ABSTRACT

A classic example of multiagent coordination in a shared environment involves the use of pheromone deposits as a communication mechanism. Due to physical limitations in deploying actual pheromones, we propose a sparse representation of the pheromones using movable beacons. There is no communication between the beacons to propagate pheromones; instead, robots make movement and update decisions based entirely on local pheromone values. Robots deploy the beacons throughout the environment, and subsequently move them and update them using a variation of value iteration. Simulation results show that our approach is effective at finding good trails, locally improving them, and adapting to dynamic changes in the environments.

Categories and Subject Descriptors
I.2 [Distributed Artificial Intelligence]: Multiagent Systems

General Terms
Algorithms

Keywords
Beacon, Ant, Pheromone

1. INTRODUCTION

One challenge in swarm robotics is performing effective communication. Broadcast communication may be unattractive due to environmental factors, limited range and power, lack of global communications infrastructure, or overly congested channels. Instead, swarm robotics research has often focused on local interaction and forms of indirect communication such as pheromone deposits. In this paper we show a formal approach to pheromone deposit communication via beacons deployed in the environment by the robots. (We will use the terms “robot” and “ant” interchangeably.)

In previous work [16] we presented a technique to enable a large swarm of ant-like agents to perform foraging and other trail-following tasks by depositing, updating, and following pheromones in the environment. Unlike much previous work in the multiagent pheromone-based foraging literature, our approach cast pheromones as state utility values, and applied formal utility update equations based loosely on value iteration and temporal differencing. The work also employed multiple simultaneous pheromones to direct various aspects of the task at hand. Using these equations, we were able to demonstrate collaborative foraging using two pheromones (one which ultimately defined a gradient to food, and one back to the ants’ nest). We went further than this, demonstrating trail optimization, discovery of self-intersecting tours involving many waypoints (and many pheromones), and adaptation to dynamic changes in the environment.

That work was motivated in part by our research in swarm robotics: we sought a formal language and substrate with which swarms of agents might perform a range of cooperative tasks. The approach was also of value to swarm robotics because it did not require odometry or localization, and had a light computational load. However, the agents lived in a 2d discretized grid-world environment, with pheromones stored in the grid of the environment itself.

The present paper represents a significant advance over this early work, adapting the pheromone idea to a more realistic robot scenario through the use of deployable and modifiable beacons as a sparse representation of the pheromone environment. The ant robots deploy the beacons, and subsequently follow, update, move, and remove them. From the perspective of the robots, the beacons are a graph of states with utility values, just as the earlier grid-world was a denser graph of the same. Each beacon stores one value per pheromone employed by the swarm. Beacons only represent local information, and do not communicate with one another to spread pheromones. Rather, after deploying the beacons, robots update pheromone values using similar equations as [16], and likewise the robots make routing decisions based on the pheromones in nearby beacons. After discovering a trail, the robots may additionally move the beacons in order to optimize it. The work presented uses robots in simulation, but its approach is specifically designed to be deployable to actual robots in our laboratory. Our purpose here is to extend the approach taken in [16] to environments with more realistic assumptions, taking a significant step towards deploying to physical robots in our laboratory.

We begin the paper with a discussion of previous work in pheromone-based multiagent interaction, marker-based robotic navigation and optimization, and other related topics. We then introduce and discuss the proposed model, followed by experimental work with the model.

2. PREVIOUS WORK

Indirect communication through pheromone deposits is an example of stigmergy, a term coined by Pierre-Paul Grassé in the 1950s to describe a mechanism in which colonies of a certain genus of termites collaborate to build a nest [4, 7]. Many examples of stigmergy appear in nature, including the foraging behavior of ants.
Ants leave pheromone deposits as they move in the environment, and their navigation is stimulated through local observations of pheromone strength and gradient. The global ant foraging behavior emerges through these simple local pheromone interactions. Other familiar kinds of indirect communication through the environment include leaving footsteps in the snow; or leaving trails of pebbles or breadcrumbs along a journey in order to help find a way home.

The collective behavior of ants that emerges from communication through pheromones has been widely studied in both artificial life and in robotics. Beyond providing a robust, decentralized, and distributed means of communication, pheromone-based swarms have also shown the ability to optimize trails, as observed in [6]. The networks of paths constructed by ants have been compared to minimal spanning trees [18] which may emerge from the rapid accumulation of pheromone strength across shorter paths [2, 3]. By incorporating stochastic exploration and evaporation of old pheromones, swarms can be shown to adapt to dynamic situations such as changing goals or obstacles [12, 20]. We have demonstrated several of our own examples in [16].

Some pheromone-based reinforcement learning algorithms have been proposed for foraging problem domains. Most of these involve agents that use a fixed pheromone depositing procedure, and incorporate the existing pheromone values into their action selection and update mechanisms [8, 10, 11].

Several works have explored agent behaviors that take advantage of larger vocabularies of pheromones. For example, multiple-pheromone models have been used to establish separate gradients for different tasks [5, 16, 25], rather than relying on arbitrary a priori mechanisms to augment a one-pheromone model. Vaughan et al. have proposed an alternative involving pheromones that additionally indicate direction. Parunak et al. [17, 18] have exploited a variety of techniques such as using pheromones with different semantics, pheromones with different dynamics (e.g. different rates of evaporation and propagation) and using history (e.g. weighing pheromones more strongly in the direction of motion).

Several of the ant-inspired methods described so far rely on the ability of agents to modify the environment. While there has been some work on a robotic mechanism to deposit and sense chemical trails [21], this could be impractical in many real scenarios. One common approach is to rely on existing communications mechanisms to share internal pheromone models. For example, in [22, 23, 24], Vaughan et al. simulated stigmergic communication by making it possible for agents to share trails of waypoints over a wireless network. Similarly, Payton et al. experimented with “virtual pheromones” based on propagating discrete messages between mobile robots with short-range omni-directional infrared communications [19]. O’Hara et al. introduced the GNATs, a pre-deployed line-of-sight sensor network to support robotic navigation tasks [15]. They demonstrated successful distributed path planning in this infrastructure using variants of the wavefront planning algorithm [13] and the distributed Bellman-Ford algorithm [14]. Barth et al. proposed a dynamic programming approach to swarm navigation based on deployable, but immobile relay markers [1]. This approach also relies on communication between the markers in order to establish and update the pheromone values.

Ziparo et al. used deployable, non-mobile RFID tags to help path planning and exploration with large groups of robots [26].

It is important to note that the goal of our work is not the analysis of the network structure created by the topology of the beacons in the environment, as this has been well covered by other researchers. Instead, we focus on swarm robot behaviors which may successfully employ a collective model in the face of severe communications constraints (in this case, for the foraging task).

3. MODEL

Our foraging world is a bounded two-dimensional continuous environment. A nest is located somewhere in the environment, holding some number of robots (we will refer to them simply as ants). The environment also holds some N food source locations (in our experiments, we set N = 1). The environment may also contain obstacles through which the ants may not travel.

The ants’ task is to bring as much food back to the nest as possible within a given time frame. This consists of several subtasks. First, the ants must find a remote food location, and establish a trail there. Second, the ants must repeatedly follow the trail to the food location, harvest some food, ferry it back along the trail to the nest, and deposit it there. Third, the ants should optimize the trail, globally adopting new routes when they are found to be shorter, or attempting to straighten (and shorten) the current trail locally.

To assist them in this task, the ants deposit three pheromones and read them at various beacons the ants have deployed and organized throughout the environment. Each beacon represents the state of pheromones in the environment at that location. The three pheromones are:

- The foraging pheromone, used to build a gradient to food.
- The ferrying pheromone, used to build a gradient to the nest.
- The wandering pheromone, used to indicate the how often a state has been visited.

Beacons and ants have the same communications range. Beacons within the range of one another, and not occluded by an obstacle, are defined as neighbors. Each ant will associate itself with the nearest non-occluded beacon within its range, if any, and this beacon is called the ant’s current beacon. Ants can read and update pheromone information stored in their current beacon, and (optionally by temporarily moving to the current beacon) its immediate neighbors. Ants can also detect if food or the nest are within range. Ants do not directly communicate with each other, and likewise beacons do not communicate with each other. In fact, beacons need not be wireless or even active: they could be buckets, RFID tags, or other markers with limited visibility range. Ants only need to be able to identify and home in on individual beacons, and to store and retrieve data in them.

Each ant has a mode, either FORAGING or FERRying, indicating the ant’s current task. All ants start FORAGING and located at the nest. At each iteration an ant updates the pheromones of its current beacon (if any), then performs one action (such as moving to the food or nest; exploring randomly; moving, deploying, or deleting a beacon; etc.), then the ant updates the pheromones again for good measure. The action decision process is described later. Afterwards, each beacon depletes (evaporates) its foraging and ferrying pheromone values by multiplying them by a fixed constant 0 < β ≤ 1. Except in the case of moving obstacles, depletion is not required (β can be 1). We set β = 0.9. The wander pheromone is not depleted; it is updated as described below.

3.1 States and Pheromones

Ants spread pheromone values from beacon to beacon using roughly the same method as in the (beacon-less) [16]: a form of value iteration. From a dynamic programming perspective, beacons are the states in the environment, and pheromone values are the utilities \( U_b \) of those states, one per pheromone \( p \). To illustrate state (beacon) transitions, consider Figure 1, where the ant has recently transitioned from state (beacon) \( s \) to state \( s' \) and may transition to any of the \( s^{(t)} \) states in the immediate future. After reaching a
goal via a series of transitions, an ant will receive a positive reward \( R_p \) and change its mode from FORAGING to FERRYING or vice versa. Specifically, when the ant has just changed to (or is starting as) FORAGING, \( R_{\text{ferrying}} \) is set to a positive constant \( R_f \), and when the ant has just changed to FERRYING, \( R_{\text{foraging}} \) is set to \( R_f \). In all other cases, \( R_p = 0 \) for all \( p \).

**Updating.** To be maximally general, let us start with the assumption that for each \( s' \), the ants have available a model \( T(s', a, s'') \) indicating the probability of transitioning to \( s'' \) from \( s' \) if the ant chooses to perform some action \( a \). The update rule for each pheromone \( p \in \{ \text{foraging}, \text{ferrying} \} \) is a variation the Bellman Equation in which \( U_p \) does not decrease:

\[
U_p(s') = \max \left( U_p(s'), R_p(s') + \gamma \max_{a \in A} \sum_{i} T(s', a, s''') U_p(s''') \right)
\]

where \( \gamma \) is a value between 0 and 1. However, in actuality transitions are deterministic (or at any rate, \( T \) is not available to the ants), which reduces the update equation to:

\[
U_p(s') = \max \left( U_p(s'), R_p(s') + \gamma \max_{a \in A} U_p(s'') \right)
\]  

The **wandering pheromone** is updated so that the ant will move away from commonly-visited beacons. Specifically:

\[
U_{\text{wandering}}(s') = U_{\text{wandering}}(s') - 1
\]

**Following.** When it is in the mode of transitioning from beacon to beacon, an ant’s policy is fixed: if the ant is FORAGING, or FERRYING, or in certain situations wandering, it will move to the neighboring state beacon whose pheromone for that mode is highest. That is, it will select the beacon

\[
\arg\max_i U_p(s'_i)
\]

where \( p \) is foraging, ferrying, or wandering, as appropriate. We refer to this action as following the pheromone trail.

It’s important to note here that the ants are updating *all* the pheromones, but are only transitioning along *one* pheromone at a time. This has an important quality which crucially differentiates this method from plain value iteration. Because the ant is updating all pheromones as it transitions from the food to the nest (or back), it can build up a gradient in \( O(n) \) time, where \( n \) is the number of transitions. Contrast this to traditional dynamic programming or reinforcement learning approaches, whose repeated backups require \( O(n^2) \) time. This significant improvement in complexity is made possible by the symmetry of the environment in the ant foraging problem: the probability of transitioning from state \( s_j \) to state \( s_i \) is the same as the probability of transitioning from \( s_i \) to \( s_j \).

### 3.2 Algorithm with Pre-Positioned Beacons

To introduce the algorithm more gently, we first present a version that considers the beacons as fixed in the environment. This algorithm is the sparse analogue of the dense pheromone grid structure in [16] and is shown here to demonstrate its close similarity. In the next section we will then extend this to the full algorithm, with beacon deployment, movement, and deletion.

Robustness can be accomplished with some degree of randomness. We add this in two ways. First, with small probability the ant will initiate a temporary exploration mode in which it performs random transitions some number of times before resuming normal operation. Second, if the ant does not (with some probability) decide to do its standard “pheromone following” procedure, it will instead follow the wandering pheromone, which pushes it to areas not well explored.

The algorithm relies on certain constants: \( R_{\text{reward}} \), \( R_{\text{exploration}} \), \( R_{\text{foraging}} \), and \( R_{\text{ferrying}} \). How much reward the ant receives when it reaches the nest or food: \( R_{\text{reward}} \), How long the ant will stay in exploration mode: \( R_{\text{exploration}} \), how likely the ant will enter exploration mode: \( R_{\text{foraging}} \), how likely the ant will do pheromone following: \( R_{\text{ferrying}} \). An ant will always have a current beacon unless no beacons are yet placed near the nest. The algorithm looks like this:

1. **global variables:**
   2. \( \text{mode} \leftarrow \text{FORAGING}, \text{count} \leftarrow 0, \text{and } R_{\text{reward}} \leftarrow R_{\text{foraging}} \)
   3. **loop**
   4. \( c \leftarrow \text{compute current beacon, if any} \)
   5. **if** \( c \text{ exists} \) **then**
   6. UpdatePheromones(c)
   7. **if** food within range of me and \( \text{mode}=\text{FORAGING} \) **then**
   8. Move to food, \( \text{mode} \leftarrow \text{FERRYING}, \text{count} \leftarrow R_{\text{ferrying}} \)
   9. **else if** nest within range of me and \( \text{mode}=\text{FORAGING} \) **then**
   10. Move to nest, \( \text{mode} \leftarrow \text{FORAGING}, \text{count} \leftarrow R_{\text{foraging}} \)
   11. **else if** \( \text{Rand}(R_{\text{exploration}}) \) **then**
   12. **else if** c exists and CanFollow(c) **then**
   13. Move to random neighbor of c, \( \text{count}=\text{count}–1 \)
   14. **else if** Rand(1) **then**
   15. **else if** c exists **then**
   16. **else if** CanFollow(c) **then**
   17. Follow(mode, c)
   18. **else if** c exists **then**
   19. **else if** c exists **then**
   20. Follow(WANDERING, c)
   21. \( c \leftarrow \text{recompute current beacon, if any} \)
   22. **if** c exists **then**
   23. UpdatePheromones(c)

This algorithm directly extends [16] to the more general connectivity graph case. The functions used in the algorithm are:

**UpdatePheromones(current beacon).** Each pheromone is updated using Equations 1 and 2, where \( s' \) is the current beacon, and where either \( R_{\text{foraging}} \) or \( R_{\text{ferrying}} \) is set to the current reward, depending on whether the mode is presently FERRYING or FORAGING respectively. The other \( R_p \) is set to 0.

**Rand(prob).** Return true with probability prob, else false.
3.3 Deploying, Moving, and Deleting Beacons

This initial algorithm is sensitive to beacon location. If the beacons are positioned poorly, the ant trail will be suboptimal; and if the graph is disconnected, the ants may not be able to find a trail to food at all. For these reasons it is advantageous for the ants to be able to deploy the beacons on their own, then later move them and ultimately remove excess or unnecessary beacons to optimize the graph. We now extend the algorithm to include these cases.

This requires a few new constants: $p_{\text{Deploy}}$ is the probability of deploying a new beacon, and $p_{\text{Move}}$ is the probability of moving the current beacon. Beacon deletion always occurs if it is feasible. Certain other constants are described later. The algorithm only differs from the previous one in certain lines, denoted with $\lceil \rceil$.

The revised algorithm is:

$$
\begin{align*}
1: & \quad \text{global variables:} \\
2: & \quad \text{mode} \leftarrow \text{FORAGING}, \text{count} \leftarrow 0, \text{and reward} \leftarrow \text{REWARD} \\
3: & \quad \text{loop} \\
4: & \quad \text{c} \leftarrow \text{compute current beacon, if any} \\
5: & \quad \text{if c exists then} \\
6: & \quad \text{UpdatePheromones(c)} \\
7: & \quad \text{if food within range of me and mode=FORAGING then} \\
8: & \quad \text{Move to food, mode=FERRYING, reward=REWARD} \\
9: & \quad \text{else if nest within range of me and mode=FERRYING then} \\
10: & \quad \text{Move to nest, mode=FORAGING, reward=REWARD} \\
11: & \quad \text{else if c exists and CanRemove(c) then} \\
12: & \quad \text{Remove(c)} \\
13: & \quad \text{else if count>0 and c exists and has neighbors then} \\
14: & \quad \text{Move to random neighbor of c, count=\text{count}−1} \\
15: & \quad \text{else if Rand(p_{\text{Explore}}) then} \\
16: & \quad \text{count = COUNT} \\
17: & \quad \text{else if c exists and CanMove(c) and Rand(p_{\text{Move}}) then} \\
18: & \quad \text{Move(c)} \\
19: & \quad \text{else if c exists and CanFollow(c) and Rand(p_{\text{Follow}}) then} \\
20: & \quad \text{CanFollow(mode, current beacon). Let beacon B be the neighbor of the current beacon with the highest pheromone value corresponding to mode. If B exists and its pheromone value is > 0, return true, else false.} \\
21: & \quad \text{Follow(mode, current beacon). Move to the neighbor of the current beacon with the highest pheromone value corresponding to mode (break ties randomly).}
\end{align*}
$$

Figure 2: Example trace of the algorithm in action.
else if 22: Follow(mode,c)
else if CanDeploy() and Rand(PDeploy) then
   Deploy()
else if c exists then
   Follow(WANDERING,c)
else move to the closest beacon, breaking ties randomly
if c ← recompute current beacon, if any
if c exists then
UpdatePheromones(c)

Note line 27: deleting or moving a beacon can cause an ant associated with that beacon to become stranded such that there are no beacons within its range. On line 27 the ant searches (for example, in a spiral) to find and move to the closest beacon.

The new deployment, deletion, and movement functions are:

CanDeploy(). The goal is to only deploy a beacon into an uncrowded region, and only if there are beacons left to deploy.

This requires three new constants: DEPLOYTRIES (10), DEPLOY RANGE (0.9), and DEPLOYCROWD (0.6). If the maximum number of beacons has been reached, return false. Otherwise the ant tries DEPLOYTRIES times to find a random non-occluded location no further than DEPLOY RANGE × range away from the current beacon, or from the ant (if there is no current beacon), such that there is no beacon within DEPLOYCROWD × range of that location. If a location was found return true. Else return false.

The maximum number of beacons controls the overall beacon density. Since we are most interested in sparse beacon deployment, we’ve set this low (3).

Deploy(). Deploy a new beacon at the location computed by CanDeploy(). Move the ant to that location.

CanMove(current beacon). The goal is to move the beacon precisely in-between neighbors likely to be on the ant trail, so as to straighten the trail, without breaking other possibly important trails.

Let locations B₁ and B₂ be the positions of the neighbors of the current beacon with the highest foraging and ferrying pheromones respectively, breaking ties randomly; let P₁ and P₂ be those foraging and ferrying pheromone values; and let W be the minimum wander pheromone of the two. If the food is within range of the ant, replace B₁ with the food location and set P₁ = REWARD; likewise if the nest is within the range of the ant, replace B₂ with the nest location and set P₂ = REWARD.

Compute a new location that is the midpoint between B₁ and B₂. Return false if any of the following are true:

- If B₁ or B₂ do not exist, or B₁ = B₂, or P₁ = 0, or P₂ = 0.
- If after relocating to the move location, the set of “interesting” neighbors of the current beacon, as defined next, would not be a subset of the neighbors of the beacon at the new location. Notionally this can be done by moving the ant to the location, testing, then moving back.
- If the new move location is not reachable from the current location due to an obstacle or environmental border. Notionally this can be done by moving the ant to the location, testing, then moving back.

Else return true.

An “interesting” beacon is one which is likely to be part of an important path. We’d prefer to not damage such paths. At present we test for such beacons conservatively based on how often they’ve been used (their wander pheromone). We define two new constants: WANDERFRACTION (0.7) and MINWANDER (200). A beacon is “interesting” if its wander pheromone is ≤ WANDERFRACTION × W (it’s not been very much) and if the wander pheromone of the current beacon is ≤ MINWANDER (the region is old enough to have reliable wander statistics).

Move(current beacon). Move the beacon, and the ant, to the midpoint location computed by CanMove().

CanRemove(current beacon). There are two cases in which we presently remove beacons: first, if the beacon appears to be stranded or at the end of an abandoned string, and second, if the beacon is redundant.

For the first test, we return true if the neighborhood of the current beacon contains more than 2 other beacons, and the current beacon is not within the range of the food or nest, and it’s old enough (its wander pheromone is ≤ MINWANDER).

For the second test, we return true if there is another beacon (breaking ties randomly), called the merge beacon, within the range of the current beacon which has both higher foraging and ferrying pheromones, and which is within range of everything (food, nest, other beacons) that are within the range of the current beacon.

If we fail both tests, we return false.

Remove(current beacon). Remove the current beacon. Set the wander pheromone of the merge beacon (if any) to the minimum of the merge beacon and the current beacon.
3.4 Example

An example trace of the algorithm is shown in Figure 2. The ants optimize the trail in two basic ways. First, they may adopt a new route through the established beacons, or newly deployed beacons. Second, they can move beacons and eventually remove beacons which are “sufficiently close” to one another. Eventually the ants establish a reasonably optimized route between the food and nest, abandoning suboptimal routes and tightening up the best-discovered route.

Note that the path will likely never be fully optimized in this example because our present beacon-deployment and beacon-movement rules are overly conservative: the beacon-movement rule tries at all cost to avoid breaking chains; and the beacon-deployment rule winds up refusing to deploy in certain situations it perceives as overly crowded, even though they are needed to improve the current route.

4. EXPERIMENTS

We tested our algorithm to demonstrate the ant’s ability to find the food, to discover optimal transition sequences between the nest and the food, and to recover when obstacles are added to the environment. Our metric was the amount of food collected by all the ants every 50 timesteps.

Obstacles. To perform our experiments, we constructed several different obstacle environments, which are shown in Figure 3. We chose these environments to test two aspects of the algorithm: adaptively searching through the beacon graph (Exploration), and moving beacons so as to optimize the path when allocated only a limited budget of them (Optimization).

The L obstacle allowed ants to deploy beacons near the food source but forced them to create a path around the edge of the environment. The ants could effectively explore the majority of the landscape. The Block and Block2 obstacles forced ants into narrow corridors, and occupied much of the landscape. Finally, the Ant Clock With Obstacle environment, discussed later, included a large obstacle to complicate a dynamic environment. We compared each of these obstacles with a Blank environment as a control. For L, Block, Block2, and Blank the nest was placed in the upper left corner at (10, 10), and the food in the lower right corner at (90, 90).

Minutiae. We ran 63 independent runs in the MASON simulator [9] for 14,000 timesteps each. $p_{\text{Follow}}$ and $p_{\text{Deploy}}$ were fixed at 0.9. All experiments used 100 ants, a 100x100 bounded, continuous world, and a beacon range of 10. We limited the number of available beacons to 60 for Block, 100 for Block2, and 400 for all other environments. These limits provide just enough beacons to establish a trail for the given environment. We compared results after 2,000 and 14,000 timesteps using an ANOVA test with a Tukey post hoc comparison at the 95% confidence level.

4.1 Exploration

We began by examining the ants’ ability to find better routes through the graph given beacons which, once deployed, could not be moved or deleted. Accordingly, we set $p_{\text{Move}} = 0$, and $p_{\text{Explore}}$ varied over 0.0001, 0.001, 0.01, 0.1, 0.5, and 0.9. We studied two cases: removing obstacles (which created new situations to exploit) and adding new obstacles (which had to be worked around).

Removing obstacles. We began by letting the ants discover a suboptimal trail around some obstacle for 3,000 timesteps, then removed the obstacle to see how the ants would find new routes to take advantage of their revised situation. Our obstacle of choice for this experiment was the L obstacle. Figure 4 shows the performance with $p_{\text{Explore}} = 0.1$ before and after removing the L obstacle. As can be seen, the ants rapidly adapt to the new situation. After the obstacle is removed, performance converges rapidly to approximately the same performance of the ants on the Blank environment, as the ants find a superior path through the beacon graph. Changing $p_{\text{Explore}}$ does not significantly alter this rate of adaptation: though larger values of $p_{\text{Explore}}$ generally result in significantly lower total food collection as more time is spent exploring.

Adding obstacles. Next we examined how the ants would react to environmental changes that made previously good trails no longer viable. We let the ants explore an empty environment for 3000 time steps, and then introduced obstacles. When an obstacle was introduced, it might collide with a number of beacons and ants. We treated these as destructive events to the beacons and ants. Specifically, a beacon in collision was automatically removed. The ants in collision were “killed” — it was eliminated entirely, and so the total count of ants was reduced by one. We chose to do this rather than artificially “restart” the ant at the nest or “move” it to a safe location.

Figure 5 shows the ants’ performance for two values of $p_{\text{Explore}}$. In both cases we see that the ants can recover, but their performance after the obstacle is introduced is reduced proportionally to...
the number of resources (ants and beacons) that were destroyed by the obstacle. The L obstacle, which exhibits the best recovery, covers a smaller area than either Block or Block2, both of which show more limited performance.

In comparing the two graphs, it is important to note the scale. While the ants recover faster with a higher \( p_{\text{Explore}} \) value, their overall performance is less than the ants with a lower \( p_{\text{Explore}} \). This is essentially the same situation as noted in the previous experiment (removing obstacles): as the ants spend more time exploring (represented by higher \( p_{\text{Explore}} \) values), they spend less time ferrying food to the nest. As \( p_{\text{Explore}} \) approaches 1, the performance of the ants drops dramatically as they spend more time exploring.

### 4.2 Dynamic Food Location

Having shown that the ants could adopt better trails where the food location was not moving, we next tested to see if this held when the food was moving. To do this, we recreated an experiment performed in [16], called the “Ant Clock”. In this experiment, the nest started in the center of the environment (50, 50), and the food was initially placed due east, 10 distance units from the right edge of the environment. At each timestep, the food would rotate about the nest in a clockwise direction at one-quarter of a degree per timestep. We placed an obstacle north of the nest such that the food would just clear the left and right edges of the obstacle in its orbit about the nest (see Figure 3). As a control we had the food source rotate about the nest, but without the obstacle to the north. We set \( p_{\text{Move}} = 0 \), and varied \( p_{\text{Explore}} \).

Without the obstacle, the ants ably adapted to the constantly moving obstacle, maintaining an approximately straight-line path at all times. With the obstacle, the ants’ path would effectively “bend” around the obstacle as the food passed by it, but eventually the exploration would enable the ants to reestablish an optimized straight path. This “bending” is reflected in the periodic drops in performance in Figure 6. Figures 6(a) and 6(b) illustrate another tradeoff of more exploration: higher values of \( p_{\text{Explore}} \) decrease the absolute amount of food returned to the nest; but higher \( p_{\text{Explore}} \) values also decrease the severity of the periodic drops. The increased exploration prevents the ants from spending too much time on the established, suboptimal “bent” trail.

### 4.3 Optimization

In our final experiment, we set \( p_{\text{Explore}} = 0 \) and varied \( p_{\text{Move}} \) to test the ants’ ability to optimize trails with a limited number of beacons. Here we used the Block and Block2 environments. We sought to test one of the key ideas behind our algorithm: that a sparse representation of pheromones benefits from physical revisions and updates. In the experiment, let the ants establish a trail for 3,000 timesteps and then removed the obstacle.

After removing the obstacle, the ants would begin to move the beacons so as to straighten the trail, and eventually straighten the trail entirely, deleting redundant beacons as the trail became shorter. For \( p_{\text{Move}} = 0.1 \), Figure 7 shows that while performance statistically improves after the obstacle is removed, it does not converge to the performance of Blank. We can surmise two possible reasons for this: first, though beacons would be removed, the trail was ultimately still denser with beacons than if the ants had been (as in Blank) free to deploy beacons in the space. Second, moving and deleting beacons would occasionally trap ants in “islands”—small disjoint beacon groups — and unable to participate.

Even so, the results on the Block2 environment verified our visual inspection that the trail line was rapidly straightened out and optimized. In the Block environment this effect is not seen, largely because the number of beacons remained approximately the same before and after optimization. Similar performance is seen with \( p_{\text{Move}} = 0.5 \) and \( p_{\text{Move}} = 0.9 \).

### 5. CONCLUSION

We presented an approach to establishing trails among swarm ant robots using a non-invasive and non-destructive stigmergic communication in the form of deployable and movable beacons. The
robots use those beacons as a sparse representation of a pheromone map embedded in the environment. The algorithm uses a variation of value iteration to update pheromones and make transitions from beacon to beacon. We demonstrated the efficacy of the technique and explored its present robustness, and optimization capabilities.

This work is intended as a stepping stone to actual deployment on swarm robots, and using sensor motes as beacons. This deployment is our first task in future work. We will also examine extending the beacon model to more collaborative tasks than simply establishing trails. For example, in other experiments we have demonstrated sophisticated self-crossing, multi-waypoint tours. We believe we can also employ beacons in this model to define regions to avoid or requests for assistance (to move objects or establish formations, for example).

6. REFERENCES


