Experiments in Planning and the “ISLAND MISSION” Game

Dr. Anthony H. Dekker
Defence Science and Technology Organisation, Canberra, Australia
dekker@ACM.org

Abstract. Today’s complex military endeavours require collaborative planning in a multi-agency context. Collaborative games are useful tools for experimenting with such planning. The “ISLAND MISSION” game, previously presented at SimTecT 2007 as a board game, is one such tool. “ISLAND MISSION” is based on a peace enforcement/humanitarian relief scenario, and players take the role of representatives of Army, Navy, Air Force, medical, media, and other agencies, each of which has goals that are partially shared and partially in conflict. To succeed, participants must negotiate to produce a plan satisfactory to all parties, just as must be done in real-world multi-agency operations. “ISLAND MISSION” has now been implemented as an RMI-based software tool for experiments with human beings, and as an intelligent agent system for *in silico* experiments. This paper briefly describes the implementation, and gives the results of initial experiments.

1. INTRODUCTION

Military operations are activities which mix competition with cooperation. From antiquity, competitive games such as Chess and Go have provided an abstraction of warfare [1].

In the modern era, cooperative aspects of military operations have become particularly important, with multiple people, services, agencies, and countries collaborating to achieve common goals. The individual people, services, agencies, and countries involved may also have goals of their own, which may differ from those of other actors. This presents a potential challenge when the actors are expected to self-synchronise [2],[3] to achieve coordinated actions.

Team games such as SCUDHunt [4],[5],[6],[7], Go*Team [8], and CAFFEINE [9] are useful tools for exploring team phenomena such as information sharing, collaboration and self-synchronisation.

Building on previous experience with SCUDHunt [10], we have designed a team game called “ISLAND MISSION” [11]. Initially a purely paper-based game, it has now been implemented in software, and this paper describes the recent implementation of the game.

A client/server architecture based on Java and RMI permits multi-player operation. Since the client software has a web-deployable applet version, players may participate from different locations around the world.

We have also developed automated intelligent agents for playing the game. These agents can be mixed with human players, in order to fill out the required number of players. In addition, the automated agents can be used for preliminary testing of experimental designs, and for *in silico* investigation of automated planning techniques.

The second half of this paper presents some initial results of experiments with the implemented game, and the implications of those results for conducting future experiments. Some minor improvements to the game design are also recommended.

2. THE “ISLAND MISSION” SCENARIO

The game board for ISLAND MISSION, shown in Figure 1, represents a small island nation. The game is based on a multi-agency peace enforcement and humanitarian relief scenario, involving several conflicting priorities:

- The island is occupied by rebel forces, whose locations are initially uncertain.
- There is a hurricane warning for the East Lowlands, which is expected to result in significant damage and casualties.
- To speed economic recovery, it is important to reopen the mine in the Hill Country and the route to the nearby port as soon as possible.
- It is necessary to stop smuggling of drugs by boat from the forest villages numbered 20 and 23, either using land forces, or offshore naval blockade.
- The central mountain stronghold is inaccessible to friendly forces, and any rebels there would need to be contained.
- Navy, Air Force, land combat, engineering, and medical forces are available.
- The desired end state is that no rebels remain (except possibly contained in the mountain stronghold), all casualties have been treated by medical units, and all damage has been repaired by engineering units.
- The operation will be conducted under close media scrutiny, and media reports of actions will have a significant (positive or negative) impact on mission success.
- A rapid resolution of these issues is important, and game scoring includes a penalty for slow completion, and for delays in dealing with medical cases.

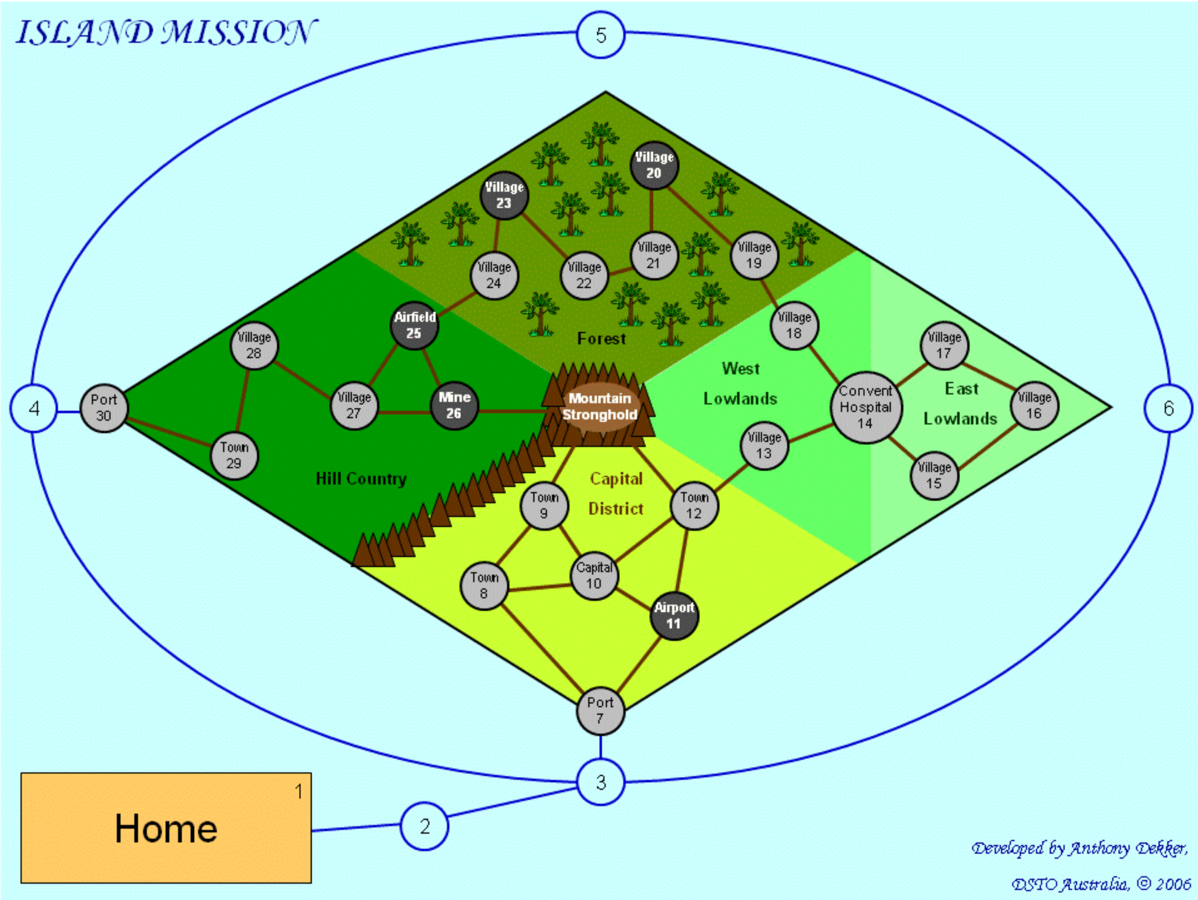

Figure 1: Game board layout for ISLAND MISSION

3. GAME IMPLEMENTATION

The ISLAND MISSION game can be played with either 4 or 7 players, but testing has been mainly of the 4-player version. With 4 players, the roles represented are:

- Naval/Air task force, responsible for:
  - One "ship" token representing a naval force, which can carry out transport, blockade, or naval gunfire actions.
  - Two "air" tokens, representing Air Force units, which can carry out transport, surveillance, or air strike actions.
- Army task force, responsible for:
  - Three "combat" tokens, which can carry out combat or protection actions.
  - One unarmed "engineer" token, which can carry out damage repair.
- Medical task force, responsible for:
  - Two unarmed "medical" tokens, which can treat casualties.
- Media, responsible for:
  - One "media" token, which moves around the island, and reports on actions at its location.

The 7-player version splits some of these roles, such as Navy and Air, as well as adding some new ones, such as coalition partners and non-military medical.

As presented at SimTecT 2007 [11], game play used the cards and tokens shown in Figure 2, but the game has now been implemented in Java software.

Individual players can receive or lose points for actions, and there are also shared points. Achieving the end state for the game takes several rounds, and each player's final score is obtained by adding his/her individual points to the total shared points. The overall team score is the average of these final player scores.

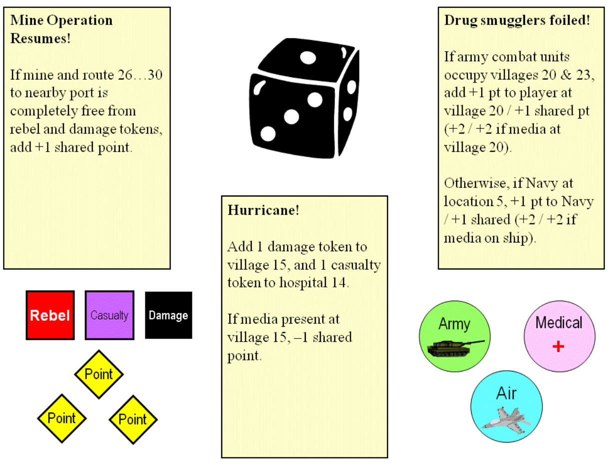

Figure 2: Sample event cards, tokens, and die for paper version of the ISLAND MISSION game

Figure 3 illustrates the solitaire (Java applet) version of the game, where a single player plays all 4 game roles. This version is intended mostly for training purposes, and the ordinary multi-player version of the game uses a Java RMI client/server architecture. The clients in this framework can be either Java applets or standalone Java programs, but have the same look and feel as Figure 3.

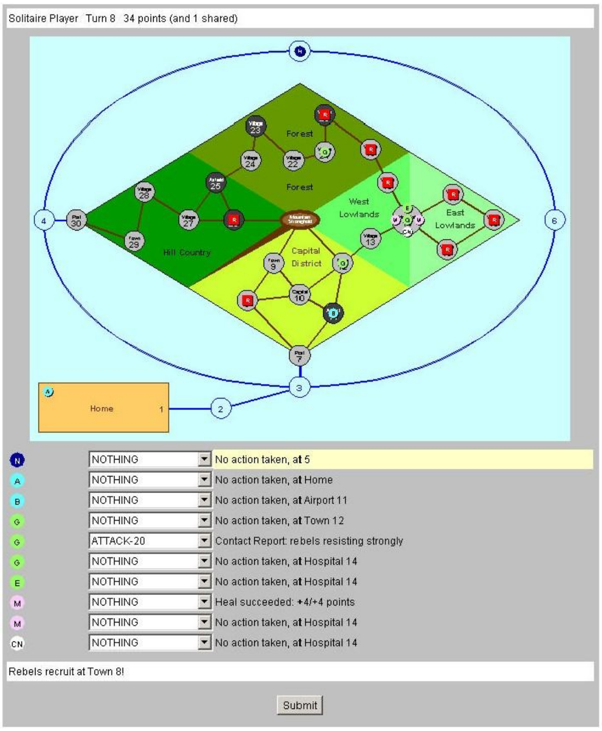

Figure 3: Snapshot of the ISLAND MISSION software

The ISLAND MISSION client program or applet displays a map of the island, showing the location of all units, and the known locations of rebels (red squares in Figure 3). The location of rebels in unexplored areas remains unknown until friendly forces first arrive in those areas.

Below the map display are token symbols for the entities controlled by the player (for the solitaire version, this is all 10 entities). For each entity, a choice of moves is given. The default choice is the previous move, if that is still a valid option, or else a do-nothing move. The player selects choices for all the entities under his or her control, and then clicks the "Submit" button. Beside the list, the result of each move is shown, with the result of moves in the previous turn displayed on a grey background, and the result of moves in the current turn displayed on a green background. A yellow background marks the next entity to be moved.

In Figure 3, the human player is 8 turns into the game. The whole island has been explored, the immediate medical crisis has been handled, and a naval blockade is preventing drug smuggling. However, eight rebel units are still active on the island, including one recently recruited rebel unit. Rebel activity is preventing operation of the mine, and is also likely to interfere with responses to the hurricanes expected in the East Lowlands. Nevertheless, the player is performing quite well at this stage, and has accumulated a healthy score.

4. GAME PLAY BY A HUMAN

Table 1 summarises the author's experiences playing 20 games of the 4-player version, with the author playing all four roles. The average score obtained ranged from 18.45 for the media role to 30.35 for the Army/Engineering role, and this provides an indication of the score which can be achieved by moderately good play.

Table 1: Author's experiences over 20 games

Role	Average Score	Range
Air/Navy	20.65	2…34
Army/Engineering	30.35	7…52
Medical	25.85	13…45
Media	18.45	–3…33
Average	**23.8**	

The differences between scores for the different roles are statistically significant (at the 0.2% level, by analysis of variance). In other words, some roles inherently obtain somewhat higher scores than others. This is a potential (and deliberately designed) source of conflict within a team setting, particularly since there are some game options which can increase individual players scores, while significantly decreasing the overall team score. One of the challenges for teams of human players is to avoid the temptation of this kind of non-cooperative behaviour.

Since the game was designed to support experiments on team behaviour, it has been designed so that overall team success requires collaboration between participants. Collaborative actions within the game include:

- air and naval assets providing transport to other players;
- army units protecting unarmed engineer or medical assets belonging to another player (by sharing the same location, or by containing rebels in the mountain stronghold); and
- army units clearing the way for engineer or medical assets belonging to another player (by removing rebels located between them and the casualties or damage they wish to respond to).

We have conducted some preliminary trials of the multi-player version of the game, and a number of formal team experiments are planned. One point which we wish to investigate is whether military experience will be reflected in higher players scores, as was the case with SCUDHunt [10]. We also hope to test the effect of different communication technologies, as was also done with SCUDHunt [5],[6],[7]. For example, do player's scores differ if they communicate by voice or by chat? Or does the shared map display communicate sufficient information that no additional discussion is necessary?

We would also like to test the effect of restricting membership of the team which plans the game strategy, prior to beginning actual game moves. A significant factor in modern multi-agency operations is the degree to which coalition partners and non-military agencies are considered "part of the team" and involved in planning. ISLAND MISSION allows for non-military players (medical and media players, as well as for coalition partners (in the 7-player version). We plan to conduct experiments comparing the effect of including or excluding those players from initial team planning activities. Even more realism would be added if the players were selected from different nations, perhaps using volunteers from the student body of a military training institute.

5. AUTOMATED AGENTS FOR THE GAME

We have implemented three versions of automated agents for playing ISLAND MISSION. These automated agents can be mixed with human agents, and the control panel for the game server allows the selection of which roles will be played by automated agents, and which by human beings. This option allows experiments to be conducted with 1, 2, 3, 5, and 6 human participants, by automating the other roles in a team of 4 or 7.

The simplest agents use fixed tactics developed by the author. These are encoded in about 60 lines of software for each of the 6 different kinds of game entity (air, naval, army, engineering, medical, and media). These tactics allow the entities to respond to the current situation, but with no forward planning or active cooperation. Moreover, these agents can "dither," wasting time by repeatedly beginning a movement across the island to deal with a crisis, and then returning to another crisis closer at hand.

The second category of agent uses planning [13], within the Belief-Desire-Intention (BDI) agent model [14]. Planning agents have **beliefs** (corresponding to their picture of the island, and their picture of other agent's intentions), **desires** (corresponding to individual and shared goals), and **intentions** (planned sequences of actions produced by the agent). The agents have a simple brute-force planner which looks 8 moves ahead, and attempts to maximise the overall team score. The planner is limited in that it does not have an accurate model of how the rebels move, and is not able to coordinate activities with other players. Agents commit to their plan, which consists of up to 8 moves in sequence, until the situation on the island changes to such an extent that these moves are no longer valid. This commitment to a plan avoids "dithering," but this can also mean that agents are slow to respond to changing circumstances.

The third category of automated agent combines planning with communication of plans between agents. This means that, during the planning process, when agents calculate the impact of the moves they make, the actions of other agents are also taken into account. For example, medical units can take into account the fact that one disease outbreak is already being handled by another agent, and concentrate their attention on a second one. Similarly, army agents will prioritise defeating rebels which are preventing medical entities from responding to disease outbreaks.

6. GAME PLAY BY AUTOMATED AGENTS

Automated agents permit a number of experiments comparing performance against each other and against human beings, in a similar way to work with the SCUDHunt game [12]. The results of such experiments can also be used to fine-tune the game design.

The 20 games played by the author were complemented by 200 games played by each of the three kinds of automated agent. Table 2 shows the average scores for the different (automated and human) players. Unfortunately, the substantial level of variation in scores means that these differences are not statistically significant.

Table 2: Average player scores for human and automated players (differences are not statistically significant)

Player	Games	Average Score	Std Dev
Author	20	23.8	8.0
Fixed Tactics	200	16.1	24.8
Planning	200	13.9	33.2
Coordination	200	18.9	18.8
Average		**16.5**	

What is the explanation for the high level of variation in game scores? It is associated with variation in the speed with which agents achieve their goals. Table 3 summarises the average number of turns required by different players in order to finish the game. The planning agents without coordination are slower to complete the game than the other two kinds of automated agent and the author (these differences are statistically significant at the 1% level). The overall average time taken to complete the game is 34.4 turns.

Table 3: Duration of games for human and automated players (uncoordinated planning is slower than the other three kinds of player)

Player	Games	Average Turns	Std Dev
Author	20	29.5	9.6
Fixed Tactics	200	32.0	13.6
Planning	200	38.6	14.9
Coordination	200	33.0	12.1
Average		**34.4**	

A high variation in scores is associated with lengthier games. A longer game can give players many opportunities to shine, if they are fortunate, but can also reflect a player who takes so long to deal with one crisis that he, she, or it is repeatedly overtaken by other crises.

Figure 4 shows a scatter-plot of scores against turns required to complete the game. This graph shows that, beyond the average of 34 turns, the variation in player scores increases significantly. On average, scores decrease with number of turns, although the effect is slight.

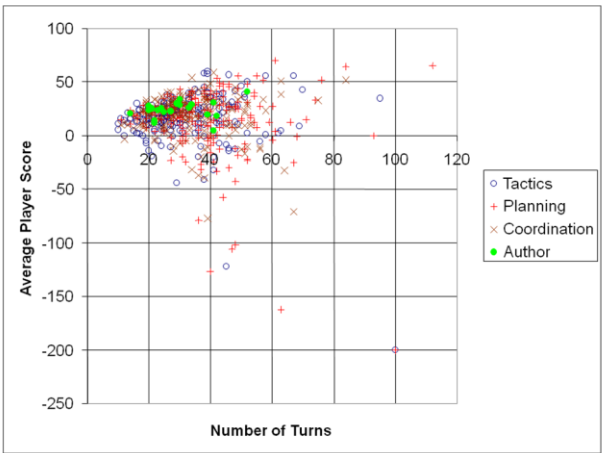

Figure 4: Scatter-plot of scores against turns required to complete the game

Given the high variability in scores beyond 34 turns, it is instructive to look at the average scores for games completed in at most 34 turns. Table 4 shows these averages. We can assert (with statistical significance at the 2% level), that the fixed tactics perform less well (an average of 16.6) than coordinated planning (21.0) or the author (24.2). However, the status of uncoordinated planning (19.0) is unclear.

Table 4: Average player scores for games ending in at most 34 turns (fixed tactics perform less well than coordinated planning or the author)

Player	Games	Average Score	Std Dev
Author	15	24.2	5.7
Fixed Tactics	125	16.6	12.5
Planning	88	19.0	13.4
Coordination	127	21.0	12.5
Average		**19.1**	

A better way of adjusting for the variability in scores with game length is to divide the scores by the number of turns taken. Table 5 shows the average of these quotients. We can again assert (with statistical significance at the 4% level), that the fixed tactics perform less well than the author.

We can also assert (with statistical significance at the 2% level) that uncoordinated planning performs less well than the author, and that (with statistical significance at the 0.5% level), coordinated planning performs better than uncoordinated planning. The other three differences are not statistically significant.

Table 5: Average player scores divided by number of turns (fixed tactics and uncoordinated planning perform less well than the author, uncoordinated planning performs less well than coordinated planning)

Player	Games	Average Score/Turns	Std Dev
Author	20	0.88	0.33
Fixed Tactics	200	0.58	0.61
Planning	200	0.46	0.75
Coordination	200	0.65	0.56
Average		**0.576**	

7. DISCUSSION

We can draw four general conclusions from these experimental results:

- There is a clear advantage for the planning agents which coordinate, compared to those which do not. Consequently, the game succeeds as a way of testing whether participants (agent or human) are coordinating adequately. Effective coordination will be reflected in higher team scores, and so the game can be used to test for effective teamwork.
- The planning agents with coordination perform almost as well as a good human player. Consequently, the design of experiments with human beings can be first tested with such agents in order to identify potential problems. Experiments with human beings can be costly, both in terms of staff time and the goodwill of potential participants, and it is very helpful to have a method for pre-testing potential experiments.
- Although the fixed tactics perform less well than the author, their performance is still reasonably good (an average score of 16.1 over all games). Consequently, the simple non-planning agents are adequate for automating "missing roles" in an experiment with human beings.
- There is considerable random variation in player scores, due to the random nature of events in the game. This poses a problem in experiments with human beings – obtaining statistically significant experimental effects will be difficult.

8. IMPLICATIONS FOR GAME DESIGN

Successful use of the "ISLAND MISSION" game for human experimentation requires a reduction in the random elements of the game. The elements of particular concern include:

- the initial distribution of rebels, which results from (effectively) throwing a die for each possible location, and placing a rebel token when a 1, 2, or 3 is thrown;
- the random nature of medical action, which succeeds when a 4, 5, or 6 is thrown on a die;
- the random nature of ground combat, where the outcome is also decided on the throw of a die.

We have now modified the game to reduce these random elements. First, we have replaced the die by a process equivalent to having a stack of cards marked with the numbers 1…6. Replacing the die in this way ensures a fairer distribution of "good" and "bad" outcomes in any particular game.

We have also removed the random elements from medical action, by reprogramming the software so that every medical intervention succeeds the second time it is performed. This ensures that the average time taken to deal with medical cases does not change.

Table 6 shows the equivalent of Table 2 for the modified game. The game modifications have increased the scores, but decreased the standard deviations. The superiority of coordinated planning over uncoordinated planning is now significant at the 2% level, and the superiority of fixed tactics over uncoordinated planning is now significant at the 10% level. The modifications have therefore reduced variability sufficiently to improve statistical significance.

Table 6: Average player scores for modified game with automated players (uncoordinated planning performs less well than the other two strategies)

Player	Games	Average Score	Std Dev
Fixed Tactics	200	20.3	15.3
Planning	200	16.2	31.5
Coordination	200	22.6	16.0
Average		**19.7**	

9. FINAL REMARKS

In this paper, we have presented the software implementation of the "ISLAND MISSION" team game, together with some preliminary experimental data. The ISLAND MISSION game is based on an abstraction of real multi-agency military operations, and it allows the exploration of joint, coalition, and multi-agency issues, since players need to balance individual and team goals. The software implementation is designed to be web-deployable.

The ISLAND MISSION game can be played by a single human being, by teams of human beings, by teams of intelligent agents, or by a mixed team of human beings and agents. The experimental data demonstrates that the intelligent agents perform sufficiently well to be mixed with human beings, and for experimental designs to be tested with intelligent agents.

The experimental data also demonstrates that the game successfully tests for coordination between participants, and hence is suitable for its intended purpose.

The large amount of random variation in the results is a problem for future experiments with human beings, and we have identified a number of changes which should help to resolve this problem.

In future work, we intend to conduct experiments with teams of military personnel playing the modified game.

REFERENCES

1. Leigh, E. (2006) "21st Century Uses of Simulation: Three Basic Key Influences," *SimTecT Simulation Technology and Training Conference*.

2. Alberts, D.; Garstka, J. & Stein, F. (1999) *Network Centric Warfare: Developing and Leveraging Information Superiority*, CCRP Press: Washington. www.dodccrp.org/publications/pdf/Alberts_NCW.pdf

3. van Bezooije, B.; Essens, P. & Vogelaar, A. (2006) "Military Self-Synchronization: An Exploration of the Concept," *Proc. 11th International Command and Control Research and Technology Symposium*, CCRP Press: Washington, DC.

4. ThoughtLink Inc (2003) SCUDHunt website www.scudhunt.com updated 23 March 2003.

5. Perla, P.; Markowitz, M.; Nofi, A.; Weuve, C.; Loughran, J. & Stahl, M. (2000) *Gaming and Shared Situation Awareness*, Center for Naval Analyses. www.cna.org/documents/D0002722.A2.pdf

6. Loughran, J.; Stahl, M. & Perla, P. (2001) *Key Drivers for C2 Performance: Data Mining SCUDHunt Experiment Data*, ThoughtLink Inc. Available at www.thoughtlink.com/publications/TLI-SCUDHuntAnalysis01/TLI-SCUDHuntAnalysis01.zip

7. Stahl, M. & Loughran, J. (2002) *Exploring Joint Force Command and Control Concepts Using SCUDHunt – Final Report*, ThoughtLink Inc. www.thoughtlink.com/publications/SCUDHuntFinalWeb.zip

8. Hart, D.; Eronen, M.; Jagiello, J. & Warne, L. (2006) "Go*Team: A Team Version of Go for Investigating Cooperation, Coordination and Information Sharing in Network-Centric Warfare Environments," *SimTecT 2006 Simulation Technology and Training Conference*.

9. Huber, R.K.; Eggenhofer, P.; Römer, J.; Schäfer, S. & Titze, K. (2006) "Developing an analytical framework for cognitive and social experimentation in team decision-making and collaboration," *Proc. 11th International Command and Control Research and Technology Symposium*, CCRP Press: Washington, DC.

10. Dekker, A.H. (2006) "Revisiting 'SCUDHunt' and the Human Dimension of NCW," *Proc. 11th International Command and Control Research and Technology Symposium*, CCRP Press: Washington, DC.

11. Dekker, A.H. (2007), "Studying Multi-Agency Planning Using Team Games, " *SimTecT 2007 Simulation Technology and Training Conference*.

12. Perla, P. & Loughran, J. (2003) "Using Gaming and Agent Technology to Explore C2," *Proc. 8th International Command and Control Research and Technology Symposium*, CCRP Press: Washington. www.dodccrp.org/events/8th_ICCRTS/pdf/077.pdf

13. Dekker, A.H. & de Silva, L. (2006) "Investigating Organisational Structures with Networks of Planning Agents," *Proc. International Conference on Intelligent Agents, Web Technologies and Internet Commerce*. www.acm.org/~dekker/iawtic.pdf

14. Rao, A.S. & Georgeff, M.P. (1991) "Modeling rational agents within a BDI-architecture." *Proc. 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pp. 473–484.