How Does Simulating Human-Like Curiosity Through Exploration Extensions to Agent Behaviour Influence Perception of Smartness in Games?

Tomas Mazurkevic - 1702208

Abstract—This paper examines the value of exploration extension for simple artificial intelligence behaviour that potentially enhances its perception of smartness. There are many identifiers for *smartness*, however, this will be discussed in the paper with proposed measurements for *smart* AI in todays' games. Due to the nature of this project, curiosity will also be included both in psychological and computational aspects. Moreover, this paper will look into AI importance for games and why it is considered one of the best ways to test real life applications, such as robotics, autonomous driving cars, data mining and much more. The application this paper is proposing could possibly be considered for video game genre walking simulator or robotic in particular. Finally, data will be discussed, providing analysis on whether proposed exploration extension is capable of improving AI's smartness element.

I. INTRODUCTION

THROUGHOUT the last 30 years, researchers have been trying to create clever artificial intelligence (AI) using different methods and algorithms. Recent achievements include AlphaGo being able to beat world champions in Go [1] or OpenAI playing Atari games and, in most of them, getting higher scores than human players [2]. AI in games, however, is used to create non-player characters (NPCs), which can have very different patterns varying from AI as Adversary to AI as Spectacle [3].

AI in video games is crucial when it's used as one of the main mechanics. Intelligent non-player characters (NPCs) are extremely hard to make due to the nature of its complexity. The more complex it is the more outcomes developers will have to consider and the more problems will occur during development [4]. Moreover, it tends to use lots of computer resources [4], such as memory, poor run-time, scalability, and others, which brings more challenges.

Some of the very complex algorithms, such as reinforcement learning (RL), deep Q-learning, curiosity-driven, evolution-based and a lot more are being used in video games as testbeds for different disciplines that vary from training algorithms [5][6] to robotics [7][8][9][10] to Go [11][1]. However, due to the specificity of games, it is not always the best solution. Depending on the game's design, simple methods or less complex ones can be used instead of previously mentioned algorithms. As an example, Finite State Machine (FSM, Figure 3) became a common one to use after the success of the game F.E.A.R with only using three FSM states [12] or Monte Carlo Tree Search (MTCS, Figure 2) [13], which has been used in creation of a new paradigm for Go [14][15].
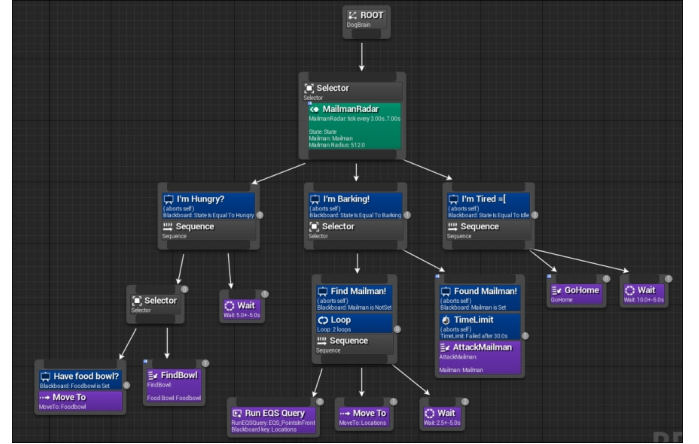

Fig. 1. Behaviour Tree in Unreal Engine

A more simple yet very effective technique for AI creation is behaviour tree (BT). It has first been introduced by Isla [4] in videogames such as Halo and is now commonly integrated into game engines like Unreal Engine[1] (UE4), which is what this project is going to use. It is believed, that BTs are "highly modular, flexible and reusable" [16]. Particularly in UE4, BTs are supported by a visual representation of nodes - tasks, decorators and services - which is being introduced in Figure 1. In fact, this is how BTs work - have a root node where the behaviour starts that leads to sequences and selectors, which are getting executing tasks depending on the current state of the game. This can be seen in Figure 4.

II. RELATED WORK

The main focus of this paper is to demonstrate a potentially valuable AI behaviour that is not complex yet efficient in enhancing smart AI. In section different AI implementation methods will be discussed, both complex and simple as well as why and how AI is used both in games and real life. This will include human curiosity and discuss it from a psychological view, why it's one of the main factors towards human's willingness to learn and gain new knowledge and some relevant learning algorithm implementations to make interesting and smart AI. Turing's Test will also be included in this section, which will describe it's use, advantages, and disadvantages and why it's used as a testing method for this
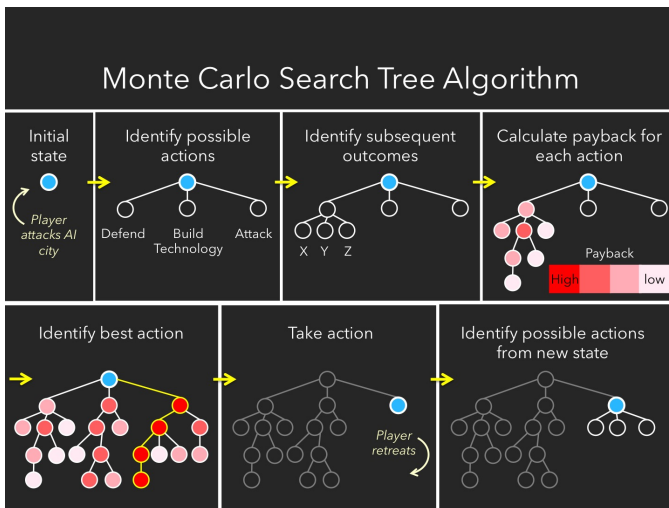
---

[1]Unreal Engine 4: https://www.unrealengine.com/en-US/

Fig. 2. Monte Carlo Tree Search [17]

Fig. 3. Finite State Machine [17]

Fig. 4. Behaviour Tree [19]

but they are still very valuable. Solving AI problems using games has become pretty common nowadays and is used for many areas as a testing purpose [11]. The testbeds can alter from simplicity to diversity [11] and can include testing natural language understanding and/or production, theorem proving, physical and simulated robotics and the original Turing's test [11].

AI for games is mainly developed to increase a player's engagement and/or satisfaction feelings [18]. Whether that's bosses or companions, friends or enemies. Given the flexibility of AI, games use it very differently. Thus, some games decide to enhance players' positive emotions and some might make the game feel terrifying. It is the designers' aspiration to create such feelings depending on the game itself. The satisfaction from hard challenges that required patience and dedication over many hours to practice to overcome, which makes them unforgettable as well as the fright from the ever building up atmosphere enhanced by terrifying creatures. All of this makes games unique and entertaining for people whose desire is to rest from our real world.

Of course, not every game must have AI, but what's the most frustrating a game can have is experience ruining AI. In which case, immersion breaks and players start complaining about the game being "unplayable" even though it's only enemies or friends behaving inappropriately or silly. Indeed, when players want to have a relaxing time playing a game, which is broken by AI only is upsetting. This again shows how important AI is for games, which requires a considerable amount of effort put into to create fun and engaging experience. So, to prove this point, the next couple of paragraphs will present and discuss the successful use of AI creating either positive or negative emotions in (1) Alien: Isolation and (2) Halo 2.

1) Alien: Isolation: Artificial intelligence is very commonly used to create various feelings alongside an atmosphere of the game. A good example of this is a game called Alien: Isolation created by Creative Assembly, Feral Interactive and awarded a BAFTA Games Award for Audio Achievement. The alien enemy is the main threat that makes the player fear it to create a horrifying experience. However, the atmosphere around it accompanied by audio also plays a huge role. Moreover, specifically in this game, the AI is very complex and smart. In fact, it has two personalities, in other words - brains [20]. The first one playing the role of the Alien "having perfect information about the environment" [20] and the second one responsible for movement and immediate stimuli response [20]. Indeed, this together creates a very interesting and exceptional behaviour of the AI. The alien reacts to any sudden
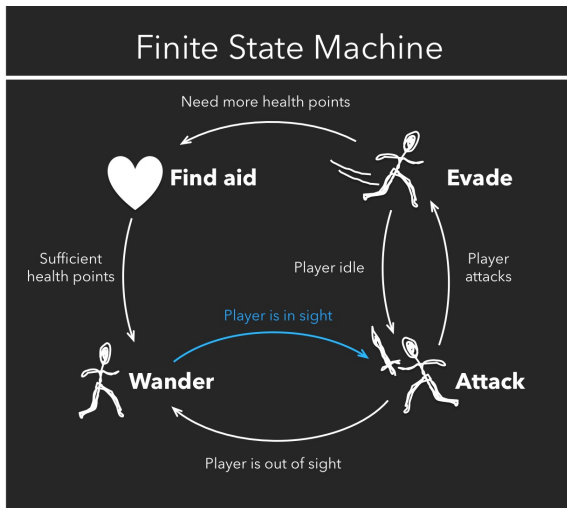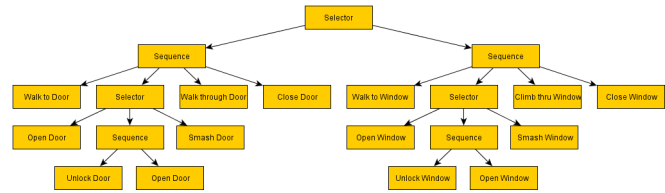
project. Moreover, it will describe how smartness is going to be measured in this context. Methodology and hypotheses will be reviewed in sections III and IV respectively. The methodology will also include research artefact, introduce to what the participants will be asked to do and how data will be collected. The proposed behaviour implementation will be handled in Unreal Engine 4.22.3 using our team's game, which can be found in appendix A, for easier testing purpose as there will be enough assets to create a unique world which AI could be tested on.

A. AI Importance in Games

Artificial intelligence has a big place in the modern world. Robots are being created to ease human life, machines that can beat the best players in the world in most complex games such as Go [1] or chess. But AI for games is not created to overcome a player's abilities. It's there to enhance player's experience, to make the game more engaging and fun to play [5] while acting realistically [13]. Therefore, other difficult problems are being solved, which may not fit well with AI used in real life,

movements, noises and other characters that are not a player. In fact, there are other enemies besides aliens, which the alien itself attacks. This creates different scenarios and entertaining gameplay where a player can bait the alien to kill any hazard on the player's way that is another NPC.

2) Halo 2: Another excellent example of smart and flexible AI in Halo 2 developed by Bungie Studios, which used the previously mentioned technique - behaviour trees [21]. AI in Halo 2 "lives in a simulated world" [18] - emphasizes Chris Butcher, one of four Engineering Leads at Bungie Studios, who was responsible for certain sections of Halo's development - and there's a very small distinction between NPCs and the player [18]. Although every creature in Halo 2 feels unique they all function the same. NPCs perceive the world using their "senses", interpret all this data and make decisions based on them [18]. This is mostly the same capabilities human players have. Therefore, this makes AI feel very smart, fun to play against and, more importantly, feel like a real human player. In addition, players can witness AI characters getting surprised, making mistakes and questionable decisions [18] because they perceive their environment the same way a human player would. This particularly makes AI predictable, which was the aim of Halo 2 making NPCs consistent [18] and is considered to be "smart" AI [22]. Thus, players can predict NPCs' actions, but the consequences will be unpredictable.

Even more impressive feature AI in Halo 2 has is their joint behaviour system [18]. It allows NPCs to communicate with each other and inform other allies of any threats or request a tactical move. This wouldn't be possible for players to notice if developers didn't add a speech to it. Indeed, whenever an enemy does any action, it is followed by it saying it out loud. This is a simple if/then statement, but it improves AIs predictability, which therefore increases its smartness as was mentioned before. In fact, small features that don't exactly relate to AI implementation might also enhance the way it feels, making it "smart" and unique.

B. Curiosity

Curiosity is a driving force for human exploration, which consists of exploration, investigation and learning behaviours [23][24]. It makes people chase for knowledge and investigate anything new and potentially valuable. Moreover, curiosity is beneficial for people on two levels: the individual and the social [25]. The first one is represented as the "innate love of learning and of knowledge... without the lure of any profit" [26]. The social level, on the other hand, is presented as "an ingredient for enhancing personal relationships" [23].

In computational form, curiosity can be split into different designs [23]. However, in this case, we are going to look at a curious exploratory agent. It can reach high learning effectiveness in an unknown environment [23][27]. Even though the behaviour this project is going to test will be hard-coded, it gives opportunity for further learning, such as having a learning algorithm like curiosity-driven instead of hard-coding. Moreover, due to curiosity being proposed as algorithmic principles [23][28][29] it enhances machines'

"exploration" and allows it to own "the desire to improve the model network's knowledge about the world." [30]. It also had success in unsupervised developmental robotics [7][10].

Curiosity-driven learning in games can show some fascinating results. In [31], researchers train AI using curiosity-driven learning without extrinsic rewards and test it on dozens of games, 48 of which are Atari games. Previously, this kind of learning was using an either extrinsic or intrinsic reward system, however, this paper applies no rewards for AI and yet demonstrates positive results. The research has also been using Random Network Distillation (RND), which showed significant progress in *Montezuma's Revenge* [32][33]. It scored more than twice as high as an average human player, which shows potential in future development. However, this kind of learning doesn't always score as high as in *Montezuma's Revenge*. This shows that it still requires improvements to be used in today's games and potentially in robotics.

The way curiosity-driven learning works is that the AI is given curiosity element into its learning and therefore it start to explore its environment. However, having only curiosity-driven learning is not always beneficial. (citation here) uses next-state prediction extension to improve curiosity-driven AI. The algorithm behind next-state prediction is that AI tries to predict what state of the game the AI will be in in the next frame. This kind of extension had success in exploring games and eventually getting higher scores than human players. However, there's one problem researchers stumble upon - noisy-TV problem. In fact, it totally break the AI while the algorithm is still working as intented. The meaning behind it is when the AI sees a random sequence of images in front of it, it tries to predict what the next image will be, therefore, getting stuck watching it. Thus, it is called noisy-TV problem, due to old TVs black and white screen noise it has when not working.

C. Turing's Test

"I believe that in about fifty years it will be possible to program computers... to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning" [34]. These are the words of Alan Turing, who inspired the creation of machine learning [35]. He created an imitation game, which is a testing game that considers whether a machine can think [34]. The way it originally works is an interrogator, that stays apart from the other two people that are a different gender, questions these two subjects. The goal of the interrogator is to identify which of the two is the man and which is the woman [34]. However, there are different interpretations of this game. Specifically for AI testing, one of the subjects is a machine and the other is human. A computer terminal is used in this case to question subjects [36].

However, this has been criticized in many ways [35][37][38][36]. In fact, some consider, that researchers are trying to only pass this test no matter what, thus excluding the important details of their implementation [35]. It has become a standard for AI these days and therefore researchers tend

	Hypothesis	Null Hypothesis	Data Source
H1	Perception of AI smartness in games is enhanced with exploration extension compared to AI without the extension	Perception of AI smartness in games is not enhanced with exploration extension compared to AI without the extension	Questionnaire results
H2	AI with exploration extension's score of interest and/or curiosity is higher than or the same as player's behaviour	AI with exploration extension's score of interest and/or curiosity is lower than player's behaviour	Questionnaire results

<div align="center">

TABLE I

HYPOTHESIS

</div>

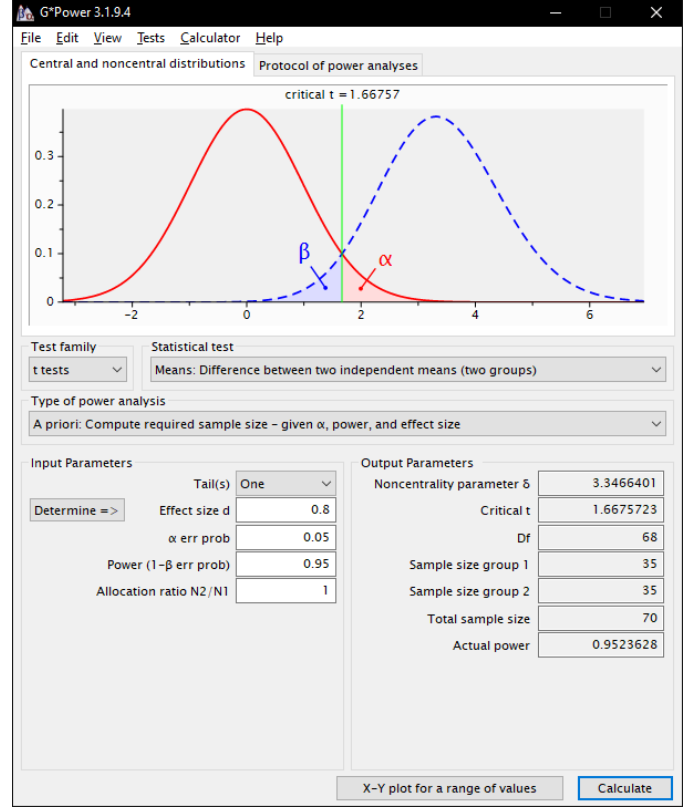| Groups/Recordings	Recording 1	Recording 2
A1	Human Player	Non-Curious AI
A2	Non-Curious AI	Human Player
B1	Human Player	Curious AI With Exploration Extension
B2	Curious AI With Exploration Extension	Human Player

<div align="center">

TABLE II

METHODOLOGY

</div>

to lean towards confirming the null hypothesis [35]. Another issue proposed in [35] is that it is not always clear what the study is trying to detect. It has to be clear at all times what is being measured and what exactly that measurement represents.

Even though Turing's test is being criticised very broadly, this project is still going to use this kind of test. However, it will not be identical to the original one nor be similar to the imitation game. In fact, there will be no subjects and only an interrogator. Therefore, the interrogator, or in other words, participant, will not be able to ask questions, but only view the AI and human behaviours and try to distinguish them. Thus, most of the criticism presented in [35][37][38][36] will not be mitigated. So, how will this paper measure proposed "smartness" of AI? Practically, there are no certain measurements for intelligence thereby researchers make up their own measurements for AI depending on their study. Though one statement sums up all the aspects of intelligent agents pretty clearly - "Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience." [39]. Paper [40] proposes measures for an intelligent agent to be the "ability to achieve goals in a wide range of environments.". As this project is looking particularly into exploration extension for AI behaviour, this statement becomes appropriate.

III. HYPOTHESES

The aim of this project is to implement a simple exploration extension for AI behaviour based on human's curiosity. Therefore, this project's first hypothesis is that exploration extension will perceive AI smartness compared to non-curious AI. The second hypothesis aims the participants to distinguish the player's behaviour from AI's behaviour, which will be provided by the questionnaire data. After analysing this data, the results will show whether the curious behaviour is leaning towards human-like, which will prove or disprove this project's hypothesis.

These hypotheses are mainly aimed to see if exploration behaviour based on human's curiosity enhances smartness for AI in games. Due to AI in games primarily being used for NPCs and concentrated on combat systems, idle states are often excluded from having "smart" behaviour. This is done in order to save the computer's resources for a more important matter, such as combat. However, having a smart looking AI for walking simulators can, therefore, enhance a player's experience in such genres. This could become a decisive factor for developers that want to create a gorgeous walking simulator using a fantastic looking environment enhanced with smart and interesting artificial intelligence.

Fig. 5. Screenshot of a-priori power analysis to calculate sample size using G*Power available at: http://www.gpower.hhu.de/

IV. METHODOLOGY

The testing of the hypotheses will be handled by separating participants into four groups, two main groups and two sub-groups, randomly chosen and shown video recordings selected at random using random number generator, see Table II. The first group will be presented with non-curious AI behaviour and the human player's behaviour or vice versa. The second group, on the other hand, will be introduced with curious AI behaviour, which was enhanced by exploration extension, and human player's behaviour or vice versa. All of this will be done randomly to minimise bias towards either side as much as possible.

A. Participants

Due to the need for a human player's behaviour, seven to eight people will be asked to play the game and record their gameplay. This necessity does not require a lot of participants and, moreover, they will not be taking part in providing research data and are not counted towards the required sample size presented in Figure 5. Ideally, they will have no to little experience with such games to generate a more curious nature

ID	Question	Data
1	How often do you play games?	Multiple Choice
2	Have you played first-person exploration walking simulators before?	Yes or No
3	Do you enjoy exploring environments in games?	1-7 Scale
4	What group were you assigned to?	Group Number
5	What behavior felt more curious, interesting?	First or Second
6	Could you distinguish which recording was of a player and which of an AI?	Yes or No

TABLE III
QUESTIONS OF THE FIRST PART OF THE QUESTIONNAIRE

ID	Question	Data
1	Which video was of a player?	First or Second
2	How confident you are about your choice to distinguish one from another?	1-10 Scale
3	What made it distinguishable?	Text Answer
4	How smart did you feel the AI was compared to player?	1-7 Scale
5	How curious did the player feel about the environment?	1-7 Scale
6	How curious did the AI feel about the environment?	1-7 Scale

TABLE IV
QUESTIONS OF THE SECOND PART OF THE QUESTIONNAIRE IF THE
ANSWER TO THE LAST QUESTION OF PART ONE WITH POSITIVE

ID	Question	Data
1	Behaviour in which video felt more curious compared to the other?	First or Second
2	Behaviour in which video felt smarter?	First or Second

TABLE V
QUESTIONS OF THE SECOND PART OF THE QUESTIONNAIRE IF THE
ANSWER TO THE LAST QUESTION OF PART ONE WITH NEGATIVE

of humans. The goal for these participants will be to explore the game's environment. However, they will not be told so to avoid bias towards having unnatural curiosity behaviour, which could influence the project's testing. Instead, they will have the ability to do whatever they desire. This will instead generate more natural behaviour and thought processes from these participants and will benefit this project's outcome. It will be handled in the first place before collecting the necessary data to analyse later in the paper. Player recordings will also help with AI development as they give an ability to observe player behaviours that could potentially be implemented for AI to imitate human-like actions.

The participants that are going to take part in the research and answer the questionnaire will likely be from different disciplines both inside and outside of Falmouth University's Games Academy due to bigger audience variety and experience in such field. There will be two groups of participants, each group being divided into another two sub-groups, see Table II, observing different recordings of human player and AI behaviours that are selected randomly as shown in Figure 5. The study will require a minimum of 70 participants, 35 of each main group, given the effect size of 0.8, which is considered large by Cohen's guidelines [41]. However, if the project brings more attention, the sample size could be increased accordingly, giving more data to analyse and potentially have more sufficient results.

B. Data Management Plan

After the participants have viewed the recordings, they will be asked to complete a survey. The survey will provide the necessary data to analyse and later discuss the efficiency of the proposed exploration behaviour enhancing the perception of AI's smartness. Data accumulated from participants will give a general conception of their experience with games, concretely with first-person exploration simulators, and their rate of perspicacity for AI and player behaviours. As shown in Table III, the first survey part intends to give a construal of their experience with games and first impression of the provided recordings. Moreover, enjoyment from exploring game environments is withal expected to affect how they rate both AI and player deportments. The query for group number provides a possibility to connect subsequent questions in the

analysis. The last two questions intend to gain a first thought answer, which is considered a genuine one [42] due to not having any explicit impact. Thus, having a first thought answer gives a veritable impression on both behaviours. Moreover, depending on the last answer, the particpant will be given different last questions.

Data from the second part of this survey commences with an overview of participator's ability to distinguish AI from player.It also provides rates of curiosity and smartness of AI and player deportments, see Table IV. The first few questions inquire for the potential to distinguish AI from player as well as how facilely they could distinguish them and how confident they are about their faculty to differentiate the two. It enables for a more rigorous analysis by granting some extra data to expand from. It is not as decisive as the rest of the queries, but it may, however, influence the final results. The last questions are the most imporant ones granting data to analyse the perception of AI's smartness. However, if the participant could not distinguish AI from player it prooves the AI is human-like. Also, a different second part of the questionnaire is given if the answer was negative, which can be seen in Table V. It provides additional data on behaviours' curiosity and smartness level.

Due to the nature of this study, linear scales from the second part are the main source of hypothesis validation. Moreover, this rating method will allow in a precise analysis of the provided data for hypothesis validation. The main reason for using a linear scale is that it grants an ideal and easy measurement for proving or disproving research's hypothesis. Also, data analysis code can be found in the appendix B.

C. Ethics

Ethics had to be concerned due to the Falmouth Research Ethics Policy. For both parts of data collection no personal data was collected that could identify a specific participant. Hence, the data is considered low-risk. However, having two parts of data collection, two different consent forms and informations sheets had to be written depending on the data collected during that part. Information sheets had to be writted to give participants' a brief idea of what the research study is about. It also states that the participation is entirely voluntary and could be withdrawn at any time without giving any reason. These can be seen in the study's repository[2].

For the first part of data collection a consent form stating that participant's gameplay will be recorded had to be mentioned. Thus, a different consent form requesting their agreement of participation had to be made. However, as mentioned before, this part still has low-risk for personally-identifiable data, as nothing except for game screen was recorded. These participant were also asked to answer a few questions after playing, which were not necessary, but potentially valuable to improve or imitate player actions for AI as well as used in analysis. This is further discussed in Section IV-B.

Second part of data collection would also require these papers, yet no personal data is required. The only questions that were asked to answer are mentioned in Section IV-B. This data is also stored anonymously, same as the previously mentioned one, and will be retained for approximately 6-month period. Afterwards, all data will be deleted to prevent any information leakage outside this study anytime.

Regarding game demo there are no legal, social or professional concerns. The demo is made exceptionally to test AI, where environment to observe is vivid and reach, and doesn't serve any purpose for human player gameplay. The assets used in the demo are free and were downloaded from Unreal Engine Marketplace[3] and are credited in the text file in the research artefact's repository[4].

V. Computing Artefact

The computing artefact is a first-person walking simulator that has no intristic goals due to the main objective being implementation of artificial intelligence and it's testing. There are two AI behaviours - simple and extended. Simple having only waypoint movement, whereas extended having both simple's AI behaviour and exploration extension, see Figure 6. This is done due to the nature of this study, where a basic AI behaviour had to be implemented and afterwards enhanced with exploration extension for Turing's Test, that was described in Section II-C. However, for a proper Turing's test a human behaviour had to be recorded. It was not only used for testing but also for AI implementation, attempting to mimic player's actions.

[2]Dissertation repository: https://gamesgit.falmouth.ac.uk/users/tm200066/repos/dissertation/browse

[3]Unreal Engine 4 Marketplace: https://www.unrealengine.com/marketplace/en-US/store

[4]Computing artefact repository: https://gamesgit.falmouth.ac.uk/users/tm200066/repos/computing-artefact/browse
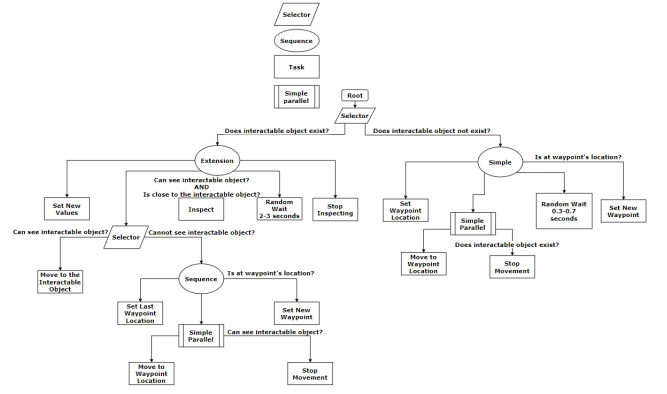

Fig. 6. UML diagram of implemented AI behaviour: Simple AI having only right behaviour, whereas extended AI having both.

The players as well as AI are granted several features for a better imitation of environment exploration game. The game is going to be referred as **demo** in the paper. Players were given full freedom so as to produce a more natural gameplay, which is used for Turing's Test and AI's development. Due to all players being students of Games Academy in Falmouth's University, their behaviour in the demo has been extraordinary. They have not been going through a predefined path, therefore, it generated some enthralling data. However, entirely emulate player's behaviour is impossible, so only a few changes were applied. This is due to a limited amount of time given for the artefact's development.

A. AI Behaviours

- Put design of the behaviours
- Explain these behaviours
- Why did I choose to use these specific features
- Why did I choose to have head rotation when the view is in first-person
- What were the difficulties??
- Using Unreal Engine and Behaviour Trees

The simple AI behaviour has been used as a base for the extended behaviour as well as tested alongside in Turing's Test. The only action it has is waypoint movement. A very simple design has been chosen to implement, therefore, dedicating more time on exploration extension performance. The testing of this extension is also essential to verify the AI is functional and is ready for Turing's Test. The extended behaviour can be seen in Figure 6. Moreover, it has other configurations implemented to augment the overall look and feel. These are:

- **Head rotation**, which is used to allow agent for a directional movement while observing the environment.
- **Sprint** - used to imitate player sprinting around the level.
- Ordained **custom navigation areas** inside waypoints that are used by an agent to emulate chaotic and unpredictable movement.

B. Life Cycle

The computing artefact has been completed within approximately 3 months, including testing, AI and level designs. It used a combination of V-model and star-model [43].

This is done with the fact that one is more structured regarding design, implementation and testing, whereas the other gives flexibility towards feature implementations. V-model also lacks freedom of choice in terms of task allocation. It was beneficial for the artefact to allow more versatility in development due to variety of behaviours required and some features for both AI and player to grant for better imitation of game exploration. However, consistent testing was required to verify that all the important features are working as well as AI behaviours are acting as expected, which V-model allows very well. If any of the features or behaviours break during the development process, an act of redesigning or reworking ir required.

A story board was being updated throughout the development whenever a task was in progress, completed or a new one supplemented. In addition to that, notes have been taken during the development process to track current bugs and to-do tasks at the time. These will be available in the artefact's repository. A testing plan was also being updated every time a test has been written and performed. Testing plan is further introduced in Section V-C and in Appendix C.

C. Testing

Artefact testing was handled regularly to ensure that the most important AI features were being functional. In order to keep track of these tests, a test plan was written beforehand to provide the necessary tests for the project. The test plan was written with a step-by-step guide of how a particular function will be tested as well as updated whenever the test was first handled. These tests were managed by two different methods:

- Unreal Engine's built-in plugin - Automation system
- Unit test
- Observing??

Automation system is a system that automatically tests any features and/or functions if an appropriate code has been written. It drastically decreases the time of discovering bugs, which keeps the software with as fewer problems as possible. Due to complexity of keeping AI fully functional, such systems are very beneficial for both small and big projects. It had greatly helped detect any crucial issues regarding AI, especially having limited time for the development. How was unit tests handled? were they even done? Testing by observing. Is it valuable?

VI. RESULTS AND ANALYSIS

A. AI Improvements

B. Player Behaviour

VII. DISCUSSION

VIII. CONCLUSION

So, in conclusion, artificial intelligence for games is very important in today's games, however, terminology "smart" is interpreted differently. It is not about winning human player's all the time how it is to maintain the game's fun and satisfaction levels. Moreover, games are also used as testbeds for AI, which can further improve real-life robotics. Robots, however, are becoming more and more advanced hence making curiosity for robots more feasible. Even though curiosity is becoming more possible for robots there are still other major problems to be solved before it becomes real.

REFERENCES

[1] D. Silver and D. Hassabis, "Alphago: Mastering the ancient game of go with machine learning," https://ai.googleblog.com/2016/01/alphago-mastering-ancient-game-of-go.html, 2016.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, p. 529, 2015.

[3] M. Treanor, A. Zook, M. P. Eladhari, J. Togelius, G. Smith, M. Cook, T. Thompson, B. Magerko, J. Levine, and A. Smith, "Ai-based game design patterns," 2015.

[4] D. Isla, "Gdc 2005 proceeding: Handling complexity in the halo 2 ai," *Retrieved October*, vol. 21, p. 2009, 2005.

[5] L. E. S. Maass and A. Luc, "Artificial intelligence in video games," https://towardsdatascience.com/artificial-intelligence-in-video-games-3e2566d59c22.

[6] R. Luzgin, "Video games as a perfect playground for artificial intelligence," https://towardsdatascience.com/video-games-as-a-perfect-playground-for-artificial-intelligence-3b4ebeea36ce.

[7] Schmidhuber, "Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts," *Connection Science*, vol. 18, no. 2, pp. 173–187, 2006.

[8] M. Colledanchise and P. Ögren, *Behavior Trees in Robotics and Al: An Introduction*. CRC Press, 2018.

[9] M. Brady, "Artificial intelligence and robotics," *Artificial intelligence*, vol. 26, no. 1, pp. 79–121, 1985.

[10] F. K. Pierre-Yves Oudeyer, "Intelligent adaptive curiosity: a source of self-development," 2004.

[11] T. Schaul, J. Togelius, and J. Schmidhuber, "Measuring intelligence through games," *arXiv preprint arXiv:1109.1314*, 2011.

[12] J. Orkin, "Three states and a plan: the ai of fear," in *Game Developers Conference*, vol. 2006, 2006, p. 4.

[13] G. Chaslot, S. Bakkes, I. Szita, and P. Spronck, "Monte-carlo tree search: A new framework for game ai." in *AIIDE*, 2008.

[14] S. Gelly and D. Silver, "Monte-carlo tree search and rapid action value estimation in computer go," *Artificial Intelligence*, vol. 175, no. 11, pp. 1856–1875, 2011.

[15] S. Gelly, L. Kocsis, M. Schoenauer, M. Sebag, D. Silver, C. Szepesvári, and O. Teytaud, "The grand challenge of computer go: Monte carlo tree search and extensions," *Communications of the ACM*, vol. 55, no. 3, pp. 106–113, 2012.

[16] M. Colledanchise, "Behavior trees in robotics," Ph.D. dissertation, KTH Royal Institute of Technology, 2017.

[17] H. Lou, "Video games as a perfect playground for artificial intelligence," http://sitn.hms.harvard.edu/flash/2017/ai-video-games-toward-intelligent-game/, 2017.

[18] R. Valdes, "The artificial intelligence of halo 2," https://electronics.howstuffworks.com/halo2-ai.htm, 2004.

[19] C. Simpson, "Behavior trees for ai: How they work," https://www.gamasutra.com/blogs/ChrisSimpson/20140717/221339/Behavior-trees-for-AI-How-they-work.php, 2014.

[20] M. Seller, "Horrific things: Alien isolation and the queer materiality of gender, desire and being," *Gender and Contemporary Horror in Comics, Games and Transmedia (Emerald Studies in Popular Culture and Gender), Emerald Publishing Limited*, pp. 43–55, 2019.

[21] T. Thompson, "Outsmarting the covenant," https://aiandgames.com/outsmarting-the-covenant/, 2015.

[22] M. Brown, "What makes good ai? — game maker's toolkit," https://www.youtube.com/watch?v=9bbhJi0NBkk, 2017.

[23] Q. Wu and C. Miao, "Curiosity: From psychology to computation," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, p. 18, 2013.

[24] T. B. Kashdan, P. Rose, and F. D. Fincham, "Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities," *Journal of personality assessment*, vol. 82, no. 3, pp. 291–305, 2004.

[25] T. B. Kashdan, "Curious?: Discover the missing ingredient to a fulfilling life ." 2010.

[26] G. Loewenstein, "The psychology of curiosity: A review and reinterpretation." *Psychological bulletin*, vol. 116, no. 1, p. 75, 1994.

[27] L. Macedo and A. Cardoso, "The role of surprise, curiosity and hunger on exploration of unknown environments populated with entities," in *2005 portuguese conference on artificial intelligence*. IEEE, 2005, pp. 47–53.

[28] S. Pang, S. Ozawa, and N. Kasabov, "Curiosity driven incremental lda agent active learning," in *2009 International Joint Conference on Neural Networks*. IEEE, 2009, pp. 2401–2408.

[29] C. Karaoguz, E. Potapova, D. Drix, M. Hülse *et al.*, "Curiosity driven exploration of sensory-motor mappings," in *Deliverable for the IM-CLeVer spring school at the Capo Caccia cognitive neuromorphic engineering workshop*, 2011, pp. 1–7.

[30] J. Schmidhuber, "A possibility for implementing curiosity and boredom in model-building neural controllers," in *Proc. of the international conference on simulation of adaptive behavior: From animals to animats*, 1991, pp. 222–227.

[31] Y. Burda, H. Edwards, D. Pathak, A. Storkey, T. Darrell, and A. A. Efros, "Large-scale study of curiosity-driven learning," *arXiv preprint arXiv:1808.04355*, 2018.

[32] Y. Burda and H. Edwards, "Reinforcement learning with prediction-based rewards," https://openai.com/blog/reinforcement-learning-with-prediction-based-rewards/, 2008.

[33] T. S. . R. Chen, "Learning montezuma's revenge from a single demonstration," https://openai.com/blog/learning-montezumas-revenge-from-a-single-demonstration/, 2018.

[34] M. Alan, "Turing. 1950," *Computing machinery and intelligence. Mind*, vol. 59, pp. 433–60.

[35] P. Hayes and K. Ford, "Turing test considered harmful," in *IJCAI (1)*, 1995, pp. 972–977.

[36] D. Livingstone, "Turing's test and believable ai in games," *Computers in Entertainment (CIE)*, vol. 4, no. 1, p. 6, 2006.

[37] L. Sweeney, "That's ai?: a history and critique of the field," 2003.

[38] L. Crockett, *The Turing test and the frame problem: AI's mistaken understanding of intelligence*. Intellect Books, 1994.

[39] L. S. Gottfredson, "Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography," 1997.

[40] S. Legg and M. Hutter, "Universal intelligence: A definition of machine intelligence," *Minds and machines*, vol. 17, no. 4, pp. 391–444, 2007.

[41] J. Cohen, "A power primer." *Psychological bulletin*, vol. 112, no. 1, p. 155, 1992.

[42] J. R. Huntsinger and C. T. Smith, "First thought, best thought: Positive mood maintains and negative mood degrades implicit-explicit attitude correspondence," *Personality and Social Psychology Bulletin*, vol. 35, no. 2, pp. 187–197, 2009.

[43] P. Isaias and T. Issa, "Information system development life cycle models," in *High Level Models and Methodologies for Information Systems*. Springer, 2015, pp. 21–40.

APPENDIX A
FIRST APPENDIX

As mentioned in the paper, project's computing artefact will be based of team's project game. The repository for that is gamesgit.falmouth.ac.uk/users/tm200066/repos/computing-artefact/browse and the only part that will be modified for the computing artefact is located in Content/Artificial_Intelligence. However, some unrelated to AI aspects will potentially be added, such as assets from other team members, to change the environment the AI is going to be placed in.

APPENDIX B
SECOND APPENDIX

```r
# Install the necessary packages
install.packages("readr")

library("readr")

# Import a CSV file and save it to a
    variable
data <- read.csv("C:/3Year/
    Dissertation/Research_Test/
    ResearchTest.csv")

# Assign variables to question results
    for easier readability and
    usability
Player <- data$How.curious.did.the.
    player.feel.about.the.environment.
AI <- data$How.curious.did.the.AI.feel
    .about.the.environment..

# Use one-tailed t-test to calculate
    difference between means that also
    gives p-value
t.test(Player, AI, paired = TRUE,
    alternative = "greater")
```

APPENDIX C
TESTING

Put any testing screenshots and test plan here.

APPENDIX D
SELF-REFLECTION

Some introduction here - I hate my life and all that

A. Dispositional

Keeping motivation for the dissertation writing was hard due to the high enjoyment of the team game project. Throughout this academic year, I have been enjoying our team game project development. Team composition and relationship between each other have been positive from the very beginning. Therefore, I was dedicating all my time to our team and the game. Thus, keeping my motivation for the dissertation project was much less than game development. So, I was behind the schedule and my peers, which made me nervous that I will struggle a lot later. Eventually, it happened and I had to make myself concentrate entirely on the dissertation putting the game aside, which resulted in me constantly trying to keep motivation for work. Moreover, a world pandemic happened and everyone was forced to work from home, which made it even harder to inspire myself to work. This could've potentially been avoided if only I have had a dedicated schedule for both my projects. A framework of at least a half an hour of work doing dissertation writing or artefact would've changed a lot. I would still be working on the game most of my time, however, with this structure, I could've been doing progress on my dissertation constantly. This would have made things much easier later on and grown a habit of having a dedicated schedule for working hours.

B. Cognitive

Write about my understanding of AI implementation in Unreal????? Or talk about testing features and how it was my first experience with it

C. Affective

Final year turned to be very stressful in the second study block. Computing artefact was very difficult to develop due to constant workflow for the team game project. Moreover, a world pandemic happened and it became more difficult to get enough participants for my dissertation. But the most frustrating experience was writing up the dissertation during horrible situation. Due to being forced to self-isolate, the working atmosphere changed, therefore, keeping motivation to work was hard. Trying to convince myself for work without having a dedicated schedule and having distractions around made it hard to concentrate and have things done or, atleast, started. Thus, whenever I was thinking about the upcoming deadline for the dissertation, which was really fast approaching without having to wake up early for University, made me stressful. This stress was either negative or positive. Either I would not be able to think or would start working because of the fact the deadline is so close.

Stress and frustration from just thinking about the deadline that's very quickly approaching.

D. Procedural

AI is prone to break a lot so debugging skills for such area is very useful???

E. Interpersonal

Supervisor meetings have helped greatly in keeping my head up as well as some pieces of advice from my peers. However, I felt like I had less help from these meetings due to my anxiety of asking for advice and thinking everything's alright.

F. Practise

Some of the features I thought would be useful to enhance player experience were not even used by half of the players.

G. Conclusion

Something to finish it up with?