Nonverbal Behavior Observation:
Collaborative Gaming Method
for Prediction of Conflicts
during Long-Term Missions

Natalia Voynarovskaya, Roman Gorbunov, Emilia Barakova, Rene Ahn,
and Matthias Rauterberg

Eindhoven University of Technology,
Postbus 513, 5600MB, Eindhoven, The Netherlands
{n.voynarovskaya,r.gorbunov,e.i.barakova,r.m.c.ahn}@tue.nl,
g.w.m.rauterberg@tue.nl

Abstract. This paper presents a method for monitoring mental state of small isolated crews during long-term missions (such as space mission, polar expeditions, submarine crews, meteorological stations, and etc). It combines the records of negotiation game with monitoring of the nonverbal behavior of the players. We analyze the records of negotiation game that has taken place between the crew members who were placed in isolated environment for 105 days during the Mars-500 experiment. The outcomes of the analysis, differently from the previously made conclusions, show that there was not a significant deviation of the rational choice of the players. We propose an extension of the method that includes monitoring of the nonverbal behavior of the players next to recording the game records. The method is focused on those aspects of psychological and sociological states that are crucial for the performance of the crew. In particular, we focus on measuring of emotional stress, initial signs of conflicts, trust, and ability to collaborate.

Keywords: Colored Trails, social network analysis, nonverbal communication, emotions, long-term missions, evolutionary game theory.

1 Introduction

Long-term missions, such as a space flight or polar expedition, usually are performed by small group of people, and characterized by extreme physical and psychological parameters. The psychological parameters, like small group living, extreme social isolation, working together in close proximity [2], could be crucial for the mental state of crew members as well as the whole crew. The interpersonal issues, such as crew tension, cohesion, leadership, language, cultural factors [1], could become reasons for personal conflicts which, in their turn, could affect performance of the mission. For instance, inter-crew tension can cause formation of subgroups, disruptions of cohesion, scapegoating, refusal of communication and collaboration.

For measuring, prediction and prevention of the above mentioned problems, it is crucial to have methods for automatic monitoring of psychological and social states in the group. The psychological states of the crew members, as well as different aspects of interpersonal relations, have been assessed using different approaches. In the current study we combine three approaches: strategic multiplayer game, [8, 11], techniques for automatic monitoring of nonverbal behavior [16,17,18,21,22,23], and self assessment based on questionnaires [12,13,14,15]. In this way we are aiming to overcome limitation of every component and develop a self consistent and comprehensive technique for monitoring intra- and inter-personal state of the crew members.

2   Related Work

The first concept of the AMHA project has been done within Mars-500 experiment for 105 days isolation [4, 5, 6, 7]. In 2004 the Institute for Biomedical Problems (IBMP) in Moscow and the European Space Agency have started to plan a full-scale ground based simulation of a manned mission to Mars. Such a full scale mission requires between 520 to 700 days of isolation. Referring to the lower end of this time frame the initiative was named Mars-500.

The goal of the Mars-500 study is to gather data, knowledge and expertise required to prepare a real mission to Mars. Hence, all key peculiarities expected to be present during future missions to Mars are reflected: ultra-long duration, need for autonomy, affected communication due to signal delay, and limited stock of expendables.

This ensures that psychological and physiological impacts of isolation through such an extended period of time are observed as close to reality as possible. A crew of six candidates (four Russians and two from EU countries) are sealed insight the facilities of the Institute for Biomedical Problems in Moscow. An initial 105-day isolation period took a place in spring of 2009. The present concept is going to be tested in the full 520-day study, which has launched at the 3rd of June, 2010.

The Mars-105 experiment presents the crew members encountered in strategic game interactions while data is being gathered about their interpersonal dynamics.

2.1   Evolutionary Game Theory and Colored Trails Game

In this project, we explore the use of strategic multi-player games to alleviate stress, and more importantly as an unobtrusive tool to monitor the mental capacity of astronauts as well as the development of different social interaction patterns within the crew. We are primarily interested in games that feature the following properties:

- Simple enough for analysis
- Rich enough to reflect features of real life interactions
- Grounded in a situated task domain
- Strategic (i.e. partial information that promotes reasoning)
- Suited to measure social factors such as fairness

We are also interesting to produce data, which directly involve interpersonal relation and, as a consequence, can be interesting to monitoring social atmosphere in the crew.

As such a tool, we used a three-player negotiation variation [25] of the Colored Trails framework developed at Havard University [24]. The Colored Trails game is played on a board of colored squares. One square is designed as a "goal square". Each player has a piece on the board and possesses a set of chips in colors chosen from the same palette as the squares of the board. To move his own piece into an adjacent square a player must turn in a chip of the same color as the square. Chips can be exchanged by the players if a mutual agreement is reached. The goal of the game is to move your own piece as close to the goal square as possible using as less chips as possible. Distance to the goal is more important than number of chips left after the move.
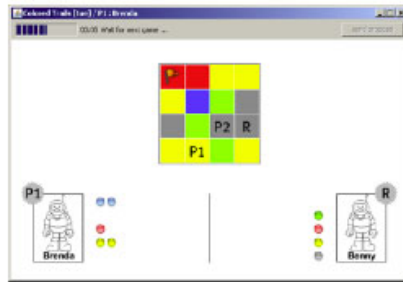

Fig. 1. Colored Trails Game

2.2 Data Analysis of Mars-105 Experiment

Design of the presented monitoring system is based on our previous analysis of the data obtained in the MARS-105 experiment. In particular behavior of Responders was analyzed. Situation, in which a responder needed to make a decision, was described by two numbers: how the Responder can improve his/her payoff by accepting the first offer and how he/she can improve his payoff by accepting the second offer. This description of the situation is a simplification because other factors can influence the decision of the Responder (for example what other players will get as a result of his/her decision or how good are the available offers in comparison with the offers which could be proposed). However, the two selected parameters are the only factors, which are determining behavior of the Responder and showing if the player's behavior is totally rational. So, the mentioned two parameters can be used to check how rational the behavior of Responder is. This check is important because any deviation from the rational behavior can potentially be an indication of psychological preferences or other interpersonal relations.

On the Figure 2 is shown the behavior of one of the Responders. The x- and y-axis correspond to the improvement of the payoff which can be achieved by accepting the first and second offers, respectively. So, generally speaking, every point in the plot represents a particular situation in which the given Responder needed to make a choice. The color of the point indicates the choice which was made. The red/blue colors mean that the first/second offer was accepted. The green color means that the both offers were rejected.

The space on the graph is divided into three different regions. If a responder is totally rational, every region should only contain points of one corresponding color.
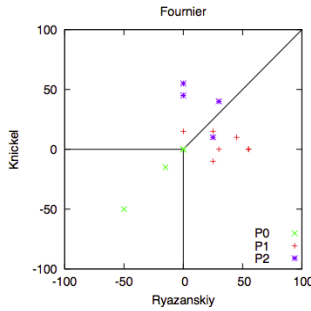
Fig. 2. Dependence of the behavior of one Responder on the accepted offer

For examples, if both offers worsen the payoff of the responder (the left bottom region) than both offers should be rejected (green points). The right-bottom region means that the first offer is better than the second one and it improves the payoff of the responder (so, in this case the first offer should be accepted (red points)). In the third region (left-top) the second offer should be accepted (blue points).

3   Method

3.1   Design Concept

Based on the previous AMHA design, the study of the Colored Trails game performed during MARS-105 [4, 5, 6, 7] and data analysis, presented in previous section and [4], we lay down conditions and requirements to extend the rules of the game, and experiment by itself.

To build the AMHA concept model, firs of all, we defined the parameters, which are crucial for successful communication and collaboration, such as stress level, pleasure, trust, dominance, and etc. Furthermore, we defined a personal mental state and interpersonal relationships as two main components of the crew mental health. The combination of different techniques is required to assess different aspects of the psychosocial states as well as to perform cross validation and correct interpretation of the collected data (see Figure 3).
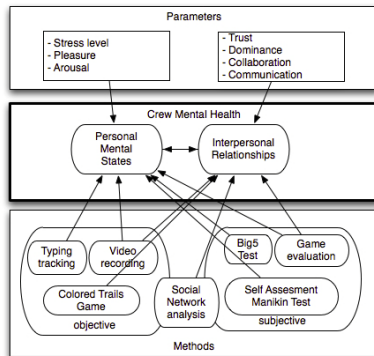

Fig. 3. AMHA conceptual model

Usage of questionnaires for a self assessment is important component of our design, since crew members can provide explicit information about their psychological states and interpersonal relations in the crew. This information is important for interpretation of the game results in terms of perception of psychosocial state of the crew.

3.2 Generalization of the Colored Trails Game

For the MARS-500 experiment we have proposed a generalization of the Colored Trails game. In the new version of the game every user plays as Proposer. Moreover, in contrast to the previous version of the game, Proposers are free to choose a player (Responder) whom he/she wants to offer a chips exchange. As a consequence of these modifications, two or three players will play the Responder role in the second stage of the game.

The game flow has four phases.

1. Choosing a partner
   In this phase each player is choosing a partner for the current session.
2. Proposition phase
   In this phase, each player, acting as a Proposer, can offer a chips exchange to his adversary. The Proposer's and his Adversary chips are always located on the left and right hand side of the window, respectively. On the same screen the player can see the game overview.
3. Response Phase
   In the Response Phase, each player plays as a Responder. After evaluation of the situation he can stay with his own proposal, or accept a proposal from another player, if it is available. If he accepts a proposal from another player, his own proposition is canceled.
4. Scoring
   The final evaluation is fully automatic; players do not need to act in this phase. The game server automatically computes the best possible sequence of moves for each player and assigns personal scores. Points are calculated as follows. For reaching the goal location a player receives 125 points. If he does not reach the goal, 25 penalty points are subtracted for every square between the goal and the player's position. In addition, for every chip the player has not used, he receives 10 extra points.
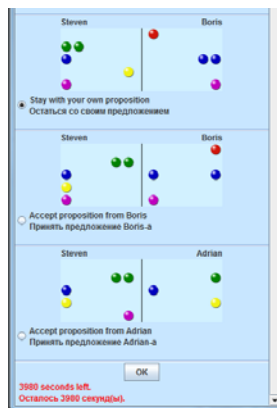

Fig. 4. Screenshot of Colored Trails Game, Response Phase used for 520days isolation

The introduced modifications provide several advantages.

First, we increased number of Proposer per game (from 2 to 3) and, as a consequence, we get more data about the behavior of Proposers.

Second, we increased the number of Responders (from 1 to 2 or 3). This way more information about behavior of Responders can be collected.

Third, we increased the variety of situations in which Responders can be. In particular, Responders have to choose from different number of offers coming from Proposers which are in different situations.

Forth, based on the analysis of the previous experiment, we added a phase at the game, which is aimed to assess irrational preferences of the players. At this stage, when each player is choosing a partner for a current game, player's behavior can not be based on rational thinking about state of the game.

We also designed the Colored Trails Questionnaire that is aimed to define which strategy each player has kept. The players are asked to fill it in after eacj gaming session.

In addition to the previous experiment we combine the usage of the negotiation game with direct monitoring techniques of the nonverbal behavior of the subjects. In that way we want to study if our conclusions about user's persistence of the game and other player can be confirmed by the direct measurement methods. In particular, the monitoring techniques can be used to detect if players are satisfied or not with offers from other players as well as decisions made by the Responders.

3.3 Self Assessment Techniques

Self assessment is an important component of mental health assistant since subjects can explicitly provide needed information about intra- and inter- personal states.

3.3.1 Big 5 Test

One of the intentions of our experiment is to find a relation between observed dynamics of the interpersonal relations with psychological parameters of the crew members. Such a relation could help us to generalize the behavior observed in particular crew and to predict, in this way, relations in groups consisting of members of similar psychological types.

At the baseline of the experiment, we are collecting data, which could give us the psychological characteristics of participants about personality and collaboration patterns at the same time. This data we can be used as a reference point in the future data analysis. For this purpose we are using the Big 5 Test [12].

The Big Five model is considered to be one of the most comprehensive, empirical, data-driven research findings in the history of personality psychology. Over three or four decades of research, these five broad factors were gradually discovered and defined by several independent sets of researchers.

To address the need for a short instrument measuring the prototypical components of the Big Five that are common across investigators, the Big Five Inventory (BFI) has been constructed [10]. The 44-item BFI was developed to represent the prototype definitions developed through expert ratings and subsequent factor analytic verification in observer personality ratings. The goal was to create a brief inventory that would allow efficient and flexible assessment of the five dimensions when there is no need for more differentiated measurement of individual facets.

For our study the results of this test are used to build a Social Network Model, and use it for the further data analysis in terms to compare and find a correlation with the Social Network Model, that we would have from records of the game.

3.3.2   Self Assessment Manikin (SAM) and Cognitive Task Load

To assess the emotional parameters during the game, players will be asked to fill in the questionnaires (see Figure 5).
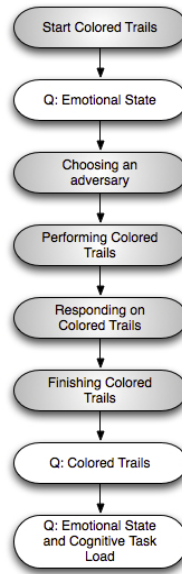

Fig. 5. Colored Trails workflow

For emotional state assessing of the participants, we are using the Self-Assessment Manikin test. The Self Assessment Manikin (SAM) Test was introduced in 1985 by P.J.Lang. It offers the ability to avoid the verbal expression of the emotion assessment, so it establishes a quick and easy to use experimental procedure [13, 14].

Usage of this technique has few advantages. The test is well established, it's aimed for the main and basic parameters, but the results are full enough for analysis. And it's simple and intuitive to perform, which is crucial for our design.

The test assesses the follow emotional states: valence (pleasure), arousal (activity level), dominance (control in the situation).

Cognitive load theory has been designed to provide guidelines intended to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance. At the end of each game, the participants are asked to fill out a Cognitive Task Load (CTL) Questionnaire (apart the SAM Questionnaire).

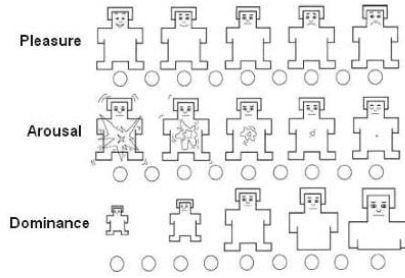Within the CTL we assess: level of information processing, task-set switches and time occupied.

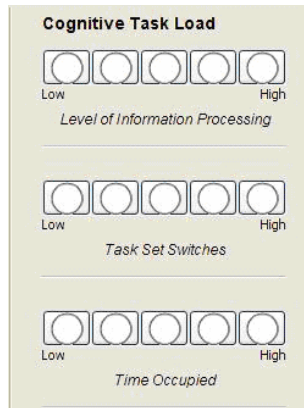**Fig. 6.** SAM Questionnaire

Fig. 7. CTL Questionnaire

Analysis of the SAM before and after each game will show how game could affect an emotional state of the player. Also it would allow us to analyze the correlation between an emotional state of the player and his/her strategies of playing.

3.4   Measuring Nonverbal Behavior

Automatic monitoring of face expression, voice intonation and body movements were proven to be useful techniques for monitoring of emotional states of subjects. These techniques are important since humans naturally express their emotions throw non-verbal behavior.

In our study we use records of facial expressions during the time people are playing the Color Trails Game. Facial expressions give important clues about emotions. In contrast to self assessment techniques, the tracking of nonverbal behavior provides information about emotions and psychological state unconscious. It would allow us to get more objective data compare to results of questionnaires.

Using the software for analysis video (like Visual Recording by University of Amsterdam), we would be able to quantify the 6 main emotions [9].

3.5   Social Network Analysis

Finally, to describe the relation between participants and analyze how the relations are developing through the time, we use Social Network Analysis (SNA). The objective of SNA is to understand the pattern and content of the interactions that take place within and between social units. For this reason, such analysis brings fresh methodological and conceptual power to tests of the contact hypothesis. SNA defines networks as sets of ties linking several individuals. Ties or contacts may be of different kinds, formal or informal, frequent or infrequent, affect-laden or purely utilitarian. Network analysts use the terms "transactional content" or "tie type" to identify the type of exchange or relationship that takes place between actors in a network [15]. SNA approach allow not only to draw a 'map of ties' between structures/ institutions/ societies as macro-level of social reality, but, in addition, to track changes that occur with these 'actors' during social transformations in the contexts of interconnections at the micro-level (meaning connections between individuals, households or organizations).

In this way SNA provides a powerful method for description and representation of results, collected by different kind of measuring techniques, used in the present study.

4   Discussion

4.1   Mars-500 Experiment: 520 Days Isolation

By now, we combine two approaches to measure mental states of the isolated groups of people in attempt to predict arising conflicts in the group. The first is based on analysis of the subjectivity during play of a negotiation game which is verified with the second approach that relies on direct measurements of the nonverbal behavior of the participants. Our initial analysis of the game subjectivity outcomes were based on the 105 days trial. The analysis showed that there were not significant deviations from the logical choices of the proposers or the responders. This may mean that not extreme situations or noticeable changes in the emotional state of the players took place during the 105 days isolation. However, the reason could be that the negotiation game outcomes by itself are not informative about the emerging conflict situations. The highly educated and performance oriented astronauts trainees may be able to suppress their emotional tension during the task of a game play.

We expect that during the longer isolation period the relations between crew members can become more intense. Related experiments have shown that typing intervals, facial expressions and gestures give clearer indication about the emotional tensions, so we can correlate the game outcomes results with these measurements. In connection with the Mission Execution Crew Assistant (MECA) by TNO [19, 20], we have an unique possibility to test our design within MARS- 500 experiment. MARS-500 provides a unique test platform, because of its setting in which a small crew is isolated for a long duration to simulate a manned Mars mission. In this setting, more prolonged or repeated usage of MECA/AMHA can be tested. A major focus of MECA/AMHA is to support team resilience by monitoring team member's performance (i.e., the effectiveness and efficiency of operations and the related refreshment trainings for nominal and off-nominal situations) and corresponding condition (i.e., the appropriateness of his or her cognitive and affective responses). In the MARS-500

program, we will select a small set of core elements of MECA/AMHA that need this type of prolonged evaluation: a diagnostic method to measure psychosocial crew status, a feedback mechanism, and support for crew planning. The objective of this evaluation is to improve the requirements baseline and its design rationale for these elements, and to refine the corresponding models and methods. The setting is as follows:

- Two groups of three astronauts who train and game once a week for half an hour (incl. procedure training and an entertainment game).
- The astronauts communicate via chat.
- MECA/AMHA collects information on crew condition (social network, emotional state) and performance (effectiveness and efficiency of operations during the training and gaming).
- MECA/AMHA provides (simple) feedback on crew condition and performance.

4.2 Web-Based Experiment

To get quantitative data, we are running an additional web-based study. For this purpose the web-site with Colored Trails Game has been developed.
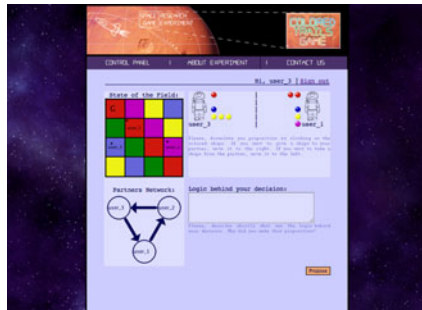

Fig. 8. Screenshot of the Colored Trails Game web site

The research questions for this experiment, apart to get quantitative data, are:

- how the CT game can develop the patterns of existing relationships between people;
- how the CT game can create the relationships between people;
- how the CT game can monitor the behavior.

In these settings we are planning:

- 5-6 groups (by 3 person each) people who know each other
- 5-6 groups (by 3 person each) people who don't know each other
- About 15-20 groups of three people, which are know each other well and/or working together.
- Three sessions per week.
- The duration of the experiment is 15-18 sessions (5-6 weeks).

By the end of the experimental part of the project, we would be able to do data analysis. Two different experimental settings provide an opportunity for cross validation

and answer the question of transferability of results obtained with a single experiment with fixed constrains.

References

1. Kanas, N., Salnitsky, V., Grund, E., Gushin, V.: Social and cultural issues during Shuttle/Mir space missions. Acta Astronautica 47(2-9), 647–655 (2000)
2. Manzey, D.: Human missions to Mars: new psychological challenges and research issues. Acta Astronautica 55, 781–890 (2004)
3. Gushin, V.: Psychological countermeasures during space missions: russian experiences. Journal of Gravitational Physiology 9(1), 311–312 (2002)
4. Hennes, D., Tuyls, K., Neerincx, M.A., Rauterberg, G.W.M.: Micro-Scale Social Network Analysis for Ultra-Long Space Flights. In: IJCAI-09/SMC-IT-09/IWPSS-09 Workshop on Artificial Intelligence in Space (2009)
5. Hennes, D., Tuyls, K., Rauterberg, G.W.M.: State-Coupled Replicator Dynamics. In: 8th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2009), pp. 789–796 (2009)
6. Hennes, D., Tuyls, K., Rauterberg, G.W.M.: Formalizing Multi-state Learning Dynamics. In: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (IAT 2008), vol. 2, pp. 266–272 (2008)
7. Rauterberg, G.W.M., Neerincx, M.A., Tuyls, K., Loop, J.: Entertainment computing in the Orbit. New Frontiers for Entertainment Computing 279, 59–70 (2008)
8. Gaal, Y., Grosz, B.J.: Colored Trails: A Formalism for Investigating Decision-making in Strategic Environments (2005)
9. Valenti, R., Sebe, N., Gever, T.: Facial Expression Recognition: a Fully integrated approach (2007)
10. John, O.P., Srivastava, S.: The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives (2005)
11. Camerer, C.F.: Behavioral Game Theory: Thinking, Learning, and Teaching (2001)
12. Harvey, R.J., Murry, W.D., Markham, S.E.: A ”Big Five” Scoring System for the Myers-Briggs Type Indicator (1995)
13. Bradley, M.M., Lang, P.J.: Measuring Emotions: the Self-Assesment Manikin and the Semantic Differential. J. Behav. Ther. and Exp. Psychiat. 25(I), 49–59 (1994)
14. Tsonos, D., Kouroupetroglou, G.: A Methodology for the Extraction of Reader's Emotional State Triggered from Text Typography. Tools in Artificial Intelligence, 139–154 (2008)
15. Nelson, E.: The Strength of Strong Ties: Social Networks and Intergroup Conflict in Organizations. The Academy of Management Journal 32(2), 377–401 (1989)

16. Busso, C., Deng*, Z., Yildirim, S., Bulut, M., Lee, C.M., Kazemzadeh, A., Lee, S., Neumann*, U., Narayanan, S.: Analysis of Emotion Recognition using Facial Expressions. Speech and Multimodal Information (2004)
17. Mase, K.: Recognition of facial expression from optical flow. IEICE Transc., E. 74(10), 3474–3483 (1991)
18. Yacoob, Y., Davis, L.: Computing spatio-temporal representations of human faces. In: Proceedings of IEEE Computer Society Conference, Computer Vision and Pattern Recognition CVPR 1994, pp. 70–75 (1994)
19. Neerincx, M.A., Lindenberg, J., Smets, N.: The Mission Execution Crew Assistant: Improving Human-Machine Team Resilience for Long Duration Missions (1998)
20. Breebaart, L., Bos, A., Grant, T., Soler, A.O., Neerincx, M.A., Lindenberg, J., Smets, N., Brauer, U., Wolff, M.: The MECA Project — Ontology-based Data Portability for Space Mission Operations (2008)
21. Barakova, E., Lourens, T.: Expressing and interpreting emotional movements in social games with robots. In: Personal and Ubiquitous Computing (2010)
22. Lourens, T., Barakova, E.: Retrieving Emotion from Motion Analysis: In a Real Time Parallel Framework for Robots. In: Leung, C.S., Lee, M., Chan, J.H. (eds.) ICONIP 2009, Part II. LNCS, vol. 5864, pp. 430–438. Springer, Heidelberg (2009)
23. Lourens, T., Barakova, E.: My Sparring Partner is a Humanoid Robot -A parallel framework for improving social skills by imitation. In: Mira, J., Ferrández, J.M., Álvarez, J.R., de la Paz, F., Toledo, F.J. (eds.) IWINAC 2009. LNCS, vol. 5602, pp. 344–352. Springer, Heidelberg (2009)
24. Gal, Y., Grosz, B.J., Kraus, S., Pfeffer, A., Shieber, S.: Colored trails: a formalism for investigating decision-making in strategic environments. In: Proc. of the IJCAI Workshop on Reasoning, Representation, and Learning in Computer Games (2005)
25. Ficici, S., Pfeffer, A.: Modeling how humans reason about others with partial information (2008)