Using Stochastic Beam Searching to Play RISK

Jacob Gillenwater
Department of Computing
East Tennessee State University
Johnson City, TN, USA
gillenwaterj@etsu.edu

***Abstract*—To play the classic board game RISK, Stochastic Beam Searching produces agents capable of optimal moves based on the current game state. Across 1500 double-elimination tournaments, the agents learned to play aggressively while balancing unnecessary risk. The agents preferred to expand into new continents but often did a poor job defending their territories from outside attack, instead choosing to recapture the territory on their next turn. The agents were often able to choose an optimal move for the given game state, but they often lacked the contextual awareness of future game states. This lack of contextual understanding occasionally leads to some short-sighted and sub-optimal strategies.**

Keywords—Stochastic Beam Search, Genetic Algorithm, Artificial Intelligence, RISK, Board Game

I. Introduction

RISK is a complex turn-based grand strategy game where multiple players compete to control countries on a world map [1]. RISK was created as a board game by Albert Lamorisse in 1957, then published in the United States by Parker Brothers in 1959 [2]. To play RISK, each player takes a turn in multiple phases: An army placement phase, an attack phase, and a movement phase. Attack success is determined by random dice rolls, with some additional rules that favor the attacker. The roll outcomes and rules are well-defined enough to predict the attacks' results within some degree of probability [3]. This project uses a stochastic beam searching algorithm to simulate thousands of games to determine optimal strategies.

Some game rules have been modified from the original version to reduce randomness in the simulation and simplify the implementation. The first change is the complete removal of game cards. Game cards provide additional randomness to the original game that gives players random bonuses for certain card combinations. Some game variants already exist that do not use the game cards, so this change is reasonable [1] [4]. The second change requires attacking territories to attack with the maximum number of armies available when attacking. This change greatly simplifies the implementation and is widely regarded as a good strategy anyway [4] [5]. The final rule change imposes a 100-turn limit on each player to prevent stalemate games and those requiring an exceptionally high turn count to complete. This change keeps game simulations relatively short, increasing runtime performance.

A. *How to play RISK*

The game rules provide each player with a predetermined number of armies to position around the game board to set up the game prior to normal play [1]. Next, the player turn order is randomly decided. After establishing the player turn order, players take turns positioning armies on unclaimed territories to claim ownership of those territories until no unclaimed territories remain. Claimed territories are under the claiming player's control at the start of normal play. The players continue taking turns placing armies onto their claimed territories until all armies have been placed. The game is now ready for normal play.

During normal play, a player's turn comes in three distinct phases: army placement, attacking, and movement. The players' turn order continues from the setup portion of the game.

1) Placement Phase

At the start of the placement phase, players are given a calculatable number of armies to place onto controlled territories on the game board. The number of armies that a player receives is the total number of territories controlled by the player divided by three, rounded down to the next whole number, then bonus armies are added. Each player is guaranteed to receive at least three armies even if they control fewer than nine territories. Bonus armies come from controlling all territories in a continent or from cards. The bonus for controlling each continent varies according to Table 1. Armies may be placed individually or in groups onto an arbitrary number of player-controlled territories.

(Table 1)

Continent Control Bonuses

Continent Name	**Bonus Armies Received**
Asia	7
North America	5
Europe	5
Africa	3
South America	2
Australia	2

2) Attack Phase

After placing all armies in the placement phase, players may proceed to the attacking phase of their turn. Players may choose to attack any territory not controlled by that player, provided that the territory is connected to a territory that is controlled by that player. Additionally, players may only attack the territory using the armies placed onto the connected territory, provided at least one army is left behind to keep control of the attacking territory.

Using the modified ruleset in this implementation, the attacker must use all armies on the attacking territory except for the one necessary to keep control of the attacking territory. The attack may continue until either no attacking armies remain or the attacking player decides to end the attack after losing too many armies.

Attacks may be summarized as a series of smaller battles between an individual or a small group of armies, whereby each army represents a collection of random dice rolls. Attacking armies are divided into groups of 3, with any remainder belonging to a separate group. Defending armies, i.e., those coming from the territory being attacked, are divided into groups of 2, with any remainder belonging to a separate group. Groups from the attacker and the defender are paired up to decide individual battles.

To decide battles, fair dice are rolled for each army in each group. The highest pair of dice from each group in a battle are compared, followed by the second highest pair, and ignoring the smallest value from the attacker should it exist. If the attacker's dice roll is higher than the defender's dice roll, then the attacker wins, and the defender's army is removed from play. If the defender's dice roll is higher, then the defender wins, and the attacker's army is removed from play. The result of a tie is the same as the defender winning.

The series of battles may continue until the attacker runs out of armies or decides to stop attacking to preserve the remaining armies. Furthermore, the series of battles may continue until the defender runs out of armies, allowing the player to take control of the territory by moving an arbitrary number of remaining armies from the attacking territory into the newly controlled territory, ensuring that at least one army remains at the attacking territory.

The player may attack an arbitrary number of times to an arbitrary number of territories, from an arbitrary number of territories.

3) *Movement Phase*

Once the player has decided to stop attacking, the player may choose to transfer an arbitrary of armies from any territory in their control to any connecting territory which is also under their control. This transfer of armies between territories must leave at least one army on the supplying territory. Furthermore, this transfer of armies between territories may only occur once per turn.

4) *Game Ending Conditions*

The primary objective of RISK is to gain control of all territories on the game board. Any player who is able to accomplish this goal is the winner of the game and, by definition, the last player with territories under their control at the end of their last turn.

A player has lost and no longer receives turns if they lose control of all their territories and have no territories to place armies in at the start of their next turn.

Following the original ruleset, games may last forever. However, this implementation of RISK sets a 100-turn limit on each player in the game. For example, a game between 4 players may last for up to 400 turns. This rule change introduces the possibility of a game resulting in a tie with multiple players still controlling territories at the end of the game.

To resolve ties, the player in control of the most territories is the winner. If there is still a tie between players, then the winner is the player with the most armies currently in play. If there is still a tie between players, then it is considered a true tie, and the tying players may all claim to be the winners of the game.

II. Approach

A. *Defining The Environment*

To teach a computer to play RISK, I first had to implement a game environment for the computer to interact with. To begin, I defined a map as a set of two collections, continents and territories, in a standard JSON format that my RISK game implementation could interpret to define custom maps. This allowed me to use simplified maps that were easy to test with and allowed me to experiment with alternative maps.

Continents are defined as a simple structure with a string name, an army bonus, and a color used for visually distinguishing continents in a GUI representation of the map. The important values for each continent can be seen in Table 1.

In contrast to continents, territories are much more complex structures. Each territory is composed of an identifying index, a continent name, a set of other territory indices that indicate connections to those territories, an owner name, and an army count. A 2D vector represents a territory's position in a GUI representation of the map. Upon reading in the map data, territories are verified such that each continent name used exists in the set of continents and that all connections are both ways. For example, if Territory A has Territory B's index in its set of connections, then Territory B must include Territory A's index in its set of connections.

B. *Selecting a Machine Learning Algorithm*

To teach a computer how to play RISK using the rules outlined in the introduction section, I opted to use Stochastic Beam Searching. Stochastic Beam Searching is related to the more popular genetic search algorithm, except there is no parental crossover, meaning that it is an asexual process in contrast to the sexual process described in a genetic algorithm [6].

Stochastic Beam Searching allows agents to adjust weighted values to various in-game actions depending on the game state. These values may be adjusted through random mutations. Agents can then compete against one another to determine values that create effective in-game strategies and those that create ineffective in-game strategies. Agents with values that create effective strategies will become the basis for a new generation of player agents, each containing a slight mutation from their basis in the previous generation.

C. *Implementing Stochastic Beam Search*

1) *Implementing Agents*

Stochastic Beam Search is an iterative process of adjusting virtual dials to achieve some optimal outcome. In the context of playing RISK, the dials are a collection of agent characteristics that define how a player may perceive the game state and how that perception affects their actions. Each characteristic is divided into four categories: placement, attack, movement, and preference.

Table 2 contains a detailed description of each characteristic. Each characteristic affects an agent's perception of the game state to help it make informed decisions about what action or actions to take during its turn. Each characteristic is defined by: its category, a numeric value that an agent may reference to make decisions, an adjustment amount that affects the value during mutations between generations of agents, as well as upper and lower bounds. The boundaries, if used, constrain the value between 0 and 1 to represent a percentage. In practice, there is a non-essential description component attached to each characteristic to assist with maintainability and debugging.

(Table 2)
Agent Characteristic Descriptions

Characteristic Name	Description
Placement Category	
Anywhere	Placing an army anywhere
Enemy Adjacent	Placing an army on a territory connected to a territory controlled by a different player
Ally Adjacent	Placing an army on a territory connected to a territory controlled by the same player
Border Adjacent	Placing an army in a territory that borders a country on a different continent
Connection Bias	Placing an army on a territory with connections to multiple other countries, +value per connection
Placement Bias Multiplier	Placing an army where there already are other armies, value^(armies on territory)
Attack Category	
Anywhere	Attacking anywhere
Ally Adjacent	Attacking a territory connected to another territory controlled by the attacking player
Border Adjacent	Attacking a territory on the border of a different continent
Capture Continent	Attacking a territory that will give this player control over all territories on a continent if the attack is successful
Destroy Bias	Estimated amount of defending armies destroyed, +value per army
Remain Bias	Estimated amount of attacking armies destroyed, -value per army
Safe Threshold	Minimal amount of estimated chance of a successful attack to consider an attack safe, below this amount is considered risky
Minimal Success Chance	Minimal amount of estimated chance of successful attack necessary for an attack to be considered viable
Minimal Remaining Percent	The number of armies lost before calling off an attack, expressed as a percentage of the number of armies at the start of the attack
Movement Category	
Anywhere	Moving an army anywhere
Enemy Adjacent	Moving an army on a territory connected to a territory controlled by a different player
Ally Adjacent	Moving an army on a territory connected to a territory controlled by the same player
Border Adjacent	Moving an army in a territory that borders a territory in a different continent
Bigger Territory	Moving armies onto a territory with more armies
Smaller Territory	Moving armies onto a territory with fewer armies
Connection Bias	Moving an army on a territory with connections to multiple other countries, +value per connection
Base Transfer Rate	Base percentage of armies to transfer should it be necessary
Risky Transfer Rate	Percentage of armies to transfer if the movement is considered risky
Safe Transfer Rate	Percentage of armies to transfer if the movement is considered safe
Preference Category	
Larger	Preference to attack larger players
Smaller	Preference to attack smaller players
Risky	Preference for risky actions
Safe	Preference for safe actions

2) Implementing a Selection Process

To determine successful characteristics, agents compete against one another in a double-elimination tournament. A double-elimination tournament was selected to minimize the impact of luck during a single game of RISK. This helps ensure that an agent producing an otherwise effective strategy is not removed from the population because of an unlucky string of random occurrences while attacking or defending.

Initially, 256 agents participated in the tournament, serving as the initial population for the Stochastic Beam Searching algorithm. The characteristics of this initial population are based upon the initialization values of each characteristic shown in Table 3. Each agent mutates with significantly higher than normal mutation rates and adjustment amounts.

Ideally, each game of RISK begins with four players, although less may be used if there are not enough players to divide by four evenly. At the end of each game of RISK, the winners and losers of the game are recorded. If an agent incurs a second loss, then that agent is removed from the tournament.

The tournament continues until 25 percent of the original number of participants remains or less, approximately 64 participants. The remaining participants become the basis for the next generation of agents. When instantiating the next generation of agents, the surviving agents from the previous tournament are automatically transferred to the next generation. Before the next tournament can begin, the population must include 256 agents. To bring the size of the population back up to 256, agents from the previous tournament are randomly copied. Each new copy mutates before joining the new population of agents.

The tournament selection process repeats 1500 times throughout the learning process. Each new group of 256 agents that competes in the tournament represents a single new generation of agents in the context of Stochastic Beam Search.

(Table 3)
Agent Characteristic Values

Characteristic Name	Initial Value	Adj. Amount	Lower Limit	Upper Limit
Placement Category				
Anywhere	5	1	-inf	inf
Enemy Adjacent	5	1	-inf	inf
Ally Adjacent	5	1	-inf	inf
Border Adjacent	5	1	-inf	inf
Connection Bias	1	0.25	-inf	inf
Placement Bias Multiplier	0.85	0.05	0	1
Attack Category				
Anywhere	5	1	-inf	inf
Ally Adjacent	5	1	-inf	inf
Border Adjacent	5	1	-inf	inf
Capture Continent	5	1	-inf	inf
Destroy Bias	1	0.1	-inf	inf
Remain Bias	-1	0.1	-inf	inf
Safe Threshold	0.95	0.05	0	1
Minimal Success Chance	0.5	0.05	0	1
Minimal Remaining Percent	0.1	0.05	0	1
Movement Category				
Anywhere	5	1	-inf	inf
Enemy Adjacent	5	1	-inf	inf
Ally Adjacent	5	1	-inf	inf
Border Adjacent	5	1	-inf	inf
Bigger Territory	5	1	-inf	inf
Smaller Territory	5	1	-inf	inf
Connection Bias	1	0.25	-inf	inf
Base Transfer Rate	0.5	0.05	-inf	inf
Risky Transfer Rate	0.3	0.05	-inf	inf
Safe Transfer Rate	0.7	0.05	-inf	inf
Preference Category				
Larger	1	0.25	-inf	inf
Smaller	1	0.25	-inf	inf
Risky	1	0.25	-inf	inf
Safe	1	0.25	-inf	inf

3) Mutation Strategy

At the lowest level of implementation, individual characteristics handle mutations. There exists a mutation function that adjusts the characteristic's value based on its adjustment amount. Furthermore, this method ensures that the value is constrained within the upper and lower bounds defined by the characteristic. Mutations can adjust the characteristic's value either positively, negatively, or in a random direction if not specified.

At the next highest level of implementation, mutations are controlled by the agent. There are three parameters that affect how mutations are affected at this level: recursive chance, major mutation chance, and mutation multiplier. The recursive chance begins at 80% by default and determines the likelihood of repeating the mutation process. The recursive chance quickly diminishes because each mutation halves the recursive chance. The major mutation chance is the likelihood of mutating not a single characteristic but every mutation in a random category. The major mutation chance is constant throughout the mutation process in contrast to the recursive chance. Lastly, the mutation multiplier is defaulted to 1 and represents some factor that a mutating characteristic will multiply its adjustment amount by. This adjustment is useful because it allows for larger mutations early in the iteration process described by Stochastic Beam Search and smaller adjustments as the learning process continues.

The highest level of mutation implementation occurs at the population level. This level is mostly concerned with the adjustment of the mutation multiplier specified at the agent level of implementation. When initializing the first generation of agents, each new agent is forced to mutate at least ten times. A new agent can potentially have more mutations depending on the random results of the recursive chance. Additionally, each mutation has a mutation multiplier of 2 to ensure that the initial population has a wide range of characteristic values.

At the end of each tournament, the population must replenish itself with new agents. Each of these new agents must have some mutation to distinguish itself from its basis. During this process, the population will adjust the mutation multiplier depending on the generation number, i.e., the number of previous tournaments. The mutation multiplier at this level linearly interpolates between 1.75 and 0.5. For example, when the second generation is being created, the mutation multiplier will be 1.75. When the last generation is created, the mutation multiplier will be 0.5. This gradual change in mutation severity allows for greater experimentation and variety among agents early in the learning process and smaller finer adjustments towards the end of the learning process.

III. Results

The experiment was run multiple times, each providing slightly different results but mainly providing the same general trends. The results from this experiment are shown as a series of graphs found in the appendix section.

A. Placement

Typically, the agents learned to be aggressive by favoring army placement near enemy territories and avoiding territories surrounded by territories controlled by the same player. Surprisingly, the agents learned to avoid placing armies near the border of a continent. This is likely because armies placed on one of these territories are likely to be destroyed because territories on the border of a continent are likely to be attacked.

Additionally, the agents learned to prefer territories that had limited connections to other territories. This is likely because many of the other connections are likely to be controlled by a different player early in the game, making attacks early in the game more common. These early losses may have influenced this change in behavior.

Lastly, the placement bias multiplier, which controls the diminishing returns of placing multiple armies on the same territory, remained remarkably high. This means that the agents preferred to keep their armies concentrated on a few key territories then march this large army across the game map. While this leaves holes in its defenses, as many territories go without additional armies, it appears to be the more successful strategy in this implementation.

B. Attack

The agents showcase their aggressive behavior when attacking by their preference for attacking anywhere. This means that the agents valued attacking every turn and that most territories made for good targets.

Attacking territories that connect to the multiple other territories controlled by the attacking player was somewhat sought-after. This is likely because the opponent has multiple options to consider attacking their turn and is therefore capable of attacking whatever the weakest territory may be.

Agents learned to prefer attacking territories along the borders of continents. This is likely closely related to the aversion to placing armies along continental borders, as armies placed there are quickly attacked and destroyed. The added benefit to this strategy is the slow conquering of continents which provides additional armies for placement on the player's next turn. Controlling these territories may provide the attacker with the bonus armies on their turn at best or spoil the reward for an opponent at worse.

Being aggressive while minimizing unnecessary risk was a difficult challenge for the agents to overcome. By the end of the simulation, the agents showed a slight preference for keeping their own armies intact over destroying the armies of their opponents. Essentially, this shows that the agents learned to favor attacks that gained them territory while opposing attacks that simply weakened their opponents' defenses for a given territory.

The agents learned how to balance risky behavior when deciding to initiate and call off an attack before becoming somewhat more conservative by the end of the simulation. For example, the agents learned to consider an attack safe, and therefore more likely to engage in an attack if their estimated chance of successfully conquering the territory was 75% or better. The minimal estimated chance of conquering a territory, below which an agent will not attack a territory, mainly fluctuated between 20% and 60%. By the end of the simulation, the minimal estimated chance of conquering a territory was approximately 45%.

The change over time to a more conservative playstyle is easily seen by reviewing the minimal remaining percent of an attacking territories army that must still be alive before discontinuing an attack. This value also fluctuated wildly for many of the earliest generations before settling in between 10% and 30%. However, towards the end of the simulation, there is a strong uptick favoring a high minimal remaining percent, reaching as high as 97.5%.

C. Movement

The agents learned how to defend their border through movement. They learned it wasn't a good strategy to move armies to territories surrounded by territories controlled by the same player because those armies were unlikely to be attacked, effectively removing those armies from the game unless an opposing player captured the surrounding territories. Instead, the agents preferred to move armies to territories connected to territories controlled by one or more of their opponents, providing an extra layer of defense to their territory.

Further evidence that movement was the agents' preferred method for bolstering defense is their preference to move into smaller territories rather than larger territories. This is mirrored in the rise of the connection bias that shows a strong preference for territories connected to many other territories. A territory connected to many other territories can find itself defending a multitude of attacks from those other territories. Therefore, that territory needs to be heavily defended to maintain control.

The last observation to make from the movement characteristic category is that all movements used the maximum number of armies allowed by the rules regardless of the tactical situation.

D. Preference

The agents learned that it is better to attack territories whose controlling player controls more territories than themselves. This is likely because players that control many territories can quickly compound their advantage to achieve a quick and decisive victory by benefiting from the additional armies provided by controlling many territories and receiving any continental bonuses.

Lastly, the agents learned to be risk-averse, as seen in other characteristics. There is a significant preference for actions that can be classified as safe behavior than those that qualify as risky behavior.

IV. Conclusions

The agents learned over the course of the simulation the basic strategy for playing RISK. They learned to balance conservative play by minimizing unnecessary risk while maintaining an aggressive posture.

Unfortunately, Stochastic Beam Search did not allow the agents to learn more complex strategies that extended beyond the immediate choices set out before it. This led to some short-sighted behavior, like improperly fortifying continental borders, allowing a continental bonus to be lost. Additionally, the agents did not learn to consistently use combinatorial attacks whereby two or more territories attack a single target territory, a powerful and common technique used by human players.

To improve upon the strategies shown by these agents, I propose a new approach that allows the agents to consider a wider range of potential moves each turn that allow turns to be considered more wholistically—something like the Minimax algorithm described by Peter Norvig [6]. However, the minimax algorithm would need to be adapted for more than 2-players to accommodate a 4-player game of RISK.

Additionally, the characteristics of each agent could be expanded and made more dependent on the game state. For example, each characteristic in the preference category is used in a variety of different ways that make the agents highly sensitive to changes in these values. Even small changes in these drastically alter a wide range of seemingly unrelated behavior, for example, army placement and attack targeting. Smaller and more narrowly defined characteristics that affect single in-game actions should be preferred by future implementations.

V. References

[1] P. Brothers, “RISK: Instructions,” Hasbro, 1993. [Online]. Available: https://www.hasbro.com/common/instruct/risk.pdf. [Accessed 14 Februrary 20201].

[2] K. Veronese, “The Origins and Evolution of the Strategy Board Game RISK,” 30 March 2012. [Online]. Available: https://gizmodo.com/the-origins-and-evolution-of-the-strategy-board-game-ri-5897532. [Accessed 17 September 2021].

[3] N. Berry, “RISK Analysis,” November 2011. [Online]. Available: https://datagenetics.com/blog/november22011/index.html. [Accessed 28 February 2021].

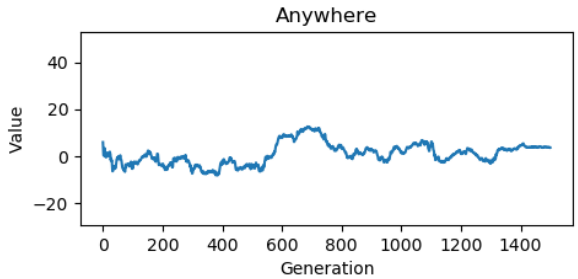
[4] Ultra BoardGames, “Tips to win Risk,” 2020. [Online]. Available: https://www.ultraboardgames.com/risk/tips.php.

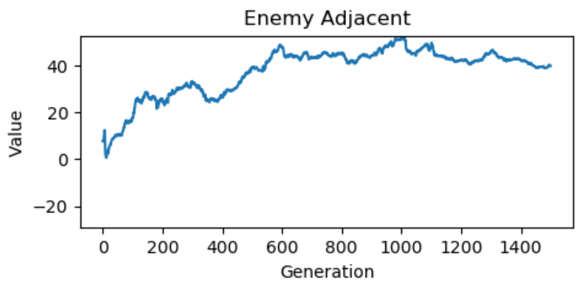
[5] I. Hart, "Risk Combat Statistics," [Online]. Available: http://www.cs.man.ac.uk/~iain/riskstats.php. [Accessed 14 Februrary 2021].

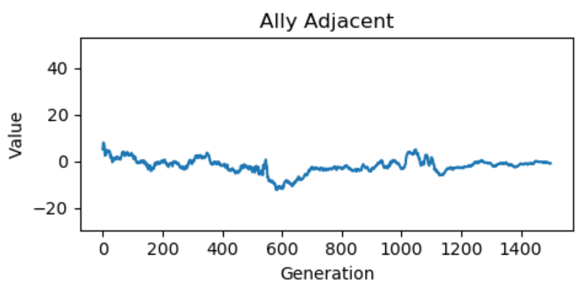
[6] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed., Pearson, 2009.

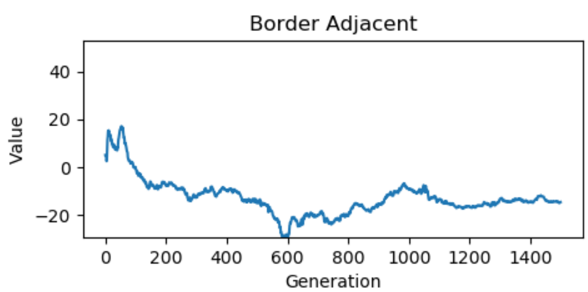
VI. Appendix

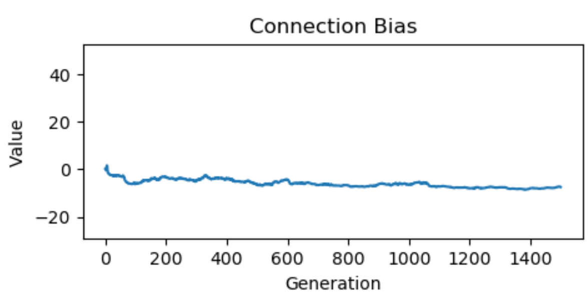
A. *Placement Charts*

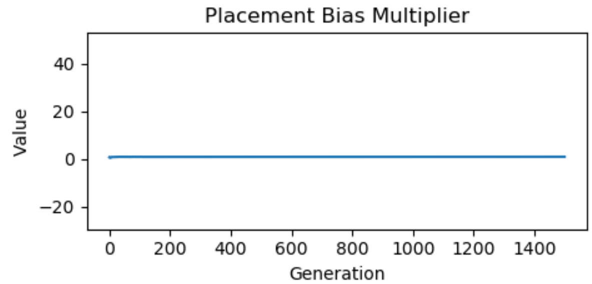




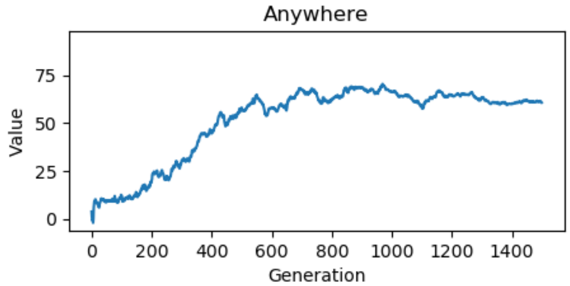


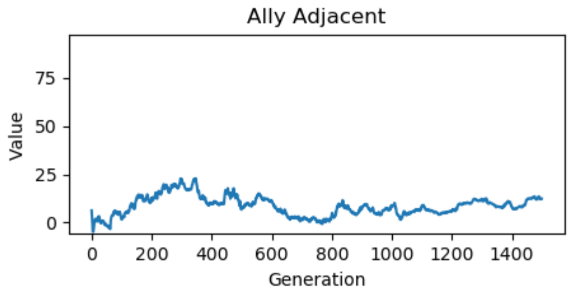


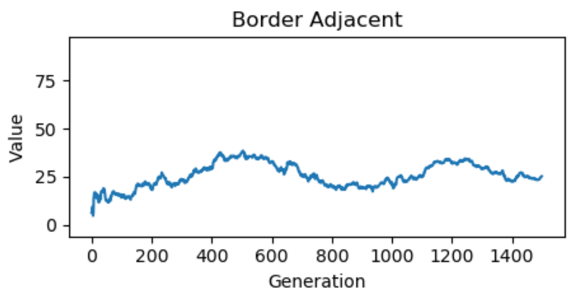


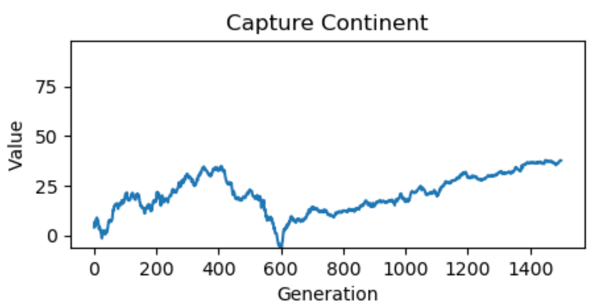


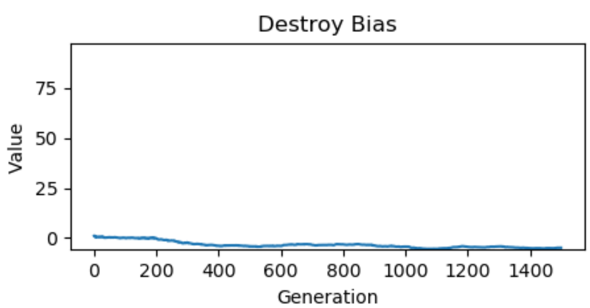
B. Attack Charts

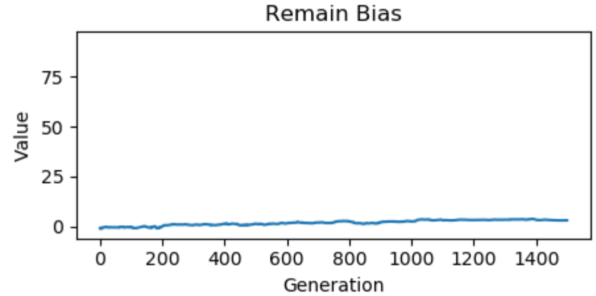


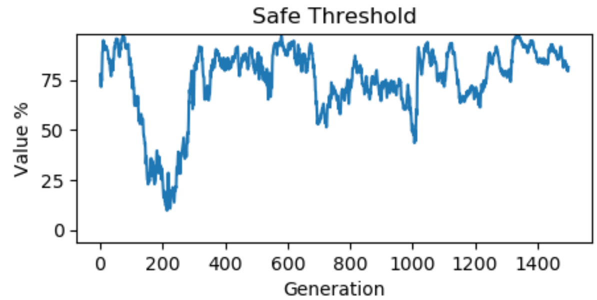


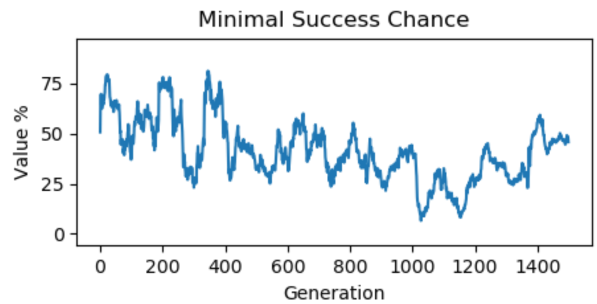


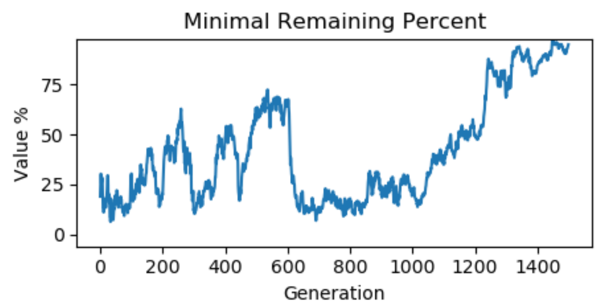




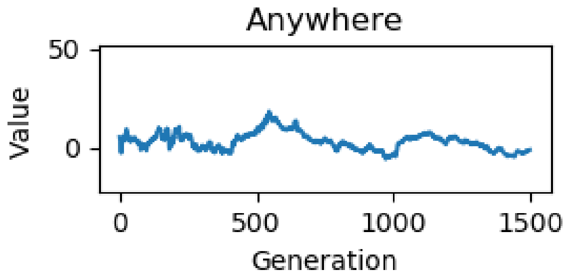


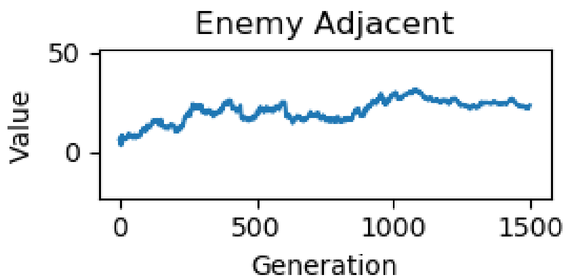


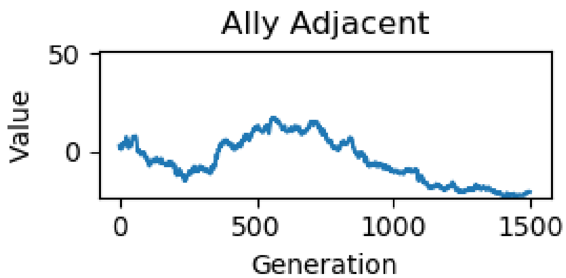


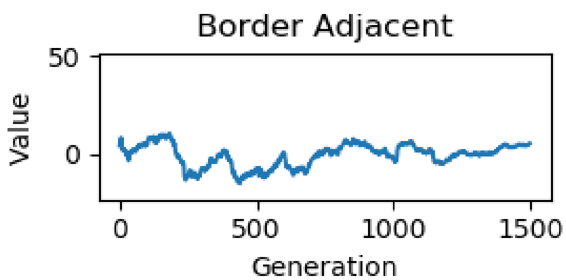


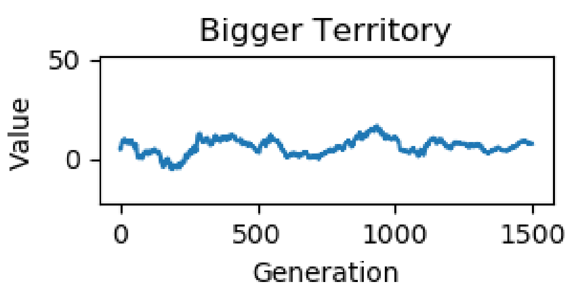
C. Movement Charts

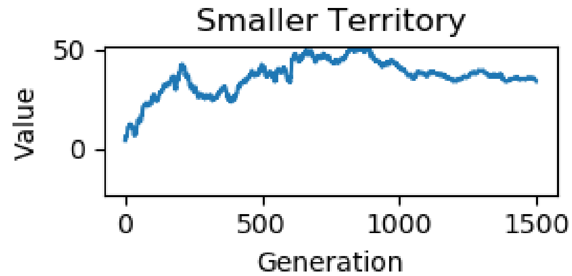


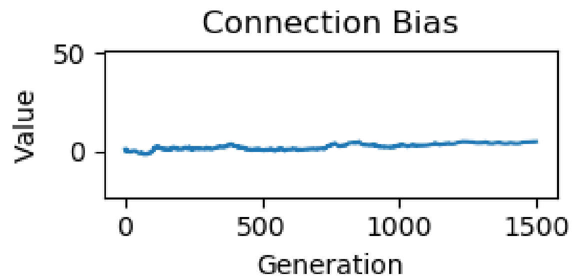


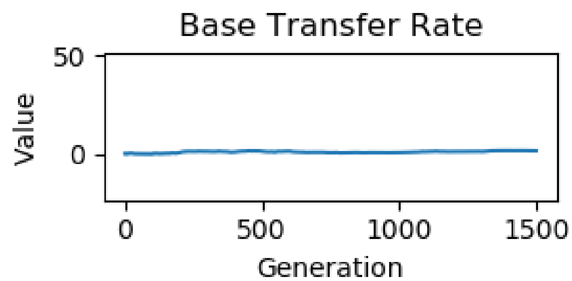


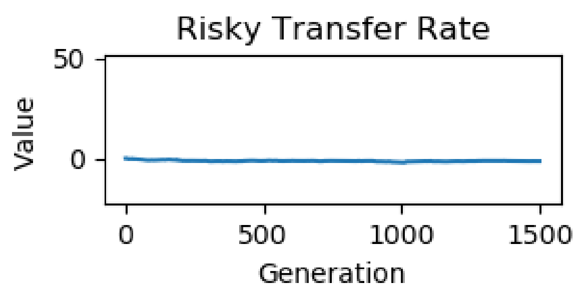


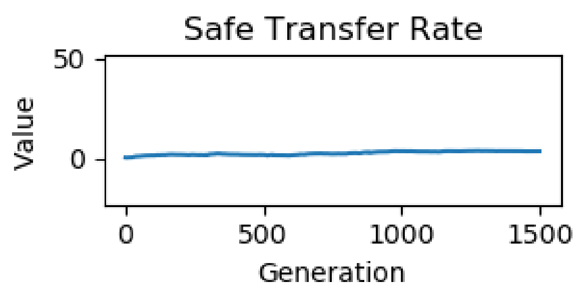




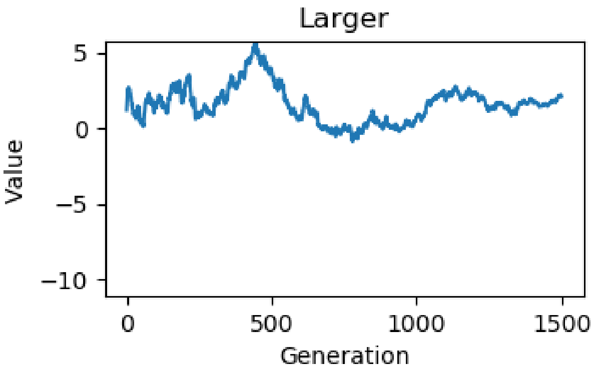


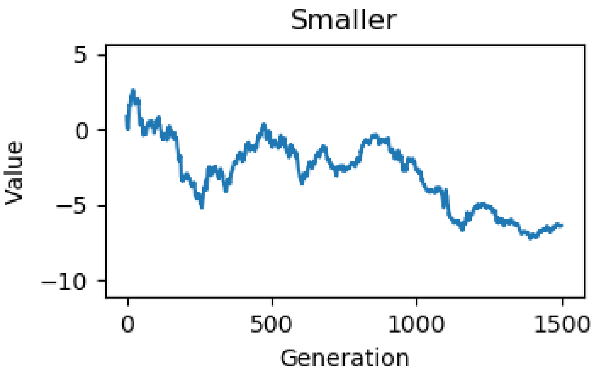


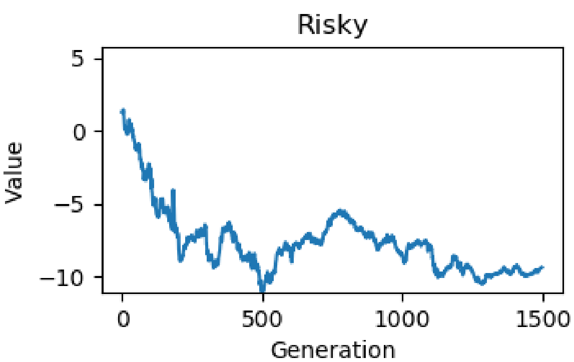


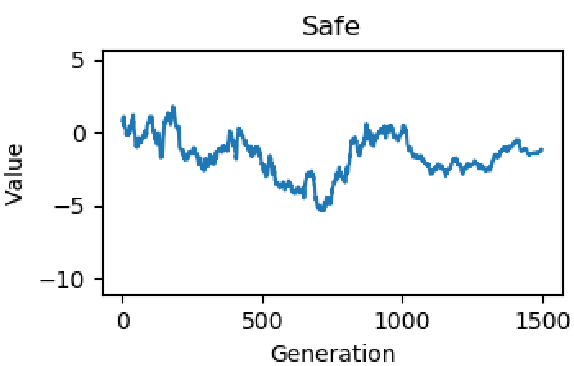


D. Preference Charts









E. Map Data

(Table 4) RISK Territory Connection Data		
Continent Name	**Territory Index**	**Connections by Index**
North America	0	1, 4, 3, 13
North America	1	0, 2, 5, 4
North America	2	1, 5, 26
North America	3	0, 4, 6
North America	4	0, 1, 3, 5, 6, 7
North America	5	1, 2, 4, 7
North America	6	3, 4, 7, 8
North America	7	5, 4, 6, 8
North America	8	6, 7, 9
South America	9	8, 10, 11
South America	10	9, 11, 12, 20
South America	11	9, 10, 12
South America	12	10, 11
Europe	13	0, 14, 17
Europe	14	13, 15, 17, 16
Europe	15	14, 16, 19, 30, 34, 37
Europe	16	14, 17, 15, 18, 19
Europe	17	13, 14, 16, 18
Europe	18	16, 17, 19, 20
Europe	19	15, 16, 18, 20, 21, 37
Africa	20	10, 18, 19, 21, 22, 23
Africa	21	19, 20, 22, 37
Africa	22	20, 21, 23, 24, 25, 37
Africa	23	20, 22, 24
Africa	24	22, 23, 25
Africa	25	22, 24
Asia	26	2, 27, 29, 31, 32
Asia	27	26, 28, 29
Asia	28	27, 29, 30, 32, 33
Asia	29	26, 27, 28, 32
Asia	30	15, 28, 33, 34
Asia	31	26, 32
Asia	32	26, 28, 29, 31, 33
Asia	33	28, 30, 32, 34, 36, 35
Asia	34	15, 30, 33, 36, 37
Asia	35	33, 36, 39
Asia	36	33, 34, 35, 37
Asia	37	15, 19, 21, 22, 34, 36
Australia	38	39, 40, 41
Australia	39	35, 38, 41
Australia	40	38, 41
Australia	41	38, 39, 40