ECAI 2010
H. Coelho et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-606-5-997

An Argumentation-Based Dialog for Social Evaluations Exchange

Isaac Pinyol and **Jordi Sabater-Mir**[1]

Abstract. In open multiagent systems, agents depend on reputation and trust mechanisms to evaluate the behavior of potential partners. Often these evaluations (social evaluate) are associated with a measure of reliability that the source agent computes. When considering communicated social evaluations, this may lead to serious problems due to the subjectivity of reputation-related information. In this paper, instead of considering only reliability measures computed from the sources, we provide a mechanism that allows the recipient according to its own knowledge, decide whether the piece of information is reliable. We do this by allowing the agents engage in an argumentation-based dialog

1 Introduction

Computational reputation models use the outcomes of past interactions and third-party communications regarding social evaluations as main sources to compute new social evaluations. Some recent reputation models also attach to the social evaluation a reliability measure that reflects how *confident* is the owner of the social evaluation about that value.

Usually, the reliability value is also transmitted together with the social evaluation when there is a communication, so the recipient agent can decide whether it is worth it to consider that piece of information. However, due to the subjectivity of reputation information, a social evaluation declared as totally reliable according to agent A may not be reliable for agent B because the bases under which A has inferred the social evaluation cannot be accepted by B. This paper offers an alternative mechanism. We suggest that, in communicated social evaluations, the reliability measure cannot be dependent on the source agent, but must be fully evaluated by the recipient agent according to its own knowledge. In our approach, rather than only allow one shot communications, we allow agents to participate in argumentation-based dialogs regarding reputation elements in order to decide on the reliability (and thus acceptance) of a communicated social evaluation.

2 The L_{Rep} Language

L_{Rep} language captures the reputation-related information that reputation models compute. Far from being a complete unifying language, it is capable of expressing the reputation information that many current state-of-the-art reputation models provide, because it is based on an ontology of reputation [6] that has been used to represent the information of some of the most popular reputation models.

[1] IIIA-CSIC - Artificial Intelligent Research Institute, Spanish National Research Council, Bellaterra, Barcelona, Spain. email:{ipinyol,jsabater}@iiia.csic.es

Social evaluations incorporate three main elements: the target, the context, and the value of the evaluation [6]. For instance, an evaluation may say that an agent A (target) is very good (value) as a car driver (context). We define L_{rep} as a first-order sorted language with special predicates that indicates the types of social evaluations. The sorts that the language includes are a finite set of individual agent identifiers and group identifiers $\mathcal{A}$ (the target of the evaluation), a countable set of contexts $\mathcal{C}$ (the context of the evaluation), a countable set of values $\mathcal{V}$ (the value of the evaluation), a countable set of discrete time units $\mathcal{T}$, and a countable set of formula identifiers $\mathcal{F}$ that includes a constant for each well-formed formula of the language.

The sort $\mathcal{V}$ represents values of a totally ordered set $M = \langle G, \leq \rangle$. It includes a constant v for each $v \in G$. Examples of M are $\langle [0,1] \cap \mathbb{Q}, \leq \rangle$, or $\langle \{VB, B, N, G, VG\}, \leq_s \rangle$ referring to the linguistic labels *Very Bad, Bad, Neutral, Good, Very Good*, and where $VB \leq_s B \leq_s N \leq_s G \leq_s VG$.

Let $i, j, \mathcal{G} \in \mathcal{A}$ (i and j individual agents, and $\mathcal{G}$ referring to a group of agents), $c \in \mathcal{C}$, $v \in \mathcal{V}$, $t \in \mathcal{T}$ and $f \in \mathcal{F}$. The predicates included in L_{rep} are $Img(j, c, v)$ (image), $Rep(j, c, v)$ (reputation), $ShV(j, c, v, \mathcal{G})$ (shared voice), $ShI(j, r, v, \mathcal{G})$ (shared image), $DE(j, c, v, t)$ (direct experience) and $Comm(j, f, t)$ (communication). The semantics of each predicate can be found at [6]

The reputation-related information that agent i holds is characterized by the tuple $\langle \Delta_i, \vdash_i \rangle$, where Δ_i is the set of ground elements (DE and Comm) gathered by i through interactions and communications (i's reputation theory), and $\vdash_i$ the consequence relation (i's reputation model).

3 The Reputation Argumentation Framework

We define the argumentative language L_{arg} (based on L_{rep}) and its associated inference relation $\vdash_{arg}$, to characterize how arguments are expressed and constructed. A formula $(\Phi{:}\alpha) \in wff(L_{arg})$ (*argument*) when $\alpha \in wff(L_{Rep})$ and $\Phi \subseteq wff(L_{Rep})$. An *argumentative theory* (adapted from [2]) is a finite set $\Gamma = \{\gamma_1, \ldots, \gamma_n\}$ where each γ_i is a formula $(\{\alpha\}{:}\alpha) \in wff(L_{arg})$ (*basic declarative unit* (bdu)). We say that $(\Phi{:}\alpha)$ is a valid argument in the bases of Γ iff $\Gamma \vdash_{arg} (\Phi{:}\alpha)$. Also, we say that a valid argument $(\Phi_2{:}\alpha_2)$ is a subargument of $(\Phi{:}\alpha)$ iff $\Phi_2 \subset \Phi$. $\vdash_{arg}$ is characterized by the following rules:

$$\text{Intro-BDU: } \frac{}{(\{\alpha\}{:}\alpha)} \qquad \text{Intro-AND: } \frac{(\Phi_1{:}\alpha_1), \ldots\ (\Phi_n{:}\alpha_n)}{(\bigcup_{i=1}^{n} \Phi_i{:}\alpha_1, \ldots, \alpha_n)}$$

$$\text{Elim-IMP: } \frac{(\Phi_1{:}\alpha_1, \ldots, \alpha_n \to \beta) \qquad (\Phi_2{:}\alpha_1, \ldots, \alpha_n)}{(\Phi_1 \cup \Phi_2 : \beta)}$$

As mention earlier, each agent i has to construct its argumentative theory Γ_i in order to build arguments. Assuming that $\vdash_i$ is defined by a finite set of natural deduction rules $\{\vdash_{i_1}, \ldots, \vdash_{i_m}\}$, (1) For all

Movement	Precondition	Postcondition
$\text{counter}^k_{PRO}(b)$	(1) k is even, $b \in C(\Gamma_{PRO} \cup X^{k-1}_{OPP})$ and b has not been issued yet (2) $\exists r \in I\!N$ s.t. $1 \leq r < \lvert S^{k-1}\rvert$, r is odd and $(b, S^{k-1}_r) \in R(\Gamma_{PRO} \cup X^{k-1}_{OPP})$ (3) $\nexists \gamma \in C(\Gamma_{PRO} \cup X^{k-1}_{OPP})$ s.t. $(\gamma, S^{k-1}_t) \in R(\Gamma_{PRO} \cup X^{k-1}_{OPP})$ where $r+1 \leq t < \lvert S^{k-1}\rvert$ and t is odd	$X^k_{PRO} = X^{k-1}_{PRO} \cup BDU(supp(b))$ $X^k_{OPP} = X^{k-1}_{OPP}$ $S^k = \langle S^{k-1}_0, \ldots, S^{k-1}_r, b\rangle$
$\text{counter}^k_{OPP}(b)$	(1) k is odd, $b \in C(\Gamma_{OPP} \cup X^{k-1}_{PRO})$ and b has not been issued yet (2)$\exists r \in I\!N$ s.t. $0 \leq r < \lvert S^{k-1}\rvert$, r is even and $(b, S^{k-1}_r) \in R(\Gamma_{OPP} \cup X^{k-1}_{PRO})$ (3)$\nexists \gamma \in C(\Gamma_{OPP} \cup X^{k-1}_{PRO})$ s.t. $(\gamma, S^{k-1}_t) \in R(\Gamma_{OPP} \cup X^{k-1}_{PRO})$ where $r+1 \leq t < \lvert S^{k-1}\rvert$ and t is even	$X^k_{PRO} = X^{k-1}_{PRO}$ $X^k_{OPP} = X^{k-1}_{OPP} \cup BDU(supp(b))$ $S^k = \langle S^{k-1}_0, \ldots, S^{k-1}_r, \beta\rangle$ (or $\langle S^{k-1}_0, b\rangle$ if $r = 0$)

Figure 1. Possible movements of the dialog game at turn k. The function $supp(b)$ returns the supporting set of b. The function $BDU(X)$ returns the set of elements from X as a BDU formula. So, if $\alpha \in X$, then $\{\alpha\}{:}\alpha \in BDU(X)$.

$\alpha \in \Delta_i$ then $(\{\alpha\}{:}\alpha) \in \Gamma_i$, and (2) for all $\alpha_1, \ldots, \alpha_n$ s.t. $\Delta_i \vdash_i \alpha_k$ where $1 \leq k \leq n$, if there exists m s.t. $\alpha_1, \ldots, \alpha_n \vdash_{i_m} \beta$, then $(\{\alpha_1, \ldots, \alpha_n \to \beta\}{:}\alpha_1, \ldots, \alpha_n \to \beta) \in \Gamma_i$. It can be proved that if $\Delta_i \vdash_i \alpha$, then there exists an argument $(\Phi : \alpha)$ such that $\Gamma_i \vdash_{arg} (\Phi : \alpha)$.

To specify the *attack* relationship among arguments, we define first the binary relation $\cong$ between L_{rep} predicates. Let α, β be well-formed non-ground formulas from L_{Rep}. Then, $\alpha \cong \beta$ iff $type(\alpha) = type(\beta)$, $\alpha.target = \beta.target$, $\alpha.context = \beta.context$ and $\alpha.value \neq \beta.value$. $\cong$ is symmetric but not reflexive nor transitive. Then, let $(\Phi_1{:}\alpha_1)$, $(\Phi_2{:}\alpha_2)$ be valid arguments in the bases of Γ. We say that $(\Phi_1{:}\alpha_1)$ *attack* $(\Phi_2{:}\alpha_2)$ iff $\exists(\Phi_3{:}\alpha_3)$ subargument of $(\Phi_2{:}\alpha_2)$ s.t. $(\alpha_1 \cong \alpha_3)$. The strength of the attack is calculated through the function w as $w(a, b) = \alpha_1.value \ominus \alpha_3.value$, where $\ominus$ is a binary function defined over the domain of the representation values used to quantify the evaluations (the total ordered set $M = \langle G, \leq\rangle$). $\ominus$ implements a *difference* function among the possible values.

We can instantiate now a weighted argument system [5] by using the constructs defined in this section. Let Γ be an argumentative theory. We define: $C(\Gamma) = \{(\Phi : \alpha) | \Gamma \vdash_{arg} (\Phi : \alpha)\}$, the set of all valid arguments that can be deduced from Γ, and $R(\Gamma) = \{((\Phi_1 : \alpha_1), (\Phi_2 : \alpha_2)) | (\Phi_1 : \alpha_1) \text{ attacks } (\Phi_2 : \alpha_2) \text{ and } (\Phi_1 : \alpha_1) \in C(\Gamma) \text{ and } (\Phi_2 : \alpha_2) \in C(\Gamma)\}$ (the set of all possible attack relations between the arguments in $C(\Gamma)$). Then, we can describe our *reputation argument framework* as $AF_\Gamma = \langle C(\Gamma), R(\Gamma), w\rangle$, where $w : R(\Gamma) \to I\!R_>$ is the strength function as defined above using the $\ominus$ difference function. We assume familiarity with Dung's abstract argument systems [3] and its weighted extension [5]. We can use any weighted acceptability semantics to decide about the reliability of arguments.

4 The Argumentation-Based Dialog

Each agent participating in the dialog will use its own argument framework to deal with possible inconsistencies. Let PRO and OPP be the proponent (who starts the dialog) and the opponent agent, then (similar to [1]):

$$AF_{PRO} = \langle C(\Gamma_{PRO} \cup X_{OPP}), R(\Gamma_{PRO} \cup X_{OPP}), w_{PRO}\rangle$$
$$AF_{OPP} = \langle C(\Gamma_{OPP} \cup X_{PRO}), R(\Gamma_{OPP} \cup X_{PRO}), w_{OPP}\rangle$$

where $\Gamma_{PRO}, \Gamma_{OPP}$ are the argumentative theories of agents PRO and OPP, which are private. w_{PRO} and w_{OPP} are the weight functions of agents PRO and OPP. X_{PRO} is the set of bdu from the arguments that results from the proponent's issued arguments. X_{OPP} the equivalent for the opponent. Both X_{PRO} and X_{OPP} are public and are the result of the exchange of arguments. Inspired in [4], a state of a dialog at the k-th turn (where $k \geq 0$) is characterized by the tuple $\langle S^k, X^k_{PRO}, X^k_{OPP}\rangle^k$ where $S^k = \langle S^k_1, \ldots, S^k_t\rangle$ is the ordered set of arguments that represents a single dispute line. X^k_{PRO}, X^k_{OPP} are the public sets of *bdu* formulas of the proponent and the opponent respectively at turn k, incrementally built after each argument exchange and that are public. The proponent is the initiator of the dialog and issues the argument $a = (\Phi{:}\alpha)$. The initial state at turn 0 is then characterized by $\langle\langle a\rangle, BDU(\Phi), \{\}\rangle^0$. The possible types of movements are summarized in figure 1. The winner is the last participant who makes a move. The protocol is a simplification of a TPI-Dispute [1]. From there, it can be deduced that if the proponent is the winer, the original argument $a = (\Phi{:}\alpha)$ belongs to a 0-preferred extensions of AF_{OPP}.

If OPP wins, OPP cannot find a 0-preferred extension that includes the argument a. In this case, OPP could choose not to update its reputation theory. However, depending on its tolerance to inconsistencies, OPP can find a 1-preferred extension that includes argument a, or even a 2-preferred extension. By increasing the inconsistency budget, the original argument may become acceptable, and thus the communicated social evaluation reliable [5].

5 Conclusions and Future Work

In this paper, we have defined an argumentation-based protocol for the exchange of reputation-related information that allows agents to judge whether a given piece of information is reliable or not. We use argumentation techniques for the semantics of the protocol. The next step regarding this work will be the inclusion of defeats among arguments. We plan to use the typology of ground elements to give strength to the arguments, independent of their attack relations. For instance, one may consider that arguments based on direct experiences are stronger than those based on communications.

6 Acknowledgments

This work was supported by the projects AT (CONSOLIDER CSD2007-0022, INGENIO 2010), LiquidPub (STREP FP7-213360), RepBDI (Intramural 200850I136) and partially supported by the Generalitat de Catalunya under the grant 2009-SGR-1434.

REFERENCES

[1] L. Amgoud, N. Maudet, and S. Parsons, 'Modelling dialogues using argumentation', in *ICMAS'00*, pp. 31–38, (2000).
[2] C.I. Chesñevar and G.R. Simari, 'Modelling inference in argumentation through labeled deduction: Formalization and logical properties', *Logica Universalis*, **1**(1), 93—124, (2007).
[3] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *AI*, **77**(2), 321–358, (1995).
[4] P.E. Dunne and T.J.M. Bench-Capon, 'Two party immediate response disputes: Properties and efficiency', *AI*, **149**(2), 221 – 250, (2003).
[5] P.E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge, 'Inconsistency tolerance in weighted argument systems', in *AAMAS'09*, pp. 851–858, (2009).
[6] I. Pinyol, J. Sabater-Mir, and G. Cuni, 'How to talk about reputation using a common ontology: From definition to implementation', in *TRUST@AAMAS'07*, pp. 90—101, (2007).