Comparing Game Tree Search Techniques for General Video Game Artificial Intelligence

Alastair Rayner
Falmouth Games Academy
Falmouth, UK
Student ID: 1507516
Email: AR185160@falmouth.ac.uk

***Abstract*—Video games have been used for benchmarking artificial intelligence techniques, however in many cases the AI use very domain specific knowledge to improve their results. The General Video Game AI (GVG-AI) competition aims to address the issue of creating an Artificial General Intelligence (AGI) which in this context means to create an AI that is able to complete video games albeit not up to a human skill level. The games within the GVG-AI competition are all arcade style game, such as *Pac-Man* and *Zelda*. This paper describes the goals of GVG-AI as well as its rules and challenges, it also covers the different types of tree search techniques used in General Game Playing (GGP) as well as other game search techniques for AGI.**

Fig. 1. Example of games in GVG-AI Framework: angelsdemons (Top Left), boulderdash (Top Right), frogs (Bottom Right), Zelda (Bottom Left).

I. Introduction

THIS paper compares tree search techniques used within the General Video Game Artificial Intelligence (GVG-AI) competition. Games play an important role in the development and bench-marking of Artificial Intelligence (AI). This paper will cover the existing work around the GVG-AI competition and compare tree search algorithms that are being used in the competition. The GVG-AI competition is a recurring video game playing competition designed to simulate and benchmark general artificial intelligence. For an AI agent to be tested, they are submitted to the competition site, then they are tested against previously unseen arcade-style video games like the ones shown in figure 1. The games that the agents are tested against change every year to stop them from becoming too domain specific.

There is a large amount of tree search techniques used within the GVG-AI competition, many of them are based around Monte Carlo Tree Search (MCTS). This project will compare the success rates of each of the different methods, as well as an in depth comparison as to where each tree search technique succeeds in a set of existing games.

Add finding and resutls here, as well as contribution

II. Research Questions

There has been considerable literature on GVG-AI and individual AI techniques within the competition, and very little covering an in depth look into the comparison of the different search algorithms.

Furthermore within the GVG-AI competition there are a few different tree search techniques used, a few of which are mentioned in section III-B.

These techniques perform quite differently within the competition, this paper will aim to answer *What are the strengths and weaknesses of different search techniques* as well as *What type of game does each search technique succeed at?*

The research gathered from this could lead to a Hyper-Heuristic [1] that can be implemented into a controller for the competition. Furthermore this paper will outline the challenges of using the different search techniques, such as issues of games that have a large search space, or multiple objectives and show what are the most challenging areas for AI within the GVG-AI competition.

III. Literature Review

A. Artificial General Intelligence in games

There have been a lot of popular games that have been used to benchmark AI, for example games such as Chess, or Go. Some of these AI programs have been improved upon until they can defeat the world champion players, such as Deep Blue which defeated Garry Kasparov in 1997 [3], [4], [5].

Deep Blue was developed at IBM during the mid-1990s. It was able to beat the world chess champion by having a massively parallel system that was able to search a very large search space concurrently [3].

The challenge of creating an AI for the game Go is that it has a huge branching factor and it lacks a good evaluation

[1] A Hyper-Heuristic contains a portfolio of algorithms and is able to select the most appropriate algorithm to use depending on the game state [1], [2].

function (these terms are described in section III-B1). There have been more recent breakthroughs within AI, such as AlphaGo [6] which was developed by Google Deepmind. It was able to beat the a professional Go player in 2015 and in 2017 was able to beat Ke Jie, the world champion player at the time[6]. AlphaGo combined neural networks with MCTS to achieve a 99.8% win rate against other Go programs. The program used a supervised learning policy network that was trained directly from expert human players, then a reinforcement learning policy network was used to improve the supervised learning policy by optimising the final outcome of self played games. Further information on how this works is described in [6]. Monte Carlo Tree Search (MCTS) has had spectacular success in the game Go [7], and is implemented in the top rated go programs.

During the match against Fan Hui, AlphaGo evaluated thousands of times fewer positions than Deep blue did against Kasparov. It did this by selecting those positions more intelligently using the policy network, then evaluating them more precisely using the value network, where as Deep blue used a more brute force approach [6], [3].

The reason for these AI's success are often as a result of very domain specific knowledge about the game it is playing and means they become highly specialized and cannot be easily ported to other games. This is what started the General Game Playing (GGP) competition (described in section III-D) which was to create general AI for board games, this is similar to what happened with GVG-AI where there was lots of specialized AI for games, which were all very domain specficic, such as Starcraft AI [8], [9]. This lead to the need for general AI for games, hence GVG-AI and Arcade Learning Environment which is described in section III-D.

A long standing goal of AI is to develop algorithms capable of completing various tasks without any need to create domain specific tailoring.

B. Game Search Techniques

1) Common Terms Used in Tree Search Algorithms: **Branching Factor:** Branching Factor is the number of children at each node of a tree. This is a large factor as to what algorithms can be used to play certain types of games, for example; the GVG-AI competition has a maximum branching factor of 7, where as the game Go has a maximum branching factor of 250 [6], [10].

Horizon Effect: This is the problem where only a small portion of the search tree can be searched, i.e. the AI can search 4 moves ahead, but the 5th move could be a detrimental move, but the AI doesn't know that because its search "*horizon*" is limited [11].

Evaluation Function: This is used to estimate the value of a position or the current game state, i.e. is the agent winning or loosing the game.

Open Loop & Closed Loop: In the closed loop MCTS the algorithm assumes it is stable to store game sates on the nodes of the tree when expansion is performed, which means that selection can navigate the tree without having to calculate new states. In Open Loop MCTS (OLMCTS) the algorithm only stores the statistics on the tree nodes and then generates the next states using the forward model [12], [13].

In the competition all the games are either deterministic or stochastic, **Stochastic** In stochastic games there is some inherent randomness, meaning any values found by a tree search using the forward model (as explained in section III-D3, may change in following ticks of the game. So an agent that executes exactly the same actions each game will lead to different outcomes.

Deterministic In deterministic games each action is predictable, and any object found using the forward model will not change in forthcoming game ticks.

2) Monte Carlo Tree Search (MCTS) : Monte Carlo Tree Search is a method for finding the optimal decisions in a specified domain by performing Monte Carlo simulations (taking random samples in the search space and building a search tree according the the results) [7].

MCTS is a class of decision tree search algorithms discovered independently by several authors [14], [15], [16].

MCTS has been demonstrated to work effectively with classic board games, modern board-games and video games [17], [18].

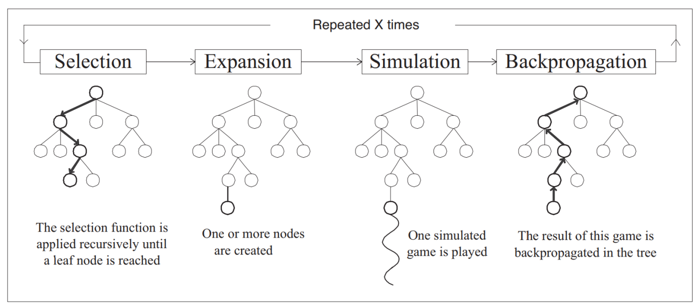

Fig. 2. Overview of Monte Carlo Tree Search. Image sourced from [17].

The basis of MCTS is the simulation of games where both the player controller and the opponent play pseudo-random moves. From playing a single game consisting of random moves, very little can be learnt about the game. However when simulating a multitude of random games, a good strategy can be inferred. This is what MCTS does, it builds a tree of possible future game states. This can described in four stages, as shown in figure 2.

Selection Starting at the root node, a selection policy is recursively applied to descend the tree until the most urgent [2] node is reached. The next action is chosen in a way that balances between exploitation and exploration. For most of the tasks, the option is to choose exploitation, which leads to the best results so far. However there may still be actions that have not been explored that could lead to good results, thus exploration is used to try and find any promising actions [17].

Upper Confidence Bound (UCB) is often a common enhancement for MCTS and is often referred to as Upper Confidence Bounds for Trees (UCT) [20], which is used by the sample MCTS controller, as described in section III-D3. UCT uses UCB1 to build a potentially asymmetric tree that

[2]Most urgent most commonly means the node with the highest UCT value.

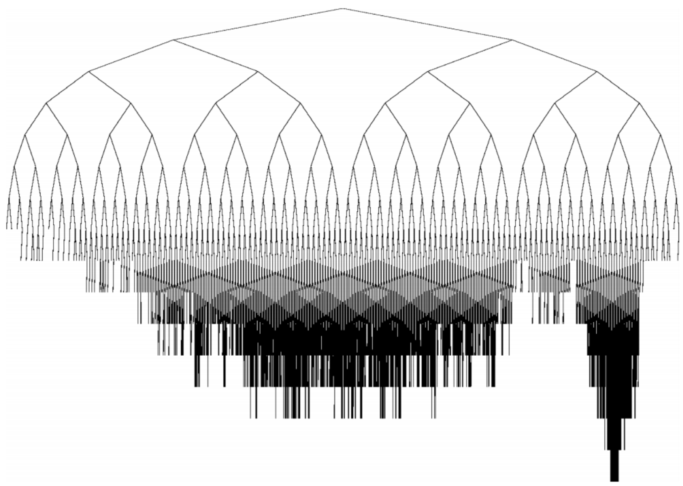

Fig. 3. An Example of an Asymmetric Tree Created by MCTS. Image sourced from [19].

grows towards the more promising parts of the search space as shown in figure 3.

UCB is based on the multi-armed bandit problem, in which it selects the optimal arm to pull in order to maximize rewards [15], [7], [21]. The main idea of UCT is to use information gathered during previous iterations of MCTS to decide what the best child node is at each level when traversing the tree. Then the child with the highest UCT value is selected.

$$UCT = \overline{X}_j + 2C_p\sqrt{\frac{2\ln n}{n_j}} \tag{1}$$

The UCT value is calculated by $\overline{X}_j$ being the average score of child j, and n is the number of times the parent node was visited and n_j is the number of times this particular child was visited. C_P is the constant value which adjusts the contribution of the second term. The second term $\sqrt{\frac{2\ln n}{n_j}}$ increases each time the parent node has been visited, but a different child was chosen.

Expansion

The major task for an expansion strategy is to choose which children should be added to the search tree.

When the selection strategy returns a node, there are a few different ways to expand the node. This step is similar to the selection step, however this step has no information available from previous iterations of MCTS for a specific child [22]. Typically only one child is added to the tree by many MCTS based agents [17], [22].

Simulation

A Monte Carlo simulation is carried out until a pre-determined depth or the end of a game. For the rest of the game, actions are selected at random until the end of the game. This means that the weighting of action selection probabilities has a significant effect on the level of play. For example if the game played with equal probability between exploration and exploitation, this will often lead to sub-optimal play [17]. To look for more promising plays, MCTS can use heuristic knowledge to give weights to actions that look more promising [23]. Silver et al. [24] describes the technique of simulation balancing using a gradient decent to bias the policy during the simulation, this aims to provide a more accurate spread of simulation outcomes.

Backpropagation

When the simulated game has played out, the tree is updated and each node in the tree that was visited is modified with the win loss ratio of that game. This informs future tree policy decisions.

The expansion and simulation stages are commonly collectively referred to as the *playout* [25]. These steps are then repeated until some predefined computational budget is reached, which most commonly are; time, memory or iteration constraint [7]. At which point the process is halted and the best performing root action is returned. This is the action that the agent will take in the game.

3) MCTS Variations: MCTS is one of the most promising baseline approaches in the literature. There has been a good deal of research on MCTS variants, each providing better results according to different domains [7], [26], [27], [28], [29], [30], [**?**].

While MCTS has been extensively applied to zero-sum games, with two players that alternate terns in discrete action spaces, it has also been applied to other domain types, such as single and multiplayer games, RTS and games with lots of uncertainty [7], [29], [30].

4) Evolutionary Algorithms: Evolutionary Algorithms (EA) are largely inspired from the biological sciences. They encode solutions to problems as individuals, part of a population which evolves over generations, until a solution is found or execution limit is reached. Figure 4 shows the stages of how Evolutionary Algorithms work [31]. Rolling Horizon Evolutionary Algorithms are one of the options available to evolve sequences of actions for planning in GVGP [32]. Furthermore within the GVG-AI competition the sampleGA controller, as described in section III-D3 uses a rolling horizon open loop algorithm [10].

Perez et al. in [32] compared MCTS with RHEA on the game Physical Traveling Salesman Problem (PTSP). The game is a modification of the popular optimization problem, the Traveling Salesman Problem [23], [33] where the player must visit a series of way points in a 2 dimensional level and the agent has up to 40ms to execute an action [32], which is similar to the GVG-AI competition. The results from that paper show that RHEA is a promising competitor to MCTS. Their approach is used in the same manner as MCTS uses roll-outs and the generative model (i.e. a simulator). Thus an agent will evolve a plan in an imaginary model for some milliseconds, then evolves a new plan repeatedly in a simulation manner, until the game is over [32]. The term Rolling Horizon comes from the fact that the planning has a certain depth that it can search within a game space (i.e. the horizon) [34], [31].

A Typical genetic algorithm is composed of several stages; First the algorithm creates a population of random individuals, then it will iteratively go though the sample and calculate the fitness (i.e. the evaluation function as described in III-B1) of every individual in the population. The most fit individuals are then used to form a new generation, and each individual

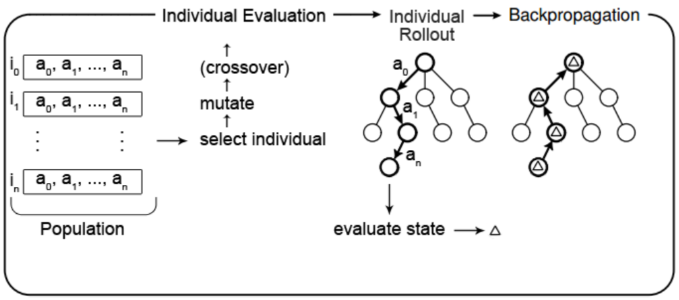

Fig. 4. RHEA statistical tree steps. Image sourced from [31].

is modified with a combination of crossover and mutation. Crossover is representative of biological reproduction where a child is produced based of the parents properties. Mutation is where a random value is changed to explore the search space and is analogous to biological mutation [35], [36], [32], [31], [34]. This process can be shown in figure 4.

Recent literature on AGI have been combining evolution and tree search in interesting ways in order to combine the benefits of both methods, such as Lucas et al. [37] applied an EA process to guide the simulation step of MCTS and improve the random default policy. Their results show a significant increase in performance in the game *Space Invaders* and the Mountain Car Problem.

5) Minimax: Minimax attempts to minimize the opponent's maximum reward at each state. The tree search is often stopped prematurely and a value function is used to estimate the outcome of the game, then the Alpha-Beta heuristic is used to prune the search tree [7], [38].

6) Depth First Search: Depth First Search (DFS) is a tree search algorithm that iteratively expands each unvisited child node of the tree until a non-expandable node is reached. The search then navigates back up the tree until it finds the next unvisited node to expand and continues the search until all the nodes have been visited [23]. Depth first search with alpha-beta pruning [38] has achieved superhuman performance in chess, checkers and orthello, however it has not been effective in the game Go [6].

7) Breadth First Search: Breadth First Search (BrFS) searches the tree one level at a time, it starts at the root node and searches neighbouring nodes first, before searching their child nodes. Within the gvg-ai competition, there have been several variations of this, one of them which uses a hash function to improve the performance of the algorithm [39], which is similar to the BrFS Extended controller used in this paper.

8) Best First Search: Best First Search (BeFS) explores a tree by selecting the most promising node and expanding it until a specified limit is reached, or a solution is found. Breadth First Search is used by YOLOBOT, which was developed by T. Joppen et al. [11] and arguably the most successful controller in the competition as of writing this paper, their approach uses Best First Search for deterministic games and MCTS for stochastic games [40].

C. Hyper-Heuristics

A hyper-heuristic approach typically combines several AI methods and automate them to be able to choose the best solution for that situation, a term used to describe them is '*heuristics to choose heuristics*' [41], [42]. During the 2015 competition, the three winning controllers all were a type of hyper-heuristic [2].

D. The General Video Game AI Competition

In most modern video games the AI is tailored specifically for that game and can't easily be modified for use in a different game type. However this is what GVG-AI aims to solve, by creating an AI that can play any game.

There have been quite a few AI competitions before in video games, such as Unreal Tournament [9], Super Mario Bros [43], Starcraft [8]. However most of the winning AI strategies used in those games are very domain specific and it is often more about knowing the game than developing good general AI [10].

Another competition that was similar to GVG-AI was the General Game Playing (GGP) competition [44], [45]. However almost all of the games in the GGP are board games, and the Game Description Language (GDL) used is not designed for video games.

The GVG-AI Competition is a competition framework that proposes the challenge of creating controllers for general video game playing. The controllers must be able to play a wide variety of video games, many of them will be completely unknown to the controller. This means the controller must have some general AI to discover the mechanics and goal of the game, so it can increase its score and win the game. [46], [10]

The framework contains a library of 2D video games some of which are based of classic arcade games, there are currently as of writing this, 82 different singleplayer games and 20 multiplayer games that AI controllers can be tested on.

Games in the GVG-AI competition are written using the Video Game Description Language [47], which is a high level description language designed to be able to create a wide variety of arcade style games, where the rules take the form of sprite movement and interaction on a 2D grid [48]. VGDL in GVG-AI is a JAVA port from pyVGDL which was programmed in python. The VGDL is a powerful tool for conducting research on computational intelligence and games [47], [45].

The Arcade Learning Environment [49] provides an interface to hundreds of Atari 2600 game environments. ALE is similar to GVG-AI in a couple of ways; firstly it provides a test bed for benchmarking AI techniques within video games and secondly it is focused around creating agents within video games, as opposed to the GGP competition [44]. The Atari 2600 is a video game console developed in 1977 , it has had over 500 original games released for the console, and nearly all popular arcade games at the time were pored to the console such as; *PAC-MAN* and *SPACE INVADERS* [49]. This provides a large test-bed for AI agents. The hardware for the Atari 2600 is very limited compared to today's standards, it had a 1.19Mhz CPU and 128 Bytes of RAM. These hardware specifications limit the complexity of the games that can be played on it, which strikes a balance of challenging but allowing search algorithms to have a small enough search space as to not have a large horizon effect [49]. ALE has been used deep reinforcement learning techniques, that have been able to reach human level of play [50], [31].

The growing interest in competitions such as the ones mentioned above clearly reflects a desire for general competency for AI within games.

1) Challenges and Goals of GVG-AI: The goal of GVG-AI is to create a generally intelligent agent that is able to win any game it is placed in, when it doesn't know the game. During the tournament a completely new set of games are used, to avoid the agents becoming too domain specific. Another challenge is the time limit that an agent can choose an action, this avoids the agent spending too long deciding a task and not making an action [22]. Furthermore this also is what defines the difference between GVG-AI and GGP, because having such a short amount of time to compute an action is what makes the competition such a challenge, otherwise the agent will be able to brute force the search space to find the best possible solution, which may take minutes or hours to compute each game tick.

2) Competition & Rules: The winning conditions are decided by three factors:

- Number of games finished with a victory
- Total sum of points
- Total time spent

The fist objective to be considered is the number of victories, however in case of a tie, the next objective is the number of points. Then if those two are a tie then the final decider is the total time spent before the win [10]. In the competition the agent will play 10 unknown games and 5 levels per game. Furthermore each level is played ten times, so each agent plays roughly 500 total games in the tournament [22]. The competition does now allow multithreading that is used by some MCTS enhancements, some MCTS enhancements that are used are discussed in section III-B3. The controllers can also use up to 1 second of CPU time for initialization and 40ms to compute an action each game tick. However if the controller takes between 40ms and 50ms, the action return will be *NIL*(Where no movement will be applied), anymore than 50ms will result in the controller automatically loosing [2], [10].

3) The GVG-AI Framework & Sample Controllers: The Framework is developed in the Java Environment and has a few differnet tracks that you can submit AI for, these are; Single Player Planning Track, 2-Player Planning Track and Level Generation Track [51]. The controllers are allowed upto 40ms to compute the agents action(s) [52], [46].

These sample agents provide useful insights into how new agents can be created for the competition by applying common AI techniques. The *HUMAN* player and the *REPLAYER* can be used for debugging the game and to help the programmer get a better understanding of how the game can be played.

The framework uses a Video Game Description Language (VGDL) to describe a wide variety of video games. The VGDL is based on a python version developed by Schaul (2014) called PyVGDL [22]. Furthermore in the GVG-AI Competition the AI agent does not have access to the whole games description, where as in GGP the agent was able to see the whole game description. This means that the agent has to analyze and simulate the game in order to figure out the rules and goal of the game.

The game uses the forward model, which allows the game to copy the current state of the game and advance it by executing a specified action.

A lot of competition winning controllers use a variation of tree search algorithms, i.e. Maast and number27 [22] both use breadth first search for deterministic games, this is due to the game states not changing during the game so extended tree searches work well [40].

The framework has an StateObservation object that has an interaction set that consists of *up, down, left, right, nil, escape* and *use*.

The GVG-AI framework comes with quite a few sample agents;

Sample MCTS The GVG-AI framework proves a sample MCTS controller, this controller has received considerable interest due to its success in the competition. The sample MCTS controller is an implementation of the vanilla MCTS algorithm, this is described in section III-B2, and the full description of MCTS algorithm is described by Browne et al. [7]. The sample MCTS controller uses a payout depth of 10 moves and an exploration-exploitation constant value of $\sqrt{2}$ (from equation 1) and selects the most visited action from the root to pick a move to return for each game cycle [10]. However MCTS isn't the sole answer to the GVG-AI competition, as it has been shown that even with a 30x computational budget, it fails to master games. However it does manage to avoid to explicitly loosing games, but does not win a lot of them either. This shows that for the AI it finds not losing is significantly easier than winning [48].

sample OLMCTS The Open Loop Monte Carlo Tree Search uses the same techniques as the sampleMCTS, however it does not store the game state within the nodes of the tree search. This is because with non-deterministic games, that contain random elements (i.e. random enemies), the states in the tree will change every frame, so the outcome of the random action will be different on every search.

sample GA The sample Genetic Algorithm(GA) controller is a rolling horizon open loop implementation for a minimalistic steady state genetic algorithm, known as microbial GA [53], [10]. A tournament takes place between two players and the loser of the tournament is mutated randomly, with the probability of 1/7. Then certain parts of its genome are recombined with parts from the winners genome, with the probability of 0.1 [10]. This repeats until the time budget has been used. The evaluation function is the same as the sampleMCTS controller. The sample GA controller came 12th in the competition [10].

Random This is a very simple controller that is provided with the GVG-AI framework, it simply returns a random action at each game cycle. The random controller came 14th in the competition, which is quite surprising as it managed to beat quite a few of the other, more complicated controllers.

The reason for this may be due to more complicated controllers not choosing an action because it isn't able to search enough of the game space, so making a random action is often better than no action.

OSLA One Step Look Ahead is another rather simple controller, it evaluates the position using the same heuristic as Sample MCTS and selects the action with the highest evaluation score, then moves the model one step ahead using the forward model. This controller came in 16th place [10].

IV. Methodology

The simulations were ran on the 2016 variation of the competition framework, which can be downloaded from the official GitHub repository [3].

The framework was modified to be able to collect the data from the different tree search controllers.

A. Hypothesis

Refactor hypothesis

Hypothesis One: MCTS has a higher win rate compared to BrFS and BeFS on large deterministic games **Null Hypothesis:** MCTS has a higher win rate compared to BrFS and BeFS on small stochastic games

Hypothesis One: MCTS has a higher win rate compared to BrFS and BeFS on larger stochastic games **Null Hypothesis:** MCTS has a lower win rate compared to BrFS and BeFS on larger stochastic games

Hypothesis Two: MCTS has a higher win rate compared to BrFS and BeFS on smaller maps **Null Hypothesis:** MCTS has a lower win rate compared to BrFS and BeFS on smaller maps

Hypothesis Three: BrFS has a higher win rate compared to MCTS and BeFS on larger maps **Null Hypothesis:** BrFS has a lower win rate compared to MCTS and BeFS on larger maps

Hypothesis Four: BrFS has a higher win rate compared to MCTS and BeFS on smaller maps **Null Hypothesis:** BrFS has a lower win rate compared to MCTS and BeFS on smaller maps

Hypothesis Five: BeFS has a higher win rate compared to MCTS and BrFS on larger maps **Null Hypothesis:** BeFS has a lower win rate compared to MCTS and BrFS on larger maps

Hypothesis Six: BeFS has a higher win rate compared to MCTS and BrFS on smaller maps **Null Hypothesis:** BeFS has a lower win rate compared to MCTS and BrFS on smaller maps

B. Research philosophy

When researching, a positivist approach [54], [55] was taken to this project. This means that the researcher was independent to the research being carried out, and maintained an objective stance.

[3] https://github.com/EssexUniversityMCTS/gvgai

C. Tools Used

For carrying out the research on the gvg-AI competition, IntelliJ IDEA 2017.3 was used. RStudio Version 1.1.419. The simulations were compiled and executed with java jdk 9, and simulated on ubuntu 16.04. Pycharm 2018.1.2 with python 3.6 was used to create the simulation launcher.

The computer used to carry out the simulations had Intel(R) Xeon(R) CPU E5-2650 24core @ 2.20GHz - 128GB RAM, 1.1TB SSD.

D. Choosing the Games

There are a lot of games within the GVG-AI framework, and most have multiple levels, so testing with all the different possible configurations is prohibitively expensive, thus a subset of games within the framework will be selected in a way that best represent the framework as a whole.

There are papers by previous authors that looked into finding a good selection of games that can be used for benchmarking the framework. [56]

In Nelsons paper [48], he presented a large scale analysis of MCTS with 62 games in the framework, which were sorted based on performance. Bontrager et al. [57] used clustering techniques on 49 games in the goal to obtain rough groups of similar games. The final 20 games that are used in this paper are based of the paper by Gaina et al. [56] where they uniformly sampled from both Nelson and Bontrager et al. papers and balanced a set of 10 stochastic and 10 deterministic games, which can be shown in table I.

Deterministic Games	Stochastic Games
Bait	Aliens
Chase	Chopper
Hungry Birds	Digdug
Missile Command	Intersection
Plaque Attack	Seaquest
Camel Race	Butterflies
Escape	Crossfire
Lemmings	Infection
Modality	Roguelike
Wait For Breakfast	Survive Zombies

TABLE I
TABLE OF GAMES SELECTED BASED OF [58], [56].

E. Choosing the Tree Search Algorithms

In this paper, a total of 6 tree search algorithms were chosen to compare.

Best First Search

This algorithm chosen because it is used in one of the best controllers submitted to this competition, as mentioned in section III-B8.

Breadth First Search

This tree search method was used because it is one of the most basic tree search algorithms, as it essentially brute-force searches the tree, and can be used as a benchmark for the other tree searching algorithms. In this paper, breadth first search starts a new tree search at each game tick.

Breadth First Search Extended

This is the same as the other breadth first search, with the only difference is that it continues the search throughout the game, so the agent will sit idle until the tree search finds a solution.

MCTS

The sample MCTS controller is the vanilla MCTS algorithm as mentioned in section III-B2, this was chosen because it has been one of the most successful algorithms in the competition, and has been used widely within the competition [10], [40]

Flat MCTS

Sample Flat MCTS is a monte carlo tree search simulation where actions at the current state are uniformly sampled and no tree is built [7], this was chosen as it is interesting to compare how the different sample MCTS controllers perform.

Open Loop MCTS

The Sample Open Loop variation of MCTS (OLMCTS) was also chosen to compare against the other sample MCTS variations.

Running G*Power with a Effect size of 0.05 and a type I error rate of 0.05, and type II error rate of 95% the sample size required for each game is 4331.

F. Design of Data Collection

When running the simulation of all the games in the framework, there were a few issues that arose, the main issue being that when running roughly a total of 600,000 games, $(1000(simulations) \cdot 5(levels) \cdot 20(games) \cdot 6(controllers) = 600,000)$ the java executable would run out of memory. To solve this a simulation launcher was developed where it would run a set amount of games concurrently, until they finished executing, then another one from the list of games would be added to the queue. This means that each java executable ran only 5000 games, compared to 1.5-2M. Moreover, the python launcher program is a lot more scalable in comparison. Typically the total games to concurrently simulate is set to the core count of the processor.

Because this solution is a lot more salable, the results were gathered using a cloud computing solution which allowed many simulations to run in parallel, for long periods without any interference.

As the data collection class outputs a json file, which was intended to be used for nested data structures, however only ended up only outputting flat json, thus a simple json to csv converter was developed in python to be used for importing large datasets into R. This specific converter can be located in the Github directory for this project [4].

To gather the data about the different search trees, I implemented a class that would visualise the time an agent spent in each cell. This gathered insights into the movement patterns within the game, to help analyze if / where the agents got stuck in a level. An example of this can be shown in figure 5

G. Implementation of tree search algorithms

There are two implementations of the breadth first search algorithm in this paper, one which was modified from the *number27* controller, and a basic one I wrote. The one that was modified from number27's controller is one that expands the game tree search over several game ticks and is named Breadth First Search Extended (BrFSE) in this paper, where as the one I wrote creates a new tree search every tick of the game and is named Breadth First Search (BrFS).

The MCTS controller implementations are provided by the competition, as described in section III-D3. The best first search algorithm was extracted and modified from the 2016 ICELab controller.

These algorithms will record various stats about the game they will play, such as; *score, level size, percentage of level explored, total cells visited* and *end game tick*.

H. Chosen Development Life Cycle

The software artifact for this project was developed using the agile approach to software development [59], [60], [61], this is a method of software development where the software is developed and tested incrementally, in small iterations. The reason why this development methodology was chosen is because this project can be easily broken into small sections and tested iteratively.

Fig. 5. Tree search locality represented in by green cells

V. Results

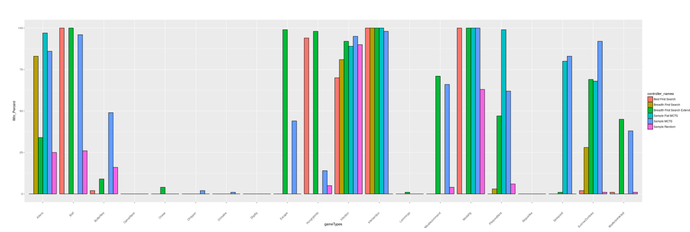

Fig. 6. The win rate of all tree search controllers on the chosen 20 games of the competition

Figure 6 shows the win percentage of all the games being tested, this shows that all the different tree search techniques generally struggle with similar games, of which are both deterministic and stochastic games. In figure 6, only the first level of each game was tested.

[4]https://github.com/Alli1223/comp320-comp360-dissertation

One notable finding is how well the random controller did in both infection and modality games, with a $> 50\%$ winrate on those games. This may suggest that those games may be not too challenging compared to the other games selected.

All the graphs were compiled using R, figure 7 shows how the large data sets were loaded into R. As this is a large image, the full size image can be viewed in the git repository for this project [5].

```
> list.filenames<-list.files(pattern=".csv$")
> list.data<-list()
> for (i in 1:length(list.filenames))
+ {
+     list.data[[i]]<-read.csv(list.filenames[i])
+ }
> names(list.data)<-list.filenames
> list.data$controllers.singlePlayer.bestFirstSearch.Agent_0.csv$Win
```

Fig. 7. R code snippet used for loading large game data sets

VI. Possible Issues

VII. Future Work

The experiments in this paper only used 20 of the games in the competition, when there are currently 82 games in the competition, so comparing all the games in the competition may lead to more interesting results. There are also quite a few other tree searching algorithms that are not evaluated in this paper, as this paper only analysed 4-5 when there are a lot of other algorithms that have not been looked at due to the project scope.

VIII. Conclusion

Re-do this

In conclusion, the literature covered in this review provides an overview of a few of the current algorithms in use in the GVG-AI competition. Based on the performance of the different search algorithms within the framework, a conclusion will be drawn between the superior algorithms for certain games. The GVG-AI competition provides a good framework to compare different tree search algorithms. The aim of this paper will be to create an agent for the GVG-AI competition based of the strengths and limitations of tree search algorithms.

Appendix A
Acknowledgements

TODO

Appendix B
Reflective Report

A. Issues with collecting the data

One of the big issues I faced when I initially started collecting the data was being able to run many games at one time, and as each game takes on average about 1-2minutes to complete, when there are roughly 600,000 games to simulate, if I was to run one game at a time, it would have taken about 1.14 years (416 days) to gather all my data. This is why the creation of the simulation launcher help so much, the only issue was that it was only created 2 weeks before the deadline, so I ended up having to spin up 4 large servers, that had a total of 56 processing cores and 224GB of RAM, which was able to parallelise the data gathered. This means that the data only took about a week to gather, however this still meant that I needed to analyse the data and write it up in only a few days. In hindsight, I should have created the simulation launcher months ago. I focused too much on gathering a large data set for statistical accuracy, when this wasn't necessary, only a small amount of games actually needed to be simulated to gather an insight to what the controllers stats were.

B. Not having a defined hypothesis before collecting the data

When I started collecting the data about the different tree searching algorithms I did not have a defined objective that I wanted to prove, this meant that at the start of the development of the artifact I spent a while working on some visualisation stuff that were not relevant to proving my hypotheses.

[5] https://github.com/Alli1223/comp320-comp360-dissertation/blob/Development/Dissertation/images/allGameBarPlot.png

References

[1] A. Mendes, J. Togelius, and A. Nealen, "Hyper-heuristic general video game playing," in *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, 2016, pp. 1–8.

[2] H. Horn, V. Volz, D. Pérez-Liébana, and M. Preuss, "Mcts/ea hybrid gvgai players and game difficulty estimation," in *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, 2016, pp. 1–8.

[3] M. Campbell, A. J. Hoane, Jr., and F.-h. Hsu, "Deep blue," *Artif. Intell.*, vol. 134, no. 1-2, pp. 57–83, Jan. 2002.

[4] C. E. Shannon, "Programming a computer for playing chess," in *Computer chess compendium*. Springer, 1988, pp. 2–13.

[5] F.-h. Hsu, M. S. Campbell, and A. J. Hoane, Jr., "Deep blue system overview," in *Proceedings of the 9th International Conference on Supercomputing*, ser. ICS '95. New York, NY, USA: ACM, 1995, pp. 240–244.

[6] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.*, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[7] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of monte carlo tree search methods," *IEEE Transactions on Computational Intelligence and AI in games*, vol. 4, no. 1, pp. 1–43, 2012.

[8] S. Ontanón, G. Synnaeve, A. Uriarte, F. Richoux, D. Churchill, and M. Preuss, "A survey of real-time strategy game ai research and competition in starcraft," *IEEE Transactions on Computational Intelligence and AI in games*, vol. 5, no. 4, pp. 293–311, 2013.

[9] P. Hingston, "A new design for a turing test for bots," in *Computational Intelligence and Games (CIG), 2010 IEEE Symposium on*. IEEE, 2010, pp. 345–350.

[10] D. Pérez-Liébana, S. Samothrakis, J. Togelius, T. Schaul, S. M. Lucas, A. Couëtoux, J. Lee, C.-U. Lim, and T. Thompson, "The 2014 general video game playing competition," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 8, no. 3, pp. 229–243, 2016.

[11] T. Joppen, M. Moneke, N. Schroder, C. Wirth, and J. Furnkranz, "Informed hybrid game tree search for general video game playing," *IEEE Transactions on Computational Intelligence and AI in Games*, 2017.

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

[13] D. Pérez-Liébana, J. Dieskau, M. Hunermund, S. Mostaghim, and S. Lucas, "Open loop search for general video game playing," in *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*. ACM, 2015, pp. 337–344.

[14] R. Coulom, "Efficient selectivity and backup operators in monte carlo tree search," in *International conference on computers and games*. Springer, 2006, pp. 72–83.

[15] L. Kocsis and C. Szepesvári, "Bandit based monte carlo planning," in *ECML*, vol. 6. Springer, 2006, pp. 282–293.

[16] G. Chaslot, J.-T. Saito, B. Bouzy, J. Uiterwijk, and H. J. Van Den Herik, "Monte carlo strategies for computer go," in *Proceedings of the 18th BeNeLux Conference on Artificial Intelligence, Namur, Belgium*, 2006, pp. 83–91.

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

[18] T. Pepels, M. H. Winands, and M. Lanctot, "Real-time monte carlo tree search in ms pac-man," *IEEE Transactions on Computational Intelligence and AI in games*, vol. 6, no. 3, pp. 245–257, 2014.

[19] P.-A. Coquelin and R. Munos, "Bandit algorithms for tree search," *arXiv preprint cs/0703062*, 2007.

[20] I. Bravi, "Evolving uct alternatives for general video game playing," 2017.

[21] S. Gelly, Y. Wang, O. Teytaud, M. U. Patterns, and P. Tao, "Modification of uct with patterns in monte carlo go," 2006.

[22] T. Schuster, "Mcts based agent for general video games," Ph.D. dissertation, Masters thesis, Department of Knowledge Engineering, Maastricht University, Maastricht, the Netherlands, 2015.

[23] D. Pérez-Liébana, E. J. Powley, D. Whitehouse, P. Rohlfshagen, S. Samothrakis, P. I. Cowling, and S. M. Lucas, "Solving the physical traveling salesman problem: Tree search and macro actions," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 6, no. 1, pp. 31–45, 2014.

[24] D. Silver and G. Tesauro, "Monte carlo simulation balancing," in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009, pp. 945–952.

[25] E. J. Powley, P. I. Cowling, and D. Whitehouse, "Information capture and reuse strategies in monte carlo tree search, with applications to games of hidden information," *Artificial Intelligence*, vol. 217, pp. 92–116, 2014.

[26] H. Park and K.-J. Kim, "Mcts with influence map for general video game playing," in *Computational Intelligence and Games (CIG), 2015 IEEE Conference on*. IEEE, 2015, pp. 534–535.

[27] D. Pérez-Liébana, S. Samothrakis, and S. Lucas, "Knowledge-based fast evolutionary mcts for general video game playing," in *Computational Intelligence and Games (CIG), 2014 IEEE Conference on*. IEEE, 2014, pp. 1–8.

[28] E. İlhan and A. Ş. Etaner-Uyar, "Monte carlo tree search with temporal-difference learning for general video game playing," in *Computational Intelligence and Games (CIG), 2017 IEEE Conference on*. IEEE, 2017, pp. 317–324.

[29] M. de Waard, D. M. Roijers, and S. C. Bakkes, "Monte carlo tree search with options for general video game playing," in *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, 2016, pp. 1–8.

[30] F. Frydenberg, K. R. Andersen, S. Risi, and J. Togelius, "Investigating mcts modifications in general video game playing," in *Computational Intelligence and Games (CIG), 2015 IEEE Conference on*. IEEE, 2015, pp. 107–113.

[31] R. D. Gaina, S. M. Lucas, and D. Pérez-Liébana, "Rolling horizon evolution enhancements in general video game playing," in *Computational Intelligence and Games (CIG), 2017 IEEE Conference on*. IEEE, 2017, pp. 88–95.

[32] D. Pérez-Liébana, S. Samothrakis, S. Lucas, and P. Rohlfshagen, "Rolling horizon evolution versus tree search for navigation in single-player real-time games," in *Proceedings of the 15th annual conference on Genetic and evolutionary computation*. ACM, 2013, pp. 351–358.

[33] M. M. Flood, "The traveling-salesman problem," *Operations Research*, vol. 4, no. 1, pp. 61–75, 1956.

[34] R. D. Gaina, J. Liu, S. M. Lucas, and D. Pérez-Liébana, "Analysis of vanilla rolling horizon evolution parameters in general video game playing," in *European Conference on the Applications of Evolutionary Computation*. Springer, 2017, pp. 418–434.

[35] T. Back, *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms*. Oxford university press, 1996.

[36] M. Črepinšek, S.-H. Liu, and M. Mernik, "Exploration and exploitation in evolutionary algorithms: a survey," *ACM Computing Surveys (CSUR)*, vol. 45, no. 3, p. 35, 2013.

[37] S. M. Lucas, S. Samothrakis, and D. Pérez-Liébana, "Fast evolutionary adaptation for monte carlo tree search," in *European Conference on the Applications of Evolutionary Computation*. Springer, 2014, pp. 349–360.

[38] D. E. Knuth and R. W. Moore, "An analysis of alpha-beta pruning," *Artificial intelligence*, vol. 6, no. 4, pp. 293–326, 1975.

[39] S. Ito, Z. Guo, C. Y. Chu, T. Harada, and R. Thawonmas, "Efficient implementation of breadth first search for general video game playing," in *Consumer Electronics, 2016 IEEE 5th Global Conference on*. IEEE, 2016, pp. 1–2.

[40] D. Perez-Liebana, J. Liu, A. Khalifa, R. D. Gaina, J. Togelius, and S. M. Lucas, "General video game ai: a multi-track framework for evaluating agents, games and content generation algorithms," *arXiv preprint arXiv:1802.10363*, 2018.

[41] E. K. Burke, M. Gendreau, M. Hyde, G. Kendall, G. Ochoa, E. Özcan, and R. Qu, "Hyper-heuristics: A survey of the state of the art," *Journal of the Operational Research Society*, vol. 64, no. 12, pp. 1695–1724, 2013.

[42] E. K. Burke, M. Hyde, G. Kendall, G. Ochoa, E. Özcan, and J. R. Woodward, "A classification of hyper-heuristic approaches," in *Handbook of metaheuristics*. Springer, 2010, pp. 449–468.

[43] N. Shaker, J. Togelius, G. N. Yannakakis, L. Poovanna, V. S. Ethiraj, S. J. Johansson, R. G. Reynolds, L. K. Heether, T. Schumann, and M. Gallagher, "The turing test track of the 2012 mario ai championship: entries and evaluation," in *Computational Intelligence in Games (CIG), 2013 IEEE Conference on*. IEEE, 2013, pp. 1–8.

[44] M. Genesereth, N. Love, and B. Pell, "General game playing: Overview of the aaai competition," *AI magazine*, vol. 26, no. 2, p. 62, 2005.

[45] N. Love, T. Hinrichs, D. Haley, E. Schkufza, and M. Genesereth, "General game playing: Game description language specification," 2008.

[46] D. Pérez-Liébana, "The general video game ai competition," http://http://www.gvgai.net/, 2017.

[47] T. Schaul, "An extensible description language for video games," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 6, no. 4, pp. 325–331, 2014.

[48] M. J. Nelson, "Investigating vanilla mcts scaling on the gvg-ai game corpus," in *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*. IEEE, 2016, pp. 1–7.

[49] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, "The arcade learning environment: An evaluation platform for general agents." *J. Artif. Intell. Res.(JAIR)*, vol. 47, pp. 253–279, 2013.

[50] 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, pp. 529–533, 2015.

[51] R. D. Gaina, D. Pérez-Liébana, and S. M. Lucas, "General video game for 2 players: framework and competition," in *Computer Science and Electronic Engineering (CEEC), 2016 8th*. IEEE, 2016, pp. 186–191.

[52] D. Pérez-Liébana, S. Samothrakis, J. Togelius, S. M. Lucas, and T. Schaul, "General video game ai: Competition, challenges and opportunities," in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[53] I. Harvey, "The microbial genetic algorithm," in *European Conference on Artificial Life*. Springer, 2009, pp. 126–133.

[54] B. H. Hanne Andersen, "Scientific method," [online] https://plato.stanford.edu/entries/scientific-method/, 2015.

[55] F. Crossan, "Research philosophy: towards an understanding," *Nurse Researcher (through 2013)*, vol. 11, no. 1, p. 46, 2003.

[56] R. D. Gaina, S. M. Lucas, and D. Pérez-Liébana, "Population seeding techniques for rolling horizon evolution in general video game playing," in *Evolutionary Computation (CEC), 2017 IEEE Congress on*. IEEE, 2017, pp. 1956–1963.

[57] P. Bontrager, A. Khalifa, A. Mendes, and J. Togelius, "Matching games and algorithms for general video game playing," in *Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2016, pp. 122–128.

[58] C. Guerrero-Romero, A. Louis, and D. Perez-Liebana, "Beyond playing to win: Diversifying heuristics for gvgai," in *Computational Intelligence and Games (CIG), 2017 IEEE Conference on*. IEEE, 2017, pp. 118–125.

[59] K. Beck, M. Beedle, A. Van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries *et al.*, "Manifesto for agile software development," 2001.

[60] A. Cockburn and J. Highsmith, "Agile software development, the people factor," *Computer*, vol. 34, no. 11, pp. 131–133, 2001.

[61] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and software technology*, vol. 50, no. 9-10, pp. 833–859, 2008.