CMPE 452 Assignment 4 – Theory Section

Name: Kieran Cosgrove | Student Number: 20226841

1. Selection

This question involves calculating the probability of individual 4 being selected using different selection methods.

(a) Roulette-Wheel Selection

In roulette-wheel selection, the probability of selecting an individual is proportional to its fitness value. The total fitness is $\mathrm{F1} + \mathrm{F2} + \mathrm{F3} + \mathrm{F4} + \mathrm{F5} = 5 + 7 + 8 + 10 + 15 = 45.$ Therefore, the probability of selecting individual 4 (with fitness 10) is $\frac{10}{45}$.

(b) Tournament Selection

For tournament selection with a size of 2 and a probability of 0.75 for selecting the best individual, we first need to calculate the probability of individual 4 being in a tournament and then being selected. Individual 4 can be paired with any of the other four individuals, so the probability of being in the tournament is $\frac{4}{5}$. Assuming individual 4 is the best in its pair, it will then be selected with a probability of 0.75. Thus, the overall probability $\frac{4}{5} \times 0.75 = \frac{3}{5}$. If the individual is the worst in its pair, it will be selected with a probability of 0.25 and therefore have an overall probability of $\frac{4}{5} \times 0.25 = \frac{1}{5}$.

(c) Roulette-Wheel Selection with Linear Ranking

Here, the lowest fitness value is set to 1 and the highest to 10. Assuming a linear ranking, the fitness values assigned would be 1, 3, 5, 7, 10 respectively for *F1* to *F5*. The total fitness is now $1 + 3 + 5 + 7 + 10 = 26$. The probability of selecting individual 4 (with ranked fitness 7) is $\frac{7}{26}$.

2. Schema Theory

This question involves calculating the order and defining length of a schema, and the average fitness of given schemas.

(a) Order and Defining Length of Schema S = 0 * * 1 * 1 * * 0 * * *

The order of a schema is the number of fixed positions (non-asterisks). Here, there are 4 fixed positions, so the order is 4. The defining length is the distance between the first and last fixed positions. In this case, it is the distance from the first 0 to the last 0, which is 9.

(b) Average Fitness of Schemas

- Schema 1***: The average fitness is the average of the values 1000, 1001, 1010, 1011, 1100, 1101, 1110, and 1111 which are 8, 9, 10, 11, 12, 13, 14, 15 respectively. So, the average fitness is $\frac{8+9+10+11+12+13+14+15}{8} = \frac{23}{2}$.
- Schema 0***: The average fitness is the average of the values 0000, 0001, 0010, 0011, 0101, 0110, and 0111 which are 0, 1, 2, 3, 4, 5, 6, 7 respectively. So, the average fitness is $\frac{0+1+2+3+4+5+6+7}{8} = \frac{7}{2}$

3. Mutation

This question involves calculating the probability of various mutation scenarios in a chromosome with binary-valued genes.

(a) No Mutation

Probability of No Mutation: This is the probability that none of the genes mutate. If the mutation rate per gene is *p* and there are *m* genes in the chromosome, then the probability that a particular gene does not mutate is $(1-p)$. Therefore, the probability that none of the genes mutate is $(1-p)^m$.

(b) Exactly One Mutation

This is the probability that exactly one gene mutates, and all others do not. There are *m* different genes that could be the one that mutates, and the probability of any single gene mutating is *p*. Thus, the probability of exactly one mutation is $m \times p \times (1-p)^{m-1}$.

(c) Less than Three Mutations

To find the total probability of having less than three mutations in a chromosome, we need to sum up the probabilities of having no mutations, exactly one mutation, and exactly two mutations. The formulas for these probabilities are as follows:

- Probability of No Mutation: $(1-p)^m$.

- Probability of Exactly One Mutation: $m \times p \times (1-p)^{m-1}$.

- Probability of Exactly Two Mutations: This is the probability that exactly two specific genes mutate, and the rest do not. The number of ways to choose 2 genes out of *m* is given by the binomial coefficient $\binom{m}{2}$, which can also be expressed as $\frac{m!}{2!(m-2)!}$, and the probability that these two genes mutate while the others do not is $p^2 \times (1-p)^{m-2}$. Therefore, the probability of exactly two mutations is $\binom{m}{2} \times p^2 \times (1-P)^{m-2}$.

The total probability of having less than three mutations is the sum of these three probabilities:
$\text{Total Probability} = (1-p)^m + m \times p \times (1-p)^{m-1} + \binom{m}{2} \times p^2 \times (1-P)^{m-2}$.