Strength calculation of rewards

Mariela Morveli-Espinoza, Ayslan T. Possebom and Cesar A. Tacla
CPGEI - Federal University of Technology - Paraná, Brazil
{morveli.espinoza, possebom, cesar.tacla}@gmail.com

Abstract

Persuasive negotiation involves negotiating using rhetorical arguments (such as threats, rewards, and appeals), which act as persuasive elements that aim to force or convince an opponent to accept a given proposal. In the case of rewards, these have a positive nature as they use the argument that something positive will happen to the opponent if he accepts to do the requirement sent by the proponent. A proponent agent can generate more than one reward depending on the information he has modeled of his opponent. The problem appears when the agent has to choose a reward, to send to his opponent, among a set of rewards. One measure that could help him in his choice is the strength each reward has. Thus, the goal of this work is to analyze the rewards components and to propose a model for calculating their strength. We propose two ways for calculating the strength of rewards depending on the kind of negotiation the agent is participating. The first proposal is to be used when the agent negotiates only with one opponent, and the second when the agent negotiates with more than one opponent.

1 Introduction

Persuasive negotiation involves negotiating using rhetorical arguments, which act as persuasive elements that aim to force or convince an opponent to accept a given proposal [9].

Although some authors argue that threats are the strongest rhetorical arguments ([11], [8]), the choice of which kind of argument will be used by a proponent agent depends on the information he has modelled about his opponents. According to Ramchurn et al.[9], it also depends on the convenience of the proposal for the proponent and the degree of trust that exists between the two agents. Therefore, as appropriate, a reward could be more effective than a threat. In this work we study rewards, which have a positive nature as a proponent agent can entice an opponent of him to do certain action by offering to do another action as a reward [1].

Let's see the following persuasive negotiation scenario where *boss* is an agent proponent, *employee* an agent opponent and the goal of *boss* is that *employee* works every weekend[1]. Taking into account the knowledge base of agent *boss*, the following rewards can be generated:

- ***boss:*** *if you work every weekend, you will receive an interim payment.*
- ***boss:*** *if you work every weekend, you will have more holidays.*

The question is which of these rewards will *boss* choose to persuade *employee* to work every weekend? One way of knowing this is by calculating the strength of the generated rewards. According to Ramchurn et al.[9], a strong argument (in this case a reward) is one that quickly convinces an opponent to do a proposal, while a weak argument is less persuasive. Therefore, calculating the strength of rewards is important in persuasive negotiation dialogues, since the quickness of persuasion depends on it.

Rewards are constructed using both proponent and opponent's goals (for example, earning more money). Some researches on this topic take into account the importance of the opponent's goal for the agent opponent and the certainty level of the beliefs used for the argument generation [1][3]. However, there are other additional criteria that are necessary. Following, some examples of situations that show this need:

1. Agent *boss* knows that "visiting his parents" (g_1) is a more important goal for *employee* than "fixing his car" (g_2). Considering only the importance, *boss* would use g_1 for generating a reward. However, what happens if *boss* knows that g_1 is less achievable than

[1] This scenario is inspired by the example presented in [1]

g_2 since visiting his parents is not possible for the moment because *employee* has a spine disorder that does not let him travel long distances? In cases like this, the importance is no longer the best or the unique criterion, related to the goal of the opponent.

2. Agent *boss* has already done rewards before and he has rarely fulfilled it, and obviously *employee* knows about it. In this case, the strength of a reward is also influenced by the execution credibility that the proponent has from the point of view of his opponent. Thus, even when the goal of an opponent is very important and/or achievable, a low level of credibility could diminish the value of the strength of a reward.

In the first case, notice that besides importance, there exists another criterion to evaluate the quality of the goal of an opponent, because it does not matter how important a goal is if it is not possible to be achieved. And in the second case, the credibility execution level of the proponent (from the point of view of the opponent, i.e. the execution level the proponent believes the opponent has about him) should also be considered.

With respect to the first criterion, to determine how achievable a goal is, we will use the belief-based goal processing model proposed by Castelfranchi and Paglieri [5]. It can be considered an extension of the belief-desire-intention model (BDI) [4] model, but unlike it, in Castelfranchi and Paglieri's model, the processing of goals is divided in four stages: (i) activation, (ii) evaluation, (iii) deliberation, and (iv) checking; and the states a goal can adopt are: (i) active (=desire), (ii) pursuable, (iii) chosen, and (iv) executable (=intention). The state of a goal changes when it passes from one stage to the next. Thus, when it passes the activation stage it becomes active, when it passes the evaluation stage it becomes pursuable, and so on. A goal is closer to be achieved when it is closer of passing the last stage.

Part of our proposal for calculating the strength of a reward considers the state of the goals of the opponent. Depending on this state, a goal can be considered more or less rewardable, a goal is considered more rewardable when it is closer of the executable state and less rewardable when its state is active. Thus, the aim of this article is to propose a model for calculating the strength of rewards by taking into account new criteria, which will lead to a more accurate calculation.

The paper is organized as follows: Section 2 presents the goal processing model on which our strength calculation model is primarily based. In Section 3 a negotiating agent architecture that considers the necessary mental states, structures and functions that support our proposal is defined. A formal definition of reward and the mechanism for its generation are presented in Section 4. Section 5 shows an analysis of the elements of a reward and our proposed strength calculation model. In Section 6, the main related works are compared with our proposal. Finally, Section 7 is devoted to some conclusions and future works.

2 Belief-based goal processing model

In this section, the four stages of the goal processing model of Castelfranchi and Paglieri are presented[2]. The aim of this section is not to present in detail the beliefs used in each stage. We focus on the goals states and make clear when a goal is considered active, pursuable, chosen, and executable, because these states will be used in the strength calculation model that is proposed in this work. Following a brief description of each stage:

1. Activation stage: In this stage, goals are activated by means of motivating beliefs. For example, if the agent has the belief that today is Thursday, it activates the goal of going to the French class, or the motivating belief that today is sunny activates the goal of playing football. When a motivating belief is satisfied, the supported goal becomes *active*. An active goal can also be seen as a desire.

2. Evaluation stage: In this stage, goals are evaluated using assessment beliefs. When there are no assessment beliefs for a certain goal, it becomes *pursuable*. Three types of assessment beliefs were defined: (i) those that control that there is no impossibility for a goal be pursued; (ii) those that control goals that are realized in the world autonomously and without the direct intervention of the agent; and (iii) those that control goals that have already realized, and that will remain as such.

3. Deliberation stage: The aim of this stage is to act as a filter on the basis of incompatibilities and preferences among pursuable goals. Goals that pass this stage are called *chosen* goals. These beliefs are concerned with the different forms of incompatibility among goals that lead an agent to choose among them. For dealing with incompatibilities, an agent uses preference beliefs.

4. Checking stage: The aim of this stage is to evaluate whether the agent knows how to achieve a goal and if it is capable of performing the required actions to achieve a chosen goal, in other words if the agent has a plan and he is capable of executing it. Goals that pass this stage are called *executable* goals and have the same characteristics of intentions.

[2]A more detailed version of this model is presented in [5].

3 The agent

In this section, we define the basic and compound structures, and functions of the agent, which are necessary in order to be able to generate rewards and calculate their strengths[3]. This architecture is based on the Castelfranchi and Paglieri's model.

Before defining the structures of an agent, let $\mathcal{L}$ be a first order logical language which will be used to represent the goals and beliefs of the agent. $\wedge$,$\vee$ and $\neg$ denote the logical connectives conjunction, disjunction and negation, and $\vdash$ stands for the classical inference.

Definition 3.1. (Basic structures) An agent has five basic structures:

- $\mathcal{K}$ is the knowledge base of the agent;
- $\mathcal{O}p$ stores the opponents of the agent;
- $\mathcal{G} = \mathcal{G}_a \cup \mathcal{G}_p \cup \mathcal{G}_c \cup \mathcal{G}_e$ is the set of goals of the agent, such that $\mathcal{G}_a$ the set of active goals, $\mathcal{G}_p$ the set of pursuable goals, $\mathcal{G}_c$ the set of chosen goals and $\mathcal{G}_e$ the set of executable goals. It holds that $\mathcal{G}_x \cap \mathcal{G}_y = \emptyset$, for $x, y \in \{a, p, c, e\}$ and $x \neq y$;
- $\mathcal{GO} = \mathcal{GO}_s \cup \mathcal{GO}_a \cup \mathcal{GO}_p \cup \mathcal{GO}_c$ is the set of goals of the opponent, such that $\mathcal{GO}_s$ is the set of sleeping goals of the opponent[4], $\mathcal{GO}_a$ is the set of active goals of the opponent, $\mathcal{GO}_p$ the set of pursuable ones, and $\mathcal{GO}_c$ the set of chosen ones. It holds that $\mathcal{G}_x \cap \mathcal{G}_y = \emptyset$ for $x, y \in \{s, a, p, c\}$ and $x \neq y$. Finally, let $State(go_i) = z$ be a function that returns the state of a given goal; for $z \in \{0, 1, 2, 3\}$ where 0 means that the goal is sleeping, 1 active, 2 pursuable and 3 that it is chosen;
- $\mathcal{R}ws$ stores the rewards constructed by the agent. The logical definition of a reward is given in Section 4.

Definition 3.2. (Compound structures) These store characteristics of the basic structures.

- $\mathcal{O}p_{det} = \{(op_i, \delta)\}$ such that $op_i \in \mathcal{O}p$ and δ is the execution credibility level of rewards the proponent has from the point of view of opponent op_i. Hereafter, we denote that $\delta \in [0, 1]$ such that δ is a real from the given interval. Let $Level\_Exec_{rw}(op_i) = \delta$ be a function that returns the execution credibility level for a given opponent agent;
- $\mathcal{GO}_{det} = \{(go_i, \delta, op_j)\}$ such that $go_i \in \mathcal{GO}$ is a goal of opponent $op_j \in \mathcal{O}p$ whose importance is given by δ. Let $Importance(go_i, op_j) = \delta$ be a function that returns the importance of a given goal; the opponent op_j is taken into account as more that one opponent may have the same goal but with different importance degree. Finally, let $Op\_Goals(op_j) = \{go_i, ...go_k\}$ be a function that returns all the goals of a given opponent.

Definition 3.3. (Function) Besides the functions defined previously, an agent is also equipped with the following function:

- $EvalRecomp(go_j)$ is a function that takes as input a rewardable goal of an opponent and returns a reward action. This function lets the proponent choose an adequate reward action, that is within the possibilities of the proponent, thereby, he can fulfill the offered reward.

4 Construction of rewards

A reward is constructed based on two goals:

1. An outsourced goal of the proponent: This kind of goal needs the opponent involvement for being achieved. For example, the goal of agent *boss* is that *employee* works every weekend, for this goal be achieved it is necessary that *employee* executes the required action. Considering the goal processing stages defined in Section 2 the state of this goal is executable.

Definition 4.1. (Outsourced goal) An outsourced goal g_i is an expression of the form $g_i(op_k, g'_i)$, such that, $op_k \in \mathcal{O}p$ and g'_i is an action that op_k has to execute. Let $first(g_i) = op_k$ and $second(g_i) = g'_i$ be the functions that return each component of g_i.

2. The goal of an opponent: $go_i \in \mathcal{GO}$ is a goal that proponent agent knows its opponent wants to achieve. For example, *boss* knows that *employee* wants to visit his parents. Besides knowing the goal of his opponent, the proponent has to know the state of that goal and its importance.

The construction of a reward begins when (i) an outsourced goal g_i passes all the goal processing stages and becomes executable and, (ii) after a failed first attempt of proponent agent to make his opponent to do the requested action g'_i. The process of construction of a reward is the following:

1. Function Op_goals returns the set of rewardable goals the proponent knows about the opponent op_j. Let $\mathcal{S}_{go} = Op\_goals(op_j)$ be the returned set.
2. If $\mathcal{S}_{go} \neq \emptyset$
    (a) For each $go_j \in \mathcal{S}_{go}$:
        i. Obtain $rop = EvalRecomp(go_j)$ to know a reward option for goal go_j.

[3] In this work, we assume that the agent has in advance the necessary information (importance and state of the goals, and the value of the credibility level of execution of rewards) for generating and calculating the strength of rewards. Some interesting works about opponent modelling related to argumentation are [7, 10].

[4] Sleeping is one of the states a goal may take that is proposed in [6]. A goal is in sleeping state when it has not been activated yet. In this work, we use the sleeping state to denote those goals of the opponent that are not necessarily active but that are important for him. For example, a mother knows that offering chocolates for her son may be a good reward, however it does not mean that the son has the goal "obtaining chocolates" always active.

ii. Generate the two rules necessary for constructing a reward. The first is an expression of the form $rrw_1 = g'_i \rightarrow rop$ and the second is $rrw_2 = rop \rightarrow go_j$.

iii. Construct a reward and save it in $\mathcal{R}ws$.

Rewards in $\mathcal{R}ws$ are called candidates. After the strength calculation, the strongest one is sent to his opponent to try to persuade him.

Following, we present the formal definition of a reward, this is based on the definition given in [1], with some modifications that consider the agent architecture proposed in the previous section.

Definition 4.2. (Reward) A reward is a triple $rw = \langle \mathcal{R}rw, g'_i, go_j \rangle$, where:
- $\mathcal{R}rw = rrw_1 \cup rrw_2$ contains both reward rules,
- $g'_i = second(g_i)$, such that $g_i \in \mathcal{G}_e$,
- $\mathcal{R}rw \cup \{g'_i\} \vdash go_j$ such that $go_j \in \mathcal{GO}$.

Let's call $\mathcal{R}rw$ and g'_i the support of the reward and go_j its conclusion.

Example 4.1. Let us define the mental state of *boss*:
$\mathcal{G}_e = g_1$, where:
$g_1 = make(employee,$'work(weekend)') is an outsourced goal, and therefore $g'_1 = work(weekend)$
$\mathcal{GO}_a = \{go_1\}$, where $go_1 = visit(parents)$,
$\mathcal{GO}_c = \{go_2\}$, where $go_2 = fix(car)$,
$\mathcal{GO}_{det} = \{(go_1, 0.8, employee), (go_2, 0.6, employee)\}$,
$\mathcal{O}p = \{employee\}$, $\mathcal{O}p_{det} = \{(employee, 1)\}$,
$\mathcal{R}ws = \{\}$

Let us suppose that agent *employee* rejected to do action $g'_1 = work(weekend)$. Therefore, *boss* begins the process of construction of candidate rewards:

1. $\mathcal{S}_{go} = Op\_goals(employee)$, so $\mathcal{S}_{go} = \{go_1, go_2\}$

2. $\mathcal{S}_{go} \neq \emptyset$, then

   (a) $EvalRecomp(go_1) = have(holidays)$
   Generate $\mathcal{R}rw_1 = \{g'_1 \rightarrow have(holidays), have(holidays) \rightarrow go_1\}$
   Construct $rw_1 = \langle \mathcal{R}rw_1, g'_1, go_1 \rangle$

   (b) $EvalRecomp(go_2) = give(inter\_paym)$
   Generate $\mathcal{R}rw_2 = \{g'_2 \rightarrow give(inter\_paym), give(inter\_paym) \rightarrow go_2\}$
   Construct $rw_2 = \langle \mathcal{R}rw_2, g'_2, go_2 \rangle$

   Finally, $\mathcal{R}ws = \{rw_1, rw_2\}$

5 Strength calculation

The strength of a reward is mainly based on the "value" that the rewardable goal has for the opponent. On the other hand, the credibility the proponent has in the face of his opponent(s) regarding his ability to fulfill his rewards is an aspect that also influences the strength calculation.

The strength calculation of a reward depends on:

1. The goal of the opponent go_i **(or rewardable goal):** two aspects are considered:

(a) Its importance: Like in some related works ([1], [3]), we will take into account the importance the rewardable goal has for the opponent.

(b) Its state: Let's recall that we use 0 to denote that a goal is sleeping, 1 to denote that it is active, 2 to denote that it is pursuable, and 3 to denote that it is chosen.

2. Execution credibility level: It is also important that the proponent agent be able to execute its rewards, from the point of view of his opponent. This value (represented in the proponent) reflects what the proponent believes the opponent thinks about the execution level of credibility of the proponent.

Considering these aspects, the formalization of our proposal is defined as follows.

Definition 5.1. (Basic strength of a reward) The basic strength of a reward depends on the importance and the state of the rewardable goal. Let $rw = \langle \mathcal{R}rw, g'_i, go_j \rangle$ be a reward, its the basic strength is obtained applying:

$$ST_{basic}(rw) = \frac{\dfrac{State(go_j)}{num\_states - 1} + Importance(go_j)}{2} \quad (1)$$

Where $num\_states = 4$ is the total number of states that an opponent goal can have and function *State* return the state of the opponent's goal (0=sleeping, 1=active, 2=pursuable, and 3=chosen).

A direct consequence of the above definition is that the value of the basic strength of a reward is a real value between 0 and 1. Formally:

Property 5.1. Let $rw = \langle \mathcal{R}rw, g'_i, go_j \rangle$ be a reward. $ST_{basic}(rw) \in [0, 1]$, where 0 represents the minimum value and 1 represents the maximum value the basic strength can have.

Proof Since the value of the basic strength is based on the values of the importance and the normalization of the states value of go_j, and this values are of between 0 and 1, then the value of the basic strength is also limited by 0 and 1.

When the proponent agent constructs a set of rewards only for one opponent, the value of the basic strength is enough to choose the reward that will be sent. Nevertheless, a more exact value can be obtained if the execution credibility level is also considered. This aspect is even more important when the proponent agent generates rewards for more than one opponent as it will let him know which opponent may be convinced more quickly when faced with one of his rewards.

Definition 5.2. (Combined strength of a reward) The combined strength of a reward depends on the basic strength of the reward and the execution credibility level of the proponent. Let $rw = \langle rrw_k, g'_i, go_j \rangle$ be a reward and $op_n \in \mathcal{O}p$ the opponent whose rewardable goal is go_j. The combined strength of rw is obtained applying:

$$ST_{comb}(rw) = ST_{basic}(rw) \times Level\_Exec_{rw}(op_n) \quad (2)$$

Property 5.2. The maximum value of the combined strength of a reward is at most the value of its basic strength: $ST_{comb}(rw) \in [0, ST_{basic}(rw)]$.

Example 5.1. Let us continue with example 4.1:

$State(go_1) = 1$, $Importance(go_1) = 0.8$

$State(go_2) = 3$, $Importance(go_2) = 0.6$

$Level\_Exec_{rw}(employee) = 1$ is the execution credibility level of agent *boss* from the point of view of agent *employee*.

Applying equation 1, the basic strengths of rewards in $\mathcal{R}ws$ are $ST_{basic}(rw_1) = 0.57$ and $ST_{basic}(rw_2) = 0.8$. Since agent *boss* generated rewards only for one opponent, he can choose the strongest one without calculating the combined strengths, even more when in this case the values of the combined strengths are the same of the basic strengths. Therefore, *boss* would sends rw_2 because it is the strongest reward.

Example 5.2. Let us suppose that *boss* has another opponent: *employee*2. Let us also suppose that *boss* has generated two rewards for *employee*2 with the following basic strengths: $ST_{basic}(rw_3) = 0.85$, $ST_{basic}(rw_4) = 0.55$, with $Level\_Exec_{rw}(employee2) = 0.7$.

Taking into consideration only the basic strengths, the strongest reward is $rw_3 = 0.85$. This means that *boss* could send this reward to *employee*2 as it seems that it would be more effective than sending a reward to *employee*. However, the execution credibility level of *boss* (from the point of view of *employee*2) is lower than the execution credibility level from the point of view of *employee*. Thus, the combined strengths (equation 2) have the same values of the basic strengths for *employee*, but different values for *employee*2:

$ST_{comb}(rw_1) = 0.57 \times 1 = 0.57$

$ST_{comb}(rw_2) = 0.8 \times 1 = 0.8$

$ST_{comb}(rw_3) = 0.85 \times 0.7 = 0.6$

$ST_{comb}(rw_4) = 0.55 \times 0.7 = 0.41$

Therefore, the best option for *boss* is to sent reward rw_2 to his opponent *employee*.

6 Related works

Servapali et al. [9] propose a model where the rhetorical strength of rewards varies during the negotiation depending on the environmental conditions. For calculating the strength value, it is taken into account a set of world states an agent can be carried to by using a certain reward. The intensity of the strength depends on the desirability of each of these states. For a fair calculation, an average over all possible states is used.

In [1], a formal definition of rewards and an evaluation system are presented. For the evaluation of strength of rewards, the certainty of beliefs that are used for the generation of the reward and the importance of the goal of the opponent are considered. The same authors have other later articles about rhetorical arguments ([2], [3]). In these works, the calculation of strength of rewards is done always by taking into account the two criteria previously mentioned. For our proposal, we made a deeper analysis of the components of a reward and defined new criteria for calculating its strength. Another difference is in relation to the reward rules, meanwhile in Amgoud and Prade [1] it is part of the knowledge base since the beginning, in our work it is constructed from an own goal of the proponent agent and the goals of the opponent, giving more flexibility for the generation of rewards.

7 Conclusions and future works

This work makes a deeper analysis of the components of a reward and considers new criteria for the calculation of their strength. Using these criteria, two forms for calculating the strength of a reward were proposed: the basic strength and the combined strength calculation. These two different ways of calculus is one of the advantages of our proposal as, depending on the need of the situation, the proponent can use either the basic or the combined equation.

We also presented a process for rewards construction and an agent architecture based on the goal processing model of Castelfranchi and Paglieri. We believe that our proposed process gives more flexibility for the generation of rewards as the rewards rules are generated dynamically from the set goals of the opponent and by applying the *EvalRecomp* function, which evaluates the rewardable goal to return an adequate reward.

As future works, we want to analyse the calculation of strength of rewards from the point of view of the opponent. Besides the importance and the state of the rewardable goal, the utility of the reward for the opponent may be considered.

We also want to work on experience-based calculation; this would be done after the proponent receives the answer of the opponent, which can be positive or negative. When it is negative, a recalculation of the initial strength should be done.

References

[1] Leila Amgoud and Henri Prade. Threat, reward and explanatory arguments: generation and evaluation. In Proceedings of the ECAI Workshop on Computational Models of Natural Argument, pages 7376, 2004.

[2] Leila Amgoud and Henri Prade. Handling threats, rewards, and explanatory arguments in a unified setting. International journal of intelligent systems, 20(12):11951218, 2005.

[3] Leila Amgoud and Henri Prade. Formal handling of threats and rewards in a negotiation dialogue. In Argumentation in Multi-Agent Systems, pages 88103. Springer, 2006.

[4] Michael Bratman. Intention, plans, and practical reason. 1987.

[5] Cristiano Castelfranchi and Fabio Paglieri. The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. Synthese, 155(2):237263, 2007.

[6] Cristiano Castelfranchi. Reasons: Belief support and goal dynamics. Mathware & Soft Computing, 3(1-2):233247, 2008.

[7] Christos Hadjinikolis, Yiannis Siantos, Sanjay Modgil, Elizabeth Black, and Peter McBurney. Opponent modelling in persuasion dialogues. In IJCAI, 2013.

[8] Sarit Kraus, Katia Sycara, and Amir Evenchik. Reaching agreements through argumentation: a logical model and implementation. Artificial Intelligence, 104(1):169, 1998.

[9] Sarvapali D Ramchurn, Nicholas R Jennings, and Carles Sierra. Persuasive negotiation for autonomous agents: A rhetorical approach. 2003.

[10] Tjitze Rienstra, Matthias Thimm, and Nir Oren. Opponent models with uncertainty for strategic argumentation. In IJCAI, 2013.

[11] Katia P Sycara. Persuasive argumentation in negotiation. Theory and decision, 28(3):203 242, 1990.