**Soft Computing Applications (IT 60108)**
Spring, 2015-2016

# Practice Sheet-II

## A. Introduction to Optimization Problems and Solving Techniques

**1**. An integer programming problem may be

(a) Unconstrained optimization.

(b) Constrained optimization.

(c) None of the above.

(d) All of the above.

**2.** Genetic algorithms are always guaranteed to give optimal solution.

(a) True.

(b) False.

**3.** Genetic algorithm is used to solve any optimization problem.

(a) True.

(b) False.

**4.** Which of the following problem cannot be solved with GA?

(a) Clustering problem (for example, a number of document are given and you have to group them according to some criteria, say similarity).

(b) Job machine assignment problem (for example, there are different machines with different throughputs and a number of different jobs with different needs of their processing times).

(c) Network cabling problem (a number of computers are there in geographically distributed locations and it is required to lay cable connecting them, so that anyone can connect to others).

(d) Matching problem (for example, given a document, it is required to find a subset of documents in Internet so that each document in the subset has at least $p\%$ score of matching)

**5.** The travelling salesman problem (TSP):

(a) Define the problem.

(b) Express the problem as an optimization problem.

(c) Show that this problem cannot be solved in polynomial time (we say that a problem can be solved in polynomial time, if time required to solve the problem is $k.N^x$, where $k$ and $x$ are some positive constant and real number, respectively. For example, $2.5n^2$ or $8n^{0.5}$ etc.

**6.** Frame the following problem as an optimization problem

(a) An order pair $(c_i, n_i)$ denoting that $n_i$ number of coins with denomination $c_i$.

(b) Given $k$ such order pairs (i.e. number of coins with denominations).

(c) Given an amount say $R$. You have to change $R$ in terms of minimum number of coins only.

(i) Identify the design parameters, objective function, constraint(s), if any in the above mentioned problem.

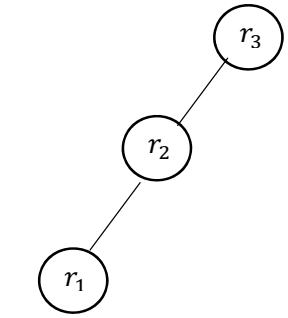
(ii) Show whether this problem can be solved in polynomial time or not. If so, how? If not, why?

**7.** Binary search tree construction problem can be illustrated as follows:

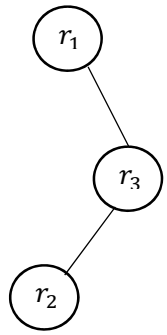
(a) Suppose, three records $r_1$, $r_2$ and $r_3$ are to be stored in a computer disk. $r_1$, $r_2$ and $r_3$ denotes their record numbers such that $r_1 < r_2 < r_3$.

(b) $\alpha, \beta$ and $\gamma$ denote the probability that the records $r_1$, $r_2$ and $r_3$ are to be accessed respectively such that $\alpha + \beta + \gamma = 1$.

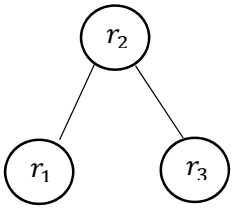
(c) Five possibilities of arranging them are shown below:



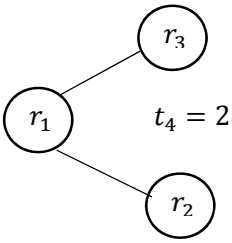

$t_1 = 3\alpha + 2\beta + \gamma$



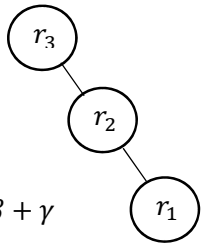

$t_2 = \alpha + 3\beta + \gamma$




$t_3 = 2\alpha + \beta + 2\gamma$




$t_4 = 2\alpha + 3\beta + \gamma$




$t_5 = 3\alpha + 2\beta + \gamma$

…. and so on

(i) Find at least three more such binary trees.
(ii) How many search trees are possible, if $n$ records $r_1, r_2, r_3, \ldots\ldots, r_n$ are to be maintained.
(iii) State this problem as an optimization problem identifying objective functions, constraints and all design variables involved in this problem.

**8.** Prove or disprove that TSP satisfies duality principle.

**9.** What is/ are the differences between Huffman algorithm and Genetic algorithm?

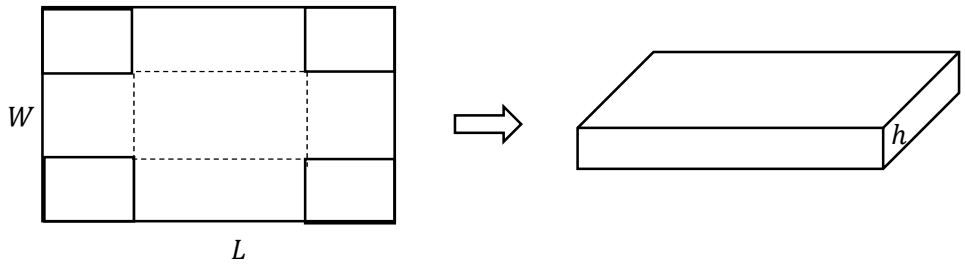
**10.** Given a perimeter of length $L$, it is possible to draw different geometrical shapes such as triangle, circle, rectangle, pentagon etc. Out of these shapes, how you can find a shape with minimum (maximum) surface area. Can you formulate this problem as an optimization problem?

**11.** Is there any scope to apply GA to obtain an ANN architecture for a given problem? If your answer is yes, then how?

**12.** Is there any scope to apply GA in the design of a fuzzy logic controller? If yes, then explain how?

**13.** For the following problems, precisely state the objective functions, constraints and design parameters involved. Also, for each suggest which encoding scheme you should consider to implement Genetic algorithm.

(a) Given a piece of cardboard of size $L \times W$. It is planned to cut out the corners as shown below and fold up the sides to form a box, also shown below. Determine the height of the box that will give maximum volume.



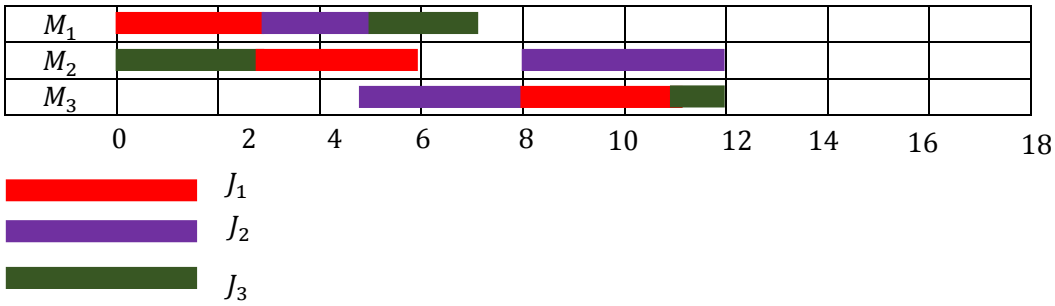

(b) **<u>Money allocation problem.</u>** In an investment scenario in equity market, $N$ different instruments are declared where an $i-th$ instrument returns $f_i(x_i)$ on investing $x_i$ amount in it. A customer has $C$ amount of capital to be invested. How he should invest his capital to the $N$ instruments to maximize his return?

(c) **<u>Job shop scheduling problem.</u>** A set of $n$ jobs $J_1, J_2, J_3, \ldots\ldots, J_n$ which are to be processed on a set of $m$ machines $M_1, M_2, M_3, \ldots\ldots, M_n$. Each job needs a strict sequence of machines which is known to the scheduler.
(i) Processing a job $J_i$ on a machine $M_j$ requires the exclusive use of $M_j$ for an uninterrupted duration $p_{ij}$, which is also known to the scheduler.
(ii) The time required to complete all jobs submitted to the scheduler is called $makespan, T$.

The objective of the scheduler is to determine schedule of jobs to machine minimizing $T$. A small instance of $3 \times 3$ job scheduling problem is given for your understanding.

| Job | Operation(time) | | |
|---|---|---|---|
| $J_1$ | $M_1(3)$ | $M_2(3)$ | $M_3(3)$ |
| $J_2$ | $M_1(2)$ | $M_3(3)$ | $M_2(4)$ |
| $J_3$ | $M_2(3)$ | $M_1(2)$ | $M_3(1)$ |

A schedule for instance (with Gantt chart)




(d) **Rod cutting problem.** Given a rod of length $L$ and it is required to be cut into maximum $p$ pieces. The cost $c_i$ of a piece $p_i$ of length $l_i$ is given.

| Length, $l_i$ | $l_1$ | $l_2$ | … | $l_m$ |
|---|---|---|---|---|
| Cost, $c_i$ | $c_1$ | $c_2$ | … | $c_m$ |

Assume the cost of cutting is free. It is asked to cut the rod into a number of small pieces so that the cost of cutting ensembles is maximum.

(e) Decide GA parameters involved different GA operators. Give a rule of thumb to decide the values of GA parameters. Also, mention how the values of GA parameters affect the performance of GA with reference to

(i) Rate of convergence

(ii) Accuracy of result

(f) Propose the framework of GA (can be termed as Hybrid GA (HGA)) combining some strategy, taking from SGA and SSGA.

(g) 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 depots are $a$, $b$ and $c$ units of 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:

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

Formulate the above problem as an optimization problem in terms of minimum number of design parameters.

## B. Introduction to GA

**1**. Which of the following statement(s) is (are) not true?

(a) GA is an optimization technique.

(b) GA is a probabilistic search technique.

(c) GA always gives optimal solution.

(d) For a problem, if there is more than one optimal solutions, GA outputs all the solutions.

**2.** GA is effective to solve only

(a) Single objective optimization problems.

(b) Mixed integer programming problem.

(c) Optimization problems involved with discrete (integer) variables.

(d) Optimization problems involved with continuous (integer/ real) variables.

**3.** In simple GA (SGS), which is/ are not allowed.

(a) Repetition in the individual selected for mating pool.

(b) Population sizes in each generations are to be varied.

(c) Overlapping generations.

(d) Two or more individuals with same fitness value.

**4.** In the context of simple GA (SGA), the convergence rate will increase if

(a) The size of mating pool ($N_p$) is comparable to that of the population size ($N$).

(b) From the mating pool of size ($N_p$), if we create $N$ offspring to replace all individual in the current population.

(c) From the mating pool of size ($N_p$), if we create a large number of offspring and replace only $N_p$ worst individuals in the current population size ($N$).

(d) Select $N_p$ best individuals for mating to produce $N$ individual for the next generation.

**5.** In the context of Steady-State GA (SSGA), which of the following statement(s) is/ are true.

(a) Generation gap is small.

(b) Convergence rate is high.

(c) Gives better result than Simple Genetic algorithm (SGA).

(d) It is susceptible to stagnation.

**6.** Which GA framework is preferable when?

(a) Population size is very large. - _____

(b) Calculation of fitness value of any individual is computationally expensive.-_____

(c) Needs to produce results with fewer cycles. - _____

(d) Needs to produce move near optimum solution. - _____

**7.** If we compare SGA and SSGA, then

(a) The time for single iteration in SGA is more than in SSGA.

(b) Overall time for termination in SGA is more than in SSGA.

(c) Population diversity in SGA is far better than that of SSGA.

(d) Selection pressure in SGA is more than that of SSGA.

**8.** Which of the following cannot be considered as convergence criterion

(a) There is no significant change in the average fitness value for $n$ successive generations.

(b) The execution of GA elapsed a pre-specified time $T$.

(c) $p$ percentage of $N$ (where $p$ is some positive constant and $N$ is the size of the population) individuals are with the same fitness values.

(d) Fitness values of all individuals are above than a certain threshold value.

**1.** An optimization problem is stated as follows:
$maximize\ f(x,y) = \frac{x^2}{2} + \frac{125}{y^2}$ where $x, y \in R^+$
The above optimization problem comes under the category of

(a) Unconstrained optimization problem.

(b) Linear optimization problem.

(c) Integer value optimization problem.

(d) Real value optimization problem.

**2.** Which of the following is <u>not</u> true for Genetic algorithms.

(a) It is a probabilistic search algorithm.

(b) It is guaranteed to give global optimum solutions.

(c) If an optimization problem has more than one solution, then it will return all the solutions.

(d) It is an iterative process suitable for parallel programming.

**3.** Which one of the following is a faster encoding scheme?

(a) Binary coded GA.

(b) Real coded GA.

(c) Order GA.

(d) Tree encoded GA.

**4.** Which of the following optimization problem(s) can be better solved with Order GA?

(a) 0-1 Knapsack problem.

(b) Travelling salesman problem.

(c) Job shop scheduling problem.

(d) Optimal binary search tree construction problem.

**5.** Which of the following GA operator is computationally most expensive?

(a) Initial population creation.

(b) Fitness evaluation.

(c) Selection.

(d) Crossover.

**6.** For a given optimization problem, and using the same GA parameters, if a GA algorithm is executed multiple times, then

(a) It will always give same solution.

(b) It will never give same solution.

(c) It is likely to give same solution.

(d) None of the above.

**7.** Principle of duality is related to transforming a maximization problem to minimization problem and vice-versa. In this context, which of the following statements is/are true?

(a) For every problem, there is a dual problem.

(b) There exists some problems for which dual problems exist.

(c) Solution of a problem $P$ and its dual problem $P'$ are same.

(d) Solution of a problem $P$ and its dual problem $P'$ are not same.

**8.** Suppose, all steps in both SGA and SSGA remain same, except instead of selecting two individuals from the current population of size $N$, $N_p$ ($N_p \ll N$) individuals as in SGA are selected. Then,

(a) Generation gap of SGA will be more than that of SSGA.

(b) Generation gap of SSGA will be more than that of SGA.

(c) Generation gap in both algorithms remains same.

(d) Nothing can be said precisely.

**9.** Genetic algorithms terminate quickly, iff

(a) In each iteration it maintains a large generation gap.

(b) In each iteration it maintains a small generation gap.

(c) In each iteration only best individuals are selected for mating.

(d) In each iteration only worst individuals are selected for mating.

**10.** GA is preferred algorithm to solving optimization problems because

(a) It can be applied to any optimization problem.

(b) It can solve any non-polynomial time optimization problem.

(c) Parallel programming is possible to implement in GA.

(d) It gives guaranteed optimal solution within reasonable time.

## C. Encoding Schemes and Crossover in GA

**1**. Which of the following statement is true?

(a) Binary coded GAs are faster than real coded GAs.

(b) Binary coded GAs are accurate than real coded GAs.

(c) All optimization problems can be encoded with binary coded GA.

(d) All GA parameters, which are applicable to binary coded GA is also applicable to order GA.

**2.** The length of chromosomes in binary coded GAs is decided by

(a) Number of design variables.

(b) Range of values of design variables.

(c) Objective functions.

(d) Constraints.

**3.** What should be the structure of chromosomes if an optimization problem is given as

$minimize\ f_i(x_1^i, x_2^i, x_3^i,..., x_n^i) \quad for\ i = 1,2, \dots, k \quad subject\ to\ g_j(x_i^j, x_2^j, x_3^j,..., x_n^j) \quad for\ j = 1,2, \dots, l$

$and\ x_1, x_2, \dots\dots, x_p$ are the set of all design parameters.

**4.** Which of the following is/are not valid chromosomes in order GA?

(a)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |

(b)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 3 | 5 | 7 | 2 | 4 | 6 | 1 |

(c)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A | B | D | E | A | F | H | G |

(d)

| | | |
|---|---|---|
| 14.6 | -23.4 | 177.23 |

**5.** Out of the following crossover techniques which do(es) not suffer(s) from “end point bias” problem?

(a) Single point crossover.

(b) Uniform crossover.

(c) Shuffle crossover.

(d) Multipoint crossover.

**6.** Hamming distance problem in crossover techniques never occur in

(a) Binary coded GA.

(b) Real coded GA.

(c) Order GA.

(d) Tree encoded GA.

**7.** Which of the following crossover techniques in order GA suffers from “position bias” proble?

(a) Single point order crossover.

(b) Precedence preservative order crossover.

(c) Position based crossover.

(d) Edge recombination crossover.

**8.** Which of the following crossover technique is known for real coded GA?

(a) Half uniform crossover.

(b) Simulated binary crossover.

(c) Three parent crossover.

(d) Position based crossover.

**9.** State whether the following statements are true or false.

(a) The most expensive operation in GA is the selection of individual for mating pool creation.

(b) Mutation operation is not mandatory in any GA based problem solving.

(c) The number of GA loop will increase if the population size is increased.

(d) Even we can implement Simple GA without performing selection operation.

10. This is with reference to the **Problem A. 13(g)**. It is proposed to solve the said optimization problem using Binary coded GA. Decide the genotype for the chromosome structure to do this.

    Suppose the problem needs to be adapted with $m$ factories and ***n*** destinations. What changes in chromosome structure you should devise?

## D. Selection operations in GA

**11.** The purpose of the fitness evaluation operation is

(a) To check whether all individuals satisfies the constraints given in the problem.

(b) To decide the termination point.

(c) To select the best individuals.

(d) To identify the individual with worst cost function.

**12.** Average fitness calculation can be used to

(a) Understand whether the optimal solution(s) has been achieved.

(b) Termination of GA execution.

(c) Generation gap between two successive GA iteration.

(d) Whether the problem has been stuck at local optima or not.

**13.** Roulette wheel selection scheme is preferable when

(a) Fitness values are uniformly distributed.

(b) Fitness values are non-uniformly distributed.

(c) Needs low selection pressure.

(d) Needs high population diversity.

**14.** Which of the following operation(s) is (are) more computationally expensive

(a) Roulette wheel selection.

(b) Rank based selection.

(c) Tournament selection.

(d) Steady state selection.

**15.** Which of the following selection scheme may select an individual more than once

(a) Roulette wheel selection.

(b) Rank based selection.

(c) Tournament selection.

(d) Steady state selection.

**16.** Which of the following statement is true in case of Rank based selection scheme

(a) Low population diversity, high selection pressure.

(b) Low population diversity, low selection pressure.

(c) High population diversity, low selection pressure.

(d) High population diversity, high selection pressure.

**17.** High selection pressure is desirable when we need

(a) Faster termination of GA.

(b) Near optimal solution.

(c) There is no improvement in successive GA iteration.

(d) Fitness values are non-uniformly distributed.

**18.** Tournament selection scheme is more preferable when

(a) Population are with very diversified fitness values.

(b) When fitness values are uniformly distributed.

(c) When fitness values are not necessarily uniformly distributed.

(d) Under all the above situations.

**19.** To make the generation gap ($G_p$) in a Steady state selection strategy a large value, which of the following can be considered?

(a) Selection of individuals according to their fitness values and replacement at random.

(b) Selection of individuals at random and replacement according to the inverse of their fitness values.

(c) Selection of both parents and replacement according to the inverse of their fitness values.

(d) All of the above.

**10.** Answer the following:

(a) Precisely state the two major steps in Tournament selection strategy.

(b) How Tournament selection strategy is comparable to Roulette wheel selection strategy?

(c) It is planned to apply Roulette wheel selection scheme into Tournament selection strategy. Give your idea to do this.

(d) Define population diversity and selection pressure.

(e) Give your idea to measure the above two while GA is in execution.

(f) Is there any relation between the two?

(g) How population diversity (or selection pressure) can be controlled?

(h) As GA leads to termination, population diversity decreases while selection pressure increases.

(i) How the two measure can be considered to decide a termination criterion?