

Applying Fuzzy Contextual Filters to Variance Assessment in the ART Testbed. The SPARTAN Appraiser Agent

Esteve del Acebo, Nicolás Hormazábal, Josep Lluís de la Rosa

Agents Research Lab
Institut d'Informàtica i Aplicacions
Universitat de Girona
acebo@ima.udg.es

Resumen

Trust modelling is widely recognized as an aspect of essential importance in the construction of agents and multi agent systems (MAS). As a consequence, several trust formalisms have been developed over the last years. All of them have, in our opinion a limitation: they can determine the trustworthiness or untrustworthiness of the assertions expressed by a given agent, but they don't supply mechanisms for correcting this information in order to extract some utility from it. In order to overcome this limitation, we introduce the concept of reliability as a generalization of trust, and present Fuzzy Contextual Filters (FCF) as reliability modeling methods loosely based on system identification and signal processing techniques. Finally we illustrate their applicability to the appraisal variance estimation problem in the Agent Reputation and Trust (ART) testbed.

Palabras clave: Trust, ART testbed, Fuzzy Corrective Filters.

1 Introduction

Trust is one of the main concepts upon which human and animal societies are built. It is evident, therefore, the importance of its formalization for the construction of artificial or electronic societies, which so vast amount of interest have caused not only in the Artificial Intelligence and Computer Science communities, but also in such different ones as Sociology, Economics and Biology. Quoting [7]:

Artificial Intelligence is quickly moving from the paradigm of an isolated and non-situated intelligence to the paradigm of situated, social

and collective intelligence. The new paradigm of the so called intelligent or adaptive agents and Multi-Agent Systems (MAS) together with the spectacular emergence of the information society technologies (specially reflected by the popularization of electronic commerce) are responsible for the increasing interest on trust and reputation mechanisms applied to electronic societies.

Over the last years, several attempts to make such formalization have been carried out from diverse points of view (recommender systems, social networks, electronic commerce...) All of them suffer, in our opinion, from a quite serious limita-

tion. While they can provide a number, category or even fuzzy statement measuring the trustworthiness of a given agent or, more precisely, the trustworthiness of the information provided by a given agent, they fail in the sense that they don't supply any filtering or correcting method in order to make the provided information useful, even if wrong. The main point of this paper is: In some cases, false information transmitted by an agent can be useful if conveniently filtered.

The aim of this paper is threefold. In the first place, we want to make evident the importance of such filtering mechanisms in order for an agent to improve its performance in a multiagent environment, introducing the concept of reliability as an extension or generalization of trust. Secondly, we present fuzzy contextual filters (FCF) as a convenient and straightforward way, loosely inspired in systems identification and signal processing techniques, to implement those filtering/correcting mechanisms. Finally, we demonstrate the usefulness of FCF by applying it to the appraisal variance estimation problem in the Agent Reputation and Trust (ART) testbed domain.

2 The Need for Trust Formalization in MAS

An essential characteristic of MAS is the existence of an information exchange between the individual agents forming the system. In the case of collaborative MAS, the aim of this communication is the improvement of the global performance of the system. Therefore agents, in general, do not lie each other consciously. In the case of competitive environments, however, individual agents are selfish, in the sense that its behavior is addressed to maximize some kind of individual utility function, even if that means a prejudice for the individual interests of the other agents or the diminution of the overall performance of the system. Communicative acts in competitive MAS are therefore addressed to obtain individual benefit and it is more suitable (because it can be profitable) the conscious communication of false information.

Both in collaborative and in competitive MAS, however, an emitter agent can communicate false information to a recipient agent because of several reasons. The main ones being:

1. The emitter agent is, simply, wrong. He is

honest, in the sense that he believes he is communicating a true statement, but the transmitted information is false.

2. Emitter and recipient agents do not use the same language. The message encloses a true statement, as understood for the emitter agent, but has a different and false meaning for the recipient agent. That's why ontologies are used, just to try to assure that all the agents in a domain speak the same language
3. A transmission error occurred. The emitted and received messages are different.
4. The emitter agent consciously transmits a false information to the recipient agent. The aim of such behavior can be supposed to be the obtaining of some benefit from the prejudicing of the recipient agent. That is the typical behavior we can expect in competitive environments

Whatever could be the reason behind the transmission of false information, individual agents need some kind of mechanism that allow them to deal with it. Agents can't afford (specially in competitive environments) to believe everything the other agents tell to them. A car vendor agent who commits itself to deliver a car "soon" and who says that the car is "fast" can be honest even if the car lasts a year to arrive and it can not run faster than 100 kilometers per hour. Perhaps he really believed what he was saying, perhaps the words "soon" and "fast" have a different meaning in the car vending language or even, perhaps, he said "late" and "slow" but somehow the sounds changed in their way from their mouth to our hears. More probably, however, he is deliberately lying to take profit from us. In either case, we need to learn from our experience in order to know what can be expected from him in further deals. Here is where trust and reputation modeling methods come in as an important field of study inside the theory of MAS.

3 Beyond Trust. Reliability

It is not inside the scope of this document to give a detailed account of the several trust and reputation formalizations that have been proposed along the last years, so we refer the interested reader to [7] for a survey of them. Nevertheless, a point seems to have been so far overseen, to our

knowledge, by these trust formalisms. *It is not necessary to trust an agent (in the sense of believing it is saying the truth) in order to get some utility from the information provided by it.* This information can be useful even if it is false, provided we had some method to correct it.

Let's put an example: a watch agent that goes two and a half hours in advance will never tell you the right time, so you will do good not trusting it. Does it implies that you can't get any utility from it?. On the contrary, you can completely rely on it. Its regularity makes possible to correct the information it provides and get the exact time, a thing that would be impossible to do accurately with a watch that goes only one minute in advance half the time and one minute in retard the other half, at random. We can say much the same thing about our car vendor agent. Better for us do not believe everything he could tell us, of course, but even if we don't trust him, we can yet extract some probably useful information from his offers, perhaps in the form of upper or lower bounds. Moreover, with the time, if we deal with him often enough, we can arrive to learn its language, that is, to capture regularities in it which can allow us to, for example, reject at once a car if he says of it to be "not very old".

The key concept in order to be able to correct messages coming from other agents is reliability¹. If an agent tends to communicate similar information under similar circumstances, a moment will arrive when we will be able to extrapolate the circumstances, more or less correctly, from the received messages. On the contrary, if an agent emits just random messages it will be very difficult, if not impossible, to obtain from them any utility at all.

The corrective mechanisms, which we will call filters, can have very different structures. The filter for a watch agent that goes two hours and a half in retard could be as simple as adding 150 minutes to the time he says it is. On the other hand, we will need a much more sophisticated filter when dealing with the car vendor agent, maybe some kind of expert system. In the following section we will show how simple filters, based on fuzzy systems, can be constructed and how they can be learned and used to improve the performance of individual agents in their environment.

4 Fuzzy Contextual Filters

Think about the following problem: An agent *A* interacts with several other agents in a multi-agent environment requesting from them some kind of information, which they supply (this information can be false because of any of the reasons exposed in section 2). Suppose also that the right answers to *A*'s requests are made available to *A* by the environment in a posterior time instant, in such a way that *A* is able to know which agents told the truth and which agents lied, and how much. Our point is: for *A* to be able to perform well in this kind of environment it has to maintain a set of filters (one of them for each agent it interacts with) which allows it to correct the information received from the other agents, as well as to assess the possible utility of the corrected information. These filters must be dynamic, in the sense that they must evolve and adapt to changes in the environment and in the behavior of the other agents. So, (see figure 1) filters act as a translative layer that eases the process of interpretation of the messages sent by other agents. They can also, on the other hand, help the agent to translate the information it wants to transmit to the language spoken by the other agents, increasing therefore the probability of being correctly understood.

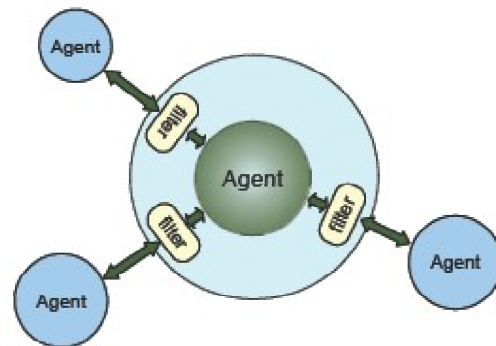


Figure 1. The set of filters of an agent act as a translative layer.

It is also very important for the agent that owns the filter to have some kind of measure of the correctness of the filtered information, that is, the degree to which it can be expected to reflect the reality. We will call this value *reliability* and the filter will compute it from the observed regularities in the behavior of the filtered agent in past interactions.

¹From reliable, in the sense of "giving the same result in successive trials". [6]

Figure 2 shows the suggested structure for the construction of these filters, which we call fuzzy contextual filters (FCF) [2]. A FCF F has two parts, the corrective module and the reliability calculation module. The corrective module is a special case of a Mamdani fuzzy inference system² where the fuzzy rules have the form:

If A_1 is S_1 and ... and A_n is S_n and V is L_1 then W is L_2

where:

- $S_1, S_2 \dots S_n$ are linguistic labels, defined by fuzzy sets on universes of discourse $X_1, X_2 \dots X_n$, respectively.
- $A_1, A_2 \dots A_n$ are fuzzy variables taking values over the fuzzy power sets of $X_1, X_2 \dots X_n$, respectively.
- L_1 and L_2 are linguistic labels defined by fuzzy sets over the universes of discourse U_1 and U_2 , respectively. U_1 and U_2 can be, and usually are, the same set.
- V and W are fuzzy variables taking values over the fuzzy power sets of U_1 and U_2 , respectively.

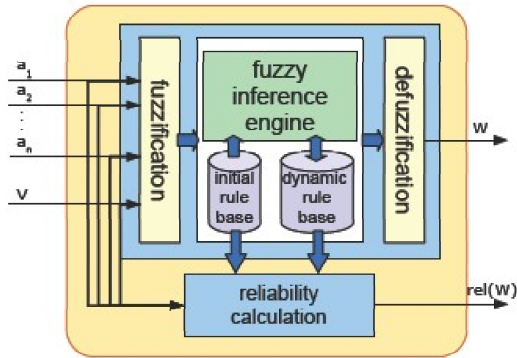


Figure 2. Structure of a fuzzy contextual filter.

We will call $A_1, A_2 \dots A_n$ the *context variables*, V the *main variable* and W the *filtered variable*. We can see the operation of the corrective module as a transformation of fuzzy sets over a certain universe U_1 to fuzzy sets over the universe U_2 (which will be usually the same as U_1) in a

way that depends on the values of the context variables as well as on the value of the main variable. The corrective module of a FCF, then, filter the values (fuzzy sets) of the main variable to obtain new values (fuzzy sets over the same universe or another one) which are expected to be more suitable for some purpose. As is the case with general Mamdani fuzzy systems, it is possible to use FCF on crisp input values to produce crisp filtered values by using appropriate fuzzification and defuzzification procedures.

The rule base of the corrective module has two components, the static and dynamic rule bases. The static rule base is fixed (and possibly the same) for every agent. It expresses the *a priori* assumptions about the behavior of the other agents in the environment and serve as a departing point in the interpretation of other agents's assertions. It can be as simple as the identity function or can, for instance, incorporate some common sense knowledge about the behavior which can be expected from certain kinds of agents. The dynamic rule base is built upon the information extracted (in the form of fuzzy rules) from the interactions between the agent which owns the filter and the filtered agent. It is dynamic in the sense that it evolves with time and can adapt itself to changes in the environment and in the behavior of the filtered agents. The construction of the dynamic rule base can be viewed as a system identification task where the behavior of the filtered agent has to be modeled from a set of examples, the results of past interactions between the modeling and the modeled agents. As a system identification problem, several modeling methods can be used, ranging from those based on a neuro-fuzzy, backpropagation-based approach (Jang's ANFIS [5] would be a good example of this) to those based on lookup tables [8] or, even, genetic algorithms [1].

The function of the second part of the FCF, the reliability calculation module, consists in computing the reliability of the filtered value obtained by the corrective module. Reliability will be a function of the input and context variables and will depend upon the number of prior similar interactions between filtering and filtered agents as well as upon the regularities observed during that interactions.

²A explanation of fuzzy sets, fuzzy logic and fuzzy inference systems' theories is beyond the scope of this paper. We refer the interested reader to [5] where several excellent introductory chapters can be found.

5 The ART Testbed

The Agent Reputation and Trust (ART) testbed [4] is a framework, based on the art appraisal domain, for experimentation and comparison of trust modeling techniques. Agents function as painting appraisers with varying levels of expertise in different artistic eras. Clients request appraisals for paintings from different eras; if an appraising agent does not have the expertise to complete the appraisal, it can request opinions from other appraiser agents. Appraisers receive more clients, and thus more profit, for producing more accurate appraisals.

Let's focus in the opinion requesting part: when an agent A does not have expertise enough to guarantee a good appraisal for a given painting, it can buy the opinion of other, more expert, agents. The process is the following: first, agent A asks all or some of the other agents to provide a value stating their confidence in the accuracy of their appraisal of the painting. Then, A decides, upon the received confidence values, which agents to trust, that is, which opinions to purchase.

This is the main point where the communication of false or misleading information can happen in the ART testbed. An agent can declare a great confidence in its appraisal just to fool the requesting agent into purchasing it, and then produce a very bad appraisal. This will result in a big error in the requesting agent's appraisal and, consequently, a big loss in its the client share. On the other hand, the requesting agent has no way to know what the confidence value provided by an agent means. It is a value over an arbitrary range that has to be interpreted. It is perfectly possible for a given confidence value to mean completely different confidence levels for different agents. In the following subsections we will examine in some detail several particular aspects of the ART testbed.

5.1 Several Considerations on the Client Share Assignment Function

ART designers propose, in [4], [7] the following equation for client share adjustment for agent a after each iteration

$$r_a = q \cdot r'_a + (1 - q) \cdot \bar{r}_a \quad (1)$$

where r'_a is agent a 's client share in the previous timestep, q is a parameter whose value lies in the $[0 \dots 1]$ interval and \bar{r}_a , depending on the mean relative error committed by a (as well as by the other agents), represents a 's preliminary client share for the current iteration. The formula in [4] for the computation of \bar{r}_a is:

$$\bar{r}_a = \left(\frac{\delta_a}{\sum_{b \in A} \delta_b} \right) \cdot |C| \quad (2)$$

where C is the set of customers or paintings to be assigned in the current iteration and

$$\delta_a = 1 - \frac{\epsilon_a}{\sum_{b \in A} \epsilon_b} = \frac{\sum_{b \in A} \epsilon_b - \epsilon_a}{\sum_{b \in A} \epsilon_b} \quad (3)$$

representing ϵ_b the mean relative error made by agent b during the previous iteration, that is:

$$\epsilon_a = \frac{\sum_{c \in C_a} \frac{|p_c^* - t_c|}{t_c}}{|C_a|} \quad (4)$$

where C_a is the set of appraiser a 's clients, p_c^* is appraiser a 's final appraisal for client c and t_c is the true value of the painting client c submitted to a for appraisal.

Substituting δ_a 's values in equation 2 and simplifying we have:

$$\begin{aligned} \bar{r}_a &= \frac{\sum_{b \in A} \epsilon_b - \epsilon_a}{\sum_{b \in A} (\sum_{c \in A} \epsilon_c - \epsilon_b)} \cdot |C| \\ &= \frac{\sum_{b \in A, b \neq a} \epsilon_b}{\sum_{b \in A} \sum_{c \in A, c \neq b} \epsilon_c} \cdot |C| \end{aligned} \quad (5)$$

this can easily be proved equivalent to:

$$\bar{r}_a = \frac{|C|}{|A| - 1} \cdot \frac{\sum_{b \in A, b \neq a} \epsilon_b}{\sum_{b \in A} \epsilon_b} \quad (6)$$

or, alternatively

$$\bar{r}_a = \frac{|C|}{|A| - 1} \cdot \frac{\Sigma}{\epsilon_a + \Sigma} \quad (7)$$

where, as stated before and following the notation in [4], ϵ_b represent the mean relative error made by agent b during the previous iteration, A and C are respectively the sets of agents and clients and

Σ is simply the sum of the mean relative errors made by the agents other than a .

It is worth to make a couple of considerations about eqs. 6 and 7, Firstly, it is clear from the equations that no agent can manage to get a preliminary client share greater than $|C|/(|A| - 1)$, never mind the accurateness of its appraisals. An agent in a seven-agent environment, for example, can't expect to obtain more than one sixth of the total client amount, even if it guesses exactly all the pictures' prices and the other agents make huge errors in their appraisals. On the contrary, a big mean relative error can dramatically diminish the preliminary client share of an agent to one single client, in the worst case (not to zero because of ART's design, see [4]). So, we could say that ART's preliminary client share assignation method penalizes the bad appraisals much more than it rewards the good ones. It will be important to take this into account when designing our appraising agent strategy, it seems to make much more sense to spend money trying to reduce the relative error in the cases when we suspect it to be large than refining opinions we can suppose to be quite accurate. On the same line, benefits from deceiving an agent by selling a deliberately wrong appraisal to it (thus causing it to loose a big portion of its customer set for the next iteration) are equally shared amongst the deceiver agent and all the other agents. This can make deceiving tactics counterproductive (mainly in populated environments, where there are more agents to share the customers lost by the deceived agent) due to the loss of reputation not to being compensated by the difference in earnings. It would be interesting to study how robustly current agents would behave if a more deception-encouraging formula for client assignment was used (something like the deceiver agent keeping all the customers lost by the deceived agent, for example).

The second consideration to be made refers to the fact that computation of the preliminary client share for agent a takes not into account the actual distribution of the relative error amongst the other agents, but the total amount of this error. So, an agent can, at the beginning of each iteration, knowing the mean relative error made by itself in the last iteration and the client share assigned to it in the current one, compute the total mean error made for all the remaining agents in the past iteration and use this value as a estimator of the total mean error which the remaining agents will made in the next iteration. This will help the agent to establish a near-optimal

strategy regarding the amount of money it has to spend refining its own appraisals as well as purchasing appraisals from other agents in the current iteration.

5.2 Appraise Combination

Following [4], in the ART testbed the appraisal error is distributed as a normal aleatory variable with mean 0 and standard deviation:

$$s = \left(s^* + \frac{\alpha}{c_g} \right) \cdot t \quad (8)$$

where s^* is the inherent standard deviation for the epoch to which the painting belongs, α is a simulation dependent constant and c_g is the amount of money spent in the appraising process. Thus, the standard deviation for the appraisals will also be s and the standard deviation of the relative errors defined as $\frac{\text{appraise}-t}{t}$ will be:

$$\sigma = \frac{s}{t} = \left(s^* + \frac{\alpha}{c_g} \right) \quad (9)$$

We will now study the following problem: given a set $A = \{a_1, a_2 \dots a_n\}$ of appraisals for a painting and given the set $V = \{\sigma_1^2, \sigma_2^2 \dots \sigma_n^2\}$ of the corresponding variances of the relative errors (we assume that the means of the relative errors of the appraisals are all zero), we want to find the set of weights $W = \{w_1, w_2 \dots w_n\}$ such that $w_i \geq 0$ and $\sum_i w_i = 1$ which minimizes the error of the combined appraisal $\bar{t} = \sum_i w_i \cdot a_i$. We will do it by minimizing the variance of the combined appraisal \bar{t} .

Let's begin with the case of two appraisals a and b , with known variances of the relative errors σ_a^2 and σ_b^2 , we can compute the variance of the appraisal $\bar{t} = p \cdot a + (1 - p) \cdot b$ as follows:

$$\begin{aligned} Var(\bar{t}) &= Var(p \cdot a + (1 - p) \cdot b) \\ &= p^2 \cdot Var(a) + (1 - p)^2 \cdot Var(b) \end{aligned} \quad (10)$$

where $Var(a)$ and $Var(b)$ are the variances of the appraisals. From Eqs. 8 and 9, and taking into account that appraisals have the same variance than errors, we can express $Var(a)$ and $Var(b)$ in terms of the known variances of the relative error σ_a^2 and σ_b^2 using the following equivalence:

$$Var(a) = t^2 \cdot \sigma_a^2 \quad (11)$$

where t is the real value of the painting. We can then restate equation 10 as follows:

$$Var(\bar{t}) = t^2 (p^2 \cdot \sigma_a^2 + (1-p)^2 \cdot \sigma_b^2) \quad (12)$$

We can now differentiate:

$$\frac{d(Var(\bar{t}))}{dp} = 2 \cdot t^2 (p \cdot \sigma_a^2 - (1-p) \cdot \sigma_b^2) \quad (13)$$

Equalling to 0 and solving for p , we have

$$p = \frac{\sigma_b^2}{\sigma_a^2 + \sigma_b^2} \quad (14)$$

So, the appraisal of minimal variance will be:

$$\bar{t} = \frac{\sigma_b^2}{\sigma_a^2 + \sigma_b^2} \cdot a + \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2} \cdot b \quad (15)$$

Finally, the variance of \bar{t} 's relative error, $\sigma_{\bar{t}}^2$ can be obtained from equations 11 and 12 as:

$$\sigma_{\bar{t}}^2 = \frac{\sigma_a^2 \cdot \sigma_b^2}{\sigma_a^2 + \sigma_b^2} \quad (16)$$

This can be easily generalized to the case of a greater number of appraisals. In the case of three appraisals a, b and c , the minimal variance appraisal will be:

$$\bar{t} = \frac{\sigma_b^2 \sigma_c^2 \cdot a + \sigma_a^2 \sigma_c^2 \cdot b + \sigma_a^2 \sigma_b^2 \cdot c}{\sigma_a^2 \sigma_b^2 + \sigma_b^2 \sigma_c^2 + \sigma_a^2 \sigma_c^2} \quad (17)$$

with a variance of \bar{t} 's relative error

$$\sigma_{\bar{t}}^2 = \frac{\sigma_a^2 \cdot \sigma_b^2 \cdot \sigma_c^2}{\sigma_a^2 \sigma_b^2 + \sigma_b^2 \sigma_c^2 + \sigma_a^2 \sigma_c^2} \quad (18)$$

and, in the general case, if we have a set $A = \{a_1, a_2, \dots, a_n\}$ of appraisals and the set $V = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2\}$ of the corresponding relative error variances:

$$\bar{t} = \frac{\sum_{i=1}^n a_i \cdot \prod_{j \neq i} \sigma_j^2}{\sum_{i=1}^n \prod_{j \neq i} \sigma_j^2} \quad (19)$$

with a variance of \bar{t} 's relative error

$$\sigma_{\bar{t}}^2 = \frac{\prod_{i=1}^n \sigma_i^2}{\sum_{i=1}^n \prod_{j \neq i} \sigma_j^2} \quad (20)$$

³In the ART testbed, paintings can only belong to one era, and they belong to it completely. It is possible, however, to imagine instances of the problem where paintings could belong, to a certain degree, to different eras. Our method is general enough to cope with this.

6 Using FCF for Appraisal Variance Estimation in the ART testbed

We have just seen how to combine optimally two or more appraisals in order to obtain the appraisal with the minimal expected relative error. The only drawback is: we do not know what the variances of the appraisals to be combined are. We need, therefore, a way to guess them departing from the confidence values supplied by the appraise-selling agents.

We solve the problem providing our agent with a set of Fuzzy Corrective Filters, one for each agent other than itself in the environment. The structure of each filter is very simple. It has, as input variable, the confidence value stated by the appraisal-selling agent, and, as context variable, the era to which the painting belongs. The filtered variable will be the expected variance of the appraisal-selling agent's appraisal relative error. The FCF will produce this output from the confidence value provided by the appraisal-selling agent and the historial of past interactions.

Let's see the structure of rule bases in the corrective module of the FCF: rules in the initial rule base are predefined by design and serve the purpose of providing a sensible starting point to the interpretation process. Rules in the dynamic rule base, on the other hand, are continuously obtained from interactions between our agent and the appraisal-selling agent. Each of the rules in the global rule base R (the union of initial and dynamic rulebases) has the same form:

$$R_i : \text{If } era = E_i \text{ and } conf = C_i \text{ then} \\ relError = E_i$$

where E_i ³ and E_i are singleton fuzzy sets over the sets of the eras and the reals, respectively and C_i is a fuzzy real number. So, for instance, if we purchase an appraisal for a cubist painting for which the appraisal-selling agent declares to have a confidence 0.5, and the provided appraised value is 20000 but the real price of the painting turns out to be 25000 (giving a relative error of 0.2), we will add to our dynamic rule base the following rule (see figure 3):

If $era = cubism$ and $conf = 0.5$ then
 $relError = 0.2$

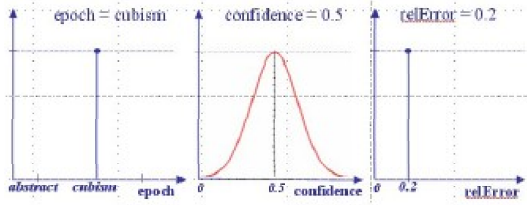


Figure 3. Fuzzy sets corresponding to the rule : If $epoch = cubism$ and $conf = 0.5$ then $relError = 0.2$.

We will have, then, a possibly large number of fuzzy rules in this form. Now suppose that we want to consider the possibility of purchasing an appraisal for a painting of a given era e from an agent which states that it has a confidence c in its appraisal. How to estimate the variance of the relative error of the appraised value?. We know that the variance is defined as the expectation of the quadratic error, and also that the mean of the relative error is zero by design. Given the confidence value c , then, it would be enough to gather all the interactions in which the agent has stated the very same confidence c in its appraisal of a painting of the same era and estimate the variance of the relative error for the confidence value c as the mean of the squares of the errors made. Unfortunately, confidence values will be, in general, scattered along a big range of values, so we can hardly expect to have enough rules to make the estimation accurate. We can, nevertheless, use the rules with a confidence value "close enough" to c in order to improve the estimation. This corresponds to compute the output of the fuzzy system (the corrective module) in the following way:

$$\sigma_{e,c}^2 = \frac{\sum_{R_i \in R} \pi_i(e) \cdot \mu_i(c) \cdot E_i^2}{\sum_{R_i \in R} \pi_i(e) \cdot \mu_i(c)} \quad (21)$$

where $\pi_i(e)$ will take the values 1 or 0 depending on whether the era of the painting corresponding to fuzzy rule R_i was e or not, E_i is the relative error made by the appraisal-seller agent in the interaction corresponding to fuzzy rule R_i and $\mu_i(c)$ is the degree to which the value c belongs to the fuzzy number C_i , which we define as:

$$\mu_i(c) = \exp\left(-\frac{(c - C_i^*)^2}{\alpha^2}\right) \quad (22)$$

where C_i^* is the central value of the fuzzy number C_i , the width of which can be controlled by the parameter α .

In order to deal with the dynamical nature of the testbed and the fact that agents' statements can change of meaning with time, we further modify Eq.21 by including a new term $TimeStep_i$ which represents the iteration in which the interaction corresponding to R_i happened:

$$\sigma_{e,c}^2 = \frac{\sum_{R_i \in R} \pi_i(e) \cdot \mu_i(c) \cdot TimeStep_i^K \cdot E_i^2}{\sum_{R_i \in R} \pi_i(e) \cdot \mu_i(c) \cdot TimeStep_i^K} \quad (23)$$

where we can use the real parameter K to vary the relative influence of the rules in the computed result, giving more or less importance to more recent interactions.

6.1 The Reliability Calculation Module

The implementation of the reliability calculation module for the FCF takes into account several static criteria for fuzzy rule base quality assessment, mainly the completeness of the rule bases (very roughly speaking, the number of rules that fire in the calculation of the variance). Those quality criteria are based in previous work by the authors. The interested reader can find further information in [3].

7 Results

The global behavior of agents in ART experiments is very sensitive to even small changes in the environment or in the particular behavior of single agents. In order to try to overcome this problem, two series of simulations have been carried out, using two sets of agents, a first one (Set A) with several of the best competitors in the 2006 International ART Competition (i.e. IAM, Frost, Neil, and Sabatini), and a second one (Set B) with new agents synthesized to be more trusty. Ten simulations have been done in each series. In five of them our agent, called SPARTAN, uses FCF in order to translate the certainty values provided by the other agents to variances, in the remaining five simulations SPARTAN don't uses FCF, that is, he assumes the other agents to talk

the same language than himself. A representative set of results is shown in figure 4.

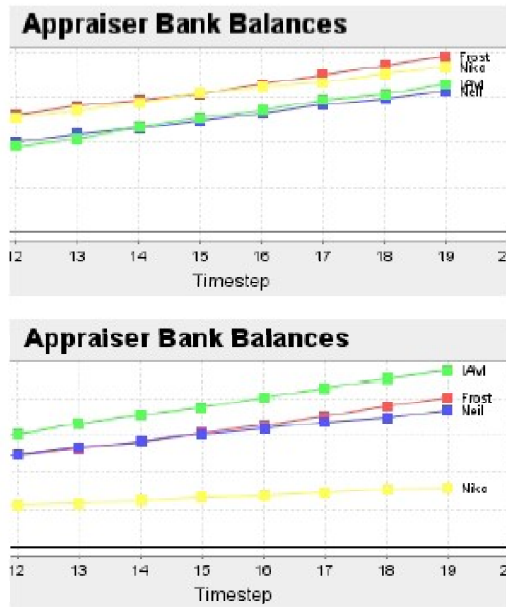


Figure 4. SPARTAN (under the nickname Niko) versus IAM, Neil and Frost. *Bottom: Results without filtering. Top: Results using FCF.*

As a consequence of ART Testbed sensitivity to initial conditions, the inherent random nature of the simulations makes the amount of money earned by the agents in every run to be very variable. Therefore one cannot simply take money as an absolute performance measure. Though other methods may apply (ranking, etc), we want a method able to keep accurately the distance between agents in the different runs, so we have decided to normalize the results by dividing the money earned by SPARTAN by the money earned by the remaining best agent. This gives us an adimensional measure of SPARTAN's *efficiency* that helps the comparison.

Table 1. Results of the experiments

	With FCF		Without FCF		%	t
	ME	SD	ME	SD		
A	0,903	0,059	0,741	0,033	21,9	5,35
B	0,89	0,096	0,75	0,125	18,7	1,99

The results of the experiments can be seen in Table 1, where ME stands for Mean Efficiency and SD stands for Standard Deviation. Results are

very similar in both series, although slightly better in the case of Set A, the more competitive agents. The improvement in efficiency is about 20% in both series. The t-values for the unequal variances Student's t-Test guarantees the statistical significance of the results with high probability (greater than 0.999 for Set A and greater than 0.95 for Set B).

Fuzzy Contextual Filters (along with several other tricks) allowed SPARTAN, to win the 4th position out of 16 participants in the 2nd international ART competition in AAMAS 2007, May 14-18, 2007 in Hawaii.

Acknowledgements

This work was supported in part by the Grant TIN2006-15111/Estudio sobre Arquitecturas de Cooperación from the Spanish government and by the EU project Num.34744 ONE: Open Negotiation Environment, FP6-2005-IST-5, ICT-for networked Businesses.

References

- [1] O. Cordon and F. Herrera. A general study on genetic fuzzy systems. In *Genetic Algorithms in Engineering and Computer Science*. John Wiley & Sons, 1995.
- [2] Esteve del Acebo and Josep Lluís de la Rosa. A fuzzy system based approach to social modeling in multi-agent systems. In *The First International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002, July 15-19, Bologna, Italy*, pages 463–464, 2002.
- [3] Esteve del Acebo, Albert Oller, Josep Lluís de la Rosa, and Antoni Ligeza. Statistic criteria for fuzzy systems quality evaluation. In Angel P. Del Pobil, José Mira, and Moonis Ali, editors, *IEA/AIE (Vol. 2)*, volume 1416 of *Lecture Notes in Computer Science*, pages 877–887. Springer, 1998.
- [4] K. Fullam, T. Klos, G. Muller, J. Sabater, A. Schlosser, Z. Topol, K. S. Barber, J. Rosen-schein, L. Vercouter, and M. Voss. A Specification of the Agent Reputation and Trust (ART) Testbed: Experimentation and Competition for Trust in Agent Societies. In

- The Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2005)*, pages 512–518. ACM Press, July 2005.
- [5] J.S.R. Jang, C.T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing*. MATLAB Curriculum Series. Prentice-Hall, first edition, 1997.
- [6] Merriam-Webster. Merriam-webster online dictionary, 2007. <http://www.m-w.com/>.
- [7] J. Sabater. *Trust and Reputation for Agent Societies*. Monografies de l'Institut d'Investigació en Intel·ligència Artificial (20). Consell Superior d'Investigacions Científiques, 2002.
- [8] Li-Xin Wang. *Adaptive Fuzzy Systems and Control. Design and Stability Analysis*. Prentice-Hall, 1994.