Opportunistic Belief Reconciliation During Distributed Interactions

Paul Martin
School of Informatics
The University of Edinburgh
Edinburgh EH8 9LE, UK
s0199422@sms.ed.ac.uk

David Robertson
School of Informatics
The University of Edinburgh
Edinburgh EH8 9LE, UK
dr@inf.ed.ac.uk

Michael Rovatsos
School of Informatics
The University of Edinburgh
Edinburgh EH8 9LE, UK
mrovatso@inf.ed.ac.uk

ABSTRACT

Successful interaction between autonomous agents is contingent on those agents making decisions consistent with the expectations of their peers — these expectations are based on their beliefs about the current state of the environment in which interaction occurs. Contradictory beliefs lead to unintended and often unjustified outcomes. Given a shared interaction protocol to which all agents agree to adhere, it is possible to identify the constraints upon which the outcome of an interaction rests as it unfolds, and so prior to resolving those constraints, agents can compare and reconcile any relevant expectations by a process of argumentation.

In this paper, we introduce a mechanism by which agents can efficiently articulate their current beliefs in order to influence the resolution by their peers of constraints imposed on a distributed interaction and thus influence its outcome. We understand this as an opportunistic process of belief synchronisation within a restricted argument space, such that all decisions can be said to be admissible given the information that it is practical for agents to share with one another. Thus, we use the distributed knowledge dispersed amongst an agent group to make better decisions in interaction without resorting to more complex, domain-specific protocols.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Coherence and coordination, multiagent systems*

General Terms

Algorithms, Theory

Keywords

Argumentation, multi-agent reasoning, logic-based methods

1. INTRODUCTION

The purpose of interaction between agents within a distributed system is to disseminate information and to assemble new behaviours by choreographing the actions of individual peers. Within an open system however, agents are intrinsically heterogeneous, being of varied provenance and design, and the behaviour of individual peers cannot be predicted without careful observation. In particular, without some kind of framework within which interaction can be controlled and directed (on the volition of those agents participating), it is extremely difficult for an agent to communicate anything but the most primitive intentions to its peers and then derive firm conclusions from any response [14].

One approach is to formalise the institutions within which certain classes of interaction are enacted, defining the protocols [11] to which those agents should adhere. Having specified these social models of interaction, it is then necessary to propagate them in a form which can be executed by amenable peers [12, 13]. This approach reduces interaction to a series of decisions based on the satisfaction of known logical constraints leading to one of a number of monitorable outcomes, but allows agents the autonomy to satisfy those constraints based on their own personal beliefs.

Our concern is with how belief affects the course of interaction, and in particular, with how decisions made on false premises based on limited knowledge of peers can be prevented. Every agent in an interaction has its own beliefs, and each agent chooses to engage in that interaction because it expects certain outcomes to arise from doing so based on those beliefs (*e.g.* an agent expecting its goals to be furthered in exchange for cooperating with a peer). More subtly, when one agent is seen to have satisfied a constraint imposed on interaction, its peers may draw conclusions which do not accurately reflect how the constraint was understood (*e.g.* an agent may expect its partner to account for a particular factor before accepting a risky contract, unaware that its peer is ignorant of this). This can lead to a failure of interaction, or lead to a longer-term miscomprehension which adversely affects future interactions (*e.g.* peers assuming a future commitment from an agent which it itself does not recognise).

To prevent such mistakes, we propose the use of a form of distributed belief maintenance in which assumption-based argumentation [1] is performed on demand as interaction between agents unfolds. Agents refer to the protocol for an interaction and posit what they believe to be satisfactory resolutions of incoming constraints into an *interaction portrayal*; its peers within the interaction can then argue for or against such resolutions, articulating their reasoning as necessary to support their claims. The agent tasked with resolving a given constraint can then interpret the resulting system of arguments [3] in order to make an informed decision, leveraging its peer group's distributed knowledge.

Naturally, there are challenges to surmount. In particular,

Cite as: Opportunistic Belief Reconciliation During Distributed Interactions, Paul Martin, David Robertson and Michael Rovatsos, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. XXX-XXX.

practicality dictates that we limit the scope of argumentation so as to not encumber dialogue with unnecessary details. To this end, we shall demonstrate the use of *argument spaces* to control the generation of arguments and project the (arbitrarily complex) internal arguments used by agents to manage their beliefs onto the (minimally sufficient) shared arguments used to decide the veracity of claims pertinent to agent coordination. A well-formed argument space allows us to *synchronise* the beliefs of agents within the space such that, based on the aggregation of information revealed, we can be confident that any decision made by an agent is justifiable to (if perhaps not favoured by) each one of its peers.

This paper can be split into two parts. Sections 2, 3 and 4 describe the kind of argumentation used by portrayals, and how it relates to the beliefs of agents — this gives us the metrics by which we evaluate our contribution. Sections 5, 6 and 7 describe how portrayals accompany interactions, how they are constructed and maintained, and how they embody the properties by which we evaluate them.

2. ABSTRACT ARGUMENTATION

Basically, a portrayal is a distributed device used to monitor the state of a multi-agent argumentation process, one conducted to determine acceptable resolutions for logical constraints imposed on an agent interaction by some interaction protocol of the type described by [11, 12]. Throughout an interaction, agents posit relevant arguments reflecting their current beliefs into a portrayal and their peers then attempt to reconcile those arguments with their own theories, positing any counter-arguments. This results in a *system of arguments* as described in [3]:

Definition 1. *A* **system of arguments** *is defined by a pair* $(\mathcal{A}, \rightharpoondown)$ *where:* $\mathcal{A}$ *is a finite set of arguments;* $\rightharpoondown$ *is an attack relation between ordered pairs of arguments* $A, B \in \mathcal{A}$ *such that "A attacks B".*

An argument is a statement about the world made to support a particular claim. The nature of an attack $A \rightharpoondown B$ is to assert that if one accepts argument A, then consequently one must reject argument B. Although arguments lack meaning without context, the status of arguments within an argument system can be evaluated independently of their internal structure or provenance by considering them purely in terms of their relationship with one another. For example:

Definition 2. *Given a system of arguments* $(\mathcal{A}, \rightharpoondown)$*, a set of arguments* $S \subseteq \mathcal{A}$ **defends** *any argument* $A \in \mathcal{A}$ *if there is an argument* $B \in S$ *for every attack* $C \rightharpoondown A$ *such that* $B \rightharpoondown C$*.*

Thus, it is possible to identify within a system of arguments particular arguments which collectively describe coherent positions which an agent might adopt when deciding the validity of certain statements:

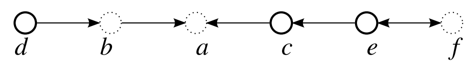
Definition 3. *Given a system of arguments* $(\mathcal{A}, \rightharpoondown)$*, an* **extension** $S \subseteq \mathcal{A}$ *of* $(\mathcal{A}, \rightharpoondown)$ *is a subset of arguments which exhibits particular properties. For instance* [3]*:*

- $\mathcal{S}$ *is* **admissible** *if there are no arguments* $A, B \in \mathcal{S}$ *such that* $A \rightharpoondown B$ *and every argument* $A \in \mathcal{S}$ *is defended by* $\mathcal{S}$*.*
- $\mathcal{S}$ *is* **complete** *if* $\mathcal{S}$ *is admissible and every argument* A *which is defended by* $\mathcal{S}$ *is in* $\mathcal{S}$*.*
- $\mathcal{S}$ *is* **preferred** *if* $\mathcal{S}$ *is maximally complete (under set inclusion).*
- $\mathcal{S}$ *is* **grounded** *if* $\mathcal{S}$ *is minimally complete (under set inclusion). The grounded extension of any system of arguments is unique.*

Given a particular acceptability semantic, a set of arguments $\mathcal{S}$ *is said to be* **acceptable** *if it exhibits the chosen quality.*

Different acceptability semantics (of which there are many [2, 3, 4]) can be more credulous (*e.g.* preferred) or sceptical (*e.g.* grounded). Choosing a particular extension over others is typically based on some heuristic preference (*e.g.* [9]), which itself may be based on (for example) the risks involved in certain interaction outcomes. All measures of acceptability are inherently defeasible — a previously acceptable argument can become untenable with new evidence. When an agent accepts an extension of an argument system, we refer to that agent's *interpretation* of the system.

Example 1. *Consider a system of arguments, illustrated as a graph* $(\mathcal{A}, \rightharpoondown)$ = *(vertices, edges):*



If argument **a** *promotes the claim "Alanna is trustworthy", then the highlighted preferred extension* {**a**, **d**, **e**} *defends that claim, whilst the alternative preferred extension* {**c**, **d**, **f**} *attacks it. Given that* {**d**} *is the grounded extension, only argument* **d** *is sceptically acceptable however.*

Fundamentally, we want to ensure that any decision made by an agent over the course of an interaction augmented by a portrayal is *at least* admissible to all peers (*i.e.* the decision is supported by an admissible extension).

3. DEFEASIBLE REASONING

The arguments which an agent inserts into a portrayal are derived from its beliefs. These beliefs are based on a combination of empirical observation and, where necessary for timely decision-making, assumption. Thus, an agent may make arguments based on assumptions which, whilst perhaps consistent with the evidence available at the time, are not in fact correct. This may be borne out by its interpretation of its peers' arguments, which can force the revision of its beliefs. In this sense, multi-agent argumentation is part of a broader process of belief revision.

The purpose of an *assumption-based argumentation framework* [1] is to provide a logical context for a system of arguments so that it can be used to drive such a process. Arguments are given formal representation and attacks are defined in terms of how the claim of one argument interferes with the support for another.

Definition 4. *An* **argumentation framework** *is described by a tuple* $(\mathcal{L}, \vdash, \Delta)$ *where:*

- $(\mathcal{L}, \vdash)$ *is a deductive framework used to construct arguments, such that:* $\mathcal{L}$ *defines all interpretable sentences such that if a sentence* φ *is interpretable by* $(\mathcal{L}, \vdash)$*, then* $\varphi \in \mathcal{L}$*;* $\vdash$ *describes a monotonic inference procedure used to derive conclusions from premises. If a conclusion* $\varphi \in \mathcal{L}$ *can be inferred from an information source* $\Theta \subseteq \mathcal{L}$ *using* $\vdash$*, then* $\Theta \vdash \varphi$*.*

- Δ *is an* **argument space** *specifying the scope of argumentation, which can be described by a pair* $(\mathcal{H}, \mathcal{F})$ *where: the* **horizon** $\mathcal{H} \subseteq \mathcal{L}$ *defines the set of sentences which can be premises for arguments in* Δ*; the* **focus** $\mathcal{F} \subseteq \mathcal{L}$ *defines the set of sentences which* $(\mathcal{L}, \vdash, \Delta)$ *has been employed to determine.*

Despite argumentation being itself defeasible, arguments are generated within a deductive logical framework as advocated in [8]. This is because the non-monotonic part of argumentation is in the generation and selection of hypotheses with which to support arguments, whilst the arguments themselves are monotonic *assuming* their supporting premises:

DEFINITION 5. *Given a framework* $(\mathcal{L}, \vdash, \Delta)$*, an* **argument** *is a pair* $\langle \Phi, \alpha \rangle$ *where* $\Phi \subseteq \mathcal{L}$ *forms the minimal consistent support for a claim* $\alpha \in \mathcal{L}$ *such that* $\Phi \vdash \alpha$*, there exists no subset* $\Psi \subset \Phi$ *such that* $\Psi \vdash \alpha$*, and* $\langle \Phi, \alpha \rangle \in \Delta$*.*

- *An argument* $\langle \Phi, \alpha \rangle$ *is* **within** Δ *(*i.e. $\langle \Phi, \alpha \rangle \in \Delta$*) if* $\Phi \subseteq \mathcal{H}$ *and either* $\alpha \in \mathcal{F}$ *or there exists an attack* $\langle \Phi, \alpha \rangle \rightharpoondown \langle \Psi, \beta \rangle$ *such that* $\langle \Psi, \beta \rangle \in \Delta$*.*

An argument $\langle \Phi, \alpha \rangle$ **attacks** *another argument* $\langle \Psi, \beta \rangle$ *if and only if there exists a* **contrary relation** (α, γ) *such that* $\{\alpha\} \vdash \neg\gamma$ *and* $\Psi \vdash \gamma$*.*

Since it is not useful to simply generate any and every argument which can be constructed within some logic, an *argument space* defines the hypothesis space in which arguments must be confined. It serves two purposes; it limits the assumptions which an agent can make when constructing arguments, and it is used to evaluate the relevance of arguments. Every argumentation process must necessarily operate within an argument space, though the bounds of that space may be refined throughout that process.

EXAMPLE 2. *Consider the argument* $\langle \{A, B\}, C \rangle$*, where* A *= "Alanna has no history of deception",* B *= "An agent with no history of deception is trustworthy" and* C *= "Alanna is trustworthy". This argument is within the argument space* $\Delta = (\mathcal{H}, \mathcal{F})$ *if* $A, B \in \mathcal{H}$ *and either* $C \in \mathcal{F}$ *or there exists another argument* $\langle \{D, E\}, F \rangle \in \Delta$ *such that* $\{C\} \vdash \neg D$ *(or anything else derivable from* $\{D, E\}$ *such that* $D \wedge E$ *must be false). If* $A \notin \mathcal{H}$ *then* $\langle \{A, B\}, C \rangle \notin \Delta$*, but if* $F, G \in \mathcal{H}$ *(*F *= "Alanna has never reneged on an agreement" and* G *= "An agent which has never reneged on an agreement has no history of deception") and* $\{F, G\} \vdash A$*, then* $\langle \{F, G, B\}, C \rangle$ *is within the argument space instead.*

The system of arguments generated within a portrayal, being a distributed device, must then be generated within a distributed argumentation framework, one which uses a common logic (such as first-order predicate logic) and deduction to generate arguments. Our attention is drawn then to how we determine the argument space. Logically, such a space would potentially be derived from a union of the hypothesis spaces in which the beliefs of the agents involved exist, but should ideally be as small as possible so as to minimise the arguments necessary to make an enlightened decision. Leaving aside the ideal case for now, the simplest way to determine the *potential* for argument is to attribute to each agent its own *personal* argumentation framework within which it can manage its own beliefs:

DEFINITION 6. *A* **theory context** $\mathcal{C}$ *can be described by a tuple* $(\Theta, (\mathcal{L}, \vdash, \Delta), (\mathcal{A}, \rightharpoondown))$ *where:* Θ *is a* **theory core**, *a set of known facts;* $(\mathcal{L}, \vdash, \Delta)$ *is an argumentation framework;* $(\mathcal{A}, \rightharpoondown)$ *is the system of arguments which has been generated thus far using* $(\mathcal{L}, \vdash, \Delta)$*.*

- *The* **accepted extension** E *of* $\mathcal{C}$ *is any consistent complete extension of* $(\mathcal{A}, \rightharpoondown)$ *consistent with* Θ*.*

A *theory context* describes a persistent argumentation process. As discussed in [5, 6], argumentation provides a generic conception for many forms of defeasible reasoning. Thus, we can understand the selection of beliefs by agents as argumentation processes regardless of whether in practice agents actually use argumentation internally. The accepted extension of a theory context attributed to an agent defines the set of beliefs held by that agent within that context:

DEFINITION 7. *Given a theory context* $\mathcal{C}$ *with a theory core* Θ *and an accepted extension* $E \subseteq \mathcal{A}$ *of* $\mathcal{C}$*, there exists a* **theory** $\Pi = \Theta \cup (\bigcup_{\langle \Phi, \alpha \rangle \in E} \Phi)$*.*

Any acceptable extension must produce a consistent theory. The *theory core* represents the 'hard core' of unassailable assertions around which a theory is formed (selected assumptions then forming a 'protective belt' around it) [7], and so any arguments which contradict the core can be dismissed (ignored) prior to interpretation:

DEFINITION 8. *Given an argumentation framework* $(\mathcal{L}, \vdash, \Delta)$ *and a theory core* Θ*, an argument* $\langle \Phi, \alpha \rangle$ *can be* **dismissed** *if* $\Phi \vdash \varphi$ *and* $\Theta \vdash \neg\varphi$ *for some sentence* $\varphi \in \mathcal{L}$*.*

It is also possible for the assumptions upon which an abstractly acceptable extension of a system of arguments is built to be inconsistent with one another. This risk can be mitigated by a more complete argument system (identifying more points of conflict); otherwise, it may be necessary to dismiss additional arguments from the argument system until a valid extension can be found.[1]

EXAMPLE 3. *Assume that Benjamin uses the system of arguments used in Example 1 to decide whether or not he accepts that Alanna is trustworthy. If we accept argument* $\mathbf{a} = \langle \{A, B\}, C \rangle$ *as part of the (consistent) preferred extension* $\{\mathbf{a}, \mathbf{d}, \mathbf{e}\}$*, then* $A, B \in \Pi$*, where* Π *describes the beliefs of Benjamin, allowing him to infer* C *= "Alanna is trustworthy". This belief may be reinforced if Benjamin knows* D *such that* $D \in \Theta$ *and argument* $\mathbf{f} = \langle \{\neg D, E\}, \neg F \rangle$ *is dismissed (in which case* $\mathbf{a}$ *is never inadmissible). Should the environment change such that* $\Theta \vdash \neg D$ *and argument* $\mathbf{e} = \langle \{F, G\}, D \rangle$ *is dismissed instead of* $\mathbf{f}$*, then argument* $\mathbf{a}$ *becomes inadmissible, and so* $A, B \notin \Pi$ *unless* A *or* B *are used in extension* $\{\mathbf{c}, \mathbf{d}, \mathbf{f}\}$ *— if then* $\Pi \nvdash C$*, Benjamin would not assume Alanna to be trustworthy any more.*

The above example illustrates an advantageous property of argument systems in interactive domains; if we treat states of the environment as we treat different interpretations of a single state, and then rely on the theory core to dismiss arguments only pertinent to particular states, then we can construct a stable description of a volatile environment and update beliefs in response to events.

Attributing theory contexts to agents allows us to more directly compare the arguments used within a portrayal with

[1] Algorithms for identifying acceptable assumption sets can be found in the literature (*e.g.* [4, 15]).

the internal reasoning of the agents which make those arguments. They also allow us to evaluate how arguments in a portrayal which an agent has not previously considered can cause a revision of its beliefs. We can then think of portrayals as acting to implement distributed belief maintenance using argumentation motivated by the requirements of interaction — however we need to understand how arguments constructed within different argument spaces can be compared first.

4. POTENTIAL ARGUMENTATION

A *potential argument* is an argument which is supported by premises which can be derived from other, more fundamental propositions, and as such could be replaced by a more concrete argument.

Definition 9. *Given a logical framework* $(\mathcal{L}, \vdash)$*, an argument* $\langle\Phi, \alpha\rangle$ *is considered to be a* **potential argument** *in relation to another argument* $\langle\Psi, \alpha\rangle$ *(written* $\langle\Phi, \alpha\rangle \sqsubseteq \langle\Psi, \beta\rangle$*) if, for every sentence* $\varphi \in \Phi$*, it is the case that* $\Psi \vdash \varphi$*. If there exists a sentence* $\varphi \in \Psi$ *such that* $\Phi \nvdash \varphi$*, then* $\langle\Phi, \alpha\rangle \sqsubset \langle\Psi, \alpha\rangle$*.*

A potential argument $\langle\Phi, \alpha\rangle$ for another argument $\langle\Psi, \alpha\rangle$ is said to be *potentially* $\langle\Psi, \alpha\rangle$. Conversely, $\langle\Psi, \alpha\rangle$ is an *elaboration* upon $\langle\Phi, \alpha\rangle$. We can now define a *potential restriction* — a simplification of arguments so as to fit within a particular argument space.

Definition 10. *A* **potential restriction** *of a set of arguments* S *into an argument space* Δ *within an argumentation framework* $(\mathcal{L}, \vdash, \Delta)$ *is a set of potential arguments* S' *where: for each argument* $\langle\Phi, \alpha\rangle \in S$*, if* $\langle\Psi, \alpha\rangle \sqsubseteq \langle\Phi, \alpha\rangle$ *for at least one argument* $\langle\Psi, \alpha\rangle \in \Delta$*, then* $\langle\Psi, \alpha\rangle \in S'$*; every argument in* S' *is potentially at least one argument in* S*.*

By restricting arguments into a more constrained argument space, we can describe an agent's beliefs in such a way as to *potentially* match a number of more nuanced viewpoints. Since for our purposes we *ideally* want to do just enough argumentation to identify any disputes as to the resolution of a given interaction constraint and then resolve them, portrayals should restrict the arguments given by agents into a quite limited argument space. Effectively, we want to describe enough about an agent's beliefs as necessary to argue its claims whilst leaving aside any explanations for those beliefs which are not relevant to the interaction at hand.

A risk inherent in restricting arguments within a smaller argument space however is that what an agent believes to be a common dependency between arguments may be concealed such that it is not clear that an attack against one argument is an attack against another. If a statement is implied by an argument in a portrayal which an agent believes must be true to justify another argument, then it will want that other argument to explicitly rely on the truth of that statement:

Definition 11. *An agent* σ *with a theory context* $\mathcal{C}$ *and an accepted extension* E *of* $\mathcal{C}$ *considers a system of arguments* $(\mathcal{A}, \rightharpoondown)$ *to be* **balanced** *within an argument space* Δ *iff: if* $\langle\Phi, \alpha\rangle \in \mathcal{A}$*, then* $\langle\Phi, \alpha\rangle \in \Delta$ *(*i.e. $\mathcal{A} \subseteq \Delta$*); there exist no two arguments* $\langle\Phi, \alpha\rangle, \langle\Psi, \beta\rangle \in \mathcal{A}$ *such that for all* $\langle\Phi', \alpha\rangle \in E$ *where* $\langle\Phi, \alpha\rangle \sqsubseteq \langle\Phi', \alpha\rangle$*, there is a sentence* φ *such that* $\Psi \vdash \varphi$ *and* $\Phi' \vdash \varphi$*, but* $\Phi \nvdash \varphi$*.*

Thus, an agent only considers an argument space to be *well-formed* if the arguments within are balanced according to its current beliefs — *i.e.* that all arguments that it thinks are attacked by a given claim can still be seen to be attacked when potentially restricted into the given space.

Example 4. *There are two arguments* $\langle\{A, B\}, C\rangle$ *and* $\langle\{D, E\}, F\rangle$ *in a portrayal. Assume that* $\{B, G\} \vdash E$*, such that* $\langle\{D, E\}, F\rangle \sqsubset \langle\{D, B, G\}, F\rangle$*. Alanna believes* B *and* G *such that she cannot derive* E *otherwise. Thus to her, any argument with a claim contradicting* B *attacks* $\langle\{D, E\}, F\rangle$ *as well as* $\langle\{A, B\}, C\rangle$*. The arguments in the portrayal are unbalanced to Alanna — however the portrayal would be balanced if* $\langle\{D, E\}, F\rangle$ *could be replaced by* $\langle\{D, B, G\}, F\rangle$*.*

An agent wants balanced arguments so that its peers understand that certain beliefs are dependent, particularly after they leave the interaction. The perceived balance of arguments can change as beliefs change however.

We state that to *synchronise* agent beliefs is to ensure that each agent's theory represents an acceptable interpretation of the same body of evidence (it is *not* necessary for them to be jointly consistent). Synchronisation can be localised within a particular hypothesis space, such that we can ignore how theories interact outside of it:

Definition 12. *A set of theories* $\Pi_1, \ldots \Pi_n$ *(where* $n > 1$*) is* **synchronised** *within an argument space* Δ *under a logical framework* $(\mathcal{L}, \vdash)$ *iff for each theory* Π_i *(where* $1 \leq i \leq n$*): there exists a system of arguments* $(\mathcal{A}, \rightharpoondown)_i$ *such that* Π_i *can be derived from a complete extension* S_i *of* $(\mathcal{A}, \rightharpoondown)_i$*; there exists a potential restriction* S'_i *of* S_i *into* Δ*; for each theory* Π_j *(where* $1 \leq j \leq n$*),* S'_i *is a potential restriction into* Δ *of a complete extension of the system of arguments* $(\mathcal{A}, \rightharpoondown)_j$*.*

Thus two theories are locally synchronised within an argument space if the arguments derived from one theory into that space are admissible in light of the arguments derived from the other into that space, and vice versa.

Example 5. *Assume Alanna believes* $\{B, D, G, H, I, J\}$*. If a portrayal has an argument space in which arguments* $\mathbf{a} = \langle\{A, B\}, C\rangle$*, and* $\mathbf{b} = \langle\{D, E\}, F\rangle$ *exist, then assuming that* $\{G, H\} \vdash A$ *and* $\{I, J\} \vdash E$*, Alanna can construct a pair of arguments* $\mathbf{a}' = \langle\{G, H, B\}, C\rangle$*,* $\mathbf{b}' = \langle\{D, I, J\}, F\rangle$ *within a theory context such that* $\{\mathbf{a}, \mathbf{b}\}$ *is a potential restriction of* $\{\mathbf{a}', \mathbf{b}'\}$*. If* $\{\mathbf{a}, \mathbf{b}\}$ *is also a potential restriction of a complete extension* $\{\langle\{K, L, B\}, C\rangle, \langle\{M, N, E\}, F\rangle\}$ *within Benjamin's theory context, then the two agents' beliefs may be synchronised within the portrayal argument space provided that an equivalent relationship exists between Benjamin's beliefs and Alanna's theory context.*

Intuitively, we would like an agent collaborating with its peers to be able to look at their decisions and find them compatible with its beliefs. Sometimes however, there is more than one admissible decision even after knowledge is shared, and the autonomy of agents essentially forbids us (in generic circumstances) from forcing a particular interpretation of facts. We must concede then that the best option is to simply ensure that an agent's decisions are justifiable to its peers. The formal purpose of a portrayal therefore is to ensure, immediately prior to the resolution of any logical constraints imposed on interaction, that the beliefs of agents are synchronised within a well-formed argument space focused on those constraints.

5. PORTRAYAL LIFECYCLE

Portrayals augment multi-agent dialogues which are conducted under the auspices of interaction protocols written using some process calculus [11]. These protocols insist that agents involved in an interaction assume well-defined roles; progress through these roles is then contingent on the satisfaction of certain declarative constraints.

EXAMPLE 6. *A simple protocol for obtaining access to a restricted resource may require the advocacy of a third party. In this case there would be three roles to play:* applicant, overseer *and* advocate*. The following constraints would also apply: (1) the applicant must be able to identify a possible advocate to present its case to the overseer; (2) the advocate can only present that case if it considers the applicant to be trustworthy — otherwise it must refuse the applicant; (3) the overseer can only grant access if the advocate has suitable influence and the applicant is eligible for access — otherwise it must refuse to do so.*

When an agent first initiates an interaction, it conceives a new portrayal within an initial, minimal argument space, positing some basic arguments; these arguments will be for resolutions of constraints which from the outset the agent believes will need to be resolved during interaction, either by it or any other peer. As other agents are inducted into the interaction, they each receive a copy of that portrayal, which becomes their own personal *portrayal instance*. Agents can then posit their own expectations and attack any claims which they believe to be unfounded into their own instances, which are kept updated with all other instances by the exchange of messages describing any manipulations of the portrayal — in order to cope with the delay of messages in an asynchronous system, portrayal operators must be essentially commutative, so allowing agents to react intelligently to updates arriving out of order:

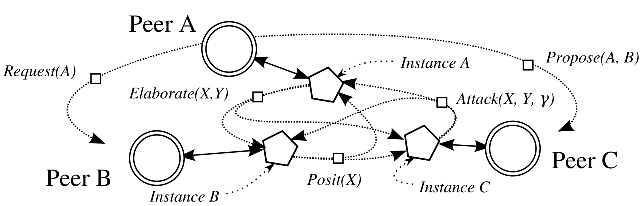


Figure 1: Portrayal instances are kept updated by exchanging messages whilst the greater interaction occurs around them.

EXAMPLE 7. *Alanna wishes to assume the role of applicant in the above protocol in order to obtain access to privileged data overseen by Charlotte. A portrayal can be used to articulate her initial expectations (that she will be found trustworthy and eligible, and that Dante would be a good advocate). If she then communicates with Dante, Dante can then contrast his own beliefs and argue with Alanna (for example, Dante may point out that he lacks the influence to be a good advocate, but Benjamin does). If Dante rejects Alanna's request, Alanna can then use the information Dante gave her and approach Benjamin instead. Benjamin might then use the portrayal to debate trust with Alanna, inviting any agent in the interaction to persuade him that Alanna is trustworthy. Further into the interaction, Charlotte can offer her opinion on Alanna's eligibility, which Benjamin and Alanna can influence by pointing out various facts. Finally, if Charlotte grants Alanna the access she requires, the resulting change in the environment can be illustrated by re-examining the assumptions on which Alanna determined that she lacked access in the first place.*

Without a portrayal, this interaction would fail — Dante is not able to fulfill the role of advocate, and Alanna is initially unaware of Benjamin. Even ignoring this, without the arguments of peers, Benjamin may not have any reason to infer that Alanna is trustworthy, and Charlotte may be unaware that Alanna fulfills eligibility criteria.

Whenever the portrayal is updated, or the local environment changes (perhaps due to actions taken by peers), agents must reconcile its content with their own beliefs, which may lead to new arguments or revised beliefs. When called upon to resolve a constraint, an agent should ensure that it is reconciled with the (current) portrayal prior to resolution according to its (revised) beliefs. As constraints are resolved, the course of interaction as described by the protocol becomes clearer, and so more constraints can be portrayed.

Because the process of reconciliation causes agents to assimilate the content of a portrayal into their own theory contexts, there is no need to preserve the portrayal at the end of the interaction — each agent can independently reinstate the arguments used in the portrayal for future interactions.

6. PORTRAYAL MECHANICS

The usefulness of synchronising beliefs within a particular argument space is very much dependent on the space itself. A *good* argument space for a portrayal is one which is expressive enough to articulate any assumptions which affect the resolution of constraints on interaction, without demanding unnecessary detail. The easiest way to ensure such a space (without precognizance of all agents' beliefs) is to initially define a very restrained argument space, and then to allow the agents themselves to expand it by requesting further elaboration upon their peer's claims — in other words, we allow the agents' self-interest to refine the space.

If we consider a portrayal instance $\mathcal{P}$ to be a container for a system of arguments, then formally, we can define the argument space of a portrayal at any particular time based on existing (ensured balanced) arguments:

DEFINITION 13. *The* **portrayal argument space** Δ *of a portrayal instance* $\mathcal{P}$ *with an existing system of arguments* $(\mathcal{A}, \rightharpoondown)$ *based on a logical framework* $(\mathcal{L}, \vdash)$ *can be described by a pair* $(\mathcal{H}, \mathcal{F})$ *as per Definition 4 where: if* S *is the set of sentences* φ *such that* $\varphi \in \Phi$ *for any argument* $\langle \Phi, \alpha \rangle \in \mathcal{A}$*, and* F' *is the set of sentences* $\varphi \in \mathcal{F}$ *such that there exists no consistent subset* $S' \subseteq S$ *where* $S' \vdash \varphi$*, then* $\mathcal{H}$ *is the union of* S*,* F' *and all sentences* $\psi \in \mathcal{L}$ *where* $S \vdash \neg\psi$*;* $\mathcal{F}$ *is the set of portrayable constraints (see below).*

We define the horizon such that we can always attack existing arguments — once an argument is then in the portrayal, it can be elaborated upon to expand the argument space further. The set of *portrayable* constraints is drawn from protocol and interaction state; initially, constraints described by the protocol may be too vague to portray (*e.g.* a constraint may require further instantiation before it can be determined which resolutions are relevant, or it might not be known whether a constraint will be encountered). We expand the focus over the course of interaction to permit arguments only for sufficiently instantiated constraints.

Example 8. *Consider the constraint "X trusts Y". It is likely that the intent is either for agents to evaluate whether a particular X trusts a particular Y, or to find an instance of Y which a particular X trusts (or vice versa). It would not generally be effective to argue about arbitrary values of X and Y, so it is clear that positing any claims of the form "X trusts Y" into a portrayal should await any variable bindings prior to actual constraint resolution (which can be determined by examination of the protocol within which the constraint is defined).*

The *initial* portrayal argument space Δ is based on the constraints expected at the start of interaction; the focus is as described above, whilst the horizon is initialised by equating it to the focus. This means that argument space will initially only accept statements satisfying or invalidating those constraints which can be predicted (*e.g.* trivial arguments of the form $\langle\{\alpha\},\alpha\rangle$) — however from this state we can rapidly expand the portrayal argument space (and more importantly, the system of arguments within) by the processes of portrayal *conception* and *reconciliation*.

Portrayal conception is the act of creating a new portrayal, and is therefore when arguments are first posited:

Algorithm 1. *An agent σ with a theory context $\mathcal{C}$ and an accepted extension E of $\mathcal{C}$* **conceives** *a new portrayal instance $\mathcal{P}$ within an initial argument space Δ with focus $\mathcal{F}$:*

1. *Whilst there exists an argument $\langle\Phi,\alpha\rangle \in E$ such that $\alpha \in \mathcal{F}$ and there is no alternative argument $\langle\Psi,\beta\rangle \in E$ such that $\beta \in \mathcal{F}$ and β subsumes α: insert $\langle\Phi',\alpha\rangle \in \Delta$ into $\mathcal{P}$, where $\langle\Phi',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$.*

2. *Until the set of arguments $\mathcal{A}$ in $\mathcal{P}$ is balanced as per Definition 11: for each elaboration $\langle\Phi',\alpha\rangle \in E$ of an argument $\langle\Phi,\alpha\rangle \in \mathcal{A}$ such that $\Phi' \vdash \varphi$ and $\Psi \vdash \varphi$ for some sentence φ and argument $\langle\Psi,\beta\rangle \in \mathcal{A}$, if $\Phi \nvdash \varphi$, then, unless there exists an alternative elaboration $\langle\Phi'_2,\alpha\rangle \in E$ of $\langle\Phi,\alpha\rangle$ such that there is no sentence ψ and argument $\langle\Psi_2,\beta\rangle \in \mathcal{A}$ for which $\Phi'_2 \vdash \psi$ and $\Psi \vdash \psi$ but $\Phi \nvdash \psi$, produce an argument $\langle\Phi'',\alpha\rangle$ such that $\langle\Phi,\alpha\rangle \sqsubset \langle\Phi'',\alpha\rangle \sqsubseteq \langle\Phi',\alpha\rangle$ and $\Phi'' \vdash \varphi$, and replace $\langle\Phi,\alpha\rangle$ with $\langle\Phi'',\alpha\rangle$ in $\mathcal{P}$.*

Algorithm 1 posits a trivially simple argument or counter-argument for any claim satisfying a (portrayable) constraint on interaction where such an argument exists in the executing agent's beliefs. Step 2 of Algorithm 1 explicates any dependencies believed necessary between arguments by elaborating them, so ensuring that they are balanced and that the portrayal argument space is thus well-formed.

The primary motivator for producing *new* arguments (and attacks against existing arguments) however is the process of portrayal reconciliation. Portrayal reconciliation is the act of ensuring that the content of a portrayal reflects the beliefs of the agents engaged in an interaction. An agent tries to reconcile its copy of the portrayal with its beliefs whenever new information is added to the portrayal (which includes when it first receives a copy of that portrayal), or when its beliefs themselves change due to external influences (*e.g.* other concurrent interactions).

Algorithm 2. *An agent σ with a theory context $\mathcal{C}$ and an accepted extension E of $\mathcal{C}$* **reconciles** *an existing portrayal instance $\mathcal{P}$ with $\mathcal{C}$:*

1. *Whilst there exists a non-empty set N of arguments $\langle\Phi,\alpha\rangle \in \mathcal{P}$ for which there exists no argument $\langle\Psi,\alpha\rangle \in \mathcal{C}$ such that $\langle\Phi,\alpha\rangle \sqsubseteq \langle\Psi,\alpha\rangle$:*
   - *(a) S is the set of all arguments $\langle\Psi',\alpha\rangle \in \mathcal{C}$ such that $\langle\Psi',\alpha\rangle \sqsubset \langle\Phi,\alpha\rangle$ for some $\langle\Phi,\alpha\rangle \in N$.*
   - *(b) If there exists any argument $\langle\Phi,\alpha\rangle \in N$ not in Δ of $\mathcal{C}$, expand Δ to include $\langle\Phi,\alpha\rangle$.*
   - *(c) Replace $\mathcal{A}$ of $\mathcal{C}$ with $(\mathcal{A}/S) \cup N$ within Δ (attacks relations are inherited by elaborations).*

2. *Whilst there exists an argument $\langle\Phi,\alpha\rangle \in E$ such that $\langle\Phi,\alpha\rangle \rightharpoondown \langle\Psi,\beta\rangle$ for some elaboration $\langle\Psi,\beta\rangle \in \mathcal{C}$ upon an argument $\langle\Psi',\beta\rangle \in \mathcal{P}$, there does not exist an alternative elaboration $\langle\Psi_2,\beta\rangle \in E$ of $\langle\Psi',\beta\rangle$ and there exists no attack $\langle\Phi',\alpha\rangle \rightharpoondown \langle\Psi',\beta\rangle$ in $\mathcal{P}$ already (where $\langle\Phi',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$):*
   - *(a) Replace $\langle\Psi',\beta\rangle$ in $\mathcal{P}$ with an elaboration $\langle\Psi'',\beta\rangle$ such that $\langle\Phi,\alpha\rangle \rightharpoondown \langle\Psi'',\beta\rangle$.*
   - *(b) Posit $\langle\Phi',\alpha\rangle$, such that $\langle\Phi',\alpha\rangle \rightharpoondown \langle\Psi'',\beta\rangle$ in $\mathcal{P}$, where $\langle\Phi',\alpha\rangle \in \Delta$ of $\mathcal{P}$.*
   - *(c) If $\langle\Phi',\alpha\rangle \rightharpoondown \langle\Upsilon,\gamma\rangle$ in $\mathcal{P}$ for any other argument $\langle\Upsilon,\gamma\rangle \in \mathcal{P}$, then ensure the attack is noted in $\mathcal{P}$.*

3. *Whilst there exists an argument $\langle\Phi,\alpha\rangle \in E$ and there exists no argument $\langle\Phi',\alpha\rangle \in \mathcal{P}$ such that $\langle\Phi',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$, but there does exists an argument $\langle\Phi'',\alpha\rangle \in \Delta$ of $\mathcal{P}$ such that $\langle\Phi'',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$, posit $\langle\Phi'',\alpha\rangle$ into $\mathcal{P}$.*

4. *Whilst there exists an argument $\langle\Phi,\alpha\rangle \in E$ and there exists an argument $\langle\Psi,\alpha\rangle \in \mathcal{P}$ such that there is a common potential argument $\langle\Psi',\alpha\rangle$ for which $\langle\Psi',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$ and $\langle\Psi',\alpha\rangle \sqsubseteq \langle\Psi,\alpha\rangle$, but $\langle\Phi,\alpha\rangle \not\sqsubseteq \langle\Psi,\alpha\rangle$ and $\langle\Psi,\alpha\rangle \not\sqsubseteq \langle\Phi,\alpha\rangle$: posit an argument $\langle\Phi',\alpha\rangle$ into $\mathcal{P}$, where $\langle\Psi',\alpha\rangle \sqsubset \langle\Phi',\alpha\rangle \sqsubseteq \langle\Phi,\alpha\rangle$ and $\langle\Phi',\alpha\rangle \not\sqsubseteq \langle\Psi,\alpha\rangle$.*

5. *Balance the set of argument $\mathcal{A}$ in $\mathcal{P}$ as per step 2 of Algorithm 1 and re-evaluate E.*

Portrayal reconciliation using Algorithm 2 has five parts: the insertion into the theory context of new (or more detailed) arguments found in the portrayal (this allows an agent to incorporate new arguments into their theory contexts and may lead to the revision of beliefs); the insertion of attacks against arguments in the portrayal (this may entail the elaboration of target arguments in order to provide an avenue of attack, which also expands the argument space of the portrayal); the insertion of new arguments relevant to the interaction as defined by the portrayal argument space; the assertion of alternative elaborations for existing arguments (typically to provide alternative support for a given claim that is unaffected by existing attacks); and the elaboration of existing arguments to balance them (which also adjusts arguments to fit within an expanded argument space). Whenever new arguments or attacks are added to a portrayal, this is communicated to every other instance, which in turn motivates peers to (re)execute the reconciliation process themselves before making any new decisions.

Example 9. *Assume that Benjamin acquires a copy of a portrayal $\mathcal{P}$. The portrayal as conceived by Alanna only contains simple arguments initially, so it is unlikely that he will learn any new concepts outright (step 1 of Algorithm 2). Benjamin does not initially accept that Alanna* *is trustworthy ($\langle\{A\}, A\rangle$), but he is aware of the elaboration $\langle\{B, C\}, A\rangle$ as defeated by his own beliefs, so he attacks $\langle\{A\}, A\rangle$ in $\mathcal{P}$ by replacing it with $\langle\{B, C\}, A\rangle$ and positing $\langle\{D, E\}, \neg B\rangle$ (step 2). He also posits any other arguments which are within the portrayal argument space — for instance, assume that Alanna did not know who the overseer was; Benjamin then posits an argument $\langle\{F\}, F\rangle$ (stating "Charlotte is the overseer"; step 3). This reconciles P for Benjamin pending Alanna's response. Alanna would then factor any new concepts into her beliefs (*e.g. *given a new argument $\langle\{G, H\}, I\rangle$, Alanna can determine whether or not her current beliefs are defended against that argument, and revise her beliefs accordingly; step 1 again). Depending on Alanna's current beliefs, Alanna would also either attack $\langle\{B, C\}, A\rangle$ (step 2), or posit an alternative argument (*e.g. *$\langle\{J, K\}, A\rangle$; step 4). There may be more arguments exchanged, which may be elaborated upon to balance the portrayal (step 5) before both Alanna and Benjamin can simultaneously reconcile P with their own beliefs.*

Once an agent has reconciled its portrayal instance with its own theories, an agent can proceed in its role in an interaction, resolving any constraints assigned to it using its revised beliefs until its beliefs change due to external interference or until another agent modifies the portrayal.

7. PORTRAYAL EVALUATION

We must now evaluate whether the processes of conception and reconciliation are sufficient to ensure the synchronisation of agent beliefs within a portrayal's argument space as described in §4. To do this, we formalise the goal state of reconciliation — an agent considers a portrayal instance to be *reconciled* if what is admissible within the argument space of the portrayal is admissible to the agent in the space of its own beliefs, and if the agent's chosen beliefs are represented within the portrayal:

DEFINITION 14. *A portrayal instance $\mathcal{P}$ is considered to be* **reconciled** *with the theory context $\mathcal{C}$ of an agent σ iff: every complete extension E_p of $\mathcal{P}$ according to σ is a potential restriction into Δ of $\mathcal{P}$ of an admissible extension E_c of $\mathcal{C}$; there exists a potential restriction R of the accepted extension E_a of $\mathcal{C}$ into Δ of $\mathcal{P}$ such that R is a complete extension of $\mathcal{P}$ according to σ.*

If a portrayal can be reconciled with the beliefs of every agent possessing a portrayal instance, then the beliefs of each agent will be locally synchronised within the argument space of the portrayal:

PROOF. For every theory Π derived from a theory context $\mathcal{C}$ held by an agent σ:

- There exists a system of arguments $(\mathcal{A}, \rightharpoondown)$ in $\mathcal{C}$ such that Π is derived from a complete extension E_a of $(\mathcal{A}, \rightharpoondown)$, where E_a is the accepted extension of $\mathcal{C}$ (being the purpose of $\mathcal{C}$).
- There exists a potential restriction R of E_a into Δ of $\mathcal{P}$ (Definition 14).
- For each theory Π_σ held by an agent σ involved in the interaction, R is a potential restriction into Δ of a complete extension of $\mathcal{C}_\sigma$ (Definition 14; we know that R is itself a complete extension, and every complete extension in $\mathcal{P}$ is a potential restriction into Δ of an complete extension in every theory context $\mathcal{C}_\sigma$ because $\mathcal{P}$ is reconciled with every $\mathcal{C}_\sigma$).

These three statements correspond to the requirements of Definition 12; therefore if every agent reconciles their portrayal instance with their beliefs, then those beliefs will be synchronised within the portrayal argument space. □

We observe that a newly conceived portrayal should always be reconciled with the agent that created it — since the arguments in the portrayal are all drawn from an accepted extension of the agent's theory context, they cannot be in conflict, and so represent a single complete extension which is a potential restriction of the original accepted extension.

Executing Algorithm 2 when an agent becomes aware of a change in the portrayal or its beliefs will ensure that its portrayal instance will become reconciled with its beliefs:[2]

PROOF. If none of the conditions for action in Algorithm 2 apply, but a potential instance $\mathcal{P}$ is not reconciled with the theory context $\mathcal{C}$ of an agent σ, then either there exists a complete extension in $\mathcal{P}$ which is not a potential restriction into Δ of $\mathcal{P}$ of an admissible extension in $\mathcal{C}$, or there exists no potential restriction into Δ of the accepted extension E_a of $\mathcal{C}$ in $\mathcal{P}$. To disprove the first of these possibilities, for every argument in a complete extension E_p of $\mathcal{P}$, there must be a complete extension E_c of $\mathcal{C}$ which is an elaboration of E_p, and for every argument in that E_c, there must be a potential argument in E_p should one exist within Δ:

- Step 1 of Algorithm 2 ensures that there will always be an elaboration in $\mathcal{C}$ of any argument in $\mathcal{P}$, so for any extension of $\mathcal{P}$, there must exist at least one set of elaborations which is a candidate for E_c.
- By Step 1 again, if E_p is attacked by an argument in $\mathcal{P}$, then that attack must exist in some form in $\mathcal{C}$. Since E_p is a complete extension, it defends itself from every attacking argument in $\mathcal{P}$, and so, given that attacks between potential arguments apply to their elaborations, any E_c must also defend against those arguments.
- We know that at least one candidate set E_c must be admissible because if it was not, then there would exist attacks against arguments in E_c which are not present in some form in $\mathcal{P}$ (thus leaving E_p admissible). By step 2 of Algorithm 2 however, given that any attacking argument with no admissible attackers itself is necessarily accepted (being part of the minimally complete grounded extension), either those attacks *would* be present in $\mathcal{P}$, or those attacks could be rejected in some admissible extension E_c of $\mathcal{C}$ subsuming E_p.

To disprove the second possibility, there must be a potential argument in $\mathcal{P}$ for every argument in the accepted extension E_a of $\mathcal{C}$ where one exists within Δ of $\mathcal{P}$, and the set of such arguments E_p must be a complete extension itself:

- Steps 2, 3 and 4 of Algorithm 2 ensure that all such potential arguments will exist in $\mathcal{P}$.
- If E_a is a complete extension in $\mathcal{C}$, then E_p should be a complete extension in P, since step 1 ensures that any attacking arguments in $\mathcal{P}$ will be in $\mathcal{C}$, and thus be already accounted for when E_a was determined.

[2]Due to space restrictions, we have omitted proof that each step in Algorithm 2 terminates — it is simple to check that each step removes any activating arguments however.

- Moreover, if there exist additional arguments in $\mathcal{P}$ defended by E_p, then step 1 ensures that they are also in $\mathcal{C}$, and so included in E_a (any attacks in $\mathcal{C}$ being then posited into $\mathcal{P}$; thus E_a and E_p are kept in harmony).

Thus, Algorithm 2 ensures that $\mathcal{P}$ is reconciled with $\mathcal{C}$. □

Therefore, provided that each portrayal instance is up to date, the beliefs of every agent using a portrayal are synchronised until further changes to the world. Arguments within the portrayal are kept balanced by identifying arguments which violate Definition 11 and replacing them immediately at the end of Algorithms 1 (step 2) and 2 (step 5).

8. CONCLUSIONS

Research into multi-agent interaction and argumentation tends to focus on specifying protocols *for* argumentation [10], rather than finding ways to augment existing interactions *with* argumentation, which is a different objective. The portrayal mechanism described here is an opportunistic system which works alongside task-oriented interactions which are *not* (necessarily) themselves concerned with promoting particular claims over others, though they may naturally exhibit the kind of negotiation / coordination tasks for which argumentation is considered to be useful.

In resolving the question of how portrayals can be used efficiently, we have had to address the question of how the space in which argumentation is conducted should be defined. This distinguishes our work from many prior theoretic treatments of assumption-based argumentation (*e.g.* [4, 6]) where the argumentation process is considered in isolation — *i.e.* the quality of the assumptions available for supporting arguments cannot be evaluated without an external reference, and so is taken as given. In the case of portrayals, we are taking an abstraction of the combined beliefs of peers, and as such we have many external references. We can use them to consider whether the hypothesis space in which arguments are created is sufficiently expressive to describe the possible claims and attacks which agents might want to make, and if not, we can manipulate that space accordingly.

The act of producing and maintaining a portrayal influences the beliefs of agents by synchronising them within the hypothesis space in which argumentation is performed, which then influences the decisions they make during interaction, being based on those beliefs. This imparts a number of practical benefits: agents are able to share their beliefs and so resolve constraints on interaction that they could not independently; if there is enough information available within the chosen argument space, agents can reach a consensus on how constraints should be resolved based on current circumstances; and even if there is *not* enough information to reach consensus, agents can reach a state wherein they can at least accept their peers' decisions as justifiable given the information available.

Moreover, because a portrayal is constructed based on the interest agents have in particular constraints, there is less need for complex, domain-specific interaction protocols which explicitly check particular factors when making a decision — if agents are concerned with a particular issue, they will elaborate upon and attack arguments which they see as contingent on that issue automatically. Thus a protocol need only concern itself with the coordination of agents based on those agents' evaluation of high-level propositions, and so can be made more generically applicable.

9. REFERENCES

[1] A. Bondarenko, F. Toni, and R. Kowalski. An assumption-based framework for non-monotonic reasoning. In *Proceedings of the Second International Workshop on Logic Programming and Nonmonotonic Reasoning*, pages 171–189, 1993.

[2] M. Caminada. Comparing two unique extension semantics for formal argumentation: Ideal and eager. In *Proceedings of the 19th Belgian-Dutch Conference on Artificial Intelligence*, pages 81–87, 2007.

[3] 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):321–357, 1995.

[4] P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, 171(10–15):642–674, July–October 2007.

[5] A. C. Kakas, R. A. Kowalski, and F. Toni. The role of adbuction in logic programming. *Handbook of Logic in Artificial Intelligence and Logic Programming*, 5:235–324, 1998.

[6] A. C. Kakas and F. Toni. Computing argumentation in logic programming. *Journal of Logic and Computation*, 9:515–562, 1999.

[7] I. Lakatos. Falsification and the methodology of scientific research programmes. In I. Lakatos and A. Musgrave, editors, *Criticism and the Growth of Knowledge*, pages 91–195. Cambridge University Press, 1970.

[8] H. Prakken. From logic to dialectics in legal argument. In *Proceedings of the 5th International Conference on Artificial Intelligence and Law*, pages 165–174, 1995.

[9] I. Rahwan and K. Larson. Pareto optimality in abstract argumentation. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 2008.

[10] I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. McBurney, S. Parsons, and L. Sonenberg. Argumentation-based negotiation. *The Knowledge Engineering Review*, 18(4):343–375, 2003.

[11] D. Robertson. Multi-agent coordination as distributed logic programming. In *Proceedings of the International Conference on Logic Programming*, 2004.

[12] D. Robertson, C. Walton, A. Barker, P. Besana, Y. H. Chen-Burger, F. Hassan, D. Lambert, G. Li, J. McGinnis, N. Osman, A. Bundy, F. McNeill, F. van Harmelen, C. Sierra, and F. Giunchiglia. Models of interaction as a grounding for peer to peer knowledge sharing. *LNCS Advances in Web Semantics*, 1, 2008.

[13] R. Siebes, D. Dupplaw, S. Kotoulas, A. P. de Pinnick, F. van Harmelen, and D. Robertson. The openknowledge system: An interaction-centered approach to knowledge sharing. In *Proceedings of the 15th International Conferece on Cooperative Information Systems*, 2007.

[14] M. P. Singh. Agent communication languages: Rethinking the principles. *Computer*, 21:40–47, December 1998.

[15] G. A. W. Vreeswijk and H. Prakken. Credulous and sceptical argument games for preferred semantics. In *Proceedings of the 7th European Workshop on Logic for Artificial Intelligence*, pages 239–253, 2000.