Multi-Agent Planning for Coordinated Robotic Weed Killing

Wyatt McAllister[1*], Denis Osipychev[2*], Girish Chowdhary[2], and Adam Davis[3]

Abstract— **This work presents a strategy for coordinated multi-agent weeding under conditions of partial environmental information. We aim to demonstrate the feasibility of coordination strategies for improving the weeding performance of autonomous agricultural robots. We show that, given a sufficient number of agents, the algorithm can successfully weed fields with various initial seed bank densities, even when multiple days are allowed to elapse before weeding commences. Furthermore, the use of coordination between agents is demonstrated to strongly improve system performance as the number of agents increases, enabling the system to eliminate all the weeds in the field, as in the case of full environmental information, when the planner without coordination failed to do so.**

As a domain to test our algorithms, we have developed an open source simulation environment, Weed World, which allows real time visualization of coordinated weeding policies, and includes realistic weed generation. In this work, experiments are conducted to determine the required number of agents are their required transit speed, for given initial seed bank densities and varying allowed days before the start of the weeding process.

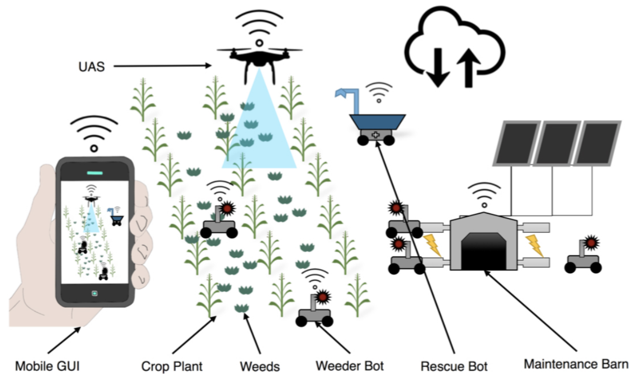

Fig. 1: Our solution for robotic mechanical weed control is a dynamically configured team of weeder bots, drones, and automated maintenance barns, which provide persistent autonomous weed-control, leveraging collaboration, as well as local and remote data sources.

I. Introduction

Weed management has historically relied on a combination of crop rotation, mechanical weed control, and the use of a variety of herbicides [23]. The evolution of herbicide-resistant weeds, coupled with the fact that new herbicide discovery has ceased in the past 30 years, has resulted in a crisis for agricultural weed management [9, 15]. Current crop losses due to herbicide resistant weeds are \$4 to \$6 billion per year, and may easily climb to \$100 billion per year when chemical control is lost [12]. Evolution of resistance to multiple sites of herbicide action is accelerating in dominant weeds, especially in the southern and north-central U.S. grain production regions [2]. Increasingly, farmers are only one site-of-action away from total loss of chemical control. For example, the five-way multiple resistant waterhemp (Amaranthus tuberculatus [Moq.] Sauer) in Illinois is now chemically controllable only by blocking the activity of enolpyruvylshikimate-3-phosphate synthase [5]. Transgenic crop cultivars with stacked resistance genes for multiple herbicides will only exacerbate resistance evolution in the many fields where herbicide resistance genes are already present [8]. To prevent a large disruption in food security, the weed management industry should move towards other forms of weed control.

*These authors contributed equally to this work
[1]Wyatt McAllister is with the Dept. of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. wmcalli22@illinois.edu
[2]Denis Osipychev and Girish Chowdhary are with the Dept. of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign. deniso2@illinois.edu, girishc@illinois.edu
[3]Adam Davis is with USDA-ARS in Urbana, IL. adam.davis@ars.usda.gov

Mechanical weed management usually targets young weeds, including germinating seeds and seedlings that are extremely vulnerable to physical damage. Before crop planting, superficial soil disturbance and subsequent soil cultivation can remove germinated weeds. However, after planting, mechanical weed control is usually limited to areas between crop rows. Hand weeding of young weeds at the two-leaf growth stage is extremely difficult and not practical for large farms. Mechanized inter-row cultivation has disadvantages, such as soil compaction due to use of heavy machinery, and am inability to work after the crop canopy closes. Due to crop canopy growth, no current mechanical weed control method is effective within the crop [16]. Our work suggests that a team of collaborative low-cost and lightweight mechanical weeding robots (termed here as *agbots* shown in Fig. 1) can be utilized for controlling herbicide-resistant weeds. The team of *agbots*, designed to target weeds within and between crop rows, possesses advantage over large agricultural equipment such as tractors, combines, and planters, which cannot be used after the crop canopy grows.

Termination of weed seedlings within several days after they emerge is critical to preventing crop yield losses in corn and soybeans [20]. For many crops, including corn, weeding may be done under a canopy, and therefore under conditions of partial environmental information. To be effective, this robotic system needs to plan robustly, operating in dynamic environments, utilizing limited information about these environment to efficiently complete the task under time constraints. This work aims to present a comprehensive study of the feasibility of a coordinated robotic weeding approach in realistic field environments. Our work aims to leverage strategies for multi-robot coordination to create a scalable weeding solution. Our goal is to ensure that robots have the ability to coordinate their actions under varying amounts of environmental information, updating a shared environmental model as they move through the field, and optimizing weeding efficiency.

Foraging, where robots move through an environment and collect objects or information, has long been considered a key problem in multi-agent robotics [4]. In our case, the foraging problem is framed in terms of recognizing and killing weeds while moving through the field. We will build on past work in coordinated robotics [11, 14, 3, 13, 24] to create a system for cooperative robotic weeding which addresses the problem of partial environmental information without relying on a separate agent for information gathering.

In this work, the solution relies on optimization over a ”reward” metric, which is chosen to be the total of the maximum height of weeds in every 0.8 square meter region of the field. This ensures that the system eliminates weeds before they grow too large for the mechanical weeding process to deal with. By optimizing over this reward metric, we find a strategy with ensures the field can be weeded completely, and prevents weeds from growing large enough to start seeding.

A. Contributions

This paper presents an approach to coordinated robotic weeding. In order to demonstrate the feasibility of our approach, we benchmark the performance of our weeding method against a method which does not utilized shared information between the agents. We find that our method is able to eliminate all the weeds in the field, as is the case when the planner has full information about the environment, when the planner without shared information is not able to do so. Furthermore, after testing our method over many trials, with a range of initial seed bank densities, and a varying number of allowed days of weed growth before weeding commences, we find that our planner is able to succeed in every case, as long as enough agents are utilized, their transit speed is large enough, and the weeds have not grown too large for the robot to kill before the weeding process starts. Based on the results in this work, we believe that our method will be feasible for collaborative robotic weeding in uncertain environments.

To efficiently test algorithms for coordinated weeding, and their performance change with respect to various parameters over time, we perform our experiments in an open source simulation environment of our own design, Weed World (shown in Fig. 2.), which enables real-time visualization of coordinated weeding policies, and incorporates a realistic weed growth model. In this environment, we discretize the field into a grid world of 85 rows, 0.8 meter wide, totaling 4047 square meters, or one acre. The simulation environment allows efficient determination of design heuristics which will inform implementations of coordinated weeding systems used in real field experiments, and will enable other researchers to test their own algorithms in the same framework.

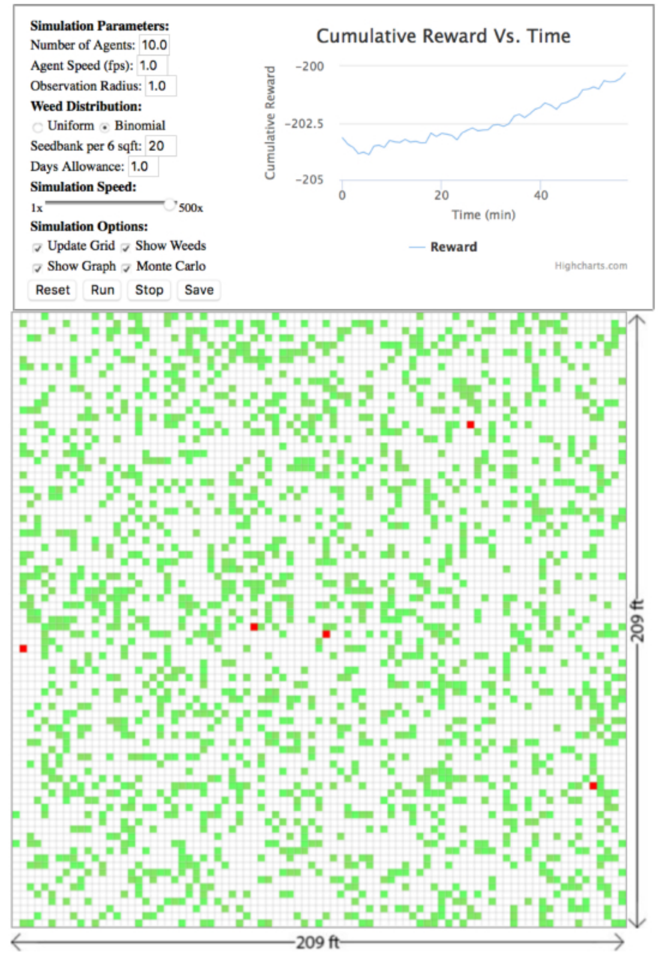

Fig. 2: Simulation environment Weed World designed in JavaScript

B. Formulation of Weeding Problem

This work aims to demonstrate the feasibility of this coordinated approach for multi-agent weeding, and showcase how the use of information collected from multiple agents may improve system performance. The problem is framed as a coordinated multi-robot task allocation problem. In this problem formulation, the agents collect environmental information during system operation, and share this information in order to plan a coordinated weeding policy which allows higher weeding performance than agents would be able to accomplish on their own.

C. Summary

The next section, Section II presents the formal taxonomy of the coordinated multi-agent task allocation problem. Section III presents the methods utilized to solve the problem. Section IV presents an interpretation of the results of the experiments conducted. Section V presents conclusions and an outline of further work. Finally, Section VI presents acknowledgments to our funding association, and collaborators.

II. Background

Before presenting the methods utilized to solve the weeding problem, we present a formal taxonomy of coordinated robotics and multi-agent task allocation. We introduce key distinctions between common problems in multi-robot coordination, and between different strategies for task allocation methods utilized to solve these problems. We ground our chosen method in this body of literature by explaining where our method fits within this taxonomy.

In, [4], a formal taxonomy of cooperative robotic planning is presented. This work presents the problem domain of foraging, where robots move through an environment and collect objects or information. In our case, the foraging problem is framed in terms of recognizing and killing weeds while moving through the field.

In, [4], three major distinctions between various cooperative robotic tasks are drawn. The first is the distinction between synchronous planning, where tasks are delegated to all agents at the same time, and asynchronous planning, where tasks are delegated a varying times when agents become available. In our problem, we assume a generalized field, where robots may not move side by side down a row, and where there are no paths to adjacent rows in the middle of the field, so that are system is generalizable to arbitrary fields and does not rely on field configuration. Under this assumption, assigning a single agent to each row is necessary. Varying density of weeds and terrain in the rows will cause the time to complete a row to be different from that which is estimated beforehand. Asynchronous planning allows us to calculate the optimum row for each agent as it becomes available, in order to complete tasks with unknown duration.

The second distinction is between homogeneous agents having identical capabilities, and non-homogeneous agents having varying capabilities. In this work, homogeneous robotic agents are utilized in order for the system to be easily mass produced and scalable. The use of homogeneous agents allows 3D printed robots to be distributed at low cost, without relying on specialized designs for agents with varying capabilities. This work aims to show that collaboration between homogeneous agents will enable a feasible weeding solution. The assumption of homogeneous agents helps us construct a factored model more easily, as agents with identical capabilities have identical value functions for the same task assignment for the same starting state.

The third distinction is between centralized planning, in which the planner optimizes task allocation for all agents using the same model, and decentralized planning, where each agent has a local planner that performs optimization based on its own environmental information. In this work, we perform experiments on a simulated environment where a set of coordinated agents performs centralized task allocation via a shared environmental model, allowing us to leverage all available environmental information to allocate tasks.

Past work has explored Multi-Robot Task Allocation (MRTA) in stochastic domains [18, 6, 17], leveraging both spatial constraints and predictive information to perform optimization. In, [7], a formal taxonomy of Multi-Robot Task Allocation (MRTA), is presented. Three major distinctions are again made. The first is between single-robot problems, where each pool of tasks is managed by a separate robot, and multi-robot problems, in which each task pool is shared between multiple robotic agents. Our problem is a multi-robot problem, since all agents cooperate the weed the field together in order to complete the weeding task more efficiently. This approach allows all the agents to adapt to changes in the environment and work together on regions of the field which have more weeds.

The next distinction is between preemptive task allocation, in which optimization is performed continuously in an on-line manner, and agents may take over another agent's task, or switch to another task before completion, and non-preemptive task allocation, in which tasks must be completed before a new task is assigned. We use a non-preemptive planning strategy, ensuring rows are completed before an agent is assigned a new row. This allows agents to plan a new row once the task has been completed, allowing them to focus computational resources on navigation and plant recognition while weeding the row.

Another distinction is between single-agent tasks, in which each task must be performed by one agent, and multi-agent tasks, in which each task must be performed by multiple agents. Here, each robot is assigned to one row, so our problem is a single-agent task scenario, with multiple agents collaborating to complete a pool of single-agent tasks.

In [7], the problem of time-extended on-line assignment, in which multiple robots need pick single-agent tasks from a pool larger than the number of agents, and complete them in a non-preemptive manner, is considered. Our algorithm is an implementation of that proposed in [7] for the time-extended on-line assignment problem, which picks an initial assignment of each robot to the most suitable task, and assigns each robots to the most suitable task from the pool as they become available.

III. Methods

This section details the methods used in this work. We first explain the weed growth model, as well as the state, action, and reward models utilized. We then introduce the optimization framework used, detailing the value function utilized for optimization. We next explain the dynamic programming algorithm used to solve the optimization problem. We then detail the algorithm for targeted information gathering used, which allows agents to simultaneously gather environmental information while performing coordinated weeding. Finally, we present an outline of the experiments conducted.

A. Weed Growth Model

The weed growth model utilized in this paper is based on Bernoulli random variables, with seeds emerging from a limited seed bank, forming a binomial distribution over time. The initial seed density of the seed bank in each square is S_0, which is uniformly distributed in the spatial domain. Upon initialization of the simulation, a certain number of days, d_0, are allowed to elapse before weeding starts. Both parameters, S_0 and d_0, are benchmarked against the number of agents and their transit speed in order to determine the feasibility of mechanical weeding with the team of small robots. The number of emerging weeds in each square, N_{emerge}, is a randomly generated Poisson variable with mean, $\lambda(x, y, t)$, such that ninety percent of the seed bank, $S(x, y, t)$, emerges in T_{total}, which is two months. This emergence rate is aligned with past work [19, 21, 25, 22, 10], which has presented detailed analysis of weed growth models, in which measurements of seed bank density for various species of weeds were conducted. Our estimate of the seed bank selected as up to 100 seeds per 0.5 meter square is realistic for some species of plants. However, we avoid limiting the paper to a specific species of weed due to the region-specific nature of this limitation.

$$\lambda_0 = \frac{0.9 \cdot d_0 \cdot S_0}{T_{\text{total}}} \tag{1}$$

$$\lambda_t(x, y, t) = \frac{0.9 \cdot \Delta t \cdot S(x, y, t)}{T_{\text{total}}} \tag{2}$$

$$N_{\text{emerge}}(x, y, t) = \text{Poi}\left(\lambda_t(x, y, t)\right) \tag{3}$$

$$S(x, y, t) = S_0 - \sum_{t=t_0}^{t_{\text{current}}} N_{\text{emerge}}(x, y, t) \tag{4}$$

The weed density at each square $\zeta(x, y, t)$, grows as seeds emerge from the seed bank. The maximum weed height at each square $\delta(x, y, t)$, increases from zero height at a fixed rate Γ inches per day.

$$\zeta(x, y, t) = \sum_{t=t_{\text{last weeded}}}^{t_{\text{current}}} N_{\text{emerge}}(x, y, t) \tag{5}$$

$$\delta(x, y, t) = \left(\frac{t_{\text{current}} - t_{\text{lastweeded}}}{60 \cdot 60 \cdot 24}\right)\left(\Gamma \frac{\text{inch}}{\text{day}}\right) \tag{6}$$

Due to limitations of mechanical weeding, it is highly important to remove weeds before they become too large to be eliminated by the specific weeding tools available to small *agbots*. We therefore define the reward for weeding each square $R_W(x, y, t)$ to be equal to the maximum height of weeds in the square $\delta(x, y, t)$.

$$R_W(x, y, t) = \delta(x, y, t) \tag{7}$$

B. State and Action Model

In the following equations, N_{dim} is the number of squares in a row (85), N_{agents} is the number of agents, Y_{dim} is the length of each row (64 meters), $R_W(x, y, t)$ is the reward for each location (x, y), and v_i is the agent velocity.

The environmental state depends on the x and y positions of each agent. The action is chosen to be the target row chosen by each agent. Here, N_{dim} is the number of rows in the field, 85, and N_{agents} is the number of agents.

$$S \equiv \{1, ..., N_{\text{dim}}\} \times \{1, ..., N_{\text{dim}}\} \tag{8}$$

$$I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{9}$$

$$\left(x_i(t), y_i(t)\right) \in S \quad \forall i \in I \tag{10}$$

$$a_i(t) = \left\{x_i(t+1), y_i(t+1)\right\} \in A \equiv S \tag{11}$$

The order of the state space is $85 \times 85 \times N_{\text{agents}}$, which is too large for efficient computation. Given our assumption that the problem is non-preemptive, and the capability of agents to observe neighboring rows, we reduce the dimensionality of the problem. We assume that since agents finish rows once starting to weed them, only the x location is relevant for the state and action. The new size of the space is $85 \times N_{\text{agents}}$.

$$S \equiv \{1, ..., N_{\text{dim}}\} \tag{12}$$

$$I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{13}$$

$$x_i(t) \in S \quad \forall i \in I \tag{14}$$

$$a_i(t) = x_i(t+1) \in A \equiv S \tag{15}$$

C. Reward Model

The total reward is composed of the reward for each square of weeds in the field.

$$S \equiv \{1, ..., N_{\text{dim}}\} \times \{1, ..., N_{\text{dim}}\} \tag{16}$$

$$I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{17}$$

$$R_W(x, y, t) \quad \forall (x, y) \in S \quad \forall i \in I \tag{18}$$

However, we plan only over the observed portion on the environment, which is composed of the rows adjacent to those previously weeded. We keep track of the estimated density and maximum height for each observed square, using this to estimate a total scalar reward for each observed row. This is the only required information on the reward.

$$A \equiv \{1, ..., N_{\text{dim}}\} \tag{19}$$

$$R_i\left(a_i(t)\right) = \sum_{y=1}^{N_{\text{dim}}} R_W\left(a_i(t), y, t\right) \quad \forall a_i(t) \in A \tag{20}$$

D. Optimization Framework

In the optimization problem of interest, we optimize the total reward for each action, time discounted by the expected operation time to complete that action. This value metric has long been used in robot foraging tasks [14]. To expedite computation, we use a factored approach [1], where the reward is an additive function of individual agent rewards.

The planned operation time is the sum of the time it takes to move to the proposed row $T_{\text{move to row}}$, the time it takes to move down it $T_{\text{move down row}}$, and the time it takes to weed all the squares in the row $T_{\text{weed row}}$.

$$T_i\left(x_i(t), a_i(t)\right) = T_{\text{to row}} + T_{\text{down row}} + T_{\text{weed row}} \tag{21}$$

$$T_{\text{to row}} = \frac{\left(a_i(t) - x_i(t)\right)}{v_i} \tag{22}$$

$$T_{\text{down row}} = \frac{Y_{\text{dim}}}{v_i} \tag{23}$$

$$T_{\text{weedrow}} = T_{\text{kill}} \cdot \sum_{y=0(t)}^{N_{\text{dim}}} \delta\left(x_i(t), y(t)\right) \tag{24}$$

For this problem, we want to maximize the overall value function, which is the sum over all agents of the planned reward, time discounted by the planed operation time.

$$V(t) = \sum_{i \in I} \gamma^{T_{\text{i}}\left(x_i(t), a_i(t)\right)} R_{\text{i}}\left(a_i(t)\right) \tag{25}$$

E. Dynamic Programming (DP) Algorithm

For the Dynamic Programming (DP) Algorithm, we plan across all the agents, evaluating the value for a transition from the agent's current state to its proposed new state. We plan a coordinated policy which sends each agent to the row with maximum value. We then assign agents asynchronously to the row with the highest value when they query the planner after completing a row.

$$V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \alpha\left(\gamma^{T_i\left(a_i(t)\right)} \cdot R_i\left(a_i\left(t\right)\right)\right) \tag{26}$$

$$a_i\left(t\right) = \underset{a_i(t)}{\arg\max}\, V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \qquad \forall i \in I \tag{27}$$

F. Information Gathering Trade-Off

The naive approach for information gathering is to simply go to the next available adjacent unexplored row. In order to improve performance, we would like to consider an approach which targets information gathering to ensure the largest increase in the total explored space.

We compute the average reward, $\bar{R}$, as the sum of rewards for all agents from the time every row was last visited $t_{\text{exp.}}$ to the current time, divided by the total number of rows weeded since every row was last visited $N_{\text{rows weeded}}$.

$$\bar{R} = \frac{\sum\limits_{t=t_{\text{exp.}}}^{t_{\text{current}}} \sum\limits_{i=0}^{N_{\text{agents}}} R_i\left(a_i\left(t\right)\right)}{N_{\text{rows weeded}}} \tag{28}$$

The information index of a row, $I\left(a_i\left(t\right)\right)$, is the number of rows which would be explored by going to that row.

$$I\left(a_i\left(t\right)\right) = \sum_{i=-r_{obs}}^{r_{obs}} \mathrm{I}_{\left\{\text{is explored}\left(x=a_i(t)+i\right)\right\}} \tag{29}$$

We compute $\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$, as the value function with $\bar{R}$ times $I\left(a_i\left(t\right)\right)$.

$$\begin{aligned}&\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \\ &= \alpha\left(\gamma^{T_i\left(x_i(t),a_i(t)\right)} \cdot \bar{R} \cdot I\left(a_i\left(t\right)\right) - \bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)\right)\end{aligned} \tag{30}$$

We denote the exploration value for each unexplored row by $E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$, which is equal to the estimated value function for that row $\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$.

$$E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \tag{31}$$

We then explore rows with exploration value greater than or equal to the maximum value for explored rows.

$$\begin{aligned}&\underset{a_i(t)}{\arg\max}\, E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \\ &\geq \underset{a_i(t)}{\arg\max}\, V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \\ &\Rightarrow a_i\left(t\right) = \underset{a_i(t)}{\arg\max}\, E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)\end{aligned} \tag{32}$$

If no rows have been explored, then we go to the next available adjacent unexplored row.

Input: $\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$: estimated value function
Output: $a_i(t)$: action for each agent
for *all rows and all agents* **do**

$$E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$$

if $\underset{a_i(t)}{\arg\max}\, E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \geq \underset{a_i(t)}{\arg\max}\, V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$ **then**

$$a_i\left(t\right) = \underset{a_i(t)}{\arg\max}\, E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$$

end
end

Algorithm 1: Information Gathering Algorithm

TABLE I: Table 1: Here, r_{obs} is the observation radius, N_{agent} is the number of agents, v_{agent} is the agent velocity, d_0 is the days of allowed weed growth before weeding, S_0 is the initial seed bank density. An 'X' denotes a parameter for a Monte Carlo run over the ranges shown in the last column

Exp.	1	2	3	4	5	6	7	**Range**
r_{obs}	∞	0	1	1	1	1	1	[0,∞]
N_{agent}	5	5	5	5	5	X	X	[3,10]
v_{agent}	1	1	1	X	X	1	1	[1,3]
d_0	3	3	3	3	X	3	X	[1,6]
S_0	20	20	20	X	20	X	20	[10,100]

G. Experiment Plan

We conduct seven experiments, each with 100 trials with varying initial parameters shown in Table I. Each trial is run for 4 days of simulated time. We first run the algorithm in the case of full environmental information, where the planner has complete knowledge of the reward for each square within the environment, which is chosen to be the maximum height of weeds within that square, in order to establish an ideal benchmark. We then run our algorithm with observation radius, $r_{\text{obs}} = 0$, to establish a worst case scenario in terms of the information the planner has available. Finally, we run the algorithm with an observation radius, $r_{\text{obs}} = 1$, to see how performance is improved in the case of partial environmental information when information about neighboring rows is used.

We then do Monte Carlo runs to determine feasibility of the method with respect to the change in the number of agents, N_{agent}, their velocity, v_{agent}, the days allowance, d_0, and the initial seed bank density, S_0. The baseline values of the parameters and the magnitude of their ranges for Monte Carlo runs, are shown in Table I. These Monte Carlo runs will allow us to determine design heuristics for coordinated robotic weeding in our simulated domain. In further work, these design heuristics may be used to refine the design of robotic agents used in real field experiments. They will also inform further work on optimizing the weeding algorithm for various application domains.

IV. Results

In this section we present a detailed interpretation of the results of the experiments detailed in Section III-G.

A. Experiments 1 - 3

As seen in Figure 3, partial environmental information, $r_{obs} = 1$, significantly improves performance over the case of zero information, $r_{obs} = 0$. The partial environmental information case eliminates all the weeds in the field by the end of the experiment, giving a terminal reward of zero, as in the case of full environment information. However, in the case of zero information, the algorithm is unable to complete the weeding of the field when weeds become sparse, as new weeds grow faster than the planner with $r_{obs} = 0$ can find and kill them. Finally, for this case of zero information, the variance is higher, as the performance planner depends heavily on the configuration of the field.

B. Experiments 4

As seen in Figure 4, for a fixed number of robots and a high seed bank density, greater than 20 seeds per square, the system will not be able to succeed in weeding the field within the time frame of the experiment, as the terminal reward is greater than zero. This is because when seed bank density grows large, every square will eventually become infested, and the speed of transit of the robot will not effect the weeding performance.

C. Experiments 5

As seen in Figure 5, for fixed seed bank density, as the days allowance, the number of days the weeds are allowed to grow before weeding commences, increases past 4 days, the system will not be able to succeed at any speed. This is because when the days allowance becomes large enough, the field is initially fully infested and the transit speed of the robot is unimportant.

D. Experiments 6

As seen in Figure 6, as the seed bank density increases, a higher number of agents are needed to complete the field. However, we observe that with 10 agents, seed bank densities of up to 60 can be handled by the system. Furthermore, there is a strong positive correlation between the initial seed bank density and the required number of agents for this density, suggesting the weeding solution will succeed on fields with varying initial seed bank density given enough agents.

E. Experiments 7

As seen in Figure 7, as days allowance increases, more agents are needed to complete the field. However, with 10 agents, the system can handle a days allowance up to 4, when weeds start to grow higher than the maximum height which the system is capable of weeding, and the system starts to fail. However, the terminal reward continues to decreases for increasing number of agents, even when the system is not able to eliminate all the weeds, suggesting that with enough agents, the system can successfully kill weeds which have not initially grown higher than the allowable height.

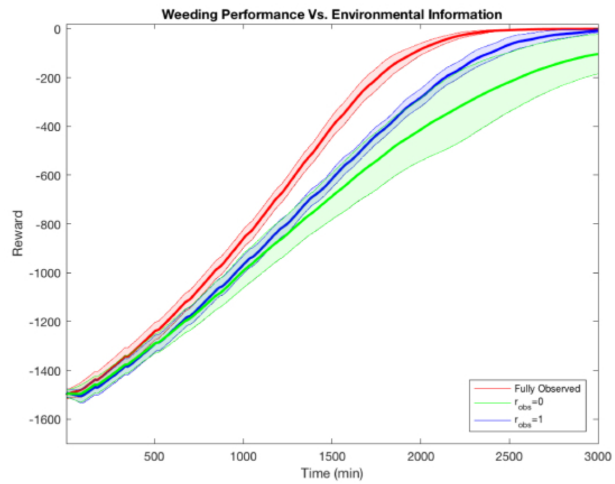

Fig. 3: Weeding Performance Vs. Environmental Information: We plot the weeding performance over time for the case of full environmental information, partial environmental information $r_{obs} = 1$, and zero environmental information, $r_{obs} = 0$. We see that the case of $r_{obs} = 1$ is able to weed the entire field, converging to a zero terminal reward corresponding to total weed elimination, as in the case of full environment information, when the case of $r_{obs} = 0$ is not able to do so.

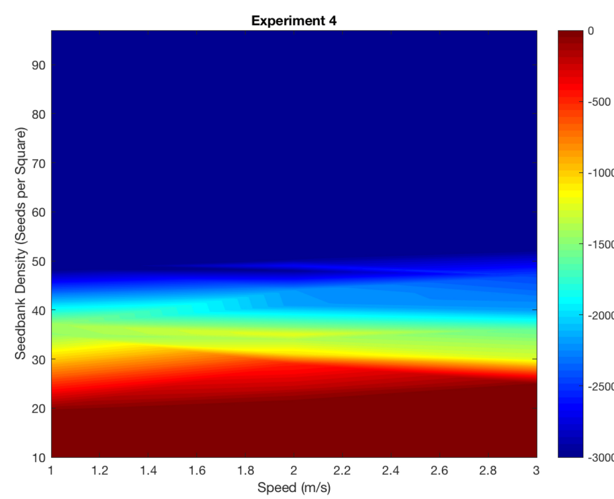

Fig. 4: Speed Vs. Seed Bank Density: The heat map of the terminal reward for 100 trials with varying agent speed and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Agent speed is not strongly correlated with the initial seedbank density, as for high enough seed bank density the field becomes fully infested and the transit speed of the robot through empty squares becomes irrelevant.

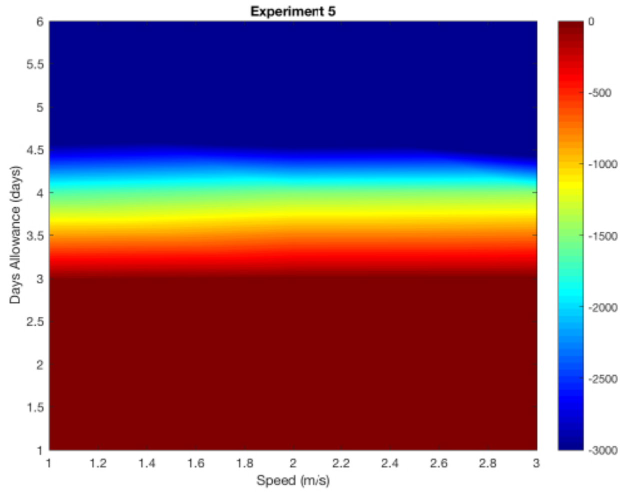

Fig. 5: Speed Vs. Days Allowance: The heat map of the terminal reward for 100 trials with varying agent speed and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Agent speed is not strongly correlated with the initial days allowance, as for high enough days allowance the field becomes fully infested and the transit speed of the robot through empty squares becomes irrelevant.

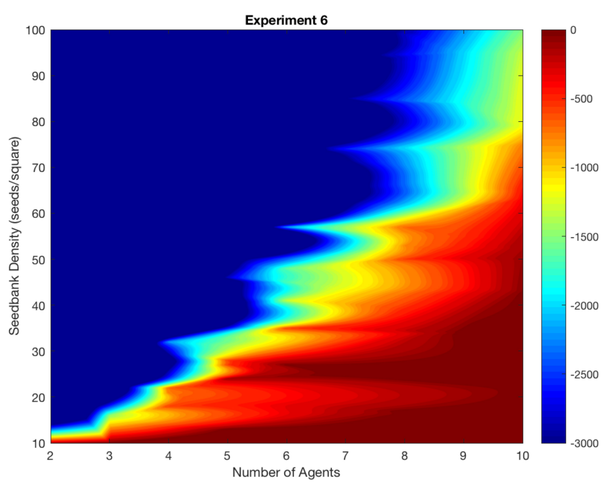

Fig. 6: Number of Agents Vs. Seed Bank Density: The heat map of the terminal reward for 100 trials with varying numbers of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. There is a strong positive correlation between the number of agents and the initial seed bank density, suggesting it is possible to weed fields with an varying initial seed bank densities with a large enough number of robots.

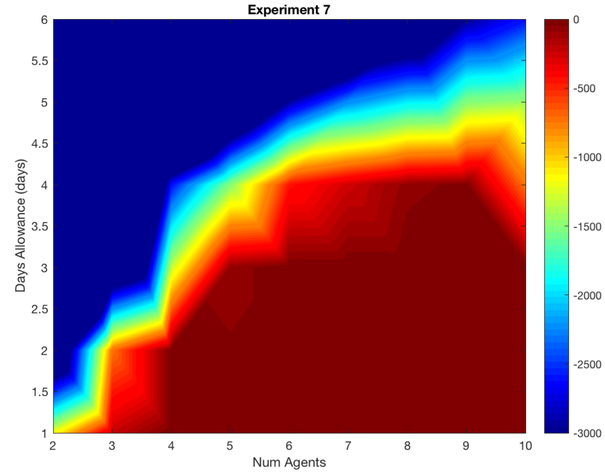

Fig. 7: Number of Agents Vs. Days Allowance: The heat map of the terminal reward for 100 trials with varying numbers of agents and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The days allowance and number of agents are strongly correlated, with a sufficient number of agents being able to cover a field with days allowance up to four days, where the trend levels off as weeds initially grow too large for weeding.

V. Conclusions and Further Work

This research demonstrates that a scale neural approach to coordinated multi-agent weeding in uncertain environments, utilizing a varying number of robots for different agricultural applications, will feasibly be able to adapt to fields with varying seed bank densities. Our approach outperforms the case in which coordination strategies were not utilized, eliminating all the weeds in the field, and exhibiting comparable performance to the case of full environmental information, when the planner without coordination failed to do so. Our results show clear improvement in performance for an increased number of agents, demonstrating the usefulness of coordination strategies for weeding fields which agents would not be able to complete on their own. We simulate trials with increasing seed bank density, and show that a larger number agents are not only able to fully weed the field when smaller teams cannot due so, but that they can drive down the weed population even after the field has become fully infested, with some weeds larger than the system is able to kill. We feel that these results clearly show that multi-robot coordination is not just useful for coordinated weeding, but that it is in fact, essential, and will be a central part of mechanical weeding solutions to the weeding crisis.

Our estimates for the range of seed bank densities hold for several species of plants. However, we will attempt to extend our work to include experiments utilizing seed bank densities within the ranges shown in the data for a larger number of species, in order to determine the optimal algorithmic parameters for applications in different environmental domains.

VI. Acknowledgements

This work was supported in part by an NSF/USDA CPS grant that is currently under contracting. Furthermore, we want to extend our thanks to all the collaborators on this project whose past work in weed recognition, design of the mechanical weeding system, and field navigation, has inspired this work.

References

[1] Christopher Amato et al. "Decentralized control of partially observable Markov decision processes". In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE. 2013, pp. 2398–2405.

[2] Muthukumar V Bagavathiannan and Jason K Norsworthy. "Multiple-herbicide resistance is widespread in roadside Palmer amaranth populations". In: *PloS one* 11.4 (2016), e0148748.

[3] Tucker Balch. "The impact of diversity on performance in multi-robot foraging". In: *Proceedings of the third annual conference on Autonomous Agents*. ACM. 1999, pp. 92–99.

[4] Y Uny Cao, Alex S Fukunaga, and Andrew Kahng. "Cooperative mobile robotics: Antecedents and directions". In: *Autonomous robots* 4.1 (1997), pp. 7–27.

[5] Cody Matthew Evans. "Characterization of a novel five-way-resistant population of waterhemp (Amaranthus tuberculatus)". PhD thesis. 2016.

[6] Emilio Frazzoli and Francesco Bullo. "Decentralized algorithms for vehicle routing in a stochastic time-varying environment". In: *Decision and Control, 2004. CDC. 43rd IEEE Conference on*. Vol. 4. IEEE. 2004, pp. 3357–3363.

[7] Brian P Gerkey and Maja J Matarić. "A formal analysis and taxonomy of task allocation in multi-robot systems". In: *The International Journal of Robotics Research* 23.9 (2004), pp. 939–954.

[8] Jonathan Gressel, Aaron J Gassmann, and Micheal DK Owen. "How well will stacked transgenic pest/herbicide resistances delay pests from evolving resistance?" In: *Pest management science* 73.1 (2017), pp. 22–34.

[9] I Heap. "The international survey of herbicide resistant weeds http://www. weedscience. org". In: *asp (http://www. weedscience. org/In. asp)* (2017).

[10] Michael J Horak and Thomas M Loughin. "Growth analysis of four Amaranthus species". In: *Weed Science* 48.3 (2000), pp. 347–355.

[11] Miao Liu et al. "Learning for Multi-robot Cooperation in Partially Observable Stochastic Environments with Macro-actions". In: *CoRR* abs/1707.07399 (2017).

[12] Michael Livingston, Jorge Fernandez-Cornejo, and George B Frisvold. "Economic returns to herbicide resistance management in the short and long run: The role of neighbor effects". In: *Weed Science* 64.sp1 (2016), pp. 595–608.

[13] Maja J Mataric. "Learning in behavior-based multi-robot systems: Policies, models, and other agents". In: *Cognitive Systems Research* 2.1 (2001), pp. 81–93.

[14] Maja J Matarić. "Reinforcement learning in the multi-robot domain". In: *Autonomous Robots* 4.1 (1997), pp. 73–83.

[15] Bruce D Maxwell, Mary Lynn Roush, and Steven R Radosevich. "Predicting the evolution and dynamics of herbicide resistance in weed populations". In: *Weed technology* 4.1 (1990), pp. 2–13.

[16] Charles L Mohler, James C Frisch, and Jane Mt Pleasant. "Evaluation of mechanical weed management programs for corn (Zea mays)". In: *Weed Technology* 11.1 (1997), pp. 123–131.

[17] Angélica Munoz-Meléndez, Prithiviraj Dasgupta, and William Lenagh. "A Stochastic Queueing Model for Multi-robot Task Allocation." In: *ICINCO (1)*. 2012, pp. 256–261.

[18] J Niko-Mora. *Stochastic scheduling. Updated version of article in Encyclopedia of Optimization, CA Floudas and PM Pardalos, eds., V: 367–372, 2001*.

[19] Dawn Nordby and Kevin Bradley. "Biology and Management of Waterhemp". In: (Feb. 2018).

[20] Eric R Page et al. "Why early season weed control is important in maize". In: *Weed Science* 60.3 (2012), pp. 423–430.

[21] Brian J Schutte and Adam S Davis. "Do Common Waterhemp (Amaranthus rudis) Seedling Emergence Patterns Meet Criteria for Herbicide Resistance Simulation Modeling?" In: *Weed technology* 28.2 (2014), pp. 408–417.

[22] Brent A Sellers et al. "Comparative growth of six Amaranthus species in Missouri". In: *Weed Science* 51.3 (2003), pp. 329–333.

[23] Dale L Shaner. "Lessons learned from the history of herbicide resistance". In: *Weed Science* 62.2 (2014), pp. 427–431.

[24] Yoav Shoham, Rob Powers, and Trond Grenager. "Multi-agent reinforcement learning: a critical survey". In: *Web manuscript* (2003).

[25] Rodrigo Werle et al. "Predicting emergence of 23 summer annual weed species". In: *Weed science* 62.2 (2014), pp. 267–279.