Coursework 2 – Tic-Tac-To: Markov Decision Processes & Reinforcement Learning (worth 25% of your final mark)

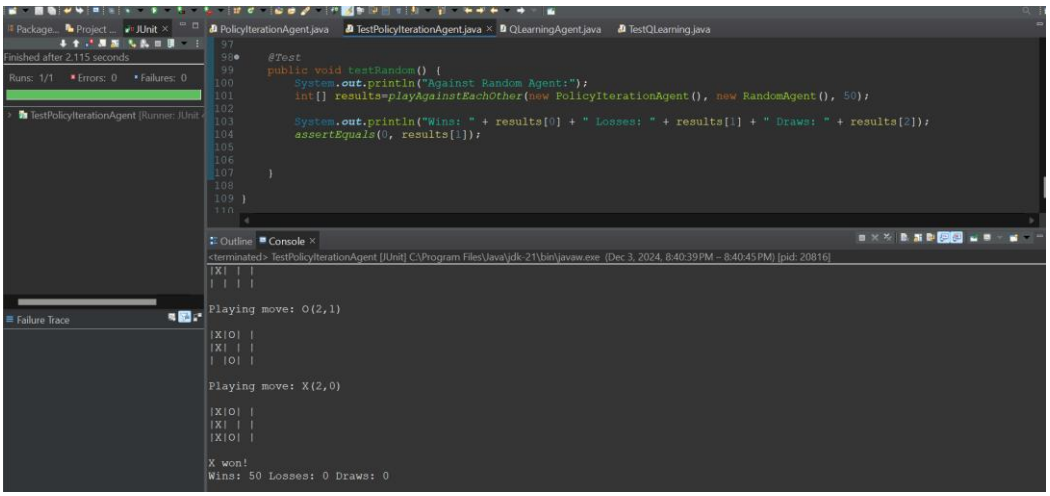
POLICY ITERATION

Question 4 (1 point): As in Question 2, this time test your Policy Iteration Agent against each of the provided agents 50 times and report on the results – how many games they won, lost & drew. The other agents are: *random*, *aggressive*, *defensive***Ans**:

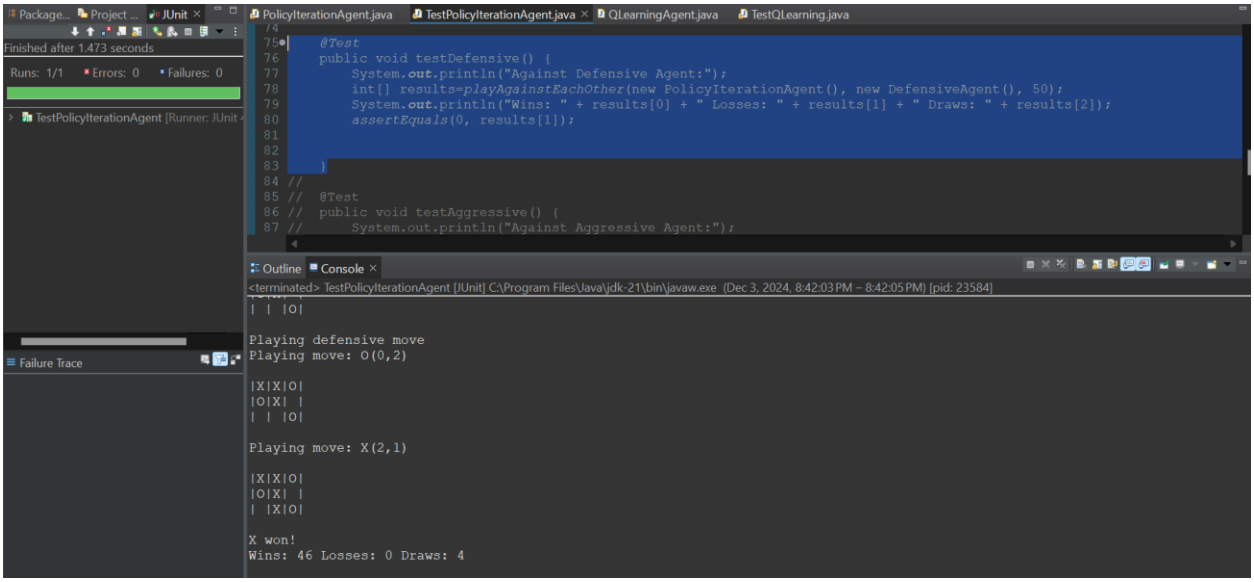
- initRandomPolicy():initializes random move generator and assigns random move to non terminal game state
- evaluatePolicy():finds maximum change less than epilson and updates policy values.
- improvePolicy():replaces actions with best moves and changes flag if no changes were made (ie if stable)
- Train() : initializes and executes the algorithm using evaluatePolicy() and improvePolicy(), terminates and updates policy when stable.

	WINS	LOSS	DRAWS
RANDOM	50	0	0
AGRESSIVE	50	0	0
DEFENSIVE	46	0	4

Random:



DEFENSIVE:



AGGRESSIVE: