We are driven by five genetic needs: survival, love and belonging, power, freedom and fun.

-- William Glasser
Question 1 (8 Marks)

Hill-climbing algorithms take the best solution seen so far and try to improve it. Genetic algorithms maintain a population of solutions, some of which may be strong and some of which may be weak. What is the purpose of keeping weak solutions in the population?

Solution: The weak solutions may contain some values or combinations of values that can combine with other solutions to give new solutions that are stronger than any of the solutions that could be found by making small improvements to the strong solutions.

For example (students are not required to provide an example) in a 01-Knapsack application, there might be a very weak solution that only contains a few very small elements of the set ... but these might be exactly the elements needed to change one of the good solutions into an optimal (or nearly optimal solution). By keeping some weak solutions in the set, we preserve the ability of the algorithm to build on the different strengths of all the current members of the population.

Marking 8/8 for understanding that the weaker solutions are essential to the ability of the algorithm to create a new generation that is significantly different from the previous generation, yet built from it. Another way to express this is that the weaker solutions help us avoid developing a monoculture.

4/8 for an answer that mentions “genetic diversity” without clarifying the value of that in this context

1/8 for trying
Question 2 (12 Marks)

Consider this Tournament system for assigning pairing points to solutions – this is based on a typical elimination sports tournament:

Randomly put the solutions in pairs to compete. Each winner gets one “point” and moves on to the next round. Each loser is eliminated. In case of a tie, arbitrarily choose one of the two solutions to be the winner.

Repeat the process until only one solution remains.

Now allocate pairings by guaranteeing each solution at least as many pairings as it has points (this stage is exactly the same as the Tournament method we discussed in class).

As an example, suppose we have four solutions $P_1$ to $P_4$, with values as shown here:

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>3</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

The tournament might play out like this.

$P_3$ ends up with 2 points, $P_1$ gets 1 point, and the others get 0 points. Given different initial matches (eg $P_1$ vs $P_3$, $P_2$ vs $P_4$), the results would be different.

This question continues on the next page.
Suggest at least one advantage and at least one disadvantage of this method in comparison to the Tournament method we discussed in class (in general, not just for this example).

Solution:

Advantages:
- No member of the current population is “left out” of the tournament
- The strongest solution in the current population is guaranteed the most points

Disadvantages:
- The number of “head-to-head” battles is \( O(n) \). If \( n \) is large, this process will be time-consuming when compared to the traditional tournament method
- Even though the strongest solution is guaranteed the most points in the tournament, the second strongest solution may not get any points at all

Marking 6 for a sensible advantage (there are probably others that I haven’t thought of)

6 for a sensible disadvantage (there are probably others that I haven’t thought of)
Question 3: (28 marks)

Because of your work on the COMPSA Canoe Trip, the Queen’s Unlimited Activities Requiring Really Excellent Lunches Society (QUARRELS) has asked you to plan their canoe trip. Fortunately they have a different problem to solve. They plan to make a simple canoe trip down Coyote Creek, with 4 people in each canoe (luckily \( n \), the number of people in the club is a multiple of 4). All you have to do is decide on the groups of 4 people for each canoe.

The problem is – as you might guess from the name of the club – the members of QUARRELS don’t get along very well. This is quantifiable: for each pair of people \( p_1 \) and \( p_2 \), the function \( grrr(p_1, p_2) \) returns an integer from 1 to 10 that indicates how strongly \( p_1 \) dislikes \( p_2 \) (1 means \( p_1 \) dislikes \( p_2 \) a little bit, and 10 means \( p_1 \) absolutely despises \( p_2 \)).

Each person’s unhappiness score is equal to the sum of their dislike for the three other people in their canoe.

Your job is to assign them to canoes (4 persons to a canoe, as mentioned) so that the total amount of unhappiness is small. You have decided that this is a perfect opportunity to use a genetic algorithm.

Design a genetic algorithm to search for a good solution to this problem. Your goal is to minimize the sum of the unhappiness of the members of the club.

There are four parts to this question. Each part is worth 7 marks.

For most parts of this question there is more than one reasonable answer. Whatever answer you give, explain your reason for choosing it.
(a) Describe how you will represent each solution as a vector. Specify clearly what the individual elements of the vector represent.

Solution:

One possibility: we can think of lining up the people in a single row - the first four people go in the first canoe, the second four people go in the second canoe, etc. The order of the people in the line determines the canoe groupings. So the vector will simply be a permutation of the people – any such vector represents a feasible solution.

This will generate some duplicate solutions (for example [1,2,3,4,5,6,7,8], [5,6,7,8,1,2,3,4] and [4,3,2,1,8,5,7,6] all represent the same 2 groups in 2 canoes).

Another possible solution is to use a vector of length n, with the \( i^{th} \) element representing the \( i^{th} \) person. Each element of the vector would be assigned the number of the canoe that person is in. Feasibility would consist of having exactly 4 persons assigned to each canoe. This method also allows for different representation of the same groups.

For each of these there are ways to arrange the information so to make it easier to identify duplicates, but the question did not ask for that!

Marking 7/7 for any workable method that makes it clear which persons are assigned to each canoe

3/7 for a representation that doesn’t work, such as representing each solution by a vector of n bits – this doesn’t give enough space to identify the groups of 4 the people are assigned to.

1/7 for trying
(b) Describe how you will reduce the size of the current population when it exceeds the population limit, and give the reason for your choice.

Solution: I would use “elitism”: preserve some percentage of the previous population, and supplement that with enough of the new generation to bring the population up to the desired value.

I would choose this because it keeps the best existing solutions but also maintains a high degree of genetic diversity in the population.

Cases can also be made for other selection methods, such as “survival of the fittest” or “kill the old folks”

Marking: 7/7 for any answer that identifies a method and gives a rational reason
3/7 for giving a method but no reason
1/7 for trying
(c) Describe the cross-over operation you will use, and explain your choice.

Solution: The “Travelling Salesman Cross-over” would be a reasonable choice: pick two values \( k_1 \) and \( k_2 \), and swap the \([k_1 \ldots k_2]\) pieces of the parents to produce the off-spring. The rest of the elements of \( \text{Child}_1 \) are the elements of \( P_1 \) that were not brought in from \( P_2 \), left in their original order. Similarly the rest of the elements of \( \text{Child}_2 \) are the elements of \( P_2 \) that were not brought in from \( P_1 \), left in their original order.

This answer is specific to the first answer given in (a). Other representations of the individual solutions could lead to other answers here. For example if the second representation described in the answer to (a) is used, the same “range-swap” operation could be used but different steps would be needed to ensure that the off-spring represent feasible solutions.

Marking 7/7 for any answer that identifies a method and gives a rational reason
3/7 for giving a method but no reason
1/7 for trying

(d) Describe the mutation operation you will use, and explain your choice.

Solution: There are a number of reasonable methods. One simple operation is to choose two people who are in different canoes, and swap them.

Marking 7/7 for any answer that identifies a method and gives a rational reason
3/7 for giving a method but no reason
1/7 for trying
Question 4: (2 marks)

True or False: When it first appeared, the term “genetic algorithm” was considered to be a typographic error for “generic algorithm”

Solution: False

Marking

2/2 if they get it right
0/2 if they get it wrong