School of Information Technology
Indian Institute of Technology Kharagpur

IT 60108 :: Soft Computing Applications
Mid-Spring Semester Test
Session 2013-2014

Full Marks: 40 | **Time: 02 hours**

Answer ALL questions
(Numbers in the right side indicate the marks for questions)

1. (a) Draw a flowchart of the steady state GA (SSGA). 3

   (b) Answer the following questions with reference to SSGA.
   - i. How SSGA is different from simple GA (SGA)? 2
   - ii. To solve what type of optimization problem the SGA can be thought for? 1
   - iii. Comment on the performance of SGA with respect to the following. (You should justify your answers.)
     - Convergence rate 1
     - Population diversity 2
     - Selection pressure 2

2. A binary search tree is a binary tree such that the key value of each node is greater than the key value of any node in its left sub-tree and less than the key value of any node in its right sub-tree. All key values in a binary search tree are assumed to be unique.

   There is a need to construct an optimum binary search tree with the following specification.

   - Given a set $K = \{K_1, K_2, \ldots., K_n\}$ with n distinct key values such that $K_i < K_{i+1}$, $i = 1, 2, \ldots., n$. We have a probability p_i such that a search will be for key K_i where $\sum_{i=1}^{n} p_i = 1$

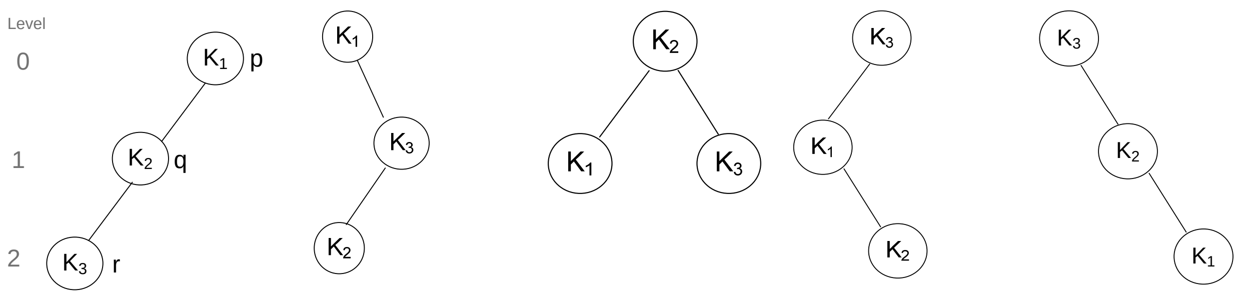
   - Assuming that the root node is at level 0. The average search time $T(n)$ for a given binary search tree is defined as

   $$T(n) = \sum_{i=1}^{n} p_i \ (level\ (K_i) \ + 1)$$

     where $level\ (K_i)$ denotes as level of K_i.

   The problem is to find an optimum binary search tree with minimum$T(n)$.

   For an example, with tree key values K_1, K_2 and K_3 (such that $K_1 < K_2 < K_3$) with probabilities p, q and r , (such that $p + q + r = 1$), there are many binary search trees with different $T(n)$ values few of which are shown below.



$T(n) = p + 2q + 3r$ $T(n) = p + 3q + 2r$ $T(n) = 2p + q + 2r$ $T(n) = 2p + 3q + r$ $T(n) = 3p + 2q + r$

With reference to the above problem answer the following questions.

(a) Whether solving the above problem comes under hard computing, soft computing or hybrid computing? 2

(b) It is proposed to solve the problem using genetic algorithm (GA). In that case state your approach to the following.

4

i. Encoding scheme (clearly explain genotype and phenotype for an instance).

4

ii. Crossover technique (explain your crossover technique with a small but clear example).

3. Selection strategy according to *Roulette Wheel* scheme is stated as follows.

Input: Given a population of size *N* with fitness values of individuals are $f_1, f_2, \ldots, f_N$

Output: Selection of a sub-population of size N_p for some N_p.

1. Calculate $p_i = \frac{f_i}{\sum_{i=1}^{N} f_i}$ for all $i = 1, 2, \ldots., N$

2. Calculate cumulative probability for each individual
   $P_i = \sum_{j=1}^{i} p_j$ for all $i = 1, 2, \ldots., N$

3. Generate a random number r (between 0 and 1, both inclusive)
4. Select the j-th individual such that $P_{j-1} < r \leq P_j$.

5. Repeat Step 3-4 until N_p number of individuals are selected.

With reference to the above algorithm, answer the following questions.

(a) How many counts the i –th individual is expected to be selected? 1

(b) Whether there is(are) case(s) that some individual(s) will be selected more than once. Argue your answer. 1

(c) It is proposed to modify the above replacing Step 3—5 as follows.

3. Choose i –th individual such that $P_{i-1} < r \leq P_{ii-1}$ as a candidate
4. Compute $e_i = p_i \times N$ with the p_i value of the candidate
5. If $\lfloor e_i \rfloor \neq 0$, then select the i –th individual
6. $r = e_i - \lfloor e_i \rfloor$
7. Repeat Step 3-6 until $N_p < N$ number of individuals are selected.

Assume that the procedure starts with an initial value of $r = 0.5$.
Here $\lfloor e_i \rfloor$ denotes the integer value not exceeding e_i.

Calculate the selection of population, if we run the modified version of *Roulette Wheel* for 8 times, with the following population with their fitness values. 8

Individual	1	2	3	4	5	6	7	8
Fitness value	1.0	2.1	3.1	4.0	4.6	1.9	1.9	4.5

4. For the following mating pairs and their crossover techniques, obtain the offspring chromosomes.

(a) P_1 :

A	B	C	D	E	F

P_2 :

C	A	B	F	D	E

with precedence Preservative Crossover (PPX) with a vector

1	2	1	1	2	2

3

(b) P_1 :

1	1	0	1	0	0	0	1

P_2 :

0	1	1	0	1	0	1	1

P_3 :

0	1	1	0	1	1	0	0

with three Three Parent Crossover. 3

(c) For the following offspring in Real-coded GA, obtain three mutated offspring chromosomes using the *Random mutation* scheme.

$P_1 = 10$ and $P_2 = 30$

Assume any other value(s) which might involve in the calculation. 3

--*--