

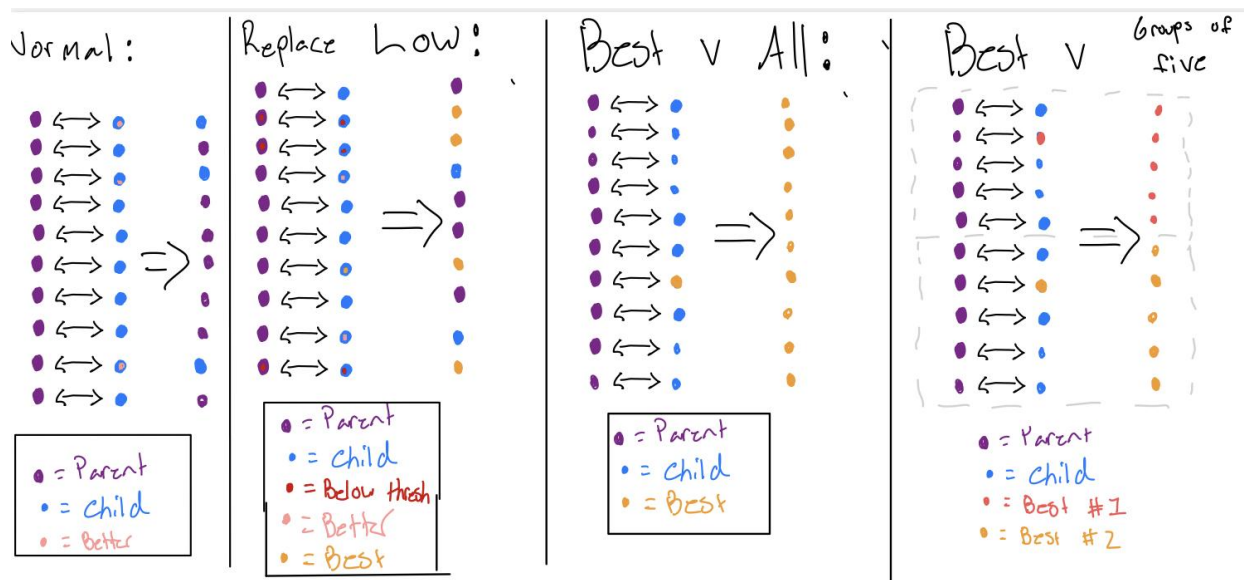
Experimental Results

Background:

The idea for the experiment presented comes from class discussions about different types of evolutionary algorithms. Specifically this experiment focuses on different levels of selection pressure which tend to either favor diversity or specialization. Changes to code were made to the parallel hill climber select function. These can be seen [here](#).

Definitions:

- “Normal Selection” -> The control for the experiment. Selection occurs by comparing the child against the parent, and replacing the parent with the child if the fitness is higher. This is the way the parallel hill climber typically selects.
- “Replace Low Select” -> If a creature had a fitness below a certain threshold (2 was used in the experiment), it was replaced by the best child in the population. Otherwise selection occurred as with normal selection.
- “Best v All Select” -> All the creatures in the population are compared and the best creature replaces every other creature. In this selection method, after the first generation all of the creatures are mutated versions of the same parent.
- “Best v Groups of 5 Select” -> The population is broken into groups of 5, each group of 5 goes through a best v all type selection process, this immediately reduces the number of unique creatures by a factor of 5.



Hypothesis:

The main hypothesis was that if different types of selection were tested, then normal selection would perform the best (meaning produce the highest fitness creatures). This is the hypothesis due to the preservation of diversity in the population that normal selection ensures.

As the research went on, it became apparent that there might be differences in the performance of the algorithms based on the ratio of number of parent creatures to the number of generations. Thus a secondary hypothesis was formed that if the number of parent creatures is significantly larger than the number of generations for evolution, that best v all selection would work best. This hypothesis makes sense since it functions as a random search for the first generation, and then only the best creature is evolved. Since there is a small number of generations, there isn't time for worse creatures to build up mutations that end up making them better in the long run.

Methods:

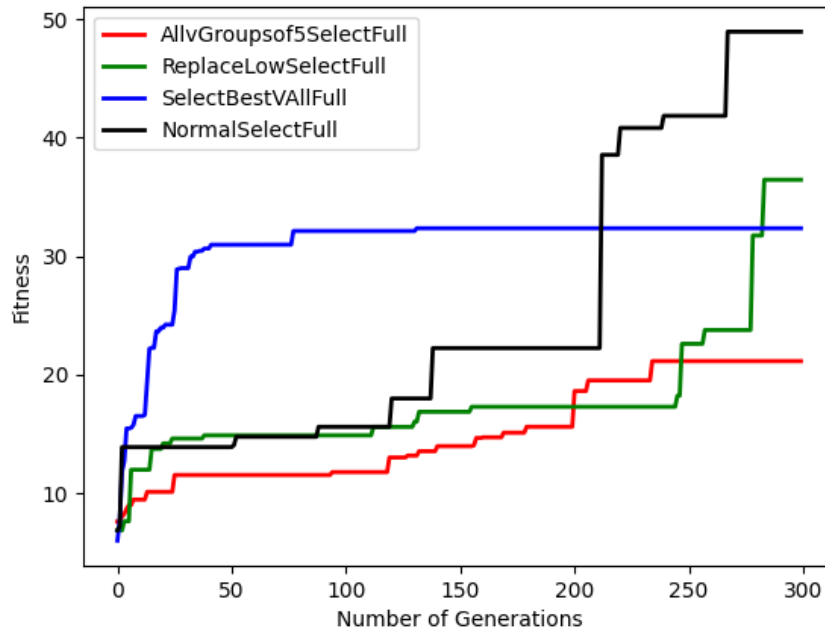
Each type of selection had 3 full runs as described below:

- "Full Run" -> 5 seeds x 25 parents x 300 generations = 37,500 sims
- "Wide Run" -> 5 seeds x 300 parents x 25 generations = 37,500 sims
- "Stat Run" -> 300 seeds x 10 parents x 10 generations = 30,000 sims

This adds up to 105,000 sims per selection method, or 420,000 sims total.

The results were all compared with t-tests for significance. The Full and Wide runs did not have enough results to compare and find significance (only 5, since it just compared the best values for each seed), which is why the stat run was conducted. Both the wide and stat run were added after the fact to answer questions about the additional hypothesis and fill in knowledge gaps.

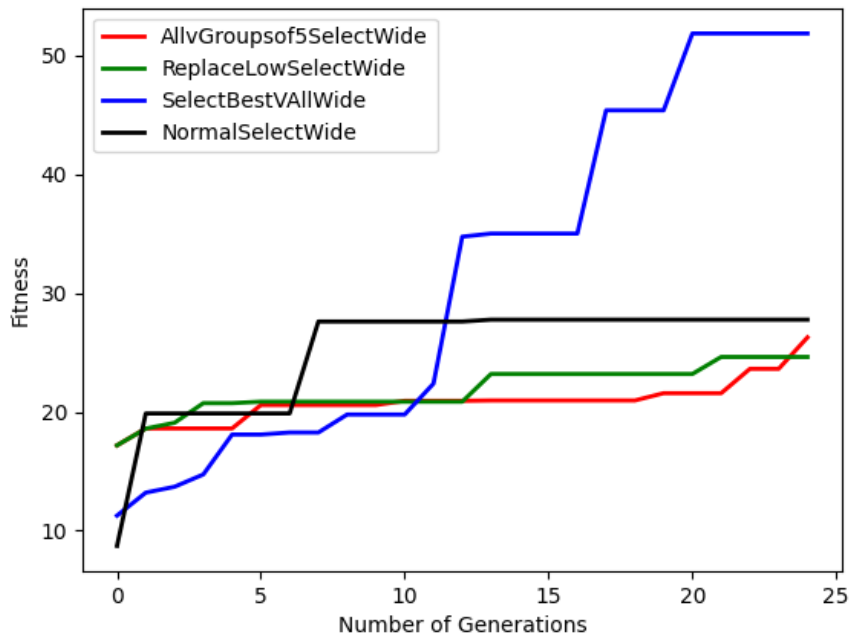
Results:



As you can see in these two graphs, the normal selection does better when compared to the other 3 types of selection. This graph is also supported by the following t-test results:

Selection Method 1	Selection Method 2	P Value	Significant?
Normal	Replace Low	0.00043345571397	Yes
Normal	Best v All	0.00000024621827	Yes
Normal	Best v Groups of 5	0.00006581881218	Yes

However when the “wide” version is run, the select best v all has a much higher fitness (as was hypothesized).



Graphs of nearly every possible combination can be found in the [graphs folder](#). The types of selection had no effect on the evolutionary dynamics, however they had huge effects on the amount of diversity. As noted above All v One selection has a diverse population of 1 after the first round, All v groups of five selection has its diverse population reduced by 5x after the first round, and even replace low select loses any creatures that fall below the threshold. This ends up having large effects on the end result of the population fitness.

Videos of evolutionary lines can be seen at the videos linked in the repository README

Conclusions:

In conclusion, both the main and exploratory hypotheses are supported. This means that when the number of generations is relatively high compared to the number of parents (such as 300 generations and 25 parents), normal selection is superior. This makes sense given the importance of diversity and small mutations building up throughout time.

On the other hand, if there are a lot more creatures than generations it makes sense to use a method that allows one creature to become very specialized. This functions more as a random search with some optimization at the end, however it can be quite effective and actually finds the highest fitness creature of any of the methods.

If there was more time I would continue to explore different selection methods, as well as test these same methods when there are both a lot of parents, and a lot of

generations (something I was unable to achieve due to computational limitations). I think random selection (or maybe different extents of randomness added to selection) would be really interesting to test as well.