■ PROBLEMS

2.1. Consider three strings A_1 = 11101111, A_2 = 00010100, and A_3 = 01000011 and six schemata H_1 = 1*******, H_2 = 0*******, H_3 = ******11, H_4 = ***0*00*, H_5 = 1*****1*, and H_6 = 1110**1*. Which schemata are matched by which strings? What are the order and defining length of each of the schemata? Estimate the probability of survival of each schema under mutation when the probability of a single mutation is $p_m = 0.001$. Estimate the probability of survival of each schema under crossover when the probability of crossover $p_c = 0.85$.

2.2. A population contains the following strings and fitness values at generation 0:

#	String	Fitness
1	10001	20
2	11100	10
3	00011	5
4	01110	15

The probability of mutation is $p_m = 0.01$ and the probability of crossover is $p_c = 1.0$. Calculate the expected number of schemata of the form 1**** in generation 1. Estimate the expected number of schemata of the form 0**1* in generation 1.

2.3. Devise three methods of performing reproduction and calculate an estimate of the expected number of schemata in the next generation under each method.

2.4. Suppose we perform a crossoverlike operation where we pick two cross sites and exchange string material between the two sites.

```
x x x | x x | x x        x x x y y x x
                    →
y y y | y y | y y        y y y x x y y
```

Calculate a lower bound on the survival probability of a schema of defining length δ and order o under this operator. Recalculate the survival probability when we treat the string as a ring (when the left end is assumed to be adjacent to the right end).

2.5. How many unique schemata exist within strings of length l = 10, 20, and 30 when the underlying alphabet is binary? How many unique schemata of order 3 exist in binary strings of length l = 10, 20, and 30? Calculate reasonable upper and lower bounds on the number of schemata processed using strings of length l = 10, 20, and 30 when the population size m = 50. Assume a significant building block length equal to 10 percent of the total string length.

2.6. Suppose a schema H when present in a particular string causes the string to have a fitness 25 percent greater than the average fitness of the current population. If the destruction probabilities for this schema under mutation and crossover are negligible, and if a single representative of the schema is contained in the population at generation 0, determine when the schema H will overtake populations of size $n = 20, 50, 100$, and 200.

2.7. Suppose a schema H when present in a particular string causes the string to have a fitness 10 percent less than the average fitness of the current population. If the destruction probabilities for this schema under mutation and crossover are negligible, and if representatives of the schema are contained in 60 percent of the population at generation 0, determine when the schema H will disappear from populations of size $n = 20, 50, 100$, and 200.

2.8. A two-armed bandit pays equal awards with probabilities $p_1 = 0.7$ and $p_2 = 0.3$. Estimate the number of trials that should be given to the observed best arm after a total of 50 trials.

2.9. Derive a more accurate formula for calculating the number of unique schemata contained in a randomly generated initial population of size m when the string length is l. (*Hint:* Consider the probability of having no schemata of a particular order and use the complementary probability to count the number of schemata represented by one or more.)

2.10. Suppose a problem is coded as a single unsigned binary integer between 0 and 127 (base 10), where $0000000_2 = 0_{10}$, $1000000_2 = 64_{10}$, and $1111111 = 127$. Sketch the portion of the space covered by the following schemata: 1******, ******0, 111111*, 10*****, *****01, **111**.

■ COMPUTER ASSIGNMENTS

A. A fitness function for a single locus genetic algorithm is given by the function f_1 = constant and f_0 = constant. Derive the recursion relationship for the expected proportion of 1's under reproduction alone and reproduction with mutation in an infinitely large population. Program the finite difference relationship and calculate the expected proportion of 1's between generation 0 and 100 assuming equal proportions of 1's and 0's initially and a ratio of $f_1/f_0 = r = 1.1, 2, 10$.

B. Redo Problem A including mutation using mutation probability values of $p_m = 0.001, 0.01, 0.1$.