Abstract—The prediction of behaviours in dynamic and subjective environments is an interesting issue in the way to obtain a complete delegation of human-like decisions in autonomous agents. Complex mathematical algorithms have been proposed by academic researchers to be applied in agent-based recommendation systems. In this paper we propose to use a classic estimation method known as Alpha-Beta filtering. This innovative approach is compared with a previous proposal (called AFRAS) from the authors that is based on fuzzy sets and that has been successfully compared with some of the most representative reputation models and with another classic estimation method (Kalman). This paper evaluates how consumer agents applying both methods predict the behaviour of provider agents focusing on the velocity of convergence of the predictions to the real behaviour where such behaviour adopts a prefixed variability.

I. INTRODUCTION

From the increasing popularity of Internet several methods of remotely recommending services and products have arisen. Most of them consists of a central entity that certifies the satisfaction of some given evaluation criteria [17]. This kind of solution relays on the assumption of the existence of objective universally accepted criteria. However, in real life many products and services are qualified according to personal subjective preferences. In those cases, the recommendations about services and products can not be considered as completely transitive. Therefore different particular views of the world computed from direct experiences with agent providers will coexist (as many as members of the system) [6]. However a certain level of cooperation is required in low-density populations of agents in order to improve the selection of trustworthy services and products. Several providers have focused on how this cooperation would be implemented in such distributed evaluation scheme. On the side of these commercial sites the problem has been tackled via very simple and fast computations: item-to-item correlation, normal or average sum of numeric evaluations, etc.

On the other hand the academic community has studied the possibility of a community of autonomous agents evaluating each other [2]. This point of view emerges from the foundations of AI: an agent represents a person, and an agent system represents a human society (this is the so called social metaphor of agents). Then the evaluation of the other agents is noted as reputation, and cooperation consists of the transmission of agents’ images among a selected subset of trustworthy agents, this is the so called witness information. From this sociobiological perspective, computations on reputation of other agents should reflect human-like reasoning. And precisely the academic proposals presented next in this document are inspired in the factors that humans are supposed to apply in the computation of reputation. From all the academic AI-inspired reputation models that were proposed, we will briefly describe the foundations of some of the most representative models of reputation.

A. SPORAS, from M.I.T.

P. Maes and other researchers of the Massachusetts Institute of Technology (M.I.T.) proposed two reputation algorithms: SPORAS and HISTOS [11]. SPORAS is inspired in the foundations of the chess players evaluation system called ELOS. The main point of this model is that trusted agents with very high reputation experience much smaller changes in reputation than agents with low reputation. SPORAS also computes the reliability of agents’ reputation using the standard deviation of such measure. On the other hand, HISTOS is designed to complement SPORAS including a way to deal with witness information (personal recommendations). HISTOS includes witness information as source of reputation through a recursive computation of weighted means of ratings. It computes reputation of agent i for agent j from the knowledge of all the chain of reputation beliefs corresponding to each possible path that connects agent i and agent j. It also plans to limit the length of paths that are taken into account. To make fair comparisons with other proposals, we value that limit as 1, since we consider that agents communicate only its own beliefs (that are obviously the result from direct experiences and the beliefs of different sources), but not the beliefs of other sources that contributed to the own belief of reputation. So in our view of reputation systems we assume that is not realistic to know the complete chain of beliefs involved along all the path from agent i to agent j.

B. REGRET, from the Spanish Artificial Intelligence Research Institute

REGRET, from the Spanish Artificial Intelligence Research Institute [1], takes into account three types of computation
of indirect reputation depending on the information source: system, neighbourhood and witness reputations. From them witness reputation is the one that corresponds to the concept of reputation that we are considering. REGRET includes a measure of the social credibility of the agent and a measure of the credibility of the information in the computation of witness reputation. The first of them is computed from the social relations shared between both agents. It is computed in a similar way to neighbourhood reputation, and it uses third parties references about the recommender directly in the computation of how its recommendations are taking into account.

We have a slightly different point of view: we assume that if recommender agent is not trusted enough, it will never be asked, but even then, we compute first the reputation of the recommender (asking whoever for references), and then we calculate, as a different action, the influence of the recommendation in the reputation of a seller. We think that mixing both operations in one does not clarify the process and it does not make any difference with the general computation of reputation with recommendations.

On the other hand, the second measure of credibility (information credibility) is computed from the difference between the recommendation and what the agent experienced by itself. The similarity is computed matching this difference with a triangle fuzzy set centered in 0 (the value 0 stands for no difference at all). The information credibility is considered as relevant and taken into account in the experiments of this present comparison.

We interpret that both decisions are also, in some way, supported by the authors of REGRET, who also assume that the accuracy of previous pieces of information (witness) are much more reliable than the credibility based on social relations (neighbourhood), and they reduce the use of neighbourhood reputation to those situations were there is not enough information on witness reputation.

C. Unnamed, from the University of South Caroline

This trust model [10] uses Dempster-Shafer theory of evidence to aggregate recommendations from different witnesses. The main characteristic of this model is the relative importance of fails over success. It assumes that deceptions causes stronger impressions than satisfactions. It then applies different gradients to the curves of gaining/losing reputation in order to lose easily reputation, if recommender agent is not trusted enough, it will never be asked, but even then, we compute first the reputation of the recommender (asking whoever for references), and then we calculate, as a different action, the influence of the recommendation in the reputation of a seller. We think that mixing both operations in one does not clarify the process and it does not make any difference with the general computation of reputation with recommendations.

II. A FUZZY REPUTATION AGENT SYSTEM (A.F.R.A.S.)

The agent system called AFRAS, proposed by the authors of this paper, intends to integrate several different features into the same agent BDI architecture: emotions through adaptive characterization of agents with sociability, susceptibility and shyness fuzzy attributes [15]; security mechanisms to introduce newcomers into already formed clusters of agents with shared interests [14]; ability to reach agreements through fuzzy counteroffers from crisp offers) [23]; privacy protection of the arguments involved in deliberative negotiations [24] and finally the focus of this paper: the extensive use of fuzzy sets, particularly to represent reputation values. This formalism makes sense since human opinions about others are vague, subjective and uncertain (in other words, reputation is a fuzzy concept, valued in fuzzy terms). This view is also assumed in [22].

In this reputation model, direct experiences and witness information are both considered. They are aggregated through a weighted mean of fuzzy sets. Aggregation of fuzzy sets is computed with Mass Assignment assumptions based on Baldwin’s theory of evidence [3]. In the case of direct experiences, weights depend on a single attribute that represents the memory of the agent \(0 < memory < 1\). This memory is a global feature of the agent, applied universally to any counterpart (just a value not a vector of \(memory_i\)). This global nature of memory intends to reflect how in real life deceptions with a particular part affects our attitude facing any others. We associated such meaning to this value because it determines the importance of past direct experiences \(\approx R_{i-1}\) (R, reputation) over a new one \(\approx (DE, Direct Experience)\).

\[
R_i = R_{i-1} + \frac{(DE_i - R_{i-1}) \cdot (1 - memory)}{2}
\]  

Memory is computed as a function of the overlapping between two fuzzy sets that represents the level of success of last prediction. If the satisfaction provided by a partner was similar to the reputation estimation assigned to such partner, the relevance of past experiences (memory) would be then increased. On the other hand, if they were different, is the relevance of the last experience what would be therefore increased (the corresponding agent is ‘forgetting’ past experiences, ‘losing’ memory).

Once similarity is computed in this way, an average sum of previous memory (in iteration \(i-1\), \(memory_{i-1}\) with similarity \(SIM) SIM(R_{i-1}, DE_i\) is applied to obtain current memory value (in iteration \(i\), \(memory_i\)):

\[
memory_i = \frac{memory_{i-1} + SIM(R_{i-1}, DE_i)}{2}
\]

This equation follows the next simple principles: If the prediction fitted well the rating \(SIM \approx 1\) then memory (the importance given to the past reputation over the last direct experience) would increase in \(1/2 + memory/2\). On the other hand, when they were not similar at all \(SIM \approx 0\), memory would become useless, and its relevance in the next estimations of reputation would be halved.

These properties avoid memory being below zero and above one. The initial value of memory associated to any agent joining the system should be minimum (zero), although it would be soon increased when there was any success in the
estimations of reputation. Reliability of reputation values is modeled through the fuzzy sets themselves. It is implicit in them, graphically we can interpret the gradient of the sides of a trapezium representing a fuzzy reputation as its reliability. A wide fuzzy set representing a given reputation represents a high degree of uncertainty over that reputation estimation, while a narrow fuzzy set implies a reliable reputation. Recommendations are aggregated directly with direct experiences in a similar way (weighted sum of fuzzy sets). But in this case, the weight given to each part (recommendation experiences) depends on the reputation of the recommender. A recommendation would (at most) count as much as a direct experience if the recommender had the highest reputation.

Finally, to update the reputation of recommenders, an agent computes similarity level of overlapping between the corresponding fuzzy sets with afterward results of the direct experience with the recommended partner. Then, reputation of recommender would be increased or decreased accordingly.

III. ALPHA-BETA FILTERING

Alpha Beta method [4] is a linear recursive algorithm to estimate an unknown state variable from noisy observations. In our application, the state variable would be the reputation, while observations would be the results from direct experiences. Alpha Beta assumes that the state variable (reputation) follows a constant velocity model, with some uncertainty characterized by a parameterized random variable (plant-noise model): starting with some initial value, the reputation’s velocity evolves through time by process noise of random accelerations, constant during each sampling interval but independent. Without any noise, reputation would have constant velocity, so we are using noise to model sudden changes of behaviour (in other words, reputations with a non-constant velocity). Alpha beta also assumes that observations only are available, subject to measurement noise of constant covariance. Clearly the more is known a priori about the motion the better predictions will be, so in our application of reputation with agents this could be considered a bit of hack since noise has in fact a constant variance, but it is not a realistic assumption to know a priori the real model of the noise.

So state variable (reputation) evolves following the next equation:

\[ x(t + \Delta t) = F(\Delta t) x(t) + q(\Delta t) \]  

where \( \Delta t \) is the time delay between last update and current observation, \( t_k \) to \( t_{k-1} \), \( F(\Delta t) \) is the transition matrix and \( q(t) \) is characterized by its covariance matrix, \( Q(\Delta t) \).

Since we assume a constant velocity model of reputation, transition matrix \( F \) adopts the next value:

\[ F(\Delta t) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \]  

And observations result from a linear operation on state variable (reputation) corrupted by additional noise:

\[ z(t) = H x(t) + n(t) \]  

being \( n(t) \) a random variable with covariance given by matrix \( R \).

In our specific model for filtering reputation, we have a dynamic linear system, with vector \( \hat{x}[k|k] \) containing both the trust estimate and its time derivative for a given agent (the notation \( k|k \) means estimation at time \( k \), considering observations until time \( k \), while \( k|k-1 \) is the prediction at time \( k \) from last update at time \( k-1 \)).

Under this context, the equations for Alpha Beta filter to compute the best estimation for \( x(t) \) are the following:

\[ \hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \left[ \frac{\alpha}{\sqrt{\Delta t}} \right] \cdot [z(k+1) - (\hat{z}(k+1|k))] \]  

So the state estimate is a weighted sum of a state \( \hat{z}[k + 1|k] \) predicted from the last estimate to be \( F(\Delta t) x(k|k) \) and innovation, computed as the difference between a predicted observation \( \hat{z}[k + 1|k] \), with the current observation, \( z(k+1) \).

We can compute the value of \( \beta \) from \( \alpha \) in order to use just \( \alpha \) as the single parameter of the estimation method:

\[ \beta = 2 \cdot (2 - \alpha) - 4 \cdot \sqrt{1 - \alpha} \]  

The values of \( \alpha \) are between 0 and 1 and represent a balance between the relevance given to the history of past observations vs. the last observation. Therefore, \( \alpha = 0 \) would mean that the last observation has no effect in next prediction. On the other hand, \( \alpha = 1 \) would mean that the history of past observations were ignored in next prediction.

Estimates for covariances \( Q \) and \( R \) are 4x4 matrices. Usually, the exact models for dynamics and observation are not known, so the design for a given application is a trade-off to adjust the parameters. Matrix \( R \) is usually adjusted from observed data variability (sample variance), while matrix \( Q \) is tuned to achieve satisfactory balance between noise filtering (when the prediction model is much better than observation noise) and reactions to sudden changes (intervals while the model fails) so the design for a given application is a trade-off to adjust the parameters. Matrix \( Q \) is usually adjusted from observed data variability (sample variance), while matrix \( R \) is tuned to achieve satisfactory balance between noise filtering (when the prediction model is much better than observation noise) and reactions to sudden changes (intervals while the model fails) so the design for a given application is a trade-off to adjust the parameters.
The resulting equation is as follows:

\[ \hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \left[ \frac{\alpha}{\alpha + \beta} \right] \cdot [z(k+1) - (\hat{x}(k+1|k))] \]

(9)

Where gamma can also be computed from alpha and beta:

\[ \gamma = \frac{\beta^2}{2 \cdot \alpha} \]

(10)

IV. EXPERIMENTS ON THE OPTIMIZATION OF PREDICTIONS: AFRAS VS. ALPHA-BETA

A. Experimental Setup

Since our trust model based on fuzzy reputation (AFRAS) has been previously compared with SPORAS, REGRET and the unnamed proposal of Singh and Yu [16], in this paper we just intend to compare AFRAS with the classic estimation method known as alpha beta filter. The comparison will follow an evolved version of the testbed proposed by Zacharia and Maes.

In those simulations the convergence of predictions were evaluated. They studied the level of success obtained from an agent continuously interacting (having direct experiences) from a particular source of reputation chosen in each iteration randomly among all the possible partners.

Our experiments have tested the results of one agent randomly having direct experiences with 10 different agents in 200 iterations. The same order of partners was selected along the 200 iterations and the same response from those agents was applied to AFRAS and Alpha Beta in order to obtain a consistent comparison. Due to this reason, we have also adapted the range of possible values of reputation to be from 0 to 100.

The initial parameters applied to \( \alpha - \beta \) and \( \alpha - \beta - \gamma \) estimation methods are computed from three values of alpha: 0.15, 0.5 and 0.85.

In order to evaluate the convergence of reputation estimations, the simulation has satisfied the next properties:

- A prefixed behaviour (uniformly distributed) was assigned to each of the partners. And all of them use such prefixed behavior along 200 iterations.
- The satisfaction provided at any time is drawn from a normal distribution. The mean of that distribution is equal to this prefixed behaviour. Three different standard deviation will be tested in the experiments: 3, 10 and 33.
- Initially the reputation of all agents is 10 (over 100) with a reliability of 1 (over 100).

The error produced in each prediction has been quantified in AFRAS through the difference between the center of gravity defuzzification of previous reputation and the result of the direct experience. Next, we will show the average error committed in the estimations computed by both alternatives in order to compare the speed of convergence. We assumed that the standard deviation of direct experiences from partners was prefixed to be 3, 10 and 33.33 (over 100).

These values represent three different scenarios. The first of them is a situation where each partner is evaluated all the time with very similar values (scenario 1). On the other hand, the third of them represents a situation where subjective evaluation may change suddenly (scenario 3). The second one is in the middle of the other two (scenario 2).

B. Scenario 1. Easy predictions

In figure 1 we can see how the curves corresponding to the little values of \( \alpha \) show a very bad convergence of predictions. On the other hand the results of \( \alpha - \beta \) with an initial value of \( \alpha = 0.85 \) even improves the results of AFRAS. The rest of the curves are slightly worst than AFRAS in the long term and more or less similar among them. Although most of them show better initial reactions than AFRAS.

C. Scenario 2. Normal predictions

In figure 2 we can see how again the curves corresponding to the little values of \( \alpha \) (\( \alpha - \beta \) and \( \alpha - \beta - \gamma \) are overlapped under the same line) show a very bad convergence of predictions. But now no \( \alpha - \beta \) estimation can improve the results of AFRAS in the long term. However their initial convergence in the first iteration is faster than AFRAS one.
D. Scenario 3. Difficult predictions

In figure 3 we can see how the curves corresponding to the little values of $\alpha$ ($\alpha - \beta$ and $\alpha - \beta - \gamma$ are overlapped under the same line) are the worst, but now in the long term they obtain similar results than the other $\alpha - \beta$ estimations. The other remarkable difference of this scenario against the others is that the error committed by first predictions of AFRAS are now very close to the corresponding error of the medium and high values of $\alpha$ in $\alpha - \beta$ estimations.

V. Conclusions

This paper is one step of a sequence of trials of applying classic estimation methods (that are extensively used in other domains) as reputation models. In fact, we are not pioneers considering this kind of methods as part of AI research [21]. The mentioned sequence includes an approach with Kalman estimation method, that seemed to generate appropriate predictions since they converge with more or less similar speed than other reputation models as Regret, Sporas/Histos, Singh-Yu and AFRAS [12]. This paper is the natural extension of that work, but here we prove the ability of alpha beta estimation method rather than Kalman.

From the results obtained we can state clearly that little values of $\alpha$ should be avoided even when predictions were easy, and that there is no relevant difference between using $\alpha - \beta$ or $\alpha - \beta - \gamma$ in our scenario of agents computing reputation. Furthermore we can see that just with easy predictions $\alpha - \beta$ estimation produces less error in the long term than AFRAS. So considering such goal, $\alpha - \beta$ methods have not improved the performance of AFRAS. However if we focus on the initial reactions of the estimators, we can see how the first predictions of $\alpha - \beta$ method are more accurate than those corresponding to AFRAS. So $\alpha - \beta$ estimators obtain a not bad global evaluation, but with the results of the simulations we can not state that $\alpha - \beta$ would be better than AFRAS. In spite of that we think that the application of more complex estimation methods may improve AFRAS performance. In this way, the next step in this line of research, is an extension with Interacting Multiple Model filters [4] that will explore the design of appropriate transitions for initialization.

REFERENCES