Oregon State University

Apple Orchards: A multitask **multiagent** approach

Alejo Velasquez, Nigel Swenson and Keegan Nave

ROB538 Fall'22 - Project Presentation

Credits: https://www.therobotreport.com/abundant-robotics-shuts-down-fruit-harvesting-business/

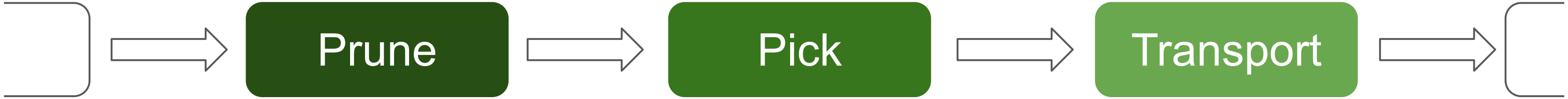
Motivation

Apple orchards are promising for robots.

Orchards have multiple tasks that need to be performed.

How to plan for multiple agents that handle different **sequential** tasks?

Credits: https://www.abc.net.au/news/2021-05-03/robots-almost-ready-to-help-pick-fruit/100107796



Prune

Pick

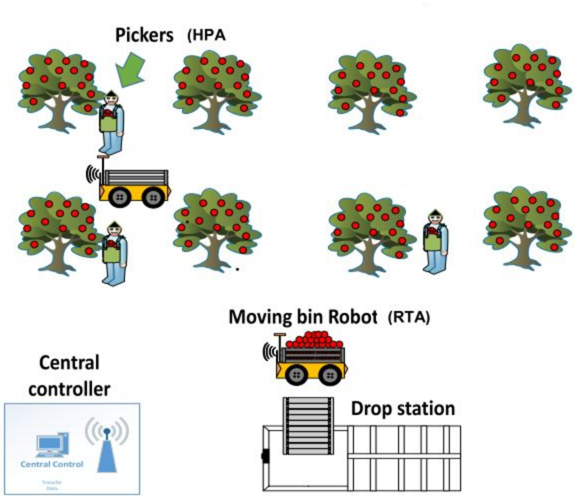
Transport

Source: https://usapple.org/news-resources/apple-stages-from-the-tree-to-the-grocery-store

Background

Hybrid (human-robot) apple picking

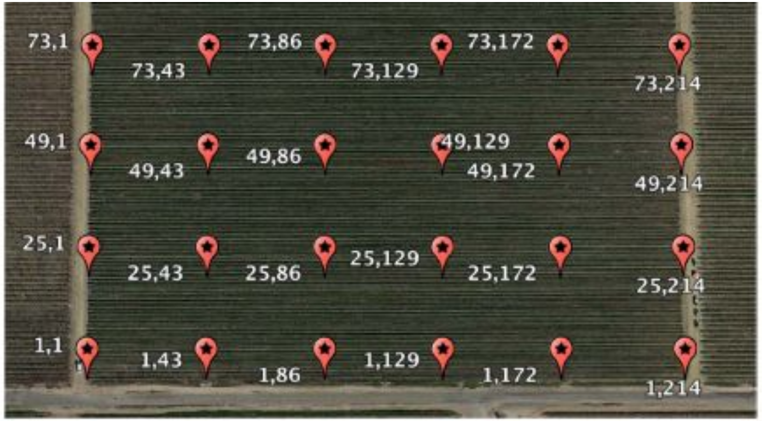
Different skills, centralized



(Salah, 2019)

Vineyard irrigation

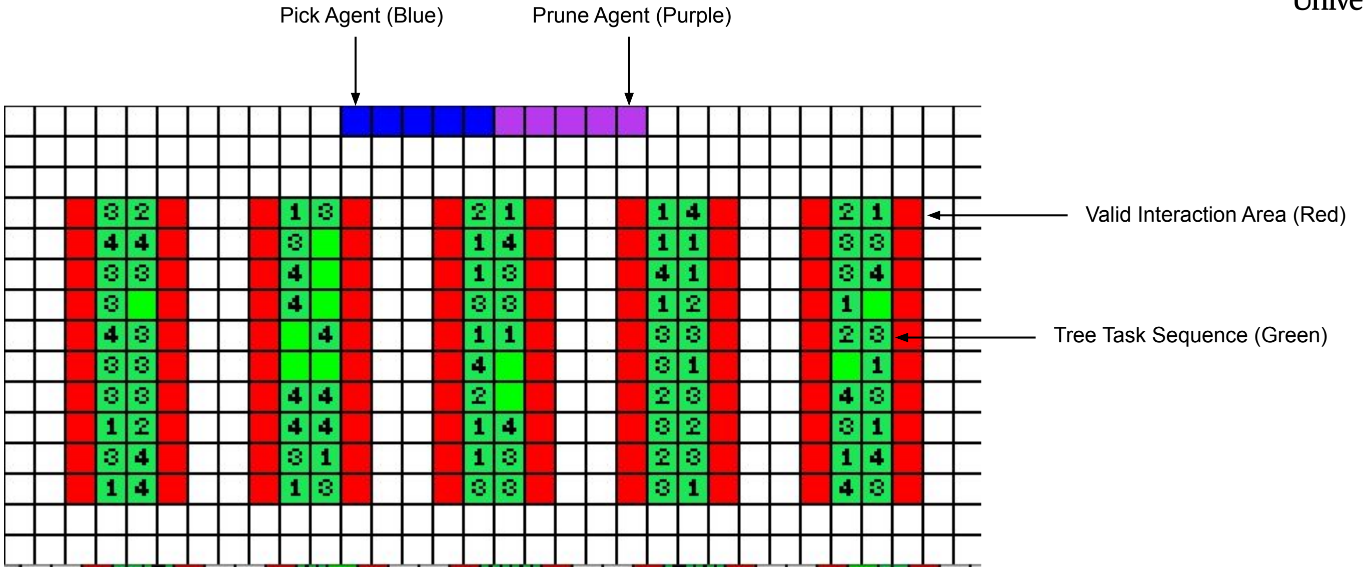
Same skill, many agents



(Thayer et al., 2020)

Problem Definition

Problem Space



Pick Agent (Blue)

Prune Agent (Purple)

Valid Interaction Area (Red)

Tree Task Sequence (Green)

3	2	1	3	2	1	1	4	2	1
4	4	3		1	4	1	1	3	3
3	3	4		1	3	4	1	3	4
3		4		3	3	1	2	1	
4	3		4	1	1	3	3	2	3
3	3			4		3	1		1
3	3	4	4	2		2	3	4	3
1	2	4	4	1	4	3	2	3	1
3	4	3	1	1	3	2	3	1	4
1	4	1	3	3	3	3	1	4	3

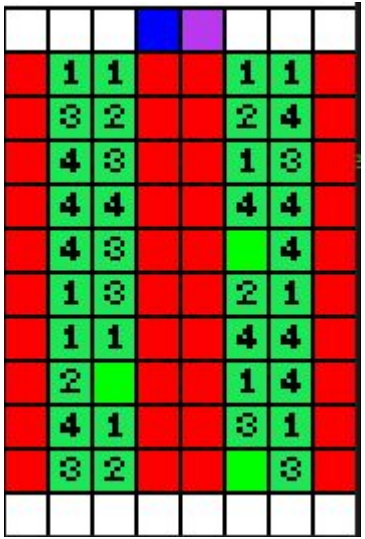
$$TaskSequences = \begin{cases} 1 : \text{Pick} \rightarrow \text{Complete} \\ 2 : \text{Prune} \rightarrow \text{Complete} \\ 3 : \text{Pick} \rightarrow \text{Prune} \rightarrow \text{Complete} \\ 4 : \text{Prune} \rightarrow \text{Pick} \rightarrow \text{Complete} \\ None : \text{Complete} \end{cases}$$

$$S_i = \{Agent_i(x,y), ValidTree_1, ValidTree_2, ..., ValidTree_n\}$$

$$A = \{up, down, left, right, interact\}$$

Goals

- Investigate tightly coupled agents in Agricultural settings.
- Improve performance on sequential task spaces.
- Maximize the amount of pick and prune tasks that are completed.



			🟦	🟪			
🟥	1	1	🟥	🟥	1	1	🟥
🟥	3	2	🟥	🟥	2	4	🟥
🟥	4	3	🟥	🟥	1	3	🟥
🟥	4	4	🟥	🟥	4	4	🟥
🟥	4	3	🟥	🟥	🟩	4	🟥
🟥	1	3	🟥	🟥	2	1	🟥
🟥	1	1	🟥	🟥	4	4	🟥
🟥	2	🟩	🟥	🟥	1	4	🟥
🟥	4	1	🟥	🟥	3	1	🟥
🟥	3	2	🟥	🟥	🟩	3	🟥

Reward Structure

Local Reward

$$L = \left\{ 10 \text{ when interacting, } -1 \text{ otherwise} \right\}$$

Global Reward

$$G = \frac{TreesPruned}{TotalLeaves} \times \frac{ApplesPicked}{ApplestoPick}$$

Potential Local Reward

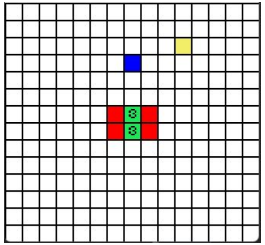
Difference Reward 1: Random Action

Difference Reward 2: Random Action with 1 step look ahead

Approach 1

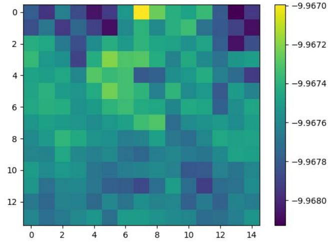
Q learning - what **didn’t** work

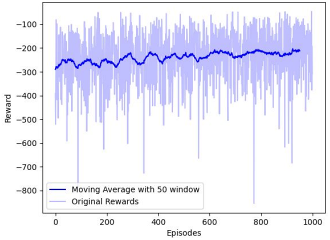
Let’s start simple



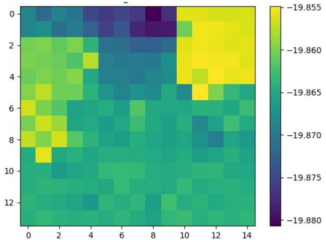
2 rows, 1 col, 2 agents
1k steps, gamma 0.9
alpha 0.05, epsilon*0.99

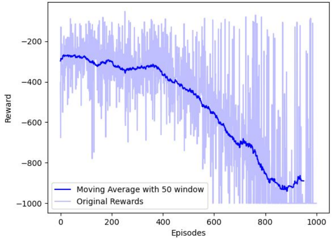
Local Rewards



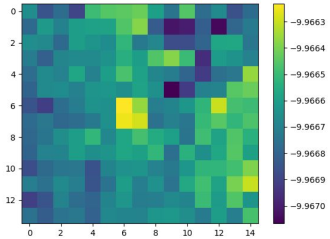


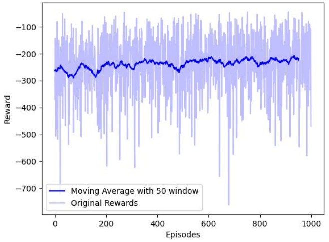
Global Rewards





Diff Rewards 1





Q learning - what **did** work

Sequential tasks can be thought as **tightly-coupled**.

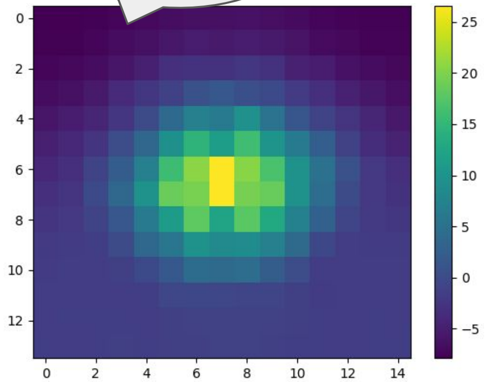
What if a tree is **not** useful for **me**, but it is for **others**?

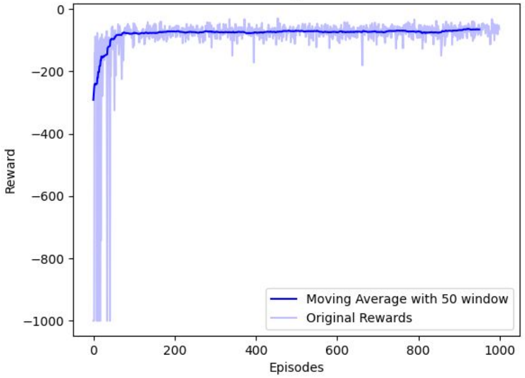
```
i: current agent
j: other agents

...
if i.reward < 0:
    for j in agents:
        # Let each agent assume the current agent's position
        j.valid_moves = get_valid_moves(i.cur_pose, j.type)

        if "interact" in j.valid_moves:
            total_rewards += 10
...
```

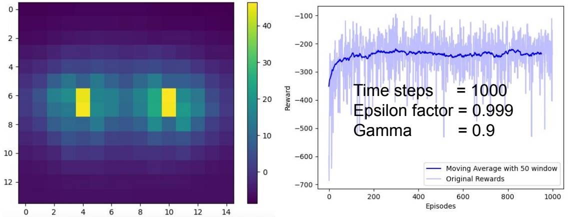
Value tables are shown for just one agent





Q tables - up to a certain point

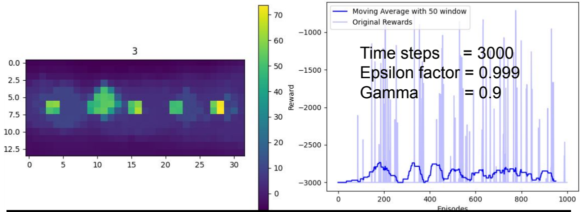
6 agents, **2 columns** and 2 rows



Time steps = 1000
Epsilon factor = 0.999
Gamma = 0.9

Sequence (3: Prune → 2: Pick → 0: Done)

6 agents, **5 columns** and 2 rows



Time steps = 3000
Epsilon factor = 0.999
Gamma = 0.9

Sequence (3: Prune → 2: Pick → 0: Done)

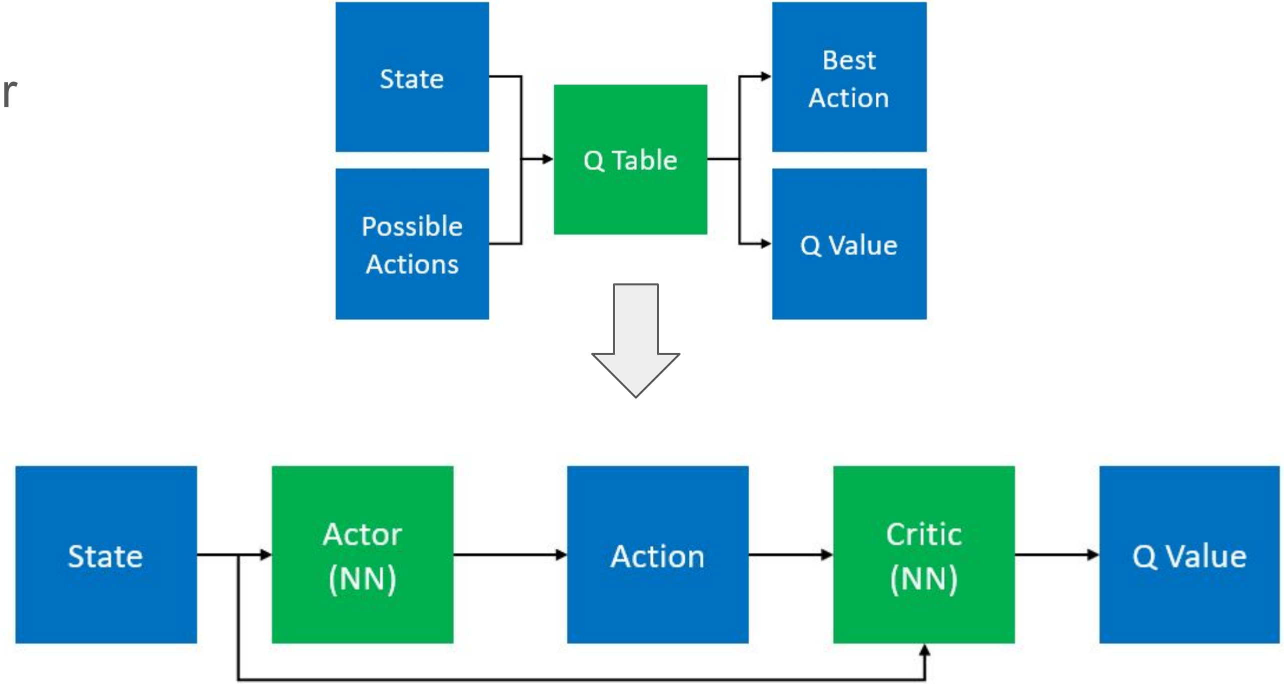
Sensitive to number of agents and size of the orchard

Approach 2

Soft Actor Critic

Replace Q table with Actor and Critic Networks

Can handle larger state spaces - don't have to memorize every point



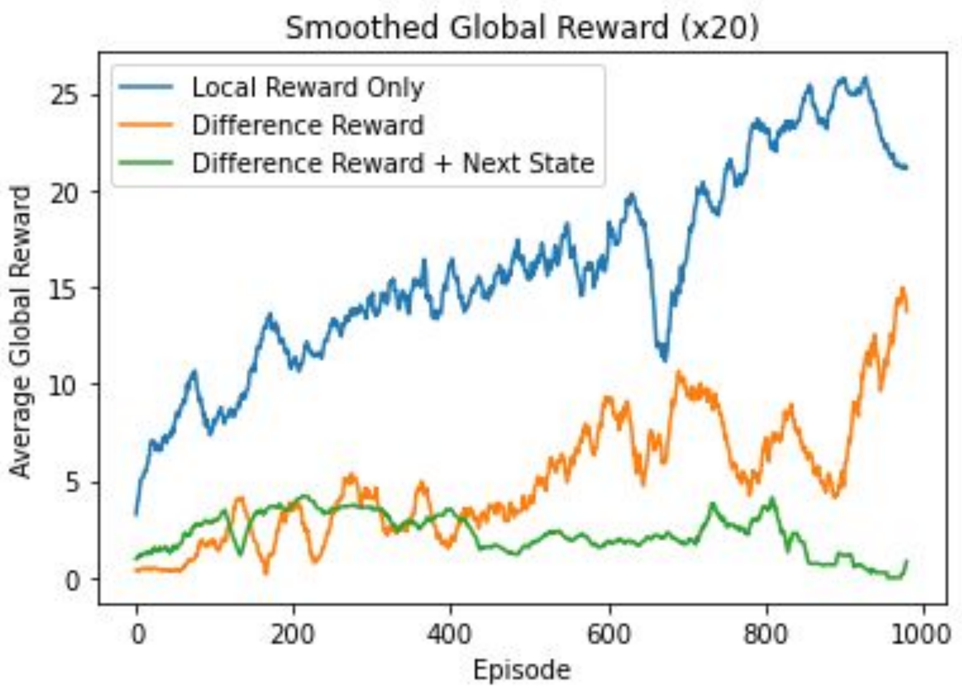
Results SAC Same Seed

1 pick agent

1 prune agent

12x9 orchard

100 steps/episode



Results SAC Different Seed

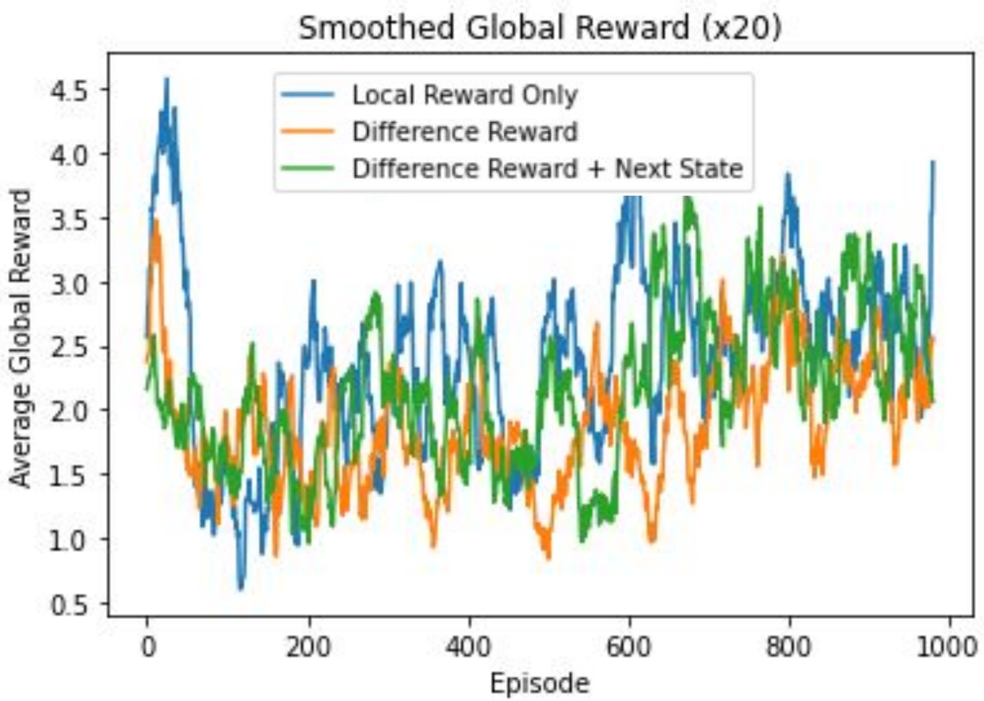
Oregon State University

1 pick agent

1 prune agent

12x9 orchard

100 steps/episode



Conclusions

Conclusions

- Qtables were useful for smaller grids. **With greater grids, specially with more columns, the system didn’t converge.**

- Local rewards outperform difference reward with current global baseline
  - Worth changing our global system to add the fractions to encourage individual actions

Future Work

Short term (paperwise)

- Explore other **reward shaping** techniques in Q-table approach.
- Explore **higher number** of sequential tasks.
- Compare with other **benchmarks** (e.g. team of pruners go first, then team of pickers)

Long term (research)

- Explore impact of **orchard structure.**
- Explore impact of **time** between tasks.

Thank you

Questions?