Towards Evolving Cooperative Behavior with Neural Controllers*

István Fehérvári and Wilfried Elmenreich
Mobile Systems Group, Lakeside Labs
Institute for Networked and Embedded Systems
University of Klagenfurt, Austria

1 Introduction

A self-organizing system achieves a global system behavior via local interactions between its entities without centralized control [1]. In other words, the entities, or agents, have to cooperate in order to achieve a global result. The problem of cooperation can be observed in different domains and on many levels, especially abundant in nature. It essentially boils down to the conflict of interest between a group of individuals and between the individual itself. In social studies and experimental economics, public goods game proved to be a standard approach to investigate cooperation in various conditions and parameters[2][3]. The basic idea of this game is that a group of subjects starting with a given number of tokens secretly choose how many of their private tokens to put into the public pot. Each subject keeps the remaining tokens plus an even split of the tokens in the pot. Usually, the pot undergoes some function (e.g., multiplied by with a factor) before it gets redistributed to encourage cooperation. In the simplest instance of this problem the following behavior can be observed: the participants start with cooperation but in the following rounds the overall donated money tends to zero. This phenomena has been extensively studied in literature proposing new methods like sanctioning between the members of a group when antisocial behavior is observed [4].

In this paper we examine a different version of the public goods game providing us a suitable testbed to evolve neural controllers as players while monitoring their willingness for cooperation.

2 Experimental Setup

As mentioned above we used a different version of the public goods game where the total number of players is 6 and each player is aware of the previous actions of the others so they are able to adjust their strategy according to the others. We implemented and tested two different revenue functions, one where each donation of a player is tripled in the pot and one where the revenue is proportional to the quadratic amount of tokens in the pot.

In our simulation the players were controlled by a fully-connected, recurrent artificial neural network (ANN) with 2 hidden neurons, 7 input and 1 output neurons, which is executed twice in order to allow for propagation of input signals over hidden neurons. The weights and biases of network were trained by a genetic algorithm with a population size of 90. The candidates are randomly allocated into groups of 6 players where they play the game for 10 rounds. In rounds 2 to 10, the players have information about the behavior of their own action and their colleagues in the previous round, so they can react accordingly. These rounds are scheduled 10 times for each player, each with different sets of players; the performance (or "fitness") of a player calculates by the average revenue over all iterations.

* This work was supported by the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant KWF 20214—18128—26673 within the Lakeside Labs project DEMESOS.

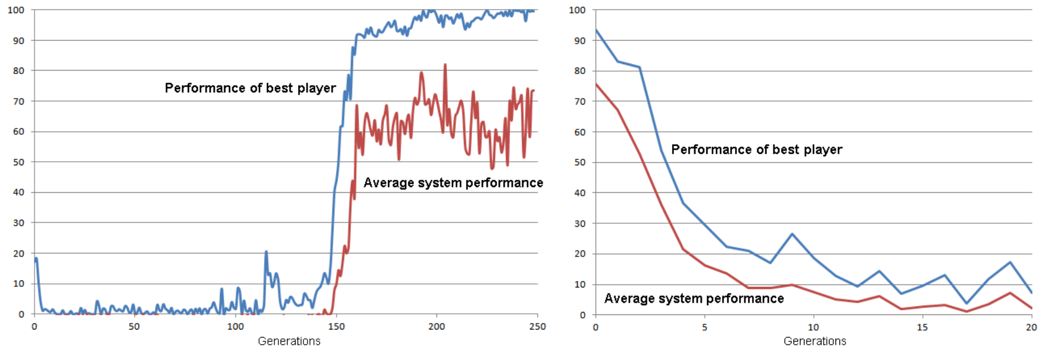

Fig. 1. Performance of the best player and the average system performance. Left: evolution of cooperation with non-linear pot function. Right: loss of cooperative behavior due with linear pot function

Clearly, in the first round there is no such information to feed to the network so all the inputs will be set to 0 except the last neuron which indicates the beginning of the round. We monitored the performance of the best player and the system's average over several hundred generations.

3 Results, Conclusion and Future Work

First we used only the linear pot function starting with randomly created ANN controllers. Here we experienced that is advantageous for a players to be non-cooperative regardless of the strategy of other players. Thus, the system has a very strong attractor[5] of selfish strategies. Even if the system is initialized with cooperative players (Fig. 1 right), the agents' behavior falls back to defection within few generations.

As a second run we used the non-linear pot function with the same setup as before which resulted into an initial overall defection however, after a hundred generations the best and average performance dramatically increased and sometimes reached the possible maximum (total cooperation) (Fig. 1 left).

Thus, the evolution of cooperation is strongly influenced by the design of the reward function. If there is a synergetic effect in cooperation, the system has an attractor towards full cooperation. However, an initial situation with a lot of non-cooperative players represent a Nash equilibrium, since changing the strategy to cooperation does not pay off, if there are too few cooperators. Thus, the system needs sufficient random influence (noise) in order to leave the range of this attractor. Further work will include a closer investigation on the influence of parameters and reward functions on possible cooperation.

References

1. W. Elmenreich and H. de Meer. Self-organizing networked systems for technical applications: A discussion on open issues. In J.P.G. Sterbenz. K.A. Hummel, editor, *Proceedings of the Third International Workshop on Self-Organizing Systems*, pages 1–9. Springer Verlag, 2008.
2. Marco Janssen and T. K. Ahn. Adaptation vs. anticipation in public-good games. In *Annual meeting of the American Political Science Association*, August 2003.
3. R. Mark Isaac, James M. Walker, and Arlington W. Williams. Group size and the voluntary provision of public goods : Experimental evidence utilizing large groups. *Journal of Public Economics*, 54(1):1–36, May 1994.
4. Nikos Nikiforakis. Punishment and counter-punishment in public goods games: Can we still govern ourselves? Experimental 0403001, EconWPA, March 2004.
5. F. Heylighen, J. Bollen, and A. Riegler. *The Evolution of Complexity*. Springer, 1999.