Decentralized K-Means Clustering and Emergent Computation

R Ryan McCune & Greg R Madey
University of Notre Dame, Department of Computer Science & Engineering

Abstract Overview

Emergent computation characterizes a distributed system that produces a computation greater than the sum of its parts (individual computers). Emergent computation is robust, scalable, adaptable, and efficient, and can potentially alleviate problems with Big Data.

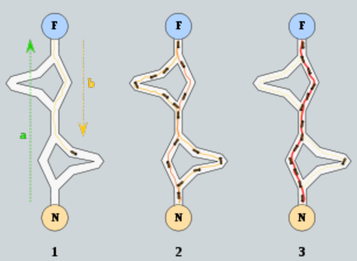
Closely related to emergent computation are swarm intelligent systems; multi-agent systems that also utilize emergent behavior to solve problems. A swarm is presented that performs emergent clustering computation, and evaluated with agent-based simulation. The clustering swarm is based off a previous swarm that computes shortest path. Both swarms are presented to further explore and develop models of emergent computation systems, or an "emergent calculus."

Emergent Computation

Emergent computation is a distributed system producing a global computation greater than the sum of its parts, where a complex computation emerges from the interaction of several networked simple computations. Emergent computation is closely related to swarm intelligent systems.

Inspired by biology, artificial swarms are multi-agent systems that exhibit emergent problem solving capability, where simple behaviors distributed across many agents give rise to complex phenomena. Examples from nature:

Birds form coordinated flocks despite no global perspective

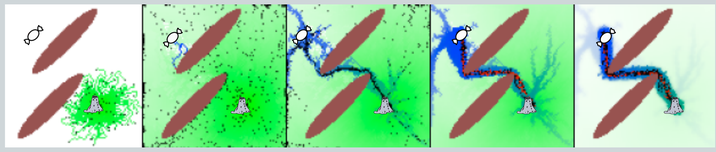


Ants uncover the shortest path to food with only local decisions

Emergent computations are robust, scalable, adaptable, and computationally efficient, potentially alleviating challenges posed by Big Data. However, how to model or engineer an emergent systems remains unresolved.

Ant Foraging Model

A previous model for ant foraging demonstrates [1] how ants uncover the shortest path to food using only local information. Ants randomly search the environment while depositing pheromones. Ants are more likely to follow a path of higher pheromone concentration. Ants follow blue and deposit green while looking for food, then follow green and deposit blue once food is found. Over time, the shortest path to food emerges.



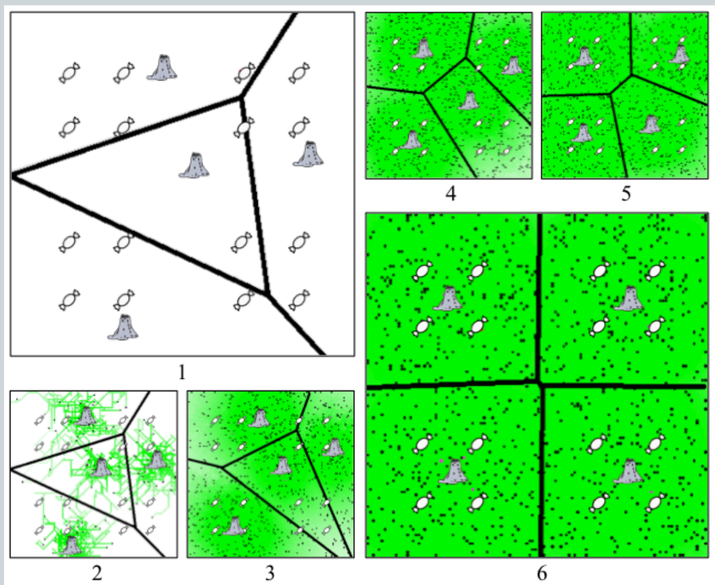
The ant model is adapted for the Decentralized Clustering system. Key differences:

Ant Foraging	Decentralized Clustering
2 pheromones – to food, to ant hill	1 pheromone – to ant hill
1 ant hill	Many ant hills
Stationary ant hill	Ant Hills are mobile

For Decentralized Clustering, when an ant returns to an ant hill with food, the ant hill moves one unit in the direction of the returning ant

Decentralized Clustering

The graphic below depicts decentralized clustering. The system is completely autonomous, with no central control. Ant hills relocate to minimize cumulative distance between ant hills and nearest food. To illustrate which foods are closest to ant hills, a Voronoi diagram is overlaid, using ant hills as seed points

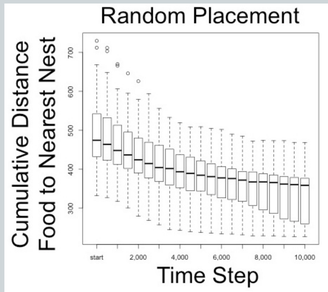


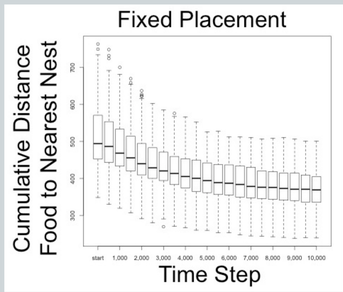
Evaluation

An agent-based simulation was developed to evaluate the decentralized clustering method. The simulation tested clustering for 2 different food distributions, each of 16 food units:

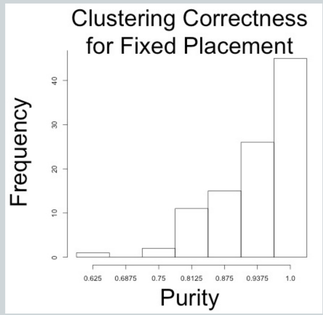
- Fixed placement into 4 squares of 4 foods
- Random placement

Simulations were comprised of 100 ant agents and 4 mobile ant hills, and were run for 10,000 time steps. Both food distribution scenarios were run 100 times.





Since the fixed placement scenario includes defined clusters, the correctness of the clustering can be evaluated using the well-known Purity metric. Out of 100 runs, the Decentralized Clustering technique correctly identified all 4 clusters 48 times.



Conclusions

The relationship between swarm intelligence and emergent computation is explored. Two swarm intelligent systems are presented, where each emergent behavior is a computation. In addition to the well-known ant foraging model that computes shortest path, an adapted Decentralized Clustering model was introduced that computes clusters. An agent-based simulation was developed to evaluate the clustering model. Future work will adapt the swarms to a distributed computing model for emergent computation.

References

[1] Panait, Liviu, and Sean Luke. "A pheromone-based utility model for collaborative foraging." *Proceedings of the Third International Joint Conference on Autonomous Agetns and Multiagent Systems-Vol 1*. IEEE Computer Society, 2004.

Acknowledgement

This work was supported by the Air Force Office of Scientific Research DDDAS program #FA9550-11-1-0351, and the Department of Education's GAANN Fellowship through the University of Notre Dame's Computer Science Department.

Spring Simulation Multi-Conference April 16-18, 2014 Tampa, FL, USA