The role of communication in cooperative hunting in a partially observable world

Pieter Bosma P.G.BOSMA@STUDENT.RUG.NL
University of Groningen, The Netherlands

Hein de Haan H.DE.HAAN.8@STUDENT.RUG.NL
University of Groningen, The Netherlands

Aliene van der Veen N.A.VAN.DER.VEEN@STUDENT.RUG.NL
University of Groningen, The Netherlands

Floris Bex F.J.BEX@RUG.NL
University of Groningen, The Netherlands

Harmen de Weerd H.A.DE.WEERD@RUG.NL
University of Groningen, The Netherlands

Keywords: cooperative hunting, multi-agent systems, emergent behavior, neural networks, genetic algorithms

Abstract

In cooperative hunting, multiple agents need to coordinate their behavior to achieve a common goal. Previous research into the emergence of cooperative hunting has found conflicting results with respect to the effectiveness of communication: some authors show that communication can be helpful for groups of artificial agents, whilst others find that communication can also decrease performance. In this paper, we model the emergence of cooperative hunting using neural networks and genetic algorithms. We find that the effectiveness of communication increases with the complexity of the environment (modeled as the vision range of the hunters). That is, while communication between agents slows down the emergence of cooperative hunting in simple environments (*i.e., when vision range is large*), communication can speed up emergence of cooperative hunting in complex environments (when vision range is small).

Preliminary work. Under review for BENELEARN 2014. Do not distribute.

1. Introduction

When designing cooperative multi-agent systems, it is useful to watch the behavior of animals and humans. One type of behavior that has proved interesting is cooperative hunting: a group of animals finds and catches a prey in a complex and dynamical world by using a cooperative strategy. Such cooperative strategies can be interesting for designing behavior of game agents or robot teams.

Nature provides some examples of cooperative hunting behavior: a pack of wolves or hyenas catching their prey. The emergence of cooperative hunting in nature is not fully understood. One possible way to study the emergence of cooperative hunting is to study emergent behavior of artificial groups. The hunting behavior of for example wolves can be described in a few simple rules (Muro et al., 2011). When the world is more complex and adaptive behavior is required, the combination of neural networks and a genetic algorithm is found to be useful to reach emergence of effective cooperative hunting behaviors in artificial teams (Haynes & Sen, 1996; Haynes & Sen, 1997a; Yong & Miikkulainen, 2009).

The main question is: how can cooperative hunting be established best? There are a number of factors that influence the emergence of cooperative behavior, for example, whether the behavior of the different agents is heterogeneous or homogeneous, whether the whole group or individual agents are rewarded when they catch the prey and whether or not the agents communicate about, for example, the location of the prey or their own location.

One possible way to study these different factors is through simulations of hunting games such as the pursuit-evasion task, in which several predators need to catch the prey. This task is introduced by Benda (1985) and is used to study the evolution of cooperative behavior (Cliff & Miller, 1996; Haynes & Sen, 1997b; Jim & Giles, 2001; Yong & Miikkulainen, 2009; Wittkamp et al., 2012; Jain et al., 2012). These studies show the advantages and disadvantages of some design choices. For example, Haynes & Sen (1996; 1997a) used genetic programming to evolve predator population strategies for the prey-capture task and found that groups of heterogeneous agents performed better on the task than groups of homogeneous agents. Yong & Miikulainen (2009) and Rajagopalan et al. (2011) showed that reward works very effectively as an incentive to evolve cooperation among predators.

With respect to the role of communication, Yong & Miikkulainen (2009) conclude that predator agents work better together when they do not communicate and only use cues from the environment (this is called *stigmergy*). However, research has shown (Jim & Giles, 2001; Rajagopalan et al., 2011; Rawal et al., 2012) that communication is useful: Rajagopalan et al. (2011) found that performance increases when predators share their locations. The question that remains is then: when is it useful for autonomous agents to communicate?

Since Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (Rawal et al., 2012) used the same techniques and nearly the same experimental setup, it is possible to compare their results. The main difference between the three studies is the complexity of the environment, which is determined by, for example, the observability and the number of preys. In Yong & Miikkulainen (2009) the environment is fully observable, in the sense that the predators always know the coordinates of the prey. In Rawal et al. (2012) the environment is sometimes fully observable and sometimes not visible for the predators. In Yong & Miikkulainen (2009) and Rawal et al. (2012) the predators need to catch just one prey. In Rajagopalan et al. (2011), there are multiple prey in the environment that the predators could catch. So, it seems that communication is more useful when the complexity of the environment increases.

In this paper, we aim to further verify whether communicative behavior starts outperforming stigmergic behavior. We choose the vision range of the predators as the parameter that reflects the complexity of the environment. Following Yong & Miikkulainen (2009) and Rajagopalan et al. (2011), we formulated the following hypothesis about the evolution of cooperative behaviors in predators: as the complexity of the world increases, the time needed for cooperative hunting to emerge increases more rapidly for non-communicative groups than for communicative groups. In figure 1, this hypothesis is shown in a graph. When the predators communicate, the evolution of cooperative behavior takes more time since the neural networks have to deal with more input information. In a simpler environment this extra information is not useful, so therefore non-communicative predator teams will perform better. In a more complex world communication is useful: the time to establish cooperation is smaller for agents that can communicate.

The rest of this paper is structured as follows. In Section 2 the experiment is discussed: world constraints, used techniques and simulation setup. In Section 3 the results are presented and in Section 4 a final conclusion is given.

2. Methods

We designed an experiment to observe cooperation in predators hunting prey that move randomly. The simulation was based on the pursuit-evasion task introduced by Benda (1985). The goal is to test whether predator communication influences evolution of cooperation as was hypothesized in Section 1.

First, we discuss the simulation environment and its constraints. The next subsection will explain how genetic algorithms and neural networks are used in this experiment. Finally, the setup of the experiment will be discussed.

2.1. The simulation environment

In this study, two different simulation environments were tested: a world with boundaries (a ‘square’ world) and a world without boundaries (a toroidal world). The worlds differ in the number of predators that are needed to catch a prey: in the square world, a prey can be caught with the help of the boundaries. In this situation, two predators can successfully catch the prey by driving the prey into a corner, while in the toroidal world at least four agents are needed. Both environments give rise to different hunting opportunities: the square world allows for an interesting chase towards an edge; the toroidal world, on the other hand,

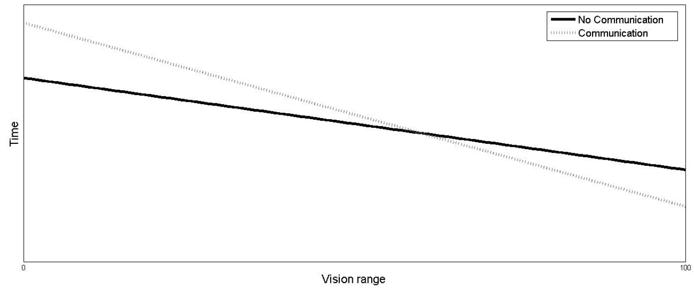

Figure 1. The hypothesized influence of vision range of the predator on performance

allows the group of predators to split, having both groups go in opposite ways to catch the prey from two sides.

In this experiment, more predator agents ($N = 20$) were used than necessary for solving the problem. We call these extra agents *collaborators*. The design of this experiment is different on this point from the experiment design as used in Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (2012). Therefore, this experiment is able to show whether more communication is useful in more complex situations. We think that collaborator agents, who do not directly help to catch the prey, could give their colleagues useful information about the complex environment.

Both worlds have 40×40 grid locations without obstacles. The prey and predators can move in four directions: east, west, north and south. In Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (2012) a 100×100 toroidal world was used. To reduce the computation time, a smaller environment was preferred for this study.

The initial positions of the prey and the predators are chosen randomly. The prey and the predators move one step at a time in one of the four cardinal directions, so to move diagonally, an agent would have to take two steps. Like in Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (2012), all predators and the prey move simultaneously. However, in our study the prey moves randomly instead of moving directly away from the current nearest predator as in Yong & Miikkulainen (2009). The predators have a neural network that will decide the next step. The neural network has four output nodes corresponding to the directions north, south, east and west. The output node with the highest activation will determine the next step.

Next to these constraints, which are constant across all simulations, the environment has some variable constraints concerning the vision range of the predators and communication. The vision range represents the maximum length of the vision field of the predator. If the vision range is 50%, the predator is able to see 50% of the diagonal of the world. A predator knows the distance to the prey when the following rule holds:

$$euclideanDistance(prey, predator) < \frac{visionRange}{100\%} * \sqrt{Width^2 + Height^2}.$$

That is, if the euclidean distance between the prey and the predator is small enough given the vision range and width and height of the world, the prey is observable to the predator. The distances in x-direction and y-direction between the prey and the predator are then given to the neural network of the predator as input. If the euclidean distance is too large, the neural network will be given input zero.

Communication between the predators is modeled as position sharing, that is, a communicating predator knows the distance between him and other predators. Different communication conditions were tested by changing the amount of neighbors that share their position (the variable k). In case of $k = 0$ no predator shares its position with another predator, and the predators can sense only prey movements, so they have to use that for deciding their next step. This is the so-called stigmergic communication situation as described in Yong & Miikkulainen (2009) and Rajagopalan (2011). If $k > 0$, a predator knows the distance between himself and k nearest neighbor predators.

2.2. Neural Network

Every predator agent has its own neural network. The neural network is a simple feed forward neural network that has 2+2*k (k is the number of nearest predators for which a predator knows the position) input nodes, one hidden layer of ten nodes and an output layer of four nodes. For calculating activation a simple sigmoid function is used.

The inputs for the neural network are the distances to the prey and to the k nearest predators. Each distance is encoded in two input nodes: one encodes the x-coordinate of the owner of the neural network (a predator) minus the x-coordinate of the prey or another predator, while the other does the same for the y-coordinates. If the prey is not visible, the inputs for the prey will be zero. The same holds for predators that do not share their position: the input for no-communicating predators will be zero. It is possible that all inputs are zero, since the agent has no knowledge at all at that moment.

In each scenario, the neural network has ten hidden units and four output units: one for the north direction, one for the east direction, one for south, and one for west. The move of a predator in a specific situation is the direction of which the corresponding output node has the highest value of all output nodes.

2.3. Genetic Algorithm

A genetic algorithm is used to optimize the weights of the neural networks. The agent population has N chromosomes, where each chromosome represents an agent. Each chromosome contains the parameters of the neural network of an agent. Each generation, all $N = 20$ agents play the predator-prey pursuit evasion task in the same world. The 16 agents having the shortest Euclidean distance to the prey at the end of the generation are passed on to the next generation.

In Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (2012) the hidden layers of neural networks evolved out of a separate population for each neuron. This technique is called multi-agent Enforced SubPopulations (ESP) (Gomez & Miikkulainen, 1997; Gomez & Miikkulainen, 1999). This allows for the evolution of heterogeneous agents that perform different tasks within the cooperative hunt. Due to the computational demands of ESP, we decided to use a different technique: neural networks were used as brains of the predators and a genetic algorithm to optimize the weights of the networks, to give rise to some cooperative behavior. In this situation there is one population of neural networks. However, since a high number of best agents stay in the population and mutation is applied to the new produced agents, there is space for heterogeneous behavior to evolve.

After each generation, the 16 predator agents with the shortest Euclidean distance to the prey pass on to the next generation. In addition, the two best agents, defined as those agents that have the shortest Euclidean distance to the prey, are selected to construct 4 new predator agents through Uniform Crossover with a mixing ratio of 0.5. That is, these new predator agents inherit each of their individual genes, which in this case correspond to one of the weights in the neural network, from a given parent with probability 0.5.

To introduce more variation in the population, mutation with a probability of 0.1% is applied to each of the four newly produced agent. In other words, to each individual weight (gene) a random value between -0.1 and 0.1 is added with a probability of 0.1%. Introducing mutation prevents getting stuck in a local minimum and ensures that the evolution will continue and new behavior can emerge.

2.4. Experiment setup

Different scenarios were tested with varying the variables for vision range and communication. The vision range ranged from 0% to 100% with a step size of 10%. The size of the communicating neighborhood k varied between $k = 0$ and $k = N - 1$. All combinations of vision range and communicating neighborhood were tested.

For each scenario we used $N = 20$ predators and one prey. We calculated the mean score of each scenario over 100 trials. Each trial ran for a number of generations. During a generation g, the number of time steps needed for the predators to catch the prey ($i_{catch}(g)$) was recorded, with a maximum of 200 time steps. After a catch or when 200 time steps had passed, the agents were randomly replaced and a new generation starts.

When the prey is caught 10 times at $g_{10catches}$ or when $g = 800$, the experiment stops and the score is calculated. The performance score of a trial is computed as the total number of time steps that have passed across all 10 generations.

$$
\begin{aligned}
score &= \sum_{g=1}^{G} (i_{end}(g)) \\
G &= \min(800, g_{10catches}) \\
i_{end} &= \min(200, i_{catch})
\end{aligned}
\tag{1}
$$

3. Results

The performance mean score (mean time for catching the prey 10 times) for the different scenarios are shown in figures 2 (square world) and 3 (toroidal world) for

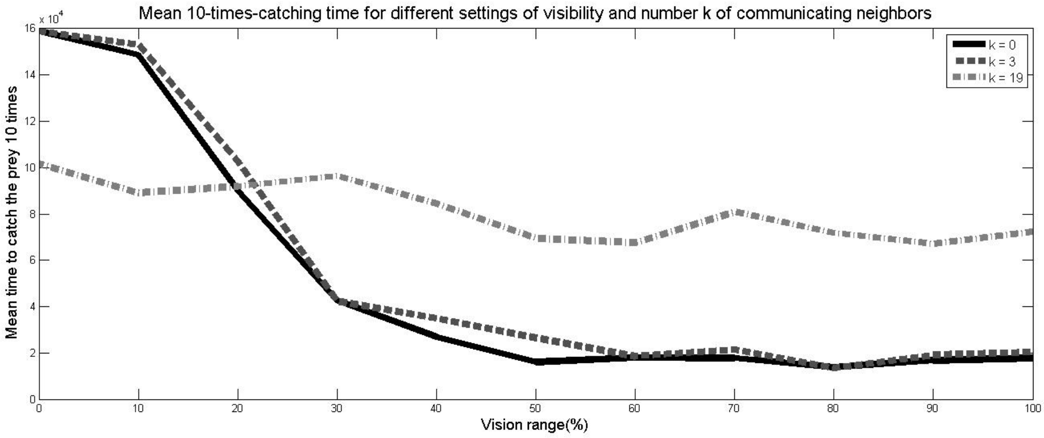

Figure 2. Mean 10-times-catching time in the square world for vision range between 0-100% and k=0 (stigmergic behavior), k=3 and k=19 neighbors. The figure summarizes results over 100 trials.

selected communication parameter k.

First, it is interesting to compare the performance for the stigmergic behavior scenarios ($k = 0$) with the performance for $k = 3$. These scenarios are used as well in Yong & Miikkulainen (2009), Rajagopalan et al. (2011) and Rawal et al. (2012). Yong (2009) concluded that $k = 3$ has a lower performance than $k = 0$, but (Rajagopalan et al., 2011) and (Rawal et al., 2012) concluded that communication makes it easier to evolve coordinated hunting behavior. Figures 2 and 3 show that $k = 0$ outperforms $k = 3$ for all different vision ranges except for a vision range of zero in the toroidal world. An ANOVA for repeated measures for performance score on k, found a significant difference between $k = 0$ and $k = 3$ for the toroidal world($F(1,2196) = 12.13$, $p < 0.05$), but not for the squared world ($F(1,2196) = 2.032$, $p = 0.15$). Therefore stigmergic behavior outperforms in both simple situations (high vision range) and more complex situations (low vision range). This is in line with our conclusion of Yong (2009), but not in line with the conclusion of Rajagopalan et al. (2012) and Rawal et al. (2011).

To see whether communication with more than a reasonable amount of agents needed to solve the problem is useful, we compare the performance for $k = 0$ and $k = 19$. In the square world the performance scores for $k = 0$ and $k = 19$ intersect each other at a vision range of 20-25%. For higher vision range the performance for $k = 0$ is higher, for lower vision ranges the performance for $k = 19$ is higher. In the toroidal world the performance scores for $k = 0$ and $k = 19$ is almost similar for higher vision range. For a vision range below 25%, the performance for $k = 19$ is much better. This is in line with the hypothesis.

The results for all values of k can be found in Figure 4(square world) and Figure 5 (toroidal world). We find that a higher value of k is generally more useful in complex situation, when the vision range is below 30%. Furthermore we find that a difference between k and $k + 1$ is small, but the difference between k and $k + 3$ or more, influences the result much more.

4. Conclusion and discussion

The goal of this experiment was to test whether communication is necessary in complex situations and whether in simple situations the problem can be solved faster by just stigmercy. Given the results as reported by Yong & Miikkulainen (Yong & Miikkulainen, 2009), it was expected that for simple situations (high vision range) stigmergic behavior outperforms communication. Given the results in (Rajagopalan et al., 2011) and (Rawal et al., 2012), it was expected that communication outperforms stigmergic behavior in more complex situations.

The results of this experiment show that stigmergic behavior outperforms communication in simple situations, but not in complex situations. However, for good performance in complex situations, communication with more predators than needed to solve the problem is needed. In this experiment, it was not enough to communicate with just three other neighbors. Therefore this experiment reinforces the conclusion of Yong & Miikkulainen (Yong & Miikkulainen, 2009) that stigmercy alone works out better than com-

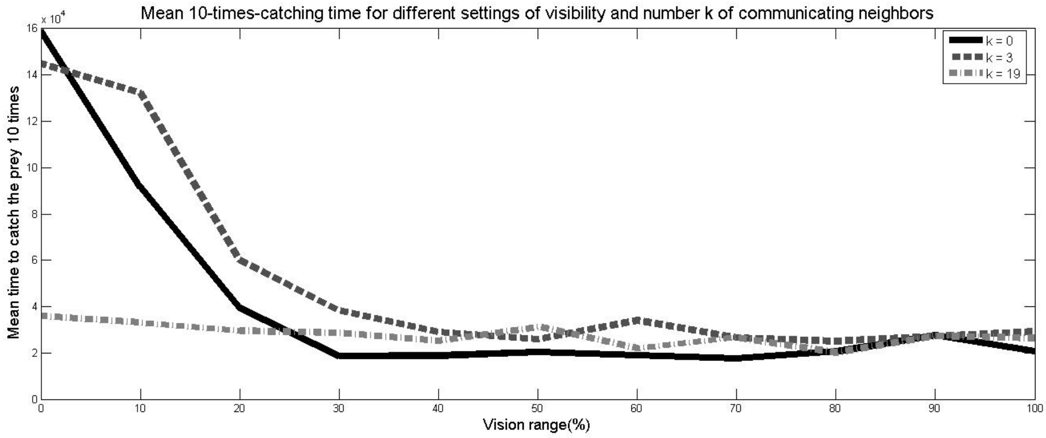

Figure 3. Mean 10-times-catching time in the toroidal world for vision range between 0-100% and k=0 (stigmergic behavior), k=3 and k=19 neighbors. The figure summarizes results over 100 trials.

munication with 3 neighbors.

When communicating with $k = 19$ neighbors in the toroidal world, the vision range seems not to matter at all. This is an interesting result. We think that the predators in this scenario found an optimal strategy without knowing the location of the prey. For future research it would be interesting to do the experiment again with a prey that always moves away from the predators and to verify whether there could still emerge an optimal strategy for 20 collaborators.

We did the experiment in two different worlds: a square world with boundaries and a toroidal world without boundaries. We saw that communication was much more helpful in the toroidal world. Probably, the agents need the positions of their colleagues to navigate through the environment. To see what the influence is of the extra added collaborators, more research needs to be done. This experiment should be done again for $N = 4$ in a toroidal world.

Acknowledgments

We would like to thank the three anonymous reviewers for their useful comments.

References

Benda, M., Jagannathan, V., & Dodhiawalla, R. (1985). *On optimal cooperation of knowledge sources* (Technical Report BCS-G2010-28). Boeing AI Center, Boeing Computer Services, Bellevue, WA.

Cliff, D., & Miller, G. F. (1996). Co-evolution of pursuit and evasion ii: Simulation methods and results. *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior (SAB96).*

Gomez, F., & Miikkulainen, R. (1997). Incremental evolution of complex general behavior. *Adaptive Behavior*, *5*, 317–342.

Gomez, F. J., & Miikkulainen, R. (1999). Solving non-markovian control tasks with neuroevolution. *IJCAI* (pp. 1356–1361).

Haynes, T., & Sen, S. (1996). Evolving behavioral strategies in predators and prey. In *Adaption and learning in multi-agent systems*, 113–126. Springer.

Haynes, T., & Sen, S. (1997a). Crossover operators for evolving a team. *Genetic programming*, *199*, 162–167.

Haynes, T. D., & Sen, S. (1997b). Co-adaptation in a team. *International Journal of Computational Intelligence and Organizations*, *1*, 231–233.

Jain, A., Subramoney, A., & Miikulainen, R. (2012). Task decomposition with neuroevolution in extended predator-prey domain. *Artificial Life 13* (pp. 341–348).

Jim, K.-C., & Giles, C. L. (2001). How communication can improve the performance of multi-agent systems. *Proceedings of the Fifth International Conference on Autonomous Agents* (pp. 584–591).

Muro, C., Escobedo, R., Spector, L., & Coppinger, R. (2011). Wolf-pack hunting strategies emerge from simple rules in computational simulations. *Behavioural Processes*, *88*, 192–197.

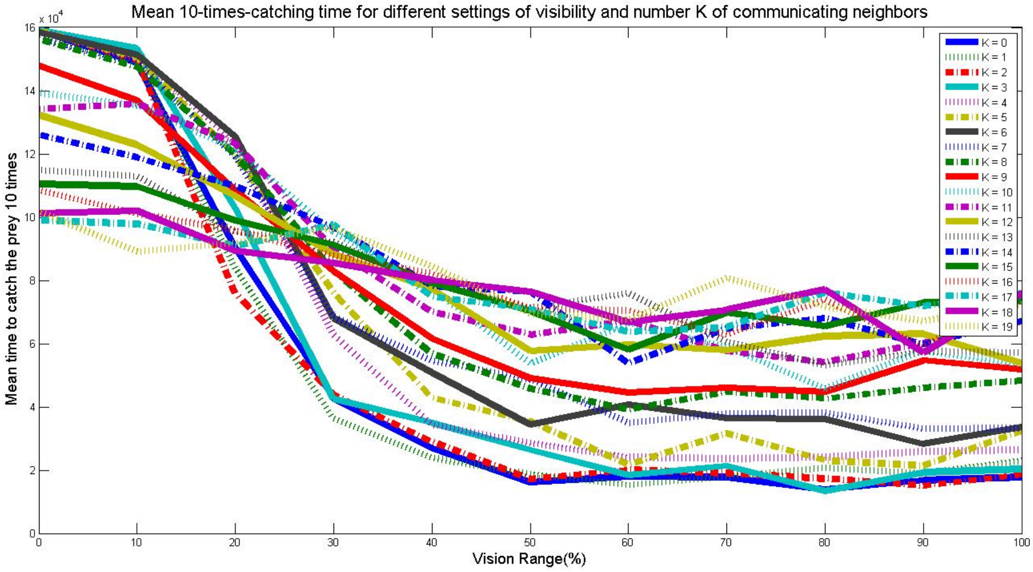

Figure 4. Mean 10-times-catching time in the square world for vision range between 0-100% and k=0-19 neighbors. The figure summarizes results over 100 trials.

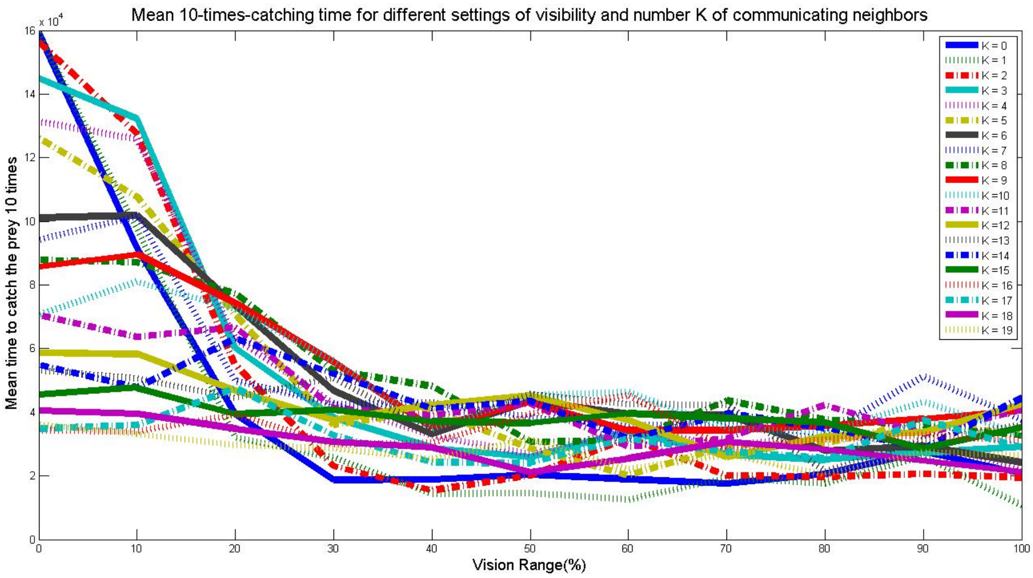

Figure 5. Mean 10-times-catching time in the toroidal world for vision range between 0-100% and k=0-19 neighbors. The figure summarizes results over 100 trials.

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

Rawal, A., Rajagopalan, P., Miikkulainen, R., & Holekamp, K. (2012). Evolution of a communication code in cooperative tasks. *Artificial Life 13* (pp. 243–250).

Wittkamp, M., Barone, L., Hingston, P., & While, L. (2012). Real-time evolutionary learning of cooperative predator-prey strategies. *Proceedings of the Thirty-fifth Australasian Computer Science Conference-Volume 122* (pp. 81–90).

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