LAB #10

Reinforcement Learning Diagram

(10 pts)

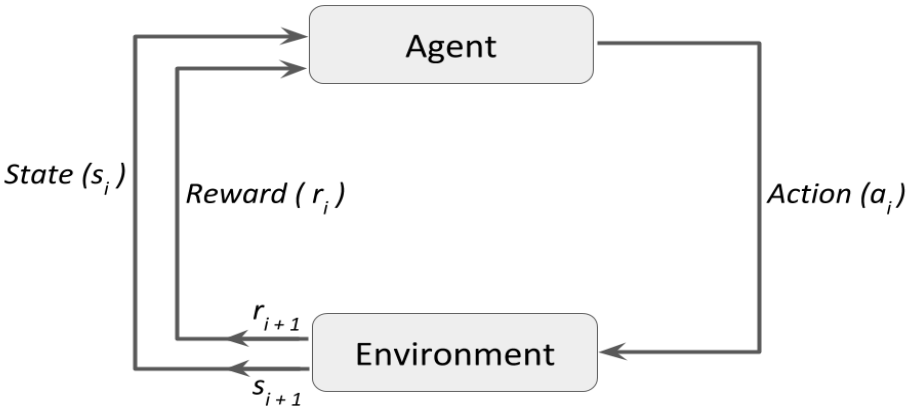
Due by Nov 1st Wed 4 pm

Constructing the environment for RL:

You and your partner are to provide **six (6)** different environments by defining and listing the items below. The environment must be in a different genre/category. For each item below, clearly define and explain so a third person can understand without asking you a question about it.

For the State (s), name at least ten states, and with each state, name the possible actions, and with each action, there must be a value and probability (you make fictitiously make these). Illustration is a MUST for clarity.

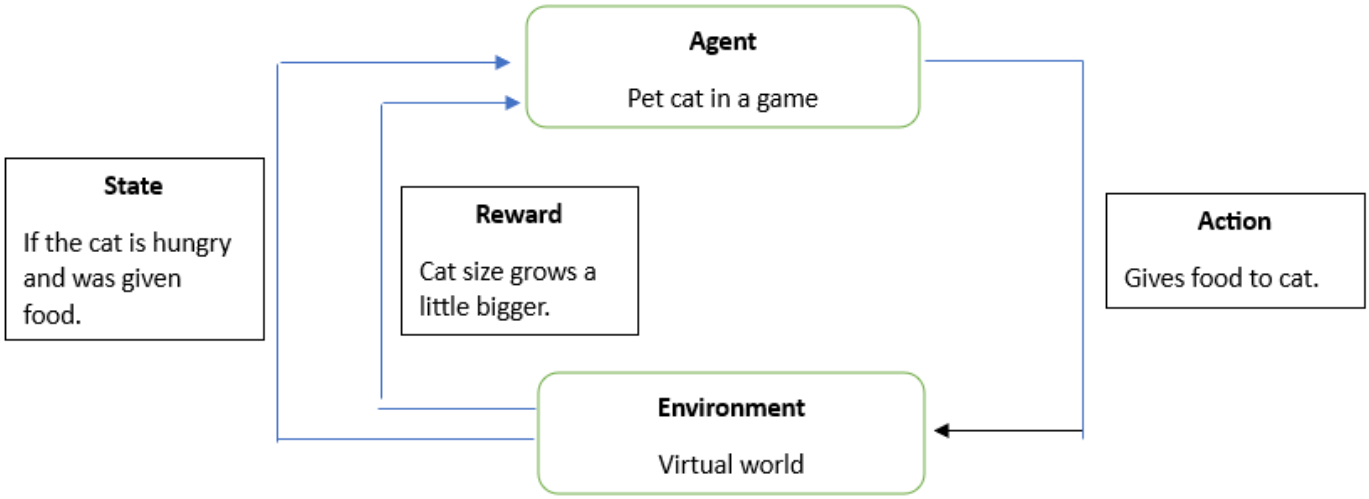
Complete your work using MS Word or PPT and submit via Canvas.



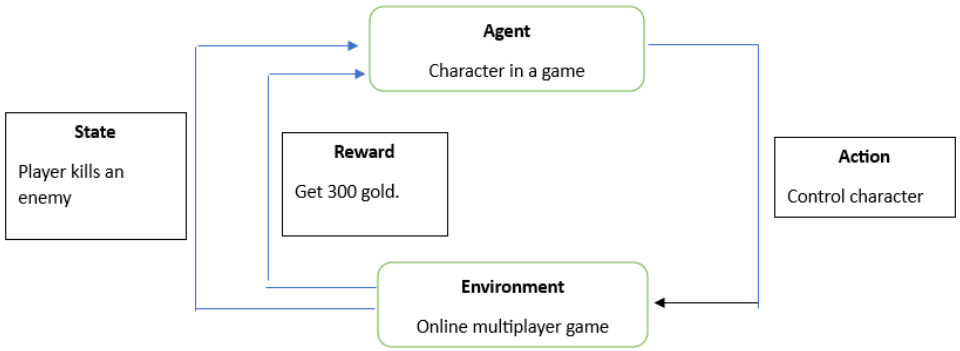
Agent:	A hypothetical entity that performs actions in an environment to gain some reward.
Action (a):	All the possible moves that the agent can take.
Environment (e):	A scenario the agent has to face.
State (s):	Current situation returned by the environment.
Reward (R):	An immediate return sent back from the environment to evaluate the last action by the agent.

Name: Zhen Rui Ng, Vannarith Om

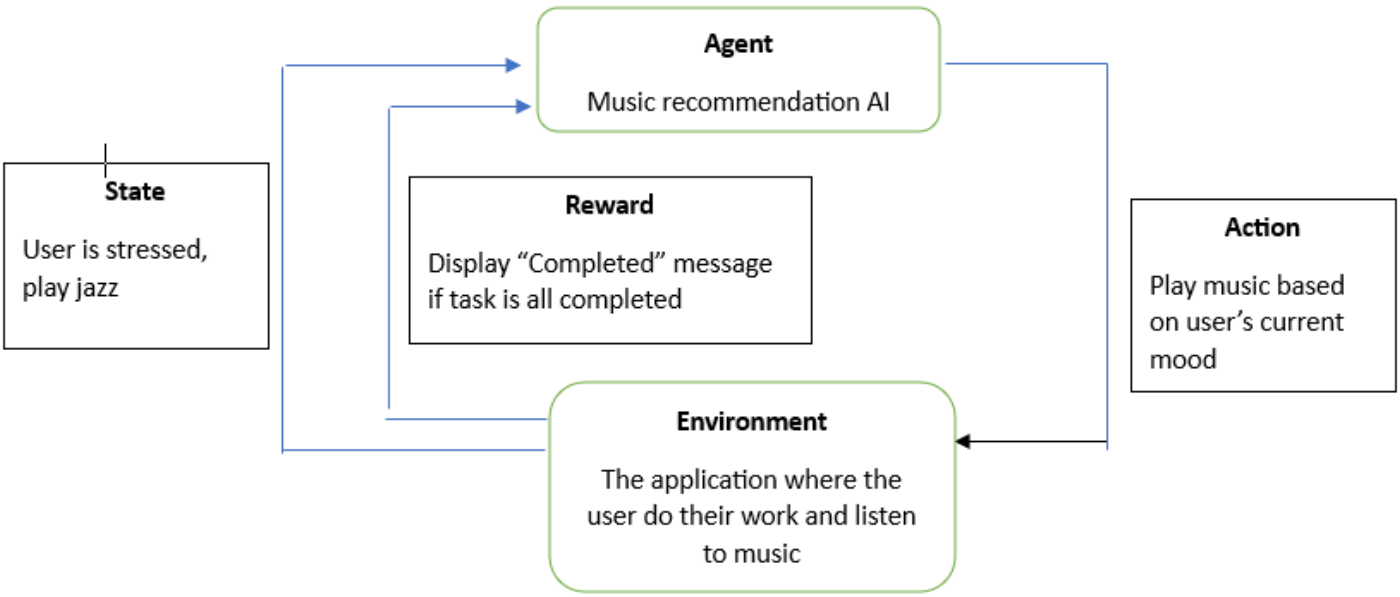
Agent	Pet cat on a cat game app.
Action	Engage with the cat by feeding and playing with it.
Environment	The virtual world that my cat lives in the app.
State	Depends on the current mood and state of the cat. 1. If the cat is hungry, it will ask for food. *Give furniture (12%)*. 2. If the cat is sad, it will ask for cuddles. *Give cat toys (1%)*. 3. If the cat is happy, it will smile. *Give cat food (44%)*. 4. If the cat is lonely, it will want to play. *Give decorations (31%)*. 5. If the cat is angry, it will run away. *Give cat poop (75%)*. 6. If the cat is afraid, it will hide. *Stay hidden (64%)*. 7. If the cat is dirty, it will ask for a bath. *Escape from bath (53%)*. 8. If the cat is hurt, it will cry. Cat gets sick (80%). 9. If the cat is tired, it will sleep. The cat is energetic (100%). 10. If the cat is curious, it will chase a butterfly. Cat gives furniture (56%).
Reward	Take care of the cat for 5 days and 5 nights in-game time, then the cat will grow bigger and bigger. Each level reached will give you more food, cat toys, and furniture for the cat.


Agent
Pet cat in a game
State
If the cat is hungry and was given food.
Reward
Cat size grows a little bigger.
Action
Gives food to cat.
Environment
Virtual world

Agent	The character you choose to play.
Action	Control your character to win the game by destroying the enemy turrets and nexus (home).
Environment	An online multiplayer game.
State	1. The player kills an enemy. Get 30 2. Player get killed 3. Player can move around to help teammates. 4. Player can kill monsters. 5. Player can take down turrets. 6. Player have to wait a couple of seconds if they get killed. 7. Player can steal the monster. 8. Player can protect turrets 9. Player can use characters' abilities 10. Player can get stronger with more weapons.
Reward	If a player kills an enemy, get gold. More enemies, more gold. Take down the enemy home, to win the game.

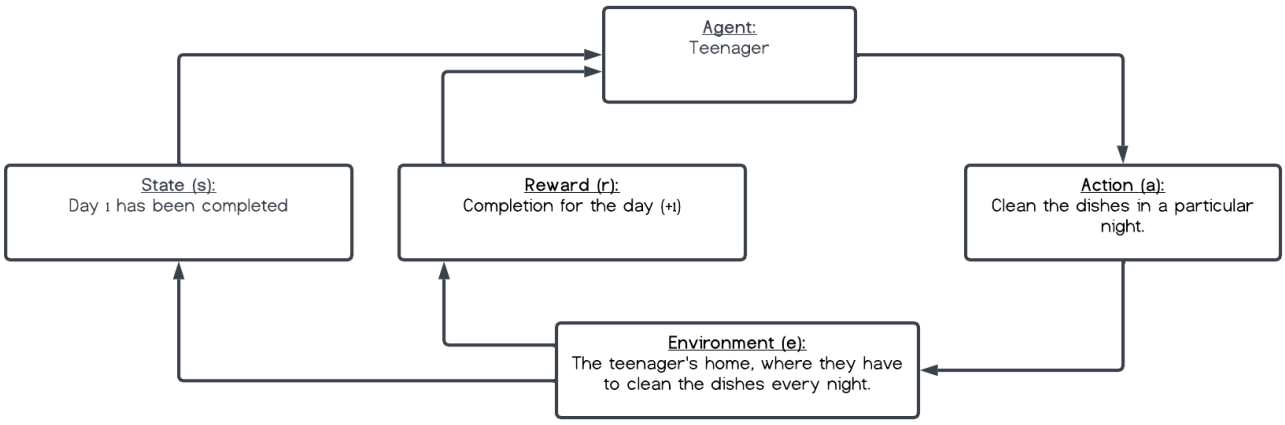


Agent	Music Recommendation AI
Action	Play different types of music based on the current mood/ productivity of the user.
Environment	The application or platform where the user is doing their work and listening to music.
State	1. User is focused on tasks, lower volume or plays no music. 2. User is multitasking, plays classical music. 3. User is stressed, playing jazz music. 4. User is creative, play pop music. 5. User is relaxed, plays the blues. 6. User is energetic, play k-pop music. 7. User is lazy, play rock music. 8. User is emotional, play Vietnamese music. 9. User is happy, play j-pop music. 10. User is nervous, playing country music.
Reward	The AI will ask the user how much work they have done to display different messages. If tasks completed <= 50% then display "Keep working!" message. If tasks completed <= 70% then display "Almost done…" message. If tasks completed = 100% then display "Completed!" message.


Agent
Music recommendation AI
State
User is stressed, play jazz
Reward
Display "Completed" message if task is all completed
Action
Play music based on user's current mood
Environment
The application where the user do their work and listen to music

Example #1: A teenager that has to clean the dishes every night in a row for a whole week (seven days) in order to have permission to attend a music concert.

Agent	The teenager who wants to attend the music concert.
Action (a)	The teenager could: • Cleaning the dishes on a particular night. • Not cleaning the dishes on a particular night.
Environment (e)	The teenager's home, where they have to clean the dishes every night.
State (s)	The state represents the current situation, which is returned by the environment. In this scenario, the state can be defined as the number of days the teenager has cleaned the dishes. There are seven possible states, one for each day of the week. We'll name the states from "Day 1" to "Day 7."
Reward (R)	• If the teenager cleans the dishes on a given day, the reward is the completion of that day. • If the teenager does not clean the dishes on a given day, he receives no rewards at all. • If the teenager cleans the dishes for seven consecutive days, on the seventh day, they receive the big reward of permission to attend the music concert.

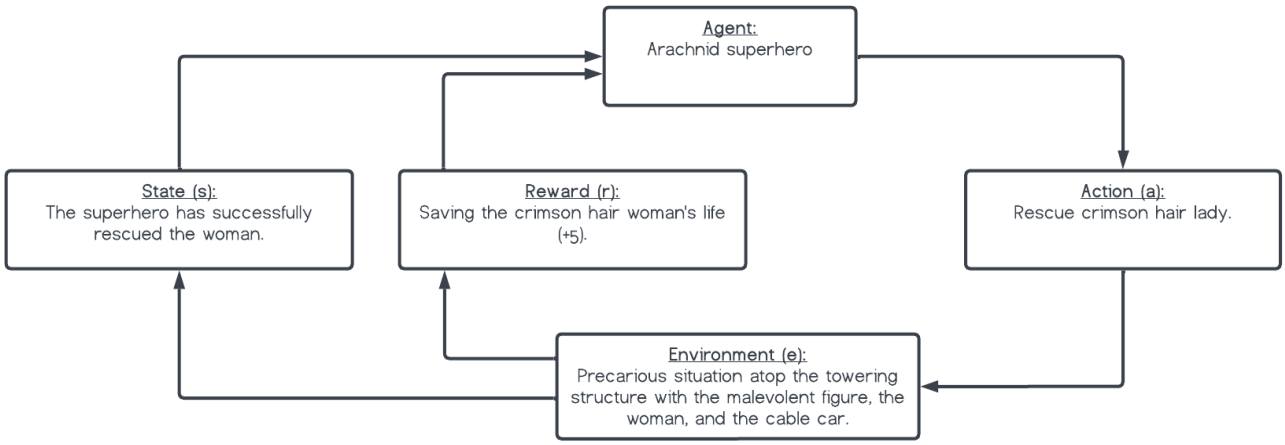


Example #2: A very green malevolent figure stands atop a towering structure. On one hand, the villain tightly clutches a woman with crimson hair, who happens to be the love interest of an arachnid-themed superhero. On the other hand, the villain grasps a cable car filled with innocent children. With a sinister grin, the villain simultaneously releases both captives. A

heart-pounding freefall ensues, offering a grim fate unless they are intercepted mid-plunge. Now, the superhero must make the harrowing decision of whom to rescue.

Agent	The superhero, who is faced with the decision of whom to rescue.
Action (a)	The possible moves/actions that the superhero can take in this scenario. The superhero's actions could be: • Rescue the love interest (woman with crimson hair). • Rescue the children in the cable car.
Environment (e)	The scenario that the superhero has to face is the precarious situation atop the towering structure with the malevolent figure, the woman, and the cable car.
State (s) State A: Save both B: Don't save either of them C: Save only one Rewards A: +10 B: - 2 C: -1 Probabilities From A to B: 0 From A to A: 1.0 From B to C: 0.6 From B to A: 0.5	In this scenario, states could represent different moments or conditions during the freefall. Here are ten example states: • The superhero is still at the top of the structure, and both the woman and the children are in the villain's grasp. • The superhero has started descending towards the woman. • The superhero has reached the woman and is attempting to rescue her. • The superhero has successfully rescued the woman. • The superhero has started descending towards the cable car. • The superhero has reached the cable car and is attempting to rescue the children. • The superhero has successfully rescued the children. • The superhero has rescued both the woman and the children. • The superhero has failed to rescue either the woman or the children. • The superhero has failed to rescue the woman but managed to save the children.
Reward (R)	The reward is the immediate return sent back from the environment to evaluate the last action by the superhero.

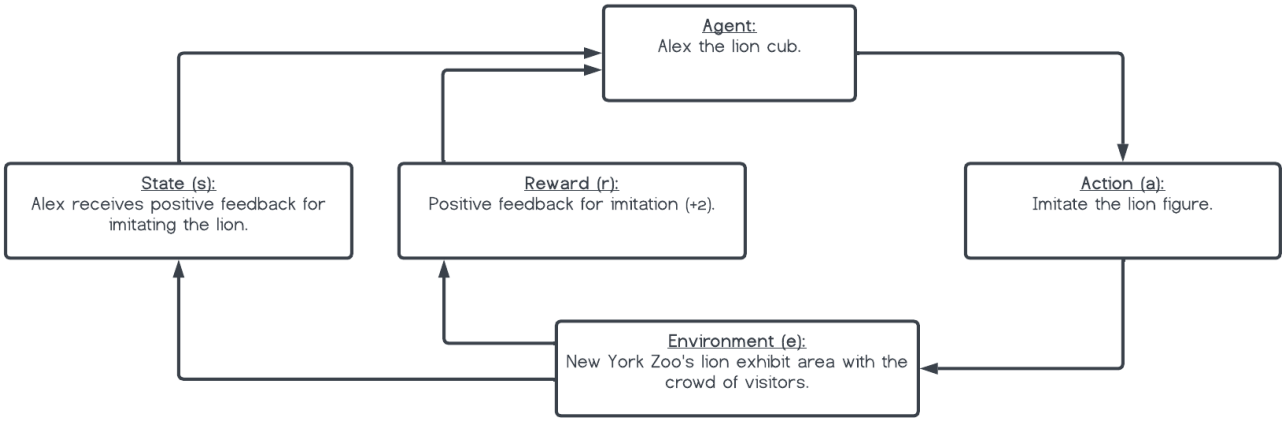
	• If the superhero successfully rescues the woman, the reward could be the life of the woman with crimson hair. • If the superhero successfully rescues the children, the reward could also be the life of the innocent children. • If the superhero fails to rescue either, there is no reward at all. • If the superhero rescues both, the reward could be both of the subject's lives.



Example #3: Alex the lion cub has recently been transferred to his own cage to be exhibited to visitors of the New York Zoo. Alex, in his eagerness to discover what he is doing in that place, tries to imitate the lion figure that is located as decoration on the stone on which he is standing. After imitating the pose, he gets a good reaction from the crowd, who are watching him from the other side of the fence. After noticing the positive feedback from the crowd, Alex decides to risk the attention that is on him by making a quick dance move. When he notices that the people's reaction turned out to be even more positive, he decides to risk himself again by making a more elaborate dance move. Fortunately, his move manages to produce even more euphoria among the public in the form of cheers and applause. With each feedback from his audience, Alex begins to make more and more complex choreography. In this way, day after day after having perfected his choreography to entertain his audience, Alex manages to earn his title as the King of the Concrete Jungle of New York.

Agent	Alex the lion cub, who is eager to entertain the visitors and gain positive feedback.
Action (a)	The possible moves/actions that Alex can take in this scenario. Alex's could: • Imitate the lion figure.

	• Perform a quick dance move. • Perform a more elaborate dance move. • Create a complex choreography.
Environment (e)	The environment in this scenario is the New York Zoo's lion exhibit area with the crowd of visitors.
State (s)	In this scenario, states could represent different stages of Alex's performance and the audience's reaction: • Alex is imitating the lion figure. • Alex receives positive feedback for imitating the lion. • Alex performs a quick dance move. • Alex receives more positive feedback for the dance move. • Alex performs a more elaborate dance move. • Alex receives even more positive feedback for the elaborate dance move. • Alex creates a complex choreography. • Alex receives enthusiastic cheers and applause from the audience. • Alex becomes known as the King of the Concrete Jungle. • Alex continues to perform his complex choreography to maintain his title.
Reward (S)	The reward is the immediate return sent back from the environment to evaluate the last action by Alex. The rewards could be based on the audience's reaction: • Positive feedback for imitation. • Positive feedback for the quick dance move. • Positive feedback for the more elaborate dance move. • Enthusiastic cheers and applause for the complex choreography. • Becoming the King of the Concrete Jungle.


Agent:
Alex the lion cub.
State (s):
Alex receives positive feedback for imitating the lion.
Reward (r):
Positive feedback for imitation (+2).
Action (a):
Imitate the lion figure.
Environment (e):
New York Zoo's lion exhibit area with the crowd of visitors.