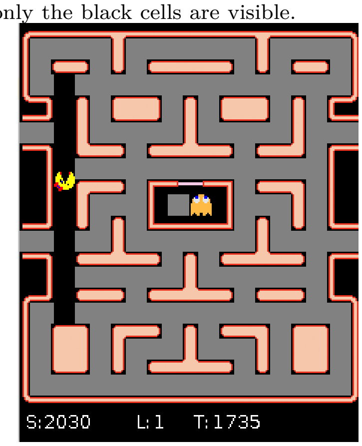
The Effect of Varying Partial Observability in Ms. Pac-Man

Piers R. Williams, Simon M. Lucas Senior Member, IEEE, and Michael Fairbank

Abstract—This paper investigates the effect on difficulty and enjoyment of varying amounts of Partial Observability within a ghost-versus-pacman game. We perform two main experiments, one with Artificial Intelligence agents to investigate difficulty, and the second with human participants to investigate enjoyment. We find that as sight limit increases in one type of Partial Observability, the Ghosts benefit more than Ms. Pac-Man; due to their ability to work as a team and collectively see more of the board. We also find unusual results from human participants that contradict Artificial Intelligence tests. Finally we conclude that although Artificial Intelligence experiments may sometimes give a rough proxy for the level of enjoyment and difficulty a human might find in a game, they cannot be relied upon in all cases to be a reliable link, and some counterexamples do exist.

Index Terms—Partial Observability, Artificial Intelligence, Game Design, Ms. Pac-Man

Figure 1. Partial observability in the Ms. Pac-Man Vs Ghost Team Competition: only the black cells are visible.



I. Introduction

The research question this paper investigates is how does varying the amount of Partial Observability (PO) in a game affect fun and difficulty. To do this, we experiment with Artificial Intelligence (AI) assisted game design using both AI agents and human players in the Ms. Pac-Man Vs Ghost Team Competition [1]. This competition was started to promote high quality research into AI in a co-operative PO environment. Competitors could submit either a Ms. Pac-Man agent or a Ghost Team of four agents. These agents were then played against each other to form rankings. The legacy of the competition is that it has provided a number of different ready-made AI agents that have already been evaluated in a competitive setting. This made the work of this paper, and other meta-studies in future, possible.

Ms. Pac-Man is a widely known and internationally successful arcade game from 1982 that has also had a large amount of research interest from the game AI community [2]. The addition of PO into Ms. Pac-Man has been done before as an AI challenge in [3] as well as in the competition framework used in this paper. Figure 1 shows the PO restriction used in the competition.

PO is the impairment of the ability of an agent to completely observe its environment. The original Ms. Pac-Man game is fully observable, since both the human player controlling Ms. Pac-Man and the ghost AI, can see the entire board. PO is implemented in the Ms. Pac-Man competition by withholding information from the players, i.e. agents do not always receive details of the locations of the other agents. In order to allow for the ghosts to operate as a team, the competition framework features a communication Application Programming Interface (API) that allows the ghosts to pass messages to each other at a restricted rate.

PO can add to the excitement in a game, with horror games typically using PO to build suspense and shock the player with sudden appearances. The use of PO could potentially induce emotional responses such as anxiety, fear, claustrophobia, or frustration during play. Getting the balance wrong and hiding too much means the player cannot make good decisions and loses agency their effect in the environment. Showing too much to the player and you miss out on some of the suspense and drama that could make a good game great.

On the other hand, the use of PO in games is often implemented in a predictable way in different game genres (Table I). This clichéd usage reduces the impact of the mechanic as it is often not an integral part of the game, and is potentially detrimental to the games industry. New and interesting mechanics are needed to stop games converging.

Symmetric PO is defined to be where all objects in the game have identical abilities to observe the environment as each other. Some First-Person Shooter (FPS) games enable asymmetric PO, for example by allowing players to choose various perks or upgrades that can alter visibility. Being able to “sneak” removes certain players from the minimap, whilst other perks allow temporary x-ray vision to add additional observations.

In Ms. Pac-Man, the Ghosts outnumber Ms. Pac-Man four to one. This could make a case for not allowing the Ghosts to see as much as Ms. Pac-Man in order to

Table I
Examples of PO in games.

First-Person Shooter	The field of view presented to the player naturally restricts sight. Often a shorter range “minimap” is provided with restrictions that can vary from game to game.
Real-Time Strategy	The game often calculates what all of the player’s units can see, and then obscures the rest.
Platform Games	The game often only presents a viewport with the character within it, and progression through the game moves the viewport.
Role-Playing Game	The field of view is often restricted to whatever the character could potentially see.
Horror	The view is often dark and poorly lit. Enemies will purposefully hide from the player until ready and combined with ominous sound effects more can be said with the unseen than the seen.

balance the game. This could be tweaked with the use of survivability analysis, followed by testing good candidate settings on human participants. For simplicity, the Ms. Pac-Man competition engine currently enforces identical restrictions on all AI agents.

Communication can be a vital tool for agents working in a team, especially if those agents are within a PO environment. Agents have to make decisions from a number of variables, and when working in a team some of those variables rely on predictions of what co-workers will do in the future. In a Complete Observability (CO) environment and working with purely reactive deterministic agents, it is possible to perfectly predict their actions. With this knowledge, an agent can then plan its own actions to best achieve the team goal. Non-determinism in co-workers can often be reasoned about with algorithms such as MCTS [4]. Non-reactive agents are harder still to co-operate with, and communication can be a solution to allow co-ordination between such agents.

When moving to a PO environment, everything becomes more difficult. When the local environment is unobservable, it becomes impossible to predict the actions even of agents which are purely reactive (i.e. have no internal state) and deterministic. In this case, information is key, and communication allows the sharing of each agent’s local environment with their co-workers. This removes most of the difficulties, although there may still be the problem of communication delays.

When an agent is non-deterministic, or when it has capabilities beyond being purely reactive (for example, when it has some form of hidden internal state upon which to base its decisions), it becomes even more difficult, but not impossible, to co-operate with. Williams et al [4] found that in a PO co-operative environment, MCTS could sometimes overcome these challenges by making random assumptions to reason against the worst-case co-worker strategy. Any improvement over this assumption in MCTS’ actual co-worker’s strategy simply resulted in better than expected results.

Creating AI agents for PO environments can often be more difficult than the same environment without restrictions. Hidden information makes various AI tasks harder such as creating accurate heuristics and forming accurate forward models for tree search methods. There has been a large amount of previous research into AI agents for PO, including PO Pac-Man [3], PO games such as poker [5], as well as co-operative PO games such as Hanabi [6].

Computer games are fundamentally games, and many techniques for designing games are equally applicable to computer games such as Ms. Pac-Man. Hunicke et al [7] define the Mechanics, Dynamics, and Aesthetics (MDA) framework. By their definition the PO restriction added into Ms. Pac-Man is a mechanic, as the restriction is a fundamental part of the control mechanism.

Koster [8] states that fun comes from the pursuit of mastery of a game. If this is true then potentially a more difficult game would be more difficult to master, but provide a longer source of fun as the player tries to master it. Koster also makes the observation that games are linked, often only changed by a single element. This is the process used to create PO Ms. Pac-Man - the same game but with the single element of visibility altered.

Nelson and Mateas [9] describe a method to formalize game mechanics as well as how to generate games that utilize those mechanics. The authors make use of both WordNet and ConceptNet to reason about which verbs and nouns make sense together. A working example is provided that produces games in the style of Nintendo’s WarioWare series.

Browne and Maire [10] create the general game system Ludi to generate and test new games. The authors are succesfull in creating intersting and publishable new games through the system.

Isaksen et al [11], [12], [13] give a method that explores the game space of the popular mobile game Flappy Bird. By varying multiple parameters such as jump height, tube spacing, and bird speed they locate “playable” games and try to find the most different but still playable locations in the design space. They use play testing to obtain a more subjective evaluation of the different games and find that they are significantly different to the original game in challenge, game feel, and theme.

Khajah et al [14] used Bayesian methods to design games and evaluated them to maximise user engagement with the generated games. The evaluation used participants that played the games for several minutes and then they had the option to either stop or continue to play the game with no further compensation. Total play time was used as well as a post-game survey. The results indicate that a user’s self-perception of competence is critical.

Perez et al [15] take agents from the successful General Video Game Artificial Intelligence (GVGAI) competition and analyse exactly how robust they are when presented with a noisy forward model as well as modified reward signals. Results show that some of the AI agents are robust to such changes. AI that can play a wide variety of games is crucial to the task of AI assisted game design.

Kunanusont et al [16] used the N-Tuple Bandit Evolutionary Algorithm to adjust game settings before evaluating the games with a series of General Video Game Artificial Intelligence (GVGAI) agents. The authors found that the Bandit algorithm was better than Random Mutation Hill Climber (RMHC) at optimising the parameters. Evaluating each game generated is an expensive process so any improvement that reduces the number of games generated that need evaluating is a great help.

Computer games have benefited from a large number of studies on the effect that they have on people. A large amount of this has been on the psychological impact of games on both adults and children, but a reasonable amount of research has been carried out on measuring enjoyment in games. Naturally there is interest in discovering why games are fun, but so far much of the research has simply focused on actually measuring fun itself.

Beume et al [17] compared algorithmic measurements from [18] with questionnaire responses. They found that the algorithmic measurements were not suitable for measuring fun, and got better answers from the questionnaires.

Ryan et al [19] measure game enjoyment in a number of studies and find that the self determination theories autonomy, competence, and relatedness factors predict enjoyment of games.

Fang et al [20] developed a questionnaire that extends the work of Nabi and Krcmar [21] to measure affective, behavioural, and cognitive reactions of respondents. This technique was then revised with the input of expert consultants as well as exploratory and confirmatory card sorting sessions. The final version resulted in 11 questions, 5 for effect, 3 for behaviour, and 3 for cognition, that measured enjoyment in games.

The rest of this paper is structured as follows:

- Section II - details the game environment that is used for the experiments as well as specific modifications made for these experiments.
- Section III - describes an experiment performed with a series of agents from the most recent iteration of the competition to evaluate the difficulty of various PO configurations.
- Section IV - describes an experiment whereby human participants play the game in two different configurations and then answer a questionnaire to determine how difficult and enjoyable they found the different games.
- Section V - gives a conclusion to the pair of experiments.
- Section VI - discusses some potential future work that could be performed.

II. Game Environment

The game environment is the Ms. Pac-Man Vs Ghost Team Competition code-base as described in [1]. This is a close approximation of the original arcade version of Ms. Pac-Man. It has a few minor deviations from the original game such as missing the bonus fruits and not giving Ms. Pac-Man the slight cornering advantage that resulted from the pixel-perfect collision system originally present in the arcade game[1]. There is a considerable API available in the engine to provide utility functions for path-finding and information gathering critical to the AI agents.

The Ms. Pac-Man engine supports three types of PO each with a sight limit enforced. They are Radius, Line-of-Sight (LOS), and Forward Facing Line-of-Sight (FF-LOS), as shown in Figure 2.

III. Artificial Intelligence Experiments

The AI experiments investigate the effect of varying PO on Ms. Pac-Man's scores, as a crude measurement of difficulty for both the ghost team and Ms. Pac-Man. A series of games were run consisting of each possible permutation of the following five independent variables:

- POType - This is the type of PO used, and can take any of the three values in Figure 2.
- Sight Limit - This is the number of nodes in the graph that the POType uses to limit the visibility. Takes values from 5 to 50 with a step of 5.
- Ms. Pac-Man agent - This is the agent controlling the Ms. Pac-Man character with the possible agents shown in Table II.
- Ghost Team agents - This is the agent controlling the ghosts with the possible agents shown in Table III.
- Communication - Communication was either on or off.

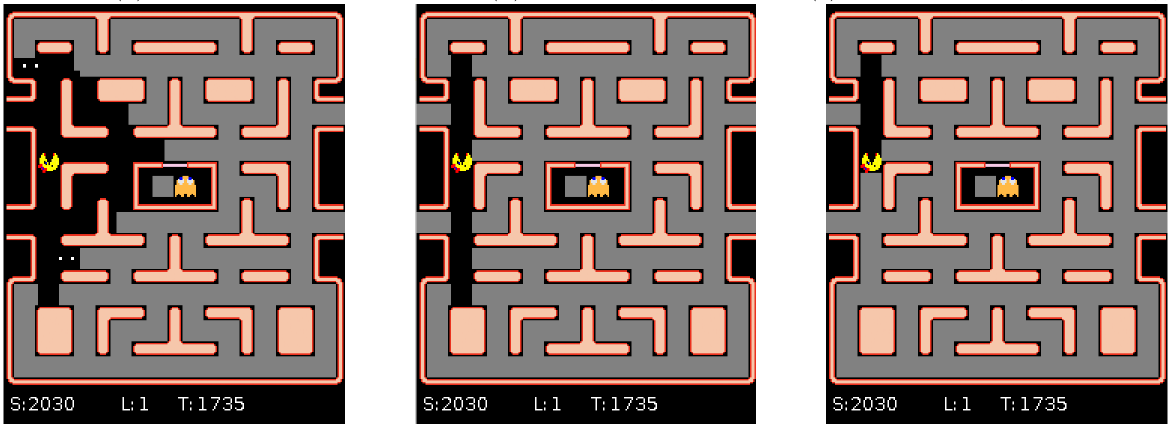
The set of agents were decided by choosing the best five agents from the 2017 Ms. Pac-Man Vs Ghost Team Competition, as well as a middle placing agent and the bottom placing agent. These agents represent a variety of good agents, middle agents, and poor agents, which have all been tested in an open public competition on the same code base. The original competition that these agents competed in was run with the LOS PO restriction and a sight limit of 50 with communication enabled for the ghost team. The entrants were aware of the restrictions that would be used on their AI agents.

A. Results

For each combination of the independent variables listed above, 100 individual games were played, and their scores recorded. This gave a target total of $3\times10\times7\times7\times2\times100 = 294,000$ games. However some games were incomplete as occasionally the AI agents crashed. This could happen since the AI agents were running beyond their design

[1]The use of pixels for collision led the rounder Ms. Pac-Man to be able to turn a corner slightly earlier than the ghosts, giving a small speed boost.

(a) Radius (b) Line-of-Sight (c) Forward Facing Line-of-Sight



(a) This restriction shows only what the agent can see in a circular fashion with the sight limit determining the radius of the circle of sight that can be seen.

(b) This restriction shows only what the agent can see in the four cardinal directions up until the next maze obstacle or the sight limit whichever is nearer

(c) This restriction is the same as LOS but only works in the cardinal direction the agent is currently facing.

Figure 2. The three modes of Partial Observability visualised.

Table II
Ms. Pac-Man agents used.

Agent	Technique
SubtleBattle	One Step Lookahead
giangrocker	MCTS
thunder	Beam Search
ToSc	State Machine
BaHe	State Machine
ImHa	Multi-Objective MCTS
MaFr	MCTS

Table III
Ghost Team agents used.

Agent	Technique
MaFr	MCTS
TiIsFePre	State Machine
thunder	Rule Based
POGC	Rule Based
NiStTiTi	State Machine
POG	Rule Based
FlBe	State Machine

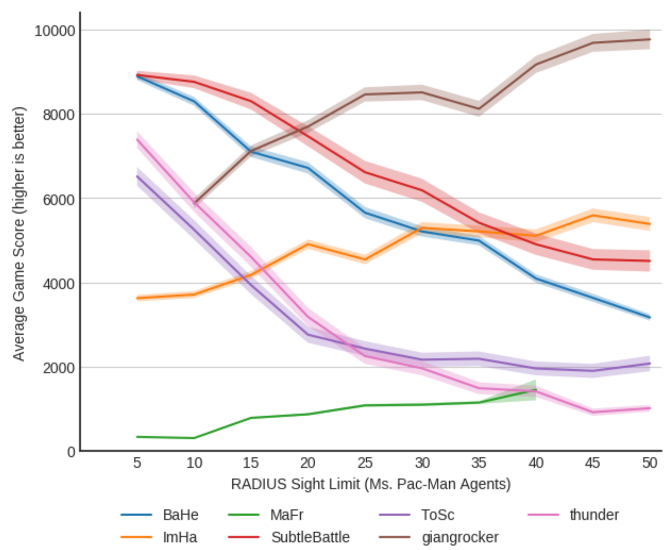


Figure 3. Effect of increasing "radius" sight limit on scores achieved by various Ms. Pac-Man agents.

specification (e.g. having communication switched off, or using a different PO mode from the original competition). The result was 174,871 successfully recorded games.

Figures 3 to 5 each show one of the three types of PO restriction, with a separate data series for each Ms. Pac-Man agent averaged over all of the Ghost agents at each sight limit. Figures 6 to 8 each show one of the three types of PO restriction, with a separate data series for each Ghost agent averaged over all of the Ms. Pac-Man agents at each sight limit

Figure 9 shows the score difference between having or not having communication at a sight limit of 50 on the LOS PO restriction.

It is important to remember at all times that the scores correspond to Ms. Pac-Man's score - so when data on Ghosts is presented a lower score indicates that the Ghost performed better.

B. Discussion

To investigate whether the difference in ability between Ms. Pac-Man and the Ghost Team is due to an advantage in increased sight, we compiled a graph showing on average how many nodes were visible to either the Ghost Team or Ms. Pac-Man in Figure 10. It is important to remember that the Ghosts cannot directly share this visibility, and there is a small delay on the information that they can share. The advantage that the agents gain

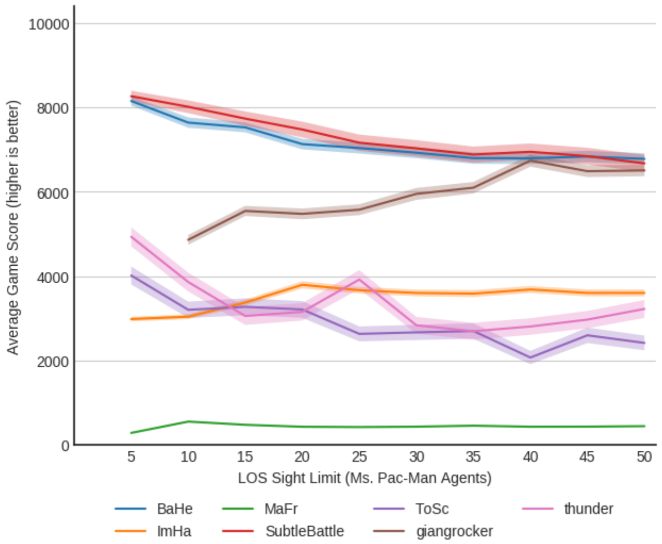


Figure 4. Effect of increasing “LOS” sight limit on scores achieved by various Ms. Pac-Man agents.

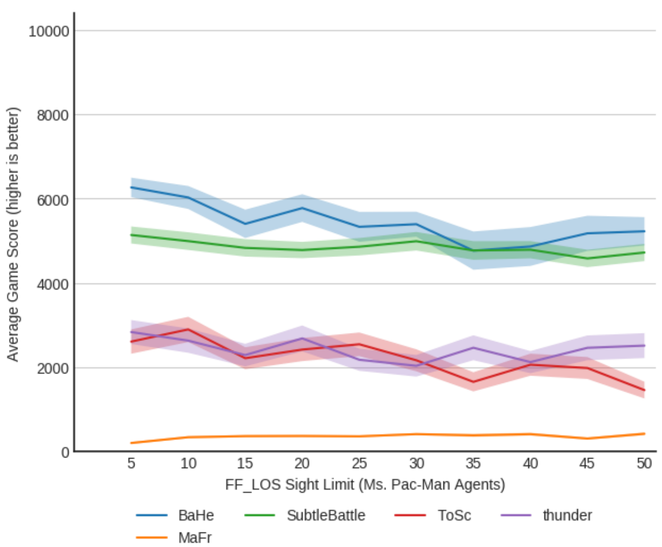


Figure 5. Effect of increasing “FF-LOS” sight limit on scores achieved by various Ms. Pac-Man agents.

from an increasing sight limit in the three modes of PO is interesting to view. The two lines for FF-LOS are indistinguishable from each other while for LOS the ghosts have on average twice the number of nodes visible at one time. For the LOS and FF-LOS modes, very little is gained as the sight limit increases. This is likely due to the fact that the agents are most often in tight little corners with obstacles being the limiting factor rather than the sight limit itself. Radius gives another story, showing a reasonable increase as the sight limit increases. It is worth mentioning that the ghosts on average can see over twice as much of the map as Ms. Pac-Man for all the sight limits.

In the case of the Radius restriction most Ms. Pac-Man agents perform worse as the sight limit increases

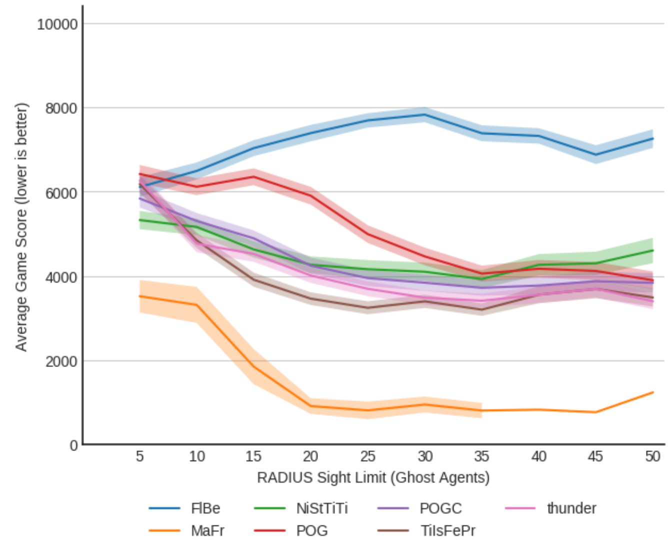


Figure 6. Effect of increasing “radius” sight limit on scores achieved by various Ms. Pac-Man agents.

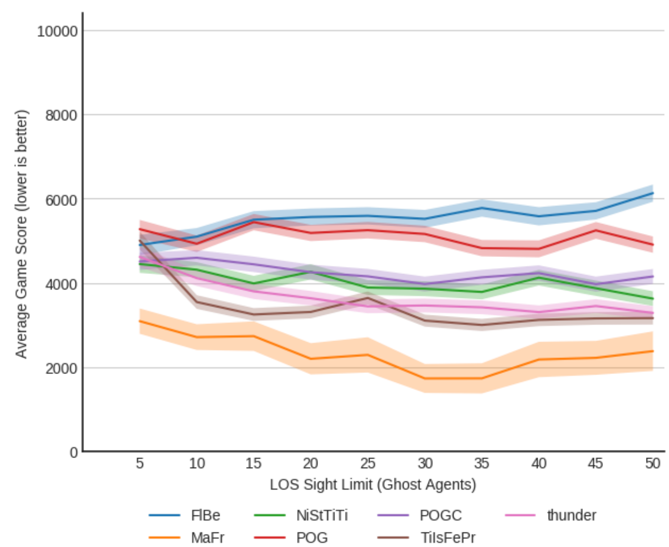


Figure 7. Effect of increasing “LOS” sight limit on scores achieved by various Ghost agents.

with a corresponding increase in performance of the Ghost agents (Figures 3 and 6). This is likely due to the ghosts enjoying over twice the benefit of extra sight than Ms. Pac-Man does (Figure 10). In LOS most Ms. Pac-Man agents perform roughly the same across the board with fairly flat lines as sight limit increases (Figure 4). The same is repeated for the five agents that completed games in FF-LOS (Figure 5).

One of the better agents from the competition, giangrocker, shows a significant benefit with additional sight limits (Figures 3 and 4). In the LOS mode, giangrocker needs the full 50-node sight limit to reach the same levels of performance as BaHe and SubtleBattle (Figure 4) whilst with the Radius PO restriction (Figure 3) giangrocker overtakes all other agents at the 20-node sight limit

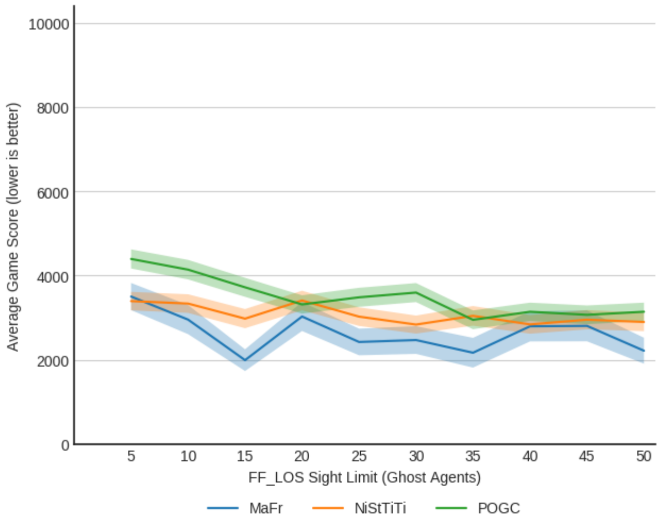


Figure 8. Effect of increasing “FF-LOS” sight limit on scores achieved by various Ghost agents.

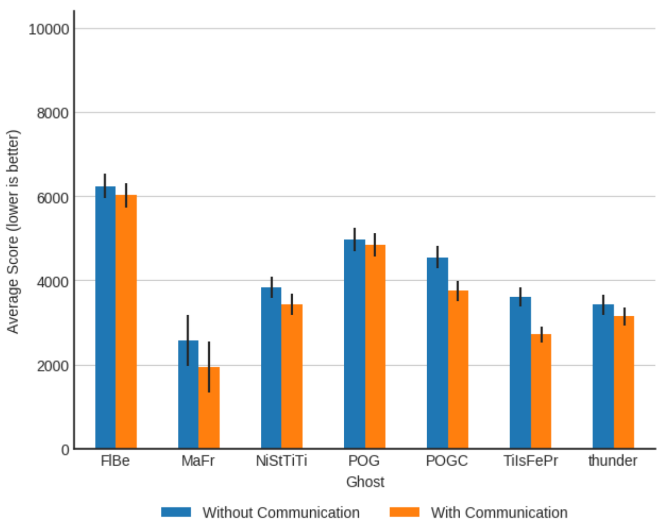


Figure 9. Effect of communication on scores by Ghost agents at a sight limit of 50 on LOS.

and gains an impressive lead. This improvement as sight increases is, to a lesser degree, shown in the other two Monte-Carlo Tree Search (MCTS) agents (ImHa and MaFr). Increasing the sight limit in the Radius restriction results in a large gain in information for the agent. MCTS in a PO environment has to make educated guesses about the environment in order to build its tree, with the accuracy of those guesses reducing the number of possible games the algorithm has to reason over. This reduction means that a higher proportion of the tree is relevant to the game leading to more accurate calculations made with that tree.

Many agents struggled to complete some games, especially those under the FF-LOS restriction (Figures 5 and 8). The error logs showed exceptions in the controller code, and it is worth mentioning that the competition

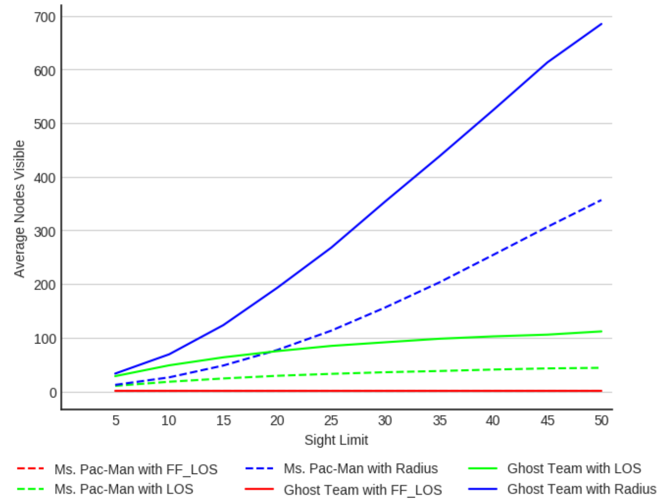


Figure 10. Average number of board nodes visible to the agents, in different PO modes, as sight limit increases.

they were designed for did not include this restriction.

The effect of communication being switched on or off (Figure 9) shows that there is a significant difference for some agents. We can see that all ghosts show a positive gain (lower score) with communication on, but the most affected agents are POGC and TiIsFePre.

IV. Human-Participation Experiments

This experiment tasks participants with playing the role of a single ghost, in a small selection of games that have altered PO or communication in them, and then filling out a questionnaire to obtain results.

A. Setup

The participants were given a short presentation containing detailed instructions on how to operate the game, as well as what the various location aids on the screen mean. Human players found it surprisingly difficult to get used to playing the role of a ghost, for example due to the inability of ghosts to reverse direction, and the difficulty for experienced Ms. Pac-Man players to focus on their ghost and pursuit, instead of on Ms. Pac-Man and fleeing. The participants were therefore allowed to play three practice games in a fully observable environment, to get accustomed to the control system and ghost behaviour, before starting the main experiment. This was to try and isolate the difficulty of the controls from the observability.

The experiment was conducted in as strict and consistent a manner as possible. All data was gathered on the same machines in the same room and configuration. The participants were tested individually, where each participant took control of the first Ghost (Blinky). The remaining three ghosts were controlled by an AI agent.

Each participant took one of two comparison experiments, where they compared two variants of the game, and then filled out a questionnaire about their experience. In comparison-experiment A, the participant played a

Table IV
Settings for the four games in the Human Experiment.

Setting: \ Game:	Practice	Radius+	LOS+	LOS-
PO	Radius	Radius	LOS	LOS
Sight Limit	∞	50	50	50
Location Aid	No	Yes	Yes	No
Ghost AI	POG	POGC	POGC	POG

version of the game with PO in Radius mode, and a version of the game with PO in LOS mode, for comparison. These are described as Game Radius with Communication (Radius+) and Game LOS with Communication (LOS+) respectively. In comparison-experiment B, the participant played a version of the game in PO LOS mode with communication switched on and a game in PO LOS mode with communication switched off. These are described as Game LOS+ and Game LOS without Communication (LOS-) respectively. The four games that participants interacted with are describe in Table IV. After completing the games, the participants were given the questionnaire.

The game used modified visuals to give the PO view required, as well as to visualise communication from other ghosts. The modifications include covering non visible areas of the map with a dark grey colouring, as well as the display of "location aids" to represent messages passed by the ghost team. These include a yellow circle to indicate the last observed location of Ms. Pac-Man, and coloured squares to indicate the locations of each ghost. A green circle was added to help the player focus on their own ghost. These are shown in Figure 11. The player used standard keyboard controls to steer their ghost (i.e. arrow keys or WASD keys).

Communication is a critical part of the ghost strategy for handling PO. When communication is switched on in the game, messages are automatically sent on behalf of the player requiring no extra effort or skill from the player. These messages are shown as location aids on screen for the user and are accessible through the API to the AI agents.

Within each comparison experiment, the order that participants played the two games was randomised. To simplify things from the participant's point of view, the two game variants they played were referred to simply as "First Game" or "Second Game" during the experiment and in the subsequent questionnaire. The participants were asked to play two games of the first game followed by two games of the second game. This was then processed into either Game Radius+, Game LOS+, or Game LOS- as appropriate.

The questionnaire given to participants differed slightly between the two experiments. Both forms contained questions taking a user id, the scores obtained in the games, the age range of the participant, and the gender of the participant. Shared questions are shown in Table V. The additional questions for Experiment A are shown in Table VI, and the additional questions for Experiment

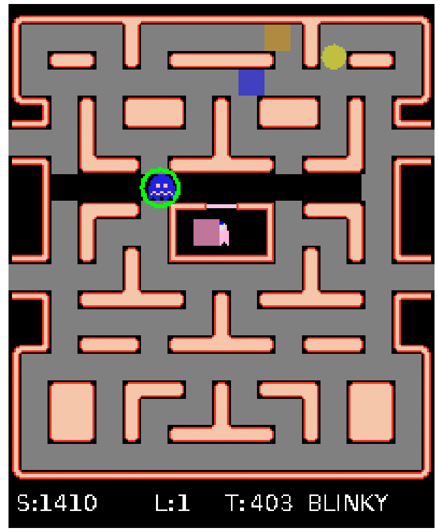


Figure 11. The graphical changes to the game environment to display required information to human participants.

B are shown in Table VII. The questionnaire focuses on asking questions comparing the player experience between the two game variants under comparison.

B. Results

The results for the various questionnaires have been compiled, calculated, and displayed in Tables VIII and IX. In total, 16 people[2] completed experiment A and 14 people[3] completed experiment B.

The results show significance on a number of the questions given to participants. Those completing Experiment A gave a significant result that they enjoyed Game Radius+ the most and found Game LOS+ the most difficult to play. In both experiments, participants did not feel claustrophobic.

In Experiment A participants found the location aid most useful in Game Radius+ rather than Game LOS+ whilst in Experiment B participants found that the location aid was useful when it was present in Game LOS+.

C. Discussion

The following subsections will individually discuss parts of the results. It is important to remember that while Game LOS+ was used in both experiments, the answers are all comparative.

1) Difficulty: Difficulty was measured in two ways for each experiment. The first method mirrors the AI experiments by obtaining score information for each participant. The second method was question 7 and partially question 8.

Looking at the scores (Table IX) obtained, Game Radius+ is the easiest for humans with an average of 4581 which is much lower than the averages for Game LOS+ (5889, 5765). Game LOS- interestingly on average scores

[2]M/F: 13/3, Age: [18-24: 8, 25-34: 5, 35-49: 2, 50-64: 1]
[3]M/F: 13/1, Age: [18-24: 8, 25-34: 3, 35-49: 1, 50-64: 2]

Table V
Questions for both tests. Some housekeeping questions are omitted.

#	Question	Options
5	In all three games the controls were the same. How difficult did you find the controls?	5 point scale "Easy"(1) → "Difficult"(5)
6	Which game did you enjoy the most?	First Game / Neither / Second Game
7	Which game did you find the most difficult?	First Game / Neither / Second Game
8	Which game did you find the most frustrating	First Game / Neither / Second Game
9	In which game did you feel the most claustrophobic?	First Game / Second Game / Didn't Feel / Equally Claustrophobic

Table VI
Additional questions for the first test. Questions re-ordered from questionnaire so that Question 10 matches between experiments better.

#	Question	Options
10	In which game did you find the location aid most useful?	First Game / Second Game / Didn't Find Useful / Equally Useful
11	In which game did you find the location aid more noticeable?	First Game / Second Game / Didn't Notice / Equally Noticeable

Table VII
Additional questions for the second test.

#	Question	Options
10	Did you find the location aid in the first game useful?	Yes / No / Didn't Notice

Table VIII
Results of the questionnaire. An * by the question name indicates p < 0.05 from a χ^2 test.

Question	Results (%)				
5A	1(25.00)	2(18.75)	3(25.00)	4(25.00)	5(6.25)
5B	1(42.86)	2(28.57)	3(14.29)	4(7.14)	5(7.14)
6A*	Radius+(56.25)	N(6.25)	LOS+(37.50)		
6B	LOS+(57.14)	N(7.14)	LOS-(35.71)		
7A*	Radius+(12.50)	LOS+(81.25)	N(6.25)		
7B	LOS+(42.86)	LOS-(50.00)	N(7.14)		
8A	Radius+(12.50)	LOS+(43.75)	N(43.75)		
8B	LOS+(14.29)	LOS-(35.71)	N(50.00)		
9A*	Radius+(6.25)	LOS+(25.00)	DF(68.75)	EC(0.00)	
9B*	LOS+(7.14)	LOS-(21.43)	DF(71.43)	EC(0.00)	
10A*	Radius+(56.25)	LOS+(25.00)	DFU(6.25)	EU(12.50)	
10B*	Yes(78.57)	No(14.29)	DN(7.14)		
11A	Radius+(37.50)	LOS+(25.00)	DN(6.25)	EN(31.25)	

Table IX
Average scores obtained in the human experiments.

Experiment	Radius+	LOS+	LOS-
A	4581	5889	
B		5765	5065

better with 5065, despite the game being theoretically harder. Game LOS- was predicted to be harder by the AI experiments (Figure 9). It is possible that the additional on-screen information proved to be more distracting than useful to people which is a problem AIs will not suffer from. The two cohorts found Game LOS+ similarly difficult when looking at scores with close averages.

Looking at the questionnaire results for Experiment A, the participants found that Game LOS+ was the most difficult with significance while Experiment B was inconclusive between Game LOS+ and Game LOS-. The AI experiments showed a much smaller gap between Game Radius+ and Game LOS+ and a large gap in scores between Game LOS+ and Game LOS-.

2) Enjoyment: While there are no AI results covering enjoyment, question 6 from the questionnaire covers enjoyment of the games for humans. Enjoyment in Experiment A showed that Game Radius+ was significantly better than Game LOS+. Experiment B proved less significant with Game LOS+ taking the lead nonetheless.

3) Communication: Communication is measured for humans through both questions as well as the scores in Experiment B. AIs were measured for their use of communication on scores alone. As mentioned previously humans performed worse with communication than without, con-

tradicting the AI results who performed better.

In Experiment A, players found the location aid more noticeable and more useful in Game Radius+ than they did in Game LOS+. These results are perhaps counter-intuitive, expecting the location aids to be more useful as vision decreases. In Experiment B the results are more one sided, with the majority saying that it was useful to have communication. Filtering the data to obtain scores for Game LOS+ and Game LOS- for those participants that stated that communication was useful yielded averages of 5112 for Game LOS+ and 4762 for Game LOS-[4]. Therefore despite believing communication was useful, on average it was not beneficial to the scores. This is an interesting result with no definitive reason why this might be the case. Comments from the participants hint at a potential information overload, perhaps making the communication hints a distraction that cost them game performance.

4) Frustration: In both experiments, more participants found the most restrictive game visually (Game LOS+ for Experiment A and Game LOS- for Experiment B) the most frustrating, but the majority found neither game to be frustrating.

5) Claustrophobia: In both experiments, the majority of participants declared that they did not feel claustrophobic in either game.

V. Conclusions

In conclusion it is clear that altering PO within the game of Ms. Pac-Man has a noticeable effect on the game itself. Firstly, it is worth re-iterating that the balance of the game changes as visibility does, with ghosts seeing on average over twice as much of the maze as Ms. Pac-Man in Radius mode. AI experiments have shown that an advantage in the amount of the map that is visible tends to be an important influence on the performance of agents depending on their particular strategies.

Secondly we found that the presence of communication for the Ghost team is beneficial for AIs but conversely found that it appears to hinder human performance despite the majority of participants declaring communication to be useful. This is a particularly interesting result because it appears to suggest that features people enjoy are not always the features that provide an advantage for them but perhaps a challenge.

Finally it is recommended that the use of AI proxies as indicators for human enjoyment should be done cautiously. There are aspects of human game playing that are difficult to emulate in AI agents such as the inability for humans to focus on all of even a small game state at once or the lightning fast reflexes of an AI. Making more human-like game AI in general is an important goal to enable more accurate estimation human player experience, and whether this is best done by building AI agents that model specific aspects of perception and cognition, or by using general learning agents trained on human behaviour is an open question.

[4]The scores are very noisy but it would take an unfeasible number of participants to bring the error to a reasonable level.

VI. Future Work

This section discusses some possible future avenues that research could take to further the work in this paper.

A. Different Games

This exploration of Ms. Pac-Man needs expanding into other games and types of games to try to correlate patterns and variations between games. If similar results are found in similar games it would be possible to derive stronger guidelines about the use of PO within games as a mechanic, opening the way offer new and interestingly different ways to play many classic games.

B. Different Alterations

This paper has explored some ways that PO could be altered in the game of Ms. Pac-Man. There are presumably a number of other interesting ways such as asymmetric PO between the ghost team and pacman. This would represent a much finer grained approach to changing the game, and potentially provide better results.

C. Other New Mechanics

This paper has discussed at length the use of PO as a game mechanic but there are other things that could be used as well. The effect of communication within a PO environment is worth deeper analysis, with different types of communication as well as tweaking various parameters such as message delay, cost of delivery, or chance of successful delivery. Altering these parameters could provide an interesting effect on the game, potentially adding some strategy to whether a person should communicate that is not currently present.

VII. Acknowledgements

The authors acknowledge the efforts of the creators of the AI entries to the Ms. Pac-Man Vs Ghost Team Competition at CIG 2017 and we are grateful for their permission to use the agents that made this work possible.The AI programmers are to be commended on the quality with which their agents functioned outside their original game parameters in new PO modes. Additional thanks to the team at CIG 2017 for hosting the competition and the IEEE CIS committee for providing prize funds to the competition. This work was funded by the EPSRC Centre for Doctoral Training in Intelligent Games & Game Intelligence (IGGI) [EP/L015846/1].

References

[1] P. R. Williams, D. Perez-Liebana, and S. M. Lucas, “Ms. Pac-Man Versus Ghost Team CIG 2016 Competition,” in CIG’2016 - IEEE Conference on Computational Intelligence and Games, ser. IEEE CIG, September 2016, pp. 420–427.

[2] P. Rohlfshagen, J. Liu, D. Perez-Liebana, and S. M. Lucas, “Pac-man conquers academia: Two decades of research using a classic arcade game,” IEEE Transactions on Games, 2017.

[3] D. Silver and J. Veness, “Monte-Carlo planning in large POMDPs,” in Advances in neural information processing systems, 2010, pp. 2164–2172.

[4] P. R. Williams, J. Walton-Rivers, D. Perez-Liebana, and S. M. Lucas, "Monte Carlo Tree Search Applied to Co-operative Problems," in CEEC'2015 - IEEE Conference on Computer Science and Electronic Engineering, ser. IEEE CEEC. IEEE Computer Society, September 2015, pp. 219–224.

[5] J. Rubin and I. Watson, "Computer poker: A review," Artificial Intelligence, vol. 175, no. 5, pp. 958–987, 2011.

[6] J. Walton-Rivers, P. R. Williams, R. Bartle, D. Perez-Liebana, and S. M. Lucas, "Evaluating and Modelling Hanabi-Playing Agents," in Congress on Evolutionary Computation, 2017. CEC'17. IEEE Conference On. IEEE, 2017, pp. 1382–1389.

[7] R. Hunicke, M. LeBlanc, and R. Zubek, "MDA: A formal approach to game design and game research," in Proceedings of the AAAI Workshop on Challenges in Game AI, vol. 4, no. 1, 2004, p. 1722.

[8] R. Koster, Theory of fun for game design. " O'Reilly Media, Inc.", 2013.

[9] M. J. Nelson and M. Mateas, "Towards automated game design," in Congress of the Italian Association for Artificial Intelligence. Springer, 2007, pp. 626–637.

[10] C. Browne and F. Maire, "Evolutionary game design," IEEE Transactions on Computational Intelligence and AI in Games, vol. 2, no. 1, pp. 1–16, 2010.

[11] A. Isaksen, D. Gopstein, J. Togelius, and A. Nealen, "Discovering unique game variants," in Computational Creativity and Games Workshop at the 2015 International Conference on Computational Creativity, 2015.

[12] A. Isaksen, D. Gopstein, and A. Nealen, "Exploring Game Space Using Survival Analysis." in FDG, 2015.

[13] A. Isaksen, D. Gopstein, J. Togelius, and A. Nealen, "Exploring Game Space of Minimal Action Games via Parameter Tuning and Survival Analysis," IEEE Transactions on Computational Intelligence and AI in Games, 2017.

[14] M. M. Khajah, B. D. Roads, R. V. Lindsey, Y.-E. Liu, and M. C. Mozer, "Designing engaging games using bayesian optimization," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. ACM, 2016, pp. 5571–5582.

[15] D. Pérez-Liébana, S. Samothrakis, J. Togelius, T. Schaul, and S. M. Lucas, "Analyzing the robustness of general video game playing agents," in Computational Intelligence and Games (CIG), 2016 IEEE Conference on. IEEE, 2016, pp. 1–8.

[16] K. Kunanusont, R. D. Gaina, J. Liu, D. Perez-Liebana, and S. M. Lucas, "The n-tuple bandit evolutionary algorithm for automatic game improvement," arXiv preprint arXiv:1705.01080, 2017.

[17] N. Beume, H. Danielsiek, C. Eichhorn, B. Naujoks, M. Preuss, K. Stiller, and S. Wessing, "Measuring flow as concept for detecting game fun in the Pac-Man game," in Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on. IEEE, 2008, pp. 3448–3455.

[18] G. N. Yannakakis and J. Hallam, "Evolving opponents for interesting interactive computer games," From animals to animats, vol. 8, pp. 499–508, 2004.

[19] R. M. Ryan, C. S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach," Motivation and emotion, vol. 30, no. 4, pp. 344–360, 2006.

[20] X. Fang, S. Chan, J. Brzezinski, and C. Nair, "Measuring enjoyment of computer game play," AMCIS 2008 Proceedings, p. 306, 2008.

[21] R. L. Nabi and M. Krcmar, "Conceptualizing media enjoyment as attitude: Implications for mass media effects research," Communication theory, vol. 14, no. 4, pp. 288–310, 2004.

Piers R. Williams is a PhD student at the University of Essex. He holds a Msci degree (2014) in Computer Science from the University of Essex. His main research is in the use of co-operation and communication within computer games and Artificial Intelligence and is the current organiser of the Ms. Pac-Man Vs Ghost Team Competition (CIG 2016, CIG 2017).

Simon M. Lucas (SMIEEE) is a professor of Artificial Intelligence and head of the School of Electronic Engineering and Computer Science at Queen Mary University of London. He holds a PhD degree (1991) in Electronics and Computer Science from the University of Southampton. His main research interests are games, evolutionary computation, and machine learning, and he has published widely in these fields with over 180 peer-reviewed papers. He is the inventor of the scanning n-tuple classifier, and is the founding Editor-in-Chief of the IEEE Transactions on Computational Intelligence and AI in Games.

Michael Fairbank is a Lecturer in Computer Science and Artificial Intelligence at the University of Essex, UK. He received his BSc degree in Mathematical Physics from Nottingham University in 1994, and a PhD in Adaptive Dynamic Programming in 2014 from City, University of London. His research interests include neurocontrol, adaptive dynamic programming, and neural-network learning algorithms, especially for recurrent neural networks.