Discovering Inconsistency through Examination Dialogues

Paul E. Dunne Sylvie Doutre Trevor Bench-Capon
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF
United Kingdom
{ped,sd,tbc}@csc.liv.ac.uk

Abstract

In this paper we introduce *examination* dialogues, an addition to the dialogue typology of Walton and Krabbe. In educational settings the purpose of dialogue is often to elicit the position of a student, e.g. to test understanding. In other settings, a frequently adopted tactic is to attack an opponent's stance by exposing internal inconsistencies in their argument. In real debate such inconsistencies will often be rather more subtle than elementary logical fallacies since they arise from contradictions apparent in the opponent's value system. Protagonists will be better positioned to judge the applicability of this tactic as more information is determined concerning the exact nature of their opponent's case, e.g. the arguments favoured and values endorsed. One obstacle, however, is that following a request to state a view, the challenged party may refuse to comment. In this paper we present an approach to modelling the evolution of examination dialogues based on the concept of value-based argument frameworks and outline some algorithmic issues regarding argument selection.

1 Introduction

One important class of dialogue not found in the influential typology of [Walton and Krabbe, 1995] is the *examination* dialogue. In such dialogues one party – the *Questioner* $\mathbf{Q}$ – elicits statements and opinions from another – the *Responder* $\mathbf{R}$ – with the aim of discovering $\mathbf{R}$'s position on some topic, either to gain insight into $\mathbf{R}$'s understanding and knowledge of the topic, or to expose an inconsistency in $\mathbf{R}$'s position. Examples include education by the Socratic method, *viva voce* examination, cross examination of witnesses, and political interviews. In contrast to information seeking or inquiry dialogues, $\mathbf{Q}$ may already have beliefs on the topic: unlike persuasion dialogues, $\mathbf{Q}$ may have no intention of converting $\mathbf{R}$ to his position. Examination dialogues may, however, be nested within information seeking dialogues: probing for inconsistency increases confidence in the veracity of $\mathbf{R}$'s beliefs; similarly, in persuasion dialogues, exposing inconsistency is a useful prelude to persuasion.

In examination dialogues the dialogue process occurs within an environment of relevant facts, arguments, and related issues that is (at least implicitly) recognised and understood by both parties, e.g. common knowledge of the subject in educational settings, matters of specific policies and legislation in political interviews; issues of evidence and witness testimony in trials; details of incidents under investigation in questioning suspects. In this paper our aim is to describe a generic abstract model within which the evolution of such examination dialogues can be analysed. In addition we outline an approach intended to aid $\mathbf{Q}$ in selecting questions.

The foundational element for our model is provided by the Value-based Argument Frameworks (VAFs) of [Bench-Capon, 2003], a development of the classical argument systems of [Dung, 1995], which offer a richer interpretation of the concept of an argument x "attacking" an argument y. In addition to the computational advantages outlined in [Bench-Capon, 2003], further indications that VAFs provide a suitable mechanism for our analysis is presented by the methods discussed in [Dunne and Bench-Capon, 2004b]. A brief overview of VAFs is given in Section 2 and our model of examination dialogue presented in Section 3. In Section 4 we outline some algorithmic aspects involving argument selection in this scheme. Conclusions are presented in Section 5.

2 Value-Based Argument Frameworks

Value-based argumentation frameworks are introduced in [Bench-Capon, 2003] as a development of the model of argument systems via directed graphs $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$ proposed in [Dung, 1995]. A *value-based argumentation framework* (VAF) is defined as $\langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta \rangle$, where $\langle \mathcal{X}, \mathcal{A} \rangle$ is an argument system, $\mathcal{V} = \{v_1, v_2, \ldots, v_k\}$ is a set of k *values*, and $\eta : \mathcal{X} \to \mathcal{V}$ associates a value $\eta(x) \in \mathcal{V}$ with each argument $x \in \mathcal{X}$. An *audience* ϑ for a VAF $\langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta \rangle$ is a transitive and irreflexive binary relation on $\mathcal{V}$. A pair $\langle v_i, v_j \rangle$ in ϑ is referred to as 'v_i is *preferred to* v_j' with respect to ϑ.

Ideas analogous to those in Dung's argument system are now defined relative to some audience: $S \subseteq \mathcal{X}$ is *conflict-free* w.r.t. ϑ if $\forall\ x, y \in S$: $\langle x, y \rangle \in \mathcal{A} \Rightarrow \langle \eta(y), \eta(x) \rangle \in \vartheta$; x is acceptable to S w.r.t. ϑ if $\forall y \in \mathcal{X}$: $\langle y, x \rangle \in \mathcal{A} \Rightarrow (\langle \eta(x), \eta(y) \rangle \in \vartheta$ or $\exists\ z \in S$ s.t. $\langle z, y \rangle \in \mathcal{A}$ and $\langle \eta(y), \eta(z) \rangle \notin \vartheta)$; S is admissible w.r.t ϑ if each $x \in S$ is acceptable to S w.r.t. ϑ.

3 Examination Dialogues in VAFs

In keeping with the observation that debates of interest take place within some recognised context, we view this context as a VAF $\mathcal{G} = \langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta \rangle$. A *debate in the context of* $\mathcal{G}$ is a sequence $\mu_1 \mu_2 \ldots \mu_j$ of *at most* $2|\mathcal{X}|+1$ *moves*: odd indexed moves being played by $\mathbf{Q}$ and even indexed moves by $\mathbf{R}$.

Suppose the next move to be made is μ_{2i-1} ($i \geq 1$). The options for the next move(s) – $\mu_{2i-1} \cdot \mu_{2i}$ – are,
(a) $\mu_{2i-1} = \text{ASK}(p)$: $\mathbf{Q}$ asks $\mathbf{R}$ their view on $p \in \mathcal{X}$; $\text{ASK}(p)$ is legal provided that it has not occurred at an earlier point in the debate, i.e. $\forall\ j < i\ \mu_{2j-1} \neq \text{ASK}(p)$. The response μ_{2i} is one of $\text{ACCEPT}(p)$ ($\mathbf{R}$ agrees with p); $\text{DENY}(p)$ ($\mathbf{R}$ does not agree with p); or $\text{NO-COMMENT}(p)$ ($\mathbf{R}$ refuses to say anything about p).
(b) $\mu_{2i-1} = \text{FOUND}(p, \text{STAT}(p))$: here p is an argument to which $\mathbf{R}$ has responded $\text{NO-COMMENT}(p)$ after $\text{ASK}(p)$ earlier. $\mathbf{Q}$ expresses a claim to know $\mathbf{R}$'s view of p as accepted ($\text{STAT} = \text{ACCEPT}$) or otherwise ($\text{STAT} = \text{DENY}$).
(c) $\mu_{2i-1} = \text{RESOLVE}(S)$: S is a set of arguments all of which have been asked about with none receiving a "NO-COMMENT" reply. $\mathbf{Q}$ asserts that *at least* one of these replies cannot be correct with respect to the context $\mathcal{G}$.
(d) $\mu_{2i-1} = \text{CONCEDE}$: $\mathbf{Q}$ can neither determine $\mathbf{R}$'s view in respect of NO-COMMENT arguments nor expose inconsistencies in the collection of "definite" responses.

The debate ends once one of the moves (b)–(d) is played. Noting that we assume (b) or (c) can be played *only if* $\mathbf{Q}$ may justify the claim made, these result in a "win" for $\mathbf{Q}$, whereas (d) results in a "win" for $\mathbf{R}$.

As a very simple example consider: $\mathcal{X} = \{x, y, z\}$; $\mathcal{A} = \{\langle x, y \rangle, \langle y, z \rangle, \langle z, x \rangle\}$; $\mathcal{V} = \{A, B\}$; and $\eta(x) = \eta(z) = A$; $\eta(y) = B$. If $\mathbf{R}$ replies $\text{NO-COMMENT}(y)$ or $\text{DENY}(y)$ to $\text{ASK}(y)$ then $\mathbf{Q}$ wins (by playing $\text{FOUND}(y, \text{ACCEPT})$ or $\text{RESOLVE}(y)$) since, regardless of whether $\mathbf{R}$'s audience includes $\langle A, B \rangle$ or $\langle B, A \rangle$, $\mathbf{R}$ must accept y. On the other hand – assuming $\mathbf{R}$ replies $\text{ACCEPT}(y)$ – the only responses regarding $\mathbf{R}$'s view of $\{x, z\}$ that could lead to a win for $\mathbf{Q}$ are when *exactly one* NO-COMMENT occurs; or it is claimed both are accepted; or both denied. Notice that may $\mathbf{R}$ *report* $\langle \text{ACCEPT}(x), \text{DENY}(z) \rangle$ (or *vice versa*) despite believing otherwise and $\mathbf{Q}$ will be unable to detect this.

In general, within such debate contexts, both players have a number of non-trivial strategic issues to resolve. In particular, assuming that $\mathbf{R}$ wishes to hide information (and therefore will either respond "NO-COMMENT" and/or misreport some views), $\mathbf{R}$ may have to form a partition of $\mathcal{X}$ determining those queries to which false and true responses will be given. We note that, as indicated in the small example above, refusing to comment on any argument or misreporting the status of every argument may fail to be a successful strategy; moreover it is an easy consequence of [Dunne and Bench-Capon, 2004a] that the decision problem faced by $\mathbf{R}$ determining if classifying p as NO-COMMENT loses, is coNP–complete.

4 Choosing Arguments

We conclude the brief technical discussion, by outlining an approach to query selection that is based on a translation to propositional formulae. Given $\langle \mathcal{X}, \mathcal{A}, \mathcal{V}, \eta \rangle$, in order for $p \in \mathcal{X}$ to be accepted w.r.t. at least one audience ϑ, it must belong to an admissible w.r.t. ϑ set $S \subseteq \mathcal{X}$. Consider disjoint sets of propositional variables Z and $W = \{w_{i,j} : 1 \leq i \neq j \leq k\}$. Define the function, $f(Z, W)$ via: $f(\alpha, \beta) = \top$ if and only if the subset S of $\mathcal{X}$ indicated by α, i.e. $x_i \in S \Leftrightarrow \alpha_i = \top$, is admissible with the binary relation ϑ encoded in β, i.e. $\langle v_i, v_j \rangle \in \vartheta \Leftrightarrow \beta_{i,j} = \top$ *and* this relationship *is* an audience, i.e. irreflexive and transitive. This function can be described by a "short" propositional formula $\Phi_f(Z, W)$ of length $O(|\mathcal{A}| + k^3)$ that with the addition of at most $O(|\mathcal{A}|)$ auxiliary variables (D) translates to an equivalent CNF formula, $\Psi_f(Z, W, D)$ with $O(|\mathcal{A}| + k^3)$ clauses, the longest of which contains at most $2 + \max_{x \in \mathcal{X}} |\{y : \langle y, x \rangle \in \mathcal{A}\}|$ literals. Using Ψ_f the strategy pursued by $\mathbf{Q}$ will be to obtain responses from $\mathbf{R}$ that result in Ψ_f evaluating to $\perp$ (i.e. $\mathbf{R}$ is inconsistent) or to infer the unique instantiation of z_i that satisfies Ψ_f with the replies already given.

5 Conclusions and Development

We have introduced a class of question-response dialogues – *examination* dialogues – and modelled them using VAF schemata as the basis for representing related knowledge. A number of issues – both theoretical and applicative – form the subject of work in progress. In addition to on-going empirical investigations of argument selection strategies employing the propositional translation briefly outlined in Section 4 there are several questions of interest regarding properties of the model when the protagonists are "limited" in different ways. Thus, to what extent is $\mathbf{Q}$ at an advantage if an upper bound is placed on the number of "NO-COMMENT" and/or false responses $\mathbf{R}$ can make? Similarly, if $\mathbf{Q}$ is restricted to some maximum number of questions (rather than being allowed to ask about the status of *every* argument in $\mathcal{X}$) it is likely that $\mathbf{Q}$ must adopt a dynamic selection mechanism (rather than choosing a set to query in advance) so as to adapt effectively to $\mathbf{R}$'s responses.

References

[Bench-Capon, 2003] T. J. M. Bench-Capon. Persuasion in Practical Argument Using Value-based Argumentation Frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[Dung, 1995] 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.

[Dunne and Bench-Capon, 2004a] P. E. Dunne and T. J. M. Bench-Capon. Complexity in value-based argument systems. In *Proc. 9th JELIA*, volume 3229 of *LNAI*, pages 360–371. Springer-Verlag, 2004.

[Dunne and Bench-Capon, 2004b] P. E. Dunne and T. J. M. Bench-Capon. Identifying audience preferences in legal and social domains. In *Proc. DEXA'04*, volume 3180 of *LNCS*, pages 518–527. Springer-Verlag, 2004.

[Walton and Krabbe, 1995] D. N. Walton and E. C. W. Krabbe. *Basic Concepts of Interpersonal Reasoning.* State Univ. of New York Press, 1995.