# Foraging in Replenishing Patches


Abhinav Joshi (20211261)
Archi Gupta (21111014)
Parth Srivastava (190592)
Samrudh B Govindaraj (20128409)
Shiven Tripathi (190816)

# Problem Statement

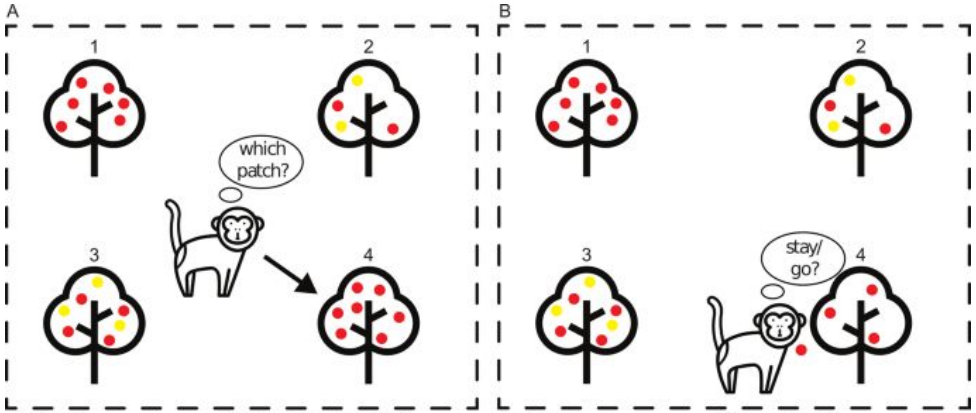
Understanding human foraging behaviour in a replenishing patches environment.

Importance

- Learn about sequential decision making processes in the face of uncertainty.
- Understand the role of working memory in sequential decision making.
- Possibly improve existing reinforcement learning algorithms.
- Different from conventional foraging tasks as patches can be revisited.

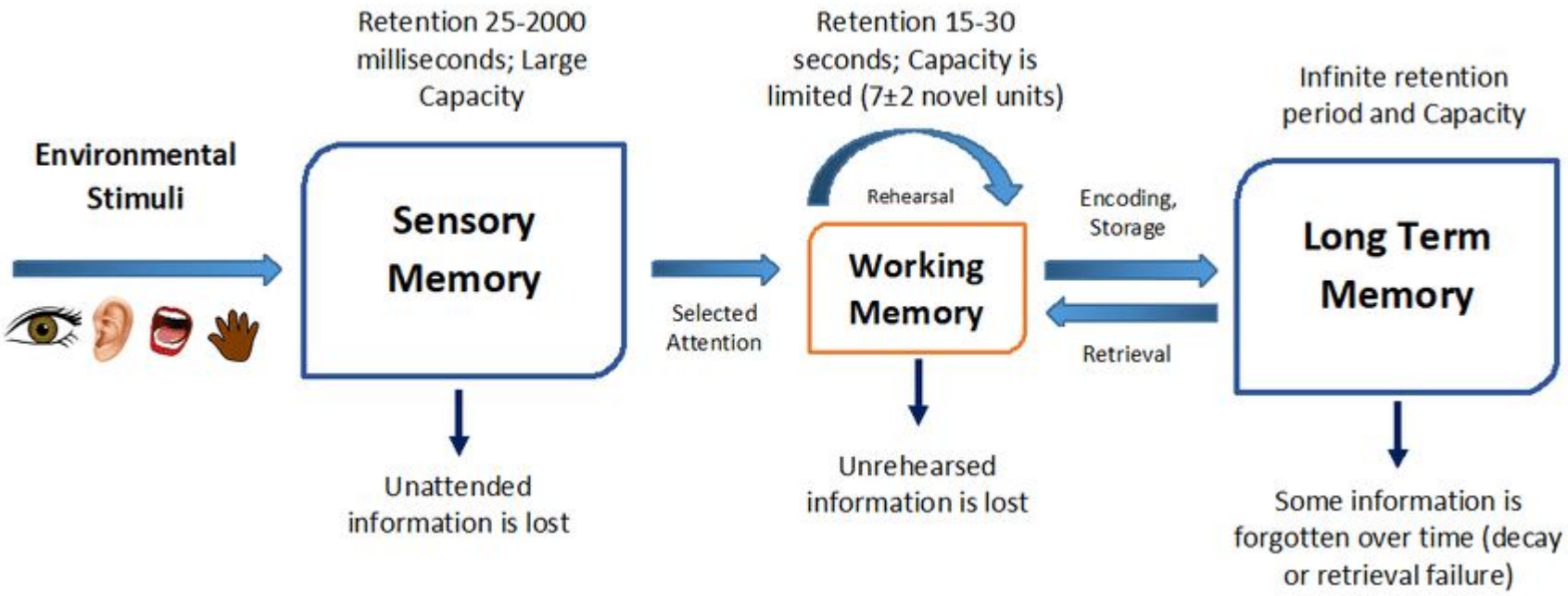
# Foraging Task

Foraging is searching for wild food resources. It affects an animal's fitness because it plays an important role in an animal's ability to survive and reproduce. [Wiki]



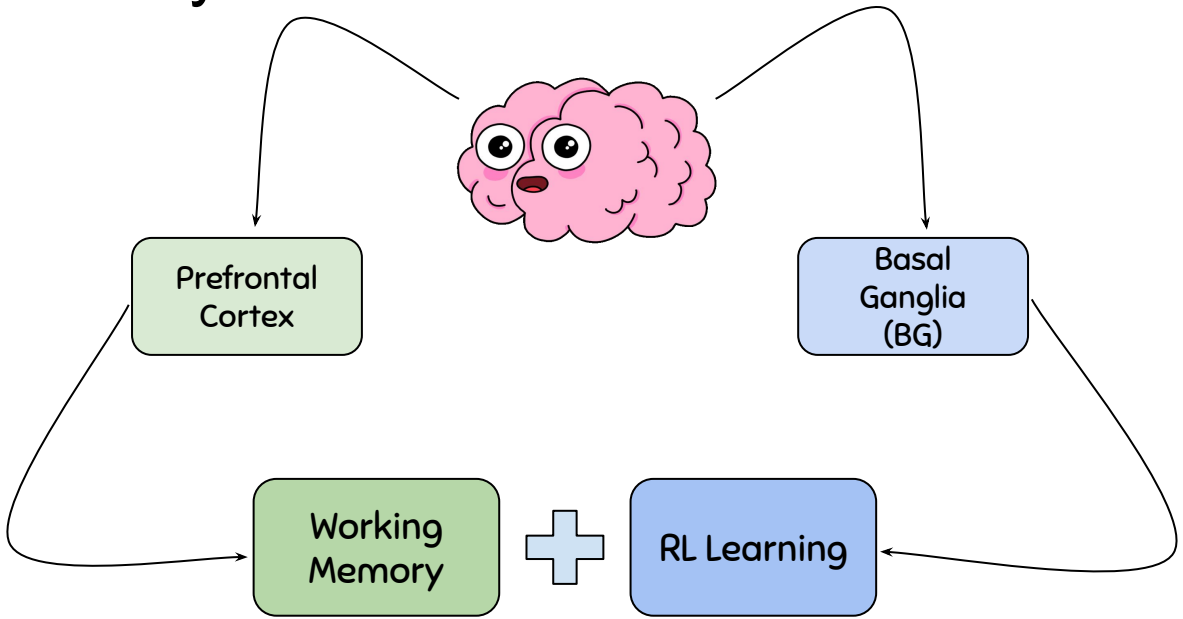

Foraging is an extensively studied tool to understand the neural computations that underlie sequential, value-based choices in animals.

# Atkinson-Shiffrin model of memory




https://www.researchgate.net/figure/Atkinson-Shiffrin-3-stage-model-of-human-memory_fig1_338116821

# Working Memory



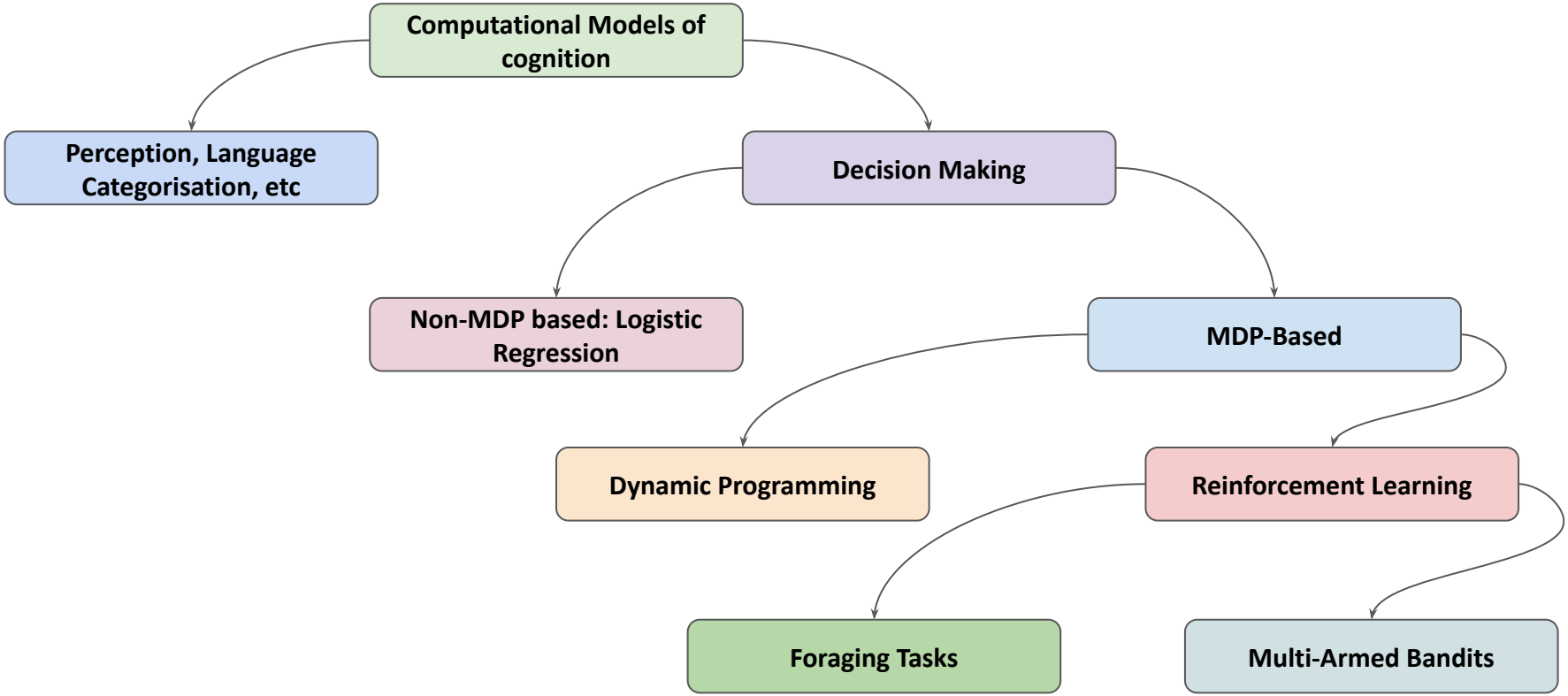

Pure reinforcement learning model

Forgetful reinforcement learning (RLF) model

Reinforcement learning + working memory model

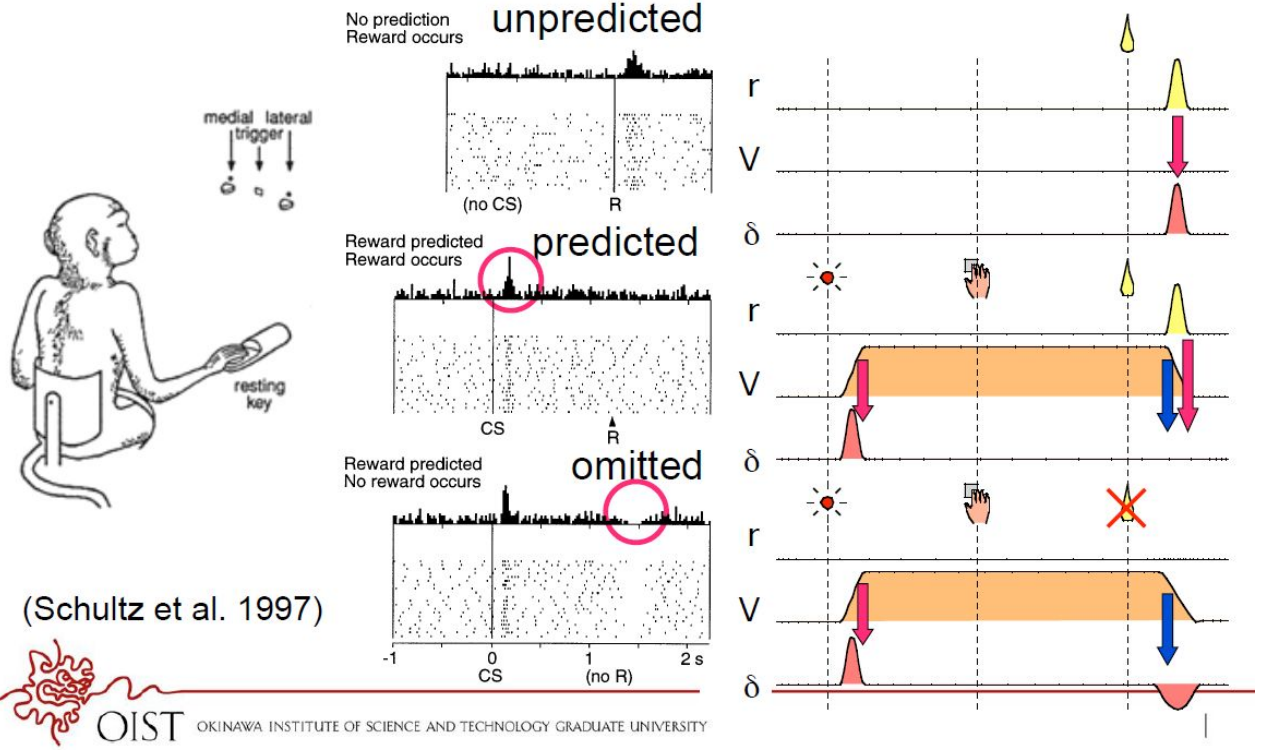
Pure working memory model (no reinforcement learning)

# Why Reinforcement Learning?



# Dopamine Neurons Code TD Error

$$\delta(t) = r(t) + \gamma V(s(t+1)) - V(s(t))$$




(Schultz et al. 1997)

OIST OKINAWA INSTITUTE OF SCIENCE AND TECHNOLOGY GRADUATE UNIVERSITY

# Summary of important concepts

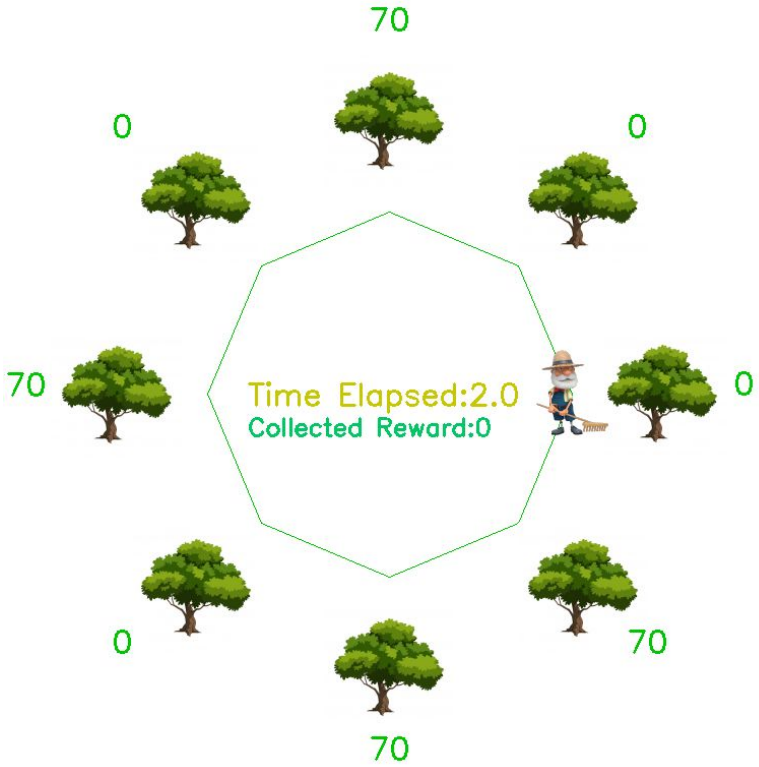
- Temporal Difference Learning algorithms.
- Marginal Value Theorem/Opportunity cost of harvesting.
- Recency Effect.
- Perseveration.
- Bayesian methods for incorporating priors and estimating uncertainty.



# Environment

Senario:

- 8 bushes forming a regular octagon
- Fixed 4 have fixed initial berries [ rewarding bushes ]
- Hungry man [ doesn't know which ones have berries, at the start ]




https://www.desmos.com/calculator/56fx5k2zol

# Environment

Aim:

- Hungry man has to maximise the berries he eats

Constraint:

- Time = 300 sec [horizon]

Actions:

- Eat/Harvest [takes 1 sec]
- Move from one bush to other [takes time ~ distance]

# Environment

Reward:

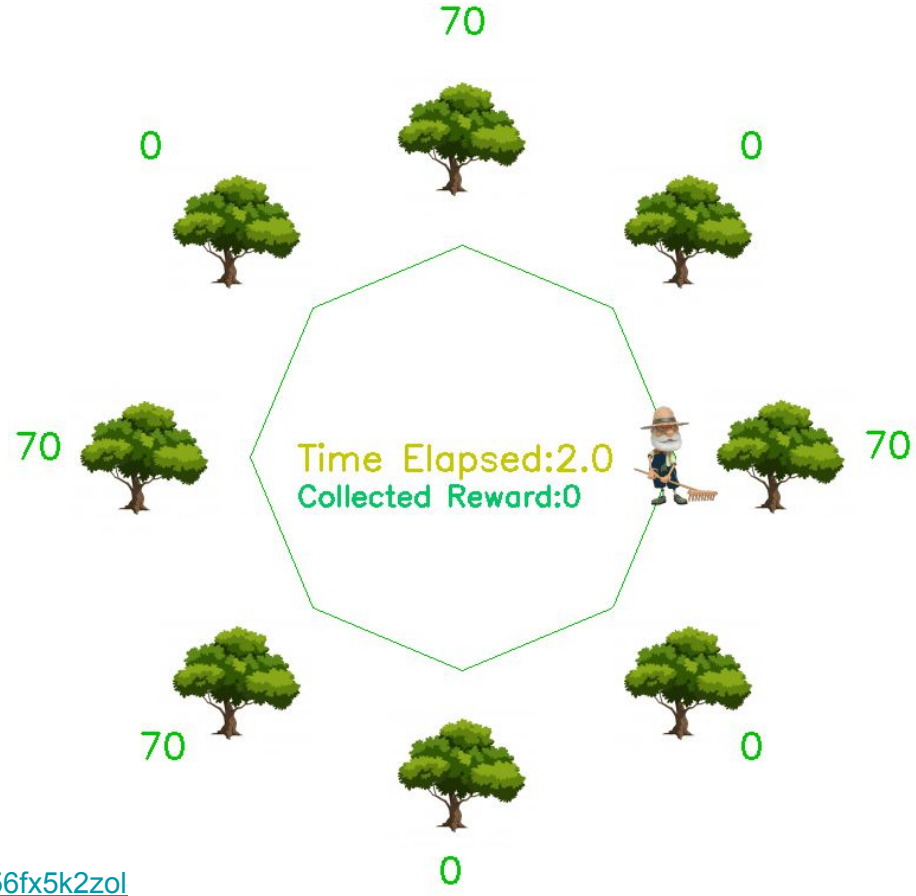
- Hungry man gets reward when he eats from a rewarding bush
- 90% of the berries of that bush

Dynamics:

- Every time man eats from a rewarding bush, the berries on that bush reduce by 0.9 times
- Every time man eats from a rewarding bush, the berries on other rewarding bush replenish by a fixed amount say x.
- Max berries at any bush = 200



# Demo




https://www.desmos.com/calculator/56fx5k2zol

# Testing Environment (Text Render of Explore Strategy)

Game start time: 2 Game start patch: 0
Time Elapsed: 4.847759065022574 Patch Visited: 2 Harvest Received: 63.0
Time Elapsed: 6.847759065022574 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 10.460884994775327 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 13.308644059797901 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 15.308644059797901 Patch Visited: 6 Harvest Received: 66.60000000000002
Time Elapsed: 18.921769989550654 Patch Visited: 2 Harvest Received: 60.300000000000001
Time Elapsed: 20.921769989550654 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 22.921769989550654 Patch Visited: 2 Harvest Received: 54.0
Time Elapsed: 25.769529054573226 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 27.769529054573226 Patch Visited: 1 Harvest Received: 77.40000000000003
Time Elapsed: 29.769529054573226 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 31.769529054573226 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 33.769529054573226 Patch Visited: 6 Harvest Received: 70.19999999999999
Time Elapsed: 34.769529054573226 Patch Visited: 6 Harvest Received: 63.0
Time Elapsed: 36.769529054573226 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 39.6172881195958 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 43.03150168196889 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 45.03150168196889 Patch Visited: 1 Harvest Received: 76.49999999999994
Time Elapsed: 48.644627611721646 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 51.49238667674422 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 54.34014574176679 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 57.75435930413988 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 60.602118369162454 Patch Visited: 6 Harvest Received: 60.299999999999955
Time Elapsed: 64.01633193153555 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 67.43054549390864 Patch Visited: 6 Harvest Received: 54.0
Time Elapsed: 70.84475905628173 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 74.45788498603449 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 76.45788498603449 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 79.30564405105706 Patch Visited: 2 Harvest Received: 70.20000000000005
Time Elapsed: 82.15340311607963 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 85.0011621811022 Patch Visited: 6 Harvest Received: 52.200000000000045
Time Elapsed: 87.84892124612477 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 90.69668031114735 Patch Visited: 6 Harvest Received: 46.799999999999955
Time Elapsed: 93.54443937616992 Patch Visited: 4 Harvest Received: 109.79999999999995
Time Elapsed: 96.95865293854301 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 99.80641200356558 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 103.22062556593868 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 106.06838463096125 Patch Visited: 6 Harvest Received: 45.0
Time Elapsed: 108.06838463096125 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 111.681510560714 Patch Visited: 1 Harvest Received: 93.59999999999991
Time Elapsed: 115.0957241230871 Patch Visited: 4 Harvest Received: 105.29999999999995
Time Elapsed: 117.94348318810967 Patch Visited: 6 Harvest Received: 47.700000000000045
Time Elapsed: 119.94348318810967 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 122.79124225313224 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 124.79124225313224 Patch Visited: 6 Harvest Received: 42.299999999999955
Time Elapsed: 127.63900131815481 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 129.6390013181548 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 131.6390013181548 Patch Visited: 6 Harvest Received: 37.799999999999955
Time Elapsed: 133.6390013181548 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 137.25212724790757 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 140.66634081028067 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 144.08055437265378 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 146.08055437265378 Patch Visited: 4 Harvest Received: 105.29999999999995
Time Elapsed: 155.75674056242258 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 156.75674056242258 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 160.17095412479568 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 161.17095412479568 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 163.17095412479568 Patch Visited: 2 Harvest Received: 102.59999999999991
Time Elapsed: 166.01871318981827 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 169.43292675219138 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 173.04605268194413 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 175.04605268194413 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 177.04605268194413 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 180.46026624431724 Patch Visited: 4 Harvest Received: 101.70000000000005
Time Elapsed: 184.07339217407 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 187.4876057364431 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 188.4876057364431 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 191.9018192988162 Patch Visited: 6 Harvest Received: 47.700000000000045
Time Elapsed: 195.3160328611893 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 198.73024642356242 Patch Visited: 6 Harvest Received: 42.299999999999955
Time Elapsed: 199.73024642356242 Patch Visited: 6 Harvest Received: 37.799999999999955
Time Elapsed: 200.73024642356242 Patch Visited: 6 Harvest Received: 33.299999999999955
Time Elapsed: 204.14445998593553 Patch Visited: 1 Harvest Received: 112.5
Time Elapsed: 206.14445998593553 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 209.55867354830863 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 212.40643261333122 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 216.01955854308397 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 218.86731760810656 Patch Visited: 1 Harvest Received: 100.80000000000018
Time Elapsed: 222.48044353785932 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 226.09356946761207 Patch Visited: 1 Harvest Received: 90.0
Time Elapsed: 228.94132853263466 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 230.94132853263466 Patch Visited: 4 Harvest Received: 116.09999999999991
Time Elapsed: 234.35554209500776 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 236.35554209500776 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 239.76975565738087 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 243.18396921975398 Patch Visited: 2 Harvest Received: 124.19999999999982
Time Elapsed: 244.18396921975398 Patch Visited: 2 Harvest Received: 111.59999999999991
Time Elapsed: 246.18396921975398 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 248.18396921975398 Patch Visited: 4 Harvest Received: 111.59999999999991
Time Elapsed: 251.03172828477656 Patch Visited: 6 Harvest Received: 54.90000000000009
Time Elapsed: 254.44594184714967 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 257.86015540952275 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 259.86015540952275 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 261.86015540952275 Patch Visited: 6 Harvest Received: 48.59999999999991
Time Elapsed: 264.7079144745453 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 268.32104040429806 Patch Visited: 4 Harvest Received: 107.09999999999991
Time Elapsed: 271.73525396667117 Patch Visited: 7 Harvest Received: 0.0
Time Elapsed: 275.1494675290443 Patch Visited: 4 Harvest Received: 96.30000000000018
Time Elapsed: 277.1494675290443 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 279.1494675290443 Patch Visited: 4 Harvest Received: 86.40000000000009
Time Elapsed: 281.1494675290443 Patch Visited: 3 Harvest Received: 0.0
Time Elapsed: 284.5636810914174 Patch Visited: 0 Harvest Received: 0.0
Time Elapsed: 287.41144015643994 Patch Visited: 2 Harvest Received: 121.5
Time Elapsed: 290.82565371881304 Patch Visited: 5 Harvest Received: 0.0
Time Elapsed: 294.4387796485658 Patch Visited: 1 Harvest Received: 117.0
Time Elapsed: 297.8529932109389 Patch Visited: 6 Harvest Received: 61.19999999999982

# Testing Environment (Patch Rewards Over Time)

Patch Rewards before harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards before harvest from Patch: 0 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 0 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 70 70 0 70 0 70 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 74 74 0 63 0 74 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 74 74 0 63 0 74 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 78 78 0 56 0 78 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 78 78 0 56 0 78 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 70 82 0 60 0 82 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 70 82 0 60 0 82 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 70 82 0 60 0 82 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 70 82 0 60 0 82 0]
Patch Rewards after harvest from Patch: 2 are: [ 0 74 73 0 64 0 86 0]
Patch Rewards before harvest from Patch: 6 are: [ 0 74 73 0 64 0 86 0]
Patch Rewards after harvest from Patch: 6 are: [ 0 78 77 0 68 0 77 0]
Patch Rewards before harvest from Patch: 3 are: [ 0 78 77 0 68 0 77 0]
Patch Rewards after harvest from Patch: 3 are: [ 0 78 77 0 68 0 77 0]

Patch Rewards after harvest from Patch: 4 are: [ 0 82 81 0 61 0 81 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 82 81 0 61 0 81 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 82 81 0 61 0 81 0]
Patch Rewards before harvest from Patch: 1 are: [ 0 82 81 0 61 0 81 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 73 85 0 65 0 85 0]
Patch Rewards before harvest from Patch: 1 are: [ 0 73 85 0 65 0 85 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 65 89 0 69 0 89 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 65 89 0 69 0 89 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 65 89 0 69 0 89 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 65 89 0 69 0 89 0]
Patch Rewards after harvest from Patch: 2 are: [ 0 69 80 0 73 0 93 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 69 80 0 73 0 93 0]
Patch Rewards after harvest from Patch: 2 are: [ 0 73 72 0 77 0 97 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 73 72 0 77 0 97 0]
Patch Rewards after harvest from Patch: 2 are: [ 0 77 64 0 81 0 101 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 77 64 0 81 0 101 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 81 68 0 72 0 105 0]
Patch Rewards before harvest from Patch: 0 are: [ 0 81 68 0 72 0 105 0]
Patch Rewards after harvest from Patch: 0 are: [ 0 81 68 0 72 0 105 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 81 68 0 72 0 105 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 85 72 0 64 0 109 0]
Patch Rewards before harvest from Patch: 1 are: [ 0 85 72 0 64 0 109 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 76 76 0 68 0 113 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 76 76 0 68 0 113 0]

Patch Rewards after harvest from Patch: 4 are: [ 0 80 80 0 61 0 117 0]
Patch Rewards before harvest from Patch: 3 are: [ 0 80 80 0 61 0 117 0]
Patch Rewards after harvest from Patch: 3 are: [ 0 80 80 0 61 0 117 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 80 80 0 61 0 117 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 84 84 0 54 0 121 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 84 84 0 54 0 121 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 84 84 0 54 0 121 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 84 84 0 54 0 121 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 88 88 0 48 0 125 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 88 88 0 48 0 125 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 92 92 0 43 0 129 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 92 92 0 43 0 129 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 92 92 0 43 0 129 0]
Patch Rewards before harvest from Patch: 1 are: [ 0 92 92 0 43 0 129 0]
Patch Rewards after harvest from Patch: 1 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards before harvest from Patch: 3 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards after harvest from Patch: 3 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 82 96 0 47 0 133 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 86 100 0 42 0 137 0]
Patch Rewards before harvest from Patch: 0 are: [ 0 86 100 0 42 0 137 0]
Patch Rewards after harvest from Patch: 0 are: [ 0 86 100 0 42 0 137 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 86 100 0 42 0 137 0]

Patch Rewards after harvest from Patch: 2 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards before harvest from Patch: 5 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards after harvest from Patch: 5 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards before harvest from Patch: 6 are: [ 0 90 90 0 46 0 141 0]
Patch Rewards after harvest from Patch: 6 are: [ 0 94 94 0 50 0 126 0]
Patch Rewards before harvest from Patch: 2 are: [ 0 94 94 0 50 0 126 0]
Patch Rewards after harvest from Patch: 2 are: [ 0 98 84 0 54 0 130 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 98 84 0 54 0 130 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 98 84 0 54 0 130 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 98 84 0 54 0 130 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 102 88 0 48 0 134 0]
Patch Rewards before harvest from Patch: 4 are: [ 0 102 88 0 48 0 134 0]
Patch Rewards after harvest from Patch: 4 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards before harvest from Patch: 0 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards after harvest from Patch: 0 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards before harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]
Patch Rewards after harvest from Patch: 7 are: [ 0 106 92 0 43 0 138 0]

# Approximate Solutions for Baseline Performance

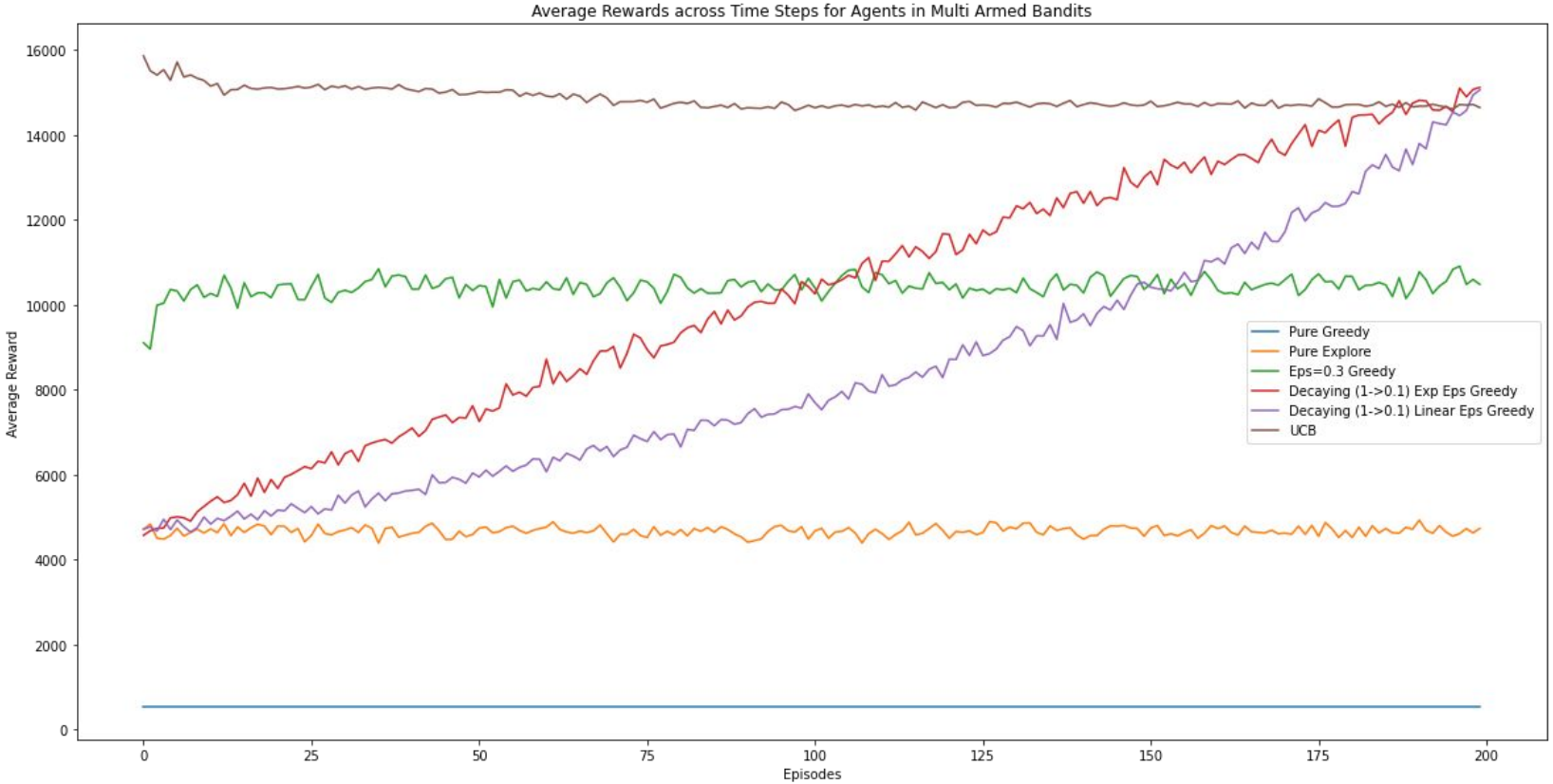
Assumptions:

- Problem can be approximated to a multi armed bandits situation
- Choice with agent to pick a patch to go to, and then it commits to harvesting
- An episode lasts until time runs out

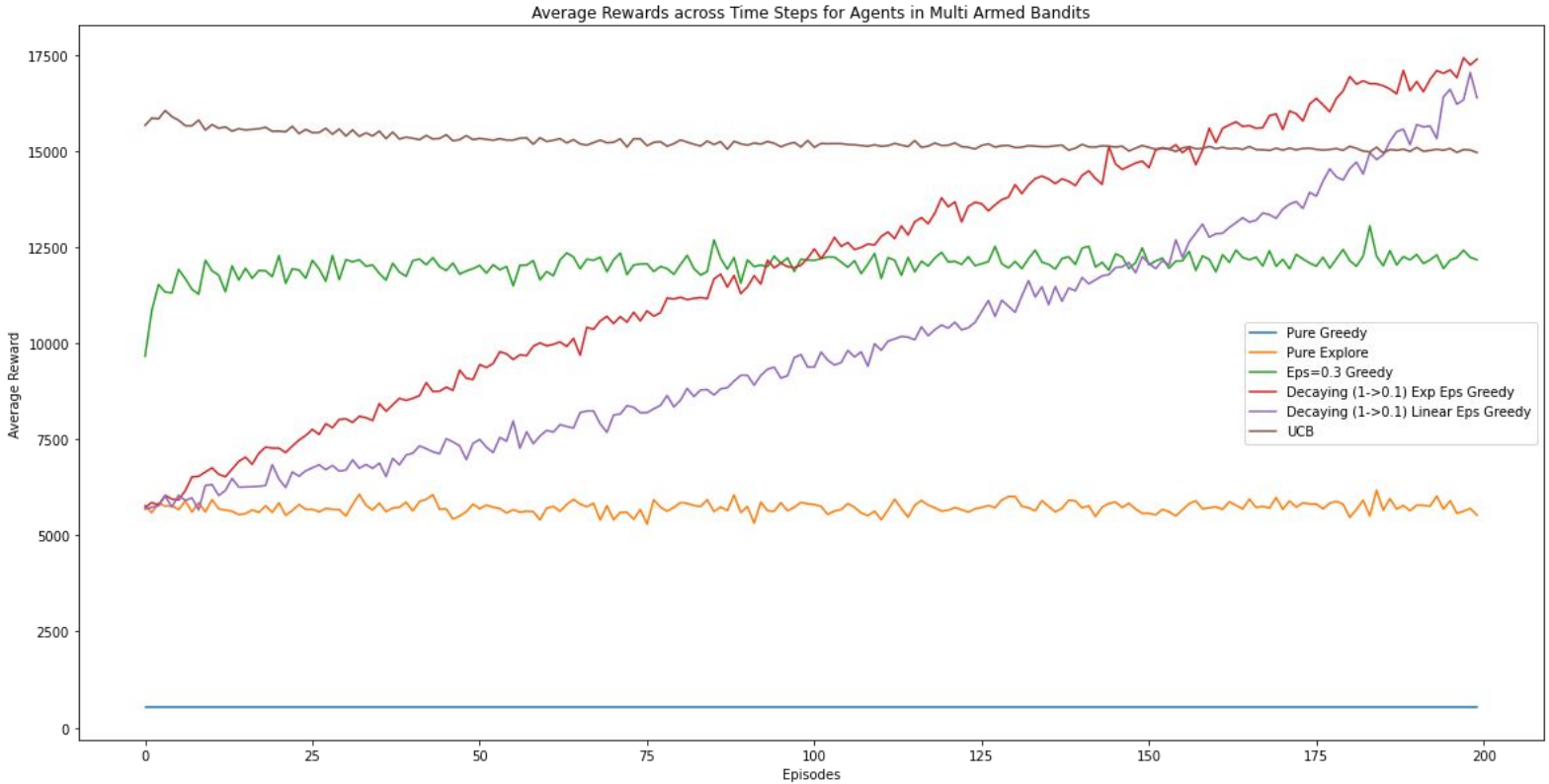
Under these simplifying constraints, following strategies were tested:

- Pure Greedy
- Pure Exploratory
- Fixed Epsilon Greedy
- Decaying Epsilon Greedy
- Uncertainty Confidence Bound (UCB)

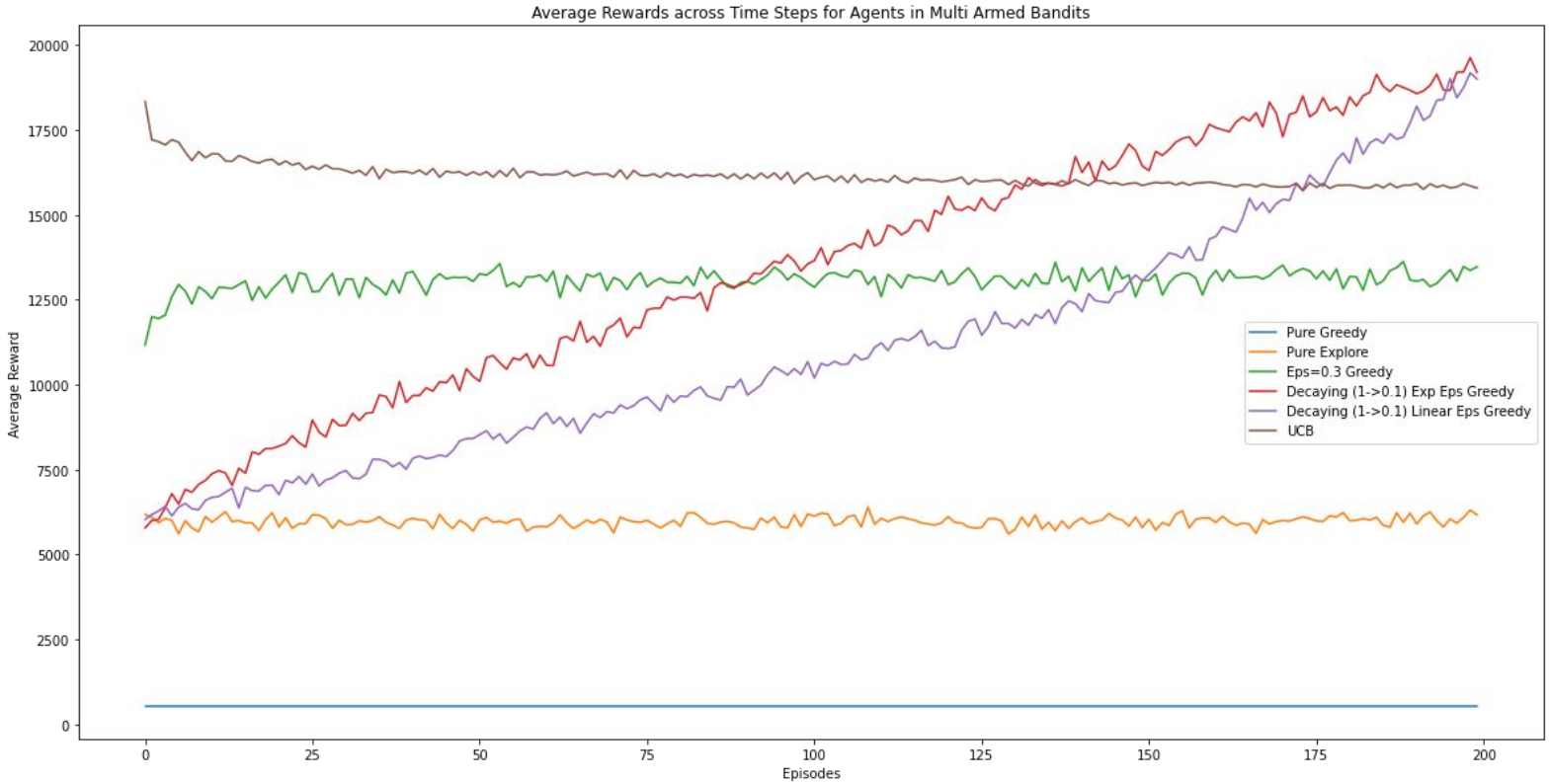
# Average Rewards for Agents in MAB across Episodes Block 1 (replenish rate = [0, 4, 4, 0, 4, 0, 4, 0])



# Block 2 (replenish rate = [0, 0, 8, 2, 0, 5, 0, 8])



# Block 3 (replenish rate = [2, 0, 0, 4, 8, 0, 16, 0])



# Initial Steps/Approaches/Models

Pure RL Strategies with minor modifications and performance of different blocks:

- Smart Exploration ( only on rewarding patches )
- Negative reward for travelling (cost ~ time or distance)
- Time component in state value function

Testing Marginal Value Theorem for optimality [1]

Determine the best algorithmic approximation for MVT for our task [2]

Generating the optimal policy when the environment is known. Thereafter using model based learning. [3]



# Approach For Human Behaviour Modelling

Aim - Dynamic integration of RL and WM processes observed in human behaviour to capture *behavioural variance*

Forgetfulness - After each value update step, we decay the values towards their initial values

$$Q(s,a) \leftarrow Q(s,a) + \varepsilon \times (Q_0 - Q(s,a))$$

Working Memory Model - Using two value functions $Q_{RL}$ (pure RL) and $Q_{WM}$ (with forgetting ) and assigning a weighted probability for action selection.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3390186/



# Challenges

- Identifying which specific events are stored in memory and which are not.
- Modeling probability of each one of events being stored.
  - Probability might dynamically change over time
- Accounting for the observed effects of memory load and time delay on performance.

# Project Timeline

Phase 1 [ done ] : Environment set up and literature review

Phase 2 [2 weeks]: RL algorithm for optimal policies guided by the Literature Survey

Phase 3 [2 weeks]: Modifying above implementations to incorporate human behaviour (WM)

Additional : Comparing human learning *trajectory* with our RLWM model



# Individual Contributions

Gym Environment - Abhinav

Baseline Solutions and Environment Testing - Shiven

Literature Review - Parth and Samrudh

Next Approaches - Archi



# References

Goldstone, R.L., Ashpole, B.C. Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review* 11, 508–514 (2004). https://doi.org/10.3758/BF03196603

Hall-McMaster, S., Luyckx, F. Revisiting foraging approaches in neuroscience. *Cogn Affect Behav Neurosci* 19, 225–230 (2019). https://doi.org/10.3758/s13415-018-00682-z

Constantino SM, Daw ND. Learning the opportunity cost of time in a patch-foraging task. Cogn Affect Behav Neurosci. 2015 Dec;15(4):837-53. doi: 10.3758/s13415-015-0350-y. PMID: 25917000; PMCID: PMC4624618.

Matt L. Miller, Kevin M. Ringelman, John M. Eadie, Jeffrey C. Schank, Time to fly: A comparison of marginal value theorem approximations in an agent-based model of foraging waterfowl, Ecological Modelling, Volume 351, 2017, Pages 77-86,

# Questions