a way that it hangs together as a coherent script about what happens when someone spills hot liquid on one's clothes. An expanded version of the story provides an explanation that helps the explainee to understand what happened. The explanation may be causal, motivational, teleological, or represent other kinds of explanations. We can represent the sequence of actions and events in this kind of story at a higher level of abstraction by fitting the script into an explanation scheme as an instance of it.

While the use of explanation schemes in argumentation has been explored recently [1], it is still unclear how they might be used in dialogue. Furthermore, what is currently also lacking is a principled exploration of different types of explanation schemes. Such explorations have been performed for argumentation schemes (e.g. [21]) and recently also for transition schemes [2].

## 5 CONCLUSIONS

In the paper, we propose the basic framework (based on the AIF ontology) for representing the difference between argumentation and explanation as a difference in illocutionary force (represented as YAnodes in the AIF ontology). Thus, we lay the basis for a principled combination of argumentation and explanation not only as reasoning structures but also in the context of reasoning processes or dialogues. We further explore some of the ramifications of combining argumentation and explanation, and how this combination is going to influence future work on reasoning schemes or patterns.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the support of EPSRC under grant EP/G060347/1 for Floris Bex, and the support from Polish National Science Center for Katarzyna Budzynska under grant 2011/03/B/HS1/04559 and the support of the Social Sciences and Humanities Research Council of Canada for Insight Grant 435-2012-0101 for Douglas Walton.

## REFERENCES

- [1] F. Bex, T. Bench-Capon, and B. Verheij, 'What makes a story plausible? the need for precedents', in *Legal Knowledge and Information Systems. JURIX 2011: The Twenty-Fourth Annual Conference*, ed., K.D. Atkinson, pp. 23–32, (2011).
- [2] F.J. Bex and C. Reed, 'Dialogue templates for automatic argument processing', in *Computational Models of Argument. Proceedings of COMMA 2012*, (2012). to appear.
- [3] F.J. Bex, P.J. van Koppen, H. Prakken, and B. Verheij, 'A hybrid formal theory of arguments, stories and criminal evidence', *Artificial Intelligence and Law*, **2**, 123–152, (2010).
- [4] T. Bylander, D. Allemang, M. C. Tanner, and J. R. Josephson, 'The computational complexity of abduction', *Artificial Intelligence*, **49**, 25–60, (1991).
- [5] A. Cawsey, *Explanation and Interaction: The Computer Generation of Explanatory Dialogues*, MIT Press, Cambridge, MA, 1992.
- [6] C.I. Chesñevar, 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**, 293–316, (2006).
- [7] J. Lawrence, F. Bex, M. Snaith, and C. Reed, 'Aifdb: Infrastructure for the argument web', in *Computational Models of Argument. Proceedings of COMMA 2012*, (2012). to appear.
- [8] Peter McBurney and Simon Parsons, 'Dialogue Games for Agent Argumentation', in *Argumentation in Artificial Intelligence*, eds., Iyad Rahwan and Guillermo Simari, volume 22, 261–280, Springer, (2009).
- [9] D. Poole, A. Mackworth, and R. Goebel, *Computational Intelligence*, Oxford University Press, 1998.
- [10] H. Prakken, 'An abstract framework for argumentation with structured arguments', *Argument and Computation*, **1**, 93–124, (2010).
- [11] 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).
- [12] I. Rahwan, F. Zablith, and C. Reed, 'Laying the foundations for a world wide argument web', *Artificial Intelligence*, **171**, 897–921, (2007).
- [13] C.A. Reed, S. Wells, K. Budzynska, and J. Devereux, 'Building arguments with argumentation : the role of illocutionary force in computational models of argument', in *Proceedings of COMMA 2010*, (2010).
- [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, 1969.
- [16] G. Tecuci, D. Marcu, M. Boicu, D.A. Schum, and Russell K., 'Computational theory and cognitive assistant for intelligence analysis', in *Proceedings of the Sixth International Conference on Semantic Technologies for Intelligence, Defense, and Security*, pp. 68–75, (2011).
- [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. Walton, 'Dialogical models of explanation', in *Papers from the 2007 AAAI Workshop*, Association for the Advancement of Artificial Intelligence Technical Report WS-07-06, pp. 1–9, Menlo Park, California, (2007). AAAI Press.
- [20] D. Walton and F.J. Bex, 'Combining explanation and argumentation in dialogue', in *Computational Models of Natural Argument - Proceedings of CMNA12*. Springer, (2012). to appear.
- [21] D.N. Walton, C.A. Reed, and F. Macagno, *Argumentation Schemes*, Cambridge University Press, Cambridge, 2008.