Multiagent Behaviors in Neural Network

Jimmy Lin
Barry Feigenbaum

Prof. Risto Miikkulainen

Department of Computer Science
The University of Texas At Austin

November 4, 2014

Table of Contents

1. Overview of Multiagent System
2. Multiagent behaviors in Neural Network
3. Discussions
4. Our Research Project
5. Appendix

Recap

What we currently have?

- Commitee Machine
- Reinforcement Learning
- Neuro-evolution
- High-level Behaviors

Motivation

- Most of works so far has focused exclusively on single agents we can extend reinforcement learning straightforwardly to multiple agents if they are all independent.
- Intuitive idea: Multiple agents together will outperform any single agent due to the fact that they have more resources and a better chance of receiving rewards.
- Today, we will touch a really broad area

"Multiagent System" (M.A.S.).

Introduction

What is multiple agent system?

- Unfortunately, it is not formally defined by M.A.S. community.
- Employment of multiple agents (10 to thousands).
- Intelligent mechanisms to address interactions between agents.

When is it proposed?

- a relatively new sub-field of computer science
- has only been studied since about 1980
- only gained widespread recognition since about the mid-1990s

M.A.S. Environments

The agents in a multi-agent system have several important characteristics:

- Autonomy: the agents are at least partially independent, self-aware, autonomous.
- Local views: no agent has a full global view of the system, or the system is too complex for an agent to make practical use of such knowledge.
- Decentralization: there is no designated controlling agent (or the system is effectively reduced to a monolithic system).

Applications/Simulations

- Crowd Simulation / Crowd Collision Avoidance
- ClearPath: Highly Parallel Collision Avoidance for Multi-agent Simulation
- MATISSE: A Multi-Agent based Traffic Simulation System

Independent v.s. Cooperative Agents

Tang Ming (1993) [**1**] studied the performance of cooperative agents, using independent agents as a benchmark. Here are the discoveries:

- Additional sensation from another agent is beneficial if it can be used efficiently.
- Sharing learned policies or episodes among agents speeds up learning at the cost of communication.
- For joint tasks, agents engaging in partnership can significantly outperform independent agents although they may learn slowly in the beginning.

Main topics of M.A.S.

Currently active research areas of M.A.S. are advanced Multiagent Behaviors, as follows:

- Communication
- Cooperation and Coordination
- Negotiation
- Distributed Problem Solving
- Multi-agent Learning
- Fault Tolerance

Communication

Communication is defined as altering the state of the environment such that other agents can perceive the modification and decode information from it.

- Direct Communication
- Indirect Communication

Cooperation and Coordination

Negotiation

Distributed Problem Solving

The state space of a large, joint multi-agent task can be overwhelming. An obvious way to tackle this is to use domain knowledge to simplify the state space, often by providing a smaller set of more powerful actions customized for the problem domain.
An alternative has been to reduce complexity by heuristically decomposing the problem, and hence the joint behavior, into separate, simpler behaviors for the agents to learn. Such decomposition may be done at various levels (decomposing team behaviors into sub-behaviors for each agent; decomposing an agents behavior into sub-behaviors; etc.), and the behaviors may be learned independently, iteratively (each depending on the earlier one), or in a bottom-up fashion (learning simple behaviors, then grouping into complex behaviors).

Multi-agent Learning

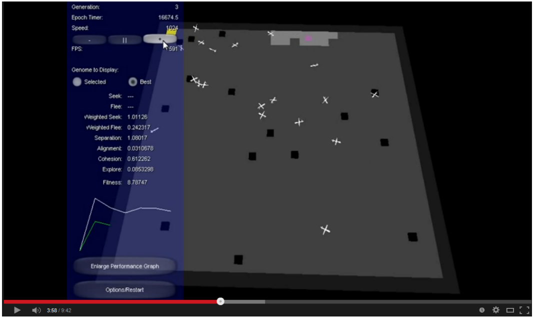
Fault Tolerance

Social Learning

- In **social learning**, agents improve their own performance by learning from observations made by other members of the population
- Traditionally, agents are partitioned into students and teachers, where teachers are chosen to be high fitness members of the population
- **Problem**: Fitness is not a guarantee of superior behavior. High fitness agents may make mistakes, and low fitness agents can still do well under certain circumstances

Cooperative search

- In this example, agents represent **Unmanned Aerial Vehicles (UAVs)** searching a grid of airspace
- Agents have limited communication and some mobility restrictions (such as slow turning)

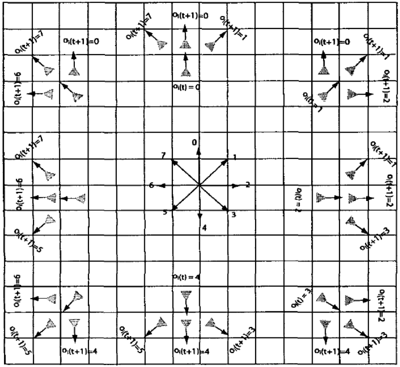


UAV Example: Goals

- Compare a *centralized* approach, where agents share a network, to a *decentralized* approach
- Determine whether opportunistic sharing of network weights improves the speed and accuracy of learning

UAV Example: Setup

- Trained via Reinforcement Learning
- The search space is divided into a grid, and each cell is given a *certainty* value between 0 and 1
- The uncertainty of a cell is reduced by half for each agent that enters that cell
- Rewards equal the decrease in uncertainty in a given cell divided evenly among the entering agents



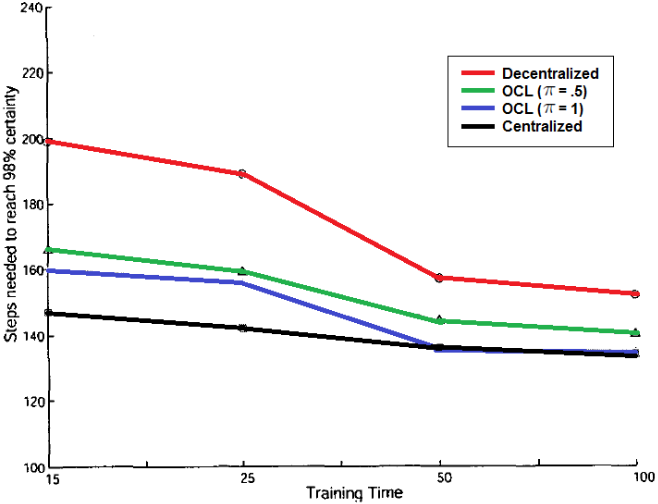
UAV Example

- Agents want to spread out and avoid entering the same cell
- At each time-step, an agent only has three options (turn left, turn right, keep straight)
- Three neural networks are used to estimate the benefit of each course

Compare three architectures:

- **Centralized Learning**: three shared networks receive input from all agents
- **Decentralized Learning**: each agent has it's own three networks
- **Opportunistic Cooperative Learning**: each agent has it's own three networks. Additionally, when two agents come close together, the less successful one copies the other's network with probability π

UAV Example: Results



Egaltarian Social Learning

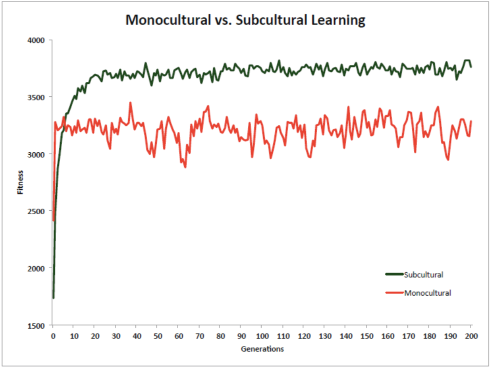
- In **Egaltarian Social Learning** (**ESL**), agents learn from each other instead of from a limited number of teachers
- The population of agents is partitioned into subcultures at the start of each generation
- An agent may probabilistically learn from any other agent in the same subculture

ESL Example: Foraging

- Agents move independently across a continuous world filled with several types of plants
- Plants are automatically consumed when an agent draws close enough; some types of plant give positive reward, others negative reward
- Agents have sensors for each type of plant, as well as a sensor for their own velocity
- They cannot see other agents, or plants they have already consumed
- Networks are trained with backpropagation and NeuroEvolution of Augmenting Topologies (NEAT)

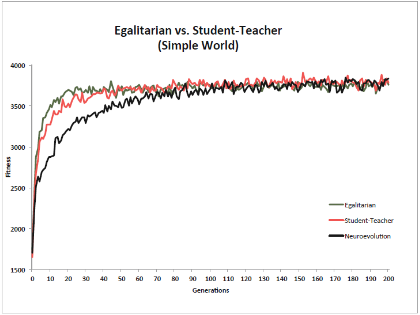
ESL Example: Foraging (Results)

- ESL without subcultures performs worse than plain neuroevolution
- Adding subcultures results in a significant improvement

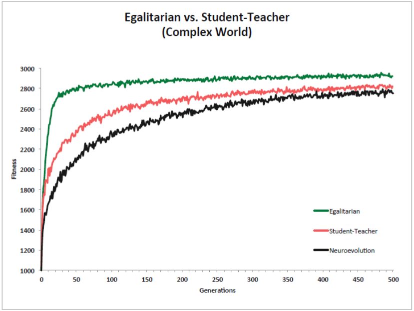


ESL Example: Foraging (Results)

Comparing ESL with student-teacher learning and plain neuroevolution



Simple environment (few plant types)



Complex environment (many plant types)

ESL Example: Foraging (Conclusion)

- ESL is advantageous because it promotes diversity and avoids premature convergence

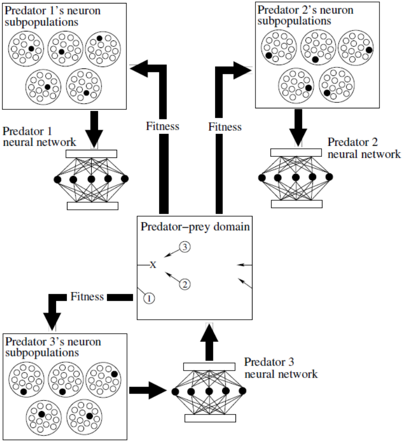
Communication

Predator / Prey

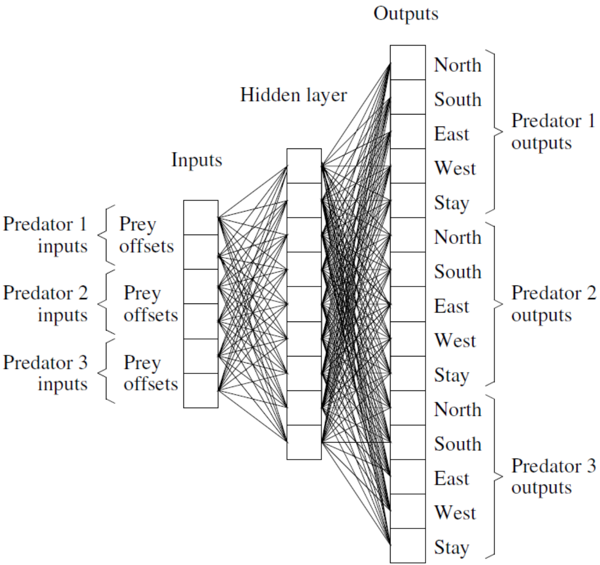
Three different examples:

- Centralized network
- Autonomous networks with communication
- Autonomous networks *without* communication

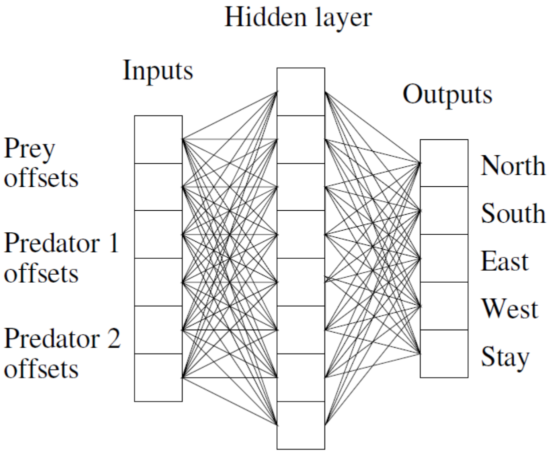
All are trained using ESP or **Multi-Agent ESP**



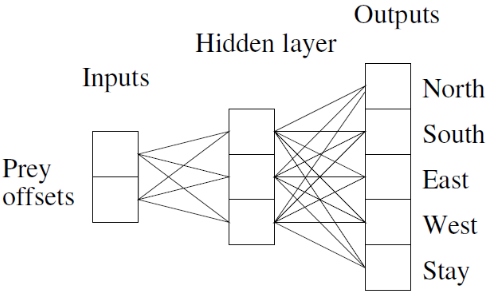
Centralized Network



Communicating Network

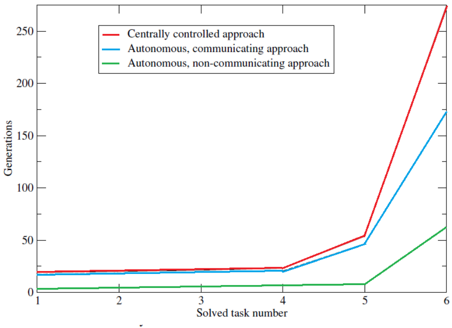


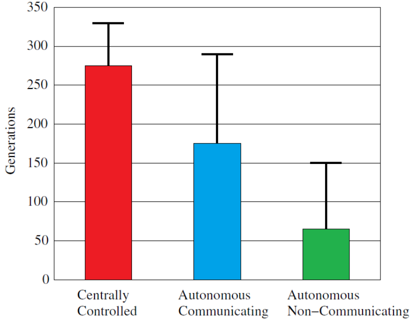
Non-communicating Network



Predator / Prey Results

How long does it take these networks to become successful?





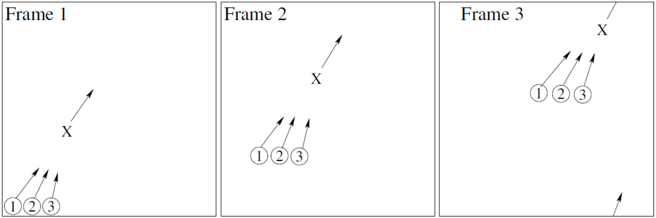
Conclusion: The predators are able to cooperate using **stigmergy** – indirect coordination based on the environment

- What if the environment is more complicated?

¿¿¿¿¿¿¿ 92faa0a3d6de879025fc240229d497b1b7d51e21

Predator / Prey Conclusion

- What happens if we use homogeneous networks? (All predators have the same network weights)
- Communication becomes a necessity



- Even with communication, homogeneous networks do poorly, catching the prey only 3% of the time

Homogeneous Networks

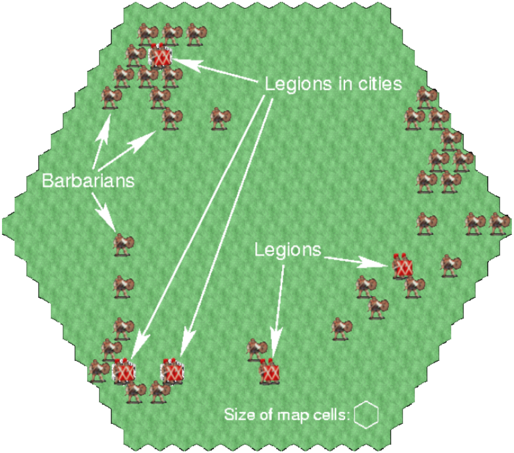
- Thus far, all multi-agent solutions have used heterogeneous networks
- In the predator prey, homogeneous networks performed very poorly
- Can we evolve agents with identical networks, which nevertheless cooperate by adopting different strategies?

Legion Game

The Legion game consists of:

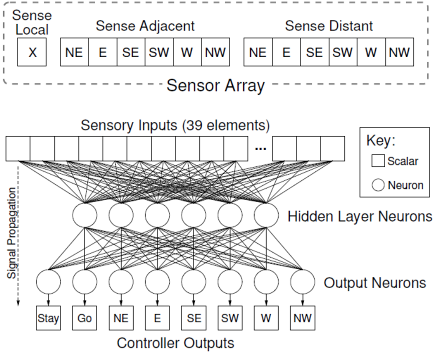
- A hexagonal grid (the countryside)
- A number of randomly placed cities
- Several legions (defenders)
- Many barbarian attackers

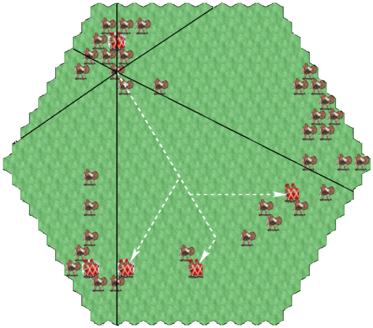
At every game tick, legions are penalized for each barbarian on the board



Legion Network

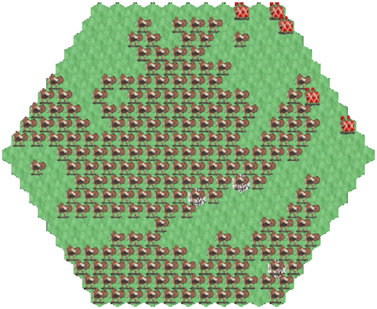
- Each legion has six sensors for each of the six directions
- Can detect immediate adjacency as well as long-distance



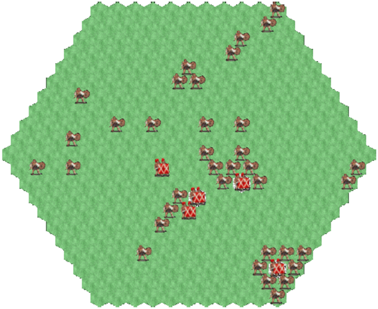


Legion Results

- Legions were trained with ESP neuroevolution
- Successfully learned to divide labor between "guard duty" and search and destroy" behavior



Before training



After training

Questions, Suggestions or Some Other Ideas?

Our Research Project

- Motivations
- Mechanisms
- Suggestions

Further Readings: Books

- W. Michael. *An introduction to multiagent systems*. John Wiley & Sons, 2009.
- S. Yoav, and K. L. Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.
- W, Gerhard, ed. *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press, 1999.

Further Readings: Courses and Labs

- Stanford CS224M: Multi Agent Systems (Spring 2013-14). HERE
- MIT CPSC689: Special Topics in Multi-Agent Systems (Spring 2006). HERE
- Stanford Multiagent Research Group. HERE
- CMU Advanced Agent-Robotics Technology Lab. HERE
- MIT Robust Open Multi-Agent Systems (ROMA) Research Group. HERE

References

[1] Tan, Ming. "Multi-agent reinforcement learning: Independent vs. cooperative agents." *Proceedings of the Tenth International Conference on Machine Learning*. Vol. 337. 1993.