DRAFT VERSION: Simulation of Cooperative Control System Tasks using Hedonistic Multi-agents

Michael Helm, Daniel Cooke, Klaus Becker, Larry Pyeatt, and Nelson Rushton
Texas Tech University, Lubbock Texas

Abstract

As control system complexity increases, scalability often becomes a limiting constraint. Distributed systems and multi-agent systems are useful approaches in designing complex systems, but communications for coordination are often a limiting factor as systems scale up.

Colonies of social insects achieve synergistic results beneficial to the entire colony even though their individual behaviors can be described by simple hedonistic algorithms, and their available communications are very limited. Cooperative actions emerge from simple fixed action patterns in these insects. Complex control systems formed from a multitude of simpler agents or subsystems, with constrained and limited communications channels may also achieve emergent cooperation. Advantages of such systems are reduced communications complexity, and reduced complexity in any single element of the systems, as well as improved robustness.

Index Terms—control system, cooperative behavior, multi-agents, simulation, emergent behavior

1. Introduction

The potential advantages of multi-agent systems include simplicity of individual agents, robustness due to elimination of single point failures, and scalability. Scalability can be a problem due to communications overhead. Communications costs can be reduced if agents are structured as simple hedonistic agents that only sense their own local environment and have very limited local communications or only stigmergy communications via sensing and modifying the state of the environment. The seeds of ideas for this approach include economic game theory as well as the behavior of social insects. In economies, individual players (agents) are privy to their own view of the world, but seldom have the benefit of full communications with other players. Cooperative behavior in some tasks can emerge even in such a restricted communications environment as the individual agents focus on their own self interests.

The purpose of this work is to develop a better understanding of emergent cooperative behaviors in multi-agent environments where communications is very limited or only via stigmergy and where each agent is working toward its own individual goals. Communications overhead for coordinating a multi-agent task can be prohibitive when the problem size scales and the number of agents required becomes large. Klavins suggests that a coordinated task could require all agents to communicate with all other agents to achieve coordination resulting in a communications complexity of $O(n^2)$ [1]. This complexity only describes the number of communications paths and does not fully address the number of messages required to achieve coordination [2]. Ants and other social insects have been shown to achieve coordinated behaviors with very limited and simple forms of communications. Ants have fairly simple fixed action patterns that are triggered by the state of the local environment and those actions can in turn result in change to the state of local environment via stigmergy [3]. Simon has described the complex appearance of the behavior of ants as being a result of the environment rather than inherent complexity in the ants themselves [4]. The behaviors of the ants form a complex dynamic environment. Kennedy and Eberhart consider emergent behavior to be a defining characteristic of such complex dynamic systems [5]. The action patterns of the individual ants appear to be simple, but cooperation emerges, e.g. when several ants work together in moving a large food item. The idea that explicit cooperation is not necessary, and that cooperative behavior emerges from actions of greedy agents is not new, it was described by Korf in 1992 [6]. In this work we will see cooperative behavior that emerges from simple behavior patterns in individual agents as a consequence of their interactions in an environment.

Manuscript received February 7, 2006.
M. Helm is with the Computer Science Dept., Texas Tech University, Lubbock, TX 79409 806-742-3527; fax: 806-742-3519; (e-mail: michael.helm@ttu.edu).
D. Cooke is with the Computer Science Dept., Texas Tech University, Lubbock, TX 79409 (e-mail: daniel.cooke@ttu.edu).
K. Becker is with the Economics Dept, Texas Tech University, Lubbock, TX 79409 (e-mail: klaus.becker@ttu.edu).
L. Pyeatt is with the Computer Science Dept., Texas Tech University, Lubbock, TX 79409 (e-mail: larry.pyeatt@ttu.edu).
N. Rushton is with the Computer Science Dept., Texas Tech University, Lubbock, TX 79409 (e-mail: nelson.rushton@ttu.edu).

Consider ants working together to move an object that is larger than a single ant could move. Each ant tries to move the object using simple actions such as grasp and lift, drag at some angle, or drag at another angle. An ant working on a single object will not likely start at the same time, nor initially be attempting the same action as other ants attempting to move that object. An ant will attempt one action or another and will continue trying different actions until movement of the object is observed or after sufficient time has elapsed the ant will give up on that particular object and move on to some other task. Some of the action patterns in ants which results in cooperative object movement has been described in some detail by Bonabeau, Dorigo, and Theraulaz [7]. The simulation developed here uses a somewhat simpler set of actions. As different ants try different actions, eventually enough ants will be attempting the same action (e.g. pulling in a particular direction) that the object may move slightly. The slight object movement is sufficient to reinforce the last action of each ant to repeat that same action again. Perhaps ten ants are working to move an object, and six pull in a particular direction at the same time resulting in a slight movement, while the other four are doing something that is not contributing to the movement. All will be reinforced to repeat their last action to cause movement again. The sum of the actions by six of the ants doing something beneficial and the other four ants doing something else is still sufficient to cause the object to move. As long as the sum of all actions is sufficient to overcome the weight or drag force of the object to be moved, the object will move.

Movement occurs if: $$\Sigma(b_i) - \Sigma(nb_i) \geq f_m \quad (1)$$

Where b_i is a force on the object to be moved from a beneficial action and nb_i is a force on the object to be moved from a non-beneficial action. The force required to move the object is given by f_m.

The four non-contributing ants may eventually adjust their actions based on feedback from the object being moved, but it is not necessary and movement will continue as long as sufficient ants are doing beneficial actions and the remaining ants are not causing too much non-beneficial actions. Cooperation emerges over time simply as a result of each agent following its own simple fixed action patterns and seeking its own goals based on the current state of the local environment. Explicit inter-agent communications among all the agents is not required. The state of the object being moved is what provides feedback to each of the agents so that they know whether to continue their present action or try another action. This minimal amount of communication via stigmergy is a significant improvement over the level of agent-to-agent communications that is typical in multi-robotic systems as described by Klavins.

2. Experimental Setup

The simulation tool for this work was developed using Visual Studio C++ for a standard desktop PC running Windows 2000. A 50x50 grid world was created that could be populated with some number of ants (agents) and some number of objects to be moved. The upper left corner area of the grid world was assigned as the nest area where ants would start their search from and where they would attempt to return objects to. The ants all had the same types of fixed action patterns based on their own state and the state of nearby (adjacent) cells in the grid world. Each object to be moved could have different characteristics of weight (amount of force required to lift it) and amount of force required to drag it if it was not lifted. The ants all had the same fixed action patterns, but started their search from random locations within the nest area.

The ant's action patterns were to search for objects to be moved (prey) while avoiding other ants and walls of the grid world. Ants could not occupy the same cell as another ant, nor the same cell as a prey object, nor cells representing the walls of the grid world. New ants could be added to the simulation one at a time or in groups of ten ants. Each new ant added to the simulation would start in some pseudo-random location within the nest area and then proceed in a quasi-random search pattern, looking for prey objects. The search pattern was given a slight bias of moving toward the southeast in the grid world so that ants would move away from the nest. Each ant also kept track of the number of time steps it had been moving away from the nest and after a sufficient time not encountering prey, the ant would move rapidly back to the nest and start a new search. Ants that encountered prey in an adjacent cell would "latch on" to the prey and remain in that cell location and begin attempting to move the prey. At first the ants would attempt to lift the prey object for a few time steps. If lifting action did not cause the prey to move, then the ant would switch to attempting to drag the prey object toward the nest and would continue that action pattern indefinitely. Once an ant switched to trying to drag the prey object, it would release recruitment "pheromone" into surrounding grid cells to attract other ants to this particular prey object. Real ants have this capability of recruiting other nearby ants to assist in a task. The pheromone was only released if the prey object was not being successfully moved at that time. The pheromone was propagated to surrounding cells for a distance of approximately ten cells away from the ant, and had an associated decay rate so that pheromone that was further away had a weaker strength. Other ants still in search mode that encountered pheromone in an adjacent cell would begin to move toward the strongest concentration of pheromone, so in time would move toward the ant releasing the pheromone. The objective was for the ant at the prey and unable to move it alone to be able to attract other ants to help out.

New ants arriving at the prey object (either via pheromone recruitment or random search) would engage in the same fixed action patterns as the other ants, i.e. first attempt to lift the prey object for some number of time steps and then eventually switch to attempting to drag the prey after some number of time steps if t success did not occur in the lift mode. Again, once an ant switched to attempting to drag the prey, it would release recruitment pheromone into the surrounding cells if the prey object was not currently moving.

Prey objects were all the same physical size within the grid world (occupying a 9x9 sub-grid), but had different characteristics of weight (when being lifted) and drag force (when being dragged). All prey objects required slightly more force to be lifted than to be moved by dragging, but the actual amounts for each prey object varied from one prey to another. Each ant could produce one unit of lift force when in lift mode and one unit of drag force when in drag mode. For simplicity at this stage of the work, all ants attempt to drag the object directly toward the nest. A prey object having a weight of 6 units would require six ants to be "attached" and attempting to lift it for it to be moved. A prey object having a drag force of 4 units would require four ants to be attached and attempting to drag it for it to be moved. For simplicity, mixed forces of lift and drag combinations were not considered.

The prey object would remain in its initial location unless a sufficient number of ants were attempting to lift it or if a sufficient number of ants were attempting to drag it. Once the prey object began to move due to the actions of the ants, it would continue to be moved toward the nest until the nest was reached. The ants would continue whatever action they had been doing when the object began to move. Ants that had succeeded in moving a prey object back to the nest were then free to begin a new search for other prey.

The simulation could start with a variable number of prey objects, each with its own characteristics in terms of position, weight, and drag force. Ants could be added one at a time or in groups of ten at a time up to a maximum of forty ants. Real-time status information was provided for the position and action of each ant as well as information on status of each prey object. The current time step was always displayed and the simulation could be run in single step mode, or three different continuous speeds. Speed could be changed on the fly. The recruitment pheromone could be displayed or not. A trail pheromone could be displayed or not that would show the recent path of each ant. Both types of pheromone had a built-in decay rate. Amount of pheromone concentration in a particular cell was shown by varying color dots in the cell. Ants were shown with red color dots about the size of a single cell, prey were yellow and covered nine adjacent cells. Nest area was shown as a light green oval. Ants changed color depending on their current action. Red ants are in search mode, ants turn yellow at attachment to prey, blue when in lift mode, then green when in drag mode. This color change was to aid in visual assessment of the progress of the simulation. The simulation could be halted at any time then continued at any speed. Exact position of individual ants could be determined by clicking on them with a mouse. The simulation was normally run with random locations and characteristics for each prey object, but for ease of characterization, some sequences of runs always started with the prey having the same location and characteristics.

3. Results

Emergent cooperation in the movement of prey objects was quickly produced in each run. The number of time steps required before such movement varied depending on location of the prey relative to the likely path of the searching ants (due to the bias to move southeast away from the nest), number of such prey, characteristics of the prey (weight and drag force) and the number of ants searching. Prey objects in the normal path of the ants moving away from the nest would be found rather quickly and surrounded by enough ants to either lift them or drag them within a relatively short number of time steps. Prey objects that were not in the central path of the searching ants generally took quite a bit longer to be moved. Eventually a single ant might encounter such a prey object and eventually release recruitment pheromone which would result in ants further away being recruited from their southeast biased search path to begin moving toward the stronger concentration of recruitment pheromone. Eventually, most recruited ants would find their way to the prey and begin trying to lift it, then drag it toward the nest. Once a minimum number of ants had arrived and were simultaneously engaged in the correct single action, the prey would be moved.

Cooperative behavior emerged from these simple agents without extensive inter-agent communications. In many cases, cooperative movement was achieved without the use of recruitment pheromone. The cooperative object movement simply emerged from simple fixed action patterns of simple agents all driven by their own local world state. Each agent could attach to a prey object, then try different actions to try and move that object. Once the object began to move, the agent could detect that fact and would continue to perform that same action that had resulted in movement. Agents found prey objects due to random search, or in some cases were recruited via pheromone by agents already attempting to move the object without success. The pheromone was not needed in all cases, but did reduce time required to achieve movement for those prey objects that were outside of the easily found zone. Note that the pheromone communications is strictly one-way and does not require any acknowledgement. Pheromones are limited to regions nearby to the agent releasing them and have a fairly rapid decay rate. The decay rate prevents the pheromone signals from saturating the region and allows the pheromone signal to disappear when it is no longer needed. Since pheromone signals are additive, if multiple agents are releasing pheromones, their strength will be greater and would have the capability of spreading to a slightly larger region. For this simulation, the region of pheromone spread was fixed to a certain number of cells from the releaser. Since only one ant could occupy a single cell, ants would wind up scattered around the prey object and each could release its own pheromone cloud. Multiple ants thus releasing pheromone would result in a slightly larger region of pheromone just due to the displacement of the multiple sources around the prey object. Thus, multiple ants releasing recruitment pheromone resulted in a slightly larger recruitment zone. Once a prey object began to move, pheromone release was stopped since additional help was not needed. It was observed that additional ants were sometimes recruited to the previous location of a prey that was already being moved due to the delay in total decay of the pheromone. It was also observed that some ants would get stuck in a single location for a while when homing in on recruitment pheromone, presumably due to equal concentrations of pheromone in a group of adjacent cells. This stuck state did not usually last very long. It was also observed that occasionally an ant that had been recruited would get stuck just outside the prey object without getting close enough to attach. To attach, an ant had to be in a cell adjacent to a cell occupied by the prey object.

Adding more ants resulted in more rapid emergence of the movement of the prey object. Occasionally, a single ant would get stuck for lengthy times, but this fact did not prevent the goals from being achieved as long as sufficient ants were available to ultimately achieve the movement task.

4. Some Observations from the Experiments

Each ant in the simulation started from a pseudo-random location in the nest area. Search patterns of the ants included pseudo-randomness. Due to this, ants would arrive at a prey item at different time steps and in different positions surrounding the prey. Due to ants not being able to occupy the space of another ant, nor space occupied by the prey, they would distribute around the prey in a somewhat random fashion. Thus, individual ants were at different positions and different phases of their attempts to move the prey at any given time step. The coordination emerged from the fact that different ants were trying different things over a series of time steps and received feedback to continue present action any time prey movement was achieved. Cooperative behavior emerged not due to explicit coordination communications, but due to the sum of actions sometimes being sufficient to achieve the goal of interest to each individual agent. Any time the sum of actions of the individual agents was sufficient, that action was “rewarded” and continued. In some sense this can be seen as an effect similar to resonance, where small “noisy” actions can be brought into a coherent phase relationship and result in very strong results.

This appears to be a key idea in the emergence of cooperative behavior in such agents. A certain amount of randomness contributes to the group as a whole eventually achieving cooperative results.

5. Conclusions and Future Work

Cooperative behavior is shown to readily emerge from simple fixed action pattern behaviors of multiple agents in this simulation. The simulation demonstrates that cooperative behavior can be achieved without extensive inter-agent communications. Each agent has simple fixed action patterns that drive it toward certain individual goals where those goals are the same goals that all of the other agents have. Specific actions of an agent at any particular time are driven by its own observation of the nearby state of the world. No agent is aware of anything outside of what it can detect in cells adjacent to it. Feedback on goal achievement is also via state of adjacent cells and is not dependent on world level global information.

Communications overhead is very minimal in this approach. Cooperative behavior emerged without extensive inter-agent communications. In many cases, no inter-agent communications were needed, in other cases, the single one-way broadcast of recruitment pheromone resulted in more agents drawn to working on a single prey object. In no case was explicit two-way communications needed. The one-way pheromone broadcast was only needed in cases that were not resolved quickly. The communications overhead in this approach is very minimal, because it is often not needed at all, and when it is used it is one-way and of a very simple format (i.e. recruitment pheromone concentration). There is of course stigmergy communications via the local world state for all agents. That is, the actions of an agent can change the local world state, and a nearby agent can sense that change in the local world state and modify its own actions. This type of communications is not treated as an explicit communication between agents since the agents are not directly attempting to communicate any information in that case and certainly there is no extra "communications cost" involved in stigmergy [8].

This approach is robust because failure of single agents or even several of the agents does not result in failure of the task as long as sufficient agents were available in the world to achieve the task. It is also robust because complexity of individual agents is limited. In the work shown, the agents are homogeneous, but the authors expect that similar behavior would be observed with heterogeneous agents in at least some cases.

This approach is scalable because goal achievement is achieved faster by adding more agents (within limits of the number of agents that can work at a single task). As the tack size or number changes, it is appropriate to change the number of agents working on those tasks. Since communications overhead is minimal even in the cases where communications is used, and because communications in any case it limited to the nearby region, communications costs never exceed a known amount. Larger problems requiring more agents do not require larger communications overhead.

This relatively simple simulation tool has shown very interesting emergent cooperative behavior results using simple multi-agents using fixed action patterns. Future extensions to this work include some expansion of those action patterns to explore more complex circumstances. The current action patterns do not have any built-in learning capability (other than to continue with the current action if movement resulted). Adding some Reinforcement Learning capability to the agents by using SARSA or Q-learning [9] would be very interesting to explore. It would also be interesting to slightly expand on the pheromone communications to allow two or possibly three different types of pheromones. These could be used singly or in combination to provide an expanded set of information to be broadcast. Ants have slightly more complex communications capability than has been explored in this simulation. Ants are able to combine multiple pheromones to achieve on the order of ten different messages [10], [11]. The multi-agent simulation shown was used in simulated object movement tasks, but the basic ideas apply to a broad range of problems. The appeal of this approach is the reduced communications costs and the improved robustness. Ongoing work by the principal author is attempting to characterize a minimal set of messages that is adequate for general problems addressed by this approach using multi-agents.

Acknowledgment

M. Helm would like to thank Texas Tech University for providing support and resources for doing this work.

References

[1] Klavins, E. *Communication Complexity of Multi-Robot Systems*; In WAFR '02, Nice, France, December 2002.

[2] Welborn, C. *Specifying a Domain Specific Language for Cooperative Robotics*, PhD dissertation Texas Tech University, Dec 2005, pp. 1, 2, 30.

[3] Holldobler, B., Wilson, E. *Journey to the Ants*, Cambridge, MA: Harvard University Press, 1994. pp. 41-58.

[4] Simon, H. *The Sciences of the Artificial 2nd ed*; Cambridge, MA: MIT Press, 1969, 1981, 8th printing 1994. p. 64.

[5] Kennedy, J., Russell, E. *Swarm Intelligence*, San Francisco, CA: Morgan Kaufmann Publishers, 2001. p. 18

[6] Balch,T. and Parker, L., Eds. *Robot Teams: From Diversity to Polymorphism*, A. K. Peters, 2002. p. 52

[7] Bonabeau, E., Dorigo, M., Theraulaz, G. *Swarm Intelligence From Natural to Artificial Systems*, New York, NY: Oxford University Press, 1999. pp. 253-261.

[8] Beckers, R., Holland, O.E., and Deneubourg, J. "From Local Actions to Global Tasks: Stigmery and Collective Robotics." in *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, R. Brooks and P. Maes, Eds., pp. 181-189, Cambridge, MA: MIT Press, 1994

[9] Sutton, R., Barto, A. *Reinforcement Learning An Introduction*, Cambridge, MA: MIT Press, 1998, 4th printing 2002. pp.179-185.

[10] Szuba, T. *Computational Collective Intelligence*, New York, NY: John Wiley & Sons, 2001. pp. 72-73.

[11] Howse, P., "Pheromonal Control of Behavior in Leaf-cutting Ants", in *Applied Myrmecology*, Vander Meer, R., Jaffe, K., and Cedeno, A., Eds., p. 429, Boulder, CO.: Westview Press, 1990.