Artificial Intelligence 216 (2014) 20–54

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

A general framework for sound assumption-based argumentation dialogues

Xiuyi Fan, Francesca Toni

Imperial College London, United Kingdom

ARTICLE INFO

Article history:
Received 6 February 2014
Received in revised form 6 June 2014
Accepted 13 June 2014
Available online 19 June 2014

Keywords:
Argumentation
Dialogue
Agents
Semantics

ABSTRACT

We propose a formal model for argumentation-based dialogues between agents, using assumption-based argumentation (ABA) as the underlying argumentation framework. Thus, the dialogues amount to conducting an argumentation process in ABA. The model is given in terms of ABA-specific utterances, debate trees and forests implicitly built during and drawn from dialogues, legal-move functions (amounting to protocols) and outcome functions. The model is generic in that it is not restricted to any specific dialogue types and can be used to support a wide range thereof. We prove a formal connection between dialogues and three well-known argumentation semantics (i.e. grounded, admissible and ideal extensions), by giving soundness results for our dialogue models with respect to these semantics. Thus, our dialogues can be seen as a distributed mechanism for successfully determining acceptability of claims (with respect to the semantics considered), while constructing argumentation frameworks and arguments for these claims.

1. Introduction

Argumentation-based dialogue systems have attracted considerable research interest in recent years (e.g. see [33,27]), largely due to the need for agents to communicate and agree in multi-agent systems. Indeed, argumentation is a powerful reasoning abstraction where conflicting positions or opinions are evaluated against one another in order to resolve conflicts. Argumentation has been used quite extensively in AI in the last two decades to support a number of applications and address a number of problems (e.g. see [4,5,38] for an overview). To support this line of research, several argumentation frameworks have been proposed through the years, including [11,7] and many more (see for example [1] for a recent overview of some approaches). The modern study of formal dialogue systems for argumentation can be deemed to have started with Charles Hamblin's work [24]. The topic was initially studied within philosophical logic and argumentation theory [26,43]. Subsequently, researchers from the field of AI & law [23,31] and multi-agent systems [3,30] have looked into dialogue systems as well.

This paper presents a two-agent argumentation-based dialogue model, using Assumption-based Argumentation (ABA) [7,12,40,42] for the representation of arguments and attacks, and for determining "success" of dialogues. ABA is well-suited as a foundation for argumentation-based dialogues for a number of reasons. It is a general purpose argumentation framework with several applications (e.g. see [12,40,42]) including applications requiring dialogues between agents [20]. It is a structured argumentation frameworks, so that a dialogue model based on it can allow the collective construction of arguments and attacks, and a distributed evaluation of "success", rather than forcing, as when using abstract argumentation [11], for example as in [36], that arguments and attacks are determined and/or constructed individually by agents or collectively but prior to dialogues. At the same time, it is an instance of abstract argumentation [15,40] and it admits abstract argumentation as an instance [40], thus allowing our dialogue model to accommodate, as a special case, the communication

http://dx.doi.org/10.1016/j.artint.2014.06.001
0004-3702/

and evaluation of abstract arguments as well. There are several other structured argumentation frameworks available as a basis for argumentation-based dialogues, notably logic-based argumentation [5], DeLP [22] and, more recently, ASPIC+ [34] (see [1] for a recent survey of all these structured argumentation frameworks). Of these, only ASPIC+ is an instance of abstract argumentation and allows abstract argumentation dialogues to be generated, if required by applications. ASPIC+ is a generalisation of ABA and has been designed to admit ABA as an instance [34]. Thus, our ABA-based dialogue model can also be seen as a dialogue model based on ASPIC+, and extending the functionalities and properties of dialogue models based on precursors of ASPIC+ (e.g. [32,33], as we will discuss later). Another essential feature of ABA, for the purposes of this article, is that it is equipped with provably correct computational mechanisms with respect to several semantics [14,15,41]. We rely upon aspects of these mechanisms, as well as their soundness, in order to prove our formal soundness results.

An ABA framework consists of rules, assumptions, and contraries, specified in a logical language. Informally, rules and assumptions form deductions (arguments); contraries of assumptions provide means of specifying counter-arguments (attacks) against arguments. Within an ABA framework, sets of arguments are deemed "acceptable" if they fulfil certain properties, e.g., under the semantics of admissible extensions [7,12], a set of arguments does not attack itself and attacks all arguments that attack it. Then, claims are deemed "acceptable" if they are supported by (are conclusions of) arguments that belong to "acceptable" sets.

Our dialogue model makes use of the same building blocks as ABA, in that a dialogue is composed of utterances whose *content* may be a rule, an assumption, a contrary, or a claim whose "acceptability" needs to be ascertained. In addition, the content of utterances may be a pass, amounting to the agent contributing no information to the dialogue at the time of the utterance. Dialogues start with an agent putting forward a claim. Our dialogue model is generic in that it does not focus on any particular dialogue type, e.g. information seeking, persuasion or negotiation [43], but can be used to support several such dialogue types [17–20].

Through dialogues, the participating agents construct a "joint knowledge base" by pooling all disclosed information to form an *ABA framework*. The ABA framework drawn from a dialogue δ, referred to as $\mathcal{F}_\delta$, contains all information that the two agents have uttered in the dialogue and gives the context for examining the acceptability of the claim of the dialogue. Conceptually, a dialogue is "successful" if its claim is "acceptable" in $\mathcal{F}_\delta$. Note that the claim of a dialogue may be a belief, and acceptability thereof an indication that the agents may legitimately uphold the belief, or a course of actions, and acceptability thereof an indication that the agents may legitimately choose to adhere to it. Indeed, acceptability has so far shown to be an important criterion for assessing the outcome of various types of dialogues [17–20], and thus "successful" dialogues can be seen as building blocks of a widely deployable framework for distributed interactions in multi-agent systems. We focus here on three forms of "acceptability" and "success", with respect to three well-known argumentation semantics.

Rather than checking "success" retrospectively, we define *legal-move functions* guaranteed to generate "successful" dialogues if a limited form of retrospective checking by means of *outcome functions* succeeds. Given a dialogue, a legal-move function returns a set of allowed utterances that can be uttered to extend the dialogue. Legal-move functions can thus be viewed as dialogue protocols. Outcome functions are mappings from dialogues to true/false. Given a dialogue, an outcome function returns true if a certain property holds for that dialogue.

In summary, the main contributions of this work are: (1) a generic formal model for ABA-based dialogues; and (2) the link between this model and standard argumentation semantics (of grounded, admissible and ideal extensions) to define success of these dialogues. We focus on these three semantics as they allow to capture general forms of credulous reasoning (admissible) and two well-understood forms of sceptical reasoning (grounded and ideal), and are thus suitable for a wide range of problems. Our soundness results are obtained by mapping the *debate tree/debate forest* generated from a dialogue onto an *abstract dispute tree* [14] that is known to sanction the "acceptability" of the claim [14,15]. These debate tree/forest can be seen as a *commitment store* [43] holding information that agents disclose and share using the dialogue.

The paper generalises and extends the initial proposal of ABA-based dialogues in [16] in several ways. Firstly, this paper shows soundness results with respect to grounded, admissible and ideal extensions, rather than just admissible extensions as in [16]. Secondly, [16] uses "dialectical trees", which are mapped onto the *concrete dispute trees* of [14] whereas this work uses a new notion of *debate trees* (see Definition 8.1), which are mapped onto the *abstract dispute trees* of [14], directly allowing to use soundness results from [15]. Moreover, [16] defines dialectical trees constructively, whereas this work defines debate trees declaratively, allowing to prove some novel results (e.g. Lemma 11.1). Thirdly, in this paper we define *debate forests* and use them to study *unrestricted* dialogues, completely absent from [16], which studies *focused* dialogues only.

The article is structured as follows. Section 2 presents background on ABA. Section 3 sets the foundation of our dialogue framework and Section 4 introduces generic notions of legal-move and outcome functions. Section 5 defines specific kinds of these functions to generate special kinds of dialogues, guaranteed to draw ABA frameworks. Section 6 defines the three notions of "successful" dialogue we are after, in a non-constructive way. Section 7 starts refining the dialogue framework by introducing new legal-move and outcome functions that enforce core properties of "successful" dialogues, constructively. Section 8 presents *debate trees* that are then used to define legal-move and outcome functions, in Section 9, allowing to construct "successful" dialogues and prove our soundness results in Sections 10 and 11, for *focused* and *unrestricted* dialogues, respectively. Section 12 discusses related works and Section 13 concludes.

The proposed dialogue model relies upon several notions and formal definitions: the most important amongst these are summarised in a glossary in Appendix A, to aid readability.