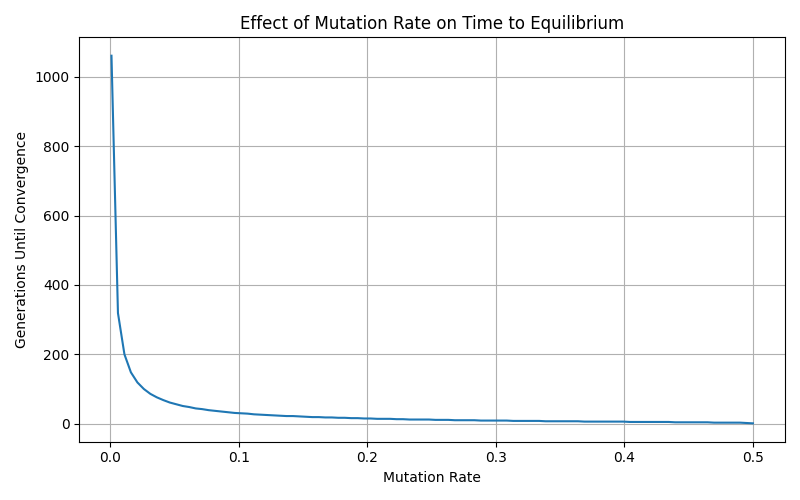
### Question 1 - Quasispecies Simulation on Binary Genomes

**Ex5 by Sarit Moses and Itamar Nini**

The objective of this assignment was to explore the synthetic fitness landscape of a binary genome of length L. For each simulation, we generated all possible options of a binary sequence at that length, assigned one of them randomly to be the founder (frequency 1 for the founder sequence and 0 for all other sequences), assigned a fitness to each of the sequences and then changed their frequencies according to the quasispecies model for a chosen number of generations. Fitness was assigned based on the number of "1" characters divided by length of the sequence. We chose to ignore the population size and work directly with frequencies.

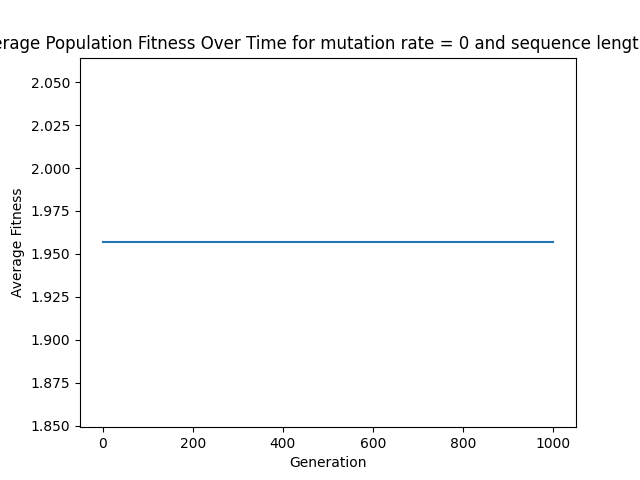
We noticed that lower mutation rate led to much longer time frames until mutation-fitness equilibrium is reached (see figure below). This may be explained by the fact that higher mutation rates lead to faster achieving a diversity of sequences (and this diversity allows an equilibrium to form), while a lower mutation rate would take longer to reach such diversity. If there is not enough diversity, the rate of mutation of each sequence into other sequences would be higher than the rate of mutation of other sequences into that sequence, which means that there is no equilibrium (Figure 1).

**Figure 1.** Effect of mutation rate on time to achieving mutation-fitness equilibrium.



We also examined the value of fixed average fitness of the population in comparison to the value of the highest possible fitness of the population (a scenario where the subspecies with the highest fitness is the entire population). Sometimes the fixed value is close to the highest possible value, but that is not always the case. For example, running the simulation with q=0 would ensure that the fixed species is the one that was randomly chosen as a founder, even if it is not the fittest (Figure 2). On the other hand, when introducing mutations at a very low rate (q=0.001 for L=5), the occurrence of mutations was enough to allow the appearance of the most fit sequence, but they were not fast enough to prevent it from fixating in the population after it appeared (Figure 3). A higher mutation rate (q=0.05 for L=5, representing one mutation per replication event) caused the creation of fixed non-zero frequencies for a few different sequences in the population, causing fixation of an average fitness that is lower than the highest possible value (Figure 4).

**Figure 2.** Fitness (left) and sequence frequencies (right) over time for a population with sequence length 5 and no mutations (q=0). The simulation was run 2 times (top and bottom figures are different runs).

תמונה שמכילה טקסט, צילום מסך, מספר, תצוגה

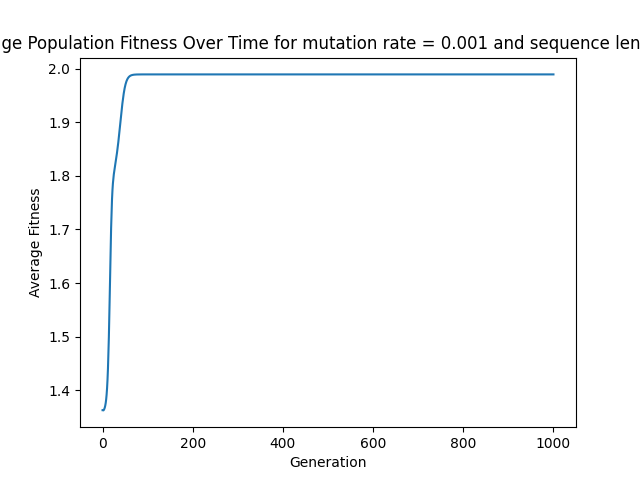
תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.

תמונה שמכילה טקסט, צילום מסך, קו, מלבן

תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.תמונה שמכילה טקסט, צילום מסך, מספר, תצוגה

תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.

**Figure 3.** Fitness (left) and sequence frequencies (right) over time for a population with sequence length 5 and mutation rate q=0.001. The simulation was run 2 times with similar results.

תמונה שמכילה טקסט, צילום מסך, מספר, תצוגה

תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.

**Figure 4.** Fitness (left) and sequence frequencies (right) over time for a population with sequence length 5 and mutation rate q=0.05. The simulation was run 2 times with similar results. In both cases, the maximum possible fitness was >1.98 while fixated fitness was ~1.74.

תמונה שמכילה טקסט, צילום מסך, קו, תרשים

תוכן שנוצר על-ידי בינה מלאכותית עשוי להיות שגוי.