

Evolution of a Communication Code in Cooperative Tasks

Aditya Rawal¹, Padmini Rajagopalan¹, Risto Miikkulainen¹ and Kay Holekamp²

¹The University of Texas at Austin

²Michigan State University

aditya@cs.utexas.edu

Abstract

Communication through vocalizations is used by spotted hyenas and chimpanzees for coordination during hunting and for raising alarm calls in defense (Bullinger et al., 2011; Holekamp et al., 2007). Vocal signals are omni-directional and are therefore more effective than visual communication in these situations. In cooperative tasks, agents use these signals to pro-actively exchange information for common good. A simulated predator-prey domain is considered in this paper - where multiple predator agents exchange real valued messages as an approximation of vocalization in nature. In artificial intelligence, the problem of coordination among multiple predator agents during prey capture is hard because of the non-Markovian environment (Panait and Luke, 2005). Experiments are carried out in this paper to show how information exchange through messaging can make the environment less non-Markovian and improve predator team performance during cooperative hunt. The values of these messages are analyzed to study the emergence of a common communication code among the predator agents. The results in this paper also provide an insight into the constraints under which language evolves in nature.

Introduction

Spotted hyenas employ vocal signals for kin-recognition (Holekamp et al., 2007), chimpanzees use it for building a consensus before embarking on a group hunt (Bullinger et al., 2011) and both hyenas and vervet monkeys (Vaughan et al., 2011) raise vocal alarm calls when under predatory threat. In nature, communication through vocalization plays an important and diverse role especially in scenarios where coordination among group of individuals is required. Coordination in teams translate into multi-agent problems in artificial intelligence (AI) domain. As in the real world, not all the information about each state in the simulated environment is known, so the agents cannot consistently select the optimal action. From the AI perspective, this non-markovian nature of the environment in multi-agent problems is a major challenge in building cooperative teams. Animals alleviate this problem of group coordination through constant and pro-active transfer of information among concerned individuals through various forms of communication - like visual, vocal, tactile

and olfactory. One advantage of vocal signalling over other modalities of communication is the fact that it is omni-directional and can travel long distances. Another interesting characteristic of vocal language is that it is usually consistent among the members of a species. Infants learn this language while growing up, and begin actively participating in societal roles like hunting and protecting.

Inspired by such instances in nature, we simulate teams of cooperative predator agents in prey capture tasks, with the goal of evolving a common communication code among them. As an approximation to the vocal signalling in nature, the predators are provided with continuous channels through which they can send and receive real-valued messages/codes among themselves. Language usually has two aspects to it - conveying meaningful information on the part of the sender and ability of the receiver to interpret this information. The first two experiments in this paper study the constraints under which such a common predator code emerges. The third experiment compares the performance of real-valued messaging with direct communication (an approximation to vision in nature). The predator agents are evolved using Multi-component ESP (a neuroevolution technique), which has previously been found to be successful in such sequential decision making (for prey capture) tasks (Rawal et al., 2010).

Background and Related Work

One of the first simulations in artificial organisms to study the emergence of language was done by (Werner and Dyer, 1992). They used discrete signals to evolve a communication protocol among agents for the task of mate selection. (Saunders and Pollack, 1996) applied both discrete and continuous for food search task and also analyzed the different evolved signals among agents. More recently, (Tuci and Vicentini, 2007) conducted an experiment in a team of 3 robots, where each robot is equipped with different sensors. With limited perception of the world, the robots are forced to cooperate by communicating sensory information among each other. A single controller is cloned and is used for controlling all the three robots in their experiments. (Jim and

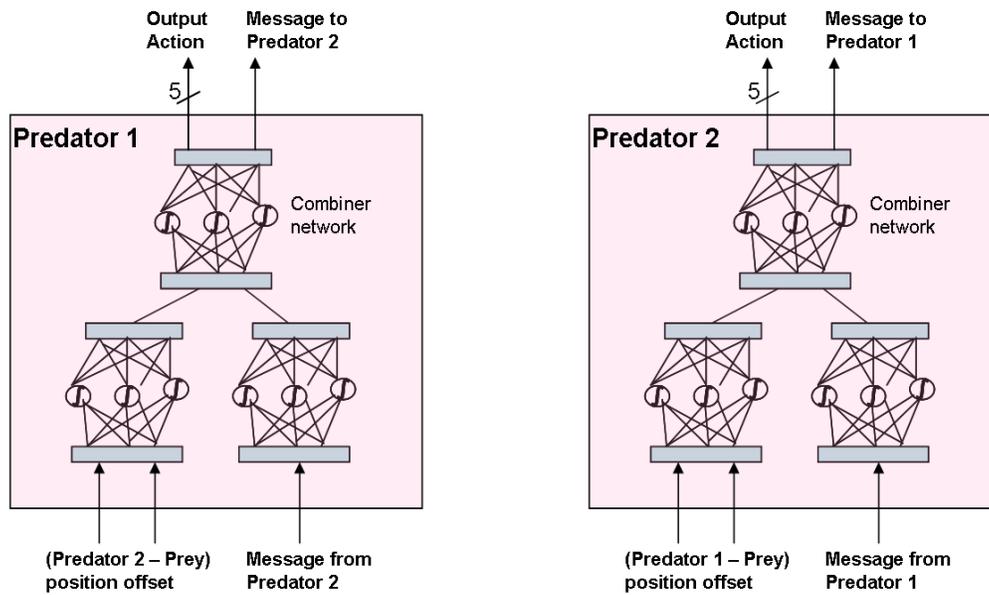


Figure 1: Experiment 1 - Predator agent architecture. Each agent controller is constructed using Multi-component ESP (Rawal et al., 2010). Each predator agent controller has two input sensory neural networks - one for tracking the position offset between the prey and the other predator and second for sensing the real-valued message from the other predator. The outputs of the two input neural networks are combined using a combiner network. The combiner network has 5 output nodes corresponding to 5 possible predator action, and one node for the output message to be sent to the other predator.

Giles, 2001) have used predator-prey domain to show how communicating agents with evolved signalling outperform non-communicating agents. (Knoester et al., 2007) evolved artificial organisms for distributed problem solving through communication. Their experiment demonstrated how information propagates in multi-agent settings.

This paper takes a different approach to study the evolution of communication in artificial agents. It first aims to study the situations under which a common communication code emerges through messaging among a team of evolving predators. We believe that the knowledge of such constraints would help in understanding the evolution of communication in nature. Second, it compares the efficacy of such a team built using messaging as compared to direct communication. Similar to (Saunders and Pollack, 1996), we use real-valued communication channels for signalling among predators.

The predator-prey domain used as a testbed in this paper is a special case of the pursuit-evasion domain. There are predators and prey on the field at the same time and the predators have to capture the prey while the prey try to evade the predators. In these experiments, a team of predators is evolved using cooperative coevolution to capture the prey. The world in this simulation is a discrete toroidal environment with 100 x 100 grid locations without obstacles, where the predators can move in four directions: east, west, north and south. They move one step at a time, and all the agents

take a step simultaneously. To move diagonally, an agent has to take two steps (one in the east-west direction and one in the north-south direction). A predator is said to have caught a prey if it moves into the same location in the world as the prey.

Multi-component ESP (Rawal et al., 2010), a hierarchical cooperative coevolution architecture is used to construct separate controllers for each predator agent. This neuroevolution architecture has previously been successfully used to coevolve a team of predators hunting prey (Rawal et al., 2010; Rajagopalan et al., 2011). It allows for a single agent controller to be composed of multiple networks - where the networks cooperate and their outputs are combined using a combiner network (figure 1). Networks within a controller can be dedicated to different subtasks that the agent must carry out, or for tracking different pieces of information it senses from the environment. Each of these networks is composed of neurons, which represent connection weights for a given node in the network. Each of these neuron is evolved separately in a subpopulation of its own. The final fitness is obtained by constructing a network out of neurons picked randomly from their respective subpopulations and evaluating it in the domain. The fitness received by the network is then assigned to its component neurons. If such a process is carried out several times, selecting neurons at random, each individual neuron's fitness, calculated by averaging over the number of times it was picked, gives a rough

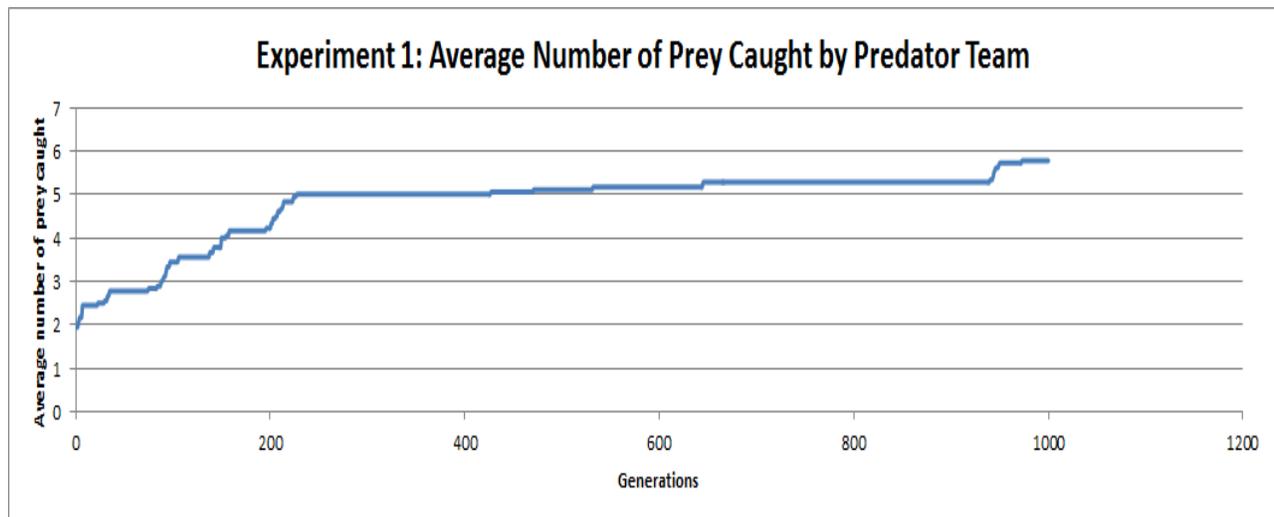


Figure 2: The average number of prey caught by the predator team at every generation in experiment 1. The predators switch between two roles - being 'blind' and mobile in one and stationary (with vision) in the other. In order to successfully capture the prey, the stationary predator should guide the 'blind' mobile predator by sending meaningful messages. The agents thus evolve communication for sustaining cooperation. The results shown are average of 10 runs.

indication of how good that component of the neural network is. This approach breaks up the task into manageable subtasks, thus making the search space smaller, and avoids competing conventions among the neurons.

All the agents on the field share rewards obtained. Reward sharing has been shown to be very effective as an incentive to evolve cooperation among agents (Yong and Miikkulainen, 2009; Rajagopalan et al., 2011). and therefore used here as well.

Experimental Setup

All the predators on the field are evolved using the Multi-Component ESP architecture (Rawal et al., 2010). Three experiments are performed in this paper for prey capture tasks. The goal of the first two experiments is to study the emergence of a communication code among the predators. The third experiment is designed to highlight the utility of this code during cooperative hunting tasks. In all the three experiments, there is a single non-evolving prey in the world. The prey re-appears randomly at a new location once it gets caught. This allows the predators to capture several prey in a single trial/episode. The prey is stationary in the first two experiments, while it moves with a speed of 0.75x in the third experiment. The environment employed is the 100x100 toroidal grid world first used successfully in the predator-prey domain in (Yong and Miikkulainen, 2009). The predators consist of one or more input sensory networks for tracking each bit (either real-valued message or discrete position offsets) of information sensed from the environment. The predators with more than one input networks also include a combiner network which combines the output of

these networks to generate the next predator action and/or predator message. All the predator networks are triggered at every time step of the episode. Each network (including the combiner) has a feedforward architecture with a single layer of 10 hidden neurons and sigmoidal activation functions. Each hidden neuron is evolved in a separate subpopulation consisting of 100 neurons; each neuron is represented as a concatenation of real-valued numbers representing full input and output connection weights.

At the beginning of each generation, 1,000 trials are conducted, and for every trial, a set of neurons is chosen at random from the subpopulations to construct the predators. Each such unique team of predators is evaluated in five simulation runs. All the predators move synchronously, taking one step at each timestep. There are five different actions possible: move up, down, left or right, or remain idle. Each simulation run consists of 500 timesteps in the first experiment, and 300 timesteps in the second and third experiments, during which the predators attempt to catch prey. Each prey gives a reward of 100 points on capture, which is shared equally among the predators. Reward sharing has previously been shown to be effective in fostering the evolution of cooperation in predators (Rajagopalan et al., 2011). The fitness obtained from averaging the total rewards earned over the five runs is then assigned to all the neurons that were used in building these predators.

After the trials, the top 25% of neurons within each hidden neuron subpopulation are selected for recombination. A chromosome is a string of real valued weights associated with each hidden neuron. Since the gene-length of chromosome is fixed, the recombination involves blending real

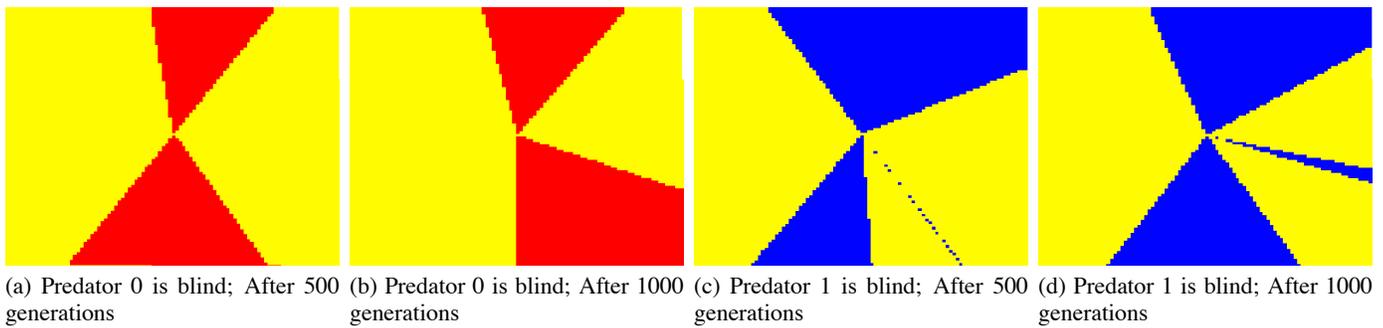


Figure 3: Experiment 1: The best predator team after several generations of evolution is picked for this plot. Each grid cell is painted a color based on the action of the blind predator when present in that cell. The other predator is stationary and fixed at (0,0) and sends messages to the blind predator based upon which it takes action. The prey is also fixed at (50,50). The color coding is as follows- red denotes go down, green denotes go right, blue denotes go up, yellow denotes go left, and black is remain idle. Each agent has evolved to send and receive only two commands - "go down" and "go left" or "go up" and "go left".

valued weights from the same position in the gene using simulated binary crossover (Agrawal and Deb, 1994). The offspring replace the bottom 50% of the neurons in the corresponding subpopulation. Mutation is carried out with a probability of 0.4 on one randomly-chosen weight on each chromosome, by adding a Cauchy-distributed random value to it. Small changes to these parameters lead to similar results.

Experiment 1

An experiment with two predators and a prey was conducted to study the evolution of a common code among predators during hunting. As shown in figure 1, each predator agent has two sensory input networks - one for sensing the offset between the other predator and prey and second to sense the message from the other predator. In order to simplify the analysis of the results, the predators switch between two states - being 'blind' (with no information about prey and other predator position) and being stationary (fixed position). For example, in the first half of an episode, predator 1 is blind i.e it receives no input in its sensory network dedicated for tracking offset between predator 2 and the prey. However, with the messages it receive at every time step from the other predator, it can make decisions to move in the world. During the same time (first half of episode), predator 2 is stationary but it can track the position offset between predator 1 and prey. These predator roles are switched in the second half of the episode. The predators are therefore required to cooperate to successfully catch the prey. One way of doing this is to perfect the system of communication by evolving a code where different real-valued messages represent different pieces of information or commands. The swapping of roles also plays an important part in the evolution of a messaging code. It ensures that both predators evolve not only the ability to send a message in the evolved

code, but also to interpret incoming messages correctly and take subsequent action.

The results of this experiment are given in figures 2 and 3. Figure 2 is a graph of the average number of prey caught by the predator team in each generation. It can be seen that the two predators evolved to catch more than one prey in every episode. Figure 3 was generated by setting one of the evolved predators at location (0,0), and the prey at location (50,50) and setting the blind predator at all the grid cells in the world. The color of any cell in the diagram represents the action taken by the blind predator at that location in the world after receiving a message from the other (stationary) predator. Here, red represents the action "go down", green represents "go right", blue is "go up", yellow is "go left", and black is "remain idle". From figure 3, it can be seen that the blind predators always take one of the two actions - "go right" or "go left", and "go up" or "go down". These actions correspond to the two fixed real values that the stationary predators have evolved to send (not shown here). Evolution has discovered that two commands are enough for prey capture. Although both predators evolved the ability to send and interpret meaningful messages, their communication code is not consistent.

Experiment 2

We have seen from the previous experiment that two predators are capable of evolving a messaging code to communicate useful information between themselves. But the messages sent by the first predator and received by the second may be in a different code from the messages sent by the second predator to be received by the first. It is a logical next step to investigate the circumstances under which different agents can evolve a common code for communication.

To this end, this second experiment was devised, where three predators coevolve to catch prey (see figure 4). In

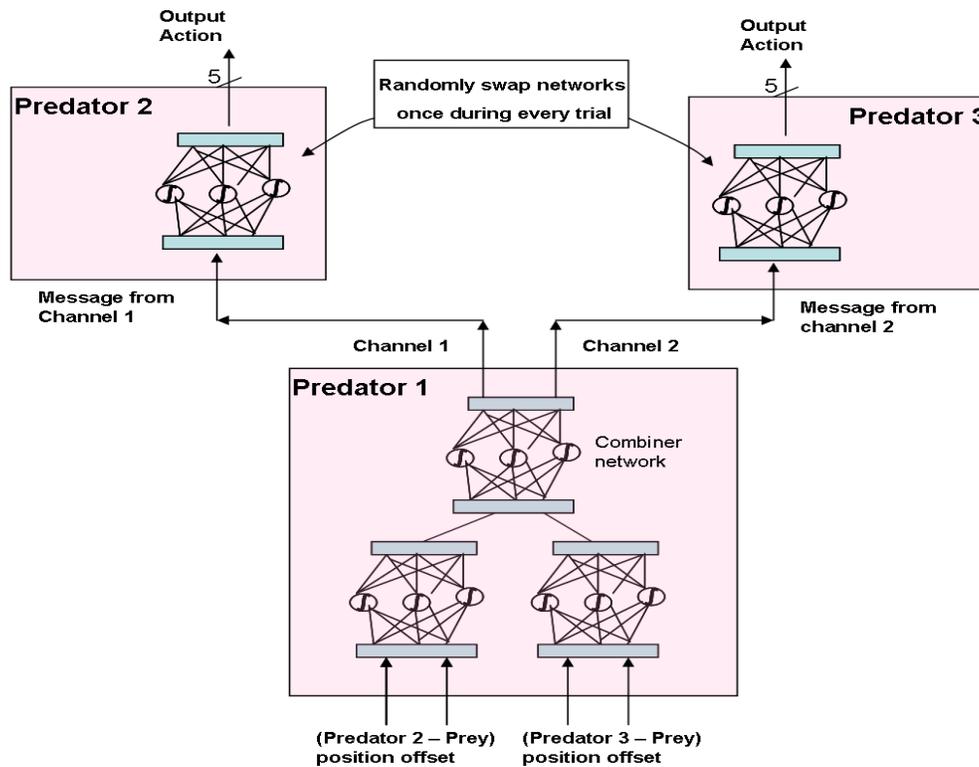


Figure 4: Experiment 2 - Predator agent architecture. Each agent controller is constructed using Multi-component ESP. Predator 1 agent controller has two input sensory neural networks - each for tracking the position offset between the prey and predator 2 and predator 3 respectively. Predator 1 is stationary and has two output channels for sending real-valued messages. Predator 2 and 3 are mobile and 'blind' and by interpreting the messages from predator 1. In order to facilitate evolution of a common communication code between the predators, the neural network controllers of predators 2 and 3 are switched once in every episode.

this case, predator 1 is always stationary and can see the prey and as well as the other two predators. The other two predators (predators 2 and 3) are blind and mobile, receiving real-valued messages from predator 1 to determine their movement. Predator 1 has two output channels of communication, that is, two output nodes to evolve two different messages to send to the two predators. In order to facilitate the two channels to evolve common message codes, a trick is used. The controllers (i.e. the neural networks) of predator 2 and 3 are swapped once in a while so that each network starts receiving messages from the other channel. Since each output communication channel of predator 1 is associated with a particular predator body and location, it will now be sending messages to the either of the two controllers.

This swapping of the neural networks in blind predators will ensure that a single common messaging code has to be evolved for any successful prey capture to occur. This is because each network must evolve to interpret the messages from both channels correctly without knowing which channel it is receiving the message from. Therefore, it evolves to treat messages from both sources as coming from the same

communication protocol. In turn, the stationary predator has to send a meaningful message from each of its channels without knowing how the receiving predator will interpret it, because it does not know which network is receiving it. Thus it will evolve to send messages in the same code from both its channels.

The results of this experiment are in figures 5 and 6. The graph in figure 5 shows that the predator team is successful in catching more than one prey on average in each run. As in the previous experiment, figures 6 were created by placing the evolved predators at certain locations in the grid world and recording the actions they take at each grid cell. The stationary predator was placed at location (0,0) and the prey at location (50,50). Each blind predator was placed at all the locations in the world and its output actions were represented by a color for that cell. As before, red denotes "go down", green denotes "go right", blue denotes "go up", yellow denotes "go left", and black is "remain idle". The input to the predator was a message from one of the two communication channels of the stationary predator. Figure 6a represents the actions of blind predator 1 as a response to

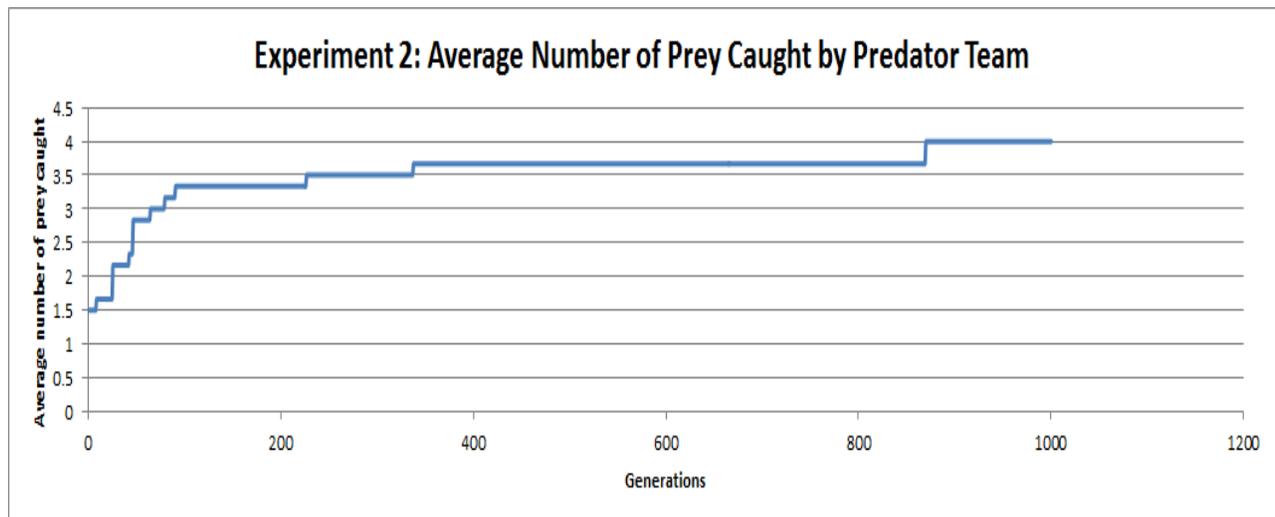


Figure 5: The average number of prey caught by the predator team at every generation in experiment 2. There are two 'blind', mobile predators and one stationary predator with vision. The controllers of the mobile predators are switched once in every episode. The predator has evolved to successfully accomplish the prey capture task.

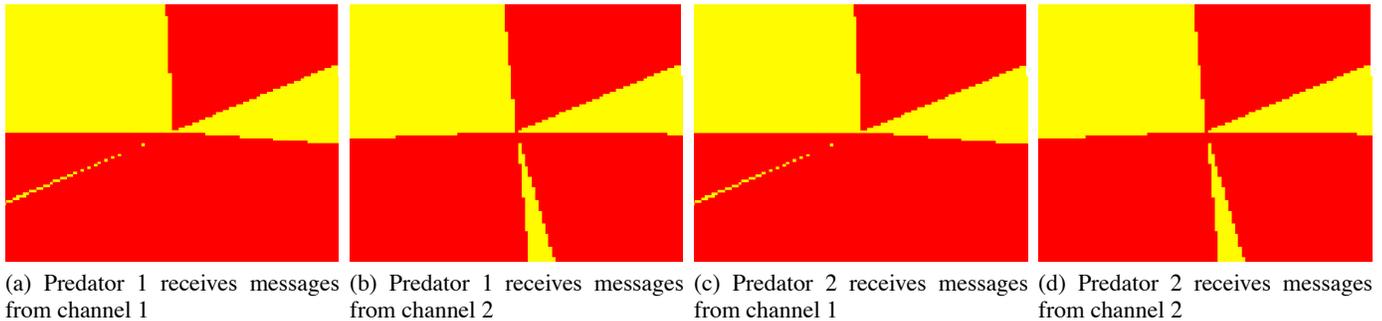


Figure 6: Experiment 2: Each grid cell is painted a color based on the action of the blind predator when in that cell after many generations of evolution. The communicating predator is stationary and fixed at (0,0) and sends messages to the blind predator. The other blind predator is not on the field. The prey is at (50,50). The color coding is same as described in earlier.

communication channel 1, figure 6b corresponds to predator 1 and channel 2, figure 6c corresponds to predator 2 and channel 1, and figure 6d to predator 2 and channel 2. Predator 1 response (action) to channel 1 message is the same as predator 2 response for channel 1. Similarly, the response of both the predators is the same for channel 2 as well.

Experiment 3

After demonstrating the evolution of a common communication code among predators in previous experiments, we now assess its utility. Two experiments, comparing messaging and direct communication, are performed here. Direct communication is analogous to vision in nature - where the predators can observe each other's position. The agent architecture used in both messaging and direct communication is same as that of experiment 1 (Figure 1), however there are changes in the inputs. For messaging, the first

input sensory network of predator 1 tracks the prey position and its second sensory network receives messages from predator 2. Similarly, the two input networks of predator 2 tracks prey position and receives messages from predator 1. For direct communication, each predator tracks the prey position (input network 1) and other predator position (input network 2). The task is made more challenging as the prey now moves in the world (prey speed = 0.75x predator speed) The prey follows a fixed policy of moving away from the nearest predator. The performance comparison of messaging and direct communication is shown in Figure 7. Messaging between predators results in slightly better performance than direct communication. Although the agents evolve a few commands during messaging, these codes are simple and flexible enough for prey capture. The videos for the evolved predator behaviors can be found at: nn.cs.utexas.edu/?alife_2012_communication

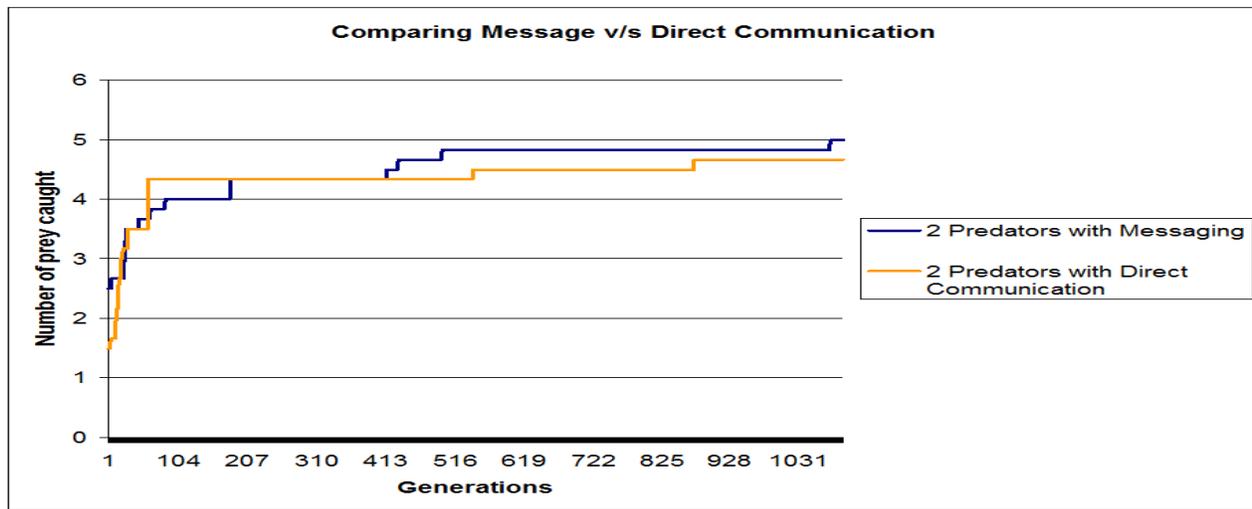


Figure 7: Performance comparison between messaging and direct communication. Two predators cooperate to capture a mobile prey (prey speed = 0.75x predator speed). The agent architecture is similar to the one in Figure 1. In messaging, agents can sense the prey as well as receive real valued signals from each other. In direct communication, the agents sense the prey and each other's position. The results are average of 5 runs. Predators in messaging set-up perform slightly better than direct communication.

Discussion and Future Work

The results of the first experiment (figure 2) demonstrate that the two predators successfully evolved to catch prey. As established in the previous section, one of them was blind and the other was stationary at every point during the simulation. Thus the stationary predator had to evolve a messaging code to communicate its knowledge to the blind predator. And since the two predators exchange their roles during the simulation run, both predators evolve the ability to both send meaningful messages and interpret them correctly. This can be seen more clearly from figure 3, which show the actions taken by the blind predator in response to a message from the stationary predator. The predators seem to evolve only two kinds of actions, one to go left or right and the other to go up or down. Since the simulated world is toroidal, these two actions are sufficient to catch the prey wherever it may be. Two actions are easier to evolve than five, especially when it also involves evolution of an interpreting system to decipher incoming messages. From the figure 3, it can further be seen that there are straight-line borders between the regions where an upward or downward action is taken and the regions where rightward or leftward actions are taken. This shows that the stationary predator has evolved to send particular messages based on the region in which the blind predator is. It is easier to evolve to recognize a large region with straight-line borders than to identify each grid cell with a different message.

In the second experiment, it can be seen from figure 5 that the predator team has been successful in catching prey. As claimed in the previous section, this indicates that a common

messaging code has been evolved. This is confirmed by analyzing the figures 6: 6a is an exact copy of 6c, and 6b is exactly like 6d. That is, both blind predators react in the same way to any given message. For a given location of a blind predator, channel 1 has evolved to send a particular message. This message is interpreted in the same way by both blind predators. This is true in spite of the fact that the two blind predators, upon close examination, have completely different connection weights in their neural networks.

Figure 6a and 6b do not look alike. This is because the two communication channels may have two different roles to play in the prey capture task. For example, one of them may guide its receiving predator to attack the prey from one direction, while the other channel may direct its predator to go after the prey from the other direction. Thus for any given location of a blind predator, the two communication channels of the stationary predator may transmit different messages. But, as discussed above, the two different messages will be from the same messaging code. Such experiments can also provide clues on the evolution of communication in nature - where perhaps organisms with common goal and similar sensory information evolved a common language.

Another important observation is that the agent controllers are evolved from separate sub-population (unlike some of earlier research). This makes the problem of evolution of consistent language more difficult, since the agents do not share any genes.

Messaging performs slightly better than direct communication as shown in experiment 3. In direct communication, the predators can track each other's position accurately and

thus have more information about the environment (Markovian). However, the predators evolve equally good cooperative behaviors through messaging. The predators can convey information (gathered by sensing prey position) about the environment to each other and thus help disambiguate the non-Markovian environment.

The evolution of language in the form of messages can be put to further use in more complex tasks where just sight (direct communication) will not be sufficient to successfully complete the task. The evolution of prey has also been shown in previous work to lead to more interesting and sophisticated behaviors on the part of both predators and prey (Rawal et al., 2010). Scaling messaging to more agents reliably is another challenge for the future.

Conclusion

The evolution of a messaging code for useful communication was successfully evolved in a team of predators using neuroevolution. If there are two predators communicating with each other, they may develop different codes for sending and receiving. To encourage them to evolve a common communication code, three predators were put on the grid world and their channels of communication with one another were frequently switched. This led to the successful evolution of a single messaging code. This approach to evolution of language can be adopted in more complex and open-ended domains, such as the evolution of realistic and interesting video game agents, or robots.

Acknowledgements

This research has been sponsored by following NSF Grants: DBI-0939454, IIS-0915038.

References

- Agrawal, R. B. and Deb, K. (1994). Simulated binary crossover for continuous search space. Technical report.
- Bullinger, A., Wyman, E., Melis, A., and Tomasello, M. (2011). Coordination of chimpanzees in a stag hunt game. *International Journal of Primatology*, 32:1296–1310. 10.1007/s10764-011-9546-3.
- Holekamp, K. E., Sakai, S. T., and Lundrigan, B. L. (2007). Social intelligence in the spotted hyena (*crocuta crocuta*). *Phil Trans R Soc B.*, 362:523538.
- Jim, K. and Giles, C. L. (2001). How communication can improve the performance of multi-agent systems. In *In Proceedings of Autonomous agents'01*, pages 584–591.
- Knoester, D. B., McKinley, P. K., Beckmann, B., and Ofria, C. (2007). Directed evolution of communication and cooperation in digital organisms. In *Proceedings of the 9th European conference on Advances in artificial life, ECAL'07*, pages 384–394, Berlin, Heidelberg. Springer-Verlag.
- Panait, L. and Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11:387–434. 10.1007/s10458-005-2631-2.

Rajagopalan, P., Rawal, A., Miikkulainen, R., Wiseman, M., and Holekamp, K. (2011). The role of reward structure, coordination mechanism and net return in the evolution of cooperation. In *2011 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 258–265.

Rawal, A., Rajagopalan, P., and Miikkulainen, R. (2010). Constructing competitive and cooperative agent behavior using coevolution. In *Proceedings of the IEEE Conference on Computational Intelligence and Games*, pages 107–114. IEEE Press.

Saunders, G. M. and Pollack, J. B. (1996). The evolution of communication schemes over continuous channels. In *In*, pages 580–589. MIT Press.

Tuci, E. and Vicentini, F. (2007). Operational aspects of the evolved signalling behaviour in a group of cooperating and communicating robots.

Vaughan, T., Ryan, M. J., and Czaplewski, N. (2011). *Mammology*. Jones and Barnett.

Werner, G. and Dyer, M. (1992). Evolution of communication in artificial organisms. In Langton, C., Taylor, C., Farmer, D., and Rasmussen, S., editors, *Artificial Life II*, pages 659–687, Redwood City, CA. Addison-Wesley Pub.

Yong, C. H. and Miikkulainen, R. (2009). Coevolution of role-based cooperation in multiagent systems. *IEEE Transactions on Autonomous Mental Development*, 1(3):170–186.