Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Abstract—Recent work has explored the use of computational argumentation-based dialogues as a mechanism to support two-party human-robot interaction and collaborative decision making. Many real-world application scenarios involve interaction between multiple humans and multiple robots—not just two parties, i.e., not just one human and one robot. In this paper, a range of social issues are considered that arise in multi-party human-robot situations. These include issues of trust, privacy and ethics with respect to sharing information and modeling the beliefs of others. Using computational argumentation and argumentation-based dialogues can provide a sound basis for addressing the mitigating circumstances presented by these issues, and for reasoning under the uncertainties that these issues present. Here, a running example is developed, along with an approach for addressing the issues that is based on the application of computational argumentation strategies. A set of challenges are identified, and some directions for future research are proposed.

I. INTRODUCTION

While in the past machines waited passively for our commands, we are moving quickly to a future where humans and machines work in partnership, where machines are proactive and guide humans’ activities [1]. These machines must be able to justify their choices and explain why a particular course of action is appropriate. Humans must be able to challenge these justifications and provide input into a machine’s decision-making process—and vice versa. As well, machines need to be able to take the initiative [2], [3] and promote new ideas or suggest new tasks or actions, without prompting from humans. Computational Argumentation theory is a natural choice to support such interactions: a key benefit of computational argumentation is its potential to act as a bridge to allow humans and machines to offer input into one another’s reasoning [4], since it formally characterises both aspects of human decision-making [5] and logical reasoning [6].

Computational argumentation [7] is a well-founded, logic-based methodology which had roots in philosophy long before being applied to the multi-agent systems (MAS) world. Argumentation provides a structured framework for reasoning in which participants state conclusions and provide evidence that supports those conclusions. Argumentation-based dialogue [8], [9], [10] is prevalent in MAS as an extension to classic negotiation. In a negotiation, agents can offer alternative claims back and forth, such as haggling over the price of a good. But in a negotiation framework, agents cannot explain their positions or exchange reasons for why or how they arrived at their conclusions. In contrast, argumentation provides the facility to do exactly that. We believe that argumentation-based exchange provides a powerful means for humans to interact with agents.

In this paper, we are particularly interested in interactions between embodied agents—robots—and humans. Previous work has proposed the use of argumentation-based dialogue to support interaction between a single robot and a single human [11]; here, we consider the issues involved in moving to a multi-party application domain, where there may be more than one robot and/or more than one human. Several of the issues we consider are also applicable to more general multi-party dialogue situations (i.e., with autonomous agents, either virtual or embodied, as in physical robots, and/or humans), and have been considered elsewhere both in the multi-agent literature (e.g., [12], [13]) and in the linguistics literature (e.g., [14], [15]). We aim to identify some of the key challenges that are especially relevant to a multi-party human-robot setting and identify some potential directions of future research to address these challenges.

In the following section, we give a motivating example of an interaction involving two robots and two humans, which we use throughout the paper to illustrate the issues under discussion. Then in Section III, we give a brief review of work on two-party argumentation-based dialogue, focussing particularly on a recent proposal to support dialogues between a single human and a single robot [11]. In Section IV, we consider some of the key challenges that must be addressed in order to support multi-party argumentation-based dialogue, particularly those between multiple humans and multiple robots. In Section V, we briefly highlight areas of related work and summarise the primary issues we have identified for applying argumentation-based dialogue to support multi-party human-robot interaction. Finally, we outline directions for future research.

II. MOTIVATING EXAMPLE

Consider an elderly woman, called “Ellie”, who has restricted mobility and lives at home, supported by a healthcare worker (“Sam”) who visits her daily. Ellie lives with two robots, which support her everyday activities, particularly when Sam is not present. One robot is a wheelchair robot, capable of navigating around Ellie’s house; the other is a mobile robot equipped with an arm and a gripper, capable of picking up objects, as well as navigation. For clarity, we’ll refer to the wheelchair robot as “Willie” and the manipulator robot as “Manny”. Both robots possess natural language generation and processing capabilities and are able
to display different degrees of autonomy as appropriate to
the situation\textsuperscript{1}.

Ellie has some pills that she must take twice a day, but she
cannot remember where they are. When Sam arrives, Ellie
tells him that she has checked the kitchen, the bedroom and
the bathroom, since those are the only rooms that she has
visited today. Here, Willie interjects to remind Ellie that she
watched television in the living room this morning, which
Ellie disputes. It is unclear whether Ellie does not remember
being in the living room or if she does not want Sam to
know that she spent some idle time watching television. Sam
believes that Willie’s information is reliable, but he knows
that Ellie does not like to appear idle or infirm, and he does
not want to upset her by questioning her remark.

Sam suggests to Ellie that they repeat her search, ex-
plaining that he often finds things in places he has already
searched, and suggests they visit the bedroom first, followed
by the bathroom and then the kitchen. Ellie agrees and they
start their search. Meanwhile, Manny takes the initiative to
visit the living room and search for the pills. Manny then
meets Sam and Ellie in the kitchen, having found the pills in
the living room and retrieved them. Manny discreetly gives
the pills to Sam, and leaves it to Sam to decide how to
communicate to Ellie that the pills have been located.

This example highlights several social issues, specifically
trust, privacy and ethics. Are Sam, Manny and Willie col-
lectively displaying lack of trust in Ellie’s judgement by
searching for the pills in locations where Ellie told them
she already looked? Is Willie revealing private information
about Ellie by telling Sam that Ellie spent time watching
television? Is it unethical for Sam to keep Ellie from knowing
that Manny found the pills?

III. TWO-PARTY ARGUMENTATION-BASED DIALOGUES

Computational argumentation is a well-studied form of
reasoning that explicitly identifies both the justifications for
any claim that is made by a particular argument, called its
support, and any conflicts that exist between arguments [7],
[16]. One can represent a set of arguments and the conflicts,
or attacks, between them as a directed graph (referred to
as an argumentation framework); such a graph can then
be evaluated according to one of a range of argumentation
semantics, in order to determine which arguments it is
coherent to accept [6]. These semantics are based on the
intuitive principles that it is not rational to accept any two
arguments that are in conflict with each other, and that an
argument which is attacked can only be accepted if all
of its attacking arguments are themselves attacked by an
accepted argument. Argumentation thus provides an intuitive
mechanism for dealing with inconsistent, uncertain and in-
complete knowledge, and, through structured argumentation-
based dialogues, supports intelligent agents in exchanging
beliefs, reaching agreements, making decisions and resolving
conflicts of opinion [4].

Much of the existing work on formal argumentation-based
dialogues comes from the multi-agent literature and defines
two-party dialogue systems for achievement of a particular
dialogue goal (e.g., to resolve some conflict of opinion [8]
or to agree on some action [17]). These are typically defined
in terms of a set of communicative acts (or moves) that the
participants can make, a set of rules that determines which
moves it is permissible to make at any point (the dialogue
protocol), rules that determine the effect of making a move
(typically defined in terms of the participants’ dialogue
commitments), and rules that determine when the dialogue
terminates and what the outcome of the dialogue is [18]. This
imposed dialogue structure goes some way towards reducing
the complexity of how to determine which utterances to make
and how to interpret received utterances, in comparison with
completely flexible natural dialogues.

Achievement of an agent’s dialogue goal typically depends
on both the arguments that the agent chooses to make
during the dialogue, determined by its strategy, and the
arguments asserted by the other participant, its interlocutor.
The strategising agent thus has the difficult problem of
having to consider not only which arguments to assert but
also the possible responses of its interlocutor [19]. Recent
works have proposed methods for determining an agent’s
dialogue strategy that take into account the strategiser’s
uncertain model of the other participant [20], [21], [22],
while others have considered how such a model might be
developed based on dialogue experience [23], [24].

Sklar & Azhar have recently developed a framework for
argumentation-based dialogues between a single robot and
a single human [11]. In this framework, a robot \( R \) holds
a set of beliefs (denoted \( R.\Sigma \)) which it uses to construct
arguments. As part of this set of beliefs, the robot maintains
a model of what it believes to be the beliefs of the human
it is interacting with; a robot \( R \)’s model of a human \( H \)
is denoted \( R.\Gamma(H) \). The robot uses its beliefs and its model
of the human to determine which type of dialogue they
should engage in (chosen from the influential Walton &
Krabbe dialogue taxonomy [25]); if the robot believes there
to be disagreement over a particular belief, they may enter
into a persuasion dialogue with the human where they can
exchange arguments to resolve this disagreement; if the robot
believes that the human knows the status of some belief
that it is unaware of (or vice-versa), they may enter into an
information-seeking dialogue with the aim of sharing this
knowledge; finally, if the robot believes that neither itself
nor the human is aware of the status of some belief, they
may enter into an inquiry dialogue where they try to jointly
construct an argument that supports the belief.

Using the Treasure Hunt Game (THG) domain, the dia-
logue framework of [11] has been implemented and evalu-
ated in [26]. In this domain, the robot is physically located
in the environment, which consists of a number of rooms
connected via a corridor, in each of which there may be
placed a treasure. The robot moves around the environment
and collects sensor data which it shares with the human
(who does not have physical access to the environment).

---

\textsuperscript{1}For the purposes of this example, we assume that these capabilities are
more advanced than the current state-of-the-art.
The human-robot team receives points for correctly identifying treasures and loses points for incorrectly identifying treasures. The robot has only a fixed amount of energy which is depleted through movement, sensing and communication. The human and robot, who interact through argumentation-based dialogues as described above, must work together in order to determine where the robot should look next and how it should get there, and then to decide whether any object present in that room is one of the searched-for treasures. In a user study reported in [26], human subjects played two games with two robots. One robot acted as a peer, collaborating with the human using the computational-argumentation-based dialogue framework described in [11]; the other robot acted as a servant, following the human’s commands without question. The interaction abilities of the first robot are greatly enhanced over traditional supervisory control, because the robot is able to disagree with the human’s action choices or opinions and suggest alternatives. The user study showed that the human-robot teams with the first style of robot performed significantly better than those with the second style of robot, across a range of objective performance measures. As well, in qualitative and survey (ranked) responses, users reported consistently greater trust in the first style of robot and greater confidence in their joint solutions.

IV. MULTI-PARTY ARGUMENT DIALOGUES: CHALLENGES

As discussed in the previous section, the majority of existing work on argumentation-based dialogue focuses on two-party situations. In this section, we consider some of the key challenges of developing frameworks for multi-party argumentation-based dialogue, particularly those that take place between multiple robots and humans in a physical environment.

A. Multiple hearer roles

In addition to the speaker, following Bell [14], we identify four different roles for participants in multi-party dialogues: addressee, auditor, overhearer and eavesdropper. Bell classifies these hearer roles, respectively, according to whether they are addressed (i.e., explicitly identified as a receiver of the message), ratified (i.e., an intended participant in the communication), or known (i.e., the speaker is aware of their receipt of the message) by the speaker. This classification is given in Table I.

<table>
<thead>
<tr>
<th>Known</th>
<th>Ratified</th>
<th>Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addressee</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Auditor</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Overhearer</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Eavesdropper</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

TABLE I

BELL’S CLASSIFICATION OF HEARER ROLES [14, P. 160].

The possible hearer roles depend on the mode of communication. With broadcast messages, there may be multiple addressees, multiple auditors, multiple overhearers and multiple eavesdroppers. If messages are sent peer-to-peer, then a single addressee is intended, and no auditors and or overhearers are possible; though eavesdroppers are possible. (We do not consider overhearers to be possible in the peer-to-peer case because we assume the speaker will only send the message to the intended receivers of the message, i.e., those that are ratified; and so any other participant who receives the message must be an eavesdropper.) When communicating via a forum, multiple addressees, auditors, overhearers and eavesdroppers are possible, however the possibility of overhearers and eavesdroppers depends on whether the forum records and makes public the participants that have accessed the forum.

In our motivating example, messages are broadcast, meaning there are potentially many addressees, auditors, overhearers and eavesdroppers of each utterance. For example, if Willie reminds Ellie that she spent some time watching television in the living room, Ellie is the addressee of this communication, but both Sam and Manny are auditors, since Willie intends them to receive the communication and infer something from it. As we discuss further in the following sections, this multiplicity of hearer roles makes it harder (than in the two-party case) to determine appropriate rules to govern a multi-party dialogue system, harder to decide which utterances to make in such a dialogue and harder to maintain a model of the other participants in the dialogue.

B. Specifying the rules of a multi-party dialogue system

Argumentation-based dialogues are typically specified in terms of axiomatic semantics: the moves that can be made and the rules that determine when moves can be made and what effect making a move has on the dialogical commitments of the participants (e.g., [18]), essentially determining a set of pre-conditions and post-conditions for each type of move. While an agent’s dialogue goal may reference the private belief state of another participant, it is not appropriate for the pre- and post-conditions of a dialogue move to do so, in order that fulfillment of the conditions can be verified. Thus, these conditions are normally specified in terms of public elements of the dialogue (e.g., past moves and dialogical commitments publicly incurred during the dialogue).

In a two-party dialogue, the only hearer role is the addressee and it is reasonable to assume that both participants (speaker and addressee) have the same view of the dialogue history and of the dialogical commitments. In a multi-party dialogue, however, this is likely not to be the case, since participants can play different hearer roles at different points in the dialogue. One might consider, for instance, that an agent does not maintain a representation of parts of the dialogue history for which it were an overhearer2, or that the dialogical commitments known by a participant may similarly depend on their hearer role at the time the

2One reason for filtering the information stored is a practical constraint on an agent being able to manage large amounts of data in real time. A strategy of “remembering to forget” [27] could be implemented to ensure that agents maintain tractable amounts of information.
commitment was incurred. Considering again our motivating example where Willie reminds Ellie that she has spent time in the living room, we might assume that Willie’s intention with this remark is for Manny to accept this claim and subsequently act on it; however, as Manny is only an auditor in this case, perhaps Willie cannot be certain that Manny is aware of Willie’s commitment to this fact.

We should also consider the hearer role of participants when determining the dialogical commitments incurred. Consider the case where the speaker poses a question to a particular addressee; it may then be that the addressee now has a dialogical commitment to answer this question, but that this commitment can also be revoked if an overhearer of the question provides an answer.

This discussion suggests that when specifying the pre- and post-conditions of dialogue moves intended for multi-party dialogues, one must consider more than just a single hearer role. Furthermore, we may need to consider the possibility that a speaker may not be able to determine if certain post-conditions have been applied, for example in the case where a hearer perceives itself as an overhearer but is actually intended to be an auditor.

C. Strategising with multiple hearers in mind

In addition to the issues discussed above regarding the specification of types of multi-party dialogues in terms of the moves that can be made and the dialogical commitments these incur, the design of dialogue strategies for a multi-party setting (i.e., how a participant determines which of the permissible moves to make) also requires us to consider the hearer roles of participants. In their Informative Hypothesis linguistic theory, Clark & Carlson propose that the design of utterances takes into account not only the addressee but also all hearers of the utterance, claiming that speakers “decide how to say what they say on the basis of what they know, believe and suppose that these hearers, in their assigned roles, know, believe, and suppose” [15, p. 342]. Clark & Carlson explain that, as we see in our motivating example, the potential difference in what is called the common ground between a speaker and a hearer in linguistic theories (i.e., their shared knowledge, assumptions and beliefs) makes it possible to convey different things to different hearers with the same utterance.

Thus, when Sam suggests to Ellie that they repeat her search, as an auditor of the communication, Manny is able to make inferences from its common ground with Sam—which includes the beliefs that Ellie does not like to appear idle or infirm, that revealing Ellie’s memory to be at fault may have this effect, that Willie’s memory is more reliable, and that it is important to find the pills. Manny believes that Sam does not believe that they will find the pills by repeating Ellie’s search, but would not object if Manny searched for them in the living room. In order to identify their common ground, Manny must maintain a model of Sam’s beliefs—represented as $\Gamma(Sam)$, as discussed in Section III.

When engaged in multi-party dialogues, we see then that one may desire to achieve different dialogue goals directed at different participants within the same dialogue, and perhaps even with the same utterance. This makes deciding which utterances to make even harder than in the two-party case. This decision is further compounded by the need to maintain and reason with models of each of the other participants, which we discuss further in the following section.

D. Higher order modelling of other parties

When participating in an argumentation-based dialogue, one must choose particular moves to make in order to try to achieve some dialogue goal. Typically, whether a particular utterance will help in the achievement of a goal will depend on the beliefs, and perhaps the goals and intentions, of the hearers. We see this in our motivating example when Sam persuades Ellie to repeat her search by letting her know that he often finds things in places he has already searched for them in. Had Sam simply suggested they repeat Ellie’s search, she would have refused, believing he was suggesting she was forgetful or infirm; however, by making reference to his own shortcomings in searching for things, Sam successfully counters this belief.

As discussed in Section III, recent works on strategy for two-party argumentation-based dialogues [20], [21], [22] have considered the use of an uncertain (probabilistic) model of one’s interlocutor to guide move selection, where that model may represent what the strategist believes its interlocutor’s beliefs to be, what its goals are, and perhaps even a nested model of what it believes to be its interlocutor’s model of itself. In the multi-party setting, not only does one have many potential interlocutors to consider, but also it can be important to consider what each believes about the other potential interlocutors. This is illustrated in our motivating example. In order to recognise Sam’s desire for Manny to search for the pills in the living room, Manny must believe that Sam believes that Ellie does not want to appear infirm and must also believe that Sam finds Willie’s claim that Ellie spent time in the living room to be more reliable than Ellie’s denial of this. Furthermore, Sam must believe that Manny has these beliefs, in order to correctly anticipate that getting Ellie to agree to retrace her search will result in Manny searching for the pills in the living room. We can represent this fourth-order belief—Sam’s belief that Manny believes that Sam believes that Ellie does not want to appear infirm—as:

$$\text{goal}(\text{not infirm}) \in \Gamma(\text{Manny}) \cap (\Gamma(\text{Sam}) \cap \Gamma(\text{Ellie}))$$

This requirement for higher-order beliefs poses several challenges, particularly for resource-bounded robots. One key challenge is how to develop and update a model of the other participants. This is an open problem within the two-party argumentation-based dialogue field—and it is further complicated in the multi-party setting by the need for beliefs about others’ beliefs about others. Different hearer roles must also be considered here, since (as discussed above) one might assume an utterance has a different effect on a hearer according to their role; and these roles may not be known (or static). Determining what depth to model the higher-order
beliefs to and the development of efficient mechanisms for reasoning with such beliefs are further challenges.

V. DISCUSSION

This paper provides a preliminary discussion of issues relevant to multi-party dialogue, applied to human-robot interaction, particularly in situations where there is more than one robot and/or more than one human. The research on multi-party dialogues is vast and interdisciplinary. From the perspective of multi-party dialogues between software agents, related work has examined the environment as an active medium for managing communication [13], [28]; publish/subscribe patterns for selectively sending and receiving messages [29]; blackboard or on-line forum approaches to sending messages to large groups (selective or open) [12], [13]; broadcast mechanisms to send messages to anyone within transmission range [30]; and insect-inspired stagmergic approaches, where agents deposit information in their environment for others to retrieve [31]. Within the computational argumentation field, a dialogue system has been proposed with which multiple agents can resolve disputes regarding the classification of particular cases [32].

From the perspective of multi-party dialogues between humans, relevant literature comes from the psychology and sociology communities, among others. This includes research on audience roles [14] and consideration of targeted speech acts and incidental hearers [15]. From the perspective of combined multi-party human-agent dialogues, early work investigates natural language approaches to interaction in virtual worlds [33]. Relevant work from a multi-party human-robot interaction perspective includes consideration of how humans’ attitudes towards robots changes as the numbers of humans and/or robots involved in the interaction changes [34], and how a data-driven perceptual system that has been trained in a group setting performs in a one-on-one setting and vice-versa [35]. Investigation of factors which impact all multi-party dialogue between humans and machines includes: natural language [36], [37], [38] to enable more fluid interaction; turn-taking [39], [40] to support flexible change in initiative; and multi-modal interaction [41], [42], to incorporate non-verbal communication such as gestures and facial expressions.

In this paper, we have presented a brief overview of two-party argumentation-based dialogue, including a description of a framework specifically intended for human-robot interaction [11], which has been shown experimentally to improve both the outcome of human-robot team performance and the human’s perception of working with the robot [26]. We discussed some of the key challenges that arise when applying argumentation-based dialogue in a multi-party setting, which are particularly prevalent in a multi-party human-robot setting where one can expect messages to often be broadcast, thus increasing the possibility of multiple hearer roles.

A common, but not always realistic, assumption in two-party argumentation-based agent dialogue, is that the knowledge available to the participants is static (a notable exception to this is Walton et al.’s model of a deliberation dialogue that accounts for dynamic knowledge [43]). In order to be practically useful in real-world settings, the specification of multi-party argumentation-based dialogues must allow for changes to the participants’ knowledge about the world. Future work will examine multi-party argumentation-based dialogue in dynamic situations, where the state of the world may change during the dialogue and belief revision is necessary for reasons other than the conclusion of a dialogue.

Degrees of trust can be associated with several different aspects relevant to communication and dialogue. These include: trust in information about the environment (such as a robot’s sensor data), trust in the communication medium (such as wireless network connectivity), trust in one’s model of another participant and trust in a participant as a knowledge source or as an actor or hearer in particular roles. The last two, in particular, can be context dependent. For example, we might trust a robot to pick up the right bottle of pills from our kitchen table, but not necessarily from a crowded shelf in a chemist. Future work will examine the notion of degrees of trust between different participants in multi-party dialogues, associated with varying contexts, initially exploring the aspects listed above.

The immediate next step with this research is to adapt the dialogue framework of Sklar & Azhar [11] in a multi-party implementation. This will require clarification of locution sequences and modification of pre- and post-conditions to include multiple roles of hearers in multi-party situations. The longterm goal is to apply this work in scenarios similar to the motivating example, with physical robots and human subjects. We have recently begun designing studies with colleagues in our university’s School of Nursing, in which we will implement and test our theoretical model with two types of patients. Our goal is to evaluate the effectiveness of argumentation-based multi-party human-robot dialogue for improving patient compliance with treatment recommendations. The first case will focus on patients with diabetic foot ulcers who have been prescribed an offloading brace and the second case will focus on geriatric patients who are participating in fall prevention programmes. With both cases, the aim is to integrate a robot and intelligent sensors as non-invasive participants in a patient’s treatment plan. The robot and the medical team (nurses and clinicians) who oversees the patient’s treatment will collaboratively track the patient’s compliance (with wearing the brace or performing strength-training exercises), and the system will employ argumentation-based dialogue to encourage the patient to adhere to the recommended treatment. The robot and intelligent sensors will work in tandem, modelling the patient’s behaviour and deciding when and how to provide feedback to the medical team, as well as incentives and rewards to the patient.

REFERENCES


