1. [5 Points] What is metaheuristic optimization?

Recent literature refers “metaheuristics” as Algorithms with stochastic components, meaning there is some randomness in the algorithm. This term was introduced by Fred Glover in his seminal paper (Glover 1986). By his convention this referred to all modern nature inspired algorithms, in his words they are a "*master strategy that guides and modifies other heuristics to produce solutions beyond those that are normally generated in a quest for local optimality*" (Glover and Laguna 1997). Metaheuristic optimization intends to seek a globally best solution to a global optimization problem using metaheuristic algorithms. It can bring about quality solutions to hard optimization problems better than gradient-based algorithms and derivative-free algorithms, but it cannot be verified if it reaches the optimal solution. The two major constituents of such algorithms are exploration and exploitation. Exploration means to explore the search space on a global scale and bring about diverse solutions; this is achieved via randomization and allows the search to escape from the local optima. On the other hand, exploitation refers to concentrating search on local region where it is known that a current good solution lies; this convergence to an optimum is ensured through selection of the best solutions. A suitable mix of these two components will usually make global optimality achievable.

1. [3 Points] What are the situations in which gradient based optimization techniques do not work?

It is very difficult to apply gradient based solutions for non-linear, multi-modal, multivariate functions. Moreover, some objective functions can have discontinuities which would hinder the performance of gradient based algorithms. It may also be computationally expensive to calculate derivatives accurately, especially with higher number of decision variables and multiple peaks in the function. Furthermore, noise or uncertainty in the objective and constraint functions can also be detrimental to gradient based approaches.

1. [5 Points] Biefly explain any one swarm based algorithm that we have NOT discussed in the class

The **Bee Algorithm** is part of the trend of algorithms that utilize swarm intelligence. It is inspired by the food forging behaviour of honey bees. Forager scouting bees from the hive explore different random areas in search of flowers with abundant and good pollen. When they discover a location with a good nectar source they bring the nectar back to the hive and communicate the direction, distance, and quality of the food source to fellow bees in the hive through a ‘waggle dance’ that is a signaling system. Using this information more bees are recruited to follow the specified locations for more nectar. This strategy allows efficient exploration of nectar over huge distances.

This algorithm works in a similar way to this principle, it initializes a population of randomly distributed solutions, and it defines neighbourhood regions around some of those solutions primarily around those of a high fitness value, which are marked as ‘elite bees’ and are saved for the next generation . It searches more solutions in each neighbourhood region and finds the solution with the highest value in each such region adds them to the next generation as well. The remaining population is filled by randomly assigning new solutions, and the process is repeated again for several iterations until the population converges to certain solutions.

This was initially created at Oxford, under the name of Honey Bee Algorithm, for allocating computers among different web-hosting servers and clients. Later on other variants were developed such as Virtual Bee Algorithm (VBA) to solve continuous optimization problems and Artificial Bee Colony (ABC) algorithm for numerical function optimization.