What Should We Grow Today so We Make Money Tomorrow?

Reinforcement Learning for Small Farmers

Aviva Prins

Context

Greenhouse-in-a-box: an affordable, modular greenhouse that uses 90% less water, grows 7 times more food and gives farmers a steady dependable income.

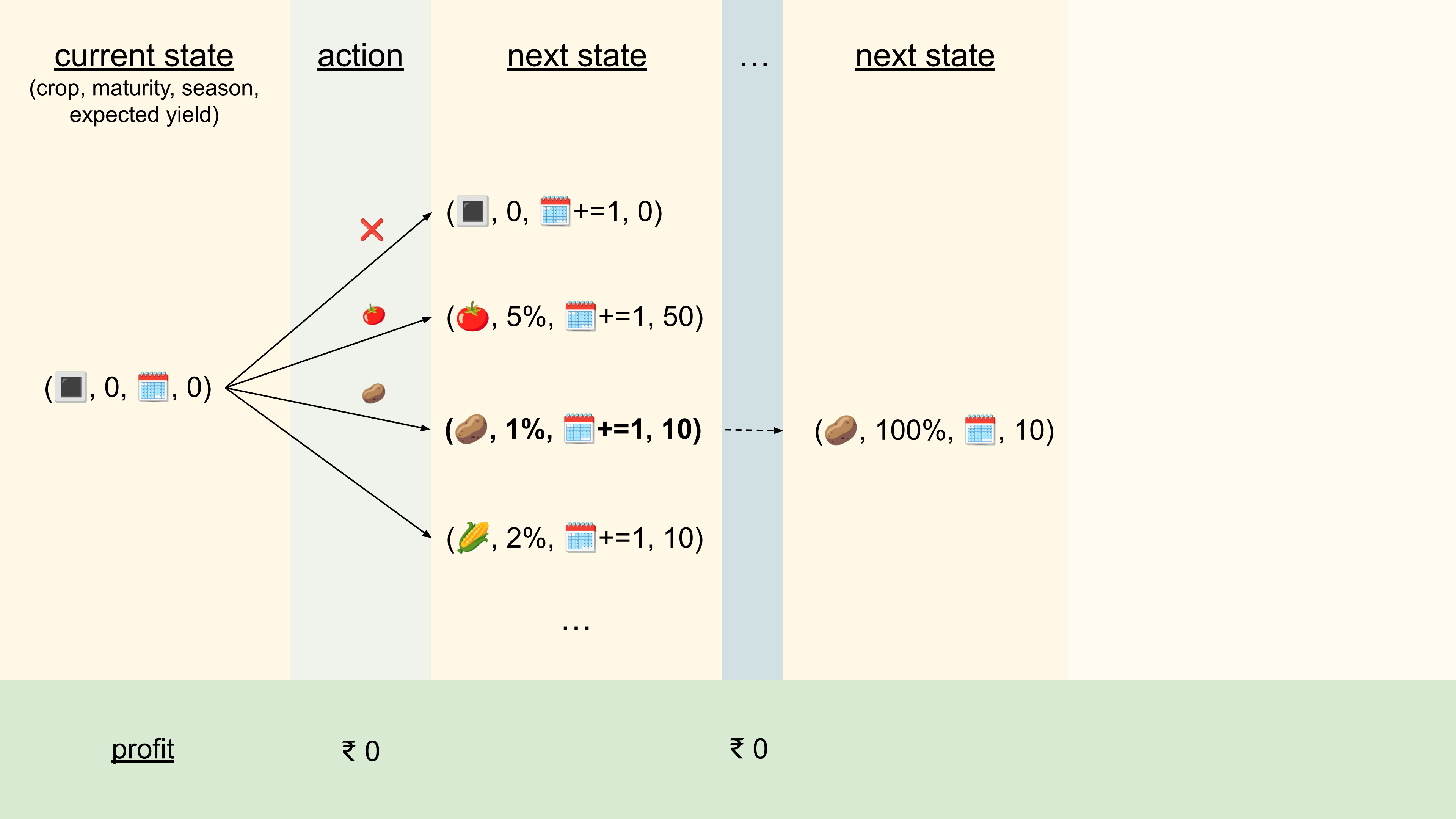
KHEYTI

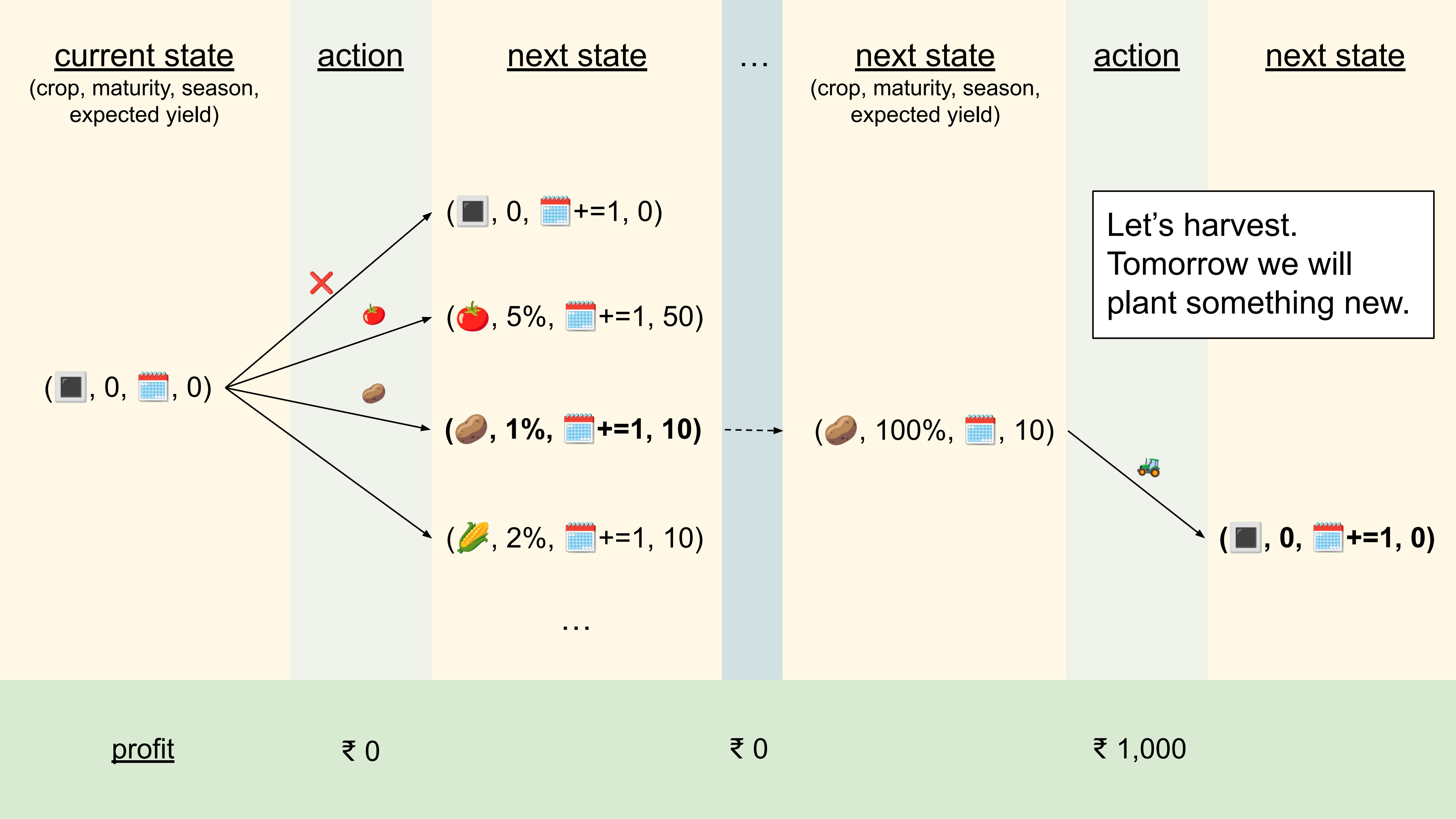
Challenge

Develop an optimization-driven **decision support system** for this low-resource sector, with a holistic eye towards real-world and multi-agent considerations.

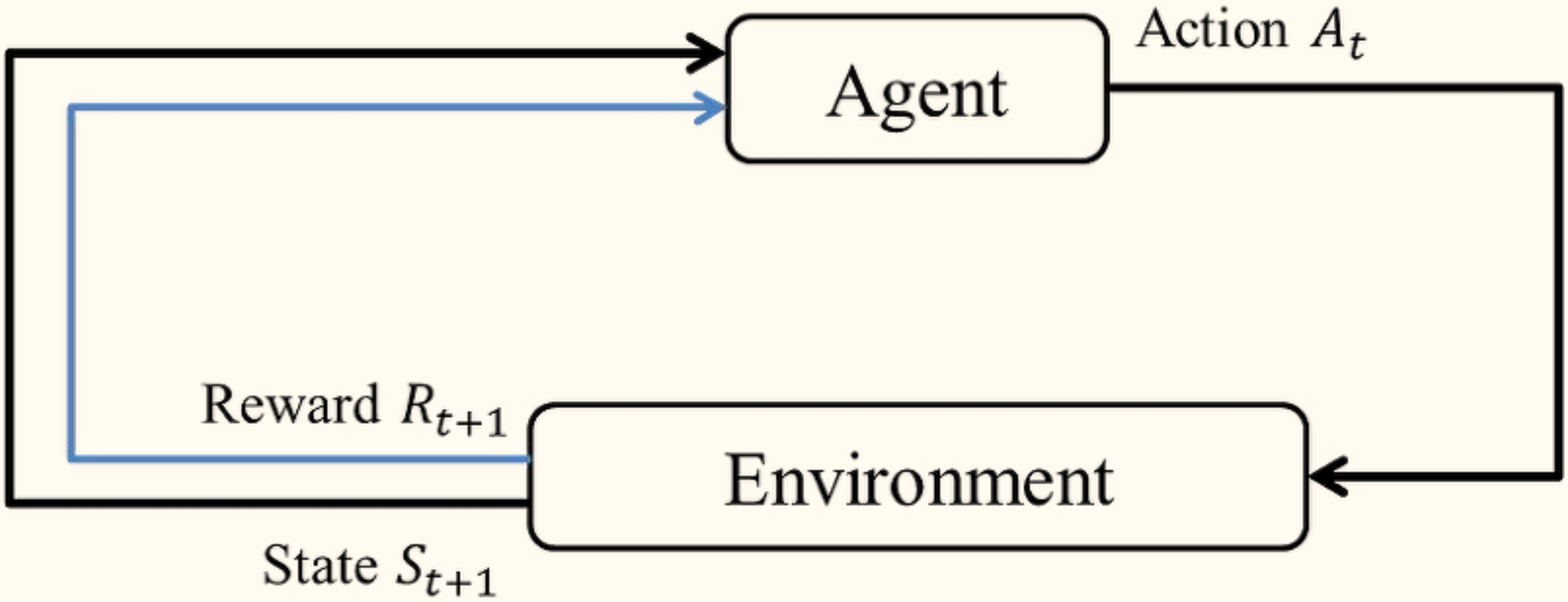
Decision support system

1. Plant **tomatoes** in December 2021
2. Plant **cucumbers** on April 24th, 2022
3. Plant **beetroot** on August 13th, 2022
4. Plant **cabbages** on November 11th, 2022


current state
(crop, maturity, season, expected yield)
action
next state
…
next state
(⬛, 0, 📅, 0)
❌
🍅
🥔
(⬛, 0, 📅+=1, 0)
(🍅, 5%, 📅+=1, 50)
(🥔, 1%, 📅+=1, 10)
(🌽, 2%, 📅+=1, 10)
…
(🥔, 100%, 📅, 10)
profit
₹ 0
₹ 0


current state
(crop, maturity, season, expected yield)
action
next state
…
next state
(crop, maturity, season, expected yield)
action
next state
(⬛, 0, 📅, 0)
❌
🍅
🥔
(⬛, 0, 📅+=1, 0)
(🍅, 5%, 📅+=1, 50)
(🥔, 1%, 📅+=1, 10)
(🌽, 2%, 📅+=1, 10)
…
(🥔, 100%, 📅, 10)
🚜
(⬛, 0, 📅+=1, 0)
Let's harvest. Tomorrow we will plant something new.
profit
₹ 0
₹ 0
₹ 1,000

Markov Decision Process



Markov Decision Process

State Space *(crop, maturity, expiry, date, flag)*

Action Space $\{N/A, harvest, plant\ c_1, plant\ c_2, \ldots\}$

Transition function $P_a(s, s')$

Reward function

$$r(s, a, s') = \begin{cases} < 0 & \text{if } a \text{ yields a constraint violation} \\ y(crop) & \text{if } a \text{ is } \textit{harvest} \\ 0 & \text{otherwise} \end{cases}$$

Solve!

Goal: maximize expected total discounted reward $\mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_{a_t}(s_t, s_{t+1})\right]$

$$V(s) := \sum_{s'} P_{\pi(s)}(s, s') \left(R_{\pi(s)}(s, s') + \gamma V(s')\right)$$

$$\pi(s) := \operatorname{argmax}_a \left\{ \sum_{s'} P_a(s, s') \left(R_a(s, s') + \gamma V(s')\right) \right\}$$

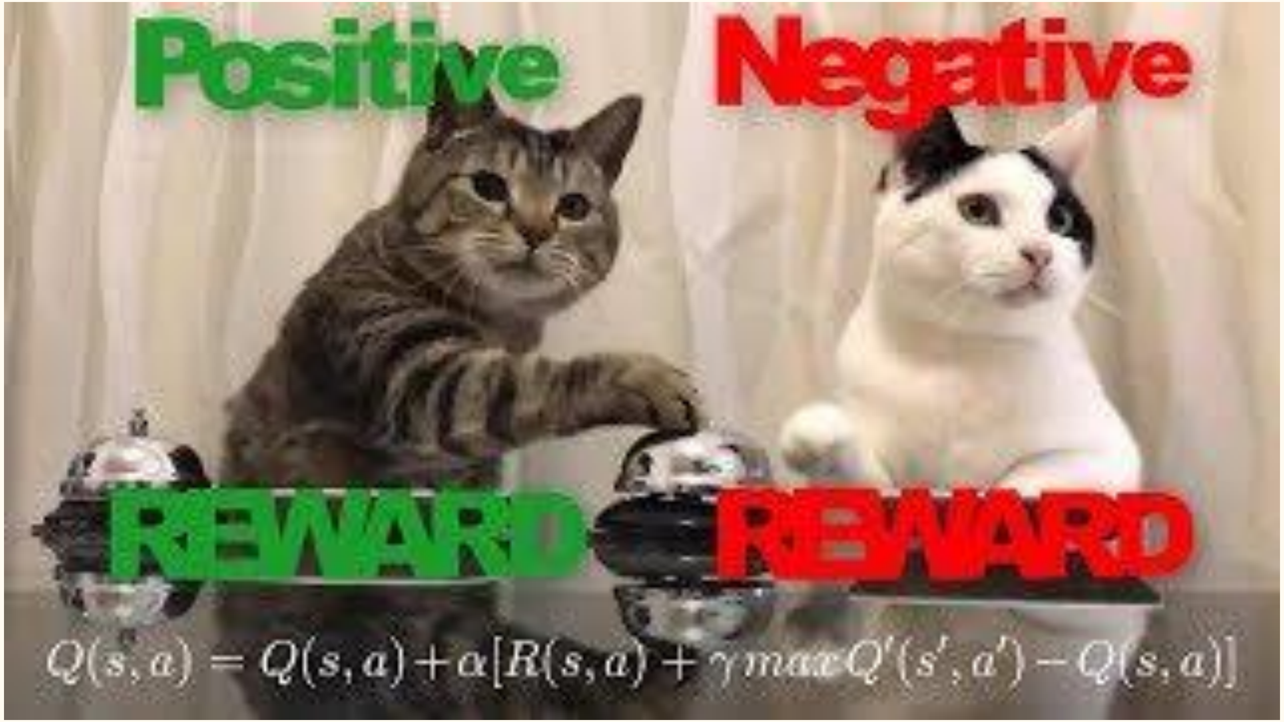
Our planning algorithm produced this!

1. Plant **tomatoes** in December 2023
   Harvest a couple of times, but before they're fully harvested
2. Plant **cucumbers** on April 24th, 2024
   Harvest a couple of times, but before they're fully harvested
3. Plant **beetroot** on August 13th, 2024
   Rip them out of the ground in November so that you can plant
4. Plant **cabbages** on November 11th, 2024

Last summer, we improved: *learning* instead of *planning*

Adjusting the plan as we interact with the environment

Reinforcement learning



Solve!

Algorithm 1 Follow the Weighted Leader for MDP

Input: Transition matrix P, parameter $\theta \in [0, 1)$, initial state s_0

Initialization: $\hat{R}_0$

for $t = 1 : H$ **do**

Update the weighted average of history rewards:

$$\hat{R}_t = (1 - \theta)\hat{R}_{t-1} + \theta R_{t-1}$$

Update estimates using observations

Solve the MDP given reward matrix $\hat{R}_t$ for the average optimal policy:

$$\pi_t \in \arg\max_{\pi} g_{\hat{R}_t}(\pi)$$

Recalculate the best action

Execute π_t, Update current State s_t

$R_t \leftarrow$ true reward matrix(from market data)

Observe what happens

end for

Output: π_t at each time step $t = 1, \ldots, H$

Decision support system

Our algorithm produced this!

1. Plant **bottle brinjal** in December 2023
   Wait until bottle brinjal are fully grown, then harvest
2. Plant **cucumber** on March 12th, 2024
   Wait until cucumbers are fully grown, then harvest
3. Plant **beetroot** on July 16th, 2024
   Wait until beetroot are fully grown, then harvest
4. Plant **cucumber** on October 22nd, 2024
   Wait until cucumbers are fully grown, then harvest

Sounds like we have a working decision support system that recommends crops.

So what are we working on this summer?

1. We could do an even better job!

    a. Do we really need to reevaluate what action to take every day?
    b. The algorithm doesn’t take into account future seasonality conditions.
    c. How would we explain the output to a farmer?

2. The **multiple** farmer universe

Is our recommendation still good if all of the farmers bring the same crops to market on the same day? **Probably not.**

Multi-agent reinforcement learning (MARL)

Each farmer is an agent motivated by their own rewards, and do actions to advance their own interests. In some MARL environments, these interests are opposed to the interests of other agents. We are in a unique position to provide small-scale coordination between agents.

Thank you