Multiple Choice Questions - write the letter of your choice on your answer paper

1. Alice is writing an EA to solve the binary knapsack constraint satisfaction problem. The sum of the item costs is 37 while the total cost is constrained to be below 36. Should she: [2]
   (a) Ignore the constraints under the motto: all is well that ends well.
   (b) Upon generating an infeasible solution, immediately kill it and generate a new solution; repeat this step until a feasible solution is generated.
   (c) Employ a penalty function that reduces the fitness of infeasible solutions, preferably so that the fitness is reduced in proportion to the number of constraints violated, or to the distance from the feasible region.
   (d) Employ a repair function that takes infeasible solutions and “repairs” them by transforming them into a related feasible solution, typically as close as possible to the infeasible one.
   (e) Employ a closed feasible solution space which guarantees that the initial population consists of feasible solutions only and all evolutionary operations on feasible solutions are guaranteed to result in feasible solutions. Typically a combination of custom representation, initialization, recombination, and mutation is employed to achieve this.
   (f) Employ a decoder function that maps genotype space to phenotype space such that the phenotypes are guaranteed to be feasible even when the genotypes are infeasible. Typically this involves mapping multiple different genotypes to the same phenotype.

2. The Iterated Prisoner’s Dilemma problem: [2]
   (a) is technically not a competitive coevolution problem because it is a single population problem
   (b) is technically not a competitive coevolution problem because it is a single species problem
   (c) both of the above
   (d) none of the above

3. Speciation is: [2]
   (a) when sub-populations of different species in the same local environmental niche adapt homogeneously to the extent that they become mating-compatible
   (b) when geographically separated sub-populations of a species adapt to their local environmental niches to the extent that they become mating-incompatible
   (c) all of the above
   (d) none of the above

4. In Diffusion Model EAs: [2]
   (a) the population is conceptually distributed on a grid and mating is restricted to demes
   (b) individuals are modeled by diffusion equations and only panmictic mating is permitted
   (c) all of the above
   (d) none of the above
5. In Fitness Sharing: [2]

(a) new individuals replace similar population members, resulting in the population sharing the niches equally
(b) the fitness of individuals immediately prior to selection is adjusted according to the number of individuals falling within some prespecified distance of each other
(c) individuals share the fitness of similar population members immediately prior to selection, resulting in the number of individuals per niche being dependent on the niche fitness
(d) none of the above

6. A shortcoming of the GPS-EA + ELOOMS hybrid in terms of population size control is: [2]

(a) that it can overshoot the optimal population size because of the stochastic nature of the doubling criterion
(b) its inability to reuse the high-quality individuals identified in preceding populations
(c) the lack of support for dynamic population size control
(d) all of the above
(e) none of the above

7. Dawkin’s concept of a “meme” is: [2]

(a) a unit of biological transmission
(b) a unit of cultural transmission
(c) a process of imitation
(d) the addition of a learning phase to the evolutionary cycle
(e) all of the above
(f) none of the above

**Regular Questions**


9. (a) What is the binary gray code for the standard binary number 111101111? [3]
    (b) What is the standard binary number encoded by the binary gray code 001001011000? [3]

10. Describe concisely three different approaches for speeding up an EA being applied to a computationally expensive simulation-based optimization problem, given that you cannot change the hardware. [6]

11. Given the following two parents with permutation representation:
    
    \[ p_1 = (829163574) \]
    \[ p_2 = (986412357) \]

    (a) Compute the first offspring with Cycle Crossover. [4]
    (b) Compute the first offspring with PMX, using crossover points between the 3rd and 4th loci and between the 7th and 8th loci. [5]
    (c) Compute the first offspring with Order Crossover, using crossover points between the 2nd and 3rd loci and between the 6th and 7th loci. [3]
    (d) Compute the first offspring with Edge Crossover, except that for each random choice you instead select the lowest element. [8]
12. Assuming a simple genetic algorithm whose global optimum has a fitness of 100.0 and given the following bit strings $v_1$ through $v_5$ and schema $S$

$v_1 = (01110110011001) \text{ fitness}(v_1) = 88.0$
$v_2 = (01110111011001) \text{ fitness}(v_2) = 1.0$
$v_3 = (01110110111001) \text{ fitness}(v_3) = 1.0$
$v_4 = (11110110011000) \text{ fitness}(v_4) = 1.0$
$v_5 = (11110110011001) \text{ fitness}(v_5) = 2.0$

$S = (*1110110011001)$

(a) Compute the order of $S$. [1]
(b) Compute the defining length of $S$ and show your computation. [1]
(c) Compute the fitness of $S$ and show your computation. [1]
(d) Do you expect the number of strings matching $S$ to increase or decrease in subsequent generations? Explain your answer! [4]

13. Say you want to purchase a new house and care most about maximizing square footage and minimizing price. You collect square footage data and pricing on ten different houses and then you normalize both the square footage data and the pricing which results in the following table, where higher square footage numbers indicate greater square footage and higher pricing numbers indicate better affordability (so lower price):

<table>
<thead>
<tr>
<th>ID</th>
<th>Square footage</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

(a) Plot the above table and use dotted lines to indicate the area of domination for each element. [2]
(b) List for each element which elements it dominates; indicate elements with their IDs. [2]
(c) Show the population distributed over non-dominated levels, like some multi-objective EAs employ, after each addition of an element, starting with element 1 and ending with element 10 increasing the element number one at a time; indicate elements with their IDs. So you need to show ten different population distributions, the first one consisting of a single element, and the last one consisting of ten elements. [6]

14. The standard linear ranking selection formula in your textbook is as follows:

$$P_{lin-rank}(i) = \frac{2 - s}{\mu} + \frac{2i(s-1)}{\mu(\mu-1)}$$

with rank 0 being the least fit individual and rank $\mu - 1$ being the fittest individual. Now assume the generational model where $\mu$ population members are selected with replacement to fill the mating pool.

(a) Show why $s$ cannot be larger than 2. [3]
(b) Show your derivation of the expected number of copies of the fittest individual in the mating pool. [4]