

Logic and Logical Philosophy Volume 27 (2018), 491–515 DOI: 10.12775/LLP.2017.021

Jacky Visser, Katarzyna Budzynska, and Chris Reed

A CRITICAL DISCUSSION GAME FOR PROHIBITING FALLACIES

Abstract. The study of fallacies is at the heart of argumentation studies. In response to Hamblin's devastating critique of the state of the theory of fallacies in 1970, both formal dialectical and informal approaches to fallacies developed. In the current paper, we focus on an influential informal approach to fallacies, part of the pragma-dialectical theory of argumentation. Central to the pragma-dialectical method for analysing and evaluating argumentative discourse is the ideal model of a critical discussion. In this discussion model, a dialectical perspective on argumentation is combined with a pragmatic take on communicative interaction. By formalising and computationally implementing the model of a critical discussion, we take a first step in the development of software to computationally model argumentative dialogue in which fallacies are prohibited along the pragmadialectical norms. We do this by defining the Critical Discussion Game, a formal dialogue game based on the pragma-dialectical discussion model, executable on an online user-interface which is part of a larger infrastructure of argumentation software.

Keywords: Argument Web; argumentation; critical discussion; Dialogue Game Description Language; fallacies; formal dialectics; pragma-dialectics

1. Introduction

The fallacies have traditionally played a central role in the study of argumentation. This is not surprising, because the research field of argumentation studies is motivated in part by the objective of improving argumentative practice. This was the case both in antiquity—where

rhetorical skills were essential for citizens in the legal and democratic process — and in modern argumentation studies — where (among others) the influential informal-logical [18] and pragma-dialectical [38] approaches were motivated by practical application. In order to improve practice it is necessary to define what should be remedied; what faults should be repaired; what mistakes should be prevented. These recurrent mishaps of argumentative practice are commonly referred to as fallacies.

After launching a devastating critique of the state of the study of fallacies in 1970, Hamblin [12] proposed to employ the machinery of formal dialectical systems as a means of studying fallacies. Within formal dialectical approaches, (argumentative) dialogue is represented in terms of highly regulated game-like interactions in accordance with a set of rules. Fallacies are considered, in such systems, as violations of the rules that constitute a reasonable dialogue, a perspective advocated by, among others, Mackenzie [27], Hintikka [15], and Walton [54].

Another response to Hamblin's critique can be found in informal approaches to fallacies. The pragma-dialectical theory, developed over the past forty years by van Eemeren and Grootendorst [46, 47, 48] and their co-authors [44], is one such informal approach and counts as one of the catalysts of modern argumentation studies. Building on the formal dialectical approach (in particular those of the Erlanger Schule [26] and of Barth and Krabbe [3]), it combines a dialectical perspective on argumentation with a pragmatic account of the everyday communicative practice that argumentation takes place in. This combination leads to a unified account of fallacies as committed through speech acts in all stages of an argumentative discussion.

We aim to bring the formal and informal approaches to fallacies closer together again by formalising part of the pragma-dialectical theory. We subsequently link the formalised model to an existing software infrastructure by implemented it as a computationally executable dialogue system. We take the pragma-dialectical approach to fallacies as our starting point for the development of a fallacy-sensitive component of the Argument Web. The Argument Web is an integrated collection of computer tools to assist in the production, analysis and evaluation of argumentative discourse [5]. The computational implementation has the added advantage of promoting the development of software in support of the pragma-dialectical theory, which has been notably absent from the ongoing rise of computational methods in the field of argumentation studies.

In the current paper, we make one step towards the twofold objective of developing a fallacy-sensitive component for the Argument Web and promoting the computational application of the pragma-dialectical theory. We do this by computationally implementing the discussion model that is central to the pragma-dialectical account of fallacies in a language that interacts with the Argument Web. In Section 2, we will introduce the theoretical backgrounds and tradition against which our work is situated, consecutively turning to: the conception of 'fallacy' that we employ, the pragma-dialectical ideal model of a critical discussion, and the relation to computational argumentation studies. In Section 3, we will formalise the pragma-dialectical discussion model in terms of a dialogue game, and then computationally implement the resulting 'Critical Discussion Game' in the Dialogue Game Description Language. In Section 4, we will turn to the execution of the implemented game on an online user-interface, and we will discuss how the Critical Discussion Game prohibits certain fallacies from being committed.

2. Theoretical backgrounds

2.1. The study of fallacies

As alluded to in the Introduction, the sense of 'fallacy' intended here is that of common forms of bad argumentative conduct (as opposed to the use of 'fallacy' as a reasoning mistake [13]). According to the *Handbook of Argumentation Theory* [43], fallacies contaminate an argumentative exchange, thereby preventing a satisfactory resolution of a difference of opinion if they go undetected.

The study of fallacies, after decades of scholarly neglect, was firmly put back on the academic map by Hamblin with his monograph Fallacies [12]. Hamblin characterised what he judged to be the mistaken traditional way of treating fallacies, primarily in textbooks of logic, as the Standard Treatment. Hamblin reconstructed the definition of a fallacy according to the Standard Treatment as "an argument that seems to be valid but is not so", and he showed that this definition is unsatisfactory. Some historically recognised fallacies are not arguments at all. It is unclear what it entails to seem valid, and to whom it should appear so. Furthermore, provided a fallacy can be classified as an argument that seems valid, it may actually be exactly that: a logically valid argument.

Let us consider two examples of dialogue contributions that are not conducive to reasonable argumentative discourse (i.e. that do not promote the reasonable resolution of a difference of opinion), but that are not covered by the definition of the Standard Treatment.

- (1) a. Bob: Dundee is a very interesting city.
 - b. Alice: Why is that?
 - c. Bob: Dundee has some interest because of its proximity to the Scottish Highlands.

In (1), Bob misrepresents his earlier standpoint that Dundee is a *very* interesting city by weakening it to a city having *some interest*. In doing so, Bob commits a fallacy of *ignoratio elenchi*, weakening his original standpoint to make it easier to defend. Bob's violation here is not found in the argument that he provides (Dundee is close to the Highlands, and this might be an acceptable reason for considering Dundee an interesting city), but rather with his misrepresentation of his standpoint. The committed fallacy itself then is not an argument, but rather lies in a procedural inconsistency in Bob's assertions.

- (2) a. Alice: We should go on vacation to Poznań.
 - b. Bob: Why is that?
 - c. Alice: Because our holiday destination should be Poznań!

In (2), the fallacy of petitio principii (also called begging the question or circular reasoning) is committed by Alice when she uses one and the same proposition as the standpoint at issue and as the argumentative defence thereof. Contrary to the rather clear-cut case in (2), in practice, the circularity might not always be immediately obvious because it is veiled by implicit entailment or implicature. Elementary circular arguments, though, are considered valid in most systems of logic, thereby defying the definition of the Standard Treatment that Hamblin reconstructs.

For Hamblin, the topic of fallacies posed an opportunity to employ the machinery of formal dialectics [28]. In formal dialectical approaches, fallacies are interpreted as violations of the procedural rules of engagement that hold in particular discursive systems or settings. Hamblin's pioneering of dialectical systems that prohibit fallacious conduct is continued by, among others, Mackenzie [27], Woods and Walton [58, 54], and Barth and Krabbe [3], who propose further systems to prohibit specific examples of fallacies. The logic textbook approach of fallacies that was criticised by Hamblin was not amended much—in some cases

merely leading to the complete disregard of the topic — but others took up the challenge and proposed various explanations of and perspectives on fallacies. An epistemological account of fallacies as failures in the rational acquisition of new beliefs is advocated by Biro and Siegel [6]. Within the informal logic movement, fallacies have been interpreted as frequently occurring incogent arguments with unacceptable, irrelevant or insufficient premises [18]. Woods and Walton, both together and independently, approach the fallacies one by one, showing how each can be explained against the background of a plurality of formal dialectical, logical, and pragmatic systems [60, 56, 59].¹

2.2. The pragma-dialectical conception of fallacies

Van Eemeren and Grootendorst [47] continue the formal dialectical approach to fallacies, but combine it with a pragmatic account of argumentation as embedded in everyday communicative interaction. In their pragma-dialectical theory, they see fallacies as potentially persuasive, but unreasonable, discussion moves that violate the rules of the *ideal model of a critical discussion* [48]. This ideal model is a proposal for a discussion procedure that is fully instrumental to the resolution of a difference of opinion in a reasonable way. True to name of the overarching pragma-dialectical theory, the discussion model conveys a pragmatic and a dialectical dimension.²

The pragmatic dimension of the ideal model does justice to the fact that argumentation forms part of communicative interactions.³ In the ideal model, all of the discussion moves are conceived of in terms of speech acts [46]. An inventory is made of the speech acts that contribute to the resolution of the difference of opinion. The Searlean [32] speech act perspective on language use is in the pragma-dialectical the-

¹ The state of the field of fallacy theory is described in more detail by, among others, Hansen [13].

² In the current paper, only the 'standard' pragma-dialectical model is considered. In an extension of the model, van Eemeren and Houtlosser [40] added a rhetorical dimension to the model, in order to account for the strategic manoeuvring of actual arguers in conventionalised argumentative activity types in their attempts to persuade their interlocutor.

³ The pragmatic dimension should not be mistaken for one aimed at practical reasoning or action-oriented decision-making. In the sense intended here, 'pragmatic' refers to the traditional distinction within linguistics between syntax, semantics and pragmatics [25].

ory integrated with a Gricean [11] perspective on verbal interaction. This integration makes it possible to use pragmatic tools from discourse analysis to perform a dialectical reconstruction. Unexpressed or implicit parts of the discussion can be reconstructed on the basis of pragmatic insight by interpreting communicative acts as part of a discussion aimed at resolving a difference of opinion in a reasonable way.

The dialectical dimension refers to the dialectical perspective on argumentation that is taken. Argumentation is interpreted as always involving a (possibly implicit) discussion between two discussants. To do justice to all parts of the text that are argumentatively relevant, the entire discussion must be taken into account, not just the advancing and criticising of arguments. By taking the full argumentative discussion into consideration, four discussion stages are distinguished [48]. First, in the confrontation stage, the interlocutors externalise their difference of opinion about an expressed opinion (e.g., an opinion, a belief, a plan of action, etc.). Second, in the opening stage, the discussants agree upon a set of mutually accepted material and procedural starting points. The discussants need at least one shared material starting point in order to have a meaningful discussion. With respect to the procedure, they need, among other things, to agree who will perform the role of protagonist defending the standpoint and who will be the antagonist criticising it.⁴ Third, in the argumentation stage, the protagonist will provide one or more arguments in defence of the standpoint, while the antagonist critically examines the arguments. Fourth, in the concluding stage, the discussants draw a conclusion on the basis of the critical testing of the arguments. If the argumentation was conclusive, both discussants will have to accept the standpoint. In case the criticism was conclusive, both discussants may no longer accept the standpoint.⁵

⁴ In the remainder of this paper, as in van Eemeren and Grootendorst's [48] introduction of the ideal model, the assumption will be that the interlocutor putting forward the standpoint will also be the one defending it as protagonist in the argumentation stage of the discussion. For reasons of clarity, the protagonist will be referred to with male pronouns, and the antagonist with female pronouns (following the standard way of referring to proponents and opponents in the literature on formal dialectics and dialogue logic). The same convention will later be followed for the players Protagonist and Antagonist in the dialogue game.

⁵ No longer accepting a standpoint should not be confused with accepting an opposite standpoint. This mistake would constitute an instance of the fallacy of 'making an absolute of the failure of the defence' or an *ad ignorantiam* [47].

The rights and obligations that discussants have in the four stages of the discussion are regulated in fifteen procedural rules [48].⁶ The fifteen rules for critical discussion specify the ways in which discussants can reasonably defend and attack standpoints. Assuming only one standpoint is at issue in the discussion, the protagonist defends it by advancing argumentation, which is critically assessed by the antagonist. If this argumentative defence is not successful at convincing the antagonist, further argumentation may be called for. Depending on the kind of criticism the additional argumentation is meant to overcome, different argumentation structures develop [36].

The pragma-dialectical model is based on a critical conception of reasonableness [39]. As such, it is a normative proposal for ideal argumentative conduct, rather than a description of actual argumentative practice. This normative perspective makes it possible to use the model as an analytical and evaluative tool in the identification of fallacies in argumentative discourse. A descriptive model would in itself be 'contaminated' by the potential imperfections of argumentative practice, and would therefore not be suitable to evaluate the same discourse it already describes. As Hamblin [12] rightly observes, a normatively oriented model should retain empirical relevance. For this reason, van Eemeren, Garssen and Meuffels [45] have performed a series of quantitative empirical studies which support the conventional (or intersubjective) validity [2] of the pragma-dialectical model. The normativity of the pragma-dialectical model is one reason for us to use it as the basis for our Critical Discussion Game for prohibiting fallacious conduct in a computational dialogue system.

A second reason is found in the comprehensive scope of the model. It covers not only inferential or formal fallacies, but allows for the identification of fallacious discussion moves that impede the reasonable resolution of the difference of opinion in other ways [47]. The model covers all stages of a discussion, and therefore deals not just with fallacies relating to the advancement and criticism of argumentation, but also to fallacies committed in putting forward standpoints, in agreeing on the starting points of a discussion, and in determining the outcome of a discussion. The pragma-dialectical interpretation of fallacies is also not restricted

⁶ The fifteen rules should not be confused with the ten 'practical' commandments for reasonable discussants [47], which are intended as practical rules of thumb for conducting, analysing and evaluating argumentative discussions.

to the role of the protagonist arguing in defence of a standpoint, but extends to the role played by the antagonist in criticising the argumentation. Lastly, using the ideal model of a critical discussion as the sole basis for deciding fallaciousness leads to a unified account, treating all fallacies as equal, and not taking an ad-hoc, historical list as the starting point to give individual solutions to individual fallacies.

Despite criticisms of the pragma-dialectical theory as a whole [41], it serves as a useful starting point for the dialogical approach to fallacies in a computational system. Rather pragmatically, we start from the assumption that if the model indeed provides a norm for reasonable argumentation, thereby prohibiting fallacies, then this norm is transferable into a formal dialogue game and subsequently into a computational implementation. The resulting computational implementation, finally, will provide a basis for the prohibition of (some classes of) fallacious moves in the production of argumentative discourse in a computational system.

2.3. A computational approach to dialogical fallacies

Over the past twenty years, computational methods have gradually found their way into the field of argumentation studies, leading to a new field of 'computational argumentation studies'. The work in this field can broadly be divided into two categories. On the one hand, insight about argumentation in natural language has inspired new approaches in Computer Science and Artificial Intelligence (i.e. applying argumentation studies to solve problems in Computer Science). On the other hand, computer tools are increasingly used to support the study of argumentation in the Humanities and Social Sciences (i.e. applying Computer Science to support argumentation studies).

The digital tools to support the production, the analysis, and the evaluation of argumentative discourse are based on computational implementations of formal models of argumentation. These models serve as the conceptual framework that determines the interpretation given to central notions such as 'argument', 'premise', 'conclusion', and 'argument scheme'. Because the various theoretical models of argumentation provide different interpretations and explanations of these notions, the computer tools that are based on these models are not theory-neutral and not fully compatible with other theoretical approaches. To mediate

⁷ Overviews of the field of computational argumentation studies are provided by, e.g., Rahwan and Simari [30], van Eemeren et al. [42], and Baroni et al. [1].

the interchange between the various tools and mitigate the (conceptual) boundaries between the approaches, the Argument Interchange Format (AIF) [8] is proposed as a theory-neutral ontology of argumentative concepts, i.e. a generic set of properties for the characterisation of the aforementioned argumentation theoretical notions. Via the AIF the notions of one theoretical framework can be translated into those of another.

The AIF also provides the foundation for the Argument Web [5], an infrastructure of computer software for argumentative tasks. The Argument Web is developed as an integrative set of computational tools to support argumentative tasks, interacting with a database of interconnected argumentative content. By building software on the basis of the AIF, the tools can be used in conjunction, with the output of one serving as the input for another. Furthermore, because the AIF serves as a theory-neutral ontology of argumentation, it facilitates the 'translation' of other computational and non-computational approaches to argumentation in order to allow interaction with the components of the Argument Web. In this way, it becomes possible to, for example, analyse an argumentative text using Rationale [50], save the result in AIFdb [24], and then diagram it in Carneades [10]. Strengths of different theoretical approaches can thus be exploited, and a lack of functionality of one tool can be overcome by using another.

The tools that are designed to be part of the Argument Web can also be used in succession, carrying over data from one application to the next. For example, an existing text can be analysed using OVA (ova.arg.tech), an online tool for the visualisation of argumentation [17]. The resulting illocutionary and inferential annotation [7] can be stored in AIFdb (aifdb.org). AIFdb is a searchable online repository of annotated argumentive texts, showing the relations of support and conflict between propositions and their discursive embedding as a directed graph [24]. Using TOAST (The Online Argument Structures Tool, toast.arg.tech), the acceptability of the argument can be computed on the basis of Dungian abstract argumentation frameworks [35, 9]. Through Arvina (arvina.arg.tech), AIFdb can be navigated, manipulated and extended by dialogically interacting with both other human users, and artificial agents [34]. The new argumentative content produced in these dialogues in turn contributes to AIFdb.

While the analysis and production of argumentation are relatively well-covered by the Argument Web, there is at present no tool for the evaluation of argumentative discourse beyond calculating the acceptability of conclusions on formal grounds. In other words, there is currently no tool for the detection of fallacies in the Argument Web and for the prohibition of fallacies in the dialogical interactions. In the current paper, we start developing a dialogue protocol for Arvina that prohibits users from committing certain fallacies in their interaction with AIFdb.

3. Formalising and computerising critical discussion

In this section, we first introduce a dialogue game that formalises part of the pragma-dialectical model of a critical discussion. Subsequently, we implement the resulting Critical Discussion Game as a computerexecutable dialogue game.

In our formalisation of the pragma-dialectical ideal model of a critical discussion, we make use of the notion of a 'dialogue game'. Dialogue games provide an abstract way of looking at communication as a game, played by interlocutors to reach an interactional goal (see, e.g., Hamblin [12]; Lorenzen and Lorenz [26]; Rescher [31]; Hintikka [14]; Walton [54]). Formality can mean various things with respect to dialogue games. Barth and Krabbe [3, 21] distinguish five senses of the term 'formal'. In line with related work by Krabbe [23] and Visser [52], our dialogue game is intended to i) provide rules for the correct assembly of moves and move sequences, ii) implement a procedural regimentation of the interaction, and iii) be normative (or a priori). These characteristics make the dialogue game, respectively, formal₂, formal₃, and formal₄ in Barth and Krabbe's classification. The ideal model of a critical discussion itself is already procedurally regimented (formal₃) and normative (formal₄) [43, 23]. These two senses should, obviously, be retained in the formalisation.

Due to the complexity of the pragma-dialectical ideal model, we construe a dialogue game based on a simplified version of the model. The resulting simplified dialogue game demonstrates the feasibility of the formalisation and can serve as a foundation for future extensions aimed at bringing the scope of the dialogue game closer to that of the original ideal model (see Visser [53] for an elaboration on this incremental

⁸ The first and the fifth sense of 'formal' distinguished by Barth and Krabbe respectively relate to Platonic forms and to non-material models. Neither of these is relevant to the present study. There are other classifications of formality, such as Johnson and Blair's [19], but the one by Barth and Krabbe provides a good point of reference and allows for the distinctions needed in the current dialogue game.

approach). The simplifications with respect to the ideal model are chosen such that a consistent dialogue game can be specified, which still exhibits the features characteristic of the pragma-dialectical model of a critical discussion. In the elementary case of critical discussion, which the dialogue game is based on, only one (positive) standpoint is at issue. This single standpoint is considered to be met with doubt in the confrontation stage, not with a concrete contradictory standpoint. This leads to a single non-mixed difference of opinion, which is the assumed starting point in the current paper. More complex differences of opinion can occur when more than one standpoint is at issue (leading to a multiple difference of opinion) or when instead of doubt a contradictory standpoint is advanced in the confrontation stage (resulting in a mixed difference of opinion). The complex differences of opinion can always be decomposed into two or more elementary differences, which is why the focus will be on the elementary case from now on [46].

To maintain a clear pragma-dialectical outset while constraining the complexity of the model, the dialogue game offers fewer choices and opportunities to players than the original model does. This restriction is most evident in the exclusion of argumentative attacks and mixed differences of opinion, only allowing a single argument in defence of a positive standpoint against doubt. Furthermore, only the argumentation stage and the concluding stage of critical discussion are explicitly reflected in the dialogue game. The other two discussion stages—the confrontation stage and the opening stage—are accounted for by means of specific assumptions. Lastly, the various argument schemes that give substance to the (defeasible) inferences underpinning argumentation are not modelled in the current dialogue game.

The resulting simplified Critical Discussion Game is played by two players, performing the discussion roles of protagonist and antagonist, respectively argumentatively defending and criticising the standpoint at issue. Assumed is that the players have agreed upon a non-empty set of material common starting points (i.e. a set of propositions they both consider acceptable). Figure 1 is a visualisation of the sequential structure of the game. The nodes of the graph represent the moves of the game; the edges represent the sequential transitions between moves.⁹

⁹ The visualisation of the sequential structure of the dialogue game in terms of a graph is similar to the dialectical profiles used as a heuristic in the pragma-dialectical

The first move, at the top of Figure 1, is made by the protagonist to advance an argument in defence of his standpoint. The antagonist can respond in one of three ways. First, she can outright accept the argumentation (the middle route in Figure 1). Second, she can indicate that she has doubts about the acceptability of the propositional content of the argumentation (left route). Third, she can indicate that she doubts whether the proposition put forward as an argument actually provides a justificatory defence of the standpoint (right route). In response, the protagonist can invoke one of two intersubjective procedures to verify whether the propositional content and the justificatory force of the argumentation are acceptable. The dialogue game terminates when either the protagonist retracts his argumentation and subsequently his original standpoint or when the antagonist accepts the argumentation and standpoint (bottom of Figure 1). Based on the commitments that the players acquire during the game, either the protagonist or the antagonist wins. ¹⁰

The Critical Discussion Game can be specified in terms of the Dialogue Game Description Language (or DGDL) [57].¹¹ Specifying the game in DGDL has several advantages. Using this standardised format facilitates the comparison to other dialogue systems, such as Walton's CB [54], the Mediation Dialogue Game [16], or Lakatos' dialogical approach to mathematical proof [29]. Another advantage, shown by Wells and Reed [57], is the avoidance of the under-specification of the dialogue rules that can occur in non-computationally-implemented formal descriptions of dialogue games. Furthermore, the DGDL-specification makes it possible to computationally execute the dialogue game on the Dialogue Game Execution Platform [4], which we will turn to in Section 4.

The specification of a dialogue game in DGDL consists of three parts. The first part, *composition*, covers the general features of the dialogue game. Listing 1 shows part of the composition in DGDL of the Critical Discussion Game. The Critical Discussion Game is played in a strict turn by turn fashion by exactly two players (Protagonist and Antagonist). Four stores are instantiated to keep track of the propo-

analysis of argumentative texts [49] and the profiles of dialogue as employed by Walton [55, 20] and Krabbe [22].

 $^{^{10}}$ Visser [51] presents a more elaborate interpretation of the pragma-dialectical ideal model as a dialogue game, and provides more justification of the adequacy of the resulting formalisation.

¹¹ We will only discuss a few excerpts of the code here; the full DGDL-specification of the Critical Discussion Game is available online at arg.tech/cdg.

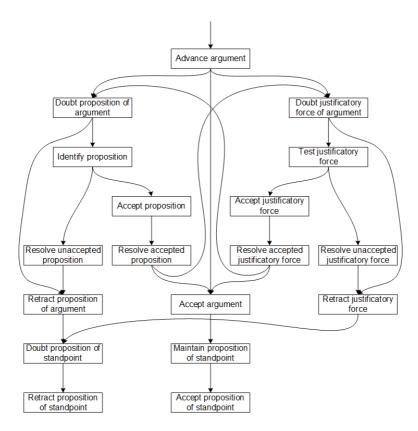


Figure 1. The sequential structure of the Critical Discussion Game

sition at issue in the discussion as the Standpoint (the proposition Britain should disarm is put in as an example), some propositions that make up the common ground (there is no moral case for using nuclear weapons, Britain has a leading role in world politics, and it is good to promote non-proliferation), and the commitments of each player (the common starting points for both, complemented with the standpoint at issue for the Protagonist — but explicitly not for the Antagonist, for obvious reasons).

Listing 1. Composition of Critical Discussion Game in DGDL

```
turns{magnitude:single, ordering:strict}
roles{Protagonist, Antagonist}
players{min:2, max:2}
```

```
store{id:Standpoint, owner:shared, structure:set,
   visibility:public, {"Britain should disarm"}}
store{id:CommonStartingPoints, owner:shared,
   structure:set, visibility:public, {"there is no
   moral case for using nuclear weapons", "Britain
   has a leading role in world politics", "it is
   good to promote non-proliferation"}}
store{id:CS, owner:Protagonist, structure:set,
   visibility:public, {"there is no moral case for
   using nuclear weapons", "Britain has a leading
   role in world politics", "it is good to promote
   non-proliferation"}}
store{id:CS, owner:Antagonist, structure:set,
   visibility:public, {"Britain should disarm",
   there is no moral case for using nuclear weapons
   ", "Britain has a leading role in world politics
   ", "it is good to promote non-proliferation"}}
```

The second part of the DGDL specification, rules, defines actions that always occur at the beginning or end of a dialogue, or at the end of turns or moves. Listing 2 shows the sole rule of the Critical Discussion Game. This specifies that the Protagonist makes the first move of the game, and that this initial move always consists of putting forward an argument in defence of the standpoint at issue, where the propositional content of the argument may not be identical to that of the standpoint.

Listing 2. Rules of Critical Discussion Game in DGDL

```
rule { id: StartingRule, scope: initial, {
    assign (Protagonist, speaker)
    & move (add, next, Argue, {p,q}, Protagonist, {
    inspect (in, {p}, Standpoint) & inspect (!in, {q},
    Standpoint}) }}
```

The third part, interactions, catalogues the moves that players can make during the game, when they can make them, and to what effect. The interactions make up the main part of the Critical Discussion Game specification. Because the game is not symmetrical in terms of the players' goals (they play different roles in either defending or attacking the acceptability of the proposition at issue), there are distinct sets of moves available to each player. Listing 3 shows three of these interactions in the Critical Discussion Game.

The Argue-move, that was already alluded to before as the mandatory first move of every game, is used by the Protagonist player

to provide argumentation in defence of his standpoint. In the code, this is represented as entailing the assertion of the standpoint p, the argument q, and establishing the justificatory relation between them (i.e. the inference of p from q). A paraphrase in natural language is provided to facilitate a user-interface: ''My reason for p is g''. Propositions p and q, and the inferential relation between them, are added to the Protagonist's commitment store CS. Finally, three subsequent moves are made available to the Antagonist: DoubtProp, DoubtJust, AcceptArg. The last of these is shown in Listing 3 as a second example of an interaction: the Antagonist can respond to the Protagonist's argumentation by accepting it with the AcceptArgmove. This entails accepting both the propositional content of the argumentation and the support this lends to the standpoint, and therefore leads to accepting the standpoint itself. The last example of an interaction shows the RetractProp-move used by the Protagonist to retract his commitment to the acceptability of a proposition. This proposition can be either an argument or a standpoint. The result is that a proposition p is deleted from the Protagonist's commitment store. This move is, furthermore, one of the two ways of terminating the game (the other being the acceptance of the standpoint by the Antagonist), if the retracted p was the proposition at issue in the discussion (i.e. if p refers to the standpoint under discussion).

Listing 3. Interactions of Critical Discussion Game in DGDL

```
interaction(Argue, {p,q}, Asserting, {p}, Asserting,
    {q}, Arguing, {<{q}, {p}>, DefaultInference}, "My
   reason for $p is $q", {
    store(add, {q, {<{q}, {p}>, DefaultInference}},
   CS, Protagonist)
    & move(add, next, DoubtProp, {q}, Antagonist)
    & move(add, next, DoubtJust,
                                   {<{q},{p}>,
   DefaultInference}, Antagonist)
    & move(add, next, AcceptArg, {p,q}, Antagonist)
   } }
interaction{AcceptArg, {p,q}, Accepting, {p},
   Accepting \{q\}, Accepting, \{\langle q \rangle, \{p \rangle\rangle,
   DefaultInference}, "I accept $q as a conclusive
   reason for $p", {
    store(add, \{p, q, \{<\{q\}, \{p\}>, DefaultInference\})\}
   }}, CS, Antagonist)
    & move(add, next, Maintain, {p}, Protagonist)}}
```

4. Fallacies and the computational Critical Discussion Game

The DGDL-specification allows the Critical Discussion Game to be executed using the Dialogue Game Execution Platform (DGEP) [4]. DGEP forms the basis for user-interfaces, such as the online user-interface Arvina (arvina.arg.tech), to allow a dialogue to be 'played out' in accordance with a protocol specified in DGDL. With Arvina, users can engage in dialogues with other users or with artificial players modelled by the computer (implementing the principle of mixed-initiative argumentative dialogue) [34]. In playing a dialogue game on Arvina, users interact with the other components of the Argument Web and manipulate the AIFdb online argumentation repository [24] (see also Section 2). Any artificial players can draw their dialogue contributions directly from the standpoints and arguments available in AIFdb.

When playing the Critical Discussion Game on Arvina, the first move consists of the advancement of argumentation by the player who has the Protagonist role, as we saw in the previous section. Due to the definition of this move in the DGDL-specification, both the standpoint and the argument in defence of it are made explicit. The StartingRule in Listing 2 ensures that the proposition that is defended by means of this advancement of argumentation is exactly what is at issue in the discussion (i.e. the standpoint): inspect(in, {p}, Standpoint). If it were possible to defend a completely different standpoint or a somewhat modified version of the original standpoint (say, Britain could disarm instead of Britain should disarm), then the fallacy of ignoratio elenchi might be committed. The Critical Discussion Game prohibits this ignoratio elenchi fallacy from occurring by requiring the

Protagonist to explicitly commit himself to the argument being a defence of the standpoint that is actually at issue.

Another way in which a fallacy might already occur at this junction is the defence of the standpoint through mere repetition. In Section 2, we already discussed *petitio principii* as one of the traditionally recognised fallacies that are not precluded by the Standard Treatment, as a result of its validity in (common) systems of logic. In the Critical Discussion Game, petitio principii is prohibited, again, as part of the StartingRule (see Listing 2). This rule requires the proposition that is put forward as an argument to be different from that which is defended: inspect (!in, {q}, Standpoint}).

Once an argument has been presented by the Protagonist, the Antagonist reacts either by accepting it, or by casting doubt on the acceptability of the content of the argument or its supporting link to the standpoint. The dialogue game is such that the Antagonist always has the opportunity to cast doubt on both of these aspects of the argumentation (thereby assuring that, e.g., no argumentum ad baculum can be employed to restrict the Antagonist's right to criticise). The Protagonist can defend his argumentation by verifying whether the proposition used as argument is part of the common ground, or by invoking an external procedure to test whether the argument actually provides support for the standpoint (i.e., whether the latter can be inferred on the basis of the former). Depending on the Antagonist's reaction to the argumentation, and on the outcome of the Protagonist's subsequent defence, the argumentation and the standpoint are accepted or retracted. If the argumentation, and in particular its justificatory force, is accepted by the Antagonist, she commits herself explicitly to the acceptability and sufficiency of the support that the argument provides to the standpoint at issue. This explicitisation of the supporting link between argument and standpoint, prohibits the fallacy of denying an unexpressed premise [47] in the Critical Discussion Game.

At the concluding stage of the discussion, if the Protagonist has to retract his standpoint because of a failure to argumentatively defend it, the risk of committing the fallacy of argumentum ad ignorantiam looms. A failure to defend the standpoint at issue does not warrant accepting the opposite standpoint, which constitutes an implicit shifting of the burden of proof. Such a fallacious move—particularly appealing to the Antagonist in this case—is prohibited in the Critical Discussion Game by terminating the game once the Protagonist performs

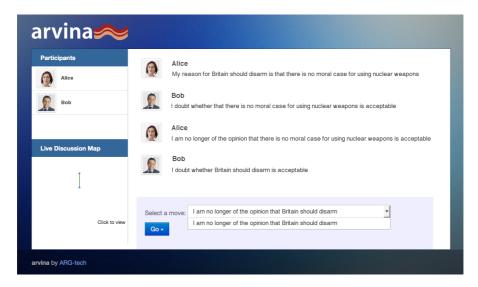


Figure 2. The Critical Discussion Game on Arvina

the RetractProp-move to retract his commitment to the standpoint, shown in Figure 2. At this point, all other commitments are left in place, thereby prohibiting, e.g., the players in our example from suddenly taking on board a commitment to Britain should not disarm, or from starting a negotiation dialogue to this effect.

There are, of course, also fallacies that are at present not captured by the rules of the Critical Discussion Game. Because of the abstraction from argument schemes (see Section 3), for example, any fallacies relating to the incorrect application of specific argument schemes are not addressed in the current (simplified) version of the Critical Discussion Game. This means that, among others, hasty generalisation, false analogy, post hoc ergo proper hoc and ad populum will only be properly handled by the Critical Discussion Game after it is extended to include argument schemes. Furthermore, because of the exclusion of mixed differences of opinion, and discussions about negative standpoints, fallacies such as evading the burden of proof and straw man are not dealt with at present. Another challenge will be posed by the extension of the Critical Discussion Game to take into account the rhetorical perspective provided by the pragma-dialectical notion of 'strategic manoeuvring' [40]. In terms of strategic manoeuvring, argumentative moves are considered to be situated on a continuum ranging from an unreasonable fallacious extreme to its reasonable counterpart — for example, ranging from a reasonable personal attack in response to an argument from expert opinion, to a fallacious *ad hominem* when the person is not the object of the dispute. How this gradual classification of reasonable and fallacious argumentative moves can be adequately implemented in a computational interpretation of the pragma-dialectical model, will require further research.

5. Conclusion

As a first venture into the dialogical identification of fallacies as part of the Argument Web, we have developed the Critical Discussion Game. The Critical Discussion Game is specified in terms of the Dialogue Game Description Language, allowing it to be played on the online Arvina user-interface. Through Arvina, players can interact with the AIFdb argumentation repository. By using the Critical Discussion Game, certain fallacies are prohibited when users engage in argumentative dialogues on Arvina.

The Critical Discussion Game is (to our best knowledge) the first executable computational implementation of the pragma-dialectical ideal model of a critical discussion. This normative discussion model forms the foundation for the pragma-dialectical method of analysing and evaluating argumentative discourse. The combination of a set of dialectical rules that prohibits the occurrence of fallacies and a pragmatic account of the communicative aspects of argumentation, makes the pragma-dialectical model a good starting point for the computational modelling of argumentative dialogue and the automated evaluation of argumentative discourse for fallacies.

The ideal model of a critical discussion was, however, not developed with such a computational application in mind. Therefore, some preparatory steps are required. A first step consists of the formalisation of critical discussion [23, 52]. A second step has been taken in the present paper by computationally implementing the formalised discussion model

¹² The formalisations of the pragma-dialectical ideal model of a critical discussion by Krabbe [23] and Visser [52] are formal in the sense we pursue here, but are not computationally implemented. Starmans [37] and Secades [33] have approached the model of a critical discussion from a computational perspective, but their contributions do not constitute, respectively, an actual formalisation of the model, and an implemented dialogue system.

in terms of DGDL. In this step, we have interpreted the normative rules of the pragma-dialectical ideal model of a critical discussion as prescriptive rules regulating the conduct of actual discussants in the construction of argumentative discourse. This interpretative transformation causes a blurring of the divide between normative ideal (the pragma-dialectical model of a critical discussion) and empirical realisation (the computationally implemented Critical Discussion Game); in principle allowing discussants to actualise an ideal critical discussion, practically behaving exactly in accordance with the normative rules. It is still to be seen whether this correspondence between actual and ideal—between practice and norm—can be maintained when the implemented Critical Discussion Game is extended to cover the full breadth of the ideal model (especially when the specific pre-conditions that are associated with particular argumentative contexts are also taken into account [40]).

In addition to being the first computational implementation of the pragma-dialectical discussion model in a dialogue system, the Critical Discussion Game is one of the first attempts at implementing a comprehensive set of fallacy-prohibiting norms for reasonable argumentative dialogue in a computational system. The goal is to use the DGDL-implementation that has been presented here for the construction of argumentation, as the foundation for the computational analysis and evaluation of argumentation—mirroring the role that the ideal model of a critical discussion plays in the analysis and evaluation within the pragma-dialectical approach (see [52]).

Acknowledgements. The work reported in this paper has been supported in part by the EPSRC in the UK under grant EP/N014871/1 and in part by the Polish National Science Centre under grant 2015/18/M/HS1/00620. We would like to thank Frans van Eemeren, Francisca Snoeck Henkemans, Bart Verheij, and John Lawrence for their contributions to the research the present paper builds on.

References

- [1] Baroni, P., T. F. Gordon, T. Scheffler, and M. Stede, Computational Models of Argument: Proceedings of COMMA 2016, IOS Press, 2016.
- [2] Barth, E.M., Evaluaties. Rede uitgesproken bij de aanvaarding van het ambt van gewoon lector in de logica met inbegrip van haar geschiedenis en de wijsbegeerte van de logica in haar relatie tot de wijsbegeerte in het

- algemeen aan de Rijksuniversiteit te Utrecht op vrijdag 2 juni 1972 (Evaluations. Address given at the assumption of duties as regular lecturer of logic including its history and philosophy of logic in relation to philosophy in general at the University of Utrecht on Friday, 2 June 1972), Assen: van Gorcum, 1972.
- [3] Barth, E.M., and E.C.W. Krabbe, From Axiom to Dialogue. A Philosophical Study of Logics and Argumentation, de Gruyter, 1982.
- [4] Bex, F., J. Lawrence, and C. Reed, "Generalising argument dialogue with the dialogue game execution platform", page 141–152 in *In Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, IOS Press, 2014.
- [5] Bex, F., J. Lawrence, M. Snaith, and C. A. Reed, "Implementing the argument web", Communications of the ACM 56, 10 (2013): 66–73. DOI: 10.1145/2500891
- [6] Biro, J., and H. Siegel, "In defence of the objective epistemic approach to argumentation", *Informal Logic* 26, 1 (2006): 91–102.
- [7] Budzynska, K., and C. Reed, "Whence inference", University of Dundee, technical report, 2011.
- [8] Chesñevar, C., J. Mcginnis, S. Modgil, I. Rahwan, G. Simari, M. South, G. Vreeswijk, and S. Willmott, "Towards an argument interchange format", Knowledge Engineering Review 21, 4 (2007): 293–316. DOI: 10.1017/S0269888906001044
- [9] Dung, P.M., "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games", *Artificial Intelligence* 77, 2 (1995): 321–357. DOI: 10.1016/0004-3702(94)00041-X
- [10] Gordon, T.F., H. Prakken, and D. Walton, "The carneades model of argument and burden of proof", Artificial Intelligence 171, 10–15 (2007): 875–896. DOI: 10.1016/j.artint.2007.04.010
- [11] Grice, H.P., "Logic and conversation", pages 41–58 in P. Cole and J.L. Morgan (eds.), Syntax and Semantics, vol. 3, "Speech acts", Academic, 1975.
- [12] Hamblin, C. L., Fallacies, Methuen, 1970.
- [13] Hansen, H., "Fallacies", in E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy, Metaphysics Research Lab, Stanford University, Fall 2017 edition. https://plato.stanford.edu/archives/fall2017/entries/fallacies/
- [14] Hintikka, J., "The logic of information-seeking dialogues. A model", pages 342–377 in W. Becker and W. K. Essler (eds.), Konzepte der Dialektik (Concepts of dialectic), Vittorio Klostermann., 1981.
- [15] Hintikka, J., "The fallacy of fallacies", Argumentation 1, 3 (1987): 211–238. DOI: 10.1007/BF00136775

- [16] Janier, M., M. Snaith, K. Budzynska, J. Lawrence, and C. Reed. "A system for dispute mediation: The mediation dialogue game", pages 351–358 in T. Scheffler, M. Stede, P. Baroni, and T.F. Gordon (eds.), Computational Models of Argument: Proceedings of COMMA 2016, 2016.
- [17] Janier, M., J. Lawrence, and C. Reed, "Ova+: An argument analysis interface", pages 463–464 in S. Parsons, N. Oren, C. Reed, and F. Cerutti (eds.), Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014), Pitlochry, IOS Press, 2014.
- [18] Johnson, R. H., and J. A. Blair, Logical Self-Defence, McGraw-Hill Ryerson, 1977.
- [19] Johnson, R. H., and J. A. Blair, "Contexts of informal reasoning: Commentary", pages 131–150 in D. N. Perkins J. F. Voss, and J. W. Segal (eds.), Informal Reasoning and Education, Lawrence Erlbaum Associates, 1991.
- [20] Koszowy, M., and D. Walton, "Profiles of dialogue for repairing faults in arguments from expert opinion", Logic and Logical Philosophy 26, 1 (2017): 79–113. DOI: 10.12775/LLP.2016.014
- [21] Krabbe, E. C. W., "Studies in dialogical logic", Doctoral dissertation, University of Groningen, 1982.
- [22] Krabbe, E. C. W., "So what? profiles for relevance criticism in persuasion dialogues", *Argumentation* 6 (1992): 271–283.
- [23] Krabbe, E. C. W., "The formalization of Critical Discussion", Argumentation 31, 1 (2017): 101–119. DOI: 10.1007/s10503-016-9401-y
- [24] Lawrence, J., M. Snaith, F. Bex, and C. Reed, "Demonstration papers on the aifdb, argublogging, arvina, and toast", Frontiers in Artificial Intelligence and Applications 245 (2012): 511–516.
- [25] Levinson, S. C., Pragmatics, Cambridge University Press Cambridge, 1983.
- [26] Lorenzen, P., and K. Lorenz, Dialogische Logik (Dialogical Logic), Wissenschaftliche Buchgesellschaft, 1978.
- [27] Mackenzie, J. D., "Question-begging in non-cumulative systems", Journal of Philosophical Logic 8, 1 (1979): 117–133. DOI: 10.1007/BF00258422
- [28] Mackenzie, J.D., "What hamblins book fallacies was about", Informal Logic 31, 4 (2011): 262–278.
- [29] Pease, A., J. Lawrence, K. Budzynska, J. Corneli, and C. Reed, "Lakatosstyle collaborative mathematics through dialectical, structured and abstract argumentation", Artificial Intelligence 246 (2017): 181–219. DOI: 10.1016/j.artint.2017.02.006
- [30] Rahwan, I., and G.R. Simari, Argumentation in Artificial Intelligence, Springer, 2009.

- [31] Rescher, N., Dialectics. A Controversy-Oriented Approach to the Theory of Knowledge, State University of New York Press, 1977.
- [32] Searle, J. R., Speech Acts: An Essay in the Philosophy of Language, Cambridge University Press, 1969.
- [33] Secades, A., "A computational model of pragma-dialectics as a tool for its analysis and evaluation", *Informal Logic* 35, 3 (2015): 342–377. DOI: 10.22329/il.v35i3.4121
- [34] Snaith, M., J. Lawrence, and C. Reed, "Mixed initiative argument in public deliberation", pages 2–13 in *Proceedings of Online Deliberation*, Fourth International Conference, OD2010, 2010.
- [35] Snaith, M., and C. Reed, "Toast: Online aspic+ implementation", pages 509–510 in S. Szeider, B. Verheij, and S. Woltran (eds.), Proceedings of the Fourth International Conference on Computational Models of Argument (COMMA 2012), IOS Press, 2012.
- [36] Snoeck Henkemans, A. F., Analyzing Complex Argumentation, SicSat, 1992.
- [37] Starmans, R. J. C. M., "Logic, argument, and commonsense", University of Tilburg, doctoral dissertation, 1996.
- [38] van Eemeren, F. H., "Taalbeheersing, argumentatie en rationaliteit: Een nadere bepaling van het begrip kritische discussiant" (Speech communication, argumentation and rationality: A further specification of the notion 'critical discussant'), pages 56–68 in Taalbeheersing 1978. Verslagen van een symposium gehouden op 13 april 1978 aan de Katholieke Hogeschool te Tilburg, VIOT, TH Twente, 1978.
- [39] van Eemeren, F. H., "The study of argumentation as normative pragmatics", Text. An Interdisciplinary Journal of Language, Discourse & Communication Studies 10, 1/2 (1990): 37–44. DOI: 10.1515/text.1.1990.10.1-2.37
- [40] van Eemeren, F. H., Strategic Maneuvering in Argumentative Discourse. Extending the pragma-dialectical theory of argumentation, John Benjamins, 2010. DOI: 10.1075/aic.2
- [41] van Eemeren, F. H., "The pragma-dialectical theory under discussion", *Argumentation* 26, 4 (2012): 439–457. DOI: 10.1007/s10503-012-9274-7
- [42] van Eemeren, F. H., B. Garssen, E. C. W. Krabbe, A. F. Snoeck Henkemans, B. Verheij, and J. H. M. Wagemans, "Argumentation and artificial intelligence", pages 615–675 in *Handbook of Argumentation Theory*, Springer, 2014. DOI: 10.1007/978-90-481-9473-5_11

- [43] van Eemeren, F. H., B. Garssen, E. C. W. Krabbe, A. F. Snoeck Henkemans, B. Verheij, and J. H. M. Wagemans, *Handbook of Argumentation Theory*, Springer, 2014. DOI: 10.1007/978-90-481-9473-5
- [44] van Eemeren, F. H., B. Garssen, E. C. W. Krabbe, A. F. Snoeck Henkemans, B. Verheij, and J. H. M. Wagemans, "The pragma-dialectical theory of argumentation", pages 517–613 in *Handbook of Argumentation Theory*, Springer, 2014. DOI: 10.1007/978-90-481-9473-5_10
- [45] van Eemeren, F. H., B. Garssen, and B. Meuffels, Fallacies and Judgments of Reasonableness. Empirical research concerning the pragma-dialectical discussion rules, Springer, 2009. DOI: 10.1007/978-90-481-2614-9
- [46] van Eemeren, F. H., and R. Grootendorst, Speech Acts in Argumentative Discussions. A theoretical model for the analysis of discussions directed towards solving conflicts of opinion, Foris, 1984. DOI: 10.1515/9783110846089
- [47] van Eemeren, F. H., and R. Grootendorst, Argumentation, Communication, and Fallacies. A pragma-dialectical perspective, Lawrence Erlbaum Associates, 1992.
- [48] van Eemeren, F. H., and R. Grootendorst, A Systematic Theory of Argumentation. The pragma-dialectical approach, Cambridge University Press, 2004. DOI: 10.1017/CBO9780511616389
- [49] van Eemeren, F. H., P. Houtlosser, and A.'ts F. Snoeck Henkemans, Argumentative Indicators in Discourse. A pragma-dialectical study, Springer, 2007. DOI: 10.1007/978-1-4020-6244-5
- [50] van Gelder, T., "The rationale for rationale", Law, Probability and Risk 6 (2007): 23–42.
- [51] Visser, J., "Speech acts in a dialogue game formalisation of critical discussion", Argumentation 31, 2 (2017): 245–266. DOI: 10.1007/s10503-016-9404-8
- [52] Visser, J. C., "A dialogue game for critical discussion: Groundwork in the formalisation and computerisation of the pragma-dialectical model of argumentation", University of Amsterdam, doctoral dissertation, 2016.
- [53] Visser, J. C., "Towards computer support for pragma-dialectical argumentation analysis", pages 1–18 in L. Benacquista and P. Bondy (eds.), Argumentation, Objectivity and Bias. Proceedings of the 11th International Conference of the Ontario Society for the Study of Argumentation (OSSA), OSSA, 2016.
- [54] Walton, D., Logical Dialogue-Games and Fallacies, University Press of America, 1984.
- [55] Walton, D. N., Question-Reply Argumentation, Greenwood Press, 1989.
- [56] Walton, D. N., A Pragmatic Theory of Fallacy, Studies in rhetoric and communication, University of Alabama Press, 1995.

- [57] Wells, S., and C.A. Reed, "A domain specific language for describing diverse systems of dialogue", *Journal of Applied Logic* 10, 4 (2012): 309–329. DOI: 10.1016/j.jal.2012.09.001
- [58] Woods, J., and D. Walton, "Arresting circles in formal dialogues", Journal of Philosophical Logic 7, 1 (1978): 73–90. DOI: 10.1007/BF00245921
- [59] Woods, J. H., The Death of Argument: Fallacies in Agent Based Reasoning, Applied Logic Series, Springer Netherlands, 2013. DOI: 10.1007/978-1-4020-2712-3
- [60] Woods, J. H., and D. Walton, Fallacies: Selected Papers 1972–1982, Foris de Gruyter, 1989.

JACKY VISSER¹, KATARZYNA BUDZYNSKA^{2,1}, and CHRIS REED^{1,2} j.visser@dundee.ac.uk; k.budzynska@dundee.ac.uk; c.a.reed@dundee.ac.uk

Centre for Argument Technology

¹ University of Dundee, DD1 4HN, United Kingdom

Polish Academy of Sciences, Nowy Świat 72, 00-330 Warsaw, Poland www.arg.tech