

A Two-Phase Dialogue Game for Skeptical Preferred Semantics

Zohreh Shams¹ and Nir Oren²

¹ University of Bath, Dept. of Computer Science, UK,
z.shams@bath.ac.uk

² University of Aberdeen, Dept. of Computing Science, UK,
n.oren@abdn.ac.uk

Abstract. In this paper we propose a labelling based dialogue game for determining whether a single argument within a Dung argumentation framework is skeptically preferred. Our game consists of two phases, and determines the membership of a single argument within the extension, assuming optimal play by dialogue participants. In the first phase, one player attempts to advance arguments to construct an extension not containing the argument under consideration, while the second phase verifies that the extension is indeed a preferred one. Correctness within this basic game requires perfect play by both players, and we therefore also introduce an overarching game to overcome this limitation.

1 Introduction

It has been argued that proof dialogues, while providing equivalent results to standard argumentation semantics, can decrease the gap between intuitive and formal accounts of argumentation [1, 11–13], and have been used in human-computer interactions to aid understanding [16, 6, 4]. While the credulous acceptance problem under preferred semantics has been modelled using dialogue games in the past [16, 7, 12, 3], the skeptical preferred semantics has received less attention.

In this paper, we propose a dialogue game for skeptical preferred acceptance that is similar to [16, 10], but (i) differs from [16] in that it is not restricted to cases when the preferred and stable semantics coincide; (ii) differs to approaches such as [12, 8] in that it does not use a meta-dialogue based approach; and (iii) also differs from existing approaches in that it uses argument labellings within the dialogue. Moreover, we believe that this dialogue is more intuitive than [8, 10, 12, 16]. The principal aim of our dialogue is to facilitate explanation as to why an argument is — or is not — skeptically preferred to a human user (similar to the work of Caminada and Podlaskowski [5]).

Our dialogue utilises two phases. In the first, one participant (the opponent) identifies an extension in which the argument under discussion is not present. In the second phase, the other participant (the proponent) attempts to prove that the opponent has (in some loose sense) cheated — that the extension advanced is not a preferred extension. Under perfect play, this dialogue will identify whether an argument is skeptically preferred or not. In the presence of imperfect play, we extend this basic game to an overarching one, which allows the two phases to repeat until both parties are satisfied as to the presence or absence of the argument within all extensions.

Next, we introduce the argumentation system and labelling based semantics. Section 3 describes our dialogue. Note that a longer version of this paper containing proofs of our results, extensions of the basic dialogue, and additional details has appeared as a technical report [14].

2 Preliminaries

We begin with basic concepts from argumentation theory. Note that throughout this paper, we consider only finite argumentation frameworks.

Definition 1 (Argumentation Framework [9]). *An argumentation framework is a pair $AF = (Arg, Def)$, where Arg is a finite set of arguments and Def is a defeat relation between arguments: $Def \subseteq Arg \times Arg$.*

Argumentation semantics focus on arguments that are justified in an argumentation framework, whereas argument labellings (c.f., [15]) consider the status of all arguments. In Caminada's approach [2], an (partial) argument labelling is described as a function $\mathcal{L} : Arg \rightarrow \{in, out, undec\}$ such that $in(\mathcal{L}) = \{a \in Arg \text{ s.t. } \mathcal{L}(a) = in\}$, $out(\mathcal{L}) = \{a \in Arg \text{ s.t. } \mathcal{L}(a) = out\}$, and $undec(\mathcal{L}) = \{A \in Arg \text{ s.t. } \mathcal{L}(a) = undec\}$. Thus, a labelling may be presented as a triple of the form $(in(\mathcal{L}), out(\mathcal{L}), undec(\mathcal{L}))$. An equivalence exists between those arguments labelled *in* according to specific labelling procedures, and the various standard argumentation semantics. To define labelling procedures corresponding to the preferred extensions, we need to first recall the definition of legal and complete labellings.

Definition 2 (Labellings [12],[2]). *Let \mathcal{L} be a labelling for $AF = (Arg, Def)$. An argument belong to Arg is legally:*

- *in iff all its defeaters are labelled out;*
- *out iff there is at least one of its defeaters that is labelled in;*
- *undec iff not all of its defeaters are labelled out and there is none of its defeaters that is in.*

An argument is labelled illegally iff its label is not legal. We say that a labelling is legal if all its arguments are labelled legally.

A complete labelling is a non-partial labelling without arguments that are illegally in, out, or undec.

A complete labelling is called a preferred labelling iff its set of in-labelled arguments is maximal (with respect to set inclusion); or equivalently, iff its set of out-labelled arguments is maximal (with respect to set inclusion).

Note that an argument can only be labelled undecided if one or more of its defeater are also undecided, and none of its defeaters are labelled in.

For a given argumentation framework, multiple preferred labellings may be found. An argument is skeptically accepted under preferred labellings (and equivalently, semantics) if it is labelled *in* within every preferred labelling. If an argument is labelled *in* within a subset of labellings, then it is credulously accepted under the labelling.

3 A Dialogue Game for skeptical Preferred Semantics

We now describe a dialogical proof procedure for the skeptically preferred semantics. Two players — P and O take part, with P seeking to prove that a single *focal argument* f is skeptically preferred. Our dialogue has two phases, and we describe the legal moves and protocols for each phase.

Intuitively, within the first phase, O aims to find a preferred labelling where f is not *in*. In the second phase, P verifies that this labelling is maximally *in* or *out*. The first phase thus allows O to identify a labelling where the focal argument is not justified, and the second phase verifies that O did not cheat in Phase one.

Note that we need to consider only a subset of arguments in the framework, namely those arguments which directly or indirectly defeat or defend the focal argument¹. We refer to the labelling of such arguments as a *sub-labelling*. Where the context is clear, we may refer to a sub-labelling as a labelling.

Definition 3 (Dialogue Moves). *The following moves are available to the dialogue participants.*

What is — $WI(a)$. *This move is used to request that a label be assigned to a , where $a \in Arg$.*

Claim — $CL(\mathcal{L}(a))$. *This move is used to assign a label $\mathcal{L} \in \{in, out, undec\}$ to a , where $a \in Args$.*

3.1 Phase One

In this phase, O seeks to create a complete sub-labelling of arguments in which the focal argument f is *undec* or *out*. If O fails to construct such a labelling, it loses the game. If on the other hand O succeeds, then the result of this phase is a complete sub-labelling, which is evaluated by P in the second phase to determine whether it is a preferred sub-labelling.

In this phase, P utters only WI moves, while O plays only CL moves. The proponent P initiates the dialogue with the move $WI(f)$, where f is the focal argument whose status is to be determined. Following this, O and P take turns to make an utterance (with O making the second move). A $CL(\mathcal{L}(a))$ move by O provides a labelling for the directly preceding $WI(a)$ move of P , where $\mathcal{L}(a)$ is the label assigned to a . Note that WI moves (and therefore CL moves) cannot be repeated. The Phase one protocol is formally specified as follows.

Definition 4 (Phase One Dialogue). *Let $AF = (Arg, Def)$ be an argumentation framework. A phase one dialogue is a sequence $\Delta_o = [\delta_1, \delta_2, \dots, \delta_n]$ ($n \geq 1$) satisfying the following conditions:*

- *odd moves ($\delta_i, 1 \leq i \leq n, i \in 2\mathbb{Z}^+ + 1$) belong to P and even moves ($\delta_i, 2 \leq i \leq n, i \in 2\mathbb{Z}^+$) belong to O .*

¹ i.e., only those arguments for which there is a directed path according to the defeat relation to the focal argument in the graph generated by the argumentation framework.

- $\delta_1 = WI(f)$, where f is the focal argument.
- each δ_i ($2 \leq i \leq n, i \in 2\mathbb{Z}^+$) is of the form $CL(\mathcal{L}(a))$, where $\delta_{i-1} = WI(a)$.
- each δ_i ($3 \leq i \leq n, i \in 2\mathbb{Z}^+ + 1$) is of form $WI(a)$ s.t. $\exists \delta_j = CL(\mathcal{L}(b))(j < i)$ and $(a, b) \in Def$.
- there exist no two WI moves δ_i and δ_j ($i \neq j$) for which $\delta_i = \delta_j$.

Phase one terminates when no more moves are possible. O loses the game if it

1. utters $CL(in(x))$, where x is the focal argument in the dialogue; or
2. labels an argument in , having previously labelled one of its defeater in or $undec$;
or
3. labels an argument $undec$, having previously labelled one of its defeaters in ; or
4. when no more moves are possible, there is an argument labelled $undec$ for which no defeaters are labelled $undec$; or
5. when no more moves are possible, there is an argument labelled out for which no defeaters are labelled in .

If O does not lose the game when Phase one terminates, Phase two begins. Alternatively, if O loses the game, then P wins. If O does not lose the game during Phase one, then we must wait until Phase two terminates to determine whether O or P win the game.

3.2 Phase Two

Here, P tries to prove that the labelling of Phase one is not preferred. If successful, P wins the game and otherwise O wins. The latter occurs when O explores the consequences of claims made by P .

This phase utilises the same moves as previously. However, now P utters CL moves while O advances WI moves. As in Phase one, the proponent (P) and opponent (O) take turns to advance arguments. P makes the first move by putting forward $CL(in(x))$, where x was labelled $undec$ in the Phase one. O responds to a CL move with a WI move, and P responds to such a move with a CL move.

As in Phase one, the argument of a WI move must be one that defeats the argument labelled by a CL move, with the additional constraint that a CL move may only be made over arguments which were labelled $undec$ in Phase one. The CL move then changes the label of such an argument. Again, WI moves cannot be repeated. Formally, the second phase of the dialogue is described as follows.

Definition 5 (Phase Two Dialogue). Let $AF = (Arg, Def)$ be an argumentation framework and $\Delta_o = [\delta_1, \delta_2, \dots, \delta_n]$ ($n \geq 1$) be a Phase one dialogue in which O did not lose. A Phase two dialogue is a sequence of moves $\Delta_t = [\delta'_1, \delta'_2, \dots, \delta'_m]$ ($m \geq 0$) satisfying the following conditions:

- odd moves ($\delta'_i, 1 \leq i \leq m, i \in 2\mathbb{Z}^+ + 1$) belong to P and even moves ($\delta'_i, 2 \leq i \leq m, i \in 2\mathbb{Z}^+$) belong to O .
- $\delta'_1 = CL(in(a))$ for $a \in Arg$ such that $\exists CL(undec(a)) \in \Delta_o$
- For any $2 \leq i \leq m, i \in 2\mathbb{Z}^+$, δ'_i is a move by O of the form $WI(a)$ where $a, b \in Arg$ and
 - $(a, b) \in Def$

- $CL(\mathcal{L}(b)) \in \Delta_t$; and
- $CL(\text{undec}(a)) \in \Delta_o$ but $CL(\mathcal{L}(a)) \notin \Delta_t$; and
- each δ'_i where $3 \leq i \leq m$, $i \in 2\mathbb{Z}^+ + 1$, is a move by P of form $CL(\mathcal{L}(a))$, where $\delta'_{i-1} = WI(a)$.
- there exists no two WI moves δ'_i and δ'_j ($i \neq j$) while $\delta'_i = \delta'_j$.

The dialogue terminates when no further moves are possible. P wins the game iff it has made at least one move during Phase 2 and the labelling at the end of Phase two is legal. Otherwise, O is the winner.

The second phase requires P to demonstrate that the labelling advanced by O in Phase one is not a preferred labelling. Since such labellings are maximally *in*, P does so by changing the label of an argument labelled *undec* in Phase one to *in*. Once such a change is made, Phase two continues by relabelling the undecided defeaters of the changed argument until no further changes are required or possible. If P is able to perform the relabelling in such a way so that the resultant labelling is legal, then it wins the game as it has shown that the labelling advanced by O in Phase one is not maximally *in*. If P fails in doing this, then O wins the game.

Note that the relabelling in Phase two does not require all undecided arguments to be relabelled. Also, if no undecided arguments exist at the start of Phase two, the game ends immediately, with O winning the game. Finally, note that while P was required to relabel an *undec* argument to *in*, it would be equivalent to require *undec* arguments to be labelled *out*.

Theorem 1. *There is a winning strategy for P (under which they will win all games) iff the focal argument is skeptically preferred. Similarly, there is a strategy for O (under which O will win all games) iff the focal argument is not skeptically preferred.*

This theorem requires perfect play by O and P is required for the dialogue to correctly identify skeptically preferred arguments. However, as we describe in [14], it is possible to introduce a strategy, together with an extended form of the game, which guarantees that the dialogue will be sound and complete even under imperfect play, though at the cost of additional computational complexity (note that a single iteration of the dialogue has complexity linear in the number of arguments). Unsurprisingly, O 's strategy in Phase one involves advancing a preferred labelling, while in Phase two, P should label the focal argument *in* and then proceed to label other *undec* arguments appropriately. Since pursuing such a strategy may be computationally infeasible, the extended form of the game (effectively) allows multiple games to take place, exploring alternative labellings until both parties are satisfied as to the outcome of the game.

4 Conclusions

In this short paper we introduced a dialogue game for the skeptical preferred semantics which exploits argument labellings. The basic game requires perfect play by both opponents, but is useful in human-computer interaction settings where argument status is explained by the computer to a human (c.f., [4]). As future work, we intend to investigate whether our approach can be more generally applied to any complete-based skeptical semantics (e.g., skeptical stable).

References

1. M. Caminada. Dialogues and hy-arguments. In Delgrande and Schaub, editors. *10th International Workshop on Non-Monotonic Reasoning*, 94–99, 2004.
2. M. Caminada. On the issue of reinstatement in argumentation. In M. Fisher, W. van der Hoek, B. Konev, and A. Lisitsa, editors, *10th European Conference on Logics in Artificial Intelligence*, 4160: 111–123, 2006.
3. M. Caminada, W. Dvořák, and S. Vesic. Preferred semantics as socratic discussion. *Journal of Logic and Computation*, 2014.
4. M. Caminada, R. Kutlak, N. Oren, and W. W. Vasconcelos. Scrutable plan enactment via argumentation and natural language generation. In A. L. C. Bazzan, M. N. Huhns, A. Lomuscio, and P. Scerri, editors, *International conference on Autonomous Agents and Multi-Agent Systems*, 1625–1626, 2014.
5. M. Caminada and M. Podlaskowski. Grounded semantics as persuasion dialogue. In B. Verheij, S. Szeider, and S. Woltran, editors, *Computational Models of Argument*, 245: 478–485, 2012.
6. M. Caminada and M. Podlaskowski. User-computer persuasion dialogue for grounded semantics. In *Benelux Conference on Artificial Intelligence*, 343–344, 2012.
7. C. Cayrol, S. Doutre, and J. Mengin. Dialectical proof theories for the credulous preferred semantics of argumentation frameworks. In S. Benferhat and P. Besnard, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, 2143: 668–679, 2001.
8. S. Doutre and J. Mengin. On sceptical vs credulous acceptance for abstract argument systems. In Delgrande and Schaub, editors, *10th International Workshop on Non-Monotonic Reasoning*, 134–139, 2004.
9. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2): 321–358, 1995.
10. P. M. Dung and P. M. Thang. A sound and complete dialectical proof procedure for sceptical preferred argumentation. In *LPNMR-Workshop on Argumentation and Nonmonotonic Reasoning*, 49–63, 2007.
11. H. Jakobovits and D. Vermeir. Dialectic semantics for argumentation frameworks. In *Seventh International Conference on Artificial Intelligence and Law*, 53–62, 1999.
12. S. Modgil and M. Caminada. Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in Artificial Intelligence*, 105–129, 2009.
13. H. Prakken. Combining sceptical epistemic reasoning with credulous practical reasoning. In P. E. Dunne and T. J. M. Bench-Capon, editors, *Computational Models of Argument*, 144: 311–322, 2006.
14. Z. Shams and N. Oren. A labelling based dialogue game for Skeptical Preferred Semantics. Technical Report ABDN-CS-2016-02 <http://homepages.abdn.ac.uk/n.oren/pages/abdn-CS2016-02.pdf>
15. B. Verheij. Two approaches to dialectical argumentation: Admissible sets and argumentation stages. In *International Conference on Formal and Applied Practical Reasoning*, 357–368, 1996.
16. G. Vreeswijk and H. Prakken. Credulous and sceptical argument games for preferred semantics. In M. Ojeda-Aciego, I. P. de Guzmán, G. Brewka, and L. M. Pereira, editors, *Logics in Artificial Intelligence, European Workshop*, 1919: 239–253, 2000.