TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA UNIVERSITY OF WELLINGTON

**EXAMINATIONS – 2021**

**TRIMESTER TWO**

**AIML 426**

**EVOLUTIONARY COMPUTATION AND LEARNING**

**Time Allowed:** ONE HOUR

**CLOSED BOOK**

**Permitted materials:** Only silent non-programmable calculators or silent programmable calculators with their memories cleared are permitted in this examination.
Non-electronic foreign language translation dictionaries may be used.

**Instructions:** You have one hour to work on the test paper.
There are a total of 60 marks on this test.
Attempt all questions, skip to the next question if the current one is hard.

## Questions

| | **Marks** |
|---|---|
| 1. Swarm Intelligence and Other EC Paradigms | [20] |
| 2. Evolutionary Multi-objective Optimisation | [20] |
| 3. Reinforcement Learning and Neuro-evolution | [20] |

1. **Swarm Intelligence and Other EC Paradigms** **(20 marks)**

(a) **(5 marks)** The following shows the particle update formulas of the standard Particle Swarm Intelligence (PSO).

$$v_{id}(t+1) = w * v_{id}(t) + c_1 * r_1 * (P_{id} - x_{id}(t)) + c_2 * r_2 * (P_{gd} - x_{id}(t)),$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1),$$

where $P_{id}$ is the dimension $d$ of the personal best of particle $i$, $P_{gd}$ is the dimension $d$ of the global best of the entire swarm. That is, the topology is fully connected.

Suggest *three* ways to *increase the exploration ability* of the standard PSO.

(b) **(5 marks)** When using Ant Colony Optimisation (ACO) to solve traveling salesman problem, the probability of going from the current node $i$ to an unvisited node $j$ is:

$$p_{ij} = \frac{[\tau_{ij}]^{\alpha} * [\eta_{ij}]^{\beta}}{C},$$

where $\tau_{ij}$ is the *pheromone*, and $\eta_{ij}$ is the *heuristic information* between nodes $i$ and $j$.

**(i)** Briefly describe the meaning of $\tau_{ij}$ and $\eta_{ij}$.

**(ii)** Suggest a good way to set $\eta_{ij}$ for each edge $(i, j)$, and briefly justify your reason.

(c) **(3 marks)** Canonical Differential Evolution (DE) has three four steps at each generation: (1) parent selection, (1) environmental selection, (3) crossover, (4) mutation. State the order of these steps at each generation of DE.

(d) **(2 marks)** State a main difference between memetic algorithm and traditional genetic algorithm.

(e) **(5 marks)** This question is about optimising the following function using cooperative co-evolution.

$$\min f(x_1, x_2, x_3, x_4, x_5) = x_1 * x_2 + x_3 / x_4.$$

There are two subpopulations, where the first subpopulation evolves $[x_1, x_2]$, and the second one evolves $[x_3, x_4]$.

Suppose the current context vector is $\mathbf{cv} = [1, 1, 1, 1]$, calculate the fitness of the following individuals. Show your steps.

**(i)** $[x_1 = 0, x_2 = 3]$ in the first subpopulation;
**(ii)** $[x_3 = 2, x_4 = 4]$ in the second subpopulation.

2. **Evolutionary Multi-objective Optimisation** **(20 marks)**

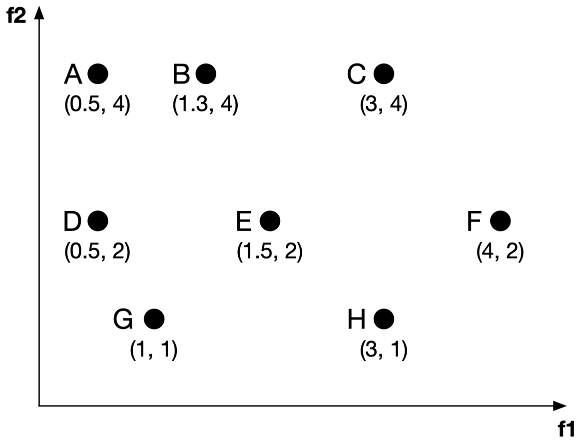
(a) **(5 marks)** Briefly describe how to calculate the *Hyper-Volume (HV)* of a set of non-dominated solutions **X**. You can draw a graph to help explanation.

(b) **(5 marks)** Fitness assignment is an important design issue in evolutionary multi-objective optimisation. Briefly describe the high-level idea of MOEA/D to address the fitness assignment issue. No mathematical formula is needed.

(c) **(5 marks)** Consider a multi-objective optimisation problem that *minimises* two objectives ($f_1$ and $f_2$). Below shows the two objective values of 8 different solutions, denoted as $A, ..., H$. For example, $A = (0.5, 4)$ means that $f_1(A) = 0.5$, and $f_2(A) = 4$.

**(i)** If we conduct non-dominated sorting (used in NSGA-II) to sort these solutions, calculate the rank of each solution in the figure (the rank starts from 1).

**(ii)** What is the crowding distance of the solution $G$?




(d) **(5 marks)** When designing multi-objective PSO, a key issue is to design the *gbest* and *pbest* update schemes. Suggest a scheme for updating *gbest* and *pbest* in multi-objective PSO, and briefly justify your design.

3. **Reinforcement Learning and Neuro-evolution** **(20 marks)**

(a) **(5 marks)** As shown in the figure below, a grid world Markov Decision Process (MDP) operates like the one we saw in class. The states are grid squares, identified by their row and column number (row first). The agent always starts in state (1,1), marked with the letter S. There are two terminal states, (2,3) with reward +5 and (1,3) with reward -5. Rewards are -0.5 in all non-terminal states. The reward for a state is received as the agent moves into the state. The transition function is such that the intended action/movement (North, South, West, or East) happens with probability 1.0. For example, if the agent performs action East in state (1,1), it will transit deterministically to the new state (1,2). If a collision with a wall happens, the agent stays in the same state.

**(i)** Draw the optimal policy for this grid world.

**(ii)** Identify the optimal state value function (i.e., $V*(s)$) with respect to every non-terminal state in the grid, assuming that the discount factor $\gamma = 1.0$.

| | | |
|---|---|---|
| | | **+5** |
| **S** | | **-5** |

Figure 1: A grid world Markov Decision Process.

(b) **(2 marks)** Suppose you have a robot trying to reach a goal and avoid cliffs in a small grid world. It can only move North, South, East, or West. If you were to solve this problem using either Q-learning or Proximal Policy Optimization (PPO), which algorithm should you consider first? Justify your answer in one sentence.

(c) **(5 marks)** Explain briefly the main difference between an *on-policy* reinforcement learning algorithm and an *off-policy* reinforcement learning algorithm. Give one *on-policy* reinforcement learning algorithm, and one *off-policy* reinforcement learning algorithm (name is enough).

(d) **(3 marks)** Describe briefly the main procedure for an actor-critic reinforcement learning algorithm to train a parameterized policy.

(e) **(5 marks)** Neuro-evolution algorithms are important evolutionary computation technologies for solving reinforcement learning problems.

**(i)** Identify two main advantages of using neuro-evolutionary algorithms for reinforcement learning.

**(ii)** The NeuroEvolution of Augmenting Topologies (NEAT) algorithm is a representative neuro-evolution algorithm. Discuss briefly a key technique used by NEAT to maintain topological innovation during the evolutionary process.

Student ID: ........................

**SPARE PAGE FOR EXTRA ANSWERS**

Cross out rough working that you do not want marked.
Specify the question number for work that you do want marked.

**SPARE PAGE FOR EXTRA ANSWERS**

Cross out rough working that you do not want marked.
Specify the question number for work that you do want marked.

* * * * * * * * * * * * * * * *