1. Suppose a delivery robot must carry out a number of delivery activities, a, b, c, d, and e. Suppose that each activity happens at any of times 1, 2, 3, or 4. Let A be the variable reprsenting the time that activity a will occur, and similarly for the other activities. The variable domains, which represent possible times for each of the deliveries, are

$$dom(A) = \{1, 2, 3, 4\}, \quad dom(B) = \{1, 2, 3, 4\}, \quad dom(C) = \{1, 2, 3, 4\},$$
$$dom(D) = \{1, 2, 3, 4\}, \quad dom(E) = \{1, 2, 3, 4\}.$$

Suppose the following constraints must be satisfied:

$$B \neq 3, C \neq 2, A \neq B, B \neq C, C < D, A = D,$$
$$E < A, E < B, E < C, E < D, B \neq D$$

The aim is the find an assignment of value to each activity, such that all the constraints are satisfied.

We define the evaluation function for any assignment by the number of constraints that it does NOT satisfy. E.g.,

- The assignment $(A = 2, B = 2, C = 3, D = 1, E = 1)$ has evaluation 4 as four constraints $A \neq B, B \neq D, A = D, C < D$ are not satisfied.
- The assignment $(A = 4, B = 2, C = 3, D = 4, E = 4)$ has evaluation 4 as four constraints $E < A, E < B, E < C, E < D$ are not satisfied.

We would like to employ genetic algorithm to solve the problem. Each candidate solution $(A = a, B = b, C = c, D = d, E = e)$ of the form can be encoded by a string $abcde$. E.g., $(A = 2, B = 2, C = 3, D = 1, E = 1)$ is encoded as 22311. Answer the following questions

(a) Perform crossover over the two candidate solutions $(A = 2, B = 2, C = 3, D = 1, E = 1)$ and $(A = 4, B = 2, C = 3, D = 4, E = 4)$ on the crossing site 4. What is the outcome and its evaluation?

(b) Using examples, briefly discuss how genetic algorithm may help desirable "traits" of a candidate solution to be preserved in the population thereby improving the solution found.

Answer.

(a) Crossing over 22311 and 42344 at crossing site 4 produces us two solutions: 22314 and 42341, with evaluation 7 and 0 respectively.

(b) Genetic algorithm helps to preserve desirable traits within a population. For example, as demonstrated in (a), the candidate solution $(A = 2, B = 2, C = 3, D = 1, E = 1)$ has low evaluation mainly because of $E = 1$, and the candidate solution $(A = 4, B = 2, C = 3, D = 4, E = 4)$ has low evaluation mainly because of the assignments to the first four variables.

- If these two were to mate, using the crossing over operation, some offspring would inherit the good traits of both.
- Using the preferential selection operation, the offspring that preserves the better traits will be fitter and more likely to survive than those that do not.

2. In a machine learning scenario, suppose we are given a dataset

x_1	x_2	x_3	y
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

where x_1, x_2, x_3 are input features, and y is the output. For example, when $x_1 = 0, x_2 = 0$ and $x_3 = 1$, the output $y = 0$.

We would like to develop a function of the form

$$y = \begin{cases} 0 & w_1x_1 + w_2x_2 + w_3x_3 < 0 \\ 1 & \text{otherwise} \end{cases}$$

to capture the association between the input features and the output. We require that each w_i to be either $-1, 0$ or 1. Hence, a candidate solution specifies the values of w_1, w_2, w_3.

Given (w_1, w_2, w_3), to compute the evaluation: For each x_1, x_2, x_3

(a) First, compute the weighted sum $s = w_1x_1 + w_2x_2 + w_3x_3$; and

(b) Then determine $y_c = 0$ if $s < 0$ and $y_c = 1$ otherwise.

(c) Then compare y_c against y, the true output.

This produces a table. For example, if $w_1 = -1, w_2 = 0, w_3 = 1$

x_1	x_2	x_3	y_c	y	$y_c = y?$
0	0	0	1	1	yes
0	0	1	1	0	no
0	1	0	1	1	yes
0	1	1	1	0	no
1	0	0	0	1	no
1	0	1	1	1	yes
1	1	0	0	1	no
1	1	1	1	1	yes

The evaluation of the candidate solution (w_1, w_2, w_3) is the number of "no"s. For example, the evaluation of $(w_1 = -1, w_2 = 0, w_3 = 1)$ is 4.

Describe how you would apply genetic algorithm to solve this problem. Use an example run-thru to illustrate the procedures.

Answer.

To apply genetic algorithm, we need to set the following:

- Chromosomes: These are just the string $w_1w_2w_3$, for example, we can have the string $(-1)01$ to denote the solution $w_1 = -1, w_2 = 0, w_3 = 1$.
- Fitness function: The fitness function is simply $8 - \mathit{valuation}$, i.e., the number of "no"s in the table above.

To illustrate the process we use a possible run-thru:

- **Step 0:** Initialization of the population. Suppose we randomly generate a population of size 4:

  $$(-1)01, (-1)00, 01(-1), 1(01)1$$

- **Step 1:**

  (a) **Selection of the mating pool:**

w_1	w_2	w_3	*fitness*
-1	0	1	4
-1	0	0	2
0	1	-1	6
1	-1	1	5

[H] Using preferential selection, suppose the following solutions are randomly selected according to the fitness probability distribution:

$$(-1)01, 01(-1), 01(-1), 1(-1)1$$

(b) **Crossover:** Suppose that the first two and last two solutions are randomly mated, with crossing sites 2 and 1 respectively. We have

$$\text{crossover}((-1)01, 01(-1), 2) = ((-1)0(-1), 011)$$

and

$$\text{crossover}(01(-1), 1(-1)1) = (0(-1)1, 11(-1))$$

(c) For simplicity, assume no random mutation takes place.

- **Step 2:**

(a) **New population:**

w_1	w_2	w_3	*fitness*
-1	0	-1	4
0	1	1	6
0	-1	1	4
1	1	-1	7

[H] Using preferential selection, suppose the following solutions are randomly selected according to the fitness probability distribution:

$$011, 0(-1)1, (-1)0(-1), 11(-1)$$

(b) **Crossover:** Suppose that the first two and last two solutions are randomly mated, with crossing sites 2 and 1 respectively.

$$\text{crossover}(011, 0(-1)1, 2) = (011, 0(-1)1)$$

and

$$\text{crossover}((-1)0(-1), 11(-1)) = ((-1)1(-1), 10(-1))$$

(c) For simplicity, assume no random mutation takes place.

- We arrive at the new population:

w_1	w_2	w_3	*fitness*
0	1	1	6
0	-1	1	4
-1	1	-1	4
1	0	-1	8

[H] The last solution, $(w_1, w_2, w_3) = (1, 0, -1)$, is the optimal solution.