Evolutionary Algorithms
Problem Set - Genetic Algorithms

1. The number of ones in a randomly generated binary string of length l is a discrete random variable, X. Suppose that each bit in the string is drawn with a probability $p \in [0, 1]$ of being a one. Let $p_i = P(X = i)$ be the probaility that a randomly generated string contains $0 \leq i \leq l$ ones.

   (a) Give an expression for p_i, taking into account that there are multiple ways to have i ones distributed among the l possible positions.

   (b) What is the expected number of ones in a randomly generated binary string ? The expectation of a discrete random variable X is defined by

   $$E(X) = \sum_{i=0}^{l} i \cdot p_i \tag{1}$$

2. Given a binary string (individual) $\vec{a} = (0101)$ and a mutation operator that inverts single bits independently of each other with probability $p \in [0, 1]$, give mathematical expressions for the probabilities of the 16 possible mutations that can happen to $\vec{a}$ (i.e., what is the probability that $\vec{a}$ is mutated into $(0000), (0001), ...$?).
   How large is the **total** probability that 2 out of the 4 bits are changed by mutation ?
   In general, what is the total probability that i $(0 \leq i \leq n)$ out of n bits are changed (if n is the vector dimension) ?

3. In the course, we discussed schema analysis as the historical attempt to analyse genetic algorithms.

   (a) Given a schema $H = (0 \ * \ * \ 1 \ * \ 1 \ * \ * \ 0 \ * \ * \ *)$, give values for the order and the defining length of H.

   (b) Given the above schema, what is the probability that it will be destroyed by one-point crossover, if $p_c = 0.3$?

   (c) Given the above schema, what is the probability that it will not be destroyed by mutation with $p_m = 1/12$?

   (d) Prove that any string of length l is an instance of 2^l different schemata.

   (e) Define the fitness of bit string $\vec{a}$ with $l = 4$ to be the integer represented by the binary number $\vec{a}$ (e.g., $f(0011) = 3$). What is the average fitness of the schema $(1 \ * \ * \ *)$ under f ? What is the average fitness of the schema $(0 \ * \ * \ *)$ under f ?