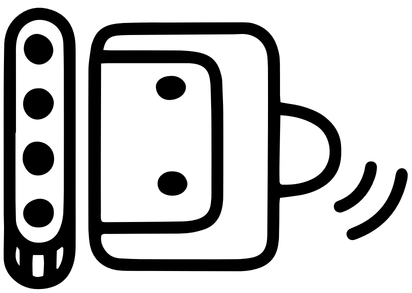
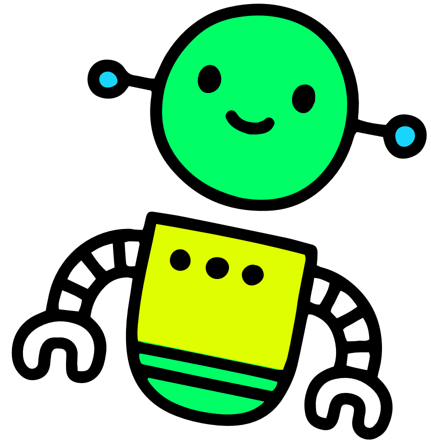
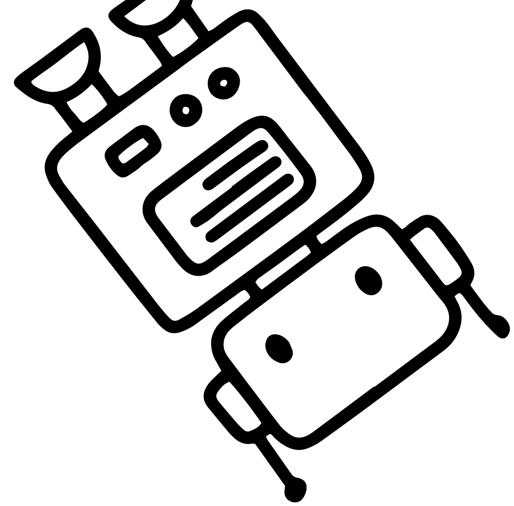
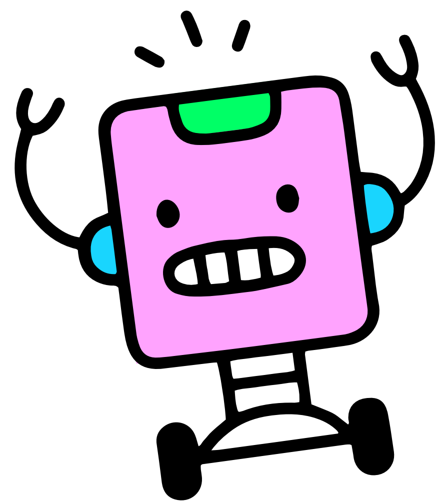
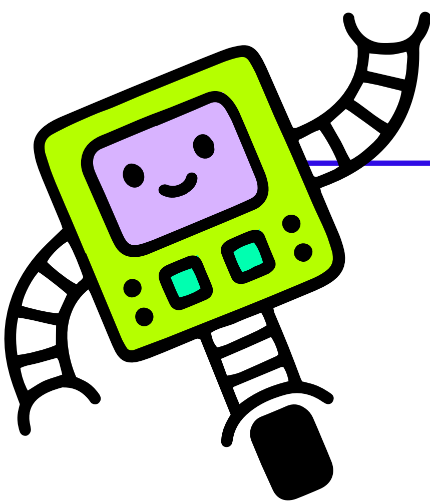






THE STRAWBERRY FARM NOVELTIES





Sude Fidan 21068639

FIRST COME SEALED BID AUCTION

A sealed-bid auction is a type of auction where all bidders submit their bids independently and without knowing the bids of the other participants. The highest bidder wins the auction and makes the payment. If there are multiple bidders with highest amount, the priority is given to the first bidder.

In the Strawberry Farm Model, each picker robot evaluates its current state to submit the bid. If the robot is free, its payload is not full and has tokens available, the bid is calculated as:

$$Bid = \frac{battery\ factor}{distance\ to\ target}$$

Where,

$$battery\ factor = \frac{battery}{maximum\ battery\ level}$$

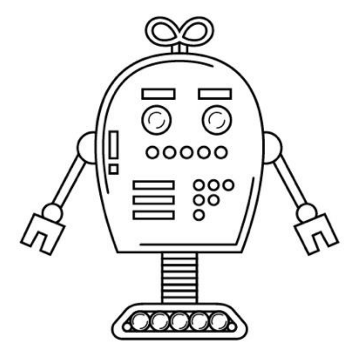
This formula ensures that robots with higher battery levels and closer proximity to the tree bid higher.

The highest bidder wins the right to pick strawberries from the tree. The winner determination calculated as:

$$Winner = \max_{i \in \{1,2,\ldots,n\}} v_i$$

Where, i goes through all the bidders, from the first bidder ($i = 1$) to the n-th bidder ($i = n$).v_i is the bid value of the i-th bidder.

When auction is done, winner makes the payment to the farm.

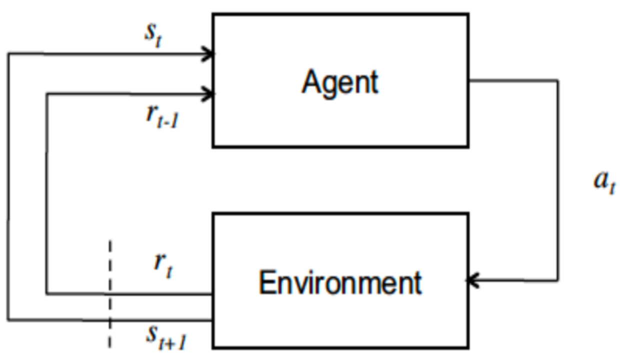


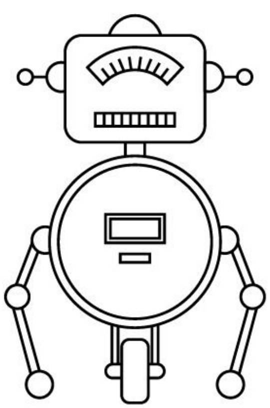
REINFORCEMENT LEARNING

Reinforcement learning (RL) is a type of machine learning where agents learn to make decisions by receiving rewards or penalties for their actions. In the Strawberry Farm Model, picker robots use Q-learning to optimise their random movement action.

State Representation: The state is represented by the robot's (x, y) position on the grid.

Action Space: The possible actions are moving up, down, left, or right. Therefore, the action space is 4.





Q-Table Update: The Q-table stores the expected rewards for each state-action pair. The Q-value is updated using the formula:

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a))$$

Where,

s is current state,
s' is next state,
a is the action taken,
a' is all possible actions,
r is the reward received after taken action(a) in state (s) ,
α is the learning rate,
γ is discount factor.

Exploration vs. Exploitation: The robot chooses between exploring new actions and exploiting known actions based on the exploration rate:

$$action = \begin{cases} random\ action & if\ random < exploration\_rate \\ \arg\ max_a Q(s,a) & otherwise \end{cases}$$

Decrementing Learning Rate: The learning rate is decremented over time to ensure that the robot's learning stabilises as it gains more experience. This helps the robot to converge to an optimal policy.

CHARGER ROBOTS

Charger robots function as mobile power banks, recharging robots with low battery levels. This allows robots to continue their tasks without needing to return to a charging station. In the basic and extended modes of the model, charger robots do not exist. Consequently, when a robot does not reach a charging station before its battery level drops below 0, it stops moving and cannot continue its tasks.

In the Strawberry Farm Model, a Charger Robot has 4 times higher battery capacity than other robots and therefore can recharge multiple robots before it needs to go to a battery station to recharge.

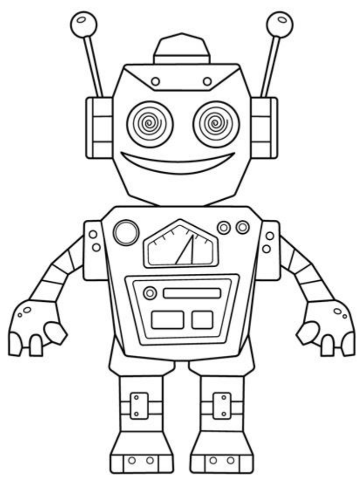
When a robot's battery level is low, it signals a charger robot.

The charger robot moves to the picker robot's location and recharges its battery. The battery charging amount is calculated as:

$$charging\ amount\ =\ Max\ battery - robot's\ battery$$

Where, Max battery is a constant determined as maximum battery level of the robots.

Charger robots transfer this calculated amount of battery power to the robots that needs charging, ensuring they can continue their tasks.



REFERENCES

Chanda, K.K. (2019) Q-Learning in Python - GeeksforGeeks GeeksforGeeks. 6 February 2019 [online]. Available from: https://www.geeksforgeeks.org/q-learning-in-python/ [Accessed 13 December 2024].

Hernandez, M.P. (2024) Agent Learning (I). Autonomous Agents and Multi-agent Systems [online]. Available From: https://xerte.uwe.ac.uk/USER-FILES/8262-me-perezhernandez-site/media/aamas2425-l8.pdf [Accessed 13 December 2024].

Hernandez, M.P. (2024) Resource Allocation (Auctions) & Collective Decision Making (Social Choice). Autonomous Agents and Multi-agent Systems [online]. Available From: https://xerte.uwe.ac.uk/USER-FILES/8262-me-perezhernandez-site/media/aamas2425-lec7.pdf [Accessed 13 December 2024].