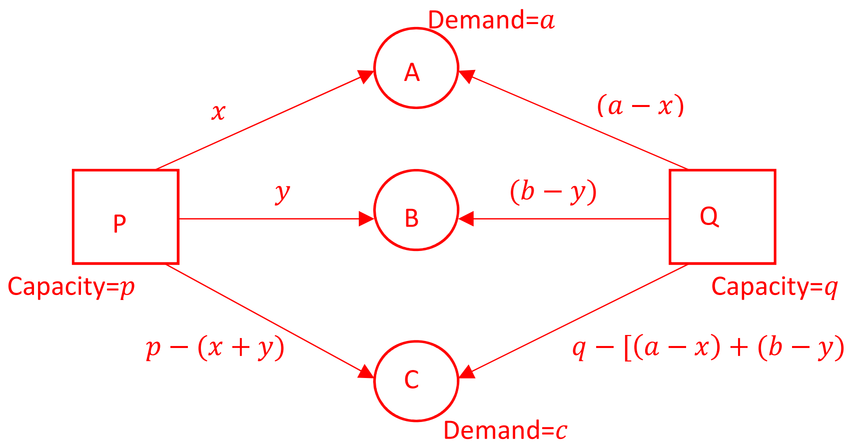
Question 1 There is a factory located at each of the two places P and Q. From these locations, a certain commodity is delivered to each of the three depots situated at A, B and C. The daily requirements of the depot are a, b and c units of the commodity, respectively while the production capacity of the factories at P and Q are p and q units, respectively. Further, the cost of transportation from any factory to any depot is given below:

	A	B	C
P	c_{pa}	c_{pb}	c_{pc}
Q	c_{qa}	c_{qb}	c_{qc}

(a) Formulate the above problem as an optimization problem. **[10]**

Answer

The given problem can be exhibited diagrammatically as follows:



Here, let the factory at P at transport x any y unit to A and B respectively. Since, the requirement are always non-negative quantity.

We have,

$x \geq 0, y \geq 0 \text{ and } p-(x+y) \geq 0$

Similarly, $(a-x) \geq 0, (b-y) \geq 0 \text{ and } q-[(a-x)+(b-y)] \geq 0$

The transportation cost from the factory at P to A, B and C are c_{pa}, c_{pb} and c_{pc}, respectively. Similarly, the transportation cost for Q to A, B and C are c_{qa}, c_{qb} and c_{qc}.

Hence, the cost of transportation is,

$$Z = c_{pa}x + c_{pb}y + c_{pc}(P - \overline{x+y}) + c_{qa}(a-x) + c_{qb}(b-y) + c_{qc}(q-(a-x)+(b-y))$$

Clearly this is to minimize the cost of transportation, hence the optimization problem can be formally stated as:

OBJECTIVE:

Minimize $Z = (c_{pa} - c_{pc} + c_{qc})x + (c_{pb} - c_{pc} - c_{qb} + c_{qc})y + p \cdot c_{pc} + a \cdot c_{qa} + b \cdot c_{qb} + q \cdot c_{qc} - (a \cdot c_{qc} + b \cdot c_{qc})$

Subject to

$(x + y) \leq p$

$x - y \geq q(a + b)$ and $0 \leq x \leq a, 0 \leq y \leq b$

Here, x and y are the design parameters only.

(b) It is proposed to solve the above optimization problem using Binary coded GA. Decide the genotype for the chromosome structure to do this. **[5]**

Answer

As there are only two design parameters, the genotype of the chromosome would look like

Genotype:

X	y

For binary encoded GA, let each parameter be represented by $n\ bits$. Since, x and y are with the same domain of values, the, phenotype for binary-coded GA will look as follows.

Phenotype for Binary GA:

$b_{11}, b_{12}, \dots \dots, b_{1n}$	$b_{21}, b_{22}, \dots \dots, b_{2n}$
for x	for y

(c) Suppose the problem needs to be adapted with m factories and n depots. What change in the chromosome structure you should devise? **[5]**

Answer

We see that for two factory problems, we need two design parameters to be considered. Thus, generalizing the above, for m+1 number of factories, we need to accommodate m design parameters. Hence, the chromosome structure would be

Genotype for $m + 1$ factory Binary coded GA

$b_{11}, b_{12}, \dots \dots, b_{1n}$	$b_{21}, b_{22}, \dots \dots, b_{2n}$	………………	$b_{m1}, b_{m2}, \dots \dots, b_{mn}$
x_1	x_2		x_m

Question 2 Answer the following:

(a) What is the use of selection operation in Genetic algorithm? **[2]**

Answer

The selection operation in Genetic algorithm is used to select individuals from the current population. The selected individuals will form a mating pool.

(b) Mention **four** criteria, which you should consider to judge the efficiency of a selection strategy? **[4]**

Answer

The four criteria which are usually considered as the measure of efficiency of any selection strategy are:

(i) Selection pressure: How good individuals are selected.

(ii) Population diversity: How varieties of individuals are selected.

(iii) Convergence rate: Selection should lead to a faster search towards optimal solution.

(iv) Solution quality: Selection should lead to global optimum or near global optimum solution.

(c) Precisely state the **two** major steps in *Tournament selection* strategy. **[4]**

Answer

Tournament selection strategy can be stated as follows:

Suppose, the size of the population is N and our task is to select N_p $(N_p \leq N)$ individuals from the current population. Further, let us consider N_u $(N_u \ll N)$ be the size of each team. The two steps in the Tournament selection is as follows.

Step 1:

Select N_u members randomly from the current population of size N.

Step 2:

Select a winner from the N_u individual from the currently chosen. Put the winner in the mating pool.

As we have to select N_p individuals, we have to repeat Step 1-2 for N_p times.

(d) It is planned to apply *Roulette wheel selection strategy* into *Tournament selection*. Give your suggestion, how the same can be realized. **[2]**

Answer

In the second step of *Tournament selection strategy*, there is a task to select a winner from a team. This selection of winner can be carried out applying *Roulette wheel scheme*. (Here only one round to select an individual from the team.)

(e) How *Tournament selection strategy* is comparable to Roulette wheel selection strategy, if they are individually applied in Genetic algorithm? You may give your comparison in the form of a table with reference to **four** efficiency measurement criteria you have mentioned as your answer to *Problem 2(b)*. **[8]**

Answer

A chart showing a summary of comparison is shown in the table given below.

Criteria / Scheme	Population diversity	Selection pressure	Convergence rate	Solution quality
Roulette Wheel selection	LOW: It favors the individual with higher fitness values.	HIGH: The chance of getting selected individual with higher fitness values.	FAST: As the population diversity is low and selection pressure is high, GA terminates at a faster rate.	Usually terminated with a local optimum solution.
Tournament selection	HIGH-MODERATE: A random selection is made to form team and hence a fair chance is given to all individuals in the population.	LOW- MODERATE: If a random selection is performed to select a winner, then selection pressure is low/ moderate.	MODERATE: As population diversity and selection pressure is low to moderate its termination rate is high to moderate.	Near optimal solution is expected.

<u>Question 3</u> Answer the following:

(a) Obtain the offspring chromosome(s), it will produce from the reproduction of three parent chromosomes P_1, P_2 and P_3 following the "Three parent crossover" technique? **[4]**

$\boldsymbol{P_1}$:

1	1	0	1	0	0	0	1

$\boldsymbol{P_2}$:

0	1	1	0	1	0	0	1

$\boldsymbol{P_3}$:

0	0	1	1	0	1	1	0

<u>Answer</u>

Three parent crossover technique: Each bit of P_1 is compared with the bit in the same position of P_2. If the two bits are same, then it is copied to offspring else the bit (at the same position) of P_3 is copied to offspring. Thus, offspring produced is

0	1	1	1	0	0	0	1

(b) Two parent chromosomes in Order GA are given as under:

$\boldsymbol{P_1}$:

1	2	3	4	5	6	7	8	9

$\boldsymbol{P_2}$:

1	2	3	4	5	6	7	8	9

Obtain the two offspring that can be obtained following the "Partially mapped crossover" technique. **[4]**

<u>Answer</u>

Let c_1 and c_2 are the two offspring

(i) Select any two points as crossover points on the parent chromosomes.

(ii) A schema in P_1 (or P_2) is directly copied in the same location and order in c_1 (or c_2).

(iii) The rest of the gene values in c_1 (or c_2) are obtained from P_2 (or P_1) in the same order as they appear.

Thus, two offspring chromosomes c_1 and c_2 will be obtained as

1	2	3	4	5	6	7	8	9

1	2	3	4	5	6	7	8	9

(c) "Linear crossover" technique in Real-coded GA takes the following form:

$$c_i = \alpha_i \cdot p_1 + \beta_i \cdot p_2$$

Where p_1 and p_2 are any two values and α_i, β_i are any two values chosen by the user.

Explain for the following two parent chromosomes, how two offspring chromosomes can be produced. **[4]**

$\boldsymbol{P_1}$:

5	10

$\boldsymbol{P_2}$:

6	9

<u>Answer</u>

Let us arbitrarily decide three set of values for $< \alpha_i, \beta_i >$ for three offspring chromosomes as follows:

$c_{11} = 0.5(p_{11} + p_{21})$; $c_{12} = 0.5(p_{12} + p_{22})$;

$c_{21} = 1.5p_{11} - 0.5p_{21}$;$c_{22} = 1.5p_{21} - 0.5p_{22}$

(d) Explain the working of "Flipping" as mutation operator in Binary-coded GA. **[4]**

<u>Answer</u>

An offspring can be mutated to a "mutated offspring" using the help of a "mutation chromosome." The mutation chromosome has a bit pattern with fewer 1's at random location. The number of 1's in "mutation chromosome" is decided by mutation operator μ, which denotes the probability that a bit is 1 at a random location in "mutation chromosome".

A bit in the offspring chromosome is flipped 0 to 1 and vice versa if a bit in the same position in mutation chromosome is 1 only.

For example,

Offspring:	0	1	1	0	0	1	1	1
Mutation chromosome:	0	1	0	0	0	1	0	0
Mutated chromosome:	0	0	1	0	0	0	1	1

(e) Consider the following is a chromosome in Order-GA encoding scheme.

Offspring:	B	H	F	G	C	E	A	D

Explain how such a chromosome can be mutated. Consider the least change in chromosomes between the offspring and the mutated offspring. **[4]**

Answer

Let us consider the change in gene values between two positions as the least change possible.

(i) First select two random locations in the offspring chromosome.

(ii) Swap the gene values between the selected locations.

For example, two random locations are shown as *

		*				*		
Offspring:	B	H	F	G	C	E	A	D
Mutated chromosome:	B	H	A	G	C	E	F	D

Question 4 Each of the following question includes a statement and possibly the statements are **wrong**. You have to **rewrite** the sentence in their correct form. **[10 X 2=20]**

(a) Vector Evaluated Genetic Algorithm (VEGA) is a multi-objective evolutionary algorithm (MOEA), which is an a priori technique and Pareto based.

Answer

Vector Evaluated Genetic Algorithm (VEGA) is a multi-objective evolutionary algorithm (MOEA), which is an a posteriori technique and non-Pareto-based.

(b) A Pareto front is also a Pareto optimal front but the reverse is not true.

Answer

A Pareto optimal front is a Pareto front but the reverse is not true.

(c) Both NSGA and NSGA-II follow the “Crowding Tournament” selection strategy to create mating pool.

Answer

NSGA follows “Stochastic remainder selection” whereas NSGA-II follows “Crowding Tournament selection” to create mating pool.

(d) The length of chromosomes in encoding scheme while solving a multi-objective optimization problem (MOOP) is proportional to the number of objective functions in the MOOP.

Answer

The chromosome length in MOOP is independent of the number of chromosomes.

(e) If c_1 and c_2 are two offspring chromosomes, then according to NPGA, c_1 will be preferable to be selected for mating pool if c_1's niche count is higher than that of c_2.

Answer

If c_1 and c_2 are two offspring chromosomes, then according to NPGA, c_1 will be preferable to be selected for mating pool if c_1's niche count is lower than that of c_2.

(f) According to MOGA, the rank of a solution is defined as the number of solutions by which it is dominated.

Answer

According to MOGA, the rank of a solution x_i is defined as

$$rank(x_i) = 1 + p_i$$

Where p_i is the number of solutions which dominates x_i.

(g) The a priori high level information that is required in "Lexicographic ordering" is the scalar weights of each objective function.

Answer

The a priori high level information that is required in "Lexicographic ordering" is the descending ordering of the rank of the importance of objective functions.

(h) Crowded comparison operator ($<_c$) (as defined in NSGA-II) to select between x and y is defined as $rank\ (x) < rank(y)$ or $rank\ (x) > rank(y)$ and $x_d > y_d$.

Answer

Crowded comparison operator ($<_c$) (as defined in NSGA-II) to select between x and y is defined as $rank\ (x) < rank(y)$ or $rank\ (x) = rank(y)$ and $x_d < y_d$.

(i) A solution x_i is said to dominate another solution x_j if

1. x_j is worse than x_i and
2. x_i is strictly better than x_j.

Answer

A solution x_i is said to dominate another solution x_j if both conditions (1) and (2) are true.

1. x_i is no worse than x_j.
2. x_i is strictly better than x_j in at least one objective.

(j) MOGA and NSGA follow their own steps to assign fitness values to all solutions in the current population whereas NPGA and NSGA-II do not require any fitness value calculation.

Answer

The statement is correct.

Question 5 Answer the following:

(a) What is niche count? What it does signify? Give an idea how niche count of a solution in a population can be calculated. **[1+1+4=6]**

Answer

Niche count: Niche count of a solution is a measure of how crowded the solution is by its neighboring solutions.

Physical meaning of Niche count: A solution with a higher niche count implies that the solution is a good representative of many other solutions with almost similar result. On the other hand, a solution with a lower niche count, if it is selected for mating pool then it ensures population diversity.

Measuring Niche count of a solution in a population: Let x_i be a solution whose Niche count is to be measured. Further, let x_j be any solution ($x_i \neq x_j$) in the given population of size N with solutions $x_1, x_2, \dots\dots.., x_N$.

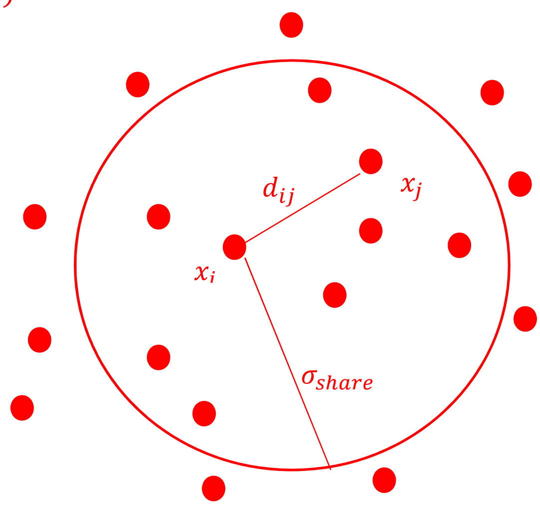
Let d_{ij} denotes the phenotype distance between x_i and x_j and σ_{share} is the maximum phenotype distance allowed between any two individuals to become members of a niche.

The, we calculate a sharing value $sh(d_{ij})$, which is as follows:

$$sh(d_{ij}) = \begin{cases} 1 - (\frac{d_{ij}}{\sigma_{share}})^2 & \text{, if } d_{ij} < \sigma_{share} \\ 0 & \end{cases}$$

Finally, Niche count of the solution x_i will be obtained as

$$Niche\ Count(x_i) = \sum_{\substack{i \neq j \\ j=1}}^{N} sh(d_{ij})$$



(b) Explain the **concept** of non-dominated sorting procedure as proposed in NSGA. What is the use of non-dominated sorting output in NSGA? **[3+5=8]**

Answer

Concept of non-dominated sorting:

Suppose, X denotes a set of solutions of a multi-objective optimization problem and, $|X| = N$.

Given X, we can find all non-dominated solution's $X_1 \subseteq X$. All solutions in X_1 lies in a front is called first non-dominated front (see Figure).

After removing X_1 from X, and repeating the same procedure we get X_2 such $X_2 \subseteq (X - X_1)$.

The above procedure can be continued until all solutions are sorted into front. Thus, given X, we obtain $X_1, X_2, \dots \dots \dots, X_N$ which are called non-dominated sorting fronts in X.

Fitness assignments with non-dominated sorting fronts

Let X_i be the $i - th$ non-dominated sorting fronts $(i = 1,2, \dots \dots, n)$. The steps to assign fitness values to all solutions in X_i are as follows.

1. Assign a dummy fitness value say f_{X_i} to each solution belongs to X_i. The dummy fitness value is a very large number and typically is proportional to the number of solutions in all fronts up to the current front. That is,

   $f_{X_i} = |X_1 \cup X_2 \cup \dots \dots \dots \cup X_i|$, a proportional constant is ≥1.
2. For each solution $x_j \in X_i$, it then calculates Niche count, $\gamma(x_j)$.
3. The fitness value of $x_j \in X_i$, is then calculated as

$$f_j = \frac{f_{X_i}}{\gamma(x_j)}$$

The above procedure is repeated to all solutions belongs to all fronts in X.

(c) The non-dominated sorting GA (NSGA-II) procedure for finding multiple Pareto optimal solutions in a multi-objective optimization problem has the following three features:

(i) It uses an elitist principle

(ii) It uses an explicit diversity preserving mechanism, and

(iii) It emphasizes the non-dominated solutions.

Briefly explain how NSGA-II accomplishes the above mentioned features. You should mention the concept only instead of procedure for each. **[2+2+2=6]**

Answer

(i) NSGA-II adopts elitist principle

A set of solutions which are very close to the optimal solutions are called elite. NSGA-II computes elite solutions with two steps:

A. Reproduce N offspring using N parents in the current population.

B. Obtain improved N solutions from the set of $2N$ solutions, which are non-dominated front by front.

(ii) NSGA-II follows explicit diversity preserving mechanism

NSGA-II proposes Crowding Tournament selection to ensure population diversity. To do this, it calculates "Crowding distance" for each solution. The crowding distance measure implies how a solution is crowded by its nearest neighbor. NSGA-II prefers a solution having less rank or crowding distance to be selected for mating.

(iii) NSGA-II emphasizes non-dominated sorting solutions

NSGA-II procedure in its each iteration obtains non-dominated sorting fronts. It then prefers the solutions in $first, second, \dots \dots \dots, i-th$ fronts until it obtains a mating pool of $size < N$. To make the size of the mating pool exactly N, it selects the remaining solutions from the $(i+1)th$ front using Crowding Tournament selection procedure.