PUC

ISSN 0103-9741

Monografias em Ciência da Computação
n° 16/08

Computing Reputation in the Art Context: Agent Design to Handle Negotiation Challenges

Andrew Diniz da Costa
Carlos José Pereira de Lucena
Viviane Torres da Silva
Sérgio Ciglione Azevedo
Fábio Azevedo Soares

Departamento de Informática

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
RUA MARQUÊS DE SÃO VICENTE, 225 - CEP 22451-900
RIO DE JANEIRO - BRASIL

Editor: Prof. Carlos José Pereira de Lucena April, 2008

Computing Reputation in the Art Context: Agent Design to Handle Negotiation Challenges

Andrew Diniz da Costa, Carlos José Pereira de Lucena, Viviane Torres da Silva[1]
Sérgio Ciglione Azevedo, Fábio Azevedo Soares

[1]Departamento de Sistemas Informáticos y Computación Universidad Complutense de Madrid, Madrid, Spain

{acosta, lucena, sazevedo}@inf.puc-rio.br, viviane@fdi.ucm.es, fazevedo@ele.inf.puc-rio.br

Abstract. Multi-Agent Systems are societies in which autonomous and heterogeneous entities can work together to achieve similar or different goals. However, it is often a difficult task to perceive when an agent can be trusted to perform a specific negotiation. In other words, to establish its reputation. For that purpose, a number of models and strategies have been studied and proposed in the literature. The AAMAS Agent Reputation Trust (ART) Testbed competition has been created to make it possible to compare different strategies. This is achieved through an environment for agent-based simulation games. In the second edition of ART that took place in 2007, new models and strategies emerged that incorporated important research contributions to the field. In this paper, we provide the description of the main set of concerns that we took into consideration to create a competitive strategy for the second version of ART. We go beyond that by providing explanations for tested good strategies that permit comparisons with the 2006 winning strategy.

Keywords: Multi-Agent System, Reputation, Trust and ART-Testbed.

In charge of publications:

Rosane Teles Lins Castilho
Assessoria de Biblioteca, Documentação e Informação
PUC-Rio Departamento de Informática
Rua Marquês de São Vicente, 225 - Gávea
22451-900 Rio de Janeiro RJ Brasil
Tel. +55 21 3527-1516 Fax: +55 21 3527-1530
E-mail: bib-di@inf.puc-rio.br
Web site: http://bib-di.inf.puc-rio.br/techreports/

Table of Contents

1 Introduction

In Multi-Agents Systems [1, 2] there are societies with heterogeneous agents created by different developers with similar or different goals. To achieve them, agents have behaviors that determine the steps of their executions.

Situations such as the need to communication with another agent, sometimes determines the success or fail of the execution. Two concepts are directly related to that: trust and reputation. For the purpose of the present work we assume as in Koogan [3] that trust can be seen as the amount of faith one is willing to assign to someone else's integrity, while reputation is a social notion associated with the observed trustfulness of its individuals.

To perceive which agents can be trusted in a particular negotiation process and what reputation should be attributed to each agent is very difficult to determine. The AAMAS Agent Reputation Trust (ART) Testbed [4, 6, 7, 8, 11] was created to search for a solution to this situation. The goal has been to stimulate the academic community to propose solutions that determine the reputation and trust of the agents, besides allowing comparison of the different strategies [9, 10, 12].

Initially, we offer an overview of the ART-Testbed (Section 2), followed by a description about the challenge of creating a powerful strategy for the second edition of the competition (Section 3). In Section 4 we describe the strategy of the Ze Carioca Les agent (finalist 2007). It proposes solutions to some of the previously referred challenges. In Section 5 a comparison is made between the Ze Carioca Les and the IAM (winner 2006). In Section 6 the main tests developed during the creation of the Ze Carioca Les are described, and in Section 7 a performance evaluation is followed by a conclusion.

2 Art-Testbed Overview

The ART-Testbed competition was created to define an environment of tests for software agents that use the reputation concept. It simulates a business environment in which clients buy opinions about paintings. Each agent on the game is a service provider (appraiser of paintings) responsible for selling its opinions when requested. Figure 1 [6] illustrates the idea of the domain.

Each painting has an era, i.e., a category of related paintings. At the beginning of each game, a simulator provided by the competition, which is the execution environment, randomly supplies the knowledge's degrees for each "era". During the game, the degrees can be changed as new values are attributed by the simulator. In some situations an appraiser can receive paintings from an era, which do not have good degrees. The agent can, thus, exchange information with other agents during the same game and try to perform good appraisals. The possible values for each era are in the interval [0,1]. These values represent the error's standard deviation of an agent with respect to the appraisal of the paintings that correspond to an era. If the value is near zero, the appraiser has more chances to make good evaluations. If it is near one, there are more chances to make bad evaluations.

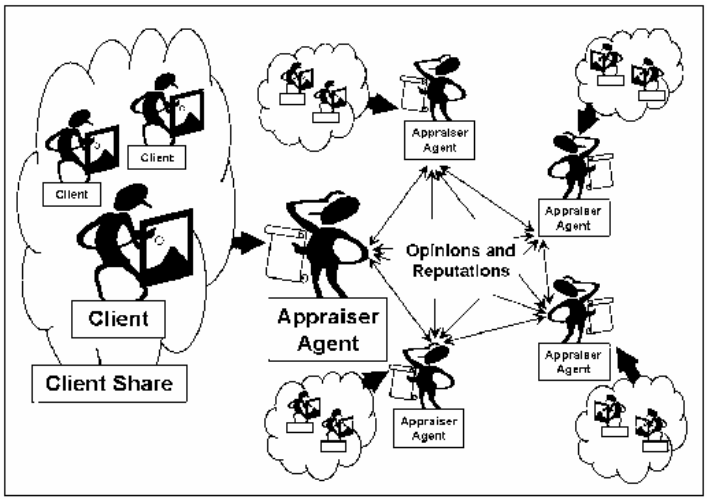

Figure 1. Domain ART-Testbed

The simulator offers two possible transaction protocols among appraiser agents: reputation and opinion. The reputation protocol, shown in Figure 2 [6], allows an agent A to request the reputation of an agent C in relation to an agent B. The information provided by agent B to the agent A is a value in [0,1]. In the papers of the competition [6, 7, 8, 9, 10], a definition is needed for what represents a bad or a good reputation. Even in the case that the requester and the provider consider, for example, zero as a bad reputation, it is difficult to perceive if the information can be trusted.

Another possible transaction is the request for opinions illustrated in Figure 3 [6]. It allows an appraiser to request an opinion about a painting from another agent. The agent, which receives the request, can reject or accept to inform about its opinion. The same happens with the reputation's transaction. If the agent accepts, its knowledge's degree about the era (Certainty Assessment) is supplied. In turn, the requester needs to decide whether to accept the opinion. In case it accepts, the payment is executed (fixed value determined by competition, $10). The provider can then offer its opinion. Note that the protocol does not guarantee that the information supplied is trustful.

The final appraisals of each agent is created (i) from the opinions provided by other agents and (ii) from the own appraisals performed. Each appraiser must specify the weight that the opinions will have in each final appraisal for each era. In addition, the agent must be aware of the time dedicated to analyze the paintings. More time for analysis means more money spent.

Each game has a set of rounds. At the end of each round, which appraiser obtained the best appraisals is established. The agent, who obtained the values nearest to the true values, receives more clients and consequently more money in the next round. The agent with the most money at the end of the game is the winner.

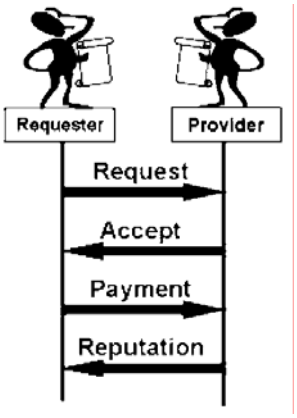

Figure 2. Reputation's transaction

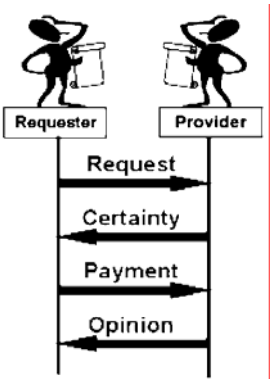

Figure 3. Opinion's transaction

To create an appraiser the class Agent offered by the competition must be extended. In it, there are nine abstract methods that should be implemented by the application classes that extend it. In these methods the intelligence of each appraiser is represented. As follows, there is a brief description of each method related with the role played by the appraiser in the transactions (reputation and opinion) as previously mentioned.

- *prepareReputationRequests*: The requester requests the reputation of some appraiser.
- *prepareReputationAcceptsAndDeclines*: The provider informs if the requested reputations are accepted or rejected.
- *prepareReputationReplies*: The provider supplies the requested reputations.
- *prepareOpinionRequests*: The requester requests opinions to other appraisers about the paintings that it should evaluate.
- *prepareOpinionCertainties*: The provider supplies the degrees of the eras related with the requested paintings by other agents.
- *prepareOpinionRequestConfirmations*: The requester informs if accept or reject some requested opinion from the certainties assessment provided.
- *prepareOpinionCreationOrders*: The appraiser defines the spent time to analyze the paintings requested by the clients and by other agents.

- *prepareOpinionProviderWeights*: The appraiser informs to the simulator the weights of the opinions provided by other agents and of its opinions in order to perform the appraisals.
- *prepareOpinionReplies*: The provider provides the opinions to the requesters.

3 Difficulties of Creating a Competitive Strategy

In this section, we analyze the main points that deserved our attention during the creation of our strategy for the ART Testbed. The challenges we have faced are presented as follows.

1 **Requesting opinions in the first round:** if the game is on its first round, agents do not have any information about each other and do not know who the trustworthy agents are. Agents can decide to request opinions to all agents, to a group of agents or to none at all. Choosing the first option, depending on the quantity of agents, the appraiser spends a lot of money. If the second possibility is chosen, the agent will be limited to the set of opinions provided by the members of the group. In the case the agent decides not to request opinions to any agent, third possibility, it will be impossible to determine if an agent can help in a future appraisal since its opinion is not known. Therefore, to determine the strategy to follow in the first round is fundamental to succeed in the game.

2 **Determining trustworthy and useful agents:** since agents can lie, it is important to determine who the trustworthy agents are. If the agent has received opinions only in one round, it cannot conclude who the trustworthy agents are. To determine if an agent is trustworthy and useful, it is necessary several rounds. One of the challenges is to determine how many rounds are necessary to adequate evaluate the behavior of the appraisers. Another challenge refers to determine if the information usually supplied by an appraiser can be useful to help the corresponding appraisal. In order to overcome such challenges is necessary to find out if the agent has a constant behavior. But note that agents can change their behavior from one round to another due to many reasons. For instance, agents can receive new degrees in a given round provoking a change on their behavior or they can simply change their strategies from one round to another. Therefore, a different behavior from one round to another does not always characterize untrustworthy or unusable opinion.

3 **Providing opinions about paintings:** this point has two implications. The first implication regards earning money. When an appraiser renders an opinion, it earns $10, which is good for the competitor. Another implication is the risk associated with the possibility of helping other agents leading one of them to be the winner. Therefore it is a challenge to decide when to provide opinions to other agents.

4 **Providing reputations:** same as above, when a reputation is supplied, the provider earns money ($1). However, depending on the information supplied, it can help opposing agent. Thus, it is also important to define when to provide reputations.

5 **Requesting reputations:** two problems are related to this point. When requesting reputation, the agent loose money. Although it is not a lot of money for each request ($1), the agent may not request reputation to every other agent. In addition, the game rules neither specify the semantics of a reputation value nor how those values are evaluated. It only states that it is a number in the range [0-1].

6 **Defining weights:** to determine which weights each agent opinion should receive is complicated since it does not only depend on the knowledge's degree in each era but also on the degree of trust defined by the agent that has received the opinions. In the first round agents do not know about the character of others, therefore, it is a challenge to define the weight to be applied to the opinion of an unknown agent with a high knowledge degree in the era of the painting being evaluated. In the other rounds some agents can change their behaviors and, thus, it is a challenge to apply a high weight to the opinion of an expert in the era if its degree of trust may have changed.

7 **Determining the time used to analyze a painting:** A big challenge is to determine the time to analyze the paintings requested by the clients and by the other agents since more time the agent uses to analyze a painting more money it will spend. In addition, to take a long time evaluating a painting does not guarantee that the evaluation will be good, because it is mostly influenced by the degree of expertise the appraiser has in the eras of the paintings.

4 A Competitive Strategy for Computing Reputation

The agent that adopted our strategy and that participated in the ART Testbed 2007 was called Ze Carioca Les. The agent was required to create algorithms that dealt with the problems referred to in the previous section. To deal with them, the following modules were defined: statistic, decision and controller. In Figure 4 the conceptual model is presented. The controller module controls the access to the other modules. The statistics module is responsible for performing a set of analyses through the use of opinion's transactions, while the decision module defines, for example, which agents are the best from which to request opinions.

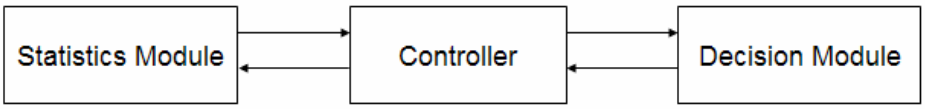

Figure 4. Conceptual Model of the Zé Carioca Les Agent

4.1 Statistics Module

When an appraiser receives bad era's degrees, it is necessary to supplement its knowledge in order to evaluate paintings included in those eras. One way to achieve this is by requesting opinions about paintings to other agents. But, who are the agents that can provide useful opinions? Agents can lie about their degrees and can also provide untrustworthy opinions.

As a consequence, there are three points to be analyzed: whether an agent can help in some era (i.e., the percentage of correctness of the opinions provided by the agent about paintings in the era), the trustworthiness of each agent (i.e., the degree of trust about the agent's opinions) and the changes in the agent behavior during the rounds.

To perform the above-mentioned analysis, our solution design created an entity that we called Oracle. Oracle consists of a data's repository with opinions provided to the Ze Carioca Les, going beyond the true values of self-assessed paintings. In Figure 5 we present the conceptual model of the Oracle. We can see that for each appraiser the a-

gent stores the received opinions. There will be a set related to each era that groups the opinions about painting of such era provided by a given agent. Each set is responsible for computing the degree of trust and the average of the correctness of the opinions provided for the era (see next subsection).

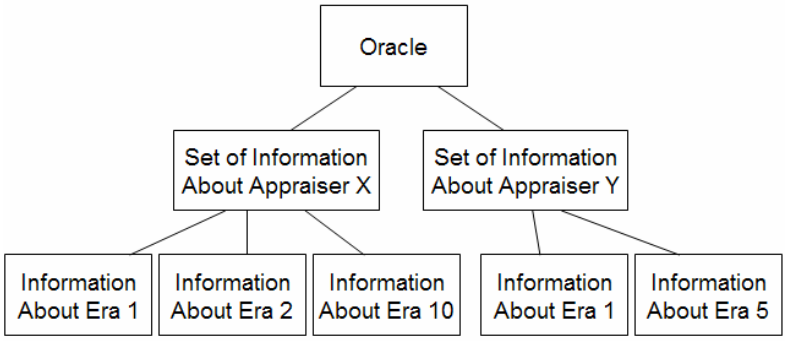

Figure 5. Conceptual model of the statistics module

4.1.1 Degree of Trust and Percentage of Correctness

Using the data stored in Oracle, it is possible to calculate the standard deviation (degree of trust), and the average of the estimates (percentage of correctness) of the opinions provided for each agent in each era. To calculate the average of the estimates the formula below was used:

$$average = \frac{\sum_{0<i<=n\_estimates} \frac{op\_supplied_i}{true\_value}}{n\_estimates} \quad (1)$$

where: op_supplied is the opinion supplied by an agent about some painting; true_value is the real value of the painting, which is informed by the simulator of the competition, and n_estimates is equal to the number of opinions the agent has received for a given era.

To calculate the degree of trust, we have used an API called Commons-Math [13] from Apache [14], which offers a method to calculate the standard deviation. The calculation takes place in two steps:

(1) The first step involves to calculate each estimate for each round related to each era using the following formula:

$$estimate = op\_supplied / true\_value \quad (2)$$

(2) With all the estimates calculated until the current round, the method *getStandardDeviation()* of the Commons-Math API is responsible of calculating the standard deviation. Therefore, the degree of trust is met by the formula:

$$trust = 1 - getStandardDeviation () \quad (3)$$

At the beginning of each round, each agent receives from the system the value of the opinion supplied by other agents in the previous rounds, besides the true values of the corresponding paintings. With these data, Oracle updates the trust and the estimates of the agents.

4.1.2 Changes in the Agent Behavior

A strategy was created to verify changes in the behavior of the agents while providing opinions to other agents. Such strategy compares the standard deviation and the average of estimates reached in the last two rounds related to an appraiser with all the previous sessions calculated. If the difference between the values calculated is significant, only the values from the last two rounds are maintained. Otherwise all the values are maintained until the current round. If the same agent continues to provide opinions in two subsequent rounds, a new analysis is performed.

For better understanding such analysis, imagine that our agent receives opinions from agent X from round zero until round 3. To verify if an agent changed its behavior, the degree of trust and the average of estimates in the two last rounds (2 and 3) are compared with the rounds from zero until 3. If a significant difference between the calculations is perceived, only the values in the last two rounds are maintained (Figure 6).

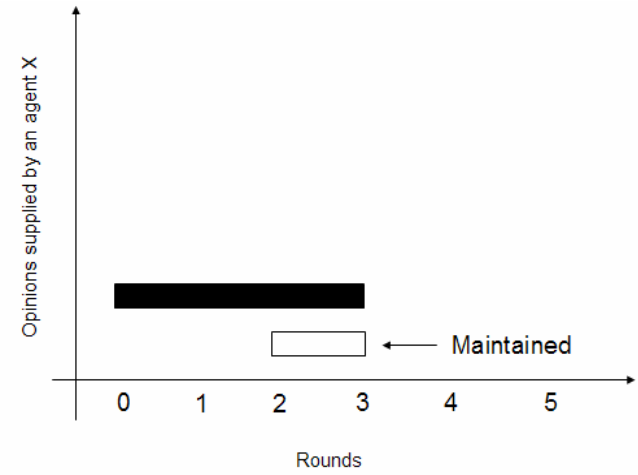

Figure 6. Detecting change of behavior

Suppose that the Ze Carioca Les continues receiving opinions from the same agent X, and even that the behavior of the agent had changed in the rounds 2 and 3. In round 5 a new analysis is performed. This time, rounds 4 and 5 are compared with rounds 2 through 5. Following the previous approach, if a big difference between the standard deviations or the average of estimates during the rounds was perceived, the values maintained are those from rounds 4 and 5. However, in the present example, there was not a big difference, and the values maintained are those from round 2 through 5 (Figure 7).

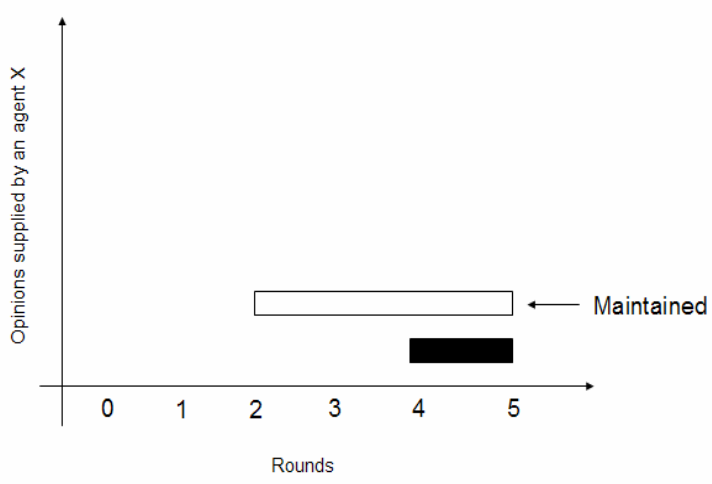

Figure 7. Without the detection of change behavior

4.2 Decision Module

The decision module provides strategies to the Ze Carioca Les to decide when opinions should be requested and provided, to establish how much time the agent should spend on the analyzes of a painting and to determine the weight to be stated to the opinions it has received.

4.2.1 Transaction of Opinions

This section focuses on presenting two different strategies. The first one determines the moment in which a request should be made to another agent while the second decides whether our agent should provide opinions to third party agents and what kind of information should be provided, true or false values.

The Ze Carioca Les agent followed the policy of requesting opinions from other appraisers in the moments when it had bad knowledge's degree in some era. In other words, value higher than 0.4 (degree reached after a set of tests).

Remember that when an agent requests an opinion from another agent, and this agent accepts to inform it, the requester must provide the payment. For this reason, one more decision had to be made: to how many agents the Ze Carioca Les should request opinions to not spend too much money. After several tests, we realized that if our agent decided to request opinions from all appraisers in every round, a good amount of money would unnecessarily be spent, especially when the number of agents in the game becomes very large. To resolve this situation, it has been decided to perform transactions with only 20% of the agents in each game. If the number of agents is lower or equal to five, then all agents will receive requests. The 20% was determined from tests with similar amounts of agents that would participate in the competition. The amount of participants in each game of the competition was mentioned in the web site of the competition (around 150 agents).

After the amount of agents is determined it is necessary to choose the agents to be requested. In the first two rounds, the agents are randomly chosen. After the selection,

the opinions of the agents are obtained. Note that such strategy is only used to evaluate the behavior of the agents trying to determine which ones are trustworthy and useful (help the appraisals of the Ze Carioca Les). The opinions received in these rounds are not used while evaluating the paintings since the appraisers' profiles are still being evaluated.

After the second round, the Oracle entity is employed to verify if one of the agents can help the Ze Carioca Les in its appraisals. To determine this fact, two pieces of information are analyzed (as already mentioned in Section 4.1): the average of the estimates using the "opinions of each agent" with the "true value of the paintings related to the opinions," and the degree of trust calculated through the standard deviation of the opinions provided.

With the average of the estimates and the degree of trust of the eras, it is possible to verify whether each appraiser selected can be useful to the Ze Carioca Les agent. For an appraiser to be considered trustful, it must have a trust (equation (3)) higher or equal than 0.85 (value obtained through tests), and to have a good estimate (equation (2)), i.e., its value should be higher or equal than the degree obtained by the formula below:

$$degree = 1 - degreeEraXZeCariocaLes \qquad (4)$$

As the degree of knowledge about each era represents the probability of obtaining a bad value in an appraisal, by subtracting one of the knowledge adopted (degreeEraXZeCariocaLes), we obtain the probability of success. If the estimate reached is higher than the degree of success, then this estimate is considered a good value.

An important implication takes place when our agent has a good average of the estimates associated with a bad trust, or a good trust associated with a bad average. When one situation like that happens, the agent is discarded. It is considered a good appraiser only when it is trustworthy and useful (provides a good average of estimates).

When there are agents who can help Ze Carioca Les, then requests continue to be made only of those that can help by maintaining the group of appraisers selected. If at some point in time it is perceived that no agents can help in the appraisals, then the collection with the names of the agents is randomly organized again. The n first names in the list provided by the simulator are selected, and the old values in the Oracle are deleted, to make room for the new values generated to the selected agents. After the new selection, the agents' behaviors are analyzed in the following two consecutive rounds to determine from which appraisers the Ze Carioca Les will continue to request opinions. If no agent is capable of helping our agent, then the set of names of the agents is randomly organized again and the analysis continues.

The strategy adopted by Ze Carioca Les when it is requested to provide opinions about a painting is to always accept the requests. Its goal is to accumulate how much money as possible. The agent does not inform the correct value of its degree of knowledge in the era to be evaluated. On the contrary, the agent randomly provides a value in the range [0.7, 1], simulating that it knows very well the era. Such behavior will probably stimulate the requester to keep on executing the protocol and pay for the opinion to be provided.

Although this strategy allows our agent to earn money, it can be a high risk to use it. The agents can perceive that the information being provided is untrustful information

and can stop to request opinions. Otherwise, it is not ease to perceive that the information being provided is not trust because our agent sometimes spends some time analyzing the requested paintings (as described in sub-section 4.2.2). Due to the time being spent, the error of the analysis is reduced, even when the knowledge of the agent is different from the knowledge provided.

4.2.2 Analysis Time

When a client requests a painting for evaluation, the Ze Carioca Les agent must decide the amount of time it will take to analyze it. As stated before in Section 3, more time for analysis means more money spent. Moreover, every appraisal has a limit "best evaluation" define by the knowledge degree in the era of the painting being evaluated.

After a set of tests, it was possible to establish that to spend money (time) for analyzing the paintings is very important, mainly in the cases of high knowledge degree agents. In the moments that an agent has a bad knowledge, it is also important to spend some money to generate a less-bad evaluation. However, in these cases, the probability of generating bad evaluations is big and, therefore, the amount of money to be spent should not be so high.

It is also important to consider the time to be spent to analyze the paintings requested by others; in other words, to generate an opinion. Since the agent spends a good amount of time (money) evaluating the paintings requested by its clients, we adopted the policy of reducing the expenses. Therefore, the opinions generated to other appraisers do not take a long time of analysis. On the other hand, it is important to take some time generating those opinions in order to not generate really bad opinions. If our agent only provides bad opinions, no agent would probably request further opinions. Thus, we decided to spend very little money, i.e, $0 or $1, to generate a not so bad opinion. To specify how much money to spend a random selection is performed between -5 and 5 using a Gaussian formula offered by JSDK 5.0 [15]. If the value is negative then no money is spent on a requested appraisal; otherwise $1 is spent.

4.2.3 Weights

In order to provide a final evaluation of a painting to the simulator, each agent must inform the weight attributed to its appraisal and the weights to be attributed to the opinions provided by other agents about the same painting. The formula below is used by the simulator of the competition to put together all the opinions informed by an agent about a painting:

$$\text{final\_appraisal} = \frac{\sum_i (w_i \,.\, p_i)}{\sum_i (w_i)} \qquad (5)$$

where; w_i and p_i are the weight of the appraiser i in relation to another agent, and the opinion provided by the same agent i, respectively.

In the two first rounds of each game where the opinions of the other appraisers are still not used, Ze Carioca Les attributes the weight 1.0 to the eras where its knowledge is lower than 0.7. To the other eras the weight is attributed following the equation below:

$$weight = 1.1 - degreeEraXZeCarioca. \quad (6)$$

The possible weights in this case are 0.4, 0.3, 0.2 and 0.1, because the possible values of "degreeEraXZeCarioca" are 0.7, 0.8, 0.9 and 1.0 respectively. The better is the knowledge degree of the agent in an era the higher is the weight.

By not considering the opinions provided by other agents in the first two rounds, i.e., by giving zero to the weights of those opinions, Ze Carioca Les obtained more money in the first and second rounds of the simulated games against the agents of 2006. More details are described in Section 6.

From the third round, it is verified if there are agents that can help Ze Carioca. If there isn't, the highest weight (1.0) is attributed to the Ze Carioca opinion, while the other appraisers receive a weight of zero. However, if it has been detected that one or more agents can help in its appraisals, the average of the estimates of the corresponding appraisals in an era provided by Oracle is used to calculate the weight of such appraisals. The agents, which cannot help continue receiving zero, while the others receive the weight based in the following formula:

$$weight = average * trust \quad (7)$$

where; "average" represents the average of estimates (equation (1)); and "trust" informs the trustfulness degree of the corresponding agent (equation(3)). To verify if some agent really is useful to Ze Carioca, the equation (4) presented in Section 4.2.1 is used. Every value represented here is based on a set of tests performed using the ART-Testbed simulators of the competitions 2006 and 2007.

5 Comparison of Zé Carioca LES and IAM (Winner 2006)

To perform the comparison between the Ze Carioca Les and IAM (winner 2006) [5], three important points were analyzed: verification of the behavior of the appraisers, detection of good ways to earn more money, and creation of appraisals. The first is about how to detect when an agent is lying, and when one is a good provider of requests for opinions. The second is about how to earn money providing opinions and reputations, and the third is about the time of analysis necessary to make a good appraisal with the knowledge received from the simulator.

Another point referred in this section is about an important characteristic of the Ze Carioca Les and which allows in the beginning of the games to earn more money with respect to the other appraisers of the competition 2006.

5.1 Analyzing Appraisers

The IAM agent, winner of the ART-Testbed 2006, applied a set of very interesting strategies to obtain the first place. The central idea is based on three main parts: (i) a lie detector that detects if an agent is malicious, (ii) a variance estimator that estimates how much variance there is in the error of the opinions of other agents, (iii) and a module responsible for calculating the weights of each agent.

First, the IAM uses a Bayesian analysis to estimate an agent's variance, to determine the agent's opinions. Subsequently, it calculates a lower boundary on the probability that an agent is lying about its appraisals and uses this to discard the opinions that probably are malicious. Using the estimated variances, an optimal method is derived to calculate the weights.

For its part, Ze Carioca Les, as already mentioned in Section 4, performs estimates using the opinions supplied by other appraisers, and the true values to define if an agent is lying. Using the estimates of the last two rounds, the agent compares such values with the other estimates through to the current round. If the difference is large, our agent concludes that the respective agent is lying because it is not maintaining a constant behavior. Therefore, when a change of behavior is detected, the values maintained are those of the last two rounds. However, if no change was detected, the values until the current round continue to be used for comparisons.

The lower boundary calculated by IAM to determine the probability that an agent is lying about the appraisals was represented by Ze Carioca Les through a fixed value met after a set of tests. It determines if the standard deviation is good or bad.

A big concern when we created the Ze Carioca agent was to determine if an agent was useful and provided trustworthy appraisals. For us, the agents we were interested in following up with a request for an opinion had to be trustworthy and useful. If an agent was not useful (did not have good knowledge) then it could not help Ze Carioca. The same was true when an agent was not trustworthy, because to believe in opinions provided through inconstant behavior could induce our agent to make bad appraisals. Below we present a table, which shows the possible combinations between the good (white) and the bad (black) values.

Useful and trust		
	not useful	useful
not trust		
trust		

Table 1. Combinations with usefulness and trustfulness

The strategies applied by Ze Carioca and IAM follow the same philosophy to try to detect lies, and to decide which agents provide trustworthy opinions. Tests performed with the finalists of 2006 prove that both agents have strong strategies. In Section 6, there is a better explanation of the tests that were performed.

5.2 Earning Money by Providing Opinion and Reputation

An important point has to do with the importance of earning money by providing opinions and reputations to other agents. Ze Carioca Les and the IAM have their respective strategies and to maximize their income they always provide opinions and reputation values.

The IAM approach provides an honest and reliable service to the agents that maintain a good business relationship. To the other agents, it provides a low quality service or cheats other agents that might initiate retaliatory behavior against them. In more detail, the IAM agent always spends $4 to generate good appraisals for others. If it realizes the existence of a cheating agent that provides false opinions or earns opinion fees

without providing an opinion, it will spend a fractional amount of time ($0.01) on all future opinions requested by that agent.

As the competition does not define the semantic for reputation values, except that their range is [0-1], the IAM provides the reputation value as an estimated variance of an agent to the agents that maintain a good business relationship. However, for the cheating agents, it randomly produces reputation values in [0,1].

Different than IAM, Ze Carioca, does not make a distinction between the agents. For the opinions, it obtains a random number between -5 and 5 using a Gaussian Formula offered by JSDK 5.0 [16]. If the value is negative, no time is spent analyzing the painting requested; otherwise, if the value is positive, our agent spends a little time ($1). For the reputations, as the competition does not define the semantics for reputation values, we inform a random value between [0.7,1].

We decided to adopt the random strategy because we simultaneously want to earn money without helping other appraisers more than necessary. Had we decided to help only the agents with whom we have a good business relationship (IAM strategy), they could receive more help than we wanted to provide them. For this reason, we decided to eventually use a reasonable value independent of the agent.

5.3 Generating the Own Appraisal

At the beginning of each game, the agents receive degrees of knowledge associated with the eras. With the degrees it is possible to perform appraisals, which are related to the concept of time of analysis (time spent by an agent to analyze a painting).

The Ze Carioca Les and the IAM agents spend a considerable amount of time analyzing their paintings, trying to use their own knowledge as much as possible. The strategy of the IAM is to spend $4 every round to create its own appraisals, while the Ze Carioca Les agent spends $10. We concluded after performing a set of tests that to spend time analyzing its own paintings allows for the generation of good appraisals, especially when the degree is good.

5.4 To Begin Winning

In any competition, to begin winning a game is an important step towards the victory. With this concept in mind, Ze Carioca Les uses a strategy tested several times with appraisers of the 2006 competition. These tests showed that our agent was capable of earning a good amount of money in the two firsts rounds in every game performed.

The secret, as has already been explained in Section 4.2.3, was to attribute a large weight to the eras, for which our agent received a good amount of knowledge, besides spending time analyzing the related paintings. In the eras with bad opinions, neither the weights nor the analysis time should be large. This strategy was adopted because, in the first two rounds, Ze Carioca does not know the other appraisers. For that reason, the information that proved to be the most secure was the agent's own knowledge.

In 10 simulated games with the four best participants of 2006 (IAM, Neil, Frost and Sabatini), Ze Carioca earned on average $3,000 more than the second place finisher. A simulated game is illustrated in Figure 8. We can see that the Ze Carioca agent earned approximately $7,000 through the second round, while the second place finisher earned $3,400, and IAM, which came in third, earned $2,500.

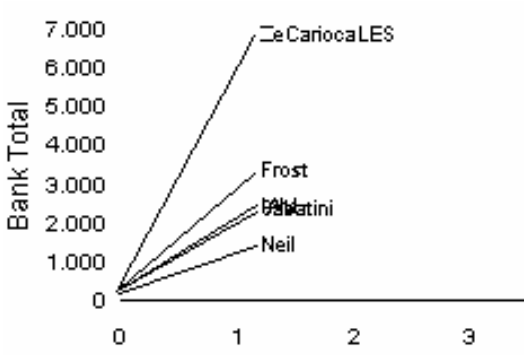

Figure 8. Firsts rounds in a game simulated

6 Tests Performed

The test with Ze Carioca Les was divided into two parts: the first involved the simulator and the agents of the 2006 competition available on the official web site [4]. The second used the simulator of 2007 with dummy agents offered by the competition.

6.1 Tests with Participants of 2006

When the Ze Carioca Les began to be developed, the simulator for the 2007 competition was not available. For that reason, the alternative was to use the simulator of the previous edition and to test our agent with the 2006 competitors, which are offered on the official web site [4].

When the tests began, we observed that the five first competitors had implemented really smart agents, and only after very well understanding each concept offered by the simulator, we were able to create a good agent. After several tests, we decided to focus the games on the top four 2006 agents: IAM, Neil, Frost and Sabatini.

In every game performed with the four above mentioned agents, during the two first rounds, on average, Ze Carioca earned $3,000 more than the first runner-ups (Figure 5). Another important point is that our agent managed to remain in the lead in a set of games and to increase the difference from the other agents. These points were really important to obtain good test performance.

Table 2 shows the performance of our agent in relation to the best four in 2006. Each simulated game involved five appraisers and 20 rounds. We can see that the Ze Carioca agent posted good performance, coming in third, in the worst two cases. Moreover, it achieved more first places (five times) than any another agent, including the winner IAM agent.

10 games simulated – 20 rounds					
Agents	1st	2nd	3rd	4th	5th
Ze Carioca Les	5	3	2		
IAM	4	2	3	1	
Neil			3	3	4
Frost	1	4	1	3	1
Sabatini		1	1	3	5

Table 2. Games simulated and the places got by each agent

One of the main reasons for the different places obtained by Ze Carioca in the ten simulated games had to do with the different degrees attributed by the simulator to their eras. Therefore, one of the first considerations was to achieve a good performance in any game.

Another important point deals with the strategies implemented by an agent that can increase or reduce the performance of some other agent. Taking this into consideration we took care to create strategies which did not have a critical performance due to other agent's strategy.

To understand the behavior of the Ze Carioca in relation with the other agents, one of the games previously mentioned is analyzed in table 3. There, our agent got the first place, followed by Frost, Neil, IAM and Sabatini.

Agents	**Bank Balance**
ZeCario-caLes	43472.19
Frost	40827.94
Neil	36845.5
IAM	36767.4
Sabatini	19962.61

Table 3. Classification of the agents

In Table 4 we can see the opinion's transactions performed during a game. The Ze Carioca received more opinions from the Neil agent compared with the other agents. For this to happen, the appraiser probably provided good opinions, and it was considered trustful and useful by our agent.

		From				
		Frost	IAM	ZeCarioca Les	Neil	Sabatini
To	Frost		17	470	470	170
	IAM	20		17	31	320
	ZeCarioca-Les	69	65		318	65
	Neil	20	20	20		20
	Sabatini	28	68	208	208	

Table 4. Opinion transactions

Another interesting point is that the Frost and Sabatini agents considered the Ze Carioca Les a good agent to get opinions from. Therefore, we can consider the strategy of generating sometimes bad and/or not so bad opinions is a good alternative. Ze Carioca Les was the second agent that provided more opinions, as can be seen in Table 6.

In Table 5 we show the transactions of reputation performed in the game. We can assume that the Ze Carioca Les, Frost, IAM and Sabatini used the same policy of providing reputations, and not to perform requests. The Neil agent is the only which requests reputations.

		From				
		Frost	IAM	ZeCarioca Les	Neil	Sabatini
To	Frost		0	0	0	0
	IAM	0		0	0	0
	ZeCarioca-Les	0	0		0	0
	Neil	40	40	40		40
	Sabatini	0	0	0	0	

Table 5. Reputation transactions

Agents	Opinions sent	Reputation sent
Frost	137	40
IAM	170	40
ZeCariocaLes	715	40
Neil	1027	0
Sabatini	575	40

Table 6. Opinions and Reputations

In summary, the Ze Carioca Les got the first place by providing a set of opinions and reputations, along with requesting opinions to other appraisers when it had a bad degree in some era.

6.2 Tests with the 2007 Simulator

After the 2007 simulator became available, we decided to test the Ze Carioca Les agent with the dummies offered by competition. In every simulated game with different rounds (from 20 until 200), Ze Carioca Les came in first.

In 10 games simulated with 100 rounds and five dummy agents, we can see in Table 7 that the agent came in first. In Table 8 a final result is shown from the ten games simulated, and we can illustrate the big difference earned (more than $300,000) by Ze Carioca in relation to others agents. It is true that the dummies do not have strong intelligence when compared to the finalists of 2006, but since they also participated in the

competition in 2007, it was very important to also compare the performance of our agent with them.

10 games simulated with 100 rounds						
Agents	1st	2nd	3rd	4th	5th	6th
ZeCariocaLes	10					
dummy_1		4	2	2	1	1
dummy_2		2	4	1	2	1
dummy_3		1	2	1	4	2
dummy_4		2	1	4	2	1
dummy_5		1	1	2	1	5

Table 7. Ten games with dummy agents 2007

Agents	Bank Balance
ZeCariocaLes	450240.86
Dummy_5	121583.0
Dummy_4	113129.0
Dummy_2	110111.0
Dummy_1	104803.0
Dummy_3	99571.0

Table 8. Final Result of a game simulated

After the comparative tests with 2006's competitors (Section 6.1) and with the dummies of 2007, the good results led us to believe that our proposed appraiser would be a strong agent in the competition, as it is demonstrated in section 7.

7 Ze Carioca LES' Performance in the ART TESTBED 2007

In the last competition of 2007 [4], there were two phases: a preliminary and a final. The first took place in May 10th and 11th with 16 agents approved by the competition organizers out of the 17 submitted. A total of 13 different institutions participated.

Each game had 100 rounds, besides the presence of eight appraiser agents and 15 more offered by the competition (five "nice" dummies, five "neutral" dummies and five "bad" dummies). For each game, three versions were performed (A, B and C). In game A no expertise (degree of an era) was changed; in game B only one expertise was changed; and in the game C two expertises were changed. The simulator chooses randomly one or two eras related to each agent and changes the degree of expertise that the agent has in such era(s).

The above rules have been applied to verify the behavior of the agents when there are changes in the degrees of the eras. In Table 9 we can see the group Game 2 of the preliminary phase. We can see in game 2A that our agent came in second (behind the Uno agent) while in the games 2B and 2C it was first. With this, we can conclude that our agent had the best intelligence in this group to deal with the situation when an

agent changes the knowledge of the eras. In Table 10 the final classification of the agents is shown.

Game 2	2A	2B	2C
ZeCarioca-Les	494891	348300	399893
Xerxes	188297	172536	178429
Uno	554646	262632	325661
Alatriste	247093	191619	193184
IMM	166370	136732	192779
AgenteVi-cente	209497	248777	249774
Reneil	335904	307724	294809
ArtGente	350937	299268	273604

Table 9. A set of games with the Ze Carioca LES on the preliminary phase

In the final phase performed on May 16th and 17th in the AAMAS conference in Honolulu Hawaii, only the best five agents in the preliminary phase were classified. The finalists performed games with 200 rounds, besides the presence of 15 agents offered by the competition, as in the previous phase.

The final games also followed the idea of three versions for each game. In Table 11 the group of games conducted in the final phase is shown. This table illustrates that the best performance of our agent occurred when only one expertise changed (4B). In Table 12, we can see that Ze Carioca Les came in fifth.

Rank	Agent	Avg Bank Balance
1	IAM2	539377
2	Jam	353700
3	Blizzard	335933
4	ZeCariocaLes	319564
5	Spartan	311777
6	ArtGente	298897
7	Uno	293324
8	Reneil	269905
9	Marmota	264356
10	Novel	229501
11	Alatriste	225276
12	Rex	211467
13	IMM	200440
14	LesMes	183655
15	AgenteVicente	181932
16	Xerxes	148610

Table 10. Average Scores in the

Game4	4A	4B	4C
ZeCariocaLes	638173	643551	542895
IAM2	1184328	1123495	1111632
Blizzard	805501	727027	753285
Spartan	754783	728991	653381
Jam	814354	869641	708821

Table 11. A set of games with the Ze Carioca LES on the final phase

Rank	Agent	Avg Bank Balance
1	IAM2	1107270
2	Jam	743077
3	Blizzard	723971
4	Spartan	674723
5	ZeCariocaLes	578524

Table 12. Average Scores in the Final Round

With the results reached by the Ze Carioca Les, we realized that our strategy is not adequate when expertises of a game are changed. Besides, an important reason to our appraiser to reach the fifth place in the final phase was that few agents had cooperate with our agent in supplying useful opinions in order to help the appraisals of the Ze Carioca.

8 Conclusions

In the present paper we outlined the challenges of creating a powerful strategy for the second edition of the competition. A comparison has been made between the Ze Carioca Les agent and the IAM (winner 2006). Not only the main tests developed during the creation of our agent have been presented but also their performance in the competition.

With the good performance of the Ze Carioca Les agent, two important conclusions can be drawn. The first is about the possibility of some ideas applied in the strategies used by Ze Carioca being reused in different domains. However, it is important to realize that some different domains can require different calculations to verify if an agent is useful and trustful. In other words, depending on the domain, some particular information can be crucial to determine if an agent is honest. The second point is about the trust and useful concepts. We could see that they were really important to the good performance of our agent. In a domain where opinions are necessary, the concepts helped us to understand the behavior of the agents, besides determining which agents were good or bad when it came time to perform requests.

The participation in the competition motivated two new works in progress. One is about the creation of a Multi-Agent System representing the Brazil stock exchange

market and applying ideas of reputation, such as, in the ART-Testbed. With this domain, we want to establish closer communication with industry, and to show how the reputation concept can be useful in more realistic scenarios. The second work is related to the behavior of agents. Our intention is to create framework capable of performing diagnosis on the execution of agents. If an agent does not achieve the desired goal, the framework reasons to understand the motive of the failure and to propose alternative ways to achieve it.

References

[1] Jennings, N. R. and Wooldridge (2000), “M. Agent-oriented software engineering”, In Bradshaw, J. (Ed.) Handbook of Agent Technology, AAAI/MIT Press.

[2] Wooldridge, M. and Jennings (1998), “N. R. Pitfalls of agent-oriented development” Proceedings of the Second International Conference on Autonomous Agents (Agents'98), ACM Press, pp. 385-391.

[3] Koogan, A.,Houaiss, A.: Encyclopedia and Dictionary. Delta Publisher, Rio de Janeiro, 1995.

[4] The Official Web site ART Testbed Team. Agent Reputation and Trust Testbed, http://www.art-testbed.net/, January 2008.

[5] Teacy, W. T. L., T. D. Huynh, R. K. Dash, N. R. Jennings, J. Patel, and M. Luck. 2007. "The ART of IAM: The Winning Strategy for the 2006 Competition," The Workshop on Trust in Agent Societies at The Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2007), Honolulu, Hawaii, USA, May 15, pp. 102-111.

[6] Fullam, K., T. Klos, G. Muller, J. Sabater, A. Schlosser, Z. Topol, K. S. Barber, J. Rosenschein, L. Vercouter, and M. Voss. 2007. The Agent Reputation and Trust (ART) Testbed Game Description (version 2.0), http://www.lips.utexas.edu/art-testbed/pdf/SpecSummary.pdf

[7] Fullam, K., T. Klos, G. Muller, J. Sabater, A. Schlosser, Z. Topol, K. S. Barber, J. Rosenschein, L. Vercouter, and M. Voss. (2005) "A Specification of the Agent Reputation and Trust (ART) Testbed: Experimentation and Competition for Trust in Agent Societies," The Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2005), Utrecht, July 25-29, pp. 512-518

[8] Fullam, K., T. Klos, G. Muller, J. Sabater, Z. Topol, K. S. Barber, J. Rosenschein, and L. Vercouter. 2005. "A Demonstration of the Agent Reputation and Trust (ART) Testbed: Experimentation and Competition for Trust in Agent Societies," The Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2005) Demonstration Track, Utrecht, July 25-29, pp. 151-152.

[9] Fullam, K. and K.S. Barber. 2006."Learning Trust Strategies in Reputation Exchange Networks," The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2006), Hakodate, Japan, May 8-12, pp. 1241-1248.

[10] Kafali, O. and P. Yolum. 2006. "Trust Strategies for ART Testbed," The Workshop on Trust in Agent Societies at The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2006), Hakodate, Japan, May 9, pp. 43-49.

[11] Sen, S., I. Goswami, and S. Airiau. 2006. "Expertise and Trust-Based Formation of Effective Coalitions: An Evaluation of the ART Testbed," The Workshop on Trust in

Agent Societies at The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2006), Hakodate, Japan, May 9, pp. 71-78.

[12] Stranders, R. 2006. Argumentation Based Decision Making for Trust in Multi-Agent Systems. Master's Thesis, Delft University of Technology.

[13] Apache Commons. 2008. "Commons-Math: The Mathematics Library", http://commons.apache.org/math/, January.

[14] The Apache Jakarta Project. 2008. Official Web site, http://jakarta.apache.org/.

[15] Java EE Downloads. 2008. Web Site of the Java EE SDK http://java.sun.com/javaee/downloads/?intcmp=1282.