25-26 COMP0215 Coursework

Reinforcement Learning in Simple Games

1. Project Overview

The objective of this coursework is for each team to develop an autonomous agent capable of learning and evolving its own strategies using Reinforcement Learning.

Core Task: Develop an RL agent for Tic-Tac-Toe (Connect 4 on a 4×4 grid, can be extended to higher dimensions).

Challenge Task (optional): Develop an RL agent for Gomoku (Five-in-a-Row) on a 9×9 or larger board.

Special Incentive: Teams that successfully implement a high-performing Gomoku agent are **exempt** from the Tic-Tac-Toe task and will receive higher weight in the final evaluation.

2. Team Structure

Size: 3-4 students per group.

Suggested Roles:

Algorithm Lead: Designing the RL logic and reward structures.

Environment Engineer: Building the game logic and API interfaces.

Data Analyst: Handling training logs, hyperparameter tuning, and visualization.

Project Manager/Editor: Managing documentation, a report and a video to demonstrate your work.

3. Requirements

The agent must improve its performance from zero knowledge or random play through a reinforcement learning mechanism.

The agent is encouraged provide a standardized function: predict(board_state)-> (x,y) to facilitate cross-group testing.

Strict Prohibitions:

Submitting pre-trained weights from online repositories without original training scripts.

(TAs will check with online repositories).

Using purely hard-coded if-else rules (Heuristics) without a learning component.

4. Deliverables

(1) Source Code: Complete training scripts, game environment, and a ReadMe for deployment.

(2) Model Weights: The final trained policy (e.g., Q-Table files or Neural Network weights).

(3) Technical Report:

Description of the chosen model architecture and RL strategy.

Quantitative analysis: Learning curves (Win Rate/Reward vs. Episodes), ablation studies, or hyperparameter analysis.

Documentation of team contributions.

(4) Short Video (Demo):

Duration: 3-5 minutes.

Content: A brief explanation of the core methodology and a live/recorded demo of the agent in action (e.g., Human vs. Al or Model Version A vs. Model Version B).

5. Grading Rubric

Criteria	Weight	Standards for Excellence
Task Difficulty	30%	Successful implementation of Gomoku (9×9 or 15×15) or 4×4 Tic-Tac-Toe.
Agent Intelligence	30%	Performance against non-random opponents; ability to execute strategic moves (offense/defense/traps).
Experimental Depth	20%	Depth of analysis regarding training stability, reward shaping, and parameter tuning.
Video & Presentation	20%	Clarity of communication, logical flow, and visual demonstration of the agent's decision-making.

6. Important Milestones

Progress Check (Preliminary Results) report (1 page): 16th March

Final Submission deadline: 30th March