Optimization Assignment 5 and 6

Shun-Lung Chang

Assignment 5 Task 1

The "Empirical Binomial" in Figure 1 shows the distribution of sums (i.e. values) of 10000 tuples, and each tuple was randomly generated as this task required. The mean and standard deviation of the distribution are around 50 and 5 respectively, and are shown in Figure 1. Furthermore, Figure 1 also shows a theoretical normal distribution and a theoretical binomial distribution, whose mean and standard deviation are both 50 and 5. As can be seen from Figure 1, "Empirical Binomial" bears resemblance to the two theoretical distributions, so that it can indeed be modeled by binomial law and approximated by normal distribution.

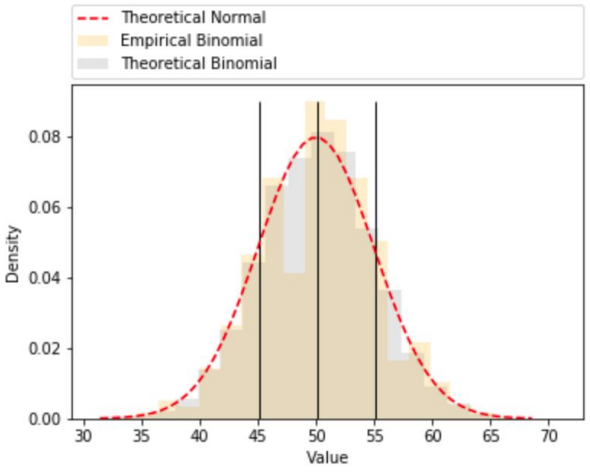

Figure 1: Distribution of Sums and Correspoding Binomial and Normal Distributions.

Assignment 5 Task 2

By setting M to 100 and following the instructions of this task, the histogram and the corresponding theoretical distribution in Figure 2 (a) can be obtained. As one can expect, the histogram lies above 50, given that it was constituted of each tuple's greatest value, which is usually larger than 50. Also, if M is larger, the histogram will move towards to the right-hand side, since larger M implies that a larger sum of a tuple can be attained. The results when M is 10000 are shown in Figure 2 (b).

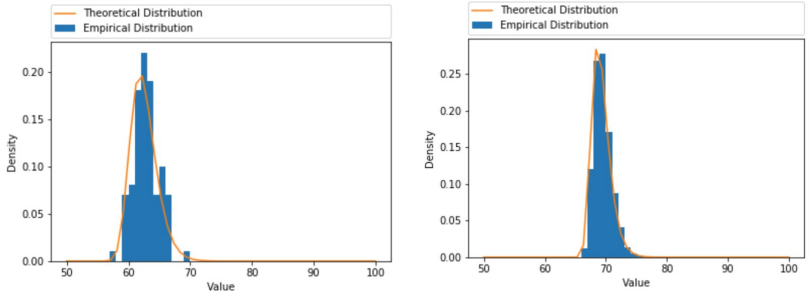

Figure 2: Distribution of Maximum Sums and Correspoding Theoretical Distributions. (a) M is 100, (b) M is 10000.

In the following three tasks, the results were obtained by doing 1000 experiments and then being aggregated by taking average.

Assignment 6 Task 1

Figure 3 shows the results when the new generation was created without taking its parent into account. The value increased progressively before 100th generation. Yet more 1's implies that it is more likely to gain an improvement by -1 (i.e. picking 1). Ultimately when the value went closer to 50, the probabilities of improving by 1 or -1 were almost the same, thus the improvement would come closer to 0 and even become negative, as Figure 3 (b) indicates.

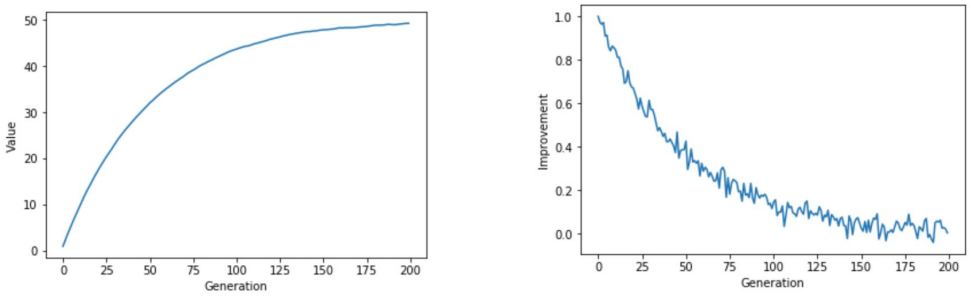

Figure 3: Results given that the parent is ignored. (a) value against generation, (b) improvetment against generation.

Assignment 6 Task 2

Figure 4 shows the average results if the next generation is determined by comparing the value of current generation and of its children, and picking the larger one. By doing so, the value is non-decreasing as generation increases, and does not reach a plateau at 50, as Figure 4 (a) shows. In fact, given enough generation, one can anticipate that the value can eventually achieve 100, namely all 1's in the tuple. Moreover, Figure 4 (b) is similar to Figure 3 (b) but with smaller oscillation, since this task's improvement in each generation is either 0 or 1. The theoretical expected improvement is also given in Figure 4 (b), and it is computed by the following formula. In this task, N is 100 and c is 1.

$1 - (\frac{k}{N})^c$, where k is the number of 1's in each generation, N is the length of tuple, c is the number of children.

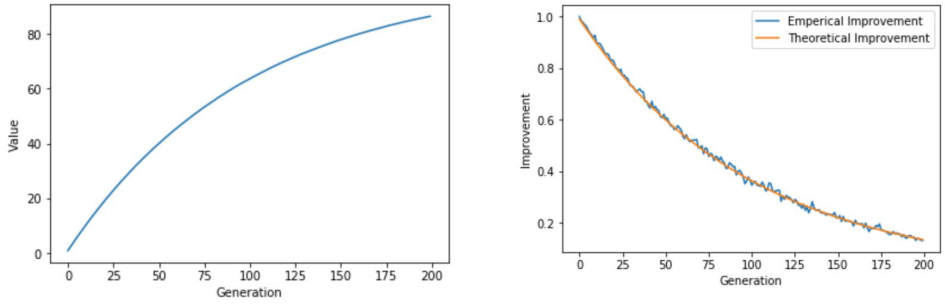

Figure 4: Results given that the parent is considered and number of children is 1. (a) value against generation, (b) improvement against generation.

If the number of children is increased to 3 and 5, the results are shown as Figure 5 and 6 respectively. The optimal value (i.e. 100) was achieved around 150th generation for 3 children per generation, and around 125th generation for 5 children per generation. The improvements in the first 50 to 70 generations remained at 1, suggesting that having more children would lead to a steady improvement in the beginning. When the number of 1's increased, the improvement by all means dropped and came closer to 0. In the case where there are only 1's in a tuple, the improvement is always 0.

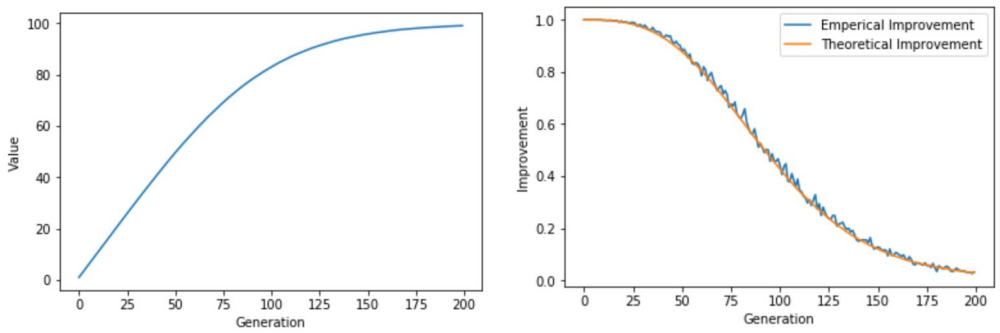

Figure 5: Results given that the parent is considered and number of children is 3. (a) value against generation, (b) improvement against generation.

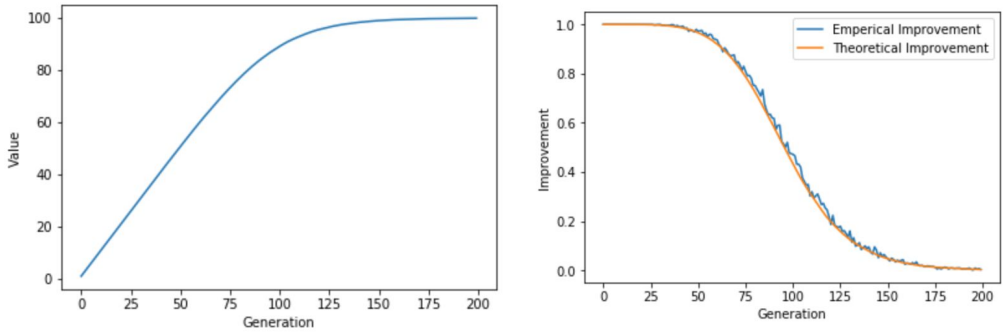

Figure 6: Results given that the parent is considered and number of children is 5. (a) value against generation (b) improvement against generation.

Assignment 6 Task 3

If the number of mutations is set to 3 with different numbers of children (1, 3, and 5), the results can be seen from Figure 7, 8 and 9 respectively. Though the value in three cases increased sharply in the beginning, yet reached a plateau after 50th generation. Also, the improvement was 3, namely three 0's are chosen, in the beginning, and then decreased to low values and fluctuated within a region. The reason is that a increased number of 1's would lead to a less chance for improvement. The improvement is more likely to be -2 or -1 if 1's outnumbered 0's, thus the parent would be select for the next generation.

To remedy the problem and to exploit the advantage of larger mutations, the number of mutations can be set to 3 before, and to 1 after 50th generation. In this setting with 5 children per generation, the optimal value could be obtained around 100th generation, as shown in Figure 10. Thus, this strategy outperformed all aforementioned settings.

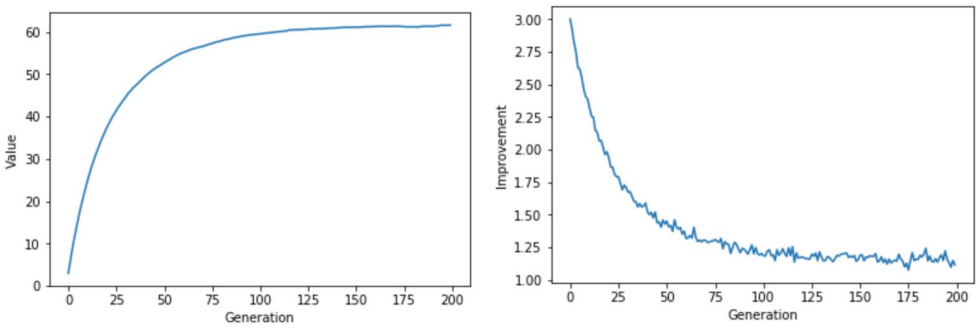

Figure 7: Results given that the parent is considered, number of mutation is 3 and number of children is 1. (a) value against generation, (b) improvement against generation.

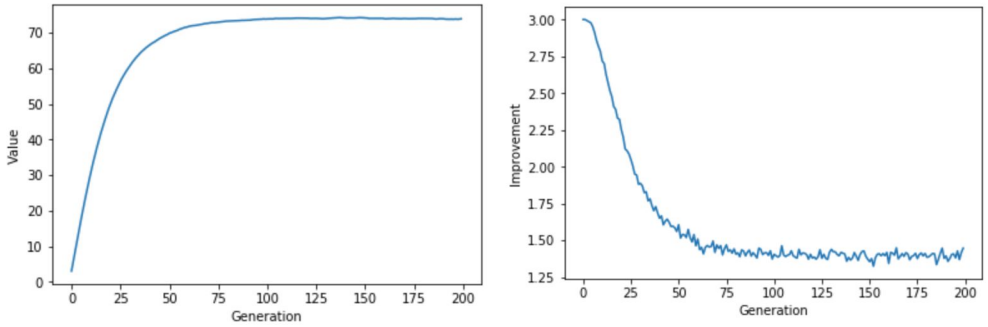

Figure 8: Results given that the parent is considered, number of mutation is 3 and number of children is 3. (a) value against generation, (b) improvement against generation.

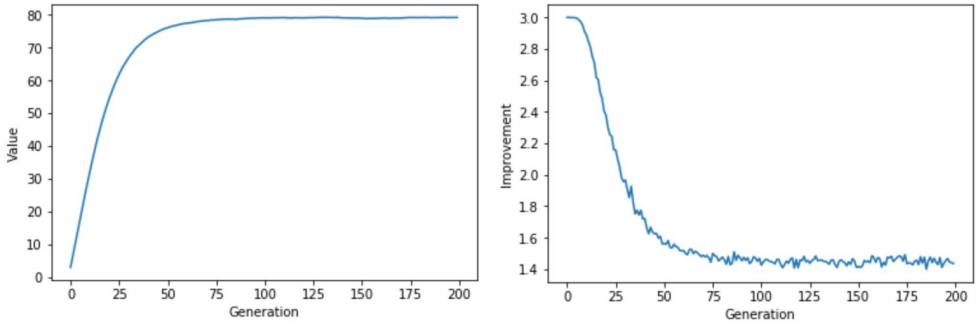

Figure 9: Results given that the parent is considered, number of mutation is 3 and number of children is 5. (a) value against generation, (b) improvement against generation.

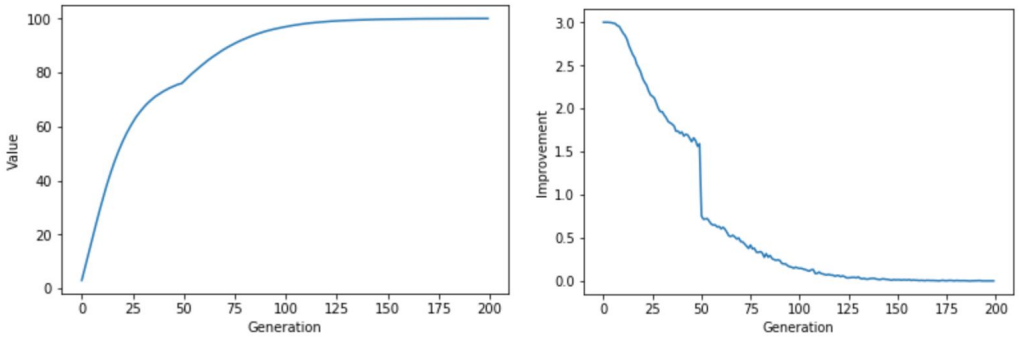

Figure 10: 3 mutations before, and 1 mutation after 50th generation, number of children is 5. (a) value against generation, (b) improvement against generation.