Multi-Agent Planning for Coordinated Robotic Weed Killing

Wyatt McAllister, Denis Osipychev, Girish Chowdhary, Adam Davis

University of Illinois at Urbana-Champaign

March 28, 2018

Robotic Weeding

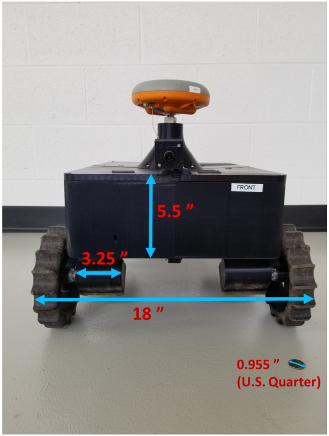

Figure 1: Robotic Platform

- Recent work on robotic weeding has focused both on the design of weeding robots and on the challenge of plant recognition [Bakker et al., 2010, Gai et al., 2015].

Robotic Weeding

Figure 2: Crop Canopy

- For many crops, including corn, weeding may be done under a canopy, and therefore with partial environmental information.
- Our goal is to design a system which allows agents to coordinate their weeding under varying amounts of environmental information, to optimize weeding efficiency.

Robotic Foraging

Figure 3: Foraging

- Foraging has long been considered a key problem in multi-agent robotics [Cao et al., 1997].
- In our case, the foraging problem is framed in terms of recognizing and killing weeds while moving through the field.

Past Work

Robot Search and Rescue

Figure 4: Search and Rescue

- Recent work on foraging in uncertain environments has solved a search and rescue problem with ground robots and UAVs [Liu et al., 2017].
- This work assumed the capability of the UAVs to gather information about possible victim locations.

Environmental Parameters

- N_{dim} is the number of squares in a row (85).
- Y_{dim} is the length of each row (209 feet).
- $R_W(x, y)$ is the reward for each weed at each location (x, y).
- v_i is the velocity of agent i.
- T_{kill} is the time to kill a square of weeds.
- S_0 is the initial seed bank density of each square.
- d_0 is the number of days allowed to elapse before robots start weeding.

Dynamic Weed Growth Model

- The number of weeds emerging, N_{weeds}, is equal to a poisson variable with mean, $\lambda\left(x_i\left(t\right), y_i\left(t\right)\right)$, so that 90 percent of the seed bank, $S\left(x_i\left(t\right), y_i\left(t\right)\right)$, emerges in two months [Nordby et al., 2007].

$$\lambda_0 = \frac{d_0 \cdot 0.9 \cdot S_0}{2 \text{ months}}, \quad \lambda\left(x_i\left(t\right), y_i\left(t\right)\right) = \frac{t \cdot 0.9 \cdot S\left(x_i\left(t\right), y_i\left(t\right)\right)}{2 \text{ months}} \tag{1}$$

$$N_{\text{weeds}}\left(x_i\left(t\right), y_i\left(t\right)\right) = \text{Poi}\left(\lambda\left(x_i\left(t\right), y_i\left(t\right)\right)\right) \tag{2}$$

$$S\left(x_i\left(t\right), y_i\left(t\right)\right) = S_0 - \sum_{t=t_{\text{lastweeded}}}^{t_{\text{current}}} N_{\text{weeds}}\left(x, y, t\right) \tag{3}$$

- The weed density at each square, $\zeta\left(x_i\left(t\right), y_i\left(t\right)\right)$, grows as seeds emerge. The maximum weed height at each square, $\delta(x_i\left(t\right), y_i\left(t\right))$, increases from an arbitrary height of one inch at a fixed rate Γ inches per day.

$$\zeta\left(x_i\left(t\right), y_i\left(t\right)\right) = \sum_{t=t_{\text{last weeded}}}^{t_{\text{current}}} N_{\text{weeds}}\left(x_i\left(t\right), y_i\left(t\right)\right) \tag{4}$$

$$\delta\left(x_i\left(t\right), y_i\left(t\right)\right) = 1 + \left(\frac{t_{\text{current}} - t_{\text{lastweeded}}}{60 \cdot 60 \cdot 24}\right)\left(\Gamma \frac{\text{inch}}{\text{day}}\right) \tag{5}$$

Problem Formulation

State, Action, Reward

- State: Current row for each agent.

$$x_i\left(t\right) \in S \quad \forall i \in I, \left\{ \begin{array}{c} S \equiv \{1, ..., N_{\mathrm{dim}}\} \\ I \equiv \{1, ..., N_{\mathrm{agents}}\} \end{array} \right\} \tag{6}$$

- Action: Target row for each agent.

$$a_i\left(t\right) = x_i\left(t+1\right) \in A \equiv S \tag{7}$$

- For each agent, the planned reward is the reward of the proposed row.

$$R_W\left(x_i\left(t\right), y_i\left(t\right)\right) \quad \forall \left(x_i\left(t\right), y_i\left(t\right)\right) \in S, \quad \forall i \in I \tag{8}$$

- The reward for the weed in each square, $R_W\left(x, y, t\right)$, is equal to the maximum height of weeds in the square, $\delta\left(x, y, t\right)$.

$$R_W\left(x_i\left(t\right), y_i\left(t\right)\right) = \delta\left(x, y, t\right) \tag{9}$$

Environmental Assumptions

Full Communication - Centralized Approach

- We assume full communication between all the agents about their current state, and action, and the total reward they have collected.

$$\{x_i(t), a_i(t), R_i(a_i(t))\} \Rightarrow \text{Known} \quad \forall \mathrm{i} \tag{10}$$

Single Agent Selection

- In our environment, we assume a field where robots can only cross rows at the edges of the field, and two robots cannot move side by side. Therefore, it is necessary to select one agent per row.

$$a_i(t) : a_i(t) \neq a_j(t) \quad \forall i \neq j \tag{11}$$

Homogeneous Agents

- In these experiments, all agents are homogeneous, meaning they have identical reward for the same rows under identical initial conditions.

$$x_i(t) = x_j(t) \Rightarrow R_i(a_i(t)) = R_j(a_j(t)) \quad \forall (i, j) \in I \tag{12}$$

Dynamic Programming (DP) Approach

- We want to maximize the overall value function, which is the sum over all agents of the planned reward, $R_i(a_i(t))$ time discounted by the planed operation time for the proposed row $T_{\mathrm{i}}(x_i(t), a_i(t))$.

$$V(t) = \sum_{i \in I} \gamma^{T_{\mathrm{i}}(x_i(t), a_i(t))} R_{\mathrm{i}}(a_i(t)) \tag{13}$$

- For the dynamic programming (DP) approach, we plan simultaneously across all the agents, evaluating the expected return for a transition from the agent's current state to any other new state.

$$V_t^i(x_i(t), a_i(t)) = \gamma^{T_{\mathrm{i}}(x_i(t), a_i(t))} R_i(a_i(t)) \tag{14}$$

- We send each agent to the row with maximum value.

$$a_i(t) = \underset{a_i(t)}{\arg\max}\, V_t^i(x_i(t), a_i(t)) \qquad \forall i \in I \tag{15}$$

Explore - Exploit Dilemma

Figure 5: Explore - Exploit Dilemma

- The naive approach to information gathering is to simply visit the next available adjacent unexplored row.
- We would like to utilize an algorithm which targets information gathering to ensure the largest increase in the total explored space.

Explore - Exploit Dilemma

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

$$\bar{R} = \frac{\sum_{t=t_{\textbf{exp.}}}^{t_{\textbf{current}}} \sum_{i=0}^{N_{\textbf{agents}}} R_i\left(a_i\left(t\right)\right)}{N_{\text{rows weeded}}} \tag{16}$$

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

$$I\left(a_i\left(t\right)\right) = \sum_{i=-r_{obs}}^{r_{obs}} \mathbb{I}_{\{\text{is explored}(x=a_i(t)+i)\}} \tag{17}$$

- We compute the estimated value function for an unexplored row, $V_t^i\left(x_i\left(t\right), a_i\left(t\right)\right)$, as the value function with $\bar{R}$ times $I\left(a_i\left(t\right)\right)$.

$$\bar{V}_t^i\left(x_i\left(t\right), a_i\left(t\right)\right) = \gamma^{T_i(x_i(t), a_i(t))} \bar{R} I\left(a_i\left(t\right)\right) \tag{18}$$

Explore - Exploit Dilemma

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

$$E_t^i(x_i(t), a_i(t)) = \bar{V}_t^i(x_i(t), a_i(t)) \tag{19}$$

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

$$\begin{aligned} \arg\max_{a_i(t)} E_t^i(x_i(t), a_i(t)) &\geq \arg\max_{a_i(t)} V_t^i(x_i(t), a_i(t)) \\ \Rightarrow a_i(t) &= \arg\max_{a_i(t)} E_t^i(x_i(t), a_i(t)) \end{aligned} \tag{20}$$

- If no rows have been explored, then visit the next adjacent unexplored row.

JavaScript Weed World

- Implemented the Weed World Environment in collaboration with Denis Osipychev, as a grid world of 85 rows of 2.5 foot squares, totaling 4400 square feet or one acre.
- Included dynamic weed-growth model, real-time visualization, and portable visualization for on-line deployment.

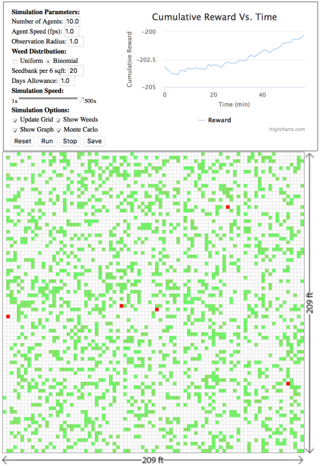

Figure 6: Simulation Environment

Experiment Outline

- We will conduct several experiments, each with 100 episodes of length 4 days.
- We first run the algorithm in the case of full environmental information, with observation radius, r_{obs}, of 0, and then observation radius, r_{obs}, of 1.
- We do Monte Carlo runs comparing the number of agents, N_{agent}, and their velocity, v_{agent}, with the days allowance, d_0, and initial seed bank density, S_0.

Table 1: Experiment Table

Exp.	1	2	3	4	5	6	7
r_{obs}	∞	0	1	1	1	1	1
N_{agent}	5	5	5	5	5	X	X
v_{agent}	1	1	1	X	X	1	1
d_0	1	1	1	1	X	1	X
S_0	10	10	10	X	10	X	10

- The above parameters take the following ranges in the Monte Carlo runs.

Table 2: Ranges of Values for Monte Carlo

r_{obs}	N_{agent}	v_{agent}	d_0	S_0
$[0,\infty]$	[3,10]	[1,3]	[1,4]	[10,100]

Full Vs. Partial Environmental Information

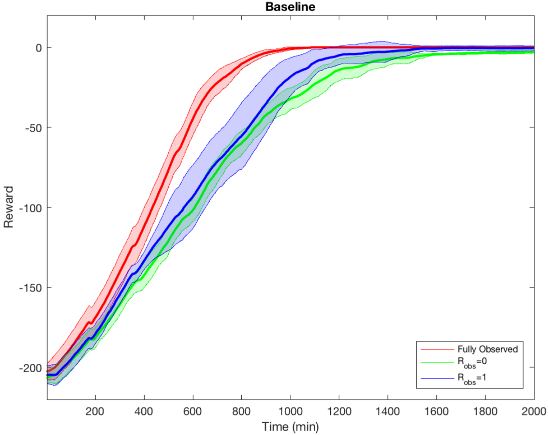

Figure 7: Full vs. Partial Environmental Information

- We see that performance drop for case of partial environmental information. Convergence rate is improved for the case of $r_{obs} = 1$.

Future Work

- Complete experiments 4 - 7 to determine design heuristics.
- Explore deep Q learning approach for planning over extended horizons.
- Incorporate estimation for the weed growth model into planner.

References I

Bakker, T., Bontsema, J., Müller, J., et al. (2010).
Systematic design of an autonomous platform for robotic weeding.
Journal of Terramechanics, 47(2):63–73.

Cao, Y. U., Fukunaga, A. S., and Kahng, A. (1997).
Cooperative mobile robotics: Antecedents and directions.
Autonomous robots, 4(1):7–27.

Gai, J., Tang, L., and Steward, B. (2015).
Plant recognition through the fusion of 2d and 3d images for robotic weeding.
In *2015 ASABE Annual International Meeting*, page 1. American Society of Agricultural and Biological Engineers.

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

Nordby, D., Hartzler, R., and Bradley, K. (2007).
Biology and management of waterhemp.