MULTI-AGENT PLANNING FOR COORDINATED ROBOTIC WEED KILLING

BY

WYATT SPALDING MCALLISTER

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Adviser:

Assistant Professor Girish Chowdhary

ABSTRACT

This work presents a strategy for coordinated multi-agent weeding under conditions of partial environmental information. Our goal is 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 and their required transit speed, for given initial seed bank densities and varying allowed days before the start of the weeding process.

To my grandmother, Edith McAllister.

ACKNOWLEDGMENTS

This work was supported in part by an NSF/USDA CPS grant that is currently under contract. Furthermore, I want to extend my 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.

I would like to thank Dr. Adam Davis and Dr. Girish Chowdhary, from the Department of Agricultural and Biological Engineering, for their valued contributions to this work in reviewing the thesis presented here. I would also like to thank Denis Osipychev, with whom we collaborated on the development of the Weed World simulation software, for his valued contribution to this work. The current simulation environment was developed from a previous JavaScript environment utilized by Denis Osipychev in previous work.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Weed management has historically relied on a combination of crop rotation, mechanical weed control, and the use of a variety of herbicides [1]. 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 [2, 3]. Current crop losses due to herbicide resistant weeds are \$4 to \$6 billion per year, and may climb to \$100 billion per year when chemical control is lost [4]. 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 [5]. 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 one gene away from total loss of chemical control [6]. Transgenic crop cultivars with "stacked" resistance genes for multiple herbicides exacerbate resistance in fields where herbicide resistance genes are already present [7]. An alternative to chemical weeding is mechanical weeding, which uses a physical system, composed of farm equipment or autonomous vehicles.

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 difficult and impractical at scale. Mechanized inter-row cultivation has disadvantages, such as soil compaction due to use of heavy machinery, and an inability to work after the crop canopy closes. Due to crop canopy growth, no current mechanical weed control method is effective within the crop [8]. Our work suggests that a team of collaborative low-cost and lightweight mechanical weeding robots (termed here as *agbots* shown in Figure 1.1) may be utilized to control herbicide-resistant weeds. The team of *agbots* targets weeds within and between crop rows, as opposed to tractors, combines, and planters, which cannot be used after the crop canopy grows. The *agbots* are ideal for working in dense fields, since they are small enough to drive over plants without damaging them, and do not compact the soil as large machinery would.

Termination of weed seedlings within several days after they emerge is critical to preventing crop yield losses in corn and soybeans [9]. For many crops, weeding may be done under a canopy, and therefore under conditions of partial environmental information. This robotic system must plan robustly, operating in dynamic environments, utilizing limited information to efficiently complete the task under time constraints. The goal of this work is to present a comprehensive study of the feasibility of a coordinated robotic weeding approach in realistic field environments. Our aim is to leverage strategies for multi-robot coordination to create a scalable weeding solution. We want to ensure 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 [10]. 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, 12, 13, 14, 15] 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, chosen to be the total of the maximum height of weeds in every 0.8 m^2 region of the field. This ensures the system eliminates weeds before they grow too large for the mechanical weeding process to deal with. By optimizing over this reward metric, we ensure the field can be weeded completely, preventing weeds from growing large enough to seed.

1.1 Formulation of Weeding Problem

The goal of this work is 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.

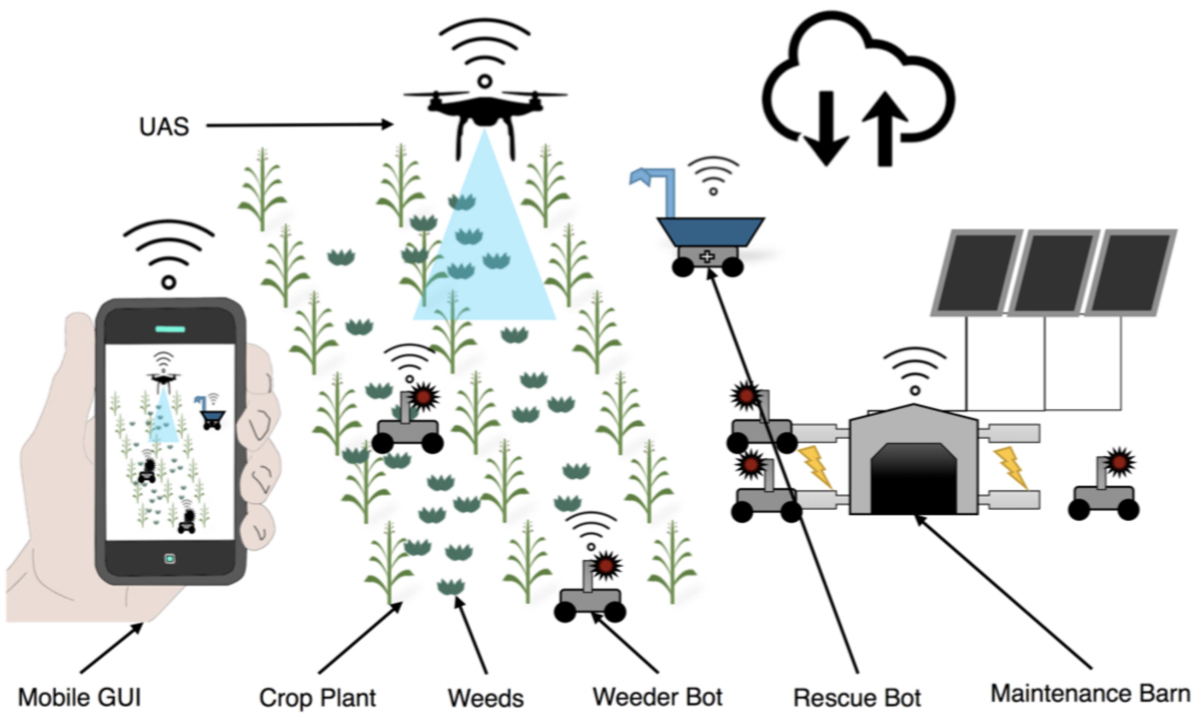

Figure 1.1: Our solution for robotic mechanical weed control is a dynamically configured team of weeder bots, drones, and automated maintenance barns, which provide persistent autonomous weed control, leveraging collaboration as well as local and remote data sources.

1.2 Mechanical Weeding

As shown in Figures 1.2, and 1.3, the number of herbicide resistant weed species has risen from 83 species in 1990 to 553 species in 2016, as detailed in the International Survey of Herbicide Resistant Weeds [2].

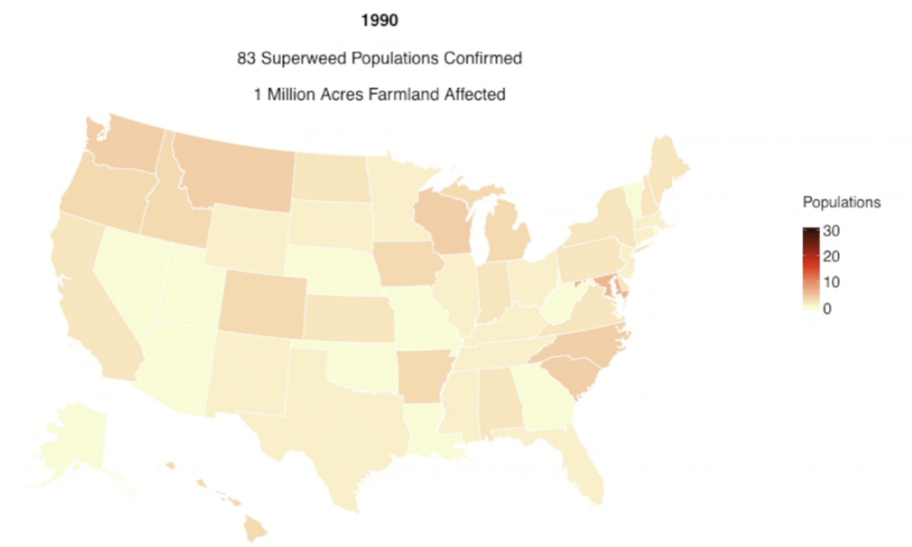

Figure 1.2: In 1990, there were 83 confirmed species of herbicide resistant weeds in 1 million acres of farmland across the continental United States.

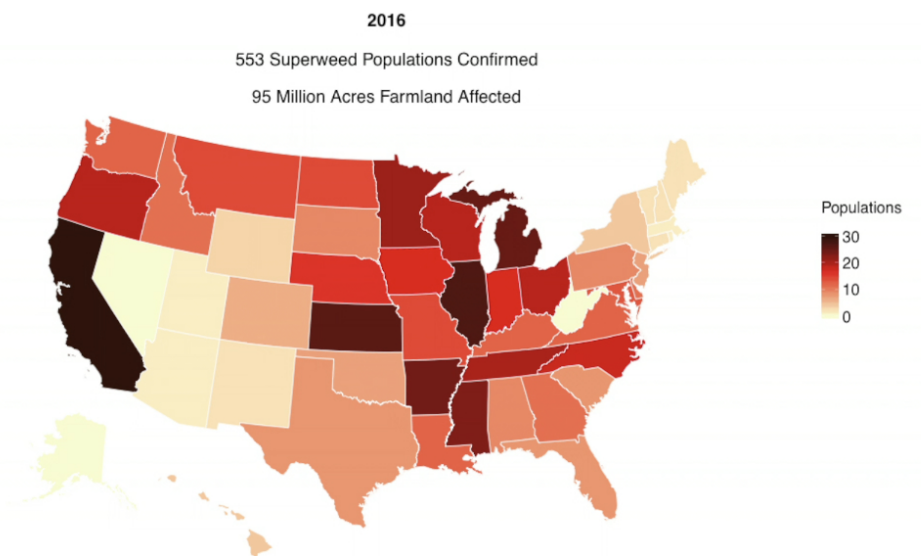

Figure 1.3: By 2016, there were 553 confirmed species of herbicide resistant weeds in 95 million acres of farmland across the continental United States.

Tractor-pulled mechanical weeding solutions [16], such as that shown in Figure 1.4, have long been employed by farmers to combat herbicide resistant weeds. However, these devices are not viable for weeding between rows, once the crop canopy has closed. To address this issue, several companies, such as TerraSentia [17] and Naio-Technologies [18], have developed small agricultural robots for autonomous weeding, shown in Figures 1.5 and 1.6 respectively. However, in order for robots like these to be employed at scale in the agricultural industry, multi-agent planning strategies applicable to real field environments must be developed.

Despite the advances in weeding robots, no comprehensive study of the feasibility of coordinated weeding has yet been presented. This work attempts to demonstrate that a scale-neutral approach to coordinated weeding, realized via collaboration between many small agricultural robots, will be feasible in a range of field environments. To do this, we develop an algorithm which finds a robust heuristic policy for weeding, based on all available information collected from the agents in the field.

Figure 1.4: Tractor-Pulled Mechanical Weeder: Note that this equipment is not viable for weeding in between the rows once the crop canopy closes [16].

Figure 1.5: TerraSentia Agricultural Robot: This robot performs autonomous weeding in between crop rows [17].

Figure 1.6: Naio-Technologies Agricultural Robot: An alternative to the TerraSentia [18].

1.3 Relevant Literature: Coordinated Robotic Planning

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 [10], 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 [10], 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 our 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 differ from what 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 shows 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. This factored model improves computational efficiency and simplifies the planning algorithm. The factored approach allows each individual agent to plan a policy maximizing its individual reward, in order to maximize the reward for the whole system, which is the total additive reward for all the agents.

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 [19, 20, 21], leveraging both spatial constraints and predictive information to perform optimization. In [22], a formal taxonomy of 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, as all agents cooperate to weed the field together in order to complete the weeding task more efficiently. This approach allows agents to adapt to changes in the environment, working together on regions with 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 once the task has been completed, allowing them to focus on navigation and plant recognition while in 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 [22], the problem of time-extended on-line assignment, in which multiple robots 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 [22] for the time-extended on-line assignment problem, which initially assigns each robot to the most suitable task, and then assigns robots to the most suitable task from the pool as they become available.

1.4 Weed-Growth Models

Past work has presented detailed analysis of weed growth models, in which measurements of seed bank density for various species of weeds were conducted [23, 24, 25, 26]. In the weed growth model used here, the rate of seedling emergence agrees with that found in the above research. A Poisson process is utilized to model the temporal evolution of emergence events, and assumes spatial uniformity in the seedbank density. These assumptions are reasonable over the short time scales in which robots may fully weed a field.

Other work in the field of crop science [27, 28] has shown that the spatial variation in the seed bank density for some species of weeds may be modeled via the Gini Coefficient of Concentration (GCC). Furthermore, in [29, 30], the relationship between seed bank emergence patterns and environmental conditions such as temperature and moisture is examined. Future work will incorporate these models into the weed growth model utilized in our simulation environment.

All of this work demonstrates conclusively that weed growth is a complex, stochastic, spatiotemporal phenomenon, dependent on a plethora of environmental factors, which will vary from field to field, depending on the time of year and the region where the crops are grown. It is clear that multi-agent planning under conditions of partial environmental information will be difficult when the reward metric is dependent on weed growth. This work develops a robust heuristic policy for coordinated weeding, which is shown to perform well in a variety of field conditions, and is agnostic to the model used for weed growth, relying only on real-time observations from each agent.

However, future algorithms could incorporate a predictive model for weed growth into the planner. These algorithms could utilize past observations of weeds in the field, sensor data from multiple agents, and predictive information gleaned from observations of fields with similar characteristics, to accurately model weed growth. As shown in Figure 1.7, this system would utilize this predictive information in the planner, yielding higher performance when deployed at scale. Our algorithm could easily be modified to utilize the predicted number of weeds in each row, instead of the observed number, allowing this future work to complement our current work.

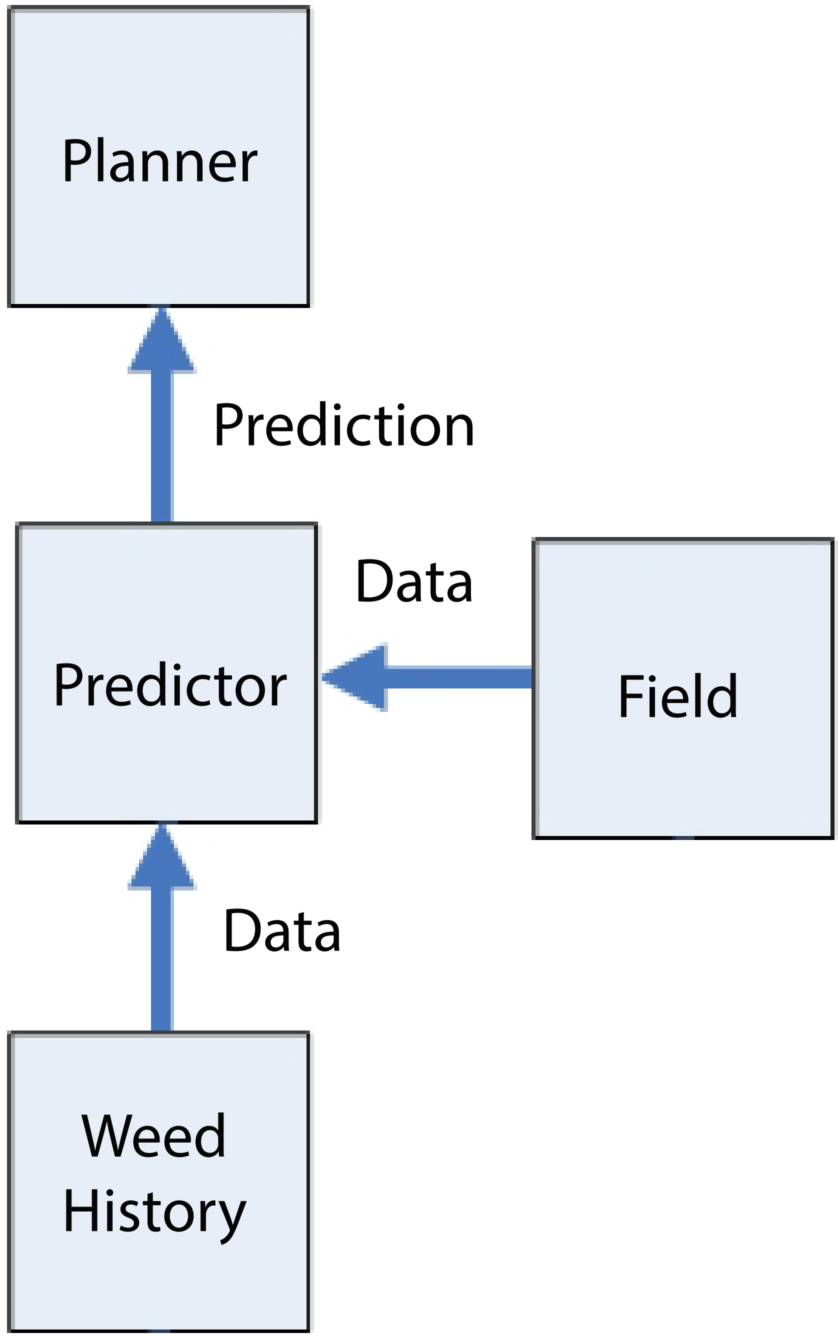

Figure 1.7: Planner Utilizing a Predictive Policy

1.5 Contributions

This thesis 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 utilize 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 Figure 2.1.), 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 m wide, totaling 4047 m^2, 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.

Summary

- Demonstrated that coordination improved weeding performance.
- Showed the feasibility of our algorithm in a variety of field conditions.
- Created a realistic simulation framework for coordinated multi-robot weeding.

1.6 Outline

Chapter 2 presents the methods utilized to solve the problem. Chapter 3 presents an interpretation of the results of the experiments conducted. Finally, Chapters 4 and 4.2 present conclusions and an outline of further work.

CHAPTER 2

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, and the value function utilized for optimization. We next explain the dynamic programming algorithm used to solve the optimization problem. We 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.

2.1 Weed Growth Model

The weed growth model utilized in this thesis 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 90 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 [23, 24, 30, 25, 26], 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 density, selected to be up to 100 seeds per 0.8 m^2, is realistic for some species of plants. However, we avoid limiting the thesis to a specific species of weed as this is region-specific.

$$\lambda_t(x,y,t) = \frac{0.9 \cdot \Delta t \cdot S(x,y,t)}{T_{\text{total}}}, \quad \lambda_0 = \frac{0.9 \cdot d_0 \cdot S_0}{T_{\text{total}}} \tag{2.1}$$

$$N_{\text{emerge}}(x,y,t) = \text{Poi}\left(\lambda_t(x,y,t)\right) \tag{2.2}$$

$$S(x,y,t) = S_0 - \sum_{t=t_0}^{t_{\text{current}}} N_{\text{emerge}}(x,y,t) \tag{2.3}$$

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.

$$\zeta(x,y,t) = \sum_{t=t_{\text{last weeded}}}^{t_{\text{current}}} N_{\text{emerge}}(x,y,t) \tag{2.4}$$

$$\delta(x,y,t) = \left(\frac{t_{\text{current}} - t_{\text{last weeded}}}{60 \cdot 60 \cdot 24}\right)\left(\Gamma \frac{\text{inch}}{\text{day}}\right) \tag{2.5}$$

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, $\delta(x,y,t)$.

$$R_W(x,y,t) = \delta(x,y,t) \tag{2.6}$$

2.2 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 m), $R_W(x,y,t)$ is the reward for each location (x,y), and v_i is the agent velocity.

The environmental state, S, depends on the x and y positions of each agent in I. The action, a, is defined 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.

$$S \equiv \{1,...,N_{\text{dim}}\} \times \{1,...,N_{\text{dim}}\}, \quad I \equiv \left\{1,...,N_{\text{agents}}\right\} \tag{2.7}$$

$$\left(x_i(t), y_i(t)\right) \in S \quad \forall i \in I, \quad a_i(t) = \left\{x_i(t+1), y_i(t+1)\right\} \in A \equiv S \tag{2.8}$$

The order of the state space is $85 \times 85 \times N_{\text{agents}}$. The order of the state-action space is then $\left(85 \times 85 \times N_{agents}\right)^2$. In our problem, the policy must be computed in real-time, and the fact that agents operate in a dynamic and uncertain environment makes the problem non-stationary. Therefore, a reinforcement learning approach which planned over the whole state space would be too computationally intensive for on-line learning. It would certainly fail to converge by the time an agent had completed a row, and new information invalidated the previous optimal policy. Thus, we attempt to reduce the dimensionality of the problem, and utilize an optimization-based approach over a one-step planning horizon, instead of using reinforcement learning to attempt to compute the policy over an extended horizon.

$$|S \times A| = \left(85 \times 85 \times N_{agents}\right)^2 \tag{2.9}$$

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 $85 \times N_{\text{agents}}$.

$$S \equiv \{1, ..., N_{\text{dim}}\}, \quad I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{2.10}$$

$$x_i(t) \in S \qquad \forall i \in I, \quad a_i(t) = x_i(t+1) \in A \equiv S \tag{2.11}$$

2.3 Reward Model

The reward is composed of the reward for each square of weeds in the field, $R_W(x, y, t)$.

$$S \equiv \{1, ..., N_{\text{dim}}\} \times \{1, ..., N_{\text{dim}}\}, \quad I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{2.12}$$

$$R_W(x, y, t) \quad \forall (x, y) \in S, \quad \forall i \in I \tag{2.13}$$

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.

$$R_i\left(a_i(t)\right) = \sum_{y=1}^{N_{\text{dim}}} R_W\left(a_i(t), y, t\right) \quad \forall a_i(t) \in A \equiv \{1, ..., N_{\text{dim}}\} \tag{2.14}$$

2.4 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 [12]. To expedite computation, we use a factored approach [31], 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, $T_{\text{to row}}$, the time it takes to move down it, $T_{\text{down row}}$, and the time it takes to weed all the squares in the row, $T_{\text{weed row}}$.

$$T_i\left(x_i\left(t\right), a_i\left(t\right)\right) = T_{\text{to row}} + T_{\text{down row}} + T_{\text{weed row}} \tag{2.15}$$

$$T_{\text{to row}} = \frac{\left(a_i\left(t\right) - x_i\left(t\right)\right)}{v_i} \tag{2.16}$$

$$T_{\text{down row}} = \frac{Y_{\text{dim}}}{v_i} \tag{2.17}$$

$$T_{\text{weed row}} = T_{\text{kill}} \sum_{y=0(t)}^{N_{\text{dim}}} \delta\left(x_i\left(t\right), y\left(t\right)\right) \tag{2.18}$$

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 at rate γ.

$$V\left(t\right) = \sum_{i \in I} \gamma^{T_{\text{i}}\left(x_i(t), a_i(t)\right)} R_{\text{i}}\left(a_i\left(t\right)\right) \tag{2.19}$$

2.5 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 assign agents asynchronously to the row with the highest value when they query the planner after completing a row.

$$V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \gamma^{T_i\left(a_i(t)\right)} R_i\left(a_i\left(t\right)\right) \tag{2.20}$$

$$a_i\left(t\right) = \arg\max_{a_i(t)} V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \tag{2.21}$$

2.6 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, $N_{\text{rows weeded}}$.

$$\bar{R} = \frac{\sum_{t=t_{\text{exp.}}}^{t_{\text{current}}} \sum_{i=0}^{N_{\text{agents}}} R_i\left(a_i\left(t\right)\right)}{N_{\text{rows weeded}}} \tag{2.22}$$

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.

$$I\left(a_i\left(t\right)\right) = \sum_{i=-r_{obs}}^{r_{obs}} \mathbb{I}_{\left\{\text{is explored}\left(x=a_i(t)+i\right)\right\}} \tag{2.23}$$

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 $I\left(a_i\left(t\right)\right)$.

$$\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \alpha\left(\gamma^{T_i\left(x_i(t),a_i(t)\right)}\bar{R}I\left(a_i\left(t\right)\right) - \bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)\right) \tag{2.24}$$

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

$$E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \tag{2.25}$$

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

$$\arg\max_{a_i(t)} E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \geq \arg\max_{a_i(t)} V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \Rightarrow a_i\left(t\right) = \arg\max_{a_i(t)} E_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) \tag{2.26}$$

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

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

end
end

Algorithm 1: Algorithm for Computing the Policy

2.7 Plan for Computational Experiments

We conduct seven experiments, each with 100 trials with varying initial parameters shown in Table 2.1. 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 2.1. 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.

Table 2.1: Initial Experimental Parameters: 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	$[0,\infty]$
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]

2.8 Simulation Environment

The simulation environment utilized for the above experiments, Weed World (shown in Figure 2.1), was developed to allow large scale simulations of coordinated weeding algorithms for multi-robot planning in uncertain environments. This environment incorporates the weed growth model and planning algorithm detailed above, as well as a framework for multi-agent collaboration, which enables a scalable amount of agents to easily share information. This framework also allows the amount of environmental information available to agents to be changed, in order to determine the effect on the algorithm performance.

Weed World allows Monte Carlo runs to be conducted, in order to test the planning algorithm's performance in fields with varying characteristics, and with agents with different capabilities, which have access to a varying amount of information about the environment. Through these Monte Carlo runs, the conditions under which full weeding of a given field is feasible are determined, informing the design considerations utilized in the development of real weeding robots, and thus expediting the design process.

The goal of developing the Weed World simulation environment was to enable future researchers to easily design and prototype algorithms for coordinated multi-agent weeding in uncertain environments. The stochastic and non-stationary nature of the evolution of the reward, as governed by the growth of weeds in a dynamic environment, may make this domain of strong theoretical interest to the robotic planning community.

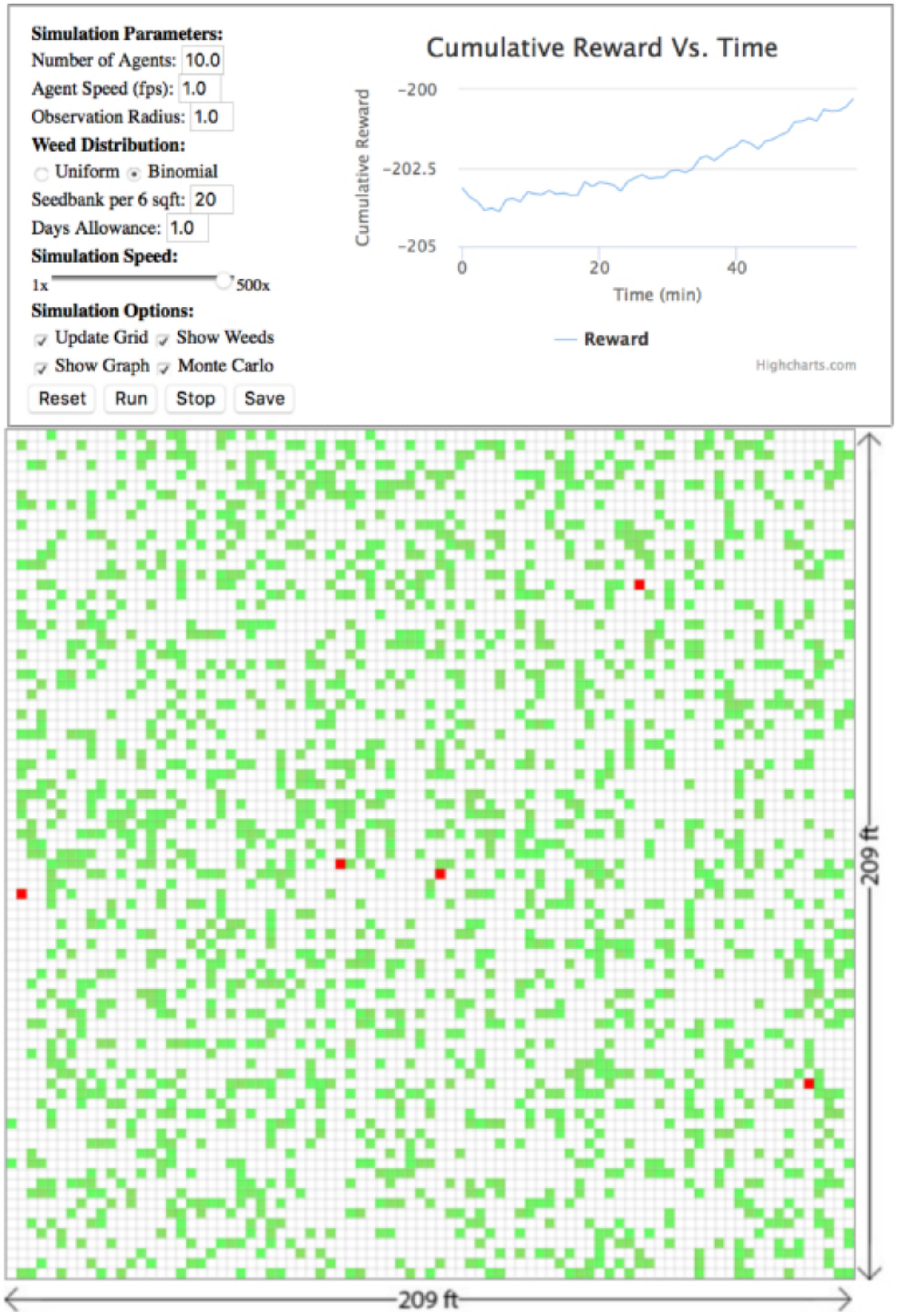

Figure 2.1: Simulation environment Weed World designed in JavaScript

CHAPTER 3

RESULTS

This chapter presents the results. In Experiments 1 - 3, we find our algorithm tracks the fully observed case under nominal field conditions, when the approach without shared information is unable to do so. In Experiments 4 - 5, we find that the algorithm performance is unaffected by the transit speed of the agents. Finally, in Experiments 6 - 7, we find that, with enough agents, the algorithm is able to successfully weed a range of fields.

3.1 Experiments 1 - 3

Comparison of The Coordinated and Uncoordinated Cases for Planning with Partial Information with the Fully Observed Case

As seen in Figure 3.1, partial environmental information, $r_{obs} = 1$, 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. For these experiments, nominal values for the seedbank density (20 seeds per m^2), days allowance (3 days), number of agents (5), and agent speed (1 m per second), were used.

3.2 Experiment 4

Performance for Varying Seedbank Density and Transit Speed of Agents

As seen in Figure 3.2, 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. *When seed bank density grows large, every square will eventually become infested, and the speed of transit of the robot will not affect the weeding performance.*

3.3 Experiment 5

Performance for Varying Days Allowance and Transit Speed of Agents

As seen in Figure 3.3, 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. *When the days allowance becomes large enough, the field is initially fully infested and the transit speed of the robot is unimportant.*

3.4 Experiment 6

Performance for Varying Seedbank Density and the Number of Agents

As seen in Figure 3.4, as the seed bank density increases, more 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. *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.*

3.5 Experiment 7

Performance for Varying Days Allowance and the Number of Agents

As seen in Figure 3.5, 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. *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.*

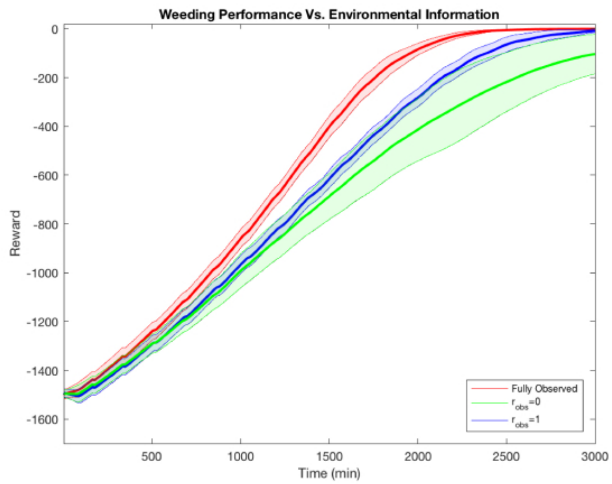

Figure 3.1: **Weeding Performance vs. Environmental Information:** We plot the weeding performance over time for the case of full environmental information, partial environmental information, $r_{obs} = 1$, and zero environmental information, $r_{obs} = 0$. *We see that the case of $r_{obs} = 1$ is able 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.*

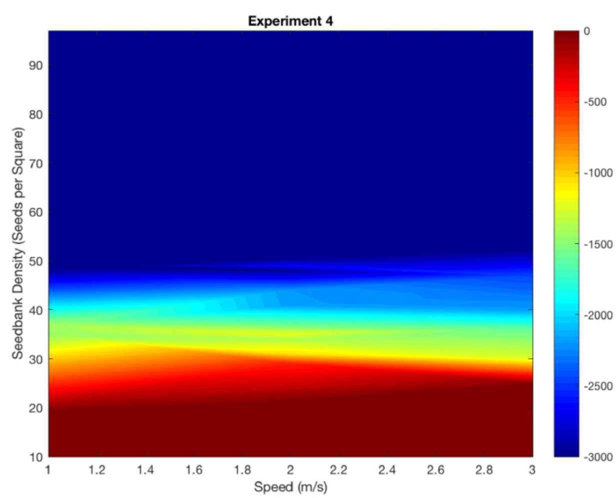

Figure 3.2: **Speed vs. Seed Bank Density:** The heat map of the terminal reward for 100 trials with varying agent speed and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. *Agent speed is not strongly correlated with the initial seed bank 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.*

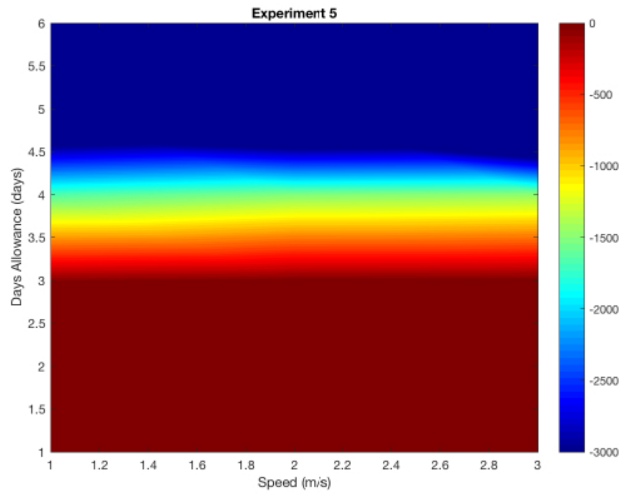

Figure 3.3: **Speed vs. Days Allowance:** The heat map of the terminal reward for 100 trials with varying agent speed and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero 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.*

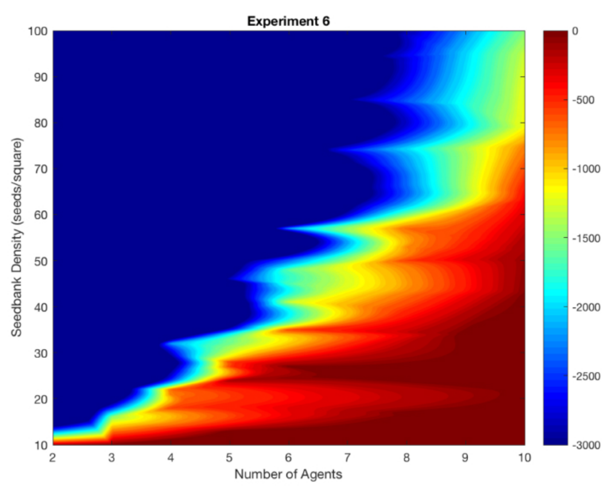

Figure 3.4: **Number of Agents vs. Seed Bank Density:** The heat map of the terminal reward for 100 trials with varying numbers of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents 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 varying initial seed bank densities with a large enough number of robots.*

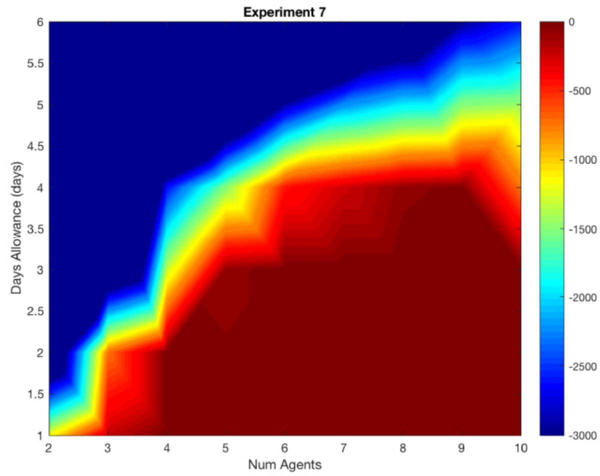

Figure 3.5: **Number of Agents vs. Days Allowance:** The heat map of the terminal reward for 100 trials with varying numbers of agents and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents 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.*

CHAPTER 4

CONCLUSIONS AND FURTHER WORK

4.1 Conclusions

This research demonstrates that a scale-neutral approach to coordinated multi-agent weeding in uncertain environments, utilizing a varying number of robots for different agricultural applications, can 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 of agents are not only able to fully weed the field when smaller teams cannot do 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 is in fact essential, and will be a central part of mechanical weeding solutions to the weeding crisis.

4.2 Further Work

Our estimates for the range of seed bank densities hold for several species of plants. However, we will extend our work to 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 various environments. Future work will include an updated model for the spatial variation within the seed bank. We will also complete further analysis of the sensitivity of algorithm performance to variations in the initial seed bank density, as well as to the initial days of allowed weed growth, for both the case of shared information between agents, and the case in which agents do not utilize a shared environmental model.

In the long-term, we will explore ways in which the emergence patterns of weeds may be affected by environmental conditions such as temperature and moisture. Furthermore, we will conduct experiments with the robots currently utilized in real field experiments to refine our parameters for weed killing time, robot transit speed, and weed recognition in neighboring rows.

4.2.1 Spatial Variation of The Seed Bank

Past work [27, 28] has shown that the spatial variation in the seed bank density for some species of weeds may be modeled via the Gini Coefficient of Concentration (GCC). In future work, we will update our model for seed bank density, previously assumed to be uniform, in order to align with the degree of dispersion predicted by measurements for the GCC given for weed species in [27, 28].

4.2.2 Sensitivity of Algorithm Performance

In this work, we have successfully determined relationships between the seed bank density of the weeds, along with the initial days of allowed weed growth, and the number of agents, along with their transit speed. Furthermore, we have characterized the performance improvement of the system which uses shared environmental information over the system which does not use a shared environmental model. In order to make these relationships more precise, after the weed growth model has been updated, we will conduct extended experiments to determine the required number of agents for various seed bank densities and days allowance, in order for the field to be weeded completely by the end of the experiment.

For the cases in which we have complete weeding, we will determine the sensitivity of the required number of agents to the seed bank density and days allowance. In order to more accurately measure performance improvement, we will conduct these experiments both when the agents operate via a shared environmental model, and when they operate independently. Due to the fact that in this work, the transit speed was found not to strongly affect algorithm performance, this parameter will not be considered in further experiments. In reality, the transit speed of the agents is determined by the physical characteristics of the robot. Future experiments will measure the average transit speed of robots currents used the the field.

4.2.3 Temporal Variation in Seed Bank Emergence

In [29, 30], the relationship between seed bank emergence patterns and environmental conditions such as temperature and moisture is examined. Our previous model for emergence, which modeled seed emergence as a Poisson process, is reasonable for many species of weeds, especially over the relatively short time scales required for robots to weed the entire field. However, we would like to run simulations over time scales of longer duration, and consideration of the effect of environmental conditions on seedbank emergence patterns will be beneficial towards this goal. We will therefore attempt to extend our model for weed growth to include a statistical model for temperature and moisture variation over the crop season, along with the relevant generating distributions for seed bank emergence given in [29, 30] for these environmental variations.

4.2.4 Field Experiments

Estimates for system parameters used in our algorithm were chosen based on the characteristics of the current weeding robots used in field experiments. In order to refine these estimates, we would like to conduct several trials to measure the average time taken to kill weeds, the average transit speed for robotic agents, and the average percentage of weeds in neighboring rows that may be measured given possible occlusion from the crops in the current row. These experiments will help us refine the performance estimates for our algorithm, so that they are more closely aligned with what may be seen in real field trials. Finally, once challenges in multi-agent communication, charging, and weed killing have been addressed in the current robotic weeding system, we hope to implement our algorithm on a team of real robots to more effectively measure its performance.

4.2.5 Predictive Planning Algorithms

Our algorithm currently utilizes only the observed values of weeds in known rows, alongside an estimate for weeds in unknown rows based on the average value of the total number of weeds in previously weeded rows. We will attempt to instead utilize the predicted value of weeds in unknown rows, based on a spatiotemporally evolving model formed from past observations of weeds in the field, and statistical models for the weed growth aligned with past work. This predictive information will allow our algorithm to perform better than when a simple estimate is used.

REFERENCES

[1] D. L. Shaner, "Lessons learned from the history of herbicide resistance," *Weed Science*, vol. 62, no. 2, pp. 427–431, 2014.

[2] I. Heap, "The international survey of herbicide resistant weeds," 2017. [Online]. Available: http://www.weedscience.org

[3] B. D. Maxwell, M. L. Roush, and S. R. Radosevich, "Predicting the evolution and dynamics of herbicide resistance in weed populations," *Weed Technology*, vol. 4, no. 1, pp. 2–13, 1990.

[4] M. Livingston, J. Fernandez-Cornejo, and G. B. Frisvold, "Economic returns to herbicide resistance management in the short and long run: The role of neighbor effects," *Weed Science*, vol. 64, no. sp1, pp. 595–608, 2016.

[5] M. V. Bagavathiannan and J. K. Norsworthy, "Multiple-herbicide resistance is widespread in roadside palmer amaranth populations," *PloS One*, vol. 11, no. 4, p. e0148748, 2016.

[6] C. M. Evans, "Characterization of a novel five-way-resistant population of waterhemp (amaranthus tuberculatus)," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2016.

[7] J. Gressel, A. J. Gassmann, and M. D. Owen, "How well will stacked transgenic pest/herbicide resistances delay pests from evolving resistance?" *Pest Management Science*, vol. 73, no. 1, pp. 22–34, 2017.

[8] C. L. Mohler, J. C. Frisch, and J. M. Pleasant, "Evaluation of mechanical weed management programs for corn (zea mays)," *Weed Technology*, vol. 11, no. 1, pp. 123–131, 1997.

[9] E. R. Page, D. Cerrudo, P. Westra, M. Loux, K. Smith, C. Foresman, H. Wright, and C. J. Swanton, "Why early season weed control is important in maize," *Weed Science*, vol. 60, no. 3, pp. 423–430, 2012.

[10] Y. U. Cao, A. S. Fukunaga, and A. Kahng, "Cooperative mobile robotics: Antecedents and directions," *Autonomous Robots*, vol. 4, no. 1, pp. 7–27, 1997.

[11] M. Liu, K. Sivakumar, S. Omidshafiei, C. Amato, and J. P. How, "Learning for multi-robot cooperation in partially observable stochastic environments with macro-actions," *Computing Research Repository*, 2017.

[12] M. J. Matarić, "Reinforcement learning in the multi-robot domain," *Autonomous Robots*, vol. 4, no. 1, pp. 73–83, 1997.

[13] T. 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.

[14] M. J. Mataric, "Learning in behavior-based multi-robot systems: Policies, models, and other agents," *Cognitive Systems Research*, vol. 2, no. 1, pp. 81–93, 2001.

[15] Y. Shoham, R. Powers, and T. Grenager, "Multi-agent reinforcement learning: a critical survey," *Web Manuscript*, 2003. [Online]. Available: http://jmvidal.cse.sc.edu/library/shoham03a.pdf

[16] R. A. University, "Mechanical weeding live," 2017. [Online]. Available: https://www.rau.ac.uk/about/news-and-events/events/mechanical-weeding-live

[17] TerraSentia, "Terrasentia earthsense, inc." 2017. [Online]. Available: https://www.earthsense.co/home/

[18] N. Technologies, "Autonomous weeding for agricultural robots - naio technologies," 2017. [Online]. Available: https://www.naio-technologies.com/en/

[19] J. Niko-Mora, "Stochastic scheduling," Updated version of article in *Encyclopedia of Optimization*, C. A. Floudas and P. M. Pardalos, eds. Kluwer, 2001., October 2005. [Online]. Available: http://halweb.uc3m.es/jnino/eng/pubs/ssche.pdf

[20] E. Frazzoli and F. 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.

[21] A. Munoz-Meléndez, P. Dasgupta, and W. Lenagh, "A stochastic queueing model for multi-robot task allocation," in *International Conference on Informatics in Control, Automation and Robotics*, 2012, pp. 256–261.

[22] B. P. Gerkey and M. J. Matarić, "A formal analysis and taxonomy of task allocation in multi-robot systems," *International Journal of Robotics Research*, vol. 23, no. 9, pp. 939–954, 2004.

[23] D. Nordby, R. Hartzler, and K. Bradley, "Biology and management of waterhemp," *Glyphosate, Weeds, and Crop Sciences, Purdue University Extension, publication GWC-13.12*, 2007.

[24] B. J. Schutte and A. S. Davis, "Do common waterhemp (amaranthus rudis) seedling emergence patterns meet criteria for herbicide resistance simulation modeling?" *Weed Technology*, vol. 28, no. 2, pp. 408–417, 2014.

[25] B. A. Sellers, R. J. Smeda, W. G. Johnson, J. A. Kendig, and M. R. Ellersieck, “Comparative growth of six amaranthus species in Missouri,” *Weed Science*, vol. 51, no. 3, pp. 329–333, 2003.

[26] M. J. Horak and T. M. Loughin, “Growth analysis of four amaranthus species,” *Weed Science*, vol. 48, no. 3, pp. 347–355, 2000.

[27] D. Mulugeta and D. E. Stoltenberg, “Seed bank characterization and emergence of a weed community in a moldboard plow system,” *Weed Science*, pp. 54–60, 1997.

[28] D. Mulugeta and C. M. Boerboom, “Seasonal abundance and spatial pattern of setaria faberi, chenopodium album, and abutilon theophrasti in reduced-tillage soybeans,” *Weed science*, pp. 95–106, 1999.

[29] A. S. Davis, S. Clay, J. Cardina, A. Dille, F. Forcella, J. Lindquist, and C. Sprague, “Seed burial physical environment explains departures from regional hydrothermal model of giant ragweed (ambrosia trifida) seedling emergence in us midwest,” *Weed science*, vol. 61, no. 3, pp. 415–421, 2013.

[30] R. Werle, L. D. Sandell, D. D. Buhler, R. G. Hartzler, and J. L. Lindquist, “Predicting emergence of 23 summer annual weed species,” *Weed Science*, 2014.

[31] C. Amato, G. Chowdhary, A. Geramifard, N. K. Ure, and M. J. Kochenderfer, “Decentralized control of partially observable Markov decision processes,” in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on.* IEEE, 2013, pp. 2398–2405.