BLACKJACK PLAYER EVOLUTION

THE PROBLEM

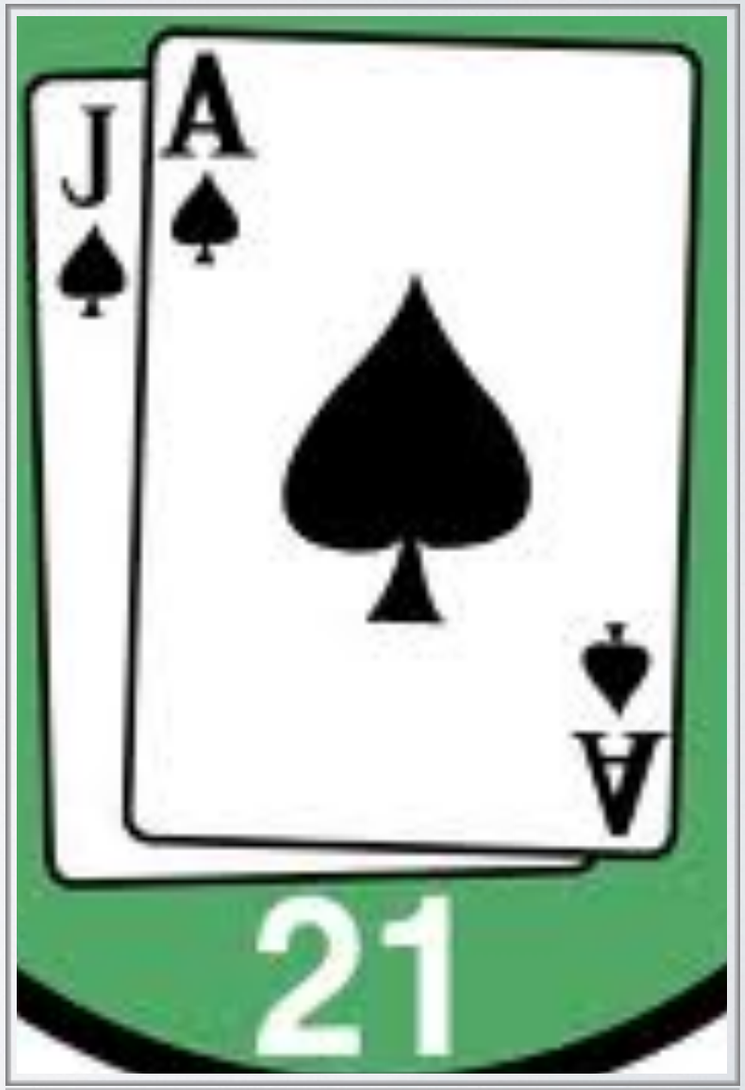
“Can I evolve a Blackjack policy better than the dealer knowing ONLY the goal of the game?”

QUESTIONS

- What does "better" mean?
- How *much* better?

ANSWERS

- Under the same conditions, I win more often
- Let's go for broke! I want to be 100% better!



ENCODING

- **Policies:** 28bit strings representing hit/stand for each of 28 states
- **States:** 4 - 22 (where 22 = BUST, and 12.5 etc. are soft)
- **Simulating Games:** Move to random higher state on hit
- **Fitness:** Who got more wins following his policy at state n

HOW MUCH BETTER CAN WE GET?

EVOLUTION PARAMETERS

- **Inheritance:** Learn from opponent i.e. The Dealer
- **Selection:** Opponent gene must be at least 50% better than mine to inherit; *Elitist Selection*
- **Crossover:** Steal genes from the opponent; *Non-Uniform*
- **Mutation:** Inversely proportional to improvement. Poor children get more mutation; *Controlled Simulated Annealing*
- **Termination:** On specified improvement target (100%)

MORE ON EVOLUTION

- With Elitist Selection & Non-Uniform Crossover, the population ultimately became the # of genes **(28)**, as selection, crossover, mutation, and comparing all took place at the gene level.

- As it turns out, this will still lead you to good results, but getting a good represention of the fitness of a policy becomes difficult.
  For the whole is more than the sum of it parts.

- Increasing the number of simulations to base fitness on, the desired improvement, or the selection criteria, all see improvement to the final result, at the expense of computation time.

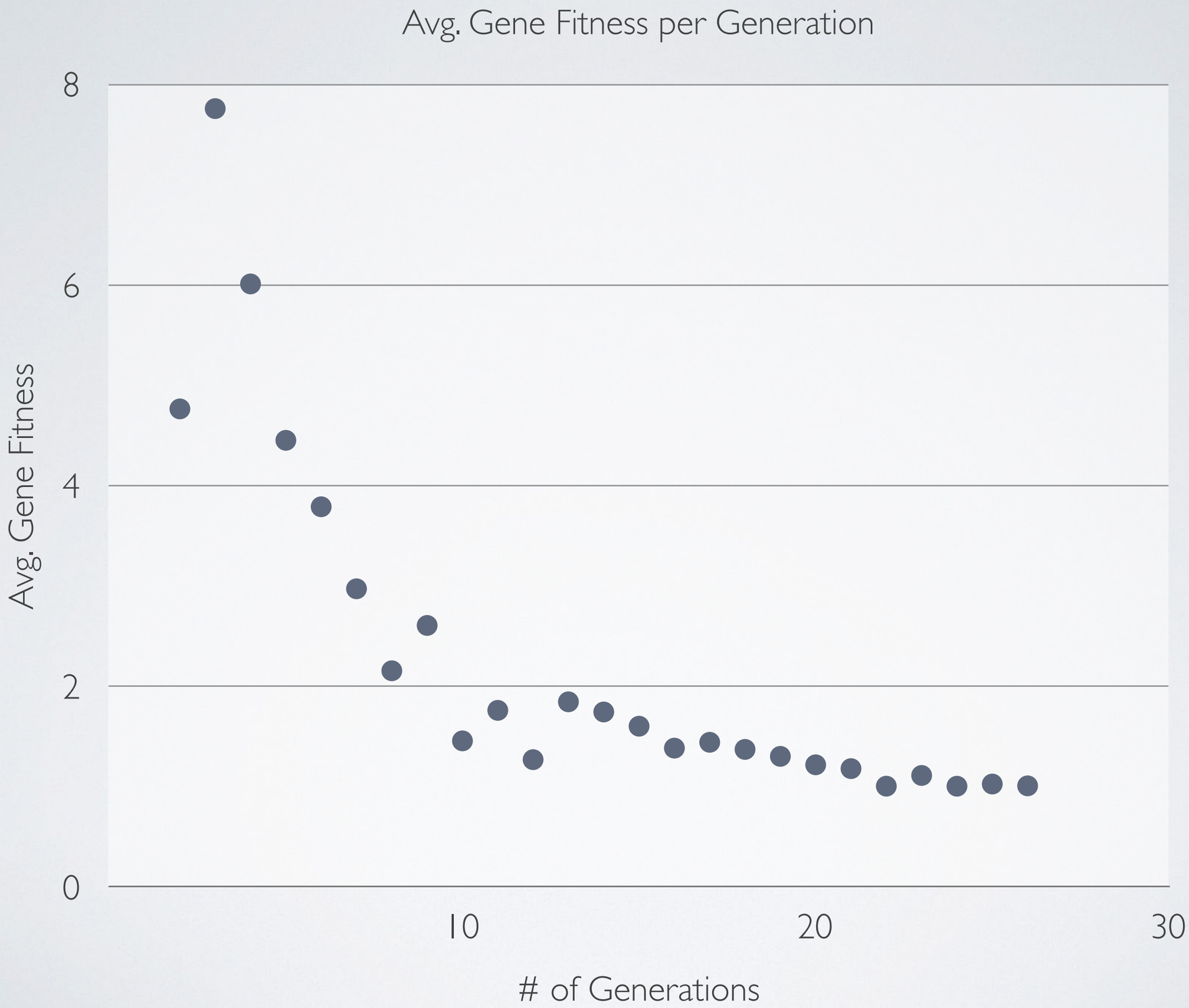
- So reaching a 100% improvement by my standards was easy, but the resulting policy shabby. Better results were seen upwards of 400%. Of course if the dealer had a terrible policy, even reaching 1000% could be done quite quickly.

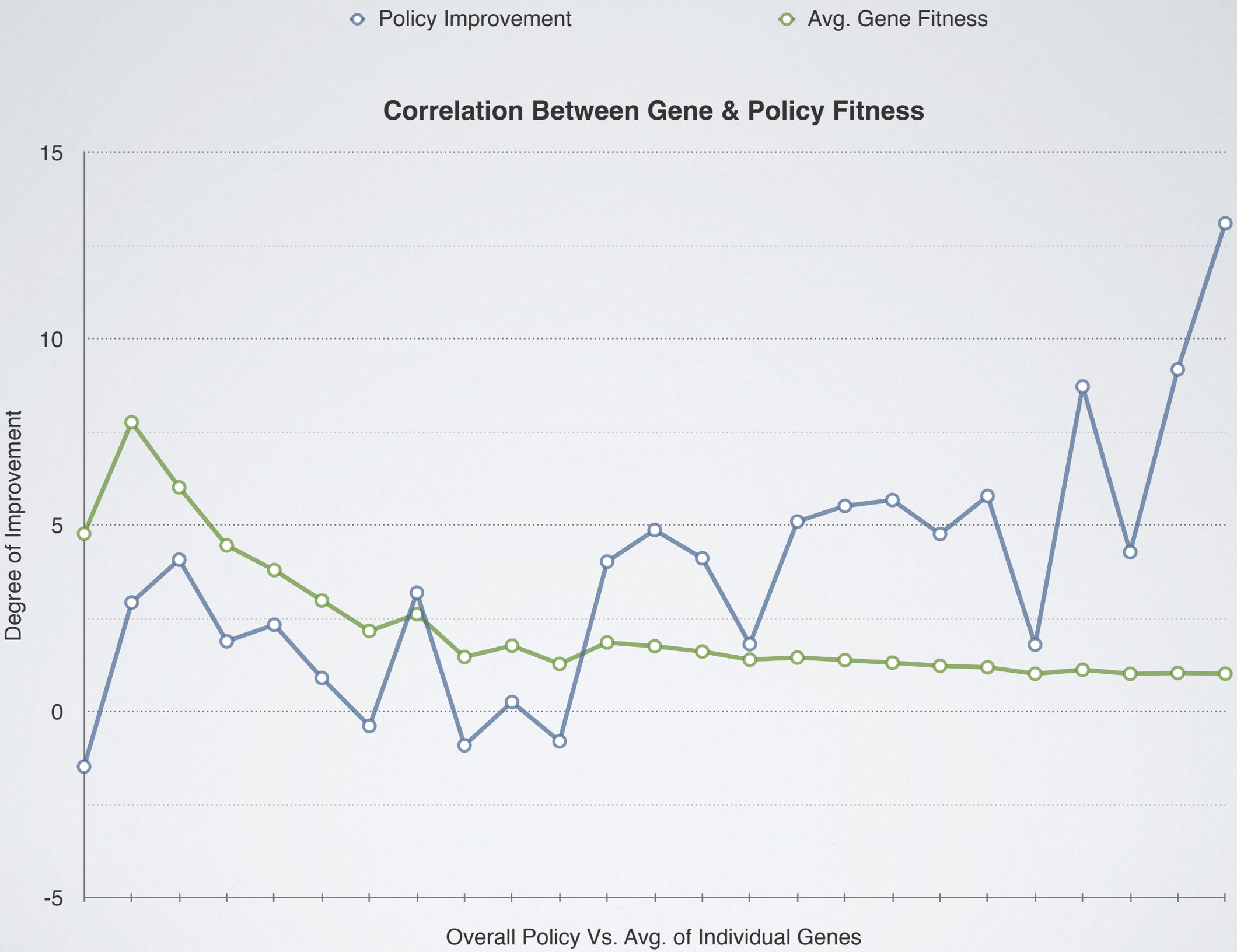
THE FOLLOWING SLIDES EXEMPLIFY MY RESULTS

With a small case-study

CASE STUDY PARAMETERS

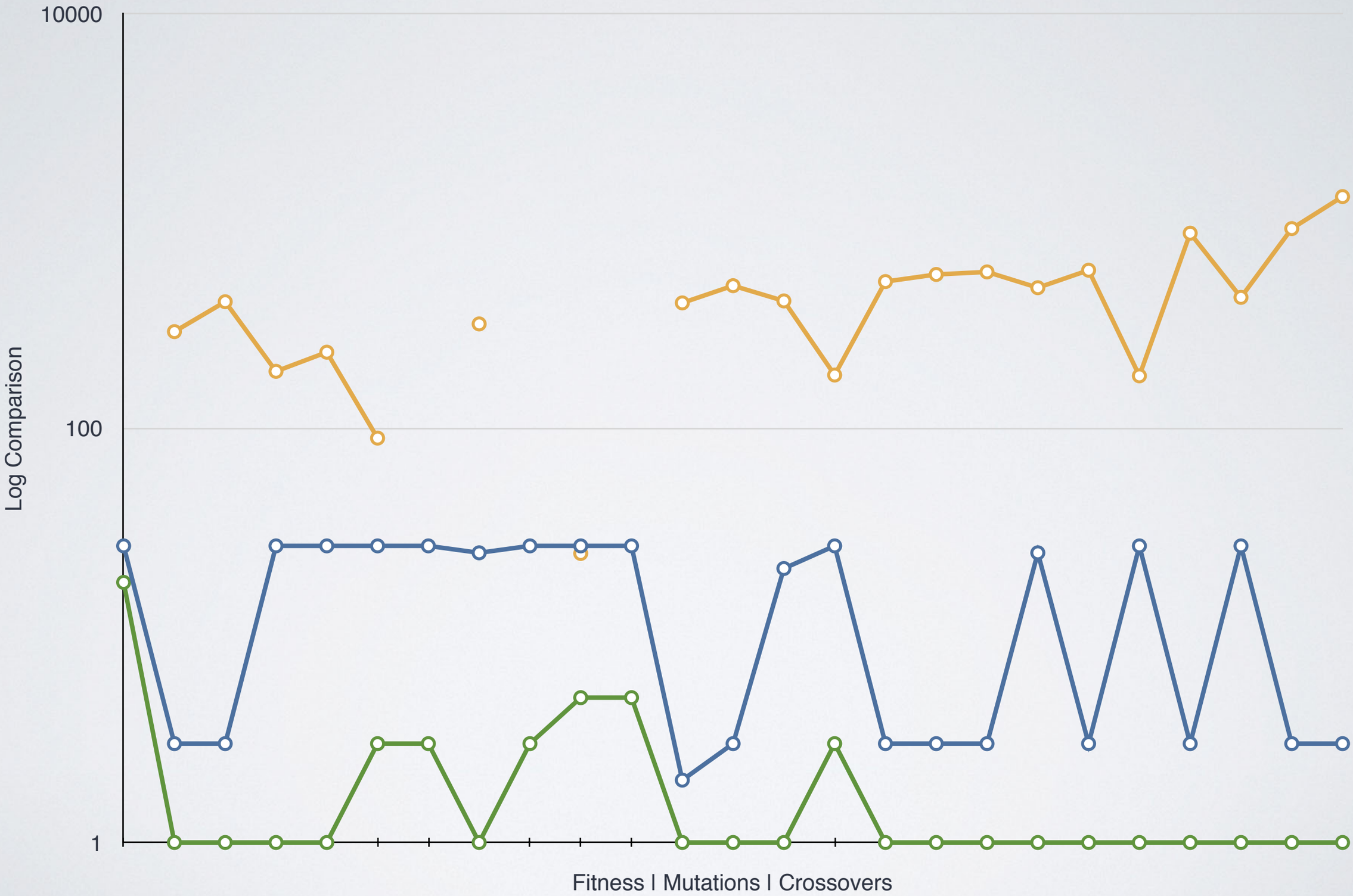
- **Dealer Policy:** Stand below 17 (Fixed)
- **Initial Player Policy:** Hit on states :- (Random)
  [S12, 13, S13, 14, S17, 18, 19, S19, S20, 21]
- **Simulations:** 100 per Genome
- **Gene Selection Criteria:** 50% improvement or better
- **Mutation Frequency:** 10% or more of the gene pool
- **Termination Criteria:** 1000% improvement over Dealer policy

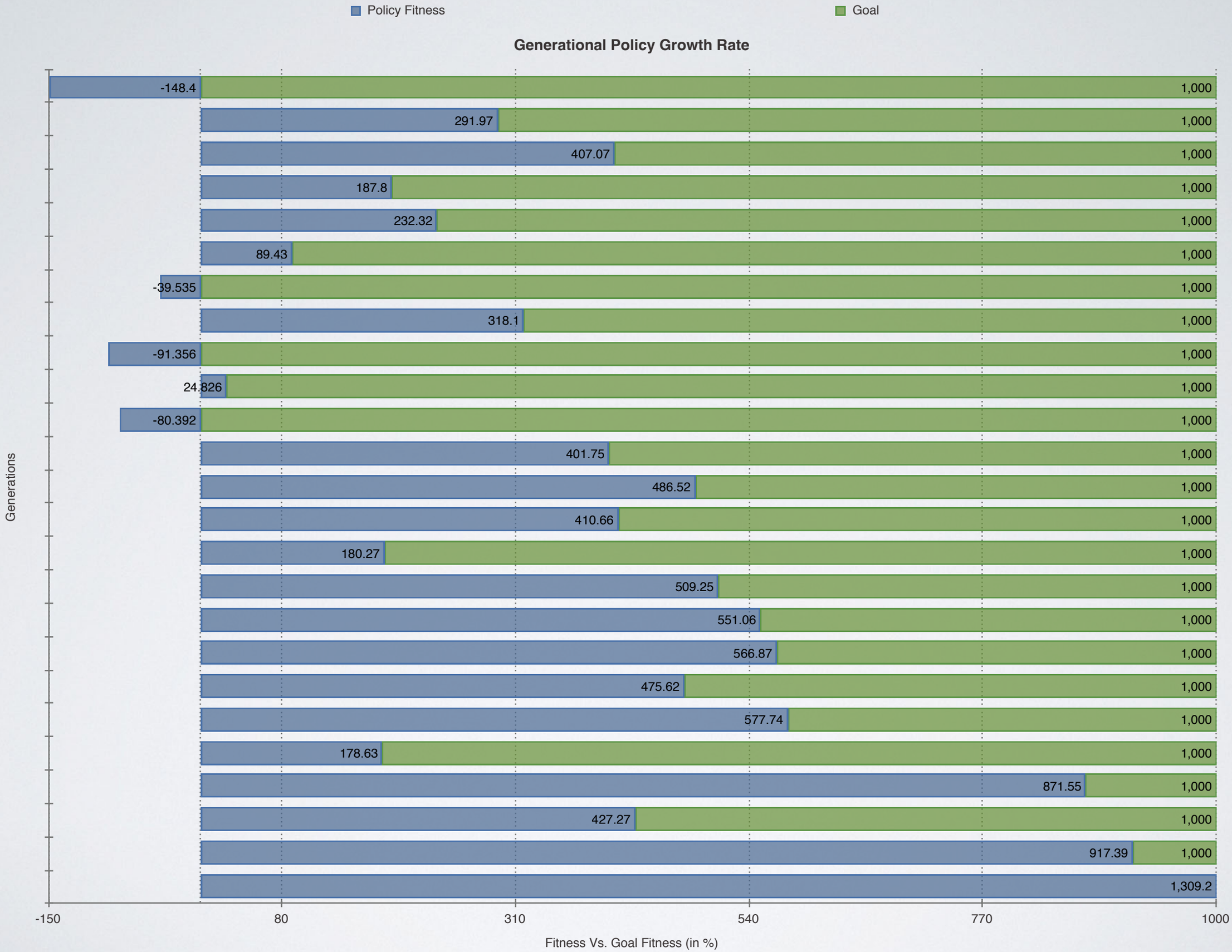

Avg. Gene Fitness per Generation
Avg. Gene Fitness
8
6
4
2
0
10
20
30
# of Generations


Policy Improvement
Avg. Gene Fitness
Correlation Between Gene & Policy Fitness
15
10
5
0
-5
Degree of Improvement
Overall Policy Vs. Avg. of Individual Genes

Mutations Crossovers Policy Fitness

Policy Fitness Vs Crossover & Mutation




Policy Fitness
Goal
Generational Policy Growth Rate
-148.4
1,000
291.97
1,000
407.07
1,000
187.8
1,000
232.32
1,000
89.43
1,000
-39.535
1,000
318.1
1,000
-91.356
1,000
24.826
1,000
-80.392
1,000
401.75
1,000
486.52
1,000
410.66
1,000
180.27
1,000
509.25
1,000
551.06
1,000
566.87
1,000
475.62
1,000
577.74
1,000
178.63
1,000
871.55
1,000
427.27
1,000
917.39
1,000
1,309.2
Generations
-150
80
310
540
770
1000
Fitness Vs. Goal Fitness (in %)

GOAL REACHED!

Case study represents the trivial case where dealer policy is terrible.
However, similar results are seen against better policies.
(The better the dealer, the lower the max improvement attainable)

FUTURE WORK

- In real blackjack, the dealer is not your only adversary. Thus we could improve on fitness by playing against multiple adversarial policies! This should lead to larger gains in policy fitness

- As seen from the correlation graph between genes and policy, the 2 are very loosely bound. A better fitness function is in order.

- A pseudo-coevolution was implemented in this program for the purpose of improving the player's policy faster (not discussed here). Upgrading this to proper co-evolution is desireable.

- Finally, having to set parameters such as 1000% i.e 10 times better is too unrealistic. This goes back to the loosely bound fitnesses, but also indicates a possible lack of correlation between my game simulations and an actual game. I might want to investigate this further.