Affective Decision-Making in Ultimatum Game: Responder Case*

Jitka Homolová, Anastasija Černecka, Tatiana V.Guy, and Miroslav Kárný

The Czech Academy of Sciences, Institute of Information Theory and Automation, Adaptive System Department, P.O. Box 18, 182 08 Prague 8,
jkratoch@utia.cas.cz, chernetskaan@gmail.com, guy@utia.cas.cz, school@utia.cas.cz

http://www.utia.cas.cz/AS

Abstract. Any artificial intelligence (AI) includes dynamic locally independent decision-makers interacting in a distributed way. Such interaction cannot be realized without solving negotiation, cooperation and coordination problems. Another important aspect of human decision making is how human feelings are integrated into decisions. Importance of this topic has been increasing in recent years. It is relevant in many modern areas. The paper focuses on prescriptive affective decision making in Ultimatum Game (UG). This simple theoretical tool allows to take into account the influence of emotions. In our approach, one of the players (responder) is modelled via Markov decision process. The responder's reward function is the weighted combination of two components: economic and emotional. The first component represents pure monetary profit while the second one reflects emotional state of the responder. The proposed model is tested on simulated data. The obtained simulation results partially reflect descriptive features of human decision making observed by psychologists.

Keywords: Decision making · Emotions in economic game · Markov Decision Process · Ultimatum Game

1 Introduction

In decision making (DM), the sequence of actions from a predefined set is considered. It ensures the system to behave in accordance with the decision maker's aim. The decision theory finds this sequence of optimal actions under the assumption of ideal decision maker, so-called "cold gain maximizer", who is fully rational. However, extensive empirical evidence indicates that people do not behave as cold gain maximizers. They consider other aspects as well. Thus to model human DM, it is necessary to optimise not only the pure economic profit of the decision maker but also human aspects such as fairness or emotions.

Majority of literature studies how decision results affect human feelings, for instance McGraw et al. [8], though influence of emotions on decisions is more

* Supported by GA16-09848S, LTC18075 and EU-COST Action CA16228

important for practical applications. Fehr and Schmidt [4] and Rabin [10] focus on fairness and include emotions only *indirectly* in fairness elements. A paper by Cox, Friedman and Gjerstad [3] contains an emotional parameter which reflects status and reciprocity between adversaries. In utility function, the parameter multiplies monetary profit of an opponent so that utility increases with increasing emotion. The emotional parameter is to be determined experimentally and could be positive as well as negative. Although this model deals with so called *emotional state function*, emotions included are very specific and related only to relationship between the players.

A model that is more closely related to emotions was introduced by Tamarit and Sanchez [12]. Their model is based on two psychological theories. This approach has good foundation, nevertheless, it has some limitations (in particular, limited ability to predict human players) [6]. Another shortcoming of the approach is that initial player's emotions are not considered. All mentioned approaches give many important insights into human decision making and provide solution of many subtasks, but none of them systematically solves affective DM.

In this paper we include an emotional component to the reward function of an optimizing responder and, based on empirical evidence, proposed a model of emotion development. The emotional state of the responder is influenced by the course of the game and vice versa. In our study we introduce 5 emotional states: 1 is the least positive and 5 is the most positive influence. The proposed model does not distinguish different qualitative types of emotions but only discrete states of them from the worst to the best reflecting their effect on DM. These states are the resulting emotional state given by a combination of different basic emotions, which thus reflects the final current emotional state of the player.

The paper outline is as follows. Section 2 introduces necessary definitions and mathematical apparatus. Section 3 describes the UG game as MDP, mainly the model of the responder, introduces the reward function and the model of the responder's emotional state development, as well as derives the corresponding DM strategy based on MDP and dynamic programming formalism. Section 4 describes the experimental part of this paper including the detailed description of all individual types of responders. The paper is concluded by a summary of findings and description of the main open problems in Section 5.

2 Preliminaries

2.1 General notation

x_t	value of random variable x at discrete time t;
$p_t(x\|y)$	conditional probability density function of a random variable x conditioned on random variable y, known at discrete time t;
$E[x], E[x\|y]$	expectation of random variable x and conditional expectation of x conditioned on random variable y;
$\delta(x, y)$	Kronecker delta function that equals 1 for $x = y$ and 0 otherwise.

2.2 Ultimatum Game Rules

The Ultimatum Game (UG) [5] is a simple game on which 'non-rational' human decision making could be illustrated and modeled. UG is an economic game for two players: proposer and responder. The proposer offers how to divide an amount of money between him/her and the responder. The task of the responder is to accept or to reject the offer. If the responder accepts, the amount is divided in accordance with the offer, otherwise both players get nothing. The DM aim of each player is to maximize his/her own profit.

This game can simulate the most common trade situations in real life: a seller of goods and a buyer have similar roles to the proposer and the responder in UG. The seller proposes the price and the buyer decides whether the deal will take place as well as the UG respondent.

2.3 Markov Decision Process

The paper models the responder in the Ultimatum Game (UG) via Markov decision process (MDP) [9]. A decision maker chooses action a_t in decision epoch $t \in \mathbf{T}$ based on observed state s_t, that evolves probabilistically based on state s_{t-1} and action a_{t-1} selected by the decision-maker. This condition is usually called *Markov assumption*.

At each decision epoch t, the system stands at state s_t. Then action a_t is realized and new system state s_{t+1} is determined stochastically by transition probability $p_t(s_{t+1}|a_t, s_t)$. After this transition, the decision-maker gains reward $r_t(s_{t+1}, a_t, s_t)$. Its value indicates a degree of reaching his/her DM preferences.

Formally, MDP is determined by the following definition. It assumes that state s_{t+1} is fully observable after choosing action a_t.

Definition 1. *(Markov decision process) The discrete-time **Markov decision process** is defined as a 5-tuple $\{\mathbf{T}, \mathbf{S}, \mathbf{A}, p, r\}$, where $\mathbf{T}$ denotes a discrete finite set of decision epochs; $\mathbf{T} = \{1, 2, ..., N\}$, $N \in \mathbb{N}$, $\mathbf{S}$ is a discrete finite set of states, $\mathbf{S} = \cup_{t\in\mathbf{T}}\mathbf{S}_t$, $\mathbf{S}_t$ is a set of possible states and $s_t \in \mathbf{S}_t \subset \mathbf{S}$ is the state at decision epoch $t \in \mathbf{T}$; $\mathbf{A}$ denotes a discrete finite set of actions, $\mathbf{A} = \cup_{t\in\mathbf{T}}\mathbf{A}_t$ is a set of possible actions and $a_t \in \mathbf{A}_t \subset \mathbf{A}$ is the action chosen at decision epoch $t \in \mathbf{T}$. The function p_t represents a transition probability function known at decision epoch t. The probability function $p_t(s_{t+1}|a_t, s_t)$ is a probability that action $a_t \in \mathbf{A}_t$ changes state $s_t \in \mathbf{S}_t$ into state $s_{t+1} \in \mathbf{S}_{t+1}$ satisfying the condition $\sum_{s_{t+1}\in\mathbf{S}_{t+1}} p_t(s_{t+1}|a_t, s_t) = 1$, $\forall t \in \mathbf{T}$, $\forall s_t \in \mathbf{S}_t$, $\forall a_t \in \mathbf{A}_t$. Finally, $r_t : \mathbf{S}_{t+1} \times \mathbf{A}_t \times \mathbf{S}_t \longrightarrow \mathbb{R}$ stands for a reward function that is received after taking action a_t and transiting from state s_t into state s_{t+1}, $r_t = r_t(s_{t+1}, a_t, s_t)$.*

Further on, time-invariant reward function $r_t(s_{t+1}, a_t, s_t) = r(s_{t+1}, a_t, s_t)$ is considered. The solution to MDP (**Optimal DM policy**) is a sequence of optimal DM rules $\left\{p_t^{opt}(a_\tau|s_\tau)\right\}_{\tau=t}^{t+T-1}$ that maximizes the expected total reward:

$$\pi_{t,T}^{opt} \in \arg\max_{\pi_{t,T} \in \Pi_t} E\left[\sum_{\tau=t}^{t+T-1} r(s_{\tau+1}, a_\tau, s_\tau) \mid s_t\right]. \quad (1)$$

The *expected reward function* is defined as:

$$E_t\big[r(s_{t+1}, a_t, s_t)|s_t\big] = \sum_{s_{t+1}\in\mathbf{S}, a_t\in\mathbf{A}} r(s_{t+1}, a_t, s_t) \cdot p_t(s_{t+1}, a_t|s_t), \quad (2)$$

where

$$p_t(s_{t+1}, a_t|s_t) = p_t(s_{t+1}|s_t, a_t) \cdot p_t(a_t|s_t). \quad (3)$$

To compute an optimal policy (1), the first factor in (3) is needed. In the inspected UG, it is model of the proposer.

3 Ultimatum Game as Markov Decision Process: The responder's Strategy

The paper considers influence of the emotional state of the responder on his/her decision making. The responder intends not only to maximize his/her monetary profit, but also to retain the best achievable emotional state or to improve it at least. The DM is thus influenced by so-called *expected emotions* [7] and as such is modelled.

3.1 Model of the Responder

The considered UG is treated as an N-round repeated game and decision epochs $t \in \mathbf{T} = \{1, 2, ..., N\}$ correspond to these game rounds. There is an constant amount of money $q \in \mathbb{N}$ to be divided in each round of the game. The proposer forms the responder's system and as such is modelled.

Definition 2. *(UG as MDP of the responder) The MDP for proposed UG is modeled by* $\{\mathbf{T}, \mathbf{S}, \mathbf{A}, p, r\}$*, see Definition 1, where* $\mathbf{A} = \{1, 2\}$ *is a set of all possible actions of the responder, where* $a_t = 1$ *is the rejection and* $a_t = 2$ *the acceptance of the actual proposer's offer;* $s_t = (o_t, m_t) \in \mathbb{S} = \mathbf{O} \times \mathbf{M} \subset \mathbb{N}^2$ *is a state of the system in round* $t \in \mathbf{T}$*;* $o_t \in \mathbf{O} = \{o_{min}, ..., o_{max}\} \in \mathbb{N}$ *denotes a proposer's offer,* $0 < o_{min}$*,* $o_{max} < q$*;* $m_t =\in \mathbf{M} = \{m_{min}, ..., m_{max}\} \in \mathbb{N}$ *represents an emotional state of the responder. The corresponding reward of the responder is defined by* $r(s_{t+1}, a_t, s_t) = (1-\omega)\cdot(a_t - 1)\cdot o_t + \omega \cdot m_{t+1}$*, where* $\omega \in [0, 1]$ *is a weight reflecting balance between importance of emotional and economic components in the responder's reward.*

The first part of the reward function formula describes an economic profit of the game round. The total economic profit of the responder after $t \in \mathbf{T}$ rounds is a sum of the actual profits at each round, i.e. $z_R(t) = \sum_{i=1}^{t} z_i = \sum_{i=1}^{t} (a_i - 1) \cdot o_i$.

The second part of the reward is a responder's "emotional" profit where m_{t+1} is a deterministic, dynamically changing emotional state. It depends on action a_t and proposer's offer o_t as follows:

$$m_{t+1}(a_t, o_t, m_t) = \begin{cases} \min\{m_t + \chi_t(a_t, o_t, m_t), m_{max}\}, & \chi_t(a_t, o_t, m_t) \geq 0, \\ \max\{m_t + \chi_t(a_t, o_t, m_t), m_{min}\}, & \chi_t(a_t, o_t, m_t) < 0, \end{cases} \tag{4}$$

where

$$\chi_t(a_t, o_t, m_t) = \begin{cases} -1, & a_t = 1, \\ 0, & a_t = 2 \wedge o_t \in \langle o_{min}, p_o) \wedge m_t \in (p_m, m_{max}\rangle, \\ 1, & a_t = 2 \wedge \Big((o_t \in \langle o_{min}, p_o) \wedge m_t \in \langle m_{min}, p_m) \Big), \\ & \qquad \vee o_t \in \langle p_o, o_{max}\rangle \Big). \end{cases}$$

Parameters $p_o \in (o_{min}, o_{max})$ and $p_m \in (m_{min}, m_{max})$ are specific to a given responder. They represent threshold values of the offers and the emotional states. These parameters cause the change of responder's behaviour.

The emotional evolution scenario is as follows. The emotional state of the responder (4) is supposed to be dependent on the value of accepted or rejected offer o_t and current emotional state m_t. In case of any rejection, the emotional state decreases anyway. The responder lost the chance to earn amount of money and this loss always worsens his/her emotional state.

Another situation arises in case of acceptance of the offer. The offers of the proposer range between values o_{min} and o_{max}. The responder sets his/her own threshold value $p_o \in (o_{min}, o_{max})$ of the proposer's offer. These thresholds determine the "emotional stability point". If accepted offer is higher than this limit, the responder's emotional state always increases. The responder has gotten more than he has dreamed of, which always improves his/her emotional state.

The situation is less unambiguous if the accepted offer is below limit p_o. The responder gets less than he/she would have liked. Yet his/her feelings may be different - the emotional state may remain the same, but may still improve. It depends on his personal parameter $p_m \in (m_{min}, m_{max})$ which determine his/her "emotional tolerance point". In case of the previous "good feeling", i.e. $m_t \geq p_m$, the responder keeps his/her emotional state at the same value. The responder attention is more denied to loss: by adopting a lower offer, the profit is smaller than he/she would theoretically have had because the actual profit is below the acceptable threshold p_o.

On the other hand, if the responder accepts the offer below the limit in "bad mood", i.e. $m_t < p_m$, his/her attention is more denied to profit: there is at least a small real profit. His/her possible loss is reduced compared to the case of the rejection. The responder gains at least something, his/her financial situation improves, and thus the emotional state improves too.

3.2 Dynamic Programming in UG

The responder chooses action $a_t \in A$ based on the *randomized DM rule* $p_t(a_t|s_t)$ in each decision epoch $t \in \mathbf{T}$. The DM rule is a non-negative function representing probability of action a_t in state $s_t \in \mathbf{S}$. The responder searches for an optimal *DM policy* $\pi_{t,T}$ (a sequence of DM rules mapping states to actions) maximizing expected reward over some horizon T through the following algorithm:

$$
\begin{aligned}
a_\tau^{opt} \in & \arg\max_{a_\tau \in \{1,2\}} E\Big[r(m_{\tau+1}, a_\tau, o_\tau) + V_{k-1}^{opt}(o_{\tau+1}, m_{\tau+1}) \mid a_\tau, o_\tau, m_\tau\Big] = \\
&= \arg\max_{a_\tau \in \{1,2\}} E\Big[[\omega \cdot (a_\tau - 1) \cdot o_\tau + (1-\omega) \cdot m_{\tau+1}(a_\tau, o_\tau, m_\tau)] + \\
&+ V_{k-1}^{opt}(o_{\tau+1}, m_{\tau+1}) \mid a_\tau, o_\tau, m_\tau\Big],
\end{aligned} \tag{5}
$$

where

$$
\begin{aligned}
V_k^{opt}(o_\tau, m_\tau) &= E\Big[r(m_{\tau+1}, a_\tau^{opt}, o_\tau) + V_{k-1}^{opt}(o_{\tau+1}, m_{\tau+1}) \mid a_\tau^{opt}, o_\tau, m_\tau\Big] = \\
&= E\Big[[\omega \cdot (a_\tau^{opt} - 1) \cdot o_\tau + (1-\omega) \cdot m_{\tau+1}(a_\tau^{opt}, o_\tau, m_\tau)] + \\
&+ V_{k-1}^{opt}(o_{\tau+1}, m_{\tau+1}) \mid a_\tau^{opt}, o_\tau, m_\tau\Big],
\end{aligned} \tag{6}
$$

$$V_0(o_{t+T}, m_{t+T}) = V_0^{opt}(o_{t+T}, m_{t+T}) = 0, k = t + T - \tau.$$

Finally, the decision policy over horizon T is built by a sequence of resulting decision rules $\forall \tau \in \{t, ..., k\}$:

$$p_\tau^{opt}(a_\tau|m_\tau, o_\tau) = \delta(a_\tau, a_\tau^{opt}(m_\tau, o_\tau)). \tag{7}$$

4 Illustrative Experiments

Several simulated experiments were designed to verify the approach proposed. The experiments supposed to analyze the game results with DM strategy influenced by the emotion component of the reward function (see Definition 2).

The following sections describe the performed experiments, initialize necessary constants and values and summarize the results obtained. The simulations were performed by using MATLAB® software [1].

4.1 Models of Proposer

As being said, the proposer's model (the first factor in (3)) is needed to compute an optimal policy (1). In order to find out the influence of the emotional component of the responder's reward function, there were considered several types of opponents in our experiments.

The designed proposers can be divided into two main groups: "open loop" proposers and "closed loop" proposers. The proposer belongs to a particular group depending on whether their next offer takes into account the previous responder's action.

Open Loop These proposers do not respect the responder's previous action, i.e. $p(o_{t+1}|a_t, o_t) = p(o_{t+1}|o_t),\ \forall t \in \mathbf{T}$, so the fixed testing sequence of the proposer's offers can be generated off-time. This significantly simplifies computations, however, a natural feedback from the responder is neglected while the next offer is generated. The offers are generated by the following algorithm:

$$p_t(o_{t+1}|a_t, o_t) \propto \begin{cases} \exp\left(-\frac{\left(o_{t+1}-(o_t+p)\right)^2}{2\sigma^2}\right), & \forall o_{t+1} \in \mathbf{O}, \text{if } o_t + p \in [b_l, b_u], \\ \exp\left(-\frac{\left(o_{t+1}-o_t\right)^2}{2\sigma^2}\right), & \forall o_{t+1} \in \mathbf{O}, \text{otherwise}, \end{cases} \tag{8}$$

where o_t is an offer in game round t and p is a random variable from set $\{-1, 1\}$. Parameters b_l and b_u are lower and upper bounds. The constant σ is a standard deviation of normal distribution.

We distinguish the three subtypes of the proposer: "greedy", "neutral" and "generous". Parameters b_l and b_u are specific for each subtype. The "greedy" proposer suggests very low and unfair offers. The "neutral" proposer's offers are more fair and offers of the "generous" proposer are very "generous" to the responder but unfair to himself/herself.

Closed Loop The more realistic proposer follows the most simple and intuitive algorithm: if the previous round is successful, the proposer tries to make more money in the next round by decreasing the offer. In case of rejection, the proposer increases the offer to make it successful.

Virtually, this proposer increases his/her next offer o_{t+1} by one if previous offer o_t has been accepted, and decreases the next offer by one otherwise:

$$p_t(o_{t+1}|a_t, o_t) \propto \begin{cases} \exp\left(-\frac{\left(o_{t+1}-(o_t+1)\right)^2}{2\sigma^2}\right), & \forall o_{t+1} \in \mathbf{O}, \text{if } a_t = 1 \\ \exp\left(-\frac{\left(o_{t+1}-(o_t-1)\right)^2}{2\sigma^2}\right), & \forall o_{t+1} \in \mathbf{O}, \text{if } a_t = 2. \end{cases} \tag{9}$$

The described model is useful for testing the designed responder. However, this model does not correspond well with human thinking because model (9) does not discriminate extremely low or high offers and does not respect fairness as humans do [2].

4.2 Experiment Setup

For each game, the number of game rounds is preset to $N = 30$ and the amount to split is $q = 10$ CZK. The proposer's offers vary from 1 to 9, $o_t \in \mathbf{O} = \{1, 2, ..., 9\}$, and the emotional states from 1 to 5, $m_t \in \mathbf{M} = \{1, 2, 3, 4, 5\}$. In relation to Section 3, it holds $o_{min} = 1$, $o_{max} = 9$, $m_{min} = 1$ and $m_{max} = 5$. The proposer is non-optimizing with a pre-defined DM algorithm, while the responder uses T-step optimization with time horizon $T = 10$ as described in Subsection 3.2. For all experiments, the personal parameters of the emotional state model (4) are set to $p_m = 3$ and $p_o = 6$. The parameter σ used in models of the proposer (see Subsection 4.1) is always set to $\sigma = 1$.

The simulations are carried out for two different types of the proposer, see Subsection 4.1. The "open loop" proposer is additionally represented by three subtypes. The "open loop" proposers differs in lower and upper bounds of the chosen offers. The specific values are: $b_l = 1$ and $b_u = 3$ for "greedy" proposer, $b_l = 3$ and $b_u = 6$ for the "neutral" proposer and $b_l = 6$ and $b_u = 9$ for the "generous" one.

Each experiment is undertaken for five different values of the weight ω, specifically $\omega \in \{0, 0.2, 0.4, 0.6, 0.8\}$, combined with different initial emotional states m_1, specifically $m_1 \in \mathbf{M} = \{1, 2, 3, 4, 5\}$. Terminal emotional state m_N, total profits of players $z_R(N)$ and success rate of the game (i.e. percentage of successful rounds) are monitored.

4.3 Results

The results for each combination of the proposer type, weight ω and initial emotional state m_1 were calculated as average of results from 100 repetitions of the given simulated game.

The experiments were carried out over the above-mentioned range of the testing emotional states, see 4.2. However, the tables below provide results only for some initial emotional states. These values give a sufficient overview of dependencies of the monitored values.

Open Loop Results The first experiments were without any feedback for the proposer. The proposer's offers were sampled from the transition probability (8). The used bounds b_l, b_u corresponded to individual types of the proposers, see 4.2. The sequences of offers generated were fixed for all performed simulations.

Greedy Proposer The "greedy" proposer makes extremely unfair offers. It could be expected that the responder would reject the majority of them. However, the responder mostly accepted. The average results, see Table 1, show "unfair" profits corresponding to the offers.

With increasing weight ω, and thus the increasing importance of the emotional component of the reward function, there also increase the success of the game and the profits of both players.

At first approach, the results seems to be surprising, as pure rational decision making without dependence on emotional state should give better economic results. However, a deeper analysis indicates that the reward of the responder is significantly increased by adding the emotional component. Moreover, the value of the emotional and the economic components of the reward are comparable in case of low offers ($o_t \leq 3$) and the emotional state component even dominates for high values of ω. Accepting everything can not produce worse results because the sequence of proposer's offers is fixed and the emotional state decreases only by the rejection.

The weight ω influences also the terminal emotional state for higher initial emotional state $m_1 > 3$ - the higher the weight, the higher the terminal emotional state. This result corresponds to the theory of expected emotions [7]. Let's recall that the personal threshold p_m for the emotional state change is 3 for all simulated games. For this reason, low initial emotional states $m_1 \leq 3$ together with low offers $o_t \leq 3$ causes the terminal emotional state of the responder is always equal to 3.

Another trend can be noticed in vertical comparison. The success rate is increasing together with increasing initial emotional state. This result is also in accordance with the observation that negative emotions are connected with a higher rejection rate of unfair offers [11]. However, the case of initial emotional state $m_1 = 1$ significantly deflects from this trend. Thus an additional analysis and experiments should be necessary.

Table 1. Game with the optimizing responder and the "greedy" proposer

Weight ω	0	0.2	0.4	0.6	0.8
Initial emotional state $\mathbf{m_1 = 1}$					
Terminal emotional state of the responder	3	3	3	3	3
Total profit of the responder (CZK)	79.32	79.51	83.10	81.92	82.84
Total profit of the proposer (CZK)	191.88	192.49	204.50	202.48	205.56
Success rate (%)	90	91	96	95	96
Initial emotional state $\mathbf{m_1 = 3}$					
Terminal emotional state of the responder	3	3	3	3	3
Total profit of the responder (CZK)	77.26	79.90	81.22	81.27	82.10
Total profit of the proposer (CZK)	185.44	195.40	199.38	200.03	203.00
Success rate (%)	88	92	94	94	95
Initial emotional state $\mathbf{m_1 = 5}$					
Terminal emotional state of the responder	4	5	5	5	5
Total profit of the responder (CZK)	78.41	82.90	85.68	86.00	86.00
Total profit of the proposer (CZK)	189.29	202.80	212.72	214.00	214.00
Success rate (%)	89	95	99	100	100

Neutral Proposer The offers of the "neutral" proposer are more fair than the offers of the "greedy" proposer. It could be expected that the success rates would be generally higher in these simulations. However, such trend is not visible in our results, see Table 2. The profits of the players correspond with the offers: they are quite balanced.

Similarly to previous results for the "greedy" proposer, increasing ω causes the improvement of all observed values. The main difference is that the terminal emotional state is equal to 5 for any initial setting. This phenomenon is due to the fair offers sufficient to maximize the responder's emotional state during 30 game rounds. Therefore the influence of the initial emotional state on the success rate is not noticeable with the "neutral" proposer. This confirms again the correspondence of the previous simulations for the "greedy" proposer [11] because the offers are fair in that case.

Table 2. Game with the optimizing responder and the "neutral" proposer

Weight ω	0	0.2	0.4	0.6	0.8
Initial emotional state $\mathbf{m_1 = 1}$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	154.68	159.96	164.32	166.81	169.66
Total profit of the proposer (CZK)	106.62	113.84	119.28	121.29	125.14
Success rate (%)	87	91	95	96	98
Initial emotional state $\mathbf{m_1 = 3}$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	159.56	161.81	165.89	167.08	170.41
Total profit of the proposer (CZK)	112.34	115.49	120.91	122.62	126.09
Success rate (%)	91	92	96	97	99
Initial emotional state $\mathbf{m_1 = 5}$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	159.67	161.91	164.29	167.05	171.15
Total profit of the proposer (CZK)	113.33	115.59	118.31	121.95	126.85
Success rate (%)	91	93	94	96	99

Generous Proposer The offers of this proposer are very high, so the game of 30 game rounds improves any initial emotional state to its maximum value 5. It is seen in Table 3. The total profits of the players correspond with the offers again. They are unfair to the proposer. The success rates are the highest from all open-loop experiments as can be expected.

The results show the clear positive correlation between the monitored quantities and the weight ω as in the previous two tests with "greedy" and "neutral" proposers. The success rate and the total economic profits of both players are directly proportional to this weight. A dependence of the monitored variables on the initial emotional state is not noticeable.

Closed Loop Results Last experiments focused on the proposer with a feedback, see Section 4.1. The offers of this proposer are stochastically dependent on his/her previous offers and the responder's decisions. The proposer respects the feedback from the responder and he/she reacts on acceptance or rejection of the offers. The transition probabilities are computed by the model (9).

The results show a significant decrease of the total responder's profit in relation with increasing weight ω, see Table 4. In the effort to maintain or improve the emotional state, the responder is willing to accept more and even worse offers. It creates a paradox that the effort to increase the emotional state ultimately leads to its reduction. This phenomenon appears in the results for the initial emotional state $m_1 \leq 4$ when comparing the terminal emotional state for $\omega = 0.6$ and $\omega = 0.8$.

Table 3. Game with the optimizing responder and the "generous" proposer

Weight ω	0	0.2	0.4	0.6	0.8
Initial emotional state $\mathbf{m_1} = 1$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	198.90	197.85	200.64	201.30	204.58
Total profit of the proposer (CZK)	90.00	89.55	92.16	92.40	94.82
Success rate (%)	96	96	98	98	100
Initial emotional state $\mathbf{m_1} = 3$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	198.76	198.97	200.38	203.45	204.23
Total profit of the proposer (CZK)	90.74	90.63	92.12	94.05	94.67
Success rate (%)	97	97	98	99	100
Initial emotional state $\mathbf{m_1} = 5$					
Terminal emotional state of the responder	5	5	5	5	5
Total profit of the responder (CZK)	196.30	199.23	199.49	203.12	204.58
Total profit of the proposer (CZK)	88.70	90.77	90.81	93.98	94.82
Success rate (%)	95	97	97	99	100

Such result seems to be surprising, but it reflects that the emotional component of the reward function becomes more significant than the economic component, especially for low offers. The responder's behavior is than influenced by the emotional state function whose value decreases with each rejection. Such condition pushes the responder to accept almost every offer. It is obvious that in the game with the closed-loop proposer the purely rational DM is better compared to the DM that is significantly influenced by the emotional state component.

The final emotional state is always at least 4. This result is probably caused by the lack of penalization of unfair offers. This conjecture will be explored in future work.

The success rate is lower for the low weights in relation to open-loop experiments. However, it increases rapidly and faster with increasing weight ω. In more detail, the offers of the proposer become very unfair with high weights $\omega \in \{0.6, 0.8\}$ and the dependence between initial emotional state and success rate is visible again. The results obtained confirm the results of research proposed by Srivastava in [11] very well: the relationship between initial emotional state and success rate exists only in the case of unfair offers.

5 Conclusions

The paper primarily proposes a reward function that accounts for influence of emotional states of the decision maker. For this reward function, the T-step-ahead optimizing policy for the responder in UG has been developed and extensively tested on simulated data. To get a relatively realistic constellation, several types of the proposer were developed and simulated.

The extensive set of experiments was run and the results were carefully analyzed. Two main groups of experiments imitated two different types of trading: i) *open-loop* when the decision maker does not influence the price and decides

Table 4. Game with the optimizing responder and the P1 proposer

Weight ω	0	0.2	0.4	0.6	0.8
Initial emotional state $\mathbf{m_1 = 1}$					
Terminal emotional state of the responder	4.48	4.47	4.56	4.81	4.39
Total profit of the responder (CZK)	118.47	112.61	107.77	96.60	70.38
Total profit of the proposer (CZK)	36.03	42.79	55.03	100.10	175.32
Success rate (%)	52	52	54	66	82
Initial emotional state $\mathbf{m_1 = 3}$					
Terminal emotional state of the responder	4.24	4.42	4.57	4.83	4.12
Total profit of the responder (CZK)	114.37	111.89	105.85	90.50	63.61
Total profit of the proposer (CZK)	36.93	40.71	54.85	110.70	201.29
Success rate (%)	50	51	54	67	88
Initial emotional state $\mathbf{m_1 = 5}$					
Terminal emotional state of the responder	4	4	5	5	5
Total profit of the responder (CZK)	114.87	112.46	108.06	82.99	49.71
Total profit of the proposer (CZK)	36.73	43.04	57.94	142.31	250.29
Success rate (%)	51	52	55	75	100

only on buying/not buying (this models, for instance, the case of trading with future) and ii) *closed-loop* when the decision maker influences the price indirectly (this models, for instance, trading with goods and services).

The main findings of the work are as follows. Generally, affective DM in game with those "smarter" proposers worsens the economic results. In some cases, decrease of the monetary profit even worsened the terminal emotional state. The success rates as well as the terminal emotional states in simulations were mostly higher. This is partially caused by the proposed reward function, in particular by a fixed weight of the value of possible monetary profit and of the emotional profit within whole game. This special phenomenon is clearly visible on games with the "greedy" proposer whose offers were very unfair to the responder. The value of emotional component in the reward function should be more penalized or the relative values of monetary profits should be used to avoid this effect.

The proposed model of the emotional state development is very intuitive and reasonably corresponds with verified psychological recommendations. Its main but removable drawback is that it does not count for unfair offers. The remedy can be reached by considering a kind of un-fairness penalization. Also, the considered 5 levels of the emotional state should be refined to 12 levels in accordance with Woodruffe-Peacock [13].

The proposed reward function includes the parameter ω balancing the monetary profit and emotional state. Another foreseen improvement should define a parameter balancing the influence of the game course on the emotional state change. This could distinguish a labile player whose emotional state can be hardly changed.

Inevitable direction of further research is creating a learning adaptive responder. Currently, the responder knows the model of the proposer. This can be interpreted as that the responder already knows usual behavior of the co-player.

However, this is not a realistic assumption and the learning ability should be added to the model of the responder. It can surely be done by employing Bayesian learning suitable for a very short learning periods

Even so is our fixed simple model good enough to describe some basic tendencies of affective DM reported in the literature. In both cases of type of the trading, the affected DM significantly changed success rate and our model confirmed that.

Acknowledgement: We would like to thank Eliška Zugarová for comments and suggestions that greatly influenced the manuscript.

References

1. MATLAB® version 7.5.0 (R2007b). The MathWorks Inc., Natick, Massachusetts, USA (2010)
2. Binmore, K.G.: Game Theory and the Social Contract: Just Playing, vol. 2. The MIT Press, Cambridge, Massachusetts (1998)
3. Cox, J.C., Friedman, D., Gjerstad, S.: A tractale model of reciprocity and fairness. Games and Economic Behavior **59**(1), 17–45 (April 2007). https://doi.org/10.1016/j.geb.2006.05.001
4. Fehr, E., Schmidt, K.M.: A theory of fairness, competition, and cooperation. The Quarterly Journal of Economics **114**(3), 817–868 (August 1999). https://doi.org/10.1162/003355399556151
5. Guth, W., Schmittberger, R., Schwarze, B.: An experimental analysis of ultimatum bargaining. Journal of Economic Behavior & Organization **3**(4), 367–388 (December 1982)
6. Haselhuhn, M.P., Mellers, B.A.: Emotions and cooperation in economic games. Elsevier **23**(1), 24–33 (April 2005). https://doi.org/10.1016/j.cogbrainres.2005.01.005
7. Loewenstein, G., Lerner, J.S.: Handbook of Affective Sciences, chap. 31. The Role of Affect in Decosion Making, pp. 619–642. Oxford University Press (2003)
8. McGraw, A.P., Larsen, J.T., Kahneman, D., Schkade, D.: Comparing gains and losses. Psychological Science **21**(10), 1438–1445 (2010). https://doi.org/10.1177/0956797610381504, `https://doi.org/10.1177/0956797610381504`, pMID: 20739673
9. Puterman, M.L.: Markov Decision Processes. John Wiley & Sons, Inc. (1994)
10. Rabin, M.: Incorporating fairness into game theory and economics. The American Economic Review **83**(5), 1281–1302 (December 1993)
11. Srivastava, J., Espinoza, F., Fedorikhin, A.: Coupling and decoupling of unfairness and anger in ultimatum bargaining. Journal of Behavioral Decision Making **22**, 475–489 (2008). https://doi.org/10.1002/bdm.631
12. Tamarit, I., Sanchez, A.: Emotions and strategic behavior: The case of the ultimatum game. Plos One **11**(7) (July 2016). https://doi.org/10.1371/journal.pone.0158733
13. Woodruffe-Peacock, C., Turnbull, G.M., Johnson, M.A., Elahi, N., Preston, G.C.: The quick mood scale: Development of a simple mood assessment scale for clinical pharmacology studies. Human Psychopharmatology Clinical and Experimental **13**(1), 53–58 (January 1998)