Pac-Man or Pac-Bot? Exploring Subjective Perception of Players' Humanity in Ms. Pac-Man

Maximiliano Miranda, Antonio A. Sánchez-Ruiz, and Federico Peinado

Departamento de Ingeniería del Software e Inteligencia Artificial
Universidad Complutense de Madrid
c/ Profesor José García Santesmases 9, 28040 Madrid (Spain)
m.miranda@ucm.es - antsanch@ucm.es - email@federicopeinado.com
http://www.narratech.com

Abstract. Simulating human behaviour when playing video games has been recently proposed as an interesting challenge for the research community on Artificial Intelligence. In the exploration on Machine Learning techniques for training virtual players (i.e. computer bots) to imitate the conduct of real (human) players, we are using the classic arcade game Ms. Pac-Man as testbed. Our research goal is to find the key features that identify human playing style in this controlled environment, so we have performed an experiment with 18 human judges on how they characterise the human likeness (or absence of it) hidden behind the movements of the player's avatar. Results suggest that, although it is relatively easy to recognize human players in this game, the factors that a judge uses to determine whether a playing style is considered human or not are multiple and very revealing for creating imitation heuristics.

Keywords: Player Simulation, Virtual Humans, Believable Characters, Artificial Intelligence, Turing Test, Pac-Man, Entertainment Computing

1 Introduction

Researchers on Artificial Intelligence (AI) are always looking for problems that are challenging but feasible at the same time, in order to progress their mission of recreating intelligence in a computer. Imitating video game players has been recently considered a stimulating challenge for the AI research community, emerging several competitions on developing believable characters during the last years [1].

The Computing Entertainment industry assumes that a gaming experience with computer-controlled characters will improve if these characters behave in a less 'robotic' and more human-like style of play, that is, with subtles mistakes and natural responses to what is happening in the environment. For this reason, player modelling in video games has become an important field of study, not only for academics but for professional developers as well [2].

Human-like computer bots, as they are called, can not only be used to confront the human player, but also to collaborate with him or to illustrate how to succeed in a particular game level to help players who get stuck. It is reasonable to think that these computer-played sequences will be more meaningful if the bot imitates human style of playing. Another possible application of these 'empathetic' bots is to help during the testing phase in the game production process. These bots could be used to test the game levels, not only checking whether the game crashes or not, but verifying if the game levels have the right difficulty, or finding new ways for solving a puzzle.

One of the biggest challenges when we talk about human-like bots is precisely to define what we understand by human behaviour. The perceived humanity of a player depends on the experiences and expectations of particular human judges, so different judges may 'pronounce' different veredicts. Because of this, although it is straightforward to define a "good player" in terms of the score reached in the game, it is much more difficult to define what we understand by a "true player" in terms of its human likeness.

Another interesting question that arises naturally is what features must have a game in order to be used as part a Turing-like test designed to tell apart human players from virtual ones. Famous AAA games, like Unreal Tournament, have been previously used as an scenario for this purpose [1], but the complexity of these games makes it difficult to create bots able to deceive human judges. As part of our work, we try to answer the following question: is it possible to characterise human likeness of players using a much more simple arcade game such as Ms. Pac-Man? In this paper in particular, we performed a first experiment using Ms. Pac-Man vs Ghosts, a Java framework designed to test different AI techniques in the game.

The rest of the paper is organized as follows. Next section reviews other relevant works related to the simulation of human players. In Section 3 we describe the features of Ms. Pac-Man, including implementation details of Ms. Pac-Man vs Ghosts framework, in order to understand correctly the results of the experiment and later discussions. Section 4 explains the design of the experiment itself, the profiles of the human judges and the resources needed. The results and our analysis are presented next in Section 5. Finally, Section 6 summarizes our conclusions on how difficult seems for a 'Pac-Bot' to pass the Turing test, and suggests some interesting lines of future work.

2 Related Work

There are several works regarding the imitation of behaviour in video games, both for imitating human players and scripting-driven agents. The behaviour of an agent can be characterised by studying its proactive actions and its reactions to sequences of events and inputs over a period of time, but achieving that involves a significant amount of effort and technical knowledge [3] in the best case. Machine Learning (ML) techniques can be used to automate the problem of learning how to play a video game either using the player game traces as input like in direct imitation, or using some form of optimization technique such as genetic algorithms or reinforcement learning to optimize a fitness function that somehow 'measures' the human likeness of an agent's playing style [4].

Traditionally, several ML algorithms, like Naive Bayes classifiers and neural networks, have been used for modeling human-like players in first person shooter (FPS) video games by using sets of examples [5]. Case-based reasoning has been used successfully for training RoboCup soccer players by observing the behaviour of other players, using traces taken from the game and without requiring much human intervention [6]. Other techniques based on indirect imitation like dynamic scripting and Neurevolution achieved better results in Super Mario Bros than direct (ad hoc) imitation techniques [7].

There have been several AI competitions to test and compare different approaches for developing virtual players. Some of these competitions included special tracks for testing the human likeness of agents using Turing-like tests. One of these competitions is the Mario AI Championship [1], which included different tracks concerning the creation of AI controllers that play Infinite Mario Bros (a Java implementation of the game) to obtain the maximum score, as well as the creation of algorithms for generating levels procedurally, and even a 'Turing test track' where submitted AI controllers compete with each other for being the most human-like player, judged by human spectators [8].

The BotPrize competition [1] focuses on developing human-like agents for the FPS game Unreal Tournament, also challenging AI programmers to create bots which cannot be distinguished from human players.

Finally, Ms Pac-Man vs Ghosts, the framework that we use for this work, has been used in different bot competitions during the last years [9]. After some years discontinued, it returned in 2016 and it is running at the IEEE Computational Intelligence and Games Conference this year [10].

3 Ms. Pac-Man vs Ghosts

Pac-Man is an arcade video game produced by Namco and created by Toru Iwatani and Shigeo Fukani in 1980. Since its launch it has been considered as an icon, not only for the video game industry, but for the 20th century popular culture [11]. In this game, the player has direct control over Pac-Man (a small yellow character), pointing the direction it will follow in the next turn. The level is a simple maze full of white pills that Pac-Man eats gaining points. There are four ghosts (named Blinky, Inky, Pinky and Sue) with different behaviours trying to capture Pac-Man, causing it to lose one live. Pac-Man initially has three lives and the game ends when the player looses all of them. In the maze there are also four special pills, bigger than the normal ones, which make the ghosts to be "edible" during a short period of time. Every time Pac-Man eats one of the ghosts during this period, the player is rewarded with several points.

Ms. Pac-Man vs Ghosts (Figure 1) is a new implementation of Pac-Man's sequel Ms. Pac-Man in Java designed to develop bots to control both the pro-

[1] http://www.marioai.org/

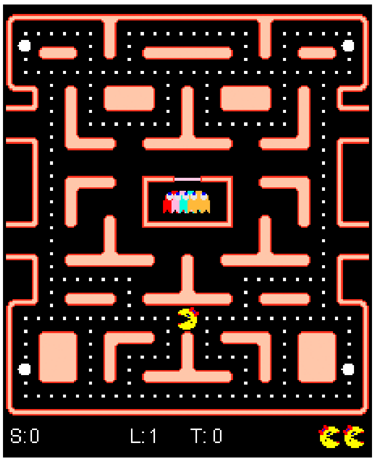

Fig. 1. A screenshot of Ms. Pac-Man vs. Ghosts

tagonist and the antagonists of the game. This framework has been used in several academic competitions during the recent years [10] to compare different AI techniques. Some of these bots are able to obtain very high scores but their behaviour is usually not very human. For example, they are able to compute optimal routes and pass very close to the ghosts while human players tend to keep more distance and avoid potential dangerous situations.

The framework provides a few examples of simple bots that we use in our experiments. Among the ghosts bots, we have selected the following behaviours:

- *Legacy* ghosts pursue Ms. Pac-Man using different ways: Blinky uses precomputed routes, Inky uses a Manhattan-based heuristic pathfinding, Pinky's pathfinding uses an heuristic based on the euclidean distance, and Sue chooses random directions when she arrives at an intersection. Just looking at the behaviour of these ghosts we see that they have 2 different states: they try to reach Ms. Pac-Man unless they are edible in which case they will try to escape.
- *StarterGhosts* show a similar behaviour: if the ghost is edible or Ms. Pac-Man is near a power pill, they escape in the opposite direction. Otherwise, they try to follow Ms. Pac-Man with a probability of 0.9, or make random movements with a probability of 0.1. Visually, they are very similar to the *Legacy* ghosts.
- *AggressiveGhosts* never run away from Pac-Man, non even when they are edible. They pursue Ms. Pac-Man no matter what.
- *RandomGhosts* are the most basic bot and they choose random directions each time they reach an intersection.

Regarding the Pac-Man bots, we use 3 basic behaviours:

- *StarterPacMan* is based on a finite state machine with 3 states: to escape from ghosts which are closer than 20 tiles, to pursue edible ghosts that are

close, and go towards the closest pill otherwise. Despite of this simple logic, this bot plays quite well.
- *NearestPillPacMan* goes always towards the closest pill, no matter where the ghosts are.
- *RandomPacMan* has a totally anarchic behaviour and decides a new random direction every game step.

The decision of using Ms. Pac-Man as a testbed to differentiate human players from automatic bots responds to several different factors. Firstly the game presents a discrete state space and a reduced number of possible actions, making the experimentation affordable. Secondly, Ms. Pac-Man is widely used as a testing ground for AI research, furthermore it is considered one of the hard games from the Arcade Learning Environment benchmark set [12], and recently Microsoft researches managed to trained a bot which managed to play a perfect game [13]. Finally, we have already worked with this game before [14, 15] and we know its architecture and implementation details.

4 Experimental Setup

The main goal of this experiment, which can be seen like a version of the famous Turing Test, first proposed by Alan Turing [16], is to determine if it is possible for human judges to distinguish between human players and automatic bots playing Pac-Man video game by simple visual evaluation. In addition, we also want to understand the reasons that lead the human judges to think the player is a human or a bot: the playing style, the precision in the actions, the skill level (in terms of score or survival skills) or some particular details in his behaviour.

4.1 Hypotheses

We work with the following initial hypotheses:

1. It is possible to distinguish a human Pac-Man player from an AI bot by simple phenomenological evaluation.
2. AI bots tend to play following fixed patterns that are relatively easy to detect.
3. Human players can be detected due to small mistakes at specific moments.
4. The more skilled a player is, the more difficult will be to distinguish him from a bot.
5. The judges tend to think the behaviour of a player is unnatural or strange when he makes decisions different from the ones the judge would make in a similar situation.

4.2 Game Settings

We used 17 Ms. Pac-Man vs Ghosts games in the experiment, 7 of which were played by automatic bots and 10 played by different human players. For each game we recorded a video of the first level of the game (maze). The videos ended either because the player completed the level or because he lost his 3 lives. We used different types of ghosts to ensure that they did not represent a determining factor for the experiment results. In particular, we used the *Starter*, *Legacy*, *Random* and *Aggressive* ghosts that were described in Section 3.

In the 7 games played by automatic bots, we used 3 different bots: *Starter-PacMan*, *NearestPillPacMan* and *RandomPacMan*. The first one has a good performance in the game and it is one of the most standard bots in the framework, the second one can be interpreted as a simplification of the first one, and the last one does not show any type of intelligence.

The other 10 games were played by 5 human players with different experience and skills. Each player played 2 different games, one against the *Legacy* and the other against the *Starter* ghosts. The first player (*P1*) is a 30-years-old expert player, *P2* and *P3* are a 25 and 33-years-old, respectively, mid-level players, *P4* is a low level 34-years-old player, and finally, *P5* player is a 11-years-old child who is playing for the first (and second) time ever to a Pac-Man video game.

In order not to alter the human games, none of the players knew that their games were going to be used in an experiment to test the human likeness of bots.

4.3 Experimental Procedure

Before the beginning of the experiment we asked the judges to complete an initial survey to gather some information about them and their background in video games (this survey will be shown in section 4.4). Next, we explained to the judges that they were going to observe various Pac-Man games and for each game they had to identity if the game was played by a human or by an automatic bot.

Once the order of the games presentation was generated randomly, each judge watched the 17 games. After each game, they had 2 minutes to fill another survey (also shown in next section) containing questions about the human likeness of the player.

Lastly, the judges had some more minutes to write their comments and impressions regarding possible aspects that leaded them to choose human or bot. It was the last opportunity to write ideas that they had not expressed in the individual games surveys.

4.4 Surveys

We use two different surveys to gather information before and during the experiment, with another final question at the end for free text with comments and explanations.

Initial Questionnaire The goal of this survey is to gather some information about the judges and their background with games. Among other information, we collect the answer to these two questions:

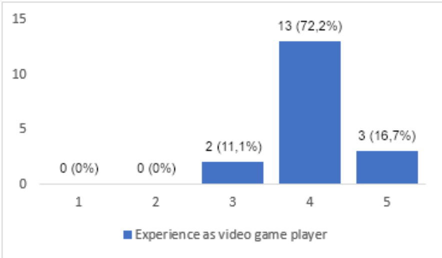

Fig. 2. Judges experience as video game players.

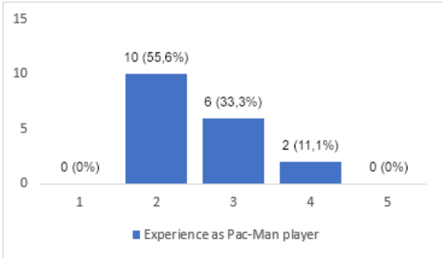

Fig. 3. Judges experience as Pac-Man players.

- How would you rate your level as a video game player? From 1 (*I do not play video games*) to 5 (*I consider myself an expert player*)
- How would you rate your level in Pac-Man? From 1 (*I've never played Pac-Man*) to 5 (*I'm an expert in Pac-Man*)

Post-Game Questionnaire This survey was filled 17 times by each judge (one after each one of the 17 games). The judges have to decide whether the game was played by a human or a bot. We offer a set of possible explanations that the judge can use (or not) to explain his decision.

- Do you think that this player is human or an AI, 3 exclusive options are given: Human player, AI (Artificial Intelligence), I don't know.
- Why? (eight options are offered plus and an additional text field where the respondent can write whatever he considers): The player makes nonsense errors, is inaccurate (sometimes he changes direction too late or too early), has one (or several) fixed behaviour (or strategy), makes cyclic movements, seems too bad (or too good) to be human, too bad (or too good) to be a bot, or too accurate to be human.

4.5 Human Judges

The experiment was carried out with 18 human judges with a certain experience in the field of video games because all of them were students in a Video game Design and Development University Degree.

Their average experience with video games was high (4.05 out of 5), with only 2 people with a score of 3 (middle), and the mode of 4 (high) (see Figure 2). Regarding their experiences as Pac-Man players, the average level was low-middle (2.5 out of 5) with a mode of 2 (with 58.82% of repetitions) (see Figure 3).

5 Results and Discussion

Table 1 shows some data of the games, and the answers percentages to the question *this player is human or an AI.*

n test	game id	time	controller type	Pac-Man controller	Ghosts controller	it is AI	it is human
1	12	0:54	Human	P4 (low level)	StarterGhosts	61.1%	**38.9%**
2	0	0:48	AI	StarterPacMan	Legacy	**72.2%**	27.8%
3	2	1:29	AI	StarterPacMan	RandomGhosts	**94.4%**	5.6%
4	15	1:19	Human	P3 (middle)	Legacy	5.6%	**94.4%**
5	8	0:56	Human	P1 (expert)	StarterGhosts	22.2%	**72.2%**
6	6	0:26	AI	RandomPacMan	StarterGhosts	**94.4%**	5.6%
7	9	0:44	Human	P5 (novice)	Legacy	50%	**50%**
8	4	0:48	AI	NearestPillPacMan	Legacy	**55.6%**	38.9%
9	5	0:45	AI	NearestPillPacMan	StarterGhosts	**44.4%**	55.6%
10	7	0:55	Human	P1 (expert)	Legacy	11.1%	**88.9%**
11	13	2:04	Human	P2 (middle)	Legacy	16.7%	**77.8%**
12	11	1:00	Human	P4 (low level)	Legacy	11.1%	**88.9%**
13	1	0:50	AI	StarterPacMan	StarterGhosts	**66.7%**	22.2%
14	3	0:39	AI	StarterPacMan	AggressiveGhosts	**72.2%**	22.2%
15	14	0:52	Human	P2 (middle)	StarterGhosts	22.2%	**72.2%**
16	16	0:58	Human	P3 (middle)	StarterGhosts	11.1%	**83.3%**
17	10	0:44	Human	P5 (novice)	Legacy	38.9%	**55.6%**

Table 1. Games data and global results. The rates indicate the relation of the judges elections (in bold those that hit).

The first game is the one with the worst results, a 61.1% of the judges points that the player is an AI but it was really played by the human player $P4$. However, after this first game the hit rate shows a significantly improvement. At the beginning the spectators suffer of a lack of knowledge how the human playing style compares against AI controller style. This is shown before with the results in the game 2, the first played by an AI. Because of this, we think that the results of game 1 should be ignored when analyzing the global results.

The average hit rate is 72%, but ignoring the first game this hit rate grows to 74%. Excluding the “I don’t know” answers, the hit rate reaches 76%.

About the games played by humans the hit rate is 76%, excluding the “I don’t know” answers the hit rate reaches 78.29%. If the game was played by an AI the hit rate is 71.43%, excluding the “I don’t know” answers, the hit rate reaches 73.69%.

Concerning the judges selections, the “Human player” option was selected 54.55% of the times, and the remaining 45,45% the “AI” option (this, without considering the “I don’t know” answers, which was selected a 2.95% of the total). This data is slightly different to the real distribution of the games: a 41.18% of them (7 out of 17) were from AI controllers, and the remaining 58.82% (10 out of 17) were games played by humans. Showing an slight variation of a 4.27% towards the “AI” option.

The main reasons marked by the judges are shown in the Table 2, which indicates the percents of the total selections for each option and the percents of the selections where the “Human player” and “AI” option were selected. For this data, the results of the first game have been discarded.

One of the factors that characterised a player is precision in its movements. Although some judges conclude that a player is an AI due to a lack of precision,

Reason	Selection %	Selected when marked as human	Selected when marked as AI
The player makes nonsense errors	28.82%	30.86%	19.84%
The player is inaccurate (sometimes he changes direction too late or too early)	20.83%	29.01%	9.52%
It is obvious that the player has one (or several) fixed behaviour (or strategy)	19.79%	9.26%	30.95%
The player makes cyclic movements	10.07%	4.94%	13.49%
The player seems too bad to be human	11.46%	6.17%	16.67%
The player seems too good to be human	4.86%	3.70%	4.76%
The player seems too bad to be a bot	8.68%	11.11%	5.56%
The player seems too good to be a bot	7.29%	12.53%	0.79%
The player is too accurate to be human	6.94%	4.32%	8.73%
Points a specific circumstance of the game	18.06%	26.54%	6.35%

Table 2. Relation of the reasons marked by the judges in the games surveys.

when the player is totally accurate, the judges generally conclude that the player is actually an AI. For example in games 2, 3 and 14, marked as AI by at least 72.2% of the judges, many of them point factors like the instantaneous change of the player's direction after eating a power pill (and starts going towards the edible ghosts), or the minimum distance the player keeps between Ms. Pac-Man and the closest ghost. Moreover, the fact that a player is perceived as accurate, is not determinant for the judges to classify him as an AI, but when the player is believed to be an AI, many judges remarks its precision. On the other hand, if a player is quite vague and carry out nonsense errors, he most probably will be classify as human. This is confirmed by the reasons selected by the judges when classifying the player as human: "The player is inaccurate" (29.01%) and "The player makes nonsense errors" (30.86%).

The other main aspect pointed by the judges for classifying a player as human is the appearance of particular situations (apart from behaviours or strategies). These could be silly mistakes like staying in the same corner for 1 or 2 game steps or heading straight for a pill, changing direction before Ms. Pac-Man could pick it, return again direction after noticing that the pill was not gained, gather the pill and finally change again to the final direction. This situations seems to make a heavy influence in the judges chose. These factors were pointed by 26.54% of the judges in the free-empty field of the surveys when marking the player as human, and appears very noticeably in the game 4, where the player makes the mistake described in the second example three times and 66,67% of the judges pointed to it.

Regarding the skill level of the human players, it seems easier to find out that they are human when they are middle or superior players, reaching a hit rate of 81.47%. Furthermore, the most indeterminate case emerges when the player is a novice in the game, with a hit rate of nearly 50% (see the games 7 and 16). So it seems that judging the playing style by the first attempts of a human is the most difficult scenario. This suggests that the relation between the skills of a human player and the facility for a spectator to identify the player as human should follow a Gaussian distribution being at the top of the bell middle-skill-players.

In that respect, the worst results emerge when the player (human or bot) is quite inaccurate, and moves without common sense (i.e. trying to avoid ghosts) or without any type of logic. This cases produce the most uncertain answers, not only as seen before in games 7 and 16, but also for AI-controllers in games 8 and 9, where these bots plays without taking care of the ghosts. In this cases it is more difficult to classify the player.

When the judges classify a player as AI, the perception of automatic behaviour patterns ("It is obvious that the player has one (or several) fixed behaviour (or strategy)" is the most selected option (30.95% of the judges).

Some judges refer to the "natural" behaviour of the player, both human and AI categories. So when the player makes "unnatural" movements (something the judge would not made in the player's current situation) looks strange to the eyes of the judges, marking them as AI. The opposite case is also valid, when the player makes natural movements according to the judges, it is marked as human (particularly clear in the game 9, where an AI player was classified as human by 55.6% of the judges).

It seems that the experiment produces some type of "learning" in the spectators, as if visualizing each game brings more knowledge about the different playing styles. As described before the worst results precisely correspond to the first one (played by an inexperienced human but marked as AI by 61.1% of the judges). Another game of the same player is presented in the 12th experiment, but this time the results are quite different, with 88.9% of the judges marking the player as human. It may be thought that this inconsistency could be due because the player plays quite differently, but the game is quite similar (practically the same duration and points earned and even the sames failures).

Another example of this situation, yet no so obvious, can be noticed in games 5 and 10, both are quite identical (same expert human player, same duration, quite the same points earned and even almost the same movements and progression through the level), in the first game (5th in the experiment), 72.2% of the judges classified the player as human, and in the second game (10th in the experiment), 88.9% did so, showing an increasing of a 16.7% in the hit rate (a growth rate of 23.13%).

It could be argued that this does not happens all the time, games 4 and 6 are from the same human player but the results are slightly better in the first one presented, however, in this case, the playing style of the player is quite different from one game to other. This was confirmed by the player himself, arguing that in the first game he was trying to test the game, and made "unnatural" mistakes (this was also pointed in some of the judges appreciations). In his second game, he just tried to "play well".

This suggests that the more games a human spectator watch, the more capable he becomes to concluded if a player is human or not successfully, even if the spectator doesn't know in real time if he is hitting or not.

6 Conclusions

As pointed in Section 4, it is possible to distinguish a human Pac-Man player from AI controllers by simple phenomenological evaluation. With a global hit rate of 74% in this experiment, we see this hypothesis closer to be proven: it seems that Ms. Pac-Man is a game complex enough to establish distinctions between human and AI players.

We also were trying to find features to characterise how a player's behaviour may seem human-like or not, and this experiment has revealed some interesting points about this. In automated controllers it seems relatively straightforward to noticed automatic and fixed behaviour patterns, and the judges marked this reason as the main one when classifying a player as AI.

About human players, results suggest that they can be discovered because of their lack of 'total precision', and for making small mistakes at specific moments. The three main reasons marked by the judges when classifying a player as human, were just the three ones regarding vagueness, nonsense errors, and errors performed at a specific moment.

The existence of a relation between the skill level of a player and the human-likeness of its playing style has been partly proven. This relation seems to follow a normal distribution, i.e. there is a point in which if the player becomes better he also becomes harder to be classified as human. This would need another experiment with much more games played by a lot of different players with diverse experience, so this is a very interesting work for the near future.

Another initial hypothesis was that judges tend to think that if behaviour of a player is unnatural or strange (making different decisions from the ones the judge itself would make in a similar situation) then the player should be classified as AI. This is confirmed by some of the reasons indicated by the judges, especially in the final survey of the experiment.

It is important to underline that the AI controllers used in the games for this experiment was not created to play in a human-like manner. We believe that the results would be much more adjusted if using human-like computer bots, so we will address this matter in the future.

Another interesting point refers to the aspects that characterised the playing style of Ms. Pac-Man players. In the future we will try to measure this parameters in more games and, with some type of automatic classifier system, verify if these could classify players in humans and bots. Furthermore, as seeing in the results analysis, the judges seem to 'learn' to distinguish more accurately, this drives us to think that we try ML techniques to develop a system that learns how to perform a Turing test of Ms. Pac-Man players.

Acknowledgements

This work is partially supported by the Spanish Ministry of Economy, Industry and Competitiveness under grant TIN2014-55006-R.

We would like to thank all volunteers who took part in our experiment for giving us their time and effort.

References

1. Hingston, P.: A new design for a turing test for bots. In: Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games, CIG 2010, Copenhagen, Denmark, 18-21 August, 2010. (2010) 345–350
2. Yannakakis, G.N., Maragoudakis, M.: Player modeling impact on player's entertainment in computer games. In: User Modeling 2005, 10th International Conference, UM 2005, Edinburgh, Scotland, UK, July 24-29, 2005, Proceedings. (2005) 74–78
3. Wooldridge, M.: Introduction to multiagent systems. Cell **757**(239) (2002) 8573
4. Togelius, J., Nardi, R.D., Lucas, S.M.: Towards automatic personalised content creation for racing games. In: 2007 IEEE Symposium on Computational Intelligence and Games. (2007) 252–259
5. Geisler, B.: An empirical study of machine learning algorithms applied to modeling player behavior in a "first person shooter" video game. PhD thesis, Citeseer (2002)
6. Floyd, M.W., Esfandiari, B., Lam, K.: A case-based reasoning approach to imitating robocup players. In: Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference, May 15-17, 2008, Coconut Grove, Florida, USA. (2008) 251–256
7. Ortega, J., Shaker, N., Togelius, J., Yannakakis, G.N.: Imitating human playing styles in super mario bros. Entertainment Computing **4**(2) (2013) 93 – 104
8. Togelius, J., Yannakakis, G., Shaker, N., Karakovskiy, S.: Assessing believability. In: Believable Bots. P. Hingston (Ed.) (2012) 219–233
9. Lucas, S.M.: Ms pac-man versus ghost-team competition. In: Proceedings of the 2009 IEEE Symposium on Computational Intelligence and Games, CIG 2009, Milano, Italy, 7-10 September, 2009. (2009)
10. Williams, P.R., Liebana, D.P., Lucas, S.M.: Ms. pac-man versus ghost team CIG 2016 competition. In: IEEE Conference on Computational Intelligence and Games, CIG 2016, Santorini, Greece, September 20-23, 2016. (2016) 1–8
11. Goldberg, H.: All Your Base are Belong to Us: How 50 Years of Videogames Conquered Pop Culture
12. Bellemare, M.G., Naddaf, Y., Veness, J., Bowling, M.: The arcade learning environment: An evaluation platform for general agents. J. Artif. Intell. Res. **47** (2013) 253–279
13. van Seijen, H., Fatemi, M., Romoff, J., Laroche, R., Barnes, T., Tsang, J.: Hybrid reward architecture for reinforcement learning. ArXiv e-prints (june 2017)
14. Miranda, M., Peinado, F.: Improving the performance of a computer-controlled player in a maze chase game using evolutionary programming on a finite-state machine. In: Proceedings 2o Congreso de la Sociedad Española para las Ciencias del Videojuego, Barcelona, Spain, June 24, 2015. (2015) 13–23
15. Miranda, M., Sánchez-Ruiz, A.A., Peinado, F.: A neuroevolution approach to imitating human-like play in ms. pac-man video game. In: Proceedings of the 3rd Congreso de la Sociedad Española para las Ciencias del Videojuego, Barcelona, Spain, June 29, 2016. (2016) 113–124
16. Turing, A.M.: Computing machinery and intelligence. Mind **59**(236) (1950) 433–460