Incorporating Emotions into Automated Negotiation

Hong Jiang, José M. Vidal, Michael N. Huhns
Computer Science and Engineering
University of South Carolina
Columbia, SC 29208, USA
jiangh@engr.sc.edu, vidal@sc.edu, huhns@sc.edu

Abstract

We introduce an emotional agent model that shows how emotions affect an agent's negotiation strategy. By adding emotions, we add the effects of these indirectly related features to the negotiation, features that are ignored in most models. Our new method, the PAD Emotional Negotiation Model, maps a nonemotional agent's strategy during negotiation to the strategy used by an emotional agent. Our evaluations show this model can be used to implement agents with various emotional states that mimic human emotions during negotiation.

1 Introduction

There are many studies of emotion in the psychological literature [Sousa, 2003]. The work of [Wright *et al.*, 1996] and [Picard, 2000] placed emotion into computational theory and has led to increasing interest in computational models of emotion. [Ekman and Davidson, 1994] reveals the central issues in emotion research and theory in the words of many of the leading scientists working in the field today. [Richard J. Davidson and Goldsmith, 2002] gives a comprehensive road-map to the burgeoning area of affective sciences, and brings together the various strands of inquiry and the latest research in the scientific study of the relationship between the mechanisms of the brain and the psychology of mind. However, we are not aware of any research that has tried to incorporate emotional models into the problem of negotiation.

Perfectly rational agents are only affected in their negotiation by features of the problem that directly impact their utility of the resulting deal. However, humans are not perfectly rational and often let their emotions, even those that are unrelated to the negotiation problem, affect their negotiation strategy.

There are many difficulties in incorporating emotional models into automated negotiation. We must determine how to measure emotions and then convert them into negotiation actions. The PAD [Mehrabian, 1995] model is an established method for modeling emotions. It uses three dimensions: Pleasure, Arousal, and Dominance. We show how to combine these dimensions to correctly reflect the effects of emotions in negotiation and use them to implement emotionally enhanced automated negotiating agents.

Section 2 starts by presenting the Emotional Worth-Oriented Domain (EWOD) over which the agents will be negotiating. Section 3 shows how we incorporate the PAD model into a negotiating agent. The evaluation of this model is given in section 4, which shows that this model reflects human experience and negotiation theory. Finally, a conclusion and ideas for future work are given in section 5.

2 Emotional Worth-Oriented Domain (EWOD)

We assume that our agents have utility functions that capture their preferences over possible deals. As such, agents can value the same item differently, which can lead to negotiation. For example, agent A offers agent B a watch with price \$10; B may think it's too expensive and that \$4 would be more reasonable, so then negotiates with A for \$4; A thinks \$4 is not acceptable and asks for \$8, and so on. Nonemotional agents are typically assumed to have fixed utility functions. However, a human's utility valuation can change due to their emotional state, and an agent's should as well.

Before we take into account emotions, let's reconsider emotional and rational behavior. As described in figure 1, a behavior is rational generally because there is some relation between the reason and the behavior. If there is no direct relation between the reasons and the behavior, we think the behavior is not rational, so we might describe it as an emotional behavior. Emotional behaviors are different from rational ones, but they are not in complete conflict. By adding emotions between the behavior and the unrelated reason, many things become easy to explain. For example, suppose an agent A gets a gift from a friend B today that makes him very happy. When people are in a happy mood they are more willing to help others. C asks A for help. Usually A rejects C, but today A gives C the help that C requests. There is no direct relation between the fact that A gets a gift from B and that A helps C, but by adding emotions we can explain it. Usually the effect of emotions is rational, in other words, there is often a reason why people are happy

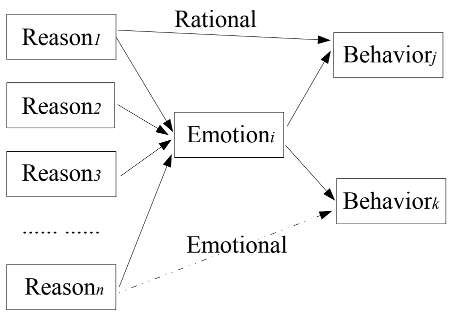

Figure 1: Rational and Emotional Behaviors Description

or sad. On the other hand, from emotion to behavior, there are also some rules to follow.

Emotions do have some effect on people's behavior. However, these effects are usually ignored in automated negotiation protocols. Correspondingly, some features that do not seem to be directly related are also ignored. By adding emotions we can better model the outcome of real human negotiations.

2.1 EWOD

Based on the above assumptions, let's consider the negotiation problem. [Rosenschein and Zlotkin, 1994] present several problem domains, including the task-oriented domain and the worth-oriented domain. In a task-oriented domain, the tasks are explicitly defined in the encounter: each agent is given a set of tasks to accomplish and associated with each there is a cost. An agent attempts to minimize the overall cost of accomplishing these tasks. The worth-oriented domain is a more general domain: the goals of an agent are specified by defining a worth function for the possible states of the environment, and the goal of the agent is thus implicitly to bring about the state of the environment with the greatest value. Unlike task-oriented domains, as mentioned in [Wooldridge, 2001], agents negotiating over worth-oriented domains are not negotiating over a single issue: they are negotiating over both the state that they wish to bring about and over the means by which they will reach this state. A task-oriented domain is a special case of a worth-oriented domain.

Without losing generality, we focus on the model of worth-oriented domains (WOD). We modify the model to include emotions and call it the Emotional Worth-oriented Domain (EWOD). Formally, an EWOD is a tuple

$$< E, Ag, J, c, r_e >$$

where

- E is the set of possible environment states;
- $Ag = \{1, ..., n\}$ is the set of possible agents;
- J is the set of possible joint plans, which are *joint* because executing one plan can require several different agents. A joint plan can be represented as $j : e_1 \to e_2$, which means that the plan j can be executed in state e_1, and when executed in this state, will lead to state e_2. If the plans are not joint, but can be done by one agent, then the problem simplifies to a task-oriented domain, and J will be the set of task assignments.
- $c : J \times Ag \to \Re$ is a cost function, which assigns to every plan $j \in J$ and every agent $i \in Ag$ a real number that represents the cost $c(j, i)$ to i of executing the plan j.
- r_e could be a function with time, or a constant. It represents the emotional degree of agent i in the range from 0 to 1. For example, for a completely rational agent, $r_e = 0$; for a completely emotional agent, $r_e = 1$.

An encounter in this model is a tuple:

$$< e, W, W_e >$$

where

- $e \in E$ is the initial state of the environment;
- $W : E \times Ag \to \Re$ is a worth function, which assigns to each environment state $e \in E$ and each agent $i \in Ag$ a real number $W(e, i)$ that represents the value, or worth, to agent i of state e.
- $W_e : S_e \times E \times Ag \to \Re$ is a emotional worth function, which gives the worth affection of current emotional status, represented by an emotional state function S_e, to each environment state $e \in E$ and each agent $i \in Ag$. It is a real number also.

Reaching agreement involves the agents negotiating over the collection of joint plans. Agents try to reach agreement on the plan that brings about the environment state with the greatest worth. The optimal plan j^i_{opt} will then satisfy the following equation:

$$j^i_{opt} = \arg \max_{j:e_0 \to e \in J} r_e \cdot W_e(S_e, e, i) + W(e, i) - c(j, i)$$

The equation involves three parts: emotional worth, rational worth, and cost. We try to find the plan that maximizes their sum.

3 PAD Emotional Negotiation Model

The PAD emotional state model is a general but precise three-dimensional approach to measuring emotions. [Mehrabian, 1995] reviews versions of the PAD scales with different dimensions, and lists sets of studies that report development and refinement of a final set of the scales and consistently yield three nearly orthogonal dimensions: Pleasure, Arousal, and Dominance. The analysis shows that these three dimensions provide a parsimonious base for the general assessment of emotional states. These scales have wide-ranging applications [Mehrabian, a]. They are used to assess consumer reactions to products, services, and shopping environments. Additionally the scales can be used to assess the emotional impact of a workplace, an advertisement, or a medical or psychotropic drug. Recently, there has been some effort to incorporate PAD in Artificial Intelligence, [Mehrabian, b] but the research is in its infancy.

PAD uses three basic dimensions of emotion: Pleasure–Displeasure (P), Arousal–Non-arousal (A) or general level of physical activity and mental alertness, and Dominance–Submissiveness (D) or feelings of control vs. lack of control over one's activities and surroundings. Our model tries to relate these measurements into automated negotiation.

First of all, we need to find out the relationships of the three dimensions with human behavior, and map them to the negotiation. We analyze the details of PAD, and the relationships with human behavior and negotiation as follows.

- P: Pleasure–Displeasure. This gives the direction of emotions, positive emotion status / negative emotion status. Generally, for humans, a positive emotional state is more conducive to a person acting in a friendly and sociable manner with others; conversely, a negative emotional state tends to heighten chances that the individual will be unfriendly, inconsiderate, or even rude to others. During negotiation, a more pleasant agent tends to cooperate with others or tends to accept others' offers; on the contrary, a more unpleasant agent tends to reject others' offers. We can reflect this relationship to the value system by assuming that pleasure makes the agent increase the evaluation value and displeasure makes the agent decrease the value.
- A: Arousal–Non-arousal. This gives the degree of effects on the above intentions as given by P. Arousal means to rouse or stimulate to action or to physiological readiness for activity. We can reflect this to the value system of negotiation by assuming that this measure magnifies or minimizes P's affection. For example, if an agent is in pleasure status this emotion makes the agent increase the evaluation value a little; if the agent is also on arousal, it increases even more. But, if the agent is in displeasure, then arousal will make the agent decrease the value more.
- D: Dominance–Submissiveness. This estimates the degree of the ability of being commanding, controlling, or prevailing over all others, or degree to yield oneself to the authority or will of another. This description is close to the idea of power in Network Exchange Theory (NET) [Willer, 1999]. The agent in a dominant state or with more power tends to persist in its own proposal and benefit more in negotiation. However, the D value in PAD is decided by emotional status, which is subjective; the power in NET is objectively decided by the network structure. We can relate this measurement to the value system of negotiation by assuming that since a dominant agent tends to persist in its own proposal it will tend to decrease the evaluation value. On the other hand, if the agent is submissive, it will tend to yield and accept the other agent's proposal.

By analyzing and combining all the above relationships together, we define the following emotional state function:

$$S_e(r_p, r_a, r_d) = r_p \cdot (1 + r_a) - r_d$$

where $r_p, r_a, r_d \in (-1, 1)$ are a measurement of the three dimensions of the PAD model. They define an emotional status. For example, anger is defined as $\{-0.51, 0.59, 0.25\}$, which means $r_p = -0.51$, $r_a = 0.59$, $r_d = 0.25$ [Mehrabian, b], fear is $\{-0.64, 0.60, -0.43\}$.

The emotional worth W_e is then defined as

$$W_e(S_e, e, i) = S_e(r_p, r_a, r_d) \cdot W(e, i).$$

We can also define the effects of emotion on the rational evaluation to be given by F where

$$F = r_e \cdot S_e(r_p, r_a, r_d).$$

F tells us how much the rational evaluation will increase or decrease due to the emotional state. For example, if $F = 0.1$ and the rational worth function is given by W then the emotional state makes the worth increase by $0.1 \cdot W$.

The agent's decision is thus based on

$$\begin{aligned} j^i_{opt} = \arg \max_{j: e_0 \to e \in J} & r_e \cdot (r_p \cdot (1 + r_a) - r_d) \\ & \cdot W(e, i) + W(e, i) - c(j, i) \end{aligned}$$

So far, we have presented a detailed equation for the optimal plan. The negotiation protocol we assume has the agent offering a new proposal or accepting the other's proposal at each time step. If the proposal is accepted then negotiation ends. The difference between the offers at successive time steps at time τ is called d_τ. Different agents may use different strategies for proposing their next offer. As such, d_τ could be a constant, as in the monotonic concession protocol [Rosenschein and Zlotkin, 1994], or it could change. We let $S(\tau)$ be the agent's strategy function. For example, an agent using the monotonic concession protocol would have $S(\tau) = d$.

Given a rational agent with a strategy function $S(\tau)$, we can convert it to an emotional one by mapping the original strategy function to a new one $S'(\tau)$, as such:

$$\begin{aligned} S'(\tau) &= M(S(\tau)) \\ &= r_e \cdot S_e(r_p, r_a, r_d) \cdot S(\tau) + S(\tau) \end{aligned}$$

where $M(\cdot)$ is our emotional mapping function that maps rational strategies to emotional strategies.

4 Evaluation

We consider a situation where two agents need to reach an agreement on a given issue. We assume the agents use a protocol based on the alternating offers negotiation model [Rubinstein, 1982].

We consider cases with only two agents: A and B. It is assumed that the agents can take actions

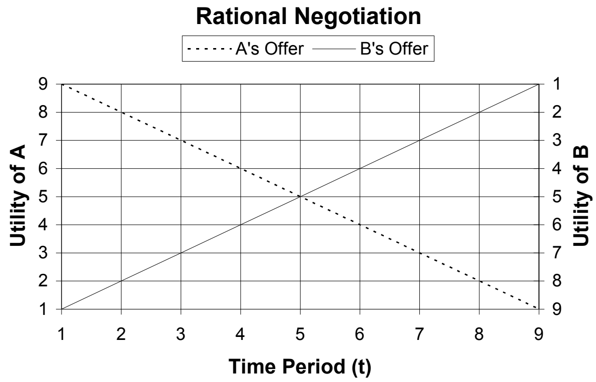

Figure 2: Rational Negotiation Process ($r_e = 0$)

in the negotiation only at certain times in the set $T = \{0, 1, 2...\}$ that are determined in advance and are known to the agents. In each period $\tau \in T$ of the negotiation, if the negotiation has not terminated earlier, the agent whose turn it is to make an offer at time τ will suggest a possible deal, and the other agent may either

1. accept the most recent offer or proposal,
2. reject it,
3. opt out of the negotiation.

If an offer is accepted by both agents, then the negotiation ends, and the offer is implemented. If at least one of the agents opts out of the negotiation, then the negotiation ends and a conflict outcome results. If no agent has opted out and at least one agent rejects the offer, then the negotiation proceeds to period $\tau + 1$ where a new offer is made.

In theory, both agents can keep rejecting offers so that an agreement may never be reached (in that case we talk about disagreement or a conflict deal). However, an agent's utility depends on the value at which an agreement is reached as well as on the time at which it is reached, hence disagreement is the worst possible outcome for both players. Our model makes the following further assumptions:

- *Agreement is preferred*: agents prefer any deal at least as much as disagreement.
- *Agents seek to maximize utility*: agents prefer to get larger utility values.
- *Agents have a reservation price*: if the utility is below the reservation price, an agent would rather not reach agreement.

Now, let's consider the following specific scenario: two agents A and B want to split \$10. To see the property of the r_p, r_a, r_d clearly, we let the strategy function of both agents be the simplest one, a constant ϵ, as in the monotonic concession protocol:

$$S(\tau) = \epsilon$$

Assuming $\epsilon = \$1$ for each time round, then rational agents with $r_e = 0$ will have their rational negotiation

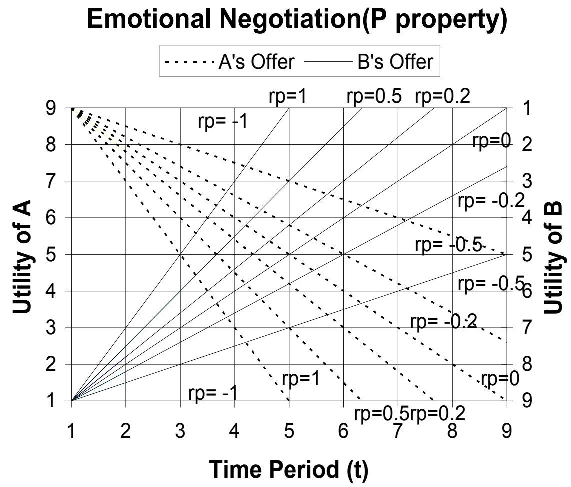

Figure 3: Emotional Negotiation Process ($r_e = 1, r_a = 0, r_d = 0$)

process described as in figure 2. A's strategy is described as the dotted line. A starts by giving its best offer: "A gets \$9 and B gets \$1". At each time step of the negotiation A will propose a new offer along this dotted linear line until it reaches its reservation price or the negotiation ends. B's strategy is similar and is described by the solid line. All the possible deals are represented as points on the y-axis. For example, A gets \$9 and B gets \$1, or A gets \$8 and B gets \$2, etc. The x-axis shows the time rounds. The cross point shows when and what deal agents A and B will agree on. In this case, they end with A getting \$5 and B getting \$5.

We now show how to add emotions to these agents. We let the agents be completely emotional, that is, $r_e = 1$, and vary their emotional dimensions r_p, r_a, r_d separately. Figure 3 shows an example where we set $r_e = 1, r_a = 0, r_d = 0$, which lets emotion features A and D be neutral and lets r_p vary from -1 to 1. From this figure, we can see that:

- A more pleasant agent ends up with a deal more quickly, and a more unpleasant agent ends up with a deal more slowly;
- Agents of the same type with the same pleasure status end up with the same benefit. That is, they will reach the deal where A gets \$5 and B gets \$5;
- If two agents of the same type but with different pleasure status engage in negotiation, then they will reach an agreement that is more favorable to the more unpleasant agent.

We can see that the above results fit our intuition as well as our theory. A pleasant person easily accepts any offer, which means he might not benefit as much. However, if there are many different negotiations, then we can imagine that a pleasant person will end up with a lot of different deals. Our model is similar. Since a more pleasant agent ends up with a deal more quickly,

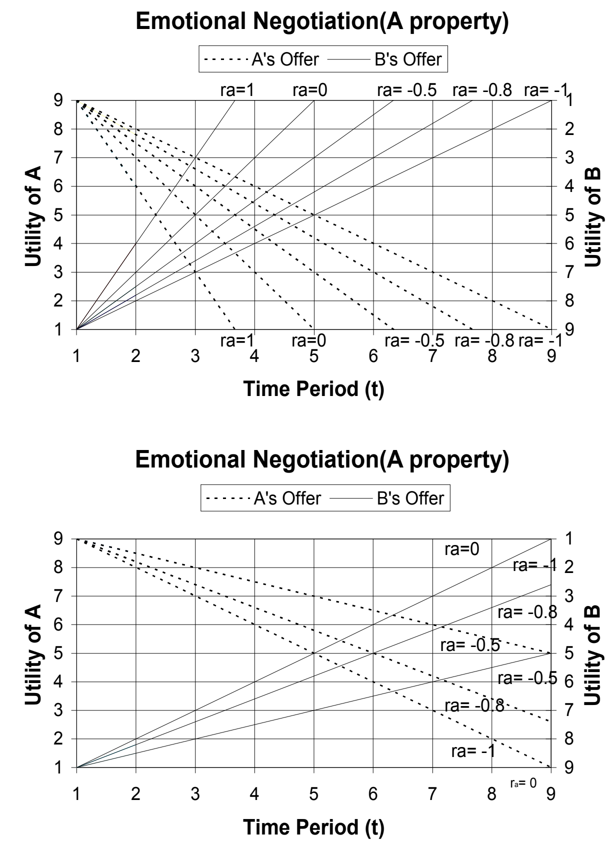

Figure 4: Emotional Negotiation Process (Top: $r_e = 1, r_p = 1, r_d = 0$; Bottom: $r_e = 1, r_p = -1, r_d = 0$)

it has time for other possible trades. Thus, its short-term loss might translate into a long-term gain.

Let's now consider emotion feature A's property by setting $r_e = 1, r_p = 1, r_d = 0$ and letting r_a vary from -1 to 1; then by setting $r_e = 1, r_p = -1, r_d = 0$ and letting r_a vary from -1 to 0. The negotiation process for these two cases is described in figure 4. Notice that in the bottom figure we do not show negotiation lines for $r_a > 0$, because the value will decrease to negative for this extreme case $r_p = -1$, and the negotiation will end immediately, since it is below the reservation price. In other words, the effects of emotion F can't be less then -1, which makes the negotiation stop. By analyzing these cases, we can find the following properties for emotion feature A:

- A more aroused agent with pleasure status will end up with a deal even more quickly, but a more aroused agent with displeasure status will end up with a deal more slowly; a more non-aroused agent with pleasure status will end up with a deal more slowly, but a more non-aroused agent with displeasure status will end up with a deal more quickly. In other words, arousal magnifies the effect of the agent's pleasure/displeasure status, and non-arousal minimizes the effect of the agent's pleasure/displeasure status.
- Two agents of the same type with the same pleasure and arousal status end up with the same benefit. That is, they reach a deal where A gets \$5 and B gets \$5;
- If two agents of the same type with the same pleasure (displeasure) status but different arousal status engage in negotiation, the result is that the one that is more aroused will benefit less (more).

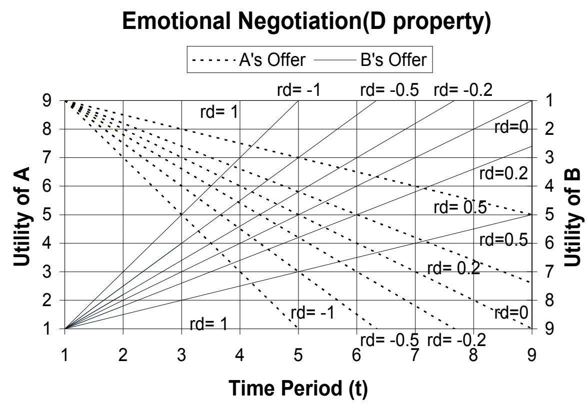

Figure 5: Emotional Negotiation Process ($r_e = 1, r_p = 0$)

Finally, let's consider emotion feature D's property by setting $r_e = 1, r_p = 0$, and letting r_d vary from -1 to 1. The negotiation process is shown in figure 5. As before, we can find the following properties for emotion feature D after analyzing the figure.

- A more dominant agent takes longer to reach an agreement, and a more submissive agent reaches an agreement faster;
- Two agents of the same type, with the same dominance / submissiveness status, end up with equal benefit. That is, they reach a deal where A gets \$5 and B gets \$5;
- If two agents of the same type, but with different dominant / submissive status, engage in negotiation, then the dominant agent will benefit more and the submissive agent will benefit less.

Above all, the properties of the emotion features P, A, D in the model reflect human experience, and agents of the same type with the same emotional status end up with a deal with equal benefit for both as in negotiation theory, which is summarized in table 1.

We note that there are some limitations to our analysis. Namely, agents with complex strategies would not be represented by a linear strategy function, but might require complex curves. Also, an agent's emotional state might change during the course of negotiation, which would have the effect of changing the strategy function. However, our analysis should still hold if we consider small enough time intervals. Any curve can be approximated by a line for small enough lengths and a changing function can be approximated by a fixed function for small enough time steps.

Table 1: Summary of P-A-D Properties

	P-property	A-property	D-property
Range	$r_p \in$[-1, 1]	$r_a \in$[-1, 1]	$r_d \in$[-1, 1]
Positive	Increase evaluation	Increase effects of P-property	Decrease evaluation
Neutral	None effects	None effects	None effects
Negative	Decrease evaluation	Decrease effects of P-property	Increase evaluation
Effect Range (F)	[-r_e, r_e]	[-1, 2$\cdot r_e$]	[-r_e, r_e]
Negotiation Theory	Match	Match	Match

5 Conclusions and Further Work

We proposed an automated negotiation model that incorporates emotions into the agents' strategies. Our model, the PAD Emotional Negotiation Model, maps a nonemotional negotiating agent to an emotional negotiating agent. We evaluated the model and showed that it reflects human experience and negotiation theory. Specifically, the P-dimension shows that a pleasant agent ends up with a deal faster but benefits less in a single trade; the A dimension magnifies or minimizes the trends of the P dimension; the D-dimension shows that a dominant agent insists on its own benefit, and it benefits more from the deal but reaches a deal slower. No matter what, agents of the same type and same emotional status will end up with a deal of equal benefit.

In the work above, we show how the emotional status of agents affects their negotiations, which is an important but very basic step. Since there are no previous numerical human experience data we could compare our model with, we simply gave the reasonable range for the model and verified that it does reflect human experience. The popular practical value ranges for each dimension are still to be collected.

We are continuing work on this research and plan to answer the following questions:

- How do agents produce emotions corresponding to changes in their environment?
- How does an agent perceive other agents' emotional status?
- How does an agent's emotional status affect other agents' emotional status?

Equally important, rather than just knowing how emotional status changes the behavior of agents in negotiation, we need to make agents produce emotions automatically as their environment changes. An agent also needs to perceive other agents' emotional status to find out their intention. This may make the agent change its own emotional status, which might affect its behavior. Much work remains to be done in practical and realistic emotional models for automated negotiation.

References

[Ekman and Davidson, 1994] Paul Ekman and Richard J. Davidson, editors. *The Nature of Emotion: Fundamental Questions*. Oxford University Press, New York, 1994.

[Mehrabian, a] Albert Mehrabian. General tests of emotion or affect for evaluating consumer reactions to services and products. `http://www.kaaj.com/psych/scales/emotion.html`.

[Mehrabian, b] Albert Mehrabian. Incorporating emotions and personality in artificial intelligence software. `http://www.kaaj.com/psych/ai.html`.

[Mehrabian, 1995] Albert Mehrabian. Framework for a comprehensive description and measurement of emotional states. *Genetic, Social, and General Psychology Monographs*, 121(3):339–361, August 1995.

[Picard, 2000] R. W. Picard. *Affective Computing*. MIT Press, Cambridge, MA, 2000.

[Richard J. Davidson and Goldsmith, 2002] Klaus R. Scherer Richard J. Davidson and H. Hill Goldsmith, editors. *Handbook of Affective Sciences*. Oxford University Press, New York, 2002.

[Rosenschein and Zlotkin, 1994] Jeffrey S. Rosenschein and Gilad Zlotkin. *Rules of Encounter*. The MIT Press, Cambridge, MA, 1994.

[Rubinstein, 1982] A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–110, 1982.

[Sousa, 2003] Ronald De Sousa. Stanford encyclopedia of philosophy: Emotion. `http://plato.stanford.edu/entries/emotion/`, 2003.

[Willer, 1999] David Willer. *Network Exchange Theory*. Praeger Publishers, an imprint of Greenwood Publishing Group, Inc., 88 Post Road West, Westport, CT, 1999.

[Wooldridge, 2001] Michael J. Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, LTD, Baffins Lane, Chichester, West Sussex, England, 2001.

[Wright *et al.*, 1996] Wright, Ian, Aaron Sloman, and Luc Beaudoin. Towards a design-based analysis of emotional episodes. *Philosophy, Psychiatry, and Psychology*, 3:101–126, 1996.