Research Design 2

Georg Grech
Institute of Information & Communication Technology
Malta College or Arts, Science & Technology
Corradino Hill
Paola PLA 9032
georg.grech.d57175@mcast.edu.mt

***Abstract*—Difficulty is considered a critical aspect in developing an enjoyable video game, with difficulty settings needing to cater to all types of gamers. This study aims to achieve distinct difficulty settings in a strategy game via the adjustment of an Artificial Intelligence (AI) opponent. This is attempted with the use of Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) technologies. By periodically extracting versions of an agent during training, they can be assigned to numerous difficulty settings to create an agent that performs differently according to the difficulty selected. By testing a game using this approach with multiple participants, it was found that the agent's performance did see improvement between difficulty settings.**

***Index Terms*—Reinforcement Learning, Deep Reinforcement Learning, Game Difficulty, Player Experience, Artificial Intelligence**

I. Chosen Research

A balanced difficulty plays a crucial part in shaping an enjoyable player experience in a video game. Many have tried to address this problem using popular emerging technologies. This research, therefore, aims to study the creation of in-game opponents via the use of Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) techniques for the generation of difficulty settings in a competitive strategy game. It is hoped that this research will provide developers with a set of recommendations for the creation of effective Artificial Intelligence (AI)-based difficulty settings that may be expanded upon in future works, such as with the use of Dynamic Difficulty Adjustment (DDA) to provide an improved player experience.

This research is positioned in the following manner on the Research Onion as described by [1].

- **Research Philosophy** - Pragmatism
- **Research Approach** - Deductive
- **Research Strategy** - Experimental Research
- **Choices** - Mixed-methods
- **Time Horizon** - Cross-sectional
- **Techniques and Procedures** - Prototype game extracting quantitative metrics and a qualitative focus group

This research was inspired by the pursuit of enjoyable player experience in video games. AI has been used prominently in games to enhance player experience, sometimes in the role of an opponent to the player. Many technologies have been used in the development of these AI opponents, or "agents" as they may be referred to. In recent years, usage of RL and DRL has been popular for developing these agents, seeing usage in all types of games. Difficulty also tends to be a major factor in forming an enjoyable game experience. Players will want to select a difficulty setting that caters to their skills, or this would be done for them automatically through the use of DDA. Therefore, there has been thought to make use of RL and DRL technologies to develop difficulty settings enjoyable for several types of players.

This research presents the following hypothesis: By using versions of an RL agent trained at varying lengths, effective difficulty settings can be created in a strategy game.

The intent of this sequential mixed-methods study is to test the effectiveness of an RL agent to create an opponent with multiple difficulty settings, In the first phase, a prototype game played by a sample of participants will be used to measure the relationship between an agent set to a particular difficulty setting and the difference between final player and opponent scores. This will be followed by a qualitative focus group with the participants to better understand how players perceived the difficulty between the available settings.

II. Review of Research Methodology

The use of Deep Reinforcement Learning (DRL) agents in games has seen a steady increase in recent years. An influential study showing the potential of DRL agents is [2]. In this study, a DRL agent was trained to play several classic Atari 2600 games. This was achieved by supplying the agent with a high-dimensional visual input at a steady rate as an observation [2]. After the completion of training, evaluation was carried out by having the agent play multiple rounds of every game, then comparing its average performance to agents using other learning methods as well as an expert human player [2].

[2] did not study, however, games in which a computer opponent directly competes with a player. Other studies have tackled this scenario. One notable study was conducted in 2017 on the use of a DRL agent in the fighting game *Super Smash Bros. Melee* [3]. Instead of raw visual input, this agent was trained on a simpler setup of directly observing in-game variables concerning itself and its opponent, such as position, velocity, and action state [3]. This AI was trained using the "self-play" technique, in which it learnt by playing against past versions of itself. Evaluation on this agent was carried out having it play against professional players in two major tournaments [3].

In a competitive strategy setting, DRL has also seen success. One of the most notable examples of this in recent years is OpenAI Five in *Dota 2* [4]. This agent was also trained using self-play, similar to [3]. Training was significantly more complex however, as the agent needed to observe around 16,000 variables, and therefore took 180 days of training [4]. Also similar to [3], this agent was evaluated by examining its performance in a professional competitive setting, playing against the world champion team in an Esports game [4].

It may be observed that the previous studies all based their research on existing games. Despite this, DRL and RL technologies can be implemented by developers in their own games. For developers using the Unity game engine, this is readily available through the Unity ML-Agents package. This has been used successfully in the methodologies of studies using DRL to develop agents that compete with players in fighting games [5], as well as board games [6].

The studies mentioned thus far, however, do not study RL and DRL agents developed for multiple difficulty settings, focusing instead on the development of a single optimal agent. Studying this requires different approaches in methodology. A 2019 study attempted this on the first-person shooter game, *Unreal Tournament 2004* [7]. Training saw the agent duel against a native bot set on the hardest difficulty, throughout which, versions of the agent were periodically saved. Outside of training, this agent was set to dynamically switch between versions, attempting to find the ideal difficulty for its opponent [7]. Evaluation was performed by pitting the agent against native bots of every difficulty, but was not tested on human players, which the study recommended for future work [7].

When testing difficulty on human players, it is best to use a sample size of multiple participants to have better chances of obtaining people with differing skills. A study testing DDA in the game *Dota 2*, for instance, made use of eleven participants [8]. After agreeing to participate in the study, the participants were given a quick tutorial on how to play the game. Participants later played two rounds of the game, one of which using a DDA system, the other with a static difficulty setting [8]. Following the gameplay, participants were interviewed to gather qualitative data on how they perceived the game's difficulty [8].

While this review focused on the methodologies of academic sources, which were this study's main points of reference, non-academic material was also studied and used as reference in smaller matters. This material included projects on GitHub, as well as blog posts and posts on game development forums.

III. Reflection on the Chosen Methodology

The following research questions were formulated when undertaking this research:

- On what data can a reinforcement learning agent be trained?
- How can difficulty levels be created based on a reinforcement learning agent?
- How can the effectiveness of difficulty levels be evaluated?

Therefore, to address these research questions, the following objectives were to be accomplished:

- Find an optimal learning setup for the agent.
- Extract models during the agent's training process.
- Evaluate and select optimal models for difficulty settings.
- Create metric saving system in prototype to use for evaluation.
- Gather qualitative data by participants.

Multiple research philosophies had to be analysed to settle on the best one for this research. The major ones included:

- **Positivism**; viewing the world as objective and can be understood with direct analysis and measurement. If knowledge cannot be confirmed as true or false, it is dismissed as meaningless.
- **Interpretivism**; viewing the world as subjective, analysing people's opinions and ideas in a social and cultural context. The aim of interpretivist study is to understand the meanings of social phenomena from people's unique points-of-view.
- **Pragmatism**; focusing on the use of multiple methods, using the most appropriate method to address each research question. Pragmatism lends itself to complex, multi-faceted research issues.

This research, therefore, gears towards a pragmatic approach, as it aims to understand both the technical quantitative data of player and agent performance, as well the subjective qualitative opinions of how difficulty is perceived from player to player.

Elements were taken from the methodologies of several of the reviewed studies to form this research's methodology. For this research, the decision was made to develop an AI opponent using DRL, considering its success in competitive strategy, as seen in [4], but unlike this study, "self-play" was not used since the agent would not be directly affected by or affect the player's actions. It was decided to make use of Unity ML-Agents, as seen in [6] and [5], to implement DRL in a custom-built game. To create difficulty levels with a DRL agent, a system similar to the one used by [7] was adopted to create an agent corresponding to different difficulty levels. For evaluating our results, a qualitative angle was pursued by gathering participant opinions on the difficulty following testing, similar to [8].

The first step in this research's methodology was the development of a prototype competitive strategy game, in which a player competes against an AI opponent to gain as many points as possible by moving around a map and gathering resources of different types, exchanging them at a home base for points. This game was developed using the Unity game engine, and used the Unity ML-Agents package to develop the opponent based on DRL. Different learning setups of the agent were trained and tested before settling on a final version, from which three models were extracted and used for each

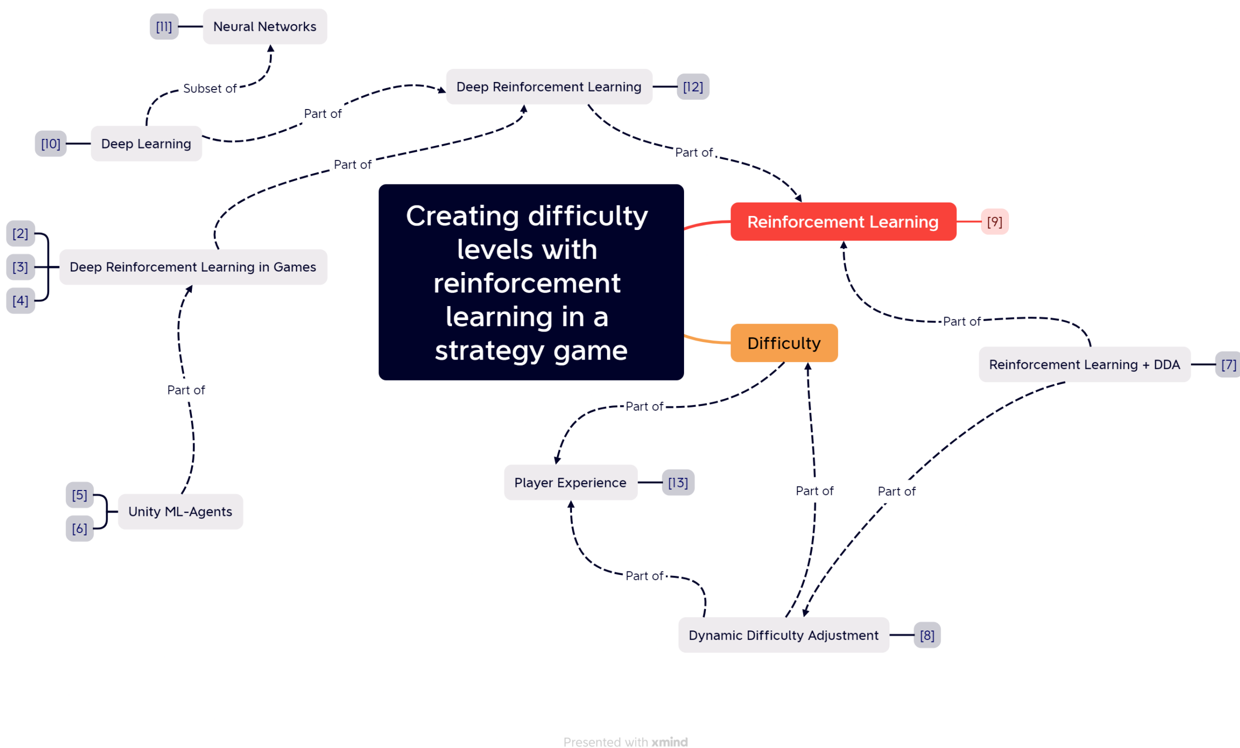

Fig. 1. Literature Map, citing additional sources: [9] [10] [11] [12] [13]

difficulty setting. The final agent was rewarded based on points achieved, and penalised based on walking distance.

The testing setup consisted of nine participants who were asked to view an in-game tutorial and play each difficulty setting, during which the game automatically saved quantitative metrics regarding player and agent performance. Following the gaming session, the participants were gathered for a focus group, during which participants could discuss their experience playing the game and their thoughts on the game's difficulty. The discussion was recorded and transcribed to be used as qualitative data in the analysis.

This methodology setup fits our research objectives, as it can prove with concrete quantitative statistics whether the agent underwent significant changes throughout difficulty settings, as well as whether these changes were effective for the purposes of difficulty by gathering qualitative opinions on the participants' experiences.

There were some ethical considerations that needed addressing. Data from participants was to be collected during testing, including images taken by the researcher during gameplay, along with a recording containing their voice that was to be transcribed. This data was handled in the most ethical way possible, limiting its use exclusively to this research, as well as keeping all participants anonymous. A participant's data was also not to be used at all if the participant decided to drop from the study, which they could do at any time. It was crucial that the participants understood all of this before agreeing, and therefore they were given a letter of information detailing what would be expected of them and how their data would be used before signing a consent form. If something wasn't clear to them, they were allowed to ask questions at any time.

IV. Findings & Discussion of Results

Following the testing session, all quantitative data from the prototype was gathered and analysed. Calculating the average score for both the players and the agent in each difficulty setting, the following result was achieved

Difficulty	Player	Enemy
Easy	489	364
Medium	540	413
Hard	584	433

A. Scores

Analysing the average agent score, it was observed that it saw an increase between difficulty levels. This supports the findings of [7], which also saw improved performance between difficulty settings using versions of an agent extracted during training. The increase was quite prominent between easy and medium, but not as significant between the medium and hard settings. This was observed as well by the participants, many of whom remarked that between easy and medium they felt a jump in difficulty, but could not easily distinguish between medium and hard.

It was observed that players generally improved as they played, gaining better scores in later games. Since most participants played difficulties in ascending order, Easy, Medium, then Hard, this can be observed by the increase in average player score between difficulties. Therefore, since players learnt alongside an improving agent, the score difference remained largely consistent between difficulties.

Difficulty	Average Score Difference
Easy	24.41%
Medium	22.41%
Hard	25.54%

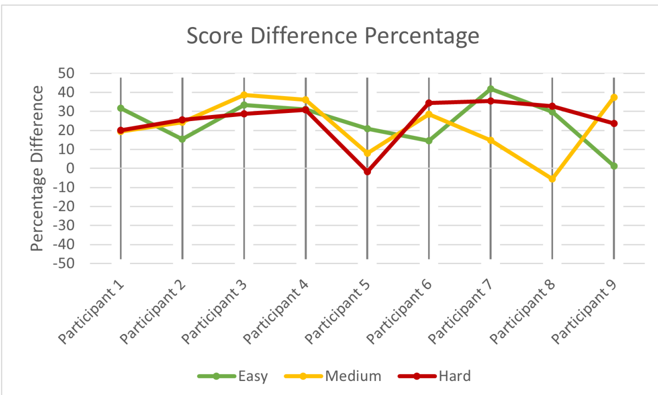

Fig. 2. Percentage score difference between player and agent in each game

B. Strategies

Improvements in both player and agent performance were owed to changes in strategies. Players admitted that their strategies changed throughout playtime. Most started off preferring the closest resources no matter the type, but soon started preferring higher value resources even if they were at greater distances. A participant noticed that the enemy displayed similar changes, but most did not perceive a distinct difference in strategy. However, participants admitted that they did not continually observe the agent, so changes may have not been detected. Analysing quantitative statistics, it was observed that agent strategy did change, as can be seen in Figure 3, taking more valuable resources (Iron and Gold) in medium and hard difficulties than in easy. Despite this, this change was not as drastic as the players', who gathered more than double the most valuable resource (Gold) in hard than in easy, as can be observed in Figure 4.

C. Different approaches to difficulty

Most participants felt that the game should automatically adjust the difficulty instead of having discrete difficulty settings, allowing both a novice and an experienced player to get an adequate challenge. This supports the findings of [8], in which players tended to enjoy dynamic adjustment in a strategy game.

Participants also suggested additions to how the game could change difficulty, such as changes in the map layout and shifting the frequency of specific resource spawns.

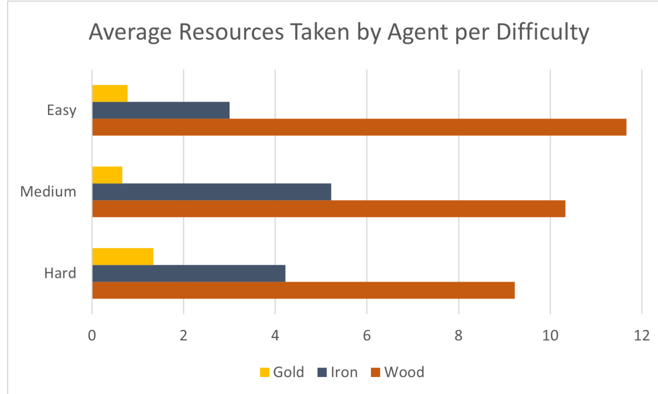

Fig. 3. Average of resource type gathered by agent in each difficulty

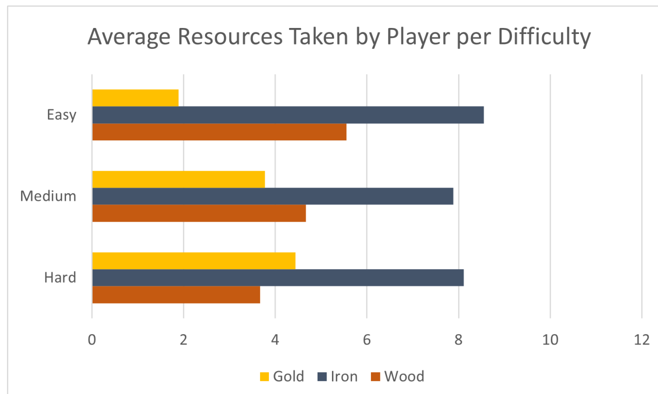

Fig. 4. Average of resource type gathered by player in each difficulty

V. Conclusion

A. Initial Conclusions

In this study, a DRL agent was used to create difficulty settings in a strategy game. It was seen how this approach managed to create an agent that performed differently throughout difficulties. In some cases, however, the differences were not distinct enough, perhaps due to limitations in the agent. It was also observed that this approach could be supported with other in-game modifications for a better difficulty. In relation to our research questions, it was found that a point-based system is an effective way to train an agent, as well penalising for wasted time. Models could be extracted during a single training run, and be used to generate difficulty settings. Lastly, difficulties can be evaluated by a mixed-methods approach, understanding how in-game performances affect subjective player opinions. The results indicate that the original hypothesis holds true, although more research may be required for better balance in difficulties

B. Limitations

The main limitation in this study was a lack of time and resources. Due to sub-optimal hardware, agents could not be trained quickly, and therefore only a small number of versions could be developed and tested within the limited time-frame.

Therefore, the final version used in this study may not be the most ideal. This limitation also resulted in a few bugs in the agent's behaviour, which a few participants experienced and pointed out.

C. Recommendations

For future research, it is recommended to attempt more thorough development and testing of the agent to possibly find an even more optimal setup that creates a greater distinction between difficulty settings, simultaneously reducing bugs and unwanted behaviours. Further research may also be attempted on how a similar system can be used in conjunction with DDA for a more enjoyable player experience, as well as attempting to expand on difficulty by changing in-game elements outside of agent behaviour.

REFERENCES

[1] M. Saunders, P. Lewis, and A. Thornhill, "Research methods," *Business Students 4th edition Pearson Education Limited, England*, vol. 6, no. 3, pp. 1–268, 2007.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013.

[3] V. Firoiu, W. F. Whitney, and J. B. Tenenbaum, "Beating the world's best at super smash bros. with deep reinforcement learning," *CoRR*, vol. abs/1702.06230, 2017. [Online]. Available: http://arxiv.org/abs/1702.06230

[4] OpenAI, C. Berner, G. Brockman, B. Chan, V. Cheung, P. Dębiak, C. Dennison, D. Farhi, Q. Fischer, S. Hashme, C. Hesse, R. Józefowicz, S. Gray, C. Olsson, J. Pachocki, M. Petrov, H. P. de Oliveira Pinto, J. Raiman, T. Salimans, J. Schlatter, J. Schneider, S. Sidor, I. Sutskever, J. Tang, F. Wolski, and S. Zhang, "Dota 2 with large scale deep reinforcement learning," 2019. [Online]. Available: https://arxiv.org/abs/1912.06680

[5] A. A. Bin Ramlan, A. M. Ali, N. H. Abdul Hamid, and R. Osman, "The implementation of reinforcement learning algorithm for ai bot in fighting video game," in *2021 4th International Symposium on Agents, Multi-Agent Systems and Robotics (ISAMSR)*, 2021, pp. 96–100.

[6] N. Baby and B. Goswami, "Implementing artificial intelligence agent within connect 4 using unity3d and machine learning concepts," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 7, no. 6S3, p. 193–200, Apr 2019.

[7] F. G. Glavin and M. G. Madden, "Skilled experience catalogue: A skill-balancing mechanism for non-player characters using reinforcement learning," in *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, 2018, pp. 1–8.

[8] M. P. Silva, V. do Nascimento Silva, and L. Chaimowicz, "Dynamic difficulty adjustment on moba games," *Entertainment Computing*, vol. 18, pp. 103–123, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1875952116300350

[9] R. S. Sutton and A. Barto, *Reinforcement learning : an introduction*. The Mit Press, 1998.

[10] L. Chen, *Deep Learning and Practice with MindSpore*. Springer Singapore, 2021.

[11] K. Gurney, *An introduction to neural networks*. London: Ucl Press, Cop, 1997.

[12] L. Lyu, Y. Shen, and S. Zhang, "The advance of reinforcement learning and deep reinforcement learning," *2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, Feb 2022.

[13] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. HarperPerennial, 1991.