PART-2

PROBLEM-3

Given initial data table D is:

income	student	rating
high	no	fair
high	yes	fair
low	yes	fair
medium	yes	excellent
low	no	fair.

Goal is to transform the data table D to TD representing the population with the fitness function $F = 0$ for all chromosomes.

PART-1

We are asked to write the set of Binary Encoding chromosomes representing the data table D.

We can call this initial population as P.

The possible values for each attribute are

Income = high, low, medium.

student = yes, no.

rating = fair, excellent.

So each chromosome would be of the format

$C_1 C_2 C_3 C_4 C_5 C_6 C_7$ as there are total of 7 unique values.

And we are also asked to calculate the fitness value of each chromosome using Definition 2.

Chromosome	In=h	In=l	In=m	Ct=Y	Ct=N	R=fair	R=Exc	F
Ch1	1	0	0	0	1	1	0	3
Ch2	1	0	0	1	0	1	0	3
Ch3	0	1	0	1	0	1	0	3
Ch4	0	0	1	1	0	0	1	3
Ch5	0	1	0	0	1	1	0	3

Note: The number of chromosomes don't change from the initial data.

The fitness function is also evaluated for all the chromosomes & we got 3 for all of them. Importantly fitness value is same for all the chromosomes.

We have also kept all these values in a single table.

PART-2

(1) We need to create ONE generations P1 from our initial population P1 using our GA operators.

We are also asked to use Single Point Crossover & Single Point Mutation.

SELECTION:

If we see, the fitness values are same for all chromosomes, so we can select randomly any two chromosomes from P.

Note: We keep the population size to be constant. So after the entire process 5 chromosomes should be present.

REPRODUCTION:

We will follow Single point crossover as mentioned in Definition 1 which is after C_4. Since the fitness values are same, we can randomly pick 2 for crossover.

We'll pick ch1 & ch2 i.e. 1000110, 1001010 & perform the cross over after c4.

ch1	1000 ! 110
ch2	1001 ! 010
off1	1000 010
off2	1001 110

Similarly next we'll pick ch3 & ch4

ch3	0101 ! 010
ch4	0011 ! 001
off3	0101001
off4	0011010

RECOMBINATION:

Since the population size needs to be constant, we'll pick ch5 as our off5. (fitness values are same)

After this step the chromosomes obtained are 1000010, 1001110, 0101001, 0011010 & 0100110

MUTATION:

Now single point mutation has to be applied on each of the obtained offsprings.

Since it's Single point Mutation, we have randomly selected 1 bit from each offspring chromosome to obtain the final set of offsprings i.e. Since mutation pro. is not mentioned.

1000010, 1001110, 0101001, 0011010 & 0100110

Inverted C_1 C_5 C_2 C_4 C_5

0000010, 1001010, 0001001, 0010010 & 0100010

TERMINATE:

Since we are asked for only one generation, the ONE generation population P_1 looks like this using the process we followed.

Chromosome	In=h	In=l	In=m	Ct=Y	Ct=N	R=fair	R=Exc	F
Ch1	0	0	0	0	0	1	0	1
Ch2	1	0	0	1	0	1	0	3
Ch3	0	0	0	1	0	0	1	2
Ch4	0	0	1	0	0	1	0	2
Ch5	0	1	0	0	0	1	0	2

(2) The best chromosome after this one generation is ch1 = 0000010 as it's fitness value is 1 which is closer to F=0 (our goal) the optimum ideal solution.

(3) Now we need to write Data Table D_1 using the population P_1.

D_1 =

Income	Student	Rating
		Fair
high	yes	Fair
	yes	Excellent
medium		Fair
low		Fair.

If we observe there are few missing values in the data table D_1.

The reasons behind them are explained below.

Ch1 : This chromosome doesn't have values for Income & student attributes.

Ch3 : No value for Income attribute

Ch4 & Ch5 : No defined values for student attribute.

So finally data table D_1 is generated of the same format as of D. (Initial data table).