# REPORT DOCUMENTATION PAGE

1

Form Approved OMB NO. 0704-0188



| 1. REPORT DATE (DD-MM-YYYY) | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
|---|---|---|
| 05-09-2014 | Conference Proceeding | - |

4. TITLE AND SUBTITLE
A Model of Trust, Moods, and Emotions in Multiagent Systems and its Empirical Evaluation

5a. CONTRACT NUMBER
W911NF-13-1-0094

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS
Anup Kalia, Nirav Ajmeri, Kevin Chan, Jin-Hee Cho, Sibel Adali, Munindar Singh

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
North Carolina State University
2701 Sullivan Drive
Suite 240, Campus Bx 7514
Raleigh, NC 27695 -7003

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
63141-CS.51

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
We study the interplay of moods, emotions, and trust in decision-making contexts characterized by commitments among agents. We develop a general approach representing the relationships among these concepts via a Bayesian network model. Our approach incorporates insights from the literature and provides a computational methodology for identifying improved Bayesian models. Based on observations from

15. SUBJECT TERMS
human behavior, trust systems, multi-agent systems

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Laurie Williams |
| UU | UU | UU | UU | | 19b. TELEPHONE NUMBER |
| | | | | | 919-513-4151 |

## Report Title

A Model of Trust, Moods, and Emotions in Multiagent Systems and its Empirical Evaluation

## ABSTRACT

We study the interplay of moods, emotions, and trust in decision-making contexts characterized by commitments among agents. We develop a general approach representing the relationships among these concepts via a Bayesian network model. Our approach incorporates insights from the literature and provides a computational methodology for identifying improved Bayesian models. Based on observations from an empirical study, we motivate a refined Bayesian model involving the above-mentioned concepts that goes beyond the relationships known in the literature. Our findings include (1) the violation of a commitment affects trust more than its satisfaction; (2) goal satisfaction affects mood and emotion more than commitment satisfaction, but the outcome of a commitment affects trust more than the outcome of a goal; and (3) an agent's prior mood and trust affect whether it satisfies its commitments.

**Conference Name:** 17th AAMAS Workshop on Trust in Agent Societies

**Conference Date:** May 05, 2014

# A Model of Trust, Moods, and Emotions in Multiagent Systems and its Empirical Evaluation

[1]Anup K. Kalia [1]Nirav Ajmeri [2]Kevin S. Chan [2]Jin-Hee Cho [3]Sibel Adalı
[1]Munindar P. Singh

[1]North Carolina State University, Raleigh, NC-27695, US
[2]US Army Research Lab, Adelphi, MD-20783, US
[3]Rensselaer Polytechnic Institute, Troy, NY-12180, US

## Abstract

We study the interplay of moods, emotions, and trust in decision-making contexts characterized by commitments among agents. We develop a general approach representing the relationships among these concepts via a Bayesian network model. Our approach incorporates insights from the literature and provides a computational methodology for identifying improved Bayesian models. Based on observations from an empirical study, we motivate a refined Bayesian model involving the above-mentioned concepts that goes beyond the relationships known in the literature. Our findings include (1) the violation of a commitment affects trust more than its satisfaction; (2) goal satisfaction affects mood and emotion more than commitment satisfaction, but the outcome of a commitment affects trust more than the outcome of a goal; and (3) an agent's prior mood and trust affect whether it satisfies its commitments.

## 1 Introduction

In a multiagent system, a *truster's* trust in a *trustee* is crucial to the truster's decision-making regarding a future interaction with the trustee. The trust is modulated by the truster's moods and emotions and affected by the outcome of the interaction. For example, suppose Alice requests Bob to perform a task for her. If Bob accepts the task, a commitment is created from Bob toward Alice. When Bob performs the task, the commitment is satisfied and Alice's trust may increase for Bob. In this case, Alice's mood may become positive and if she displays a positive emotion to Bob for helping her out, Bob's mood may become positive. The change in Bob's mood and a positive emotional response from Alice may end up improving Bob's trust for Alice and lead to continued interactions between them. Alternatively, Alice's failing to acknowledge Bob's help may result in a decrease in Bob's trust for Alice, leading to his disregarding a subsequent request from her. This example illustrates that trust is affected not only by commitments, but also by emotions and moods.

Trust and emotions have been theoretically and empirically studied both in social and cognitive psychology and in multiagent systems. De Melo et al.'s [7] Bayesian model considers emotions to predict a counterparty's intentions regarding cooperation. Burnett et al. [4] propose a trust-based decision-theoretic model for delegation. Dunn and Schweitzer [8] show that emotions such as gratitude and happiness increase trust whereas emotions such as anger decrease trust. Antos et al. [1] state that emotions influence the perception of the trustworthiness of agents to cooperate or defect in strategy-based games. The above works are promising but limited for



In: R. Cohen, R. Falcone and T. J. Norman (eds.): Proceedings of the 17th International Workshop on Trust in Agent Societies, Paris, France, 05-MAY-2014, published at http://ceur-ws.org

understanding how moods and emotions affect trust in decision-making: they (1) do not discuss how trust and emotions arise, and (2) do not clearly distinguish moods and emotions.

We address these aspects in this paper. Specifically, we introduce a novel approach for trust that builds on existing studies linking moods, trust, emotions, commitments, and goals. We adopt a limited form of appraisal theory [3, 14] assuming that moods and emotions are triggered not continuously, but at discrete points corresponding to state changes of goals and commitments.

Our empirical evaluation used Gal et al.'s [11] Colored Trails game as a decision-making platform. We collected data about subjects' commitments to each other; outcomes of their commitments; their emotions as displayed to others; their moods and trust in their opponents. We trained and evaluated Bayesian models reflecting different causal relationships. Our results yield insights into how the relevant concepts relate to decision-making. Our main findings are (1) commitment violation affects trust more than satisfaction; (2) goal satisfaction affects mood and emotion more than satisfaction of a commitment by another party, but the outcome of a commitment affects trust more than the outcome of a goal; and (3) an agent's prior mood and trust affect whether it satisfies its commitments.

In obtaining the above-mentioned findings, we begin from a model postulated based on previous findings in the literature regarding commitments, mood, trust, and emotions; obtain a superior model that highlights an unexpected relationship between trust and emotion; refine it further to introduce a relationship between goals and commitments; and refine it again to incorporate the effect of mood and trust on commitments.

## 2 Technical Approach

We treat commitment, trust, mood, and emotion as discrete random variables for our Bayesian model.

*Commitment* $\mathsf{C}_{A,B}$. A commitment $\mathsf{C}_{A,B}(r,u)$ means that a debtor $A$ commits a creditor $B$ to bring about a consequent $u$ provided an antecedent $r$ holds. In essence, a commitment provides grounds for a creditor to expect some actions from a debtor [19]. A commitment has one of two significant *outcomes*: it is *satisfied* (sat) when $u$ holds regardless of whether $r$ does; it is *violated* (vio) if $r$ holds but $u$ fails to hold. Accordingly, we distinguish two values, sat and vio.

*Trust* $\mathsf{T}_{A,B}$. Trust is described as $A$'s expectation of $B$ to bring about the specified condition, i.e., to satisfy a commitment [5]. Thus, we represent trust directed from $A$ to $B$ as $\mathsf{T}_{A,B}$. We consider the variable $\mathsf{T}_{A,B}$ as ordered with possible values of *low*, *medium*, and *high*.

*Emotion* $\mathsf{E}_{A,B}$. An emotion is a (transient) response of an agent to a significant external or an internal event [10, 21]. We consider emotions as displayed to others. In particular, in a two-party setting, one party's emotion at an event is directed at its counterparty (assumed as being responsible for the event) and is displayed to the counterparty. Accordingly, we notate $A$'s emotion toward $B$ as a variable $\mathsf{E}_{A,B}$ with possible values *positive* (pos), *negative* (neg), and *neutral* (neu).

*Mood* $\mathsf{M}_A$. A mood is a low-intensity, long-lasting condition of feeling good or bad [10, 9]. For example, a positive event occurring in the morning may leave someone in a good mood for the entire day. A mood is felt, not directly displayed. Accordingly, we represent $A$'s mood as $\mathsf{M}_A$ with possible values *negative* (neg), *neutral* (neu), and *positive* (pos).

*Goal* $\mathsf{G}_A$. A goal is a private condition of wanting something. Goals motivate agents to act but are not directly visible to others. A goal $\mathsf{G}_A$ has possible values *satisfied* (sat) or *failed* (fai).

### 2.1 Relations between Concepts

We begin by reviewing the literature to identify causal relationships between the foregoing concepts and express them in a *baseline* Bayesian model. We produce variants of the baseline by altering some of the relationships but holding constant the core ideas that commitment outcomes affect mood, emotion, and trust and emotion affects mood. Below, $\rightarrow$ indicates means causes or affects.

*Trust and Emotion.* Dunn and Schweitzer [8] state that positive emotions (happiness, gratitude) increase trust, whereas negative emotions (anger) decrease trust. We capture this intuition as $\mathsf{E}_{A,B}$:pos $\rightarrow$ $\mathsf{T}_{B,A}$:inc and $\mathsf{E}_{A,B}$:neg $\rightarrow$ $\mathsf{T}_{B,A}$:dec. Here inc means increase and dec means decrease.

*Trust and Commitments.* Chopra et al. [6] and Singh [20] relate trust to commitments. We capture the intuition that trust respectively increases or decreases in light of the outcome of a prior interaction as $\mathsf{C}_{A,B}$:sat $\rightarrow$ $\mathsf{T}_{B,A}$:inc and $\mathsf{C}_{A,B}$:vio $\rightarrow$ $\mathsf{T}_{B,A}$:dec.

*Mood and Emotions.* Sheldon and Lyubomirsky [18] state that an emotion such as gratitude triggers a positive mood. We assume that a positive emotion from $A$ to $B$ brings about a positive mood of $B$. We capture our intuition as $\mathsf{E}_{A,B}$:pos $\rightarrow$ $\mathsf{M}_B$:pos.

*Commitments vis à vis Mood and Emotions.* Guiraud et al. [12] relate positive moods (e.g., joy) and emotions (e.g., gratitude) with goal achievement, and negative moods (e.g., sadness) and negative emotion (e.g., disappointment) with goal failure. Commitments ordinarily relate to goals in that one party's goal would be the consequent of the counterparty's commitment to it. In a cooperative setting, $B$'s commitment toward $A$ corresponds to $A$'s goal. We capture Guiraud et al.'s intuition initially in terms of commitments: (1) $\mathsf{C}_{A,B}$:sat $\rightarrow$ $\mathsf{M}_B$:pos; (2) $\mathsf{C}_{A,B}$:vio $\rightarrow$ $\mathsf{M}_B$:neg; (3) $\mathsf{C}_{B,A}$:sat $\rightarrow$ $\mathsf{E}_{A,B}$:pos; and (4) $\mathsf{C}_{B,A}$:vio $\rightarrow$ $\mathsf{E}_{A,B}$:neg. Later, we introduce goals explicitly into our model.

### 2.2 Constructing a Bayesian Model

The baseline model in Figure 1 consolidates the above intuitions as causal relations in a Bayesian model [17, 16]: (1) a commitment outcome affects trust, (2) a commitment outcome affects mood, (3) a commitment outcome affects emotion, and (4) emotion affects mood. The remaining models of Figure 1 capture meaningful variants of the baseline.

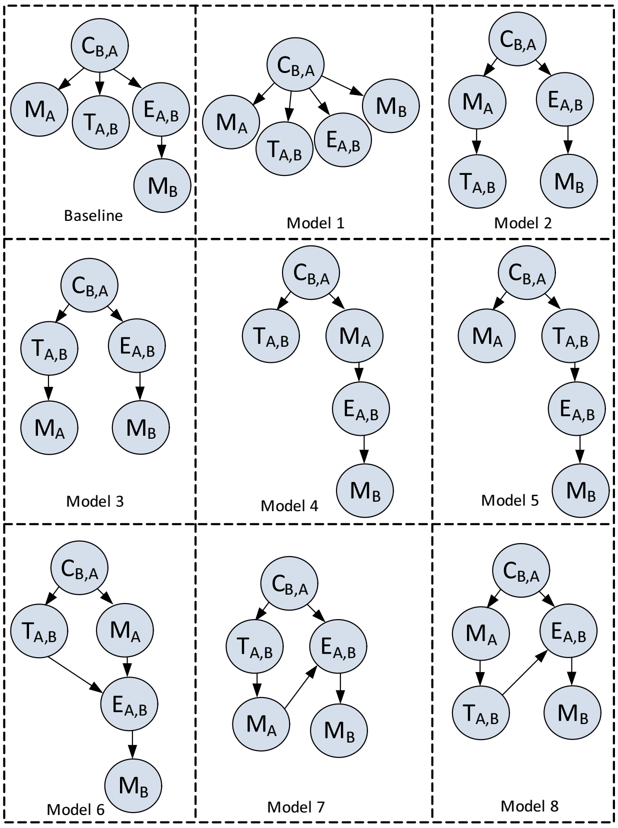


Figure 1: The baseline model (top left) and variants.

## 3 Hypotheses in Conceptual Terms

To understand the relationships between the concepts of interest modularly, we formulate and evaluate the following hypotheses. Since these hypotheses are causal statements, we first learn the parameters of a Bayesian model via expectation maximization and then use the learned parameters to evaluate each hypothesis with respect to the given trained model.

The $Group_1$ hypotheses pertain to the baseline model.

- $H_1$ Commitment satisfaction increases the creditor's trust in the debtor, whereas violation decreases trust.

- $H_2$ Commitment satisfaction yields a positive mood for the creditor and violation yields a negative mood.
- $H_3$ Commitment satisfaction makes the creditor's emotion to the debtor positive.
- $H_4$ A positive emotion causes the recipient's mood to be positive and a negative emotion causes the recipient's mood to be negative.
- $H_5$ The baseline model fits better than its variants in Figure 1.

The *Group*$_2$ hypotheses concern the centrality of commitment outcomes and their effect on moods and trust relative to goal success or failure.

- $H_6$ The outcome of a commitment affects trust, emotion, and mood more than a goal.
- $H_7$ The baseline model fits better than models that incorporate subjects' goals.

The *Group*$_3$ hypothesis concerns how commitments arise given prior mood or trust.

- $H_8$ A model that includes both trust and mood fits better than one that omits trust or mood.

The metrics we use to evaluate alternative Bayesian models are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) scores, which capture the goodness of fit of a model to the data. Given a Bayesian model, the metrics we use to evaluate our hypotheses are (1) the conditional probability distributions (CPDs) or the parameters of Bayesian models learned from the data and (2) the F-measures of a classifier based on a model to determine the accuracy of predicting the value of a specified variable.

## 4 Empirical Study Design

To help empirically ground our theory, we developed a web-based variant of Gal et al.'s [11] Colored Trails game. Figure 2 shows a typical game round that has a game board (4x4 matrix), available resources (a set of colored tiles), a chat interface to communicate with the opponent, and a transfer window to exchange tiles. In each round, the subjects are allocated a set of colored tiles. They are assigned starting positions and a common goal position on the game board. The objective of the game for a subject is to use the available tiles to reach the goal position. The subjects can advance to an adjacent cell on the game board if they have the same colored tile available. Each move consumes one tile. During the game play, the subjects are allowed to communicate with each other through the chat interface, and to trade tiles. As compared to the original Colored Trails game [11] where subjects are restricted to communicate through a fixed message protocol, we provide a free chat interface to communicate. Through the chat interface, subjects can negotiate and exchange tiles. Subjects can also express their emotions such as gratitude and anger toward their opponents by sending text messages. We map our variables to data gathered from the game play, text chats, and subject surveys. Each game involves two subjects, Alice ($A$) and Bob ($B$).

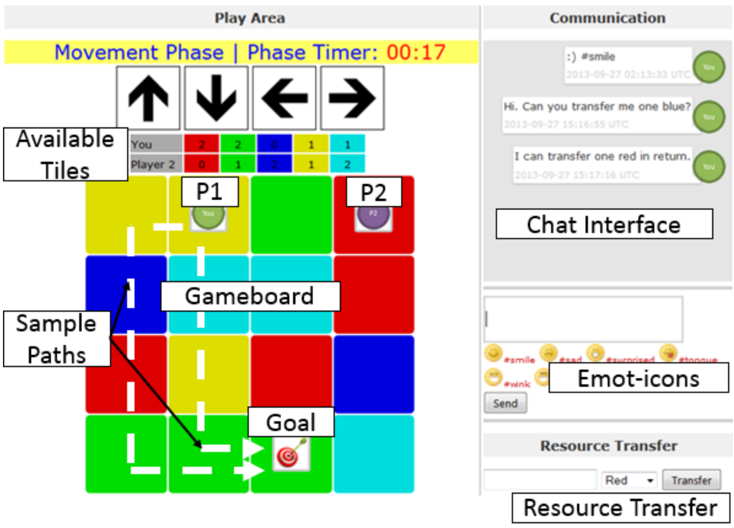


Figure 2: A screenshot of the Colored Trails game.

**Rules of the Game.** (1) A subject is required to play three games with different opponents. (2) Each game consists of five rounds. (3) Each round has a common goal position, and different starting positions for each subject. (4) In all rounds, subjects are allocated the same number but a different set (randomly selected) of colored tiles. (5) Subjects can communicate with their opponents via a chat interface, using which they can negotiate to transfer tiles to each other. (6) At the end of each round, each subject fills a survey.

*Commitment.* During the game, $A$ can create a commitment by agreeing (through chat) to transfer a specified number of tiles to $B$. If $A$ provides the tiles, $A$ satisfies its commitment, else violates it.

*Emotion.* $B$ may display an emotion toward $A$ via a chat message. We determine whether the emotion is positive, negative, or neutral by analyzing the text.

*Mood.* The subjects' mood changes as the games progress and may spill over from events unknown to the experimenter as well as from one game to the next. We determine subjects' mood from their survey answers.

*Trust.* We determine subjects' trust in their opponents from their survey answers.

**Selection of Subjects.** We advertised the study to recently admitted (to minimize the threat of prior knowledge of our study) graduate students in Computer Science. We offered a payment of 10–20 USD each, depending on success in the game. We selected 30 subjects (25 male; 5 female).

**Surveys.** (1) At the start of each game and end of each round, we asked the subjects to record their moods on a five-point scale (very negative to very positive). We mapped the responses to our mood variable. (2) Before the beginning of each round, subjects recorded on a five-point scale (very low to very high) whether they expected their opponents to cooperate. We mapped the responses to our trust variable. (3) At the end of each game, we asked the subjects to report their mood, their trust in their opponent, whether the opponent's commitments were satisfied, and whether they want to team up with this opponent in any team game.

**Data.** We collected 450 rows of data (30 subjects $\times$ 30 games $\times$ 5 rounds each game), including their survey forms, and whatever chat messages they exchanged. From their interactions we manually analyzed the chat to identify the commitments created, satisfied, and violated, and the emotions displayed. We prepared a case file that expresses a subject's change of mood, trust, and emotion. For example, if Alice reports her mood as positive before a round and negative afterward, we report this as a decrease in the mood.

**Expectation Maximization.** Based on the gathered data, we performed Expectation Maximization (EM) using the Hugin tool [15] to learn the conditional probabilities and AIC and BIC scores of the Bayesian networks given in Figure 1. We set a convergence threshold of 0.0004 and the maximum number of iterations as 100. We initialized the outcomes of a commitment as 0.5, and the outcomes of trust, emotion, and mood as 0.33.

## 5 Results

We performed three-fold training and testing on the overall data for each model we considered.

### 5.1 Group$_1$: Evaluating the Baseline Model

Below, we abbreviate increase, decrease, no change, satisfaction, and violation respectively as inc, dec, noc, sat, and vio.

**Verifying H$_1$.** The learned conditional probability distributions (CPDs) show that when $\mathsf{C}_{B,A}$ is satisfied, $A$'s trust in $B$ is mostly unaffected, but when $\mathsf{C}_{B,A}$ is violated, $A$'s trust in $B$ suffers. Specifically, $\text{P}(\mathsf{T}_{A,B}\text{=inc}|\mathsf{C}_{B,A}\text{=sat})\text{=}0.22$ and $\text{P}(\mathsf{T}_{A,B}\text{=dec}|\mathsf{C}_{B,A}\text{=vio})\text{=}0.54$. That is, commitment violation affects trust more than satisfaction.

**Verifying H$_2$.** When $\mathsf{C}_{B,A}$ is satisfied, $A$'s mood is unaffected, but when $\mathsf{C}_{B,A}$ is violated, $A$'s mood decreases. Specifically, $\text{P}(\mathsf{M}_{A}\text{=inc}|\mathsf{C}_{B,A}\text{=sat})\text{=}0.29$ and $\text{P}(\mathsf{M}_{A}\text{=noc}|\mathsf{C}_{B,A}\text{=sat})\text{=}0.43$, and $\text{P}(\mathsf{M}_{A}\text{=dec}|\mathsf{C}_{B,A}\text{=vio}) = 0.60$. That is, commitment violation affects mood more than satisfaction.

**Verifying H$_3$.** When $\mathsf{C}_{B,A}$ is satisfied, $A$ is likely to display a positive emotion to $B$, and when $\mathsf{C}_{B,A}$ is violated, $A$ is unlikely to display a negative emotion to $B$. Specifically, $\text{P}(\mathsf{E}_{A,B}\text{=inc}|\mathsf{C}_{B,A}\text{=sat})\text{=}0.89$ and $\text{P}(\mathsf{E}_{A,B})\text{=dec}|\mathsf{C}_{B,A}\text{=vio})\text{=}0.47$. That is, subjects are polite toward their opponents and do not display as many negative emotions as one might expect in the case of violation.

**Verifying H$_4$.** $A$'s display of a positive emotion toward $B$ does not affect $B$'s mood ($\text{P}(\mathsf{M}_{B}\text{=noc}|\mathsf{E}_{A,B}\text{=inc})\text{=}0.44$) whereas $A$'s display of a negative emotion toward $B$ makes $B$'s mood negative ($\text{P}(\mathsf{M}_{B}\text{=dec}|\mathsf{E}_{A,B}\text{=neg})\text{=}0.53$). That is, a negative emotion affects mood more than a positive emotion.

**Verifying H$_5$.** Figure 3 shows the negated AIC and BIC scores (higher magnitudes indicate superior fit) for the models of Figure 1. From the results, it is apparent that Model 6 is the best model.

Interestingly, the CPDs of Model 6 are somewhat slightly higher than those of the baseline. For example, P($\mathsf{T}_{A,B}$=dec|$\mathsf{C}_{B,A}$=vio)=0.54 in the baseline improves to P($\mathsf{T}_{A,B}$=dec|$\mathsf{C}_{B,A}$=vio)=0.56. Similarly, P($\mathsf{M}_B$=dec|$\mathsf{E}_{A,B}$=neg)=0.53 in the baseline improves to P($\mathsf{M}_B$=dec|$\mathsf{E}_{A,B}$=dec)=0.77. These results show not only that Model 6 fits the data better than the baseline (as the AIC and BIC scores indicate) but also that Model 6 coheres better with our intuitions about how the various concepts relate to one another.

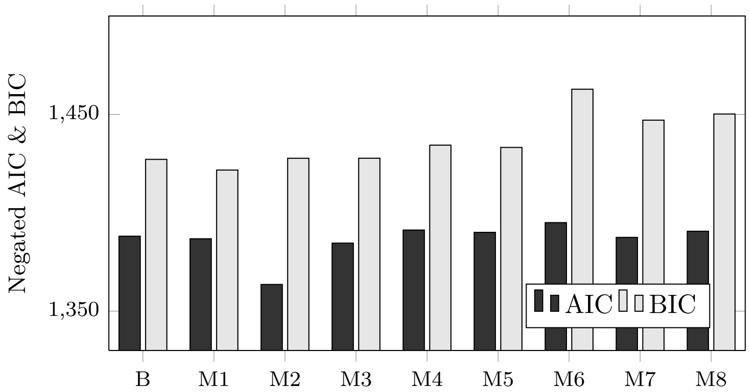


Figure 3: AIC and BIC score for each model of Figure 1.

We compared the baseline model and Model 6 in terms of predicting an unknown node (commitment, mood, trust, and emotion) from known inputs. For example, given a commitment outcome, we can predict if the mood is negative or the emotion is positive. We map a probability above a threshold $\gamma$ as categorical prediction. Table 1 shows the average of precision (P), recall (R), and F-measure (F) obtained by considering five possible values for $\gamma$, namely, $\{0.31, 0.32, 0.33, \ldots, 0.7\}$ values. The F-measures for $\mathsf{M}_A$=dec and $\mathsf{T}_{A,B}$=dec in Model 6 are slightly better than those for the baseline.

Table 1: Accuracy for the baseline model and Model 6.

| | **Baseline** | | | **Model 6** | | |
|---|---|---|---|---|---|---|
| **Prediction** | **P** | **R** | **F** | **P** | **R** | **F** |
| $\mathsf{M}_A$=dec | 0.32 | 0.32 | 0.32 | 0.32 | 0.34 | 0.33 |
| $\mathsf{T}_{A,B}$=dec | 0.19 | 0.26 | 0.22 | 0.19 | 0.49 | 0.26 |
| $\mathsf{E}_{A,B}$=inc | 0.19 | 0.09 | 0.12 | 0.06 | 0.39 | 0.10 |
| $\mathsf{M}_B$=dec | 0.84 | 0.05 | 0.09 | 0.96 | 0.00 | 0.00 |

### 5.2 Group$_2$: Evaluating the Cause of Mood, Trust, and Emotion

Since Model 6 proved superior to the baseline, we use it as a basis for further evaluation and improvements. Figure 4 shows Model 6 and three additional models to help investigate the comparative power of goals and commitments in affecting trust and emotions. Here, $\mathsf{G}_A$ represents $A$'s success with a goal. In the study, we assume a goal is achieved when a player builds a colored trail that reaches the goal position assigned to the player. Model 9 replaces $\mathsf{C}_{B,A}$ by $\mathsf{G}_A$; Models 10 and 11 respectively insert $\mathsf{G}_A$ and $\mathsf{C}_{B,A}$ as the roots on Models 9 and 10.

**Verifying H$_6$.** We ran EM on Model 9 and obtained its CPDs. These CPDs indicate that mood is affected more when a goal is satisfied than when a commitment is satisfied. Specifically, P($\mathsf{M}_A$=inc|$\mathsf{G}_A$=sat) = 0.41 > P($\mathsf{M}_A$ = inc| $\mathsf{C}_{B,A}$ = sat) = 0.29. Given trust and mood, emotion is independent of commitments and goals in Models 6 and 9, respectively. Therefore, we evaluated another model (not shown for brevity) produced by editing our baseline model by replacing commitment outcome with goal outcome. The results indicate that goal satisfaction affects emotions more than commitment satisfaction P($\mathsf{E}_{A,B}$ = inc | Goal$_A$=sat) = 0.91, P($\mathsf{E}_{A,B}$ = pos|$\mathsf{C}_{B,A}$ = sat) = 0.89). Considering trust in Models 6 and 9, we found that trust is affected more when a commitment is violated than when a goal fails. Specifically, P($\mathsf{T}_{A,B}$ = dec|$\mathsf{C}_{B,A}$ = vio) = 0.55, P($\mathsf{T}_{A,B}$ = dec|$\mathsf{G}_A$ = fai) = 0.29).

**Verifying H$_7$.** Model 9 yields better AIC and BIC scores than Model 6, indicating that goal outcome affects mood and emotion more than commitment outcome. Models 10 and 11 include both goal and commitment

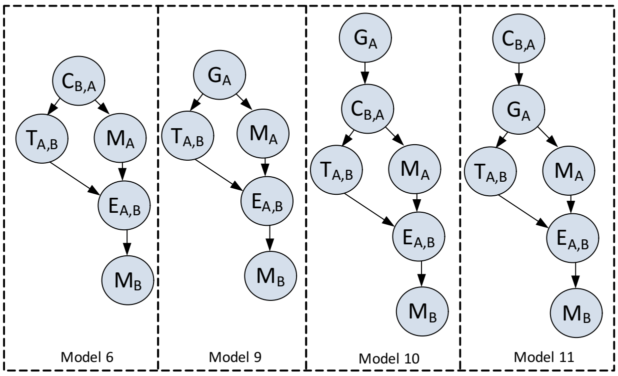


Figure 4: Model 6 and its enhancements with goals.

outcomes. These models obtain higher negated AIC and BIC scores than Models 6 and 9, suggesting that we ought to consider both commitments and goals in relation to trust, emotions, and mood. Moreover, Model 10 obtains higher negated AIC and BIC scores than Model 11, suggesting that there is a greater correlation between the outcome of a goal and of a commitment that supports the goal (Model 10) than between the outcome of a commitment and of the associated goal (Model 11). This observation supports the intuition that a goal may fail even if the relevant commitment is satisfied whereas if a goal succeeds, one expects all relevant commitments to also have been satisfied.

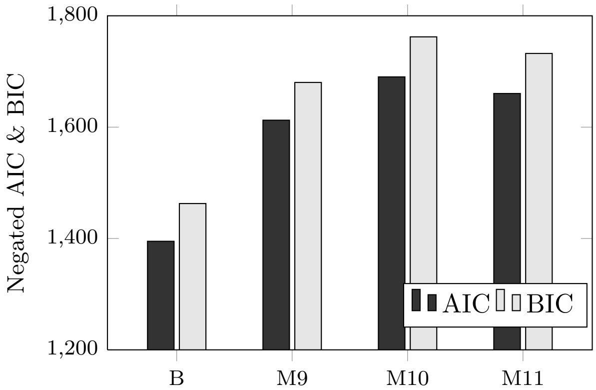


Figure 5: AIC and BIC scores for models of Figure 4.

Table 2 reports the F-measures obtained from Models 6 and 9. The results are about the same with a slight edge for Model 6.

Table 2: Accuracy for Models 6 and 9.

| | **Model 6** | | | **Model 9** | | |
|---|---|---|---|---|---|---|
| **Prediction** | **P** | **R** | **F** | **P** | **R** | **F** |
| $\mathsf{M}_A$=inc | 0.27 | 0.85 | 0.40 | 0.27 | 0.85 | 0.40 |
| $\mathsf{T}_{A,B}$=dec | 0.19 | 0.49 | 0.26 | 0.2 | 0.12 | 0.15 |

### 5.3 Group$_3$: Evaluating the Effect of Trust and Mood on Commitments

We now consider whether both mood and trust are important in predicting commitment outcomes. Figure 6 shows three alternative models relating $A$'s preexisting mood and trust with the outcome of $A$'s commitment in the current round. We do not consider emotion here because a mood is persistent whereas an emotion is transient. Our data confirms that in 75% of the cases a subject's mood carries over from one game to the next.

**Verifying H$_8$.** Models 12 and 13 that consider either mood or trust but not both obtain lower AIC and BIC scores than Models 14, 15, and 16, which consider both. Model 13, which considers only trust, obtains a significantly low score than Model 12, which considers only mood. That is, prior mood is an important predictor of a commitment outcome. Model 14 obtains better scores than Models 15 and 16. That is, mood and trust act independently in predicting the outcome of a commitment. Moreover, Model 15 obtains slightly better scores than Model 16, suggesting that mood affects trust more than trust affecting mood. Additionally, Figure 1 (with reference to Hypothesis $H_5$) shows that models that consider both mood and trust in predicting emotions obtain better scores than others. This suggests that mood and trust go hand-in-hand both for predicting a new commitment as well as an emotion. For additional verification, we computed the F-measure of each model for predicting a commitment given trust and mood, obtaining 0.29 for Models 12, 13, 15, and 16, and 0.26 for Model 14. The F-measures are not significantly different because the data for commitment outcomes is inadequate.

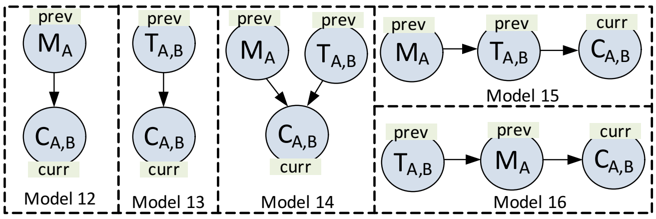


Figure 6: Models relating prior trust and mood with current commitment outcome.

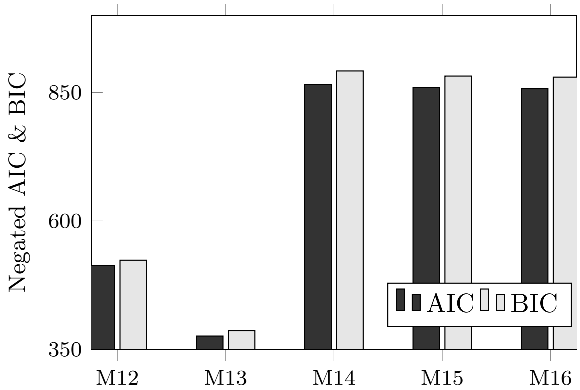


Figure 7: AIC and BIC scores for models of Figure 6.

## 6 Discussion and Conclusion

We propose a Bayesian model to describe commitments, trust, moods, and emotions between agents. We conducted an empirical study to obtain data from human subjects. We evaluated our hypotheses via three successive evaluations. Our first evaluation shows that we were incorrect in postulating the causal structure as given in the baseline model (i.e., based on the literature). Instead Model 6 obtained better negated AIC and BIC scores indicating that it fits the data better than the baseline model. In our second evaluation, we found that introducing goal satisfaction affects mood and emotion more than commitment outcome whereas commitment outcome affects trust more than goal satisfaction. We also found that considering both goal and commitment outcomes improved the negated AIC and BIC scores over Model 6, which does not consider goal outcomes. In our third evaluation, we found that both mood and trust are important in predicting the outcome of a commitment. Figure 8 shows a composite model that consolidates all the main findings of this paper.

### 6.1 Related Work

Let us now describe some related work. De Melo et al. [7] propose a Bayesian model that captures an appraisal-based mechanism for the interpersonal affect of emotion in decision-making. Their work does not consider trust and is focused on finding a relationship between emotions and decision-making behavior. Unlike their work, we emphasize finding a general relationship between emotions and trust.

Antos et al. [1] carry out a series of negotiation games between humans and computational agents, followed by a trust game where humans choose an agent from among several agents to entrust a fraction of their profit

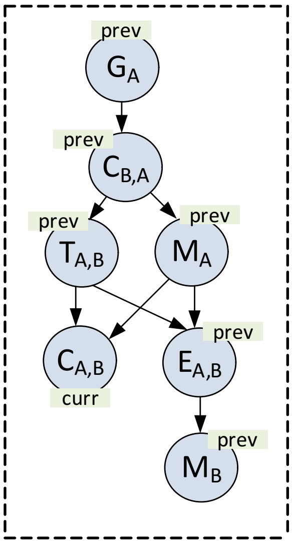


Figure 8: Composite model.

earned from the negotiation games to the agent chosen. Unlike their work where agents only display emotions, we focus on building a model that considers both emotions and trust. Using our model an agent can choose different strategies based on its emotions and trust for an agent.

Antos and Pfeffer [2] provide a methodology for decision-making by agents that leverages a computational concept of emotions. Their experimental results indicate that emotion-based agents outperform other reasonable heuristics and closely approach computationally expensive near-optimal solutions. Their work is similar to De Melo et al. [7], who establish a relation between emotions and decision-making behavior. In contrast, we focus on establishing relationship between emotions, mood, trust, and commitments.

Tanguy et al. [22] provide a model Dynamic Emotion Representation (DER) that integrates emotional responses and keeps track of emotion intensities changing over time. Our work differs from this work by dealing with multiple mental or emotional processes

### 6.2 Threats to Validity and Future Work

The topics we studied are complex themes in social psychology. Inevitably, an empirical study such as ours faces some threats to validity. Our subject population of computer scientists is not representative of typical users, though we mitigated this threat by selecting subjects who were not well-versed in the research area studied. The artificial setting of a game would tend to have lower stakes than many real-life interactions and may thus elicit limited emotional responses. To control our experiment, we prevented subjects from (1) knowing their counterparty and (2) projecting or receiving any visual or auditory signals. In real-life, past relationships are significant as is nonverbal communication. Thus our findings may not propagate well to real-life settings.

In future work, we plan to address above threats. On the theoretical side, we plan to complete the picture with regard to decision-making, especially accounting for variations in the intensity of mood, trust, and emotions. On the practical side, we plan to exploit and enhance emerging text analysis techniques, e.g., Kalia et al.'s [13] approach for extracting commitments from chats, to help automate the extraction of commitments and emotions, and facilitate studies at a larger scale.

## References

[1] Dimitrios Antos, Celso de Melo, Jonathan Gratch, and Barbara Grosz. The influence of emotion expression on perceptions of trustworthiness in negotiation. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence*, pages 772–778, San Francisco, 2011.

[2] Dimitrios Antos and Avi Pfeffer. Using emotions to enhance decision-making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 24–30, Barcelona, 2011.

[3] Magda B. Arnold. *Emotion and Personality.* Columbia University, New York, 1960.

[4] Chris Burnett, Timothy J. Norman, and Katia P. Sycara. Trust decision-making in multi-agent systems. In *IJCAI*, pages 115–120, Barcelona, 2011.

[5] Cristiano Castelfranchi and Rino Falcone. Principles of trust for MAS: Cognitive anatomy, social importance, and quantification. In *Proceedings of the 3rd International Conference on Multiagent Systems (ICMAS)*, pages 72–79, Paris, 1998. IEEE Computer Society.

[6] Amit K. Chopra, Elda Paja, and Paolo Giorgini. Sociotechnical trust: An architectural approach. In *Proceedings of the 30th International Conference on Conceptual Modeling (ER)*, volume 6998 of *LNCS*, pages 104–117, Brussels, 2011. Springer.

[7] Celso M. de Melo, Peter Carnevale, Stephen Read, Dimitrios Antos, and Jonathan Gratch. Bayesian model of the social effects of emotion in decision-making in multiagent systems. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 55–62, Valencia, 2012.

[8] Jennifer R. Dunn and Maurice E. Schweitzer. Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology*, 88(5):736–748, May 2005.

[9] Joseph P. Forgas. Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117(1):39–66, January 1995.

[10] Jay Friedenberg and Gordon Silverman. *Cognitive Science: An Introduction to the Study of Mind.* SAGE Publications, 2nd edition, 2012.

[11] Ya'akov Gal, Barbara Grosz, Sarit Kraus, Avi Pfeffer, and Stuart Shieber. Agent decision-making in open-mixed networks. *Artificial Intelligence*, 174(18):1460–1480, 2010.

[12] Nadine Guiraud, Dominique Longin, Emiliano Lorini, Sylvie Pesty, and Jérémy Rivière. The face of emotions: A logical formalization of expressive speech acts. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems*, pages 1031–1038, Taipei, 2011.

[13] Anup Kalia, Hamid Reza Motahari Nezhad, Claudio Bartolini, and Munindar P. Singh. Monitoring commitments in people-driven service engagements. In *Proceedings of the 10th IEEE Conference on Services Computing*, pages 160–167, Santa Clara, 2013.

[14] Richard S. Lazarus. *Psychological Stress and the Coping Process.* McGraw-Hill, New York, 1966.

[15] Anders L. Madsen, Michael Lang, Uffe B. Kjaerulff, and Frank Jensen. The hugin tool for learning bayesian networks. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 2711 of *Lecture Notes in Computer Science*, pages 594–605. Springer Berlin Heidelberg, 2003.

[16] Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000.

[17] Richard Scheines. An introduction to causal inference. Technical Report 430, CMU Department of Philosophy, Pittsburgh, January 1997.

[18] Kennon M. Sheldon and Sonja Lyubomirsky. How to increase and sustain positive emotion: The effects of expressing gratitude and visualizing best possible selves. *The Journal of Positive Psychology*, 1(2):73–82, 2006.

[19] Munindar P. Singh. An ontology for commitments in multiagent systems: Toward a unification of normative concepts. *Artificial Intelligence and Law*, 7(1):97–113, March 1999.

[20] Munindar P. Singh. Trust as dependence: A logical approach. In *Proceedings of the 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 863–870, Taipei, May 2011.

[21] Craig A. Smith and Phoebe C. Ellsworth. Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48(4):813–838, April 1985.

[22] Emmanuel Tanguy, Philip J. Willis, and Joanna Bryson. Emotions as durative dynamic state for action selection. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pages 1537–1542, Hyderbad, 2007.

# 7 Appendix

We asked the subjects to fill out the following survey forms.

## 7.1 Survey Form I

The first survey form was given to the subjects before they started to play the game. In the survey, the following questions were asked along with the choices for each question.

- How are you feeling today? The choices given were: very negative, low negative, normal, positive, and very positive
- Do you think your opponent will help you to achieve your goals when you ask for it? The choices given were: no, may not be, not sure, may be, and yes
- If your opponent does not help to you achieve your goal, what will be your reaction to your opponent? The choices given were: very negative, low negative, nothing, positive, and very positive
- If your opponent expresses his or her anger to you, how will you feel? The choices given were: very negative, low negative, normal, positive, and very positive
- Will you help your opponent if he or she asks for it? The choices given were: no, may not be, not sure, may be, and yes

## 7.2 Survey Form II

The second survey form was given to the subjects at the end of each game. In the survey, the following questions were asked along with the choices for each question.

- How are you feeling after the game? The choices given were: very negative, negative, normal, positive, and very positive
- During the game, did you request tiles from your opponent? The choices given were: yes, no, and not applicable
- How did you feel when your opponent helped you? The choices given were: very negative, negative, normal, positive, very positive, and not applicable
- When your opponent expressed you a positive or a very positive emotion, how did you feel? The choices given were: very negative, negative, normal, positive, very positive, and not applicable
- When your opponent expressed you a negative or a very negative emotion, how did you feel? The choices given were: very negative, negative, normal, positive, and very positive
- When your opponent expressed you nothing, how did you feel? The choice given were: very negative, negative, normal, positive, very positive, not applicable
- Do you expect your opponent will help you in the next game if you request for it? The choices given were: no, may not be, not sure, may be, and yes

## 7.3 Surver Form III

We also asked few questions before and after each round of a game. Before a round started, we asked them the following questions.

- How much do you expect your opponent to cooperate? The choices given were: very low, low, medium, high, and very high.
- How do you feel? The choices given were: very negative, negative, normal, positive, and very positive

After the end of a round we asked them the following.

- How are you feeling now? The choices given were: very negative, negative, normal, positive, and very positive