Introducing Affect into Competitive Game Play

Maha Alabduljalil, Songhua Xu and Willard Miranker

Department of Computer Science, Yale University, 51 Prospect Street, New Haven, CT, USA, 06520

maha_abduljalil@hotmail.com, songhua.xu@yale.edu, miranker@cs.yale.edu

Abstract

A dynamic notion of affect (degree of satisfaction) that an agent acquires in competitive iterated game play is developed. Simulated play against both a random environment and a competitor agent is used to study the impact of affect on game playing strategies. For definiteness, the formulation is framed in terms of stock trading. Comments on how affect in game play informs a notion of consciousness along with simulations are given.

Index terms–affect, anti-tit-for-tat, consciousness, stock trading, tit-for-two-tat

1. Introduction

A notion of affect is introduced into game playing strategies [1]. Readers are referred to [1] for general introduction to game theory. For definiteness and clarity we frame our study in terms of stock trading, so that affect may be thought of as an experience of feeling or emotion of the investor in response to his performance. Affect will play a key part in an investor's interaction with stimuli. In a direct competition between a pair of players, affect also refers to affect display, such as facial, vocal, or gestural behavior that serves as an indicator of the investor's feelings. A positive value of affect is informally characterized as the satisfaction an agent feels with the result of an action he has taken. This corresponds to a positive value of affect. Reversely, dissatisfaction is taken as a negative value of affect. For clarity we restrict affect, denoted $a(x)$ to represent these two situations, satisfaction or dissatisfaction, and so, we specify it as follows.

$$a(x) = \operatorname{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0. \end{cases} \tag{1}$$

Here x is the difference between the predicted payoff of a round of play and the actual payoff.

We consider a basic case and an advanced case and study both by simulation. The basic case involves an investor playing against an environment represented by random market behavior. If the simulation is able to produce success against a random market, playing against market behavior with a degree of predictability is likely to make success easier to achieve. The advanced case will involve two competing investors (players), each playing a strategy of his choice.

Strategy superiority is based on both a notion of accumulated affect (satisfaction with his play) of the investor and his total return. We shall see that one particular strategy of play majorizes all other strategies considered.

For clarity in describing the model, only binary moves are considered. For instance, in the basic case the rising or falling of a stock price is represented the stimulus variable s taking on the value ±1, respectively. We let r = ±1 represents an investor's response, namely his order to buy or sell a stock. The advanced case is similar; each investor either buys or sells, and one investor's action provides the stimulus for the other. In both cases, the response is a function, denoted ρ of affect, previous response, stimuli and round of play (all indexed by the discrete time variable j), namely

$$r_j = \operatorname{sgn}\ \rho(A_j, R_j, S_{j+1}, j)\,. \tag{2}$$

Here $R_j = r_0, \dots, r_{j-1}$, $A_j = a_1, \dots, a_j$ and $S_j = s_0, \dots, s_{j-1}$. $\forall i_i$, s_j is the stimulus from the environment or the competing investor as the case may be. The shift in the indices relevant to R_j and A_j reflects the timing: first a stimulus, then a response.

We shall make a number of choices for ρ. To evaluate the framework, we include some well-known strategies. These are tit-for-two-tat and anti-tit-for-tat. Some novel strategies are also introduced.

[1] This work was motivated by a set of unpublished notes on modeling affect by W. Miranker and G. Zuckerman

1082-3409/08 $25.00 © 2008 IEEE
DOI 10.1109/ICTAI.2008.63

2. STRATEGIES AND GAME PLAY

2.1 Basic Case: Player vs. Environment

Here an investor (player) employs one of several different strategies versus an environment characterized by random changes in stock price. Simple inputs are used for the specification of the response. These are the number of times a favorable buy was omitted, the number of times a favorable sale was omitted and the overall satisfaction of the investor. An omitted favorable buy is specified by the event $\{s=-1 \wedge r=-1\}$, while an omitted favorable sale corresponds to the event $\{s=1 \wedge r=1\}$.

There are four strategies illustrated. The first is the tit-for-two-tat strategy (T2T), which we take to represent a naïve or a forgiving investor. This investor keeps on either buying or selling for consecutive rounds of play. He will switch his action (buy/sell) only if he is dissatisfied for two consecutive rounds of play. The response strategy in this case is described as follows.

$$\text{T2T:} \quad r_j = \operatorname{sgn}(a_{j-1} + a_{j-2}) r_{j-1}, \quad \operatorname{sgn} 0 = 1. \tag{3}$$

Here $a_j, j > 0$ is the affect at play j (see (1)).

The second strategy is the anti-tit-for-tat strategy (ATT). This is very risky behavior, since the investor keeps alternating his response, disregarding his experience. So, he buys then sells then buys ... ATT is described as follows.

$$\text{ATT:} \; r_J = -r_{j-1}. \tag{4}$$

Third, we specify a new strategy called Developed Strategy (DS) given by

$$\text{DS:} \quad r_j = \begin{cases} r_{j-1} & \alpha_{j,N} < 0 \\ \\ -r_{j-1} & \alpha_{j,N} \geq 0. \end{cases} \tag{5}$$

Here $\alpha_{j,N}$ is the accumulated affect over some specified number N of prior plays. In particular,

$$\alpha_{j,N} = \sum_{i=j-N+1}^{j} a_j \tag{6}$$

The fourth strategy is called the DS-Factors (DSF) strategy. Suppose there are to be a trial of N rounds of play. For the first N/2 rounds of play, the investor will follow T2T (suggesting that he is new to the market place). Then, for second half of the trial his response depends on the affect (the satisfaction accumulating in his mind, so-to speak). So, the investor learns from his past performance. In particular,

$$\text{DSF:} \; r_j = \begin{cases} \operatorname{sgn}(a_{j-1} + a_{j-2}) r_{j-1}, & j \leq N/2 \\ r_{j-1}, & j \geq N/2 \wedge \alpha_{j,N/2} \geq 0 \\ -r_{j-1}, & j \geq N/2 \wedge \alpha_{j,N/2} \leq 0. \end{cases} \tag{7}$$

To account for losing trades, we replace the last two lines in (7) by

$$r_j = \begin{cases} +1 & P_j \geq \theta \wedge L_j < \theta, j \geq N/2 \\ -1 & L_j \geq \theta \wedge P_j < \theta, j \geq N/2. \end{cases} \tag{8}$$

where θ is an arbitrarily specified parameter. Here

$$P_j = \sum_{i=1}^{j} k \quad \begin{cases} k = 1 & \mathit{if}\ \{s_i = -1 \wedge r_i = -1\} \\ k = 0 & \mathit{otherwise}\ . \end{cases} \tag{9}$$

$$L_j = \sum_{i=1}^{j} m_i \quad \begin{cases} m_i = 1 & \mathit{if}\ \{s_i = 1 \wedge r_i = 1\} \\ m_i = 0 & \mathit{otherwise}\ . \end{cases} \tag{10}$$

We see that the factors (parameters) P_j and L_j express the total number of favorable trades, buy or sell, respectively, that were passed up. Other choices of factors are possible. For instance, they might depend on prices being low in summer and high in spring.

The overall satisfaction and actual gain or loss is based on a predefined payoff matrix. This matrix represents the reward for each of the possible pairs of response and stimulus.

2.2 Advanced Case: Two Players

The advanced case is a two-investor generalization of the basic case (a player versus the environment). For clarity, we require that the players employ differing strategies.

We introduce a new strategy somewhat similar to the DSF, called DS-function (DSFn). It is so named because instead of using a signum function to calculate a player's response, we use a function of the form usin (accumulated affect) + vP + wL. Then the response is equal to the function specified as follows.

$$\text{DSFn:} \quad r_{jk} = -\operatorname{sgn}(u \sin \alpha_{jk} + vP_{jk} + wL_{jk})) \tag{11}$$

where k = 1,2 indexes the players and u, v and w are parameters to be specified. The P_{jk} and the L_{jk} are two player versions of the P_j and the L_j in (9) and (10), resp.

2.3 Simulations

2.3.1 Basic Case: player vs. (a random) environment

We employ the payoff matrix shown in Table 1.

Interpretation of Table 1: If the investor places an open order to buy and the price rises, we would say he is involved in a losing situation. Similarly, for an open order to sell and the price declines. Such a situation is captured in the payoff matrix by a negative entry.

Table 1. The payoff matrix for the basic case simulations.

player \ stimulus	1	-1
1	-2	15
-1	15	-10

In Figure 1 we plot the accumulated affect (specified in (6)) for 25 different trials each of N =100 rounds of play for the four different strategies, A2T, ATT, DS and DSF.

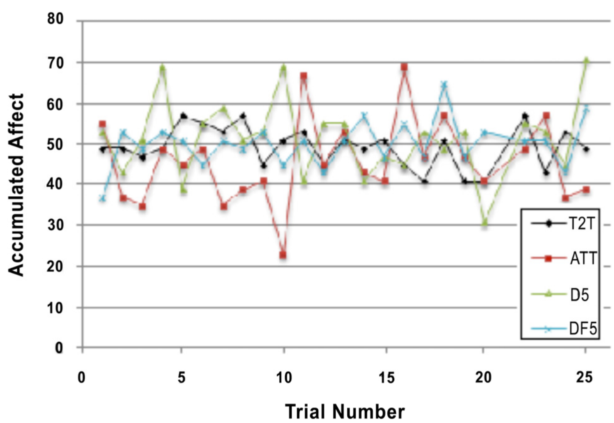

Figure 1. Accumulated affect for trials of 100 rounds of play.

From Figure 1, we see the ATT strategy variously yields the greatest satisfaction (maximum value of accumulated affect) or the least satisfaction (minimum value of accumulated affect). So its risky performance suggests that it can't be a good strategy to use in the basic case.

Some statistics for the accumulated affect, denoted α delivered by the previous simulation are shown in Fig. 2. The minimum value of α is 23. Deviations of the maximum values on the other hand, are not significant. The greater averages of accumulated affect were delivered by the developed strategies, DS and DSF.

In Figure 2 we plot the accumulated affect of 20 trials each with an increased number, namely N = 500 rounds of play.

Table 2. Simulation statistics for the accumulated affect α.

Strategy \ α	Min	Max	Avg.
Tit-for-two-tat	41	57	49.25
Anti-tit-for-tat	23	69	45.8
DS	31	69	51.5
DSF	37	65	50.25

The ATT strategy again produces the minimum and maximum values of accumulated affect (risky performance). The DSF strategy is stable since it maintains the accumulated affect in a high range with minimal variation. To quantify these behaviors, we display the averaged variance over all trials of the accumulated affect of the strategies in Table 3.

Table 3: Average variance of the accumulated affect for 20 trials each of 500 rounds of play.

T2T	**5147**
ATT	**5581**
DSF	**5539**

Let $p(r_{j-1}, s_j)$ denote the payoff (the appropriate payoff matrix entry) at round j. Then the accumulated payoff is defined as

$$\Pi_j = \sum_{i=1}^{j} p(r_{j-1}, s_j)\,. \qquad (12)$$

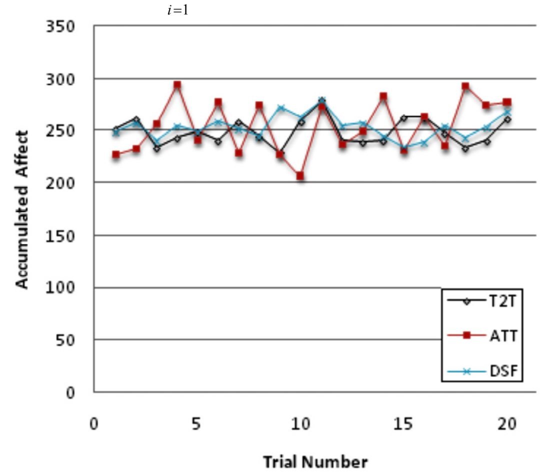

Figure 2. Accumulated affect for trials of 500 rounds of play.

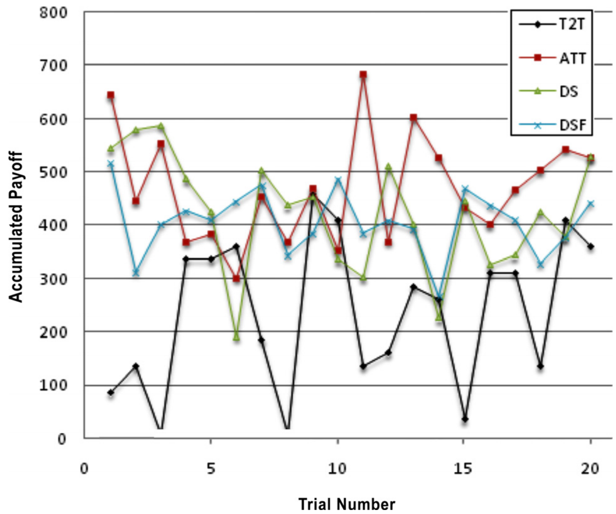

Figure 3. Accumulated payoff for trials of 100 rounds of play.

In Figure 3 we plot the accumulated payoffs for 20 trials each with 100 rounds of play. Figure 3 induces us to discard the T2T strategy for the remainder of the basic case simulation, because it displays the worst financial performance. The DSF strategy provided consistency with its high total return.

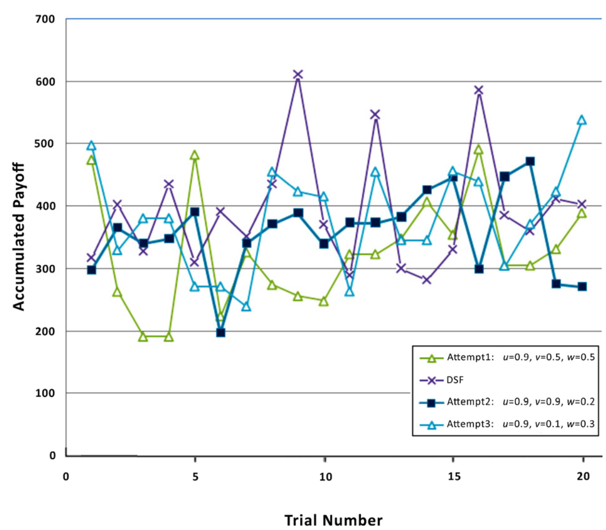

Figure 4. Accumulated payoff for trials of 500 rounds of play.

Figure 4 displays results comparing DSF and DSFn, the latter for different values u, v and w. These values, taken in the range [-0.5, 2] were specified by rule of thumb calculations.

Figure 5 is a refinement of Figure 4 obtained by increasing the trial size from 100 to 500. The results are consistent with those represented in Figure 4, suggesting the adequacy of the simulation.

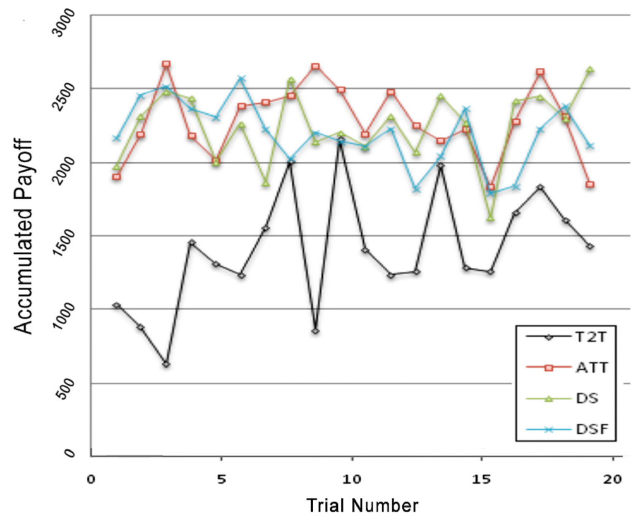

Figure 5. Accumulated payoff for trials of 500 rounds of play.

The choice of θ for the DSF strategy is also based on numerical experiments. The value $\theta = 0.4$ gave the greatest values of accumulated affect. Figure 6 illustrates this point.

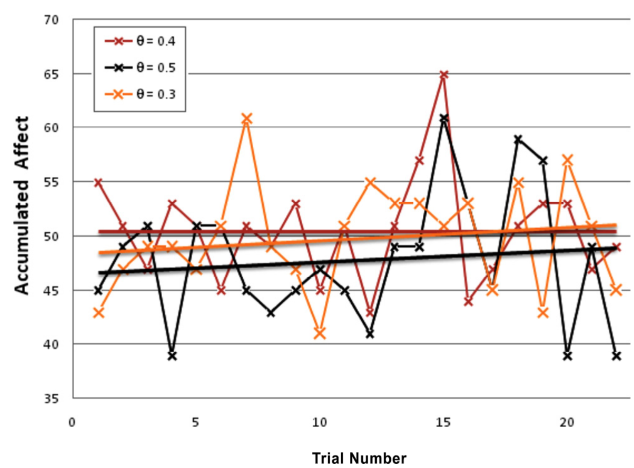

Figure 6. Accumulated affect for different values of *θ*. The linear plots are least squares fits.

2.3.2 Advanced Case: Two-player game

Table 4 displays the payoff matrix. A common payoff matrix ensures that the competition between players is fair.

The accumulated affect and the accumulated payoff resulting from simulations of play are displayed in Tables

5, 6 and 7. In each of these simulations there are 20 trials and N= 500 for all of them. The following Tables 5, 6 and 7 show clearly why we find the DSFn strategy to be superior in a two-player game.

Table 4: The payoff matrix for the advanced case.

player \ opponent	1	-1
1	-5	15
-1	20	-10

Table 5: Results of competition between T2T and ATT.

Strategies	Accumulated Payoff	Accumulated Affect
T2T	**2480**	499
ATT	**1260**	-1

Table 5 shows that T2T is superior to ATT, because it produced the larger accumulated payoff as well as the larger accumulated affect.

Table 6: Results of competition between T2T and DSF.

Strategies	Accumulated Payoff	Accumulated Affect
T2T	**2450**	499
DSF	**1235**	497

From Table 6 we see that T2T is also superior to DSF. This motivated use of DSFn in a competition with T2T. The results are shown in Table 7.

Table 7: Results of competition between T2T and DSFn.

Strategies	Accumulated Payoff	Accumulated Affect
T2T	**3510**	179
DSFn	**5180**	499

Finally the results of competition between DSFn and DSF are shown in a Table 8.

Table 8: Results of competition between DSF and DSFn.

Strategies	Accumulated Payoff	Accumulated Affect
DSF	**7385**	497
DSFn	**9830**	499

Since there are 2 players and 4 strategies, each strategy is employed 6 times by a player against his opponent. We average the accumulated payoff and accumulated affect over each such collection of 6 competitions. These results are shown in Figures 7 and 8, respectively.

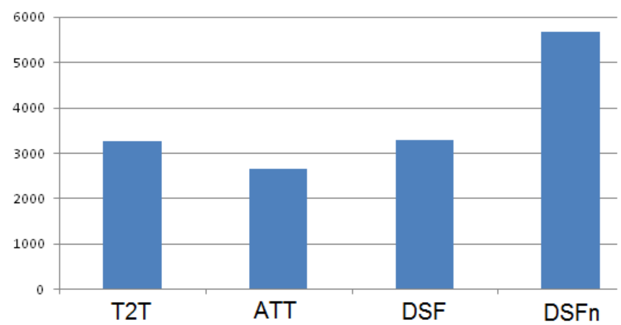

Figure 7. Accumulated payoff averaged over 6 competitions.

The investor adopting DSFn had the best performance but not the greatest satisfaction. The reduction of satisfaction of DSFn is mainly due to which player had the first play. Reversely, T2T had the greatest satisfaction, but not the greatest payoff.

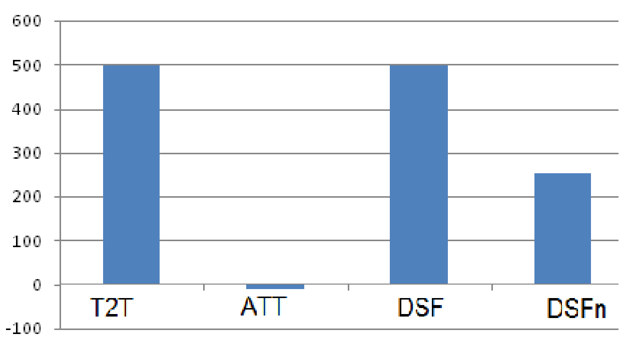

Figure 8. Accumulated affect averaged over 6 competitions.

4. References

[1] Fudenberg D., Tirole J. *Game theory*, MIT Press, ISBN 978-0-262-06141-4, 1991.