

# A Multistage Fuzzy Decision Approach for Modelling Adaptive Negotiation Strategies

Jan Richter, Matthias Klusch and Ryszard Kowalczyk

**Abstract**—Automated bilateral agent negotiation under uncertainty, that is with imprecise or uncertain information about preferences, utilities and strategies of the opponent is known to be hard. In this paper, we present the first adaptive solution that bases on multistage fuzzy decision making. The modelling of individual preferences as fuzzy goal and fuzzy constraints, and observed strategic concession behaviour of opponents during negotiation as a fuzzy Markov decision process allows the agent to adapt their negotiation strategies and implied behaviour to improve its individual payoffs. In particular, we show that such adaptive bilateral negotiation strategies can be efficiently derived by an agent from negotiation threads of only two reference cases in its respectively maintained fuzzy transition matrix. Finally, we demonstrate the benefit of applying this solution to different soft-constrained negotiation settings by an initial comparative experimental evaluation.

## I. INTRODUCTION

Automated n-agent negotiation can be considered a multistage decision process where individual rational agents seek to find the best course of actions towards an agreement under the presence of conflicting goals and preferences [1], [2], [3], [4]. Particular attention has been paid to negotiation mechanisms between two agents each of which having their own decision making process with individual utility structures and negotiation strategies, and their application to different domains such as electronic commerce [5], resource allocation and recently service-oriented computing [6]. The problem of bilateral agent negotiation with imprecise or uncertain information about utilities and strategies of the opponent is known to be hard, and many solution approaches have been proposed to cope with it ranging from simple If-then rules, heuristic tactics to more advanced learning and reasoning techniques [4]. Such adaptive negotiation mechanisms mostly assume agents to steadily explore their environment and other agents' behaviour to gain experience from past interactions, or maintain explicit beliefs about utilities, constraints and decision models of their opponents.

In this paper, we present the first adaptive solution based on modelling the bilateral negotiation process as an optimization problem using a multistage fuzzy decision approach. Individual agent preferences are expressed in terms of a fuzzy goal and time-based fuzzy constraints whilst the dynamics of the negotiation is modelled as a fuzzy Markov

decision process which represents the relation between the strategic concession behaviour between both agents involved. In this case, agents' offers and counteroffers correspond to state-action pairs in the negotiation process such that individual fuzzy (possibilistic) instead of probabilistic state transitions enable an agent to utilize uncertain knowledge or beliefs about the concession behaviour of its opponent, and, eventually, obtain a solution in form of state-action policies by fuzzy dynamic programming. We show that the resulting negotiation strategies of an agent successfully adapt to different negotiation behaviours of its opponent based on only two observed reference cases from past interactions that are required to derive its individual fuzzy transition matrix.

A first comparative experimental evaluation of our adaptive solution for different soft-constrained negotiation settings with deadlines revealed that on average agents may gain higher individual utilities compared to the use of traditional heuristic-based mixed strategies.

The remainder of the paper is structured as follows. The next section briefly recalls concepts of multistage fuzzy control deriving the concept of fuzzy state transitions from the stochastic case. Section III details how negotiation strategies are modelled using this approach while Section IV outlines the decision algorithm for a negotiation agent. We present experimental results of a comparative evaluation between heuristic-based mixed negotiation strategies in Section V. Related work is discussed in Section 5, and finally, we conclude in Section VII.

## II. MULTISTAGE FUZZY CONTROL

In this section we recall concepts and notations of multistage decision making in fuzzy environments [7], [8] deriving a system with fuzzy state transitions from the stochastic case. According to [7], [8] a fuzzy decision problem is defined as  $D = C_1 \star C_m \star \dots \star G_1 \star G_n$  where  $D$  is the fuzzy decision in a fuzzy environment specified by  $n$  fuzzy goals  $G_1, \dots, G_n$  and  $m$  fuzzy constraints,  $C_1, \dots, C_m$ . Both, fuzzy goals and constraints, are fuzzy sets in the set of options  $X$ . The  $\star$ -operator represents the aggregation between two fuzzy sets while the optimal decision  $x^* \in X$  satisfies all constraints and goals such that  $D$  with the membership function  $\mu_D(x)$  is maximized with

$$\mu_D(x^*) = \max_{x \in X} \mu_D(x). \quad (1)$$

Assume now, we have a stochastic system under control whose state transitions are governed by the conditional probability function

$$p(x_{t+1} | x_t, u_t), \quad (2)$$

Jan Richter and Ryszard Kowalczyk are with the Centre for Complex Software Systems and Services, Swinburne University of Technology, John Street, Hawthorn, 3122, Australia (email: {richter, rkowalczyk}@groupwise.swin.edu.au).

Matthias Klusch is with the German Research Centre for Artificial Intelligence, Stuhlsatzenhausweg 3, 66123 Saarbruecken, Germany (email: klusch@dfki.de).

where  $x_t, x_{t+1} \in X = \{\sigma_1, \dots, \sigma_n\}$  are the states and  $u_t \in U = \{\alpha_1, \dots, \alpha_m\}$  are the controls of the system at stages  $t$  and  $t+1$  with  $t = 0, 1, \dots, N-1$ . This corresponds to a Markov decision process where the fuzzy goals and constraints, imposed at the respective stages of the process, represent the fuzzy environment in which a decision is to be found given the time-invariant transition function and the fixed termination time. We consider the case where the final outcome at the last stage  $N$  is of most importance so that only one fuzzy goal  $G^N$  is imposed. Using the time-dependent fuzzy constraints at the control stages, the above fuzzy decision problem becomes  $D(x_0) = C^0 \star \dots \star C^{N-1} \star G^N$ , where the fuzzy sets  $C^t$  and  $G^t$  are specified by membership functions  $\mu_{C^t}(x)$  and  $\mu_{G^t}(x)$ , respectively. According to [7], the problem then is to find an optimal sequence of controls  $u_0^*, \dots, u_{N-1}^*$  that maximizes the fuzzy decision  $D$  given the initial state  $x_0$ :

$$\mu_D(u_0^*, \dots, u_{N-1}^* | x_0) = \max_{u_0, \dots, u_{N-1}} [\mu_{C^0}(u_0) \wedge \dots \wedge \mu_{C^{N-1}}(u_{N-1}) \wedge E\mu_{G^N}(x_N)]. \quad (3)$$

In this paper, we use the *min*-operation ( $\wedge$ ) as aggregation operator, though any other t-norm can be used for this purpose as well. The expected goal  $E\mu_{G^N}(x_N)$  is the probability of attaining the fuzzy goal  $\mu_{G^N}$  given the control  $u_{N-1}$  and state  $x_{N-1}$  of the previous stage given by

$$E\mu_{G^N}(x_N) = E\mu_{G^N}(x_N | x_{N-1}, u_{N-1}) = \sum_{x_N \in X} p(x_N | x_{N-1}, u_{N-1}) \mu_{G^N}(x_N). \quad (4)$$

The fuzzy goal  $\mu_{G^N}(x_N)$  can be regarded as a fuzzy event in  $X$  and the expected goal as the non-fuzzy probability of that fuzzy event. Given this goal at stage  $N$  and the constraint at stage  $N-1$ , the fuzzy decision at stage  $N-1$  selects the optimal actions for each state  $x_{N-1}$ . In particular, the fuzzy decision for each state at stage  $N-1$  may be regarded as a fuzzy goal  $\mu_{G^{N-1}}$  at stage  $N-1$  induced by the fuzzy goal  $\mu_{G^N}$  that is used for the next iteration in order to find the optimal actions at stage  $N-2$ . This backward iteration is repeated until we find all optimal actions  $u_{N-1}, u_{N-2}, \dots, u_0$ . Using Eq. (2) to (4), a dynamic programming solution for this multistage decision problem is given by the following recurrence equations [8], [7]:

$$\mu_{G^{N-i}}(x_{N-i}) = \max_{u_{N-i}} [\mu_{C^{N-i}}(u_{N-i}) \wedge E\mu_{G^{N-i+1}}(x_{N-i+1})] \quad (5)$$

$$E\mu_{G^{N-i+1}}(x_{N-i+1}) = \sum_{x_{N-i+1} \in X} p(x_{N-i+1} | x_{N-i}, u_{N-i}) \cdot \mu_{G^{N-i+1}}(x_{N-i+1}), \quad (6)$$

for  $i = 1, \dots, N$ . The solution is expressed in terms of a policy function with  $a_t^* : X \rightarrow U$  being the optimal policy at stages  $t = 0, 1, \dots, N-1$  and  $u_t^* = a_t^*(x_t)$ . The set  $A^* = \{a_0^*, \dots, a_{N-1}^*\}$  then forms the optimal control strategy. For

computing such a solution the dynamics of the system under control, i.e. the state transition function has to be known.

However, in many scenarios these state transitions might not be obtainable by an agent due to limited and uncertain information. In this paper, we therefore consider a fuzzy system under control where the state transition function is a conditional fuzzy relation with the membership function

$$\mu(x_{t+1} | x_t, u_t) \quad (7)$$

with  $\mu : X \times U \times X \rightarrow [0, 1]$ , assigning for each  $x_t \in X$  and  $u_t \in U$  a fuzzy value to the consecutive state  $x_{t+1} \in X$ . The fuzzy state transitions can also be interpreted from the viewpoint of possibility theory [9] in that the possibility degree of a particular state transition reflects how plausible it is to attain a succeeding state given the state and action at the current stage [10]. As a result, given such a fuzzy transition matrix the recurrence equation (7) can be rewritten using the max-min composition as

$$E\mu_{G^{N-i+1}}(x_{N-i+1}) = \max_{x_{N-i+1} \in X} [\mu(x_{N-i+1} | x_{N-i}, u_{N-i}) \wedge \mu_{G^{N-i+1}}(x_{N-i+1})], \quad (8)$$

for  $i = 1, \dots, N$ . Any other s-t norm composition such as the above max-product composition could be used here [8]. Since the expected goal is conditioned on states and actions at stage  $N-i$  it represents a fuzzy relation between  $x_{N-i}$  and  $u_{N-i}$  giving the maximum expected possibility over next states  $x_{N-i+1}$ . The correct notation for the expected goal is hence  $E\mu_{G^{N-i+1}}(x_{N-i}, u_{N-i})$ . In this paper, however, we also use the simplified notation introduced by Kacprzyk [8] interchangeably. In the next section, we will adapt this model to the bilateral negotiation process in order to optimize the course of counteroffers an agent proposes to achieve its desired outcome.

### III. MODELLING NEGOTIATION STRATEGIES

This section shows how to model the decision mechanism of an individual agent for its bilateral negotiation according to the model of multistage fuzzy decision making we introduced in the previous section. Before actually presenting the respective decision algorithm to use by an agent, we introduce the underlying negotiation model as well as the creation of the state and action spaces, the fuzzy transition matrix, and the preferences in form of fuzzy goal and constraints.

#### A. Negotiation Model

For modelling the process of bilateral multi-issue negotiation in which two agents  $a$  and  $b$  propose offers and counteroffers  $o_a^{k_n}$  and  $o_b^{k_{n+1}}$  at discrete time points  $k_n, k_{n+1}$  on a continuous real-valued issue such as price or delivery time, we adopt the prominent model by Faratin et al. [11] with slight modifications in notation in order to make it compliant with our multistage fuzzy decision approach. For each issue under negotiation, each agent has a negotiation interval  $[min_a, max_a]$  defined by its initial and reservation

value; if the intervals of both agents overlap an agreement is generally possible. The sequence of all offers that are exchanged until a given time  $k$  is denoted as the negotiation thread

$$NT_{k_n} = (o_a^{k_1}, o_b^{k_2}, o_a^{k_3}, o_b^{k_4}, \dots, o_b^{k_n}) \quad (9)$$

with  $n \in \mathbb{N}$  and  $o_b^{k_n}$  being the last offer of this thread at time step  $k_n$ . Besides, each agent has a scoring function  $V_{aj}: [\min_{aj}, \max_{aj}] \rightarrow [0, 1]$  associated to each issue  $j$  assigning a score to the current value of  $j$  within its negotiation interval. The additive scoring function for all issues under negotiation is denoted as:  $V_a(o) = \sum_{1 \leq j \leq p} w_{aj} \cdot V_{aj}(o)$  where the weight  $w_a$  represents the relative importance of issue  $j$  for agent  $a$  with  $\sum_j w_{aj} = 1$ . For simplicity, we assume the additive scoring function to be either monotonically increasing or decreasing. The utility functions of negotiating agents usually correspond to such kind of scoring functions possibly with discounts or negotiation costs. Offers are exchanged alternately between agents during their bilateral negotiation until one agent accepts or withdraw from the encounter, or the negotiation deadline is reached:

$$\text{response}_a(k_{n+1}, o_b^{k_n}) = \begin{cases} \text{withdraw} & \text{if } k_{n+1} > k_{max}^a \\ \text{accept}(o_b^{k_n}) & \text{if } V_a(o_b^{k_n}) \geq V_a(o_a^{k_{n+1}}) \\ \text{offer}(o_a^{k_{n+1}}) & \text{otherwise,} \end{cases} \quad (10)$$

where  $o_a^{k_{n+1}}$  is the counterproposal of agent  $a$  given agent  $b$ 's offer  $o_b^{k_n}$  at time step  $k_n$ . This negotiation model allows agents to actually employ different decision models (tactics, strategies) for proposing their offers and counteroffers[12].

### B. State and Action Space

For this purpose, we refer to agent actions instead of controls in the following. Both, the state and action space, need to be in discrete form, where the former covers the complete negotiation range and the latter the negotiation interval of the agent for the issue under negotiation. The discretization method for space  $X$  is given by

$$X = \left\{ \frac{(l-1)(uB-lB)}{n-1} + lB \mid l = 1, \dots, n \right\}. \quad (11)$$

where  $n$  is the total number of discrete points, and  $uB$  and  $lB$  represent the upper and lower boundary of the space, respectively. For example, if we assume that agent  $a$  applies this model, and agent  $b$  makes the first proposal, then for the state space, the boundaries  $uB$  and  $lB$  correspond to the first offers  $o_b^{k_1}$  and  $o_a^{k_2}$ , respectively. For the action space, the upper boundary of agent  $a$  is given by its reservation value  $RV_a$  while the lower boundary is the first offer of agent  $a$ . For simplicity, we use the same discretization factor for both spaces such that the cardinality  $m$  of the action space is given in relation to the total number of states:

$$m = \frac{|RV_a - o_a^{k_2}| \cdot n}{|o_b^{k_1} - o_a^{k_2}|} \quad (12)$$

where  $a$  can be a buyer or seller agent. Since the classical bilateral negotiation model in Section III-A and the multi-stage fuzzy decision model in Section II use different time intervals, the sequence of offers in the negotiation thread is mapped into a state-action form such that offers and counteroffers at time  $k$  and  $k+1$  correspond to states and actions at stage  $t$ . The resulting trajectory of states and actions written as  $TR = (x_0, u_0, x_1, u_1, \dots, x_{t-1}, u_{t-1}, x_t)$  is then equivalent to the offers exchanged during the encounter  $(o_b^{k_1}, o_a^{k_2}, \dots, o_b^{k_{n-2}}, o_a^{k_{n-1}}, o_b^{k_n})$  with  $o_a^{k_{n+1}}$  being the next offer of agent  $a$  responding to the last offer  $o_b^{k_n}$  of agent  $b$ . Thus, the action  $u_t$  (offer  $o_a^{k_{n+1}}$ ) is the action sought at stage  $t$ . Since the state and action space are in discrete form and offers may be proposed in a different (continuous) space, offers are mapped to states  $x_t$  and actions  $u_t$  with

$$\begin{aligned} x_t &= \arg \min_{\sigma \in X} |o_s^{k_n} - \sigma| \\ u_t &= \arg \min_{\alpha \in U} |o_b^{k_{n+1}} - \alpha|. \end{aligned} \quad (13)$$

Agent  $a$  needs at least one offer from its opponent to make a decision according to the responsive negotiation model in Section 2 while the course of actions determined by its policy function then represents its adaptive negotiation strategy in response to its opponent's strategic concession behaviour.

### C. Fuzzy State Transitions

The fuzzy transition matrix of an individual agent  $a$  encodes its fuzzy knowledge about the opponent's strategic concession behaviour and the responses of  $a$  that may lead to an agreement. Only a small number of reference cases (e.g. from past negotiations) is sufficient to obtain the fuzzy state transitions. Such reference cases reflect the range of actions over time as a response to the proposed offers of the opponent and, in that sense, define the dynamic negotiation strategy over agent  $a$ 's possible offers.

In this paper, we focus on the scenario where only two reference cases are available, and their similarity is used to create and update the agent's fuzzy transition matrix. Let  $NT[h]$  be the negotiation thread of case  $h$ , then the thread can be transformed into the state-action form (cf. Section III-B) obtaining the trajectory  $TR[h]$ :

$$TR[h] = (\sigma_{l_0[h]}, \alpha_{v_0[h]} \dots, \alpha_{v_{N[h]-1}[h]}, \sigma_{l_{N[h]}[h]}) \quad (14)$$

where  $N[h]$  is the last stage, and  $\sigma_{l_i[h]}$  and  $\alpha_{v_i[h]}$  are states and actions at stages  $i = 1, \dots, N[h]$  of case  $h$ , respectively. The indices  $l_i[h] \in \{1, \dots, n\}$  and  $v_i[h] \in \{1, \dots, m\}$  correspond to the number of the states and actions of case  $h$  at stage  $i$ . We also write the trajectory of all states of case  $h$  as  $TR_X[h] = (\sigma_{l_0[h]}, \dots, \sigma_{l_{N[h]}[h]})$  and, respectively, of all actions as  $TR_U[h] = (\alpha_{v_1}, \dots, \alpha_{v_{N[h]}})$ . As described in Section II, the policy function recommends at least one action for each state in the state space. In order to create the necessary state transitions, we therefore need to interpolate the trajectory of each case, such that it contains all states of the state space and each state is assigned a particular action. This implies, that also the last state  $\sigma_{l_{N[h]}}$  is assigned an

action  $\alpha_{v_N[h]}$  since it represents the agreement of case  $h$  with  $\sigma_{l_N[h]} = \alpha_{v_N[h]}$ . We choose linear interpolation, and obtain the interpolated states  $\sigma_{l_{i,j}[h]}$  and actions  $\alpha_{v_{i,j}[h]}$  for all  $i = 0, \dots, N[h] - 1$  with

$$l_{i,j}[h] = \begin{cases} l_i[h] + j & \text{for } l_i[h] < l_{i+1}[h] \\ l_i[h] - j & \text{for } l_i[h] > l_{i+1}[h], \end{cases} \quad (15)$$

$$v_{i,j}[h] = \begin{cases} v_i[h] + \lfloor j \cdot \delta_i[h] \rfloor & \text{for } v_i[h] < v_{i+1}[h] \\ v_i[h] - \lfloor j \cdot \delta_i[h] \rfloor & \text{for } v_i[h] > v_{i+1}[h], \end{cases} \quad (16)$$

where  $j = 0, \dots, |l_i[h] - l_{i+1}[h]| - 1$  and  $\delta_i[h]$  is the interpolation factor for two consecutive actions in the trajectory:

$$\delta_i[h] = \frac{|v_i[h] - v_{i+1}[h]| - 1}{|l_i[h] - l_{i+1}[h]| - 1}. \quad (17)$$

Index  $j$  hence depends on the number of interpolated states between two consecutive states in the state trajectory  $TR[h]$ . The interpolated state and action trajectories  $\widetilde{TR}_X[h]$  and  $\widetilde{TR}_U[h]$  can then be written as

$$\begin{aligned} \widetilde{TR}_X[h] &= (\sigma_{l_{0,0}[h]}, \dots, \sigma_{l_{0,j}[h]}, \dots, \sigma_{l_{1,0}[h]}, \dots, \sigma_{l_{N[h],0}[h]}) \\ \widetilde{TR}_U[h] &= (\alpha_{v_{0,0}[h]}, \dots, \alpha_{v_{0,j}[h]}, \dots, \alpha_{v_{1,0}[h]}, \dots, \alpha_{v_{N[h],0}[h]}). \end{aligned} \quad (18)$$

For the state transitions, we use the similarity between the trajectory of case  $h$  and the current behaviour of the opponent represented by the current trajectory  $TR_X[curr]$  at time  $t$ :

$$\begin{aligned} sim_t(TR_X[h], TR_X[curr]) &= \\ \frac{1}{t+1} \sum_{i=0}^t \frac{1 - |\sigma_{l_i[h]} - x_i|}{(\max_{h \in H}(\sigma_{l_i[h]}) - \min_{h \in H} \sigma_{l_i[h]})} \end{aligned} \quad (19)$$

for  $i \leq N[h]$  and  $H$  being the set of all cases. The similarity values provide the necessary fuzzy transitions for each case in comparison to the current negotiation and are updated at each negotiation round. If during the negotiation the current stage exceeds the last stage from a particular case its last offer is used instead. The transition matrix is then created based on an initially zero transition matrix  $\mu(x_{t+1}, x_t, u_t) = 0_{n,m,n}$  for all  $m$  actions and  $n$  states using the similarity values:

$$\begin{aligned} \mu(\sigma_{l_{i+1}[h]} | \sigma_{l_{i,j}[h]}, \alpha_{v_{i,j}[h]}) &= \\ \max[sim_t(TR_X[h], TR_X[curr]), \mu(\sigma_{l_{i+1}[h]} | \sigma_{l_{i,j}[h]}, \alpha_{v_{i,j}[h]})] \end{aligned} \quad (20)$$

for all  $i = 1, \dots, N[h] - 1$ . In scenarios, where only reference cases are used for the state transitions, the expected fuzzy goal at each stage can be derived directly from all cases:

$$\begin{aligned} E\mu_{G^{t+1}}(x_{i+1} | x'_i[h], u_i[h]) &= \\ \max_{h \in H} (sim_t(h) \wedge \mu_{G^{t+1}}(x_{i+1}[h])), \end{aligned} \quad (21)$$

for  $i = t, t+1, \dots, N[h] - 1$ , where  $t$  is the stage in the current negotiation.

This simplifies the recalculation of the expected goal at each stage with respect to the current similarity value. Thus, the computational effort is reduced, especially in scenarios where the fuzzy transition matrix is sparse due to a small

number of cases. To enable inference between the expected goal and the fuzzy constraint, the actions holding zero value in the possibility distribution over all actions need to be interpolated for each state  $\sigma_l$  as follows:

$$\begin{aligned} E\mu_{G^{t+1}}(x_{t+1} | \sigma_l, \alpha_v) &= E\mu_{G^{t+1}}(\sigma_l, \alpha_v) = \\ \frac{E\mu_{G^{t+1}}(\sigma_l, \alpha_{v_2}) - E\mu_{G^{t+1}}(\sigma_l, \alpha_{v_1})}{v_2 - v_1} * (v - v_1) + \alpha_{v_1}, \end{aligned} \quad (22)$$

with  $v, v_1, v_2 \in \{1, \dots, m\}$  under the condition that  $v_1 < v < v_2$  and  $E\mu_{G^{t+1}}(\sigma_l, \alpha_{v_1}), E\mu_{G^{t+1}}(\sigma_l, \alpha_{v_2}) > 0$ . If, however, the boundary actions  $\alpha_1$  or  $\alpha_m$  are zero, we can replace them by very small values greater zero before applying the interpolation method to obtain a non-zero possibility distribution over the whole actions space. The rationale behind is that a limited number of cases may be sufficient to propose an agent's response. Since the expected goal then holds values for all states and actions after the interpolation, the model can also propose actions not covered by any of the reference cases. Therefore, this approach provides a great flexibility towards the creation of adaptive negotiation strategies. It should be noted that other methods can be used to create and update the transition matrix, depending on the pre-existing knowledge or beliefs about the opponent.

#### D. Preferences as Fuzzy Goal and Constraints

The fuzzy goal and the set of fuzzy constraints represent the preferences of an agent over its opponent's offers (states) and its own offers (actions), respectively. Due to the qualitative nature of the multistage fuzzy decision making model, the goal specifies the agent's preferred ordering of all states in the state space for the decision process. In the context of negotiation, the degree of membership in the fuzzy goal increases for states closer to the initial value of the agent as they are more preferable to states close to the initial offer of the opponent. Thus, membership degrees for all states in the fuzzy goal have to be non-zero, as otherwise, a state might never be reached as a final or intermediate state. Whilst

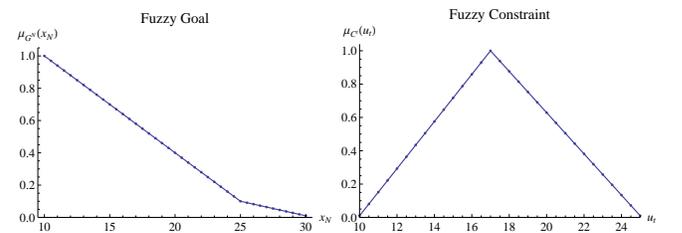


Fig. 1. Example Fuzzy Goal and Constraint

the fuzzy goal represents a time-independent preference over states, the constraints constitute a time-dependent preference over the range of possible offers of the modelling agent (an example for both a fuzzy goal and fuzzy constraint is shown in Figure 1). The effect of a constraint can vary depending on the shape and the support of its fuzzy set. In general, the larger the support and area of the fuzzy constraint the stronger the influence of the cases on the actions and vice

versa. For simplicity and easy specification, constraints are typically normalized. However, in a negotiation, the other agent may choose states close to its initial offer in the beginning with small membership degrees in the fuzzy goal, such that membership degrees for all state-action pairs in the expected goal relation (Eq. 8) become also small. As a result, constraints may have a low effect on the actions as they are normalized and may completely overlay the expected goal distribution over the actions for a particular state. The influence is increased by scaling the fuzzy constraints down, e.g. to the maximum of the expected goal, before it is applied (cf. Eq. 5):

$$\hat{\mu}_{C^t}(\alpha) = \mu_{C^t}(\alpha) \cdot \max_{\alpha \in U, \sigma \in X} (E\mu_{G^{t+1}}(\sigma, \alpha)) \quad (23)$$

for all  $\alpha \in U$ . This method ensures a high effect of the individual constraints on the transition matrix and therefore on the cases during the encounter. The scaling factor for the constraints depends on the preference of the agent and can be different from the one shown above.

---

**Algorithm 1** Decision algorithm of the buyer agent  $b$ 


---

```

1: Exchange first offers  $o_s^{k_1}$  ( $= x_0$ ) and  $o_b^{k_2}$  ( $= u_0$ )
2: Create Action and State Space from  $x_0$  and  $u_0$  ▷ cf. Eq.
3: for all reference cases  $h \in H$  do
4:    $\widehat{TR}[h] \leftarrow$  Transform  $TR[h]$  into state-action form
5: end for
6:  $end \rightarrow$  False
7:  $t \leftarrow 1$ 
8: while  $end \neq True$  do
9:   if  $t > t_{max}^b$  or timeout then
10:     Withdraw from negotiation
11:      $end \leftarrow True$ 
12:   else
13:      $x_t \leftarrow$  Next offer from opponent  $s$ 
14:     if  $s$  accepts last offer  $u_{t-1}$  of agent  $b$  then
15:        $end \leftarrow True$ 
16:     else
17:        $a^* \leftarrow$  GETPOLICY( $t, TR_X[curr]$ )
18:        $u_t^* \leftarrow a_t^*(x_t)$ 
19:       if  $U_a(x_t) \geq U_a(u_t^*)$  then
20:         Accept last offer  $x_t$  of  $s$ 
21:          $end \leftarrow True$ 
22:       else
23:         Propose counteroffer  $u_t^*$ 
24:       end if
25:     end if
26:   end if
27:    $t \leftarrow t + 1$ 
28: end while
29: end algorithm

```

---

#### IV. DECISION ALGORITHM

The question remains of how an individual agent applies the multistage fuzzy decision model from previous sections to specify its negotiation strategy. In the following, we

provide the appropriate decision algorithm for a buyer agent  $b$  negotiating with a seller agent  $s$  assuming that  $s$  proposes the first offer. Algorithm 1 details the communication mechanism with the opponent in terms of the offer exchange during negotiation (lines 9 to 33) according to Section III-A after the agent created its state and action space and transformed the reference cases into the respective form (cf. Section III-B and III-C) using first offers of both parties (lines 1 to 5). For simplicity, the agent uses the number of negotiation rounds to specify its negotiation deadline ( $t_{max}^b$ ) for its withdrawal instead of using a real time measure. A negotiation round consists of one offer proposal of both agents and thus corresponds to one stage in the multistage fuzzy decision process. However, the agent may abort the negotiation after a *timeout* period (line 9), where it receives no response after a predefined threshold time. This timeout period naturally depends on the conditions and preferences of the system and the agent.

---

**Algorithm 2** Get action policy at stage  $t$ 


---

```

1: procedure GETPOLICY( $t, TR_X[curr]$ )
2:    $E\mu_{G^N}(x_{N-1}, u_{N-1}) \leftarrow 0_{n,m}$ 
3:    $k \leftarrow N$ 
4:   while  $k > t$  do
5:      $k \leftarrow k - 1$ 
6:     for all cases  $h \in H$  do
7:        $sim[h] \leftarrow sim_t(TR_X[h], TR_X[curr])$ 
8:       for all  $i = 1, \dots, N[h] - 1$  do
9:          $g \leftarrow \min(sim[h], \mu_{G^{k+1}}(\sigma_{l_{i+1}[h]})$ 
10:        for all  $j = 0, \dots, |l_i[h] - l_{i+1}[h]| - 1$  do
11:           $E\mu_{G^{k+1}}(\sigma_{l_{i,j}[h]}, \alpha_{v_{i,j}[h]}) \leftarrow \dots$ 
12:           $\dots \max(E\mu_{G^{k+1}}(\sigma_{l_{i,j}[h]}, \alpha_{v_{i,j}[h]}), g)$ 
13:        end for
14:      end for
15:    end for
16:     $\hat{\mu}_{C^k}(u_k) \leftarrow \mu_{C^k}(u_k) \cdot \dots$ 
17:     $\dots \max_{\alpha \in U, \sigma \in X} (E\mu_{G^{k+1}}(\sigma, \alpha))$ 
18:    for all  $l = 1, \dots, n$  do
19:      Interpolate  $E\mu_{G^{k+1}}(\sigma_l, u_k)$ 
20:       $\mu_{\sigma_l}(u_k) \leftarrow \dots$ 
21:       $\dots \min_{\alpha \in U} (\hat{\mu}_{C^k}(\alpha), E\mu_{G^{k+1}}(\sigma_l, \alpha))$ 
22:       $\mu_{G^k}(\sigma_l) \leftarrow \max_{\alpha \in U} (\mu_{\sigma_l}(\alpha))$ 
23:       $a_k^*(\sigma_l) \leftarrow \arg \max_{\alpha \in U} (\mu_{\sigma_l}(\alpha))$ 
24:    end for
25:  end while
26:   $a^* \leftarrow \{a_k^*, \dots, a_{N-1}^*\}$ 
27:  Return  $a^*$ 
28: end procedure

```

---

Algorithm 2 details how the multistage fuzzy decision model is applied to obtain action policies throughout all stages of the negotiation encounter for the agent. It represents the fuzzy dynamic programming method including the creation of the expected goal matrix from the reference cases (line 2 to 15) and its interpolation. It should be noted that the linear interpolation to transform the reference cases into the

state-action form (line 4, Algorithm 1) and in the expected goal matrix (line 19, Algorithm 2) is straightforward (cf. Section III-C) and therefore not detailed here due to limitations in space.

## V. EXPERIMENTAL EVALUATION

In the evaluation we test our approach against static mixed strategies in a bilateral, single-issue negotiation environment proposed in [13] and [11] with partial overlap of negotiation intervals. The tactics employed by the mixed strategies are the time- or behaviour-dependent tactics introduced by Faratin et al [12]. The set of all possible strategies is

TABLE I  
EXPERIMENT SETTINGS

| Mixed Strategies    |   |
|---------------------|---|
| Time-dependent      | Boulware: $B = \{\beta   \beta \in \{0.1, 0.3, 0.5\}\}$<br>Linear: $L = \{\beta   \beta \in \{1\}\}$<br>Conceder: $C = \{\beta   \beta \in \{4, 6, 8\}\}$ |
| Behaviour-dependent | Absolute TFT: $a: \delta = 1, R(M) = 0$<br>Relative TFT: $r: \delta = 1$  |
| Weights             | Small: $\gamma \in \{0.1, 0.2, 0.3\}$<br>Medium: $\gamma \in \{0.4, 0.5, 0.6\}$<br>Large: $\gamma \in \{0.7, 0.8, 0.9\}$                                  |

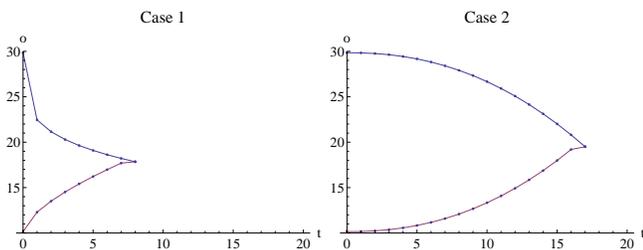


Fig. 2. Reference cases

constructed with  $ST = (C \cup L \cup B) \times \{a, r\} \times (S \cup M \cup L)$  using the settings in Table I. We compare our multistage fuzzy decision strategy with a random selected strategy of the set  $ST$  from the viewpoint of a buyer agent while the opponent (seller) plays a particular subset of static mixed strategies. In that sense the buyer with the random selected strategy plays all strategies of the set  $ST$  against one subset of  $ST$  (seller) where the average gained utility is compared with the utility of the agent using the adaptive strategy. We use simple negotiation environment settings where deadlines are equal for both agents with  $k_{max}^b = k_{max}^s = 20$ . Negotiation intervals are equal for both agents with  $\theta = max_b - min_b = max_s - min_s = 15$  and have a partial overlap with  $\phi \in \{0.5\}$ . With  $min_b = 10$  and  $max_b = 25$  the seller's interval values can be calculated as  $min_s = min_b + \phi\theta = 15$  and  $max_s = min_s + \theta = 30$ . In order to measure the performance of our agents we use the cost-adjusted utility  $cost(t) = \tanh(t * c)$  with a small communication cost  $c = 0.01$ . This takes negotiation time into account. Scoring functions are assumed to be linear and result in a value of 0.1 for the agent if an agreement is reached at its reservation value, e.g. with  $V_b(x) = 0.9 * (max_b - x) / (max_b - min_b) + 0.1$  for

the buyer. Thus, successful negotiations are scored higher as failed negotiations. In order to create a scenario where limited information about the opponent is available we use two reference cases. For simplicity, a discretization factor of 0.2 is applied to create the action and state spaces. Figure 2 shows the chosen cases for this experiment. All fuzzy constraints are specified by an isosceles triangle membership function where, similar to the polynomial time-dependent tactics, three types of constraints behaviour over time are used in the experiment: conceder ( $\beta = 3$ ), linear ( $\beta = 1$ ) and boulware ( $\beta = 0.3$ ). Figure 3 shows all three types of constraint behaviour including the spread of each fuzzy constraint (i.e. the upper and lower boundaries). The membership function of the fuzzy goal represents the scoring function, but, as explained in Section III-D, goal values for states above the reservation value can not be zero and therefore further decrease. The experimental results without and with communication costs are shown in Figure 3 below the corresponding applied constraints. The light bars represent the average value of utility of a negotiation agent using the random selected strategy and the opponent playing a particular strategy group. The dark (blue) bars show the gain in utility for the agent using the multistage fuzzy decision model. As we can see our adaptive strategy performs better than the random strategy selection in almost all scenarios, whereas the improvement is higher when the seller chooses linear or boulware mixed strategies. The reason for that is that the multistage fuzzy decision strategy is able to adapt to the behaviour of the opponent over time whereas the static mixed strategy is behaviour-dependent based on the weight and the pre-defined imitative tactic. As the two cases represent a weak boulware and conceder behaviour of the agent, it is able to choose a different course of actions when the opponent changes its behaviour.

It should be noted that the gain in utility depends to a high degree on the choice of the reference cases as they constitute the course of actions that can possibly lead to an agreement. The agents can adjust or add reference cases in future negotiations in order to increase their utility gain or the number of agreements. The main advantage is the flexibility of adding new negotiation patterns to the transition matrix and adjusting the fuzzy constraints to achieve a large range of different adaptive behaviours.

## VI. RELATED WORK

The subject of adaptive negotiation in general, and how to create negotiation strategies that adapt to the opponent behaviour in particular is not new to the field of AI. For example, Matos and Sierra [14] present a case-based reasoning-driven approach that lets agents use past successful interactions to negotiate similar agreements by respectively (case-based) adjusting combined decision function parameters. In fact, alongside the negotiation thread of each case the parameter values of the applied strategies are required. However, that inhibits the use of cases by agents with different individual decision models.

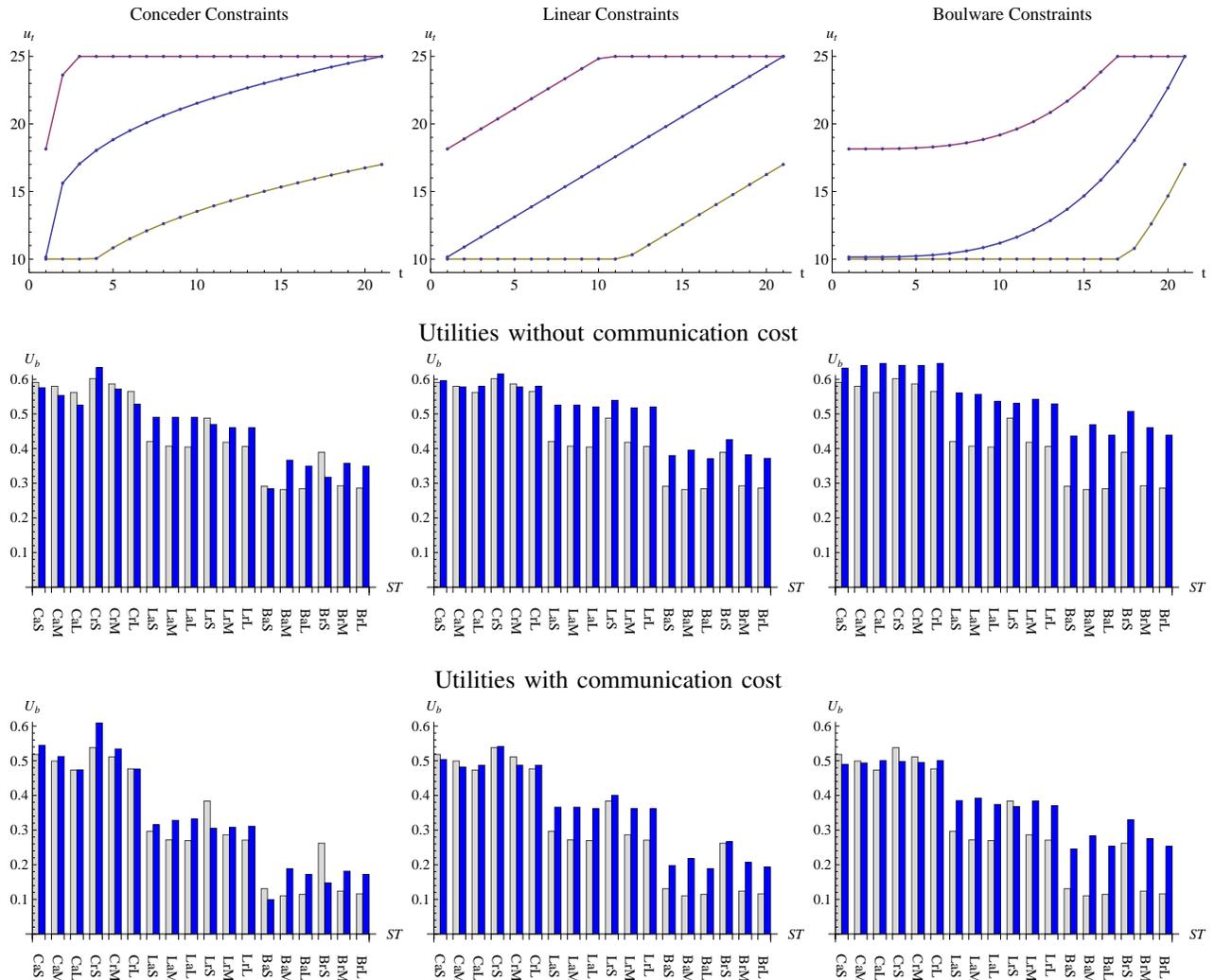


Fig. 3. Evaluation results for different fuzzy constraints

Similar to our approach, Wong et al [15] use observed concessions to capture past negotiation cases and apply certain filters to select the best one. It differs from our approach in that they do not allow for reasoning on and interpolation between the cases and the preferences of an agent. For reasons of space, we omit to cover relevant work more comprehensively, but in general, most of them require an agent to extensively explore the behaviour of partners during negotiation, or assume that it has been equipped with prior knowledge about the opponent such as its utility structure and decision model [4].

The application of possibility theory to negotiation has been proposed in, for example, [16] where the decision on potentially beneficial negotiation partners bases on the expected qualitative utility but without modelling the negotiation process itself as a fuzzy (or possibilistic) Markov decision process. In fact, only a few approaches exist so far. For example, Narayanan and Jennings [17] model the agent's behaviour by defining the states in terms of resource availability, deadlines and reservation values where counteroffers

are proposed based on the opponent's offers and changes in those three realms. It is shown that agreements can be achieved much faster when both agents use this algorithm, but no results for cases are provided where only one agent uses this strategy. Similar to our method, Teuteberg [18] models the behaviour of the opponent, but uses a probabilistic approach to generate the transition matrix based on a predefined set of opponent tactics. The major disadvantage of such an approach is the large number of negotiations required to obtain sufficient empirical data for reliable state transitions.

Negotiation has also been modelled as a fuzzy constraint satisfaction problem [19] where constraints, preferences and objectives are represented uniformly as fuzzy sets which are distributed among the agents and iteratively relaxed during the exchange of offers [4]. The search process is guided by ordering and pruning the search space but still requires negotiation strategies for proposing offers [20]. Based on the seminal paper of Bellmann and Zadeh [7] decision making in fuzzy environments has been studied and extended by many researchers, such as Kacprzyk [8], Iwamoto [21] and Dubois

et al [22], and has been applied in many areas including resource allocation, planning or scheduling [8].

However, to the best of our knowledge, multistage fuzzy decision making has not been used to model bilateral agent-based negotiation strategies so far.

## VII. CONCLUSIONS

We have presented the first adaptive solution to bilateral negotiation under uncertainty that bases on multistage fuzzy decision making, where agents preferences are expressed by a fuzzy goal and fuzzy constraints. A fuzzy Markov decision process represents the uncertain and limited knowledge of an agent about the strategic concession behaviour of its opponent. Our initial experimental evaluation revealed that the respective agent decision algorithm successfully adapts to different negotiation behaviours of opponents and achieves on average higher utilities than the traditional heuristic tactic-based negotiation strategies. The multistage fuzzy decision making-based negotiation model can be applied to many real world scenarios where available information about agents' behaviours, preferences and constraints are soft. Future work includes multistage fuzzy decision-based negotiation with fuzzy or implicit termination times, and evaluation of scenarios where both agents use this decision model.

## REFERENCES

- [1] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, M. Wooldridge, and C. Sierra, "Automated Negotiation : Prospects, Methods and Challenges," *Group Decision and Negotiation*, vol. 10, pp. 199–215, 2001.
- [2] H. Raiffa, *The Art and Science of Negotiation*. Cambridge, Mass.: Harvard University Press, 1982.
- [3] J. S. Rosenschein and G. Zlotkin, *Rules of encounter: designing conventions for automated negotiation among computers*. Cambridge, MA, USA: MIT Press, 1994.
- [4] P. Braun, J. Brzostowski, G. Kersten, J. Kim, R. Kowalczyk, S. Strecker, and R. Vahidov, *e-Negotiation Systems and Software Agents: Methods, Models, and Applications*. Springer, 2006, pp. 271–300.
- [5] J. R. Oliver, "A Machine Learning Approach to Automated Negotiation and Prospects for Electronic Commerce," *Journal of Management Information Systems*, vol. 13, no. 3, 1997.
- [6] P. Faratin, "Negotiation Among Groups of Autonomous Computational Agents," *Int. Journal of Robotics and Autonomous Systems*, 1998.
- [7] R. E. Bellman and L. A. Zadeh, "Decision Making in a Fuzzy Environment," *Management Science*, vol. 17, no. 4, pp. 141–164, December 1970.
- [8] Kacprzyk, *Multistage Fuzzy Control: A Model-based Approach to Fuzzy Control and Decision Making*. New York, NY, USA: John Wiley & Sons, Inc., 1997.
- [9] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 100, no. Supplement 1, pp. 9–34, 1999.
- [10] R. Sabbadin, "Possibilistic Markov decision processes," *Engineering Applications of Artificial Intelligence*, vol. 14, no. 3, pp. 287–300, June 2001.
- [11] P. Faratin, "Automated Service Negotiation Between Autonomous Computational Agents," Doctor of Philosophy, University of London, 2000.
- [12] P. Faratin, C. Sierra, and N. R. Jennings, "Negotiation Decision Functions for Autonomous Agents," *Robotics and Autonomous Systems*, vol. 24, no. 3-4, pp. 159–182, 1998.
- [13] J. Brzostowski and R. Kowalczyk, "On Possibilistic Case-based Reasoning for Selecting Partners for Multi-attribute Agent Negotiation," in *Proceedings of AAMAS'05*. New York, NY, USA: ACM, 2005, pp. 273–279.
- [14] N. Matos and C. Sierra, "Evolutionary Computing and Negotiating Agents," in *First International Workshop on Agent Mediated Electronic Trading on Agent Mediated E-Commerce*. London, UK: Springer, 1999, pp. 126–150.
- [15] W. Y. Wong, D. M. Zhang, and M. Kara-Ali, "Towards an Experience-Based Negotiation Agent," in *Proceedings of the 4th International Workshop on Cooperative Information Agents IV*. London, UK: Springer, 2000, pp. 131–142.
- [16] J. Brzostowski and R. Kowalczyk, "On Possibilistic Case-Based Reasoning for Selecting Partners in Multi-agent Negotiation," in *17th Australian Joint Conference on Artificial Intelligence*, ser. LNCS, X. Y. Geoffrey I. Webb, Ed. Springer, 2005, pp. 694–705.
- [17] V. Narayanan and N. R. Jennings, "An Adaptive Bilateral Negotiation Model for E-Commerce Settings," in *7th Int. IEEE Conference on E-Commerce Technology*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 34–41.
- [18] F. Teuteberg and K. Kurbel, "Anticipating Agents' Negotiation Strategies in an E-marketplace Using Belief Models," in *Proceeding of the 5th International Conference on Business Information Systems*, W. Abramowicz, Ed., Poznan, Poland, 2002.
- [19] X. Luo, N. R. Jennings, N. Shadbolt, H.-f. Leung, and J. H.-m. Lee, "A Fuzzy Constraint Based Model for Bilateral, Multi-issue Negotiations in Semi-Competitive Environments," *Artificial Intelligence*, vol. 148, no. 1-2, pp. 53– 102, 2003.
- [20] R. Kowalczyk, "Fuzzy e-negotiation agents," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 6, no. 5, pp. 337–347, 2002.
- [21] S. Iwamoto, K. Tsurusaki, and T. Fujita, "Conditional Decision-Making in Fuzzy Environment," *Journal of the Operations Research*, vol. 42, no. 2, pp. 198–218, 1999.
- [22] D. Dubois, H. Fargier, J. Lang, H. Prade, and R. Sabbadin, "Qualitative decision theory and multistage decision making - A possibilistic approach," in *Proceedings of the European Workshop on Fuzzy Decision Analysis for Management, Planning and Optimization*, 1996.