Specialization with NeuroEvolution in a Collective Behavior Task

D. W. F. van Krevelen
Department of Computer Science
Vrije Universiteit Amsterdam
De Boelelaan 1081a
1081 HV Amsterdam, The Netherlands
krevelen@cs.vu.nl

ABSTRACT

In Nature, behavioral specialization is ubiquitous. Groups benefit from complementary and specialized behaviors in individuals, especially in tasks requiring collective behavior. We apply four multiagent NeuroEvolution approaches to such a task: *Enforced SubPopulations* [5], *Parallel* and *Coevolutionary Enforced SubPopulations* [16] and *Collective NeuroEvolution* [11]. Rather than just single controllers we evolve teams of simulated robots to search an unexplored area and gather certain object types for collective construction of a specific sequence. Teams are composed of agents that may evolve from initially *homogeneous* behavior into *specialists* that effectively complement each other. Results show that CONE outperforms in the collective behavior task when assisted with target behavior heuristics for lifetime learning to speed up the search. Some evolved specialists however become what we call all-rounders, taking on some more tasks to compensate for their lack in number.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents

General Terms

Algorithms

Keywords

NeuroEvolution, Collective Behavior, Specialization

1. INTRODUCTION

Increasingly, research is directed into finding good combinations or *teams* of controllers with behaviors. NeuroEvolution (NE) has also been investigated as a means of developing coordinated behavior in robot teams [4]. Yet most of this work on multi-agent learning is based on homogeneous teams, while the potential benefits of division of labor [13] through diversity or specialization have been investigated in only a few studies in neuroevolution. Furthermore, most NE methods lack measures to exchange learned experience between individuals in a multiagent collective such that specialization may still emerge as an efficient solution to collective behavior tasks. One recent approach that does facilitate and utilize specialization is the recently developed *Collective NeuroEvolution* (CONE) method introduced by [11].

This study aims to show that CONE facilitates emergent specialization so as to achieve higher levels of task performance comparative to similar methods in a collective behavior task. We compare performance of four NE approaches in a specially designed gathering and collective construction (GACC) task that combines the foraging and pursuit-evasion tasks. Here, simulated rovers are to search and carry simple objects towards a drop zone located on a hill. Challenges in reaching this goal are that (i) the objects must be returned in a particular sequence to achieve the task's goal, that is, to construct one complex object, and that (ii) carrying objects uphill and delivering them takes at least two agents (hence they must occasionally find and assist each other).

Besides evolving fixed, reactive behaviors that must suit an entire lifetime, there is also an opportunity for the agents' neural network controllers to learn during their lifetime but only if this training is *supervised* using some target rules of behavior. For complex tasks such as GACC this approach is a bit unusual because the target behaviors reduce the flexibility and creativity inherent in evolution and it requires understanding of the specific task domain. However, early experiments with lifetime learning showed dramatic performance increases in the limited time available to us, making differences in algorithmic advantages more apparent. Furthermore, the lack of behavioral variety due to lifetime learning of a specific target behavior increases visibility of what effect (lack of) specialization has on collective task performance.

HYPOTHESIS 1. *Agent (controller) specialization improves collective behavior task performance in the* GACC *task.*

HYPOTHESIS 2. CONE *yields a higher higher collective behavior task performance comparative to similar methods used for evolving controllers to accomplish the* GACC *task.*

2. MULTIAGENT NEUROEVOLUTION

NE methods create artificial neural networks (ANNs) for a wide variety of applications by evolving the network's hidden neurons, its topology, or both. The algorithms compared here all descend from the *Symbiotic Adaptive Neuro-Evolution* or SANE method [10] which evolves a population of neurons from which it constructs a neural network that has

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-131-6/08/07.

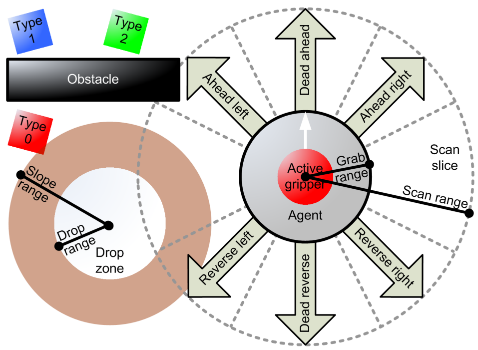

Figure 1: Task domain definitions.

a fixed, single-layer topology. The more advanced *Enforced SubPopulations* or ESP method [5] evolves neurons within their own subpopulations which greatly improves the ability to evolve and combine neurons that each specialize in particular features or subtasks. *Collective NeuroEvolution* or CONE [11] extends ESP to multiple agents and allows gene transfer between them based on a Genetic Distance metric that measures the genetic similarity of their neuron subpopulations.

Here we compare CONE with three multiagent variants of ESP, which are: cloned or MONO-ESP, parallel or POLY-ESP and coevolutionary or TEAM-ESP [16]. These methods were selected as they are very similar to CONE, yet they lack particular features regarding the ability to efficiently evolve a cooperative team of controllers. That is, MONO-ESP evolves just one controller for all agents in the team thus producing homogeneous behavior, POLY-ESP evolves each team member separately but focuses only on individual performance, while TEAM-ESP also evolves each agent separately, this time focusing on team performance (cooperative coevolution), but agents cannot learn from each other as they can with CONE. The CONE method was selected given the hypothesis that it encourages emergent specialization and such specialization is beneficial for the GACC task. MONO-ESP (coevolved clones) and POLY-ESP (individualists) however may evolve specialist behaviors but without an efficient division of labor for successful cooperation. TEAM-ESP on the other hand will evolve successful collective behaviors, but finding them will take more effort as the method lacks CONE's ability of gene transfer among agents.

Experiments performed to answer these questions require a common base for comparison, preferably based on open source standards to allow future reproduction and extension. Unfortunately no such framework was available that could accommodate the GACC task domain, provide the selected NE algorithms and link them together by means of agents that also allow lifetime learning. To have a common framework, the NeuroEvolution Simulation Toolkit[1] was constructed from scratch combining well-known open-source packages and standards such as the *Evolutionary Computation for Java* toolkit [8] for the neuroevolution algorithms, the *Multi-Agent Simulation* toolkit [7] for the agent task domain, and the WEKA data mining toolkit [15] for the neural network lifetime learning algorithm.

[1]http://gforge.cs.vu.nl/projects/nest/

3. THE COLLECTIVE BEHAVIOR TASK

Agents are placed in a square environment and start around a central drop point. Their task is to search and collect objects $\mathcal{O}$ of different types $\mathcal{A}$ placed at random in an unexplored environment before returning and dropping them again within a certain range of the drop point. As shown in Figure 1, the circular drop-zone resides on a hill which agents can negotiate fine when empty. However, carrying objects uphill and onto the drop-zone requires assistance from another agent to help push or pull them up the slope.

The objects $o \in \mathcal{O}$ must be gathered in a particular order as defined by the current scenario. The order concerns a sequence of object types $a \in \mathcal{A}$. Type 0 is distributed around the drop point while all other types > 0 are distributed in the environment's corners. The reward $r \in [0, 1]$ for moving and delivering objects is calculated as the ratio of distance objects traveled towards the drop-zone, multiplied by the ratio of (in order) deliveries $\mathcal{O}_{delivered} \subset \mathcal{O}$:

$$r = \frac{\sum_{o \in \mathcal{O}} d(o)}{\sum_{o \in \mathcal{O}} d_{\max}(o)} \times \frac{|\mathcal{O}_{delivered}| + 1}{|\mathcal{O}| + 1} \quad (1)$$

where $d(o)$ denotes the distance between an object and the drop-zone's edge.

Shaping or incremental evolution is applied to the task of finding controllers for agents that collectively deliver objects in the correct order. Similar to other NE research [5, 16, 3] the task is decomposed into several difficulty levels which the NE algorithm can increase upon completion, guiding it towards good solutions. The following conditions must be met to complete each level: (0) each bot gathered one object of any type, no obstacles; (1) gathered one object of each type, no obstacles; (2) gathered one object of each type while avoiding obstacles; (3) gathered half the sequence in order while avoiding obstacles; and (4) gathered the entire sequence in order while avoiding obstacles.

To control their sensors and motors, agents use an artificial neural network with sigmoid aggregation functions in a topology of one hidden layer and two connection layers. Agents can adapt their behavior during a trial (lifetime learning) by adapting perceptron weights (53 per neuron) each iteration using a predefined target behavior (see Figure 2) and the BACKPROPAGATION algorithm [12, 14].

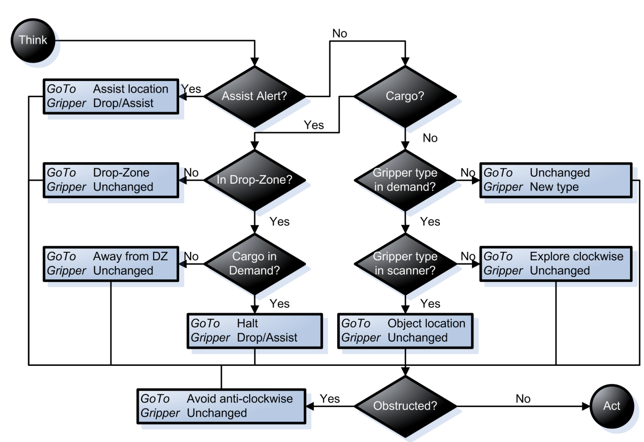

Figure 2: Target behavior for lifetime learning.

Method	*N*	*I. Best Fitness*	*II. Difficulty Level*	*III. Action Entropy*	*IV. Degree of Specialization*	*V. Neurons*	*VI. Burst Mutations*
Randomized							
N/A	53	3.4% ± 0.1%	1.0 ± 0.0	$0.0000G + 0.6783$	$-0.0000G + 0.0004$	$0.0268G + 9.6032$	$0.0664G - 1.0028$
Reactive							
MONO-ESP	11	10.7% ± 3.5%	0.6 ± 0.2	$0.0012G + 0.1779$	$-0.0012G + 0.4131$	$0.0013G + 8.7279$	$0.0712G + 0.1905$
POLY-ESP	16	8.8% ± 1.2%	0.3 ± 0.6	$-0.0005G + 0.4189$	$0.0004G + 0.0784$	$-0.0040G + 9.6615$	$0.0813G - 1.2219$
TEAM-ESP	29	7.5% ± 1.0%	0.1 ± 0.1	$0.0002G + 0.2713$	$0.0002G + 0.2017$	$-0.0112G + 9.2324$	$0.0850G - 0.3551$
CONE	27	8.1% ± 0.7%	0.1 ± 0.1	$0.0003G + 0.2459$	$-0.0000G + 0.3390$	$-0.0048G + 5.0701$	$0.0877G + 0.0562$
Adaptive							
MONO-ESP	12	14.9% ± 3.1%	1.0 ± 0.3	$0.0003G + 0.5038$	$-0.0001G + 0.1363$	$0.0008G + 9.3352$	$0.0679G + 0.2035$
POLY-ESP	27	12.2% ± 1.0%	0.3 ± 0.2	$-0.0004G + 0.4612$	$0.0001G + 0.2500$	$0.0115G + 9.3161$	$0.0723G - 1.2169$
TEAM-ESP	14	10.3% ± 2.0%	0.2 ± 0.1	$-0.0007G + 0.4569$	$0.0003G + 0.3174$	$0.0096G + 8.3262$	$0.0750G + 0.2297$
CONE	50	13.0% ± 1.2%	0.6 ± 0.2	$0.0001G + 0.3183$	$-0.0003G + 0.4241$	$0.0066G + 5.5104$	$0.0775G - 0.0108$
Lamarckian							
MONO-ESP	12	42.6% ± 11.0%	1.3 ± 0.3	$-0.0002G + 0.4565$	$-0.0001G + 0.1988$	$-0.0064G + 7.4195$	$0.0838G - 0.0820$
POLY-ESP	28	32.3% ± 4.5%	2.6 ± 0.3	$-0.0000G + 0.4603$	$-0.0003G + 0.2170$	$0.0029G + 8.2605$	$0.0755G - 1.7044$
TEAM-ESP	37	31.3% ± 2.8%	2.4 ± 0.4	$0.0005G + 0.4688$	$-0.0005G + 0.1854$	$-0.0051G + 8.4171$	$0.0792G - 0.8925$
CONE	14	33.3% ± 4.8%	3.5 ± 0.3	$0.0007G + 0.5396$	$-0.0001G + 0.0420$	$0.0137G + 5.4645$	$0.0634G - 1.3839$

Table 1: Results with 95% confidence intervals at $G = 240$ generations or linear trends for $120 \leq G \leq 240$.

4. EXPERIMENTS

We applied the MONO-ESP, POLY-ESP, TEAM-ESP and CONE methods to evolve robot controller teams for accomplishing the GACC task in three learning setups: Reactive, Adaptive and Lamarckian. *Reactive* agents have fixed behaviors and only improve through the slow evolutionary process. *Adaptive* agents adapt their behavior during a trial so NE is finding individuals that learn quickly, a process biologists call the Baldwin Effect after [1]. *Lamarckian* agents pass on lifetime experience directly through their genomes, a process not actually found in Nature but suggested by [6] and fashionable before Darwin's *Origin of Species* [2] and Mendel's laws of inheritance [9]. Table 1 presents results after $G = 240$ generations, averaged over $N \geq 10$ runs. Using time series data from all runs, we performed regression trend analysis and determine the Pearson product-moment correlation coefficients.

Correlation pair / *Expected*	*I/III* (−)	*II/III* (−)	*I/IV* (+)	*II/IV* (+)
Reactive				
MONO-ESP	-0.10	-0.33^{**}	-0.23^{**}	-0.05
POLY-ESP	-0.67^{**}	-0.42^{**}	0.53^{**}	0.13^{*}
TEAM-ESP	-0.73^{**}	-0.29^{**}	0.73^{**}	0.35^{**}
CONE	-0.47^{**}	-0.39^{**}	0.26^{**}	0.62^{**}
Adaptive				
MONO-ESP	-0.30^{**}	-0.36^{**}	-0.55^{**}	-0.52^{**}
POLY-ESP	-0.73^{**}	-0.49^{**}	0.21^{**}	0.29^{**}
TEAM-ESP	-0.53^{**}	-0.67^{**}	0.46^{**}	0.43^{**}
CONE	-0.52^{**}	-0.59^{**}	0.46^{**}	0.49^{**}
Lamarckian				
MONO-ESP	0.44^{**}	0.29^{**}	-0.83^{**}	-0.73^{**}
POLY-ESP	-0.03	-0.06	-0.68^{**}	-0.68^{**}
TEAM-ESP	0.28^{**}	0.57^{**}	-0.68^{**}	-0.88^{**}
CONE	0.37^{**}	0.74^{**}	-0.80^{**}	-0.85^{**}

$^{*}p < 0.05$ $^{**}p < 0.01$

Table 2: Performance/specialization correlations.

Hypothesis 1 states that agent specialization improves collective task performance in the GACC task. This should be reflected by coefficients of correlation in Table 2 between collective task performance (*I. Best Fitness* or *II. Difficulty Level*) and agent specialization (*III. Action Entropy* or *IV. Degree of Specialization*). Correlations I/III and II/III are expected to be negative as performance should drop when action unpredictability (III) rises, whereas correlations I/IV and II/IV are expected to be positive as performance should increase when specialization (IV) also increases.

According to the values in Table 2, *Reactive* and *Adaptive* agents perform as expected, although in the MONO-ESP cases we see opposite correlations in *Degree of Specialization* (I/IV and II/IV). These MONO-ESP runs showed increased performance when agents became all-rounders, switching more (decreasing the *Degree of Specialization*) between fewer jobs (decreasing *Action Entropy*). *Lamarckian* agents show results opposite from *Reactive* and *Adaptive* agents. Given these results, we can accept Hypothesis 1 for *Reactive* and *Adaptive* cases, but must reject Hypothesis 1 for *Lamarckian* cases where agents perform significantly better when their actions are less predictable and less focused.

Hypothesis 2 states that CONE outperforms similar methods in the GACC task. That means *Best Fitness* and especially *Difficulty Level* should be significantly higher for CONE. Table 1 shows that for *Reactive* and *Adaptive* agents CONE is outperformed only by MONO-ESP and performs about the same as POLY-ESP or TEAM-ESP (see Figure 3(a), 3(b), 3(c) and 3(d)). Slightly higher POLY-ESP performance values with *Reactive* agents are caused by a single lucky run that evolved a team which completed all levels. With *Adaptive* agents, CONE's advantage over POLY-ESP and TEAM-ESP becomes somewhat better, but MONO-ESP with its small search space still outperforms CONE (see Figure 3(e), 3(f), 3(g) and 3(h)). For *Lamarckian* agents, CONE reaches a significantly higher *Difficulty Level* while *Best Fitness* is not significantly higher (see Figure 3(i), 3(j), 3(k) and 3(l)) indicating agents did not move objects much further but cooperated better to make deliveries. We can thus accept Hypothesis 2 for the *Lamarckian* cases.

5. CONCLUSIONS

Results indicate that the CONE method supported by heuristics derives a higher collective behavior task performance in the GACC task comparative to that of the multiagent variants of ESP. The MONO-ESP method for deriving homogeneous teams is sufficient for simple task levels that require little cooperation, whereas for tasks that require cooperative and coordinated behavior coevolutionary methods TEAM-ESP and CONE were more appropriate. Their support for heterogeneity among team members allowed more complementary and complex roles to emerge. With the CONE method, teams were able to learn target behaviors more quickly thanks to CONE's gene transfer capability.

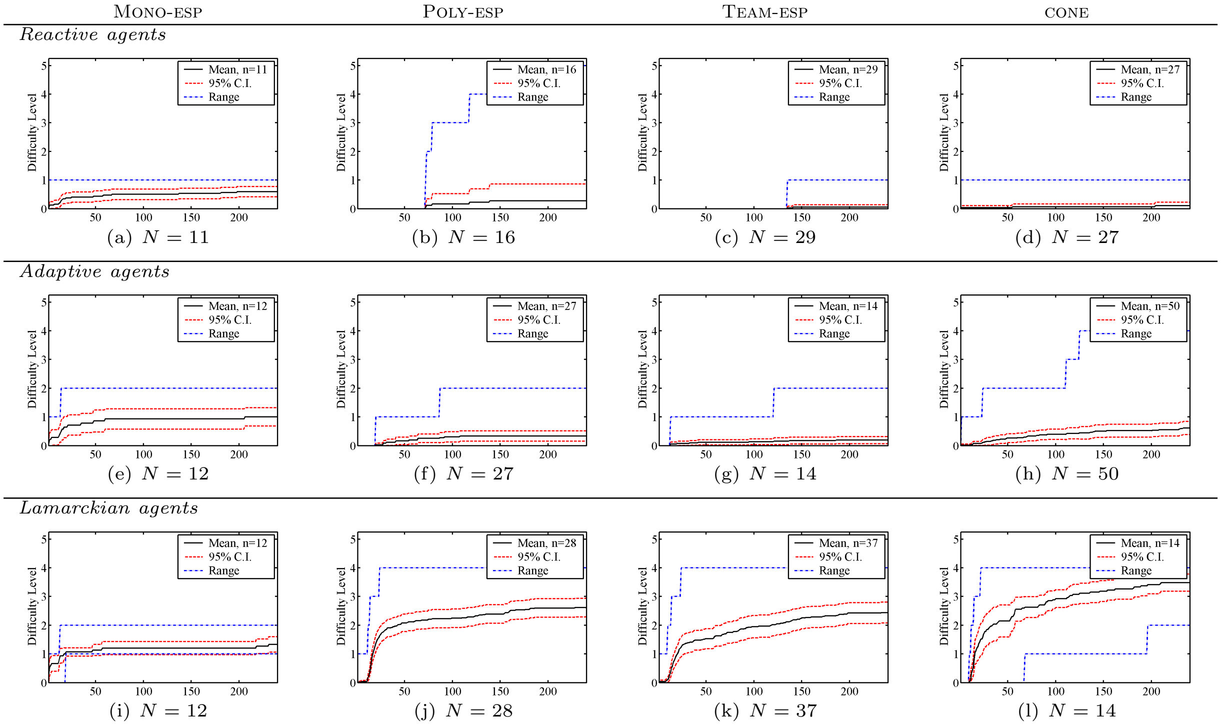

Figure 3: Difficulty level results for first $G = 240$ generations.

6. REFERENCES

[1] M. J. Baldwin. A new factor in evolution. *The American Naturalist*, 30(354):441–451, Jun. 1896.

[2] C. Darwin. *On The Origin of the Species by Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London, UK, 24 Nov. 1859.

[3] F. J. Gomez. *Robust Non-linear Control through Neuroevolution*. PhD thesis, University of Texas, Austin, TX, USA, Aug. 2003.

[4] F. J. Gomez, D. Burger, and R. P. Miikkulainen. A neuroevolution method for dynamic resource allocation on a chip multiprocessor. In *IJCNN'01: Proc. 2001 Int'l Joint Conf. on Neural Networks*, volume 4, pages 2355–2360. IEEE Computer Society Press, Jul. 2001.

[5] F. J. Gomez and R. P. Miikkulainen. Incremental evolution of complex general behavior. *Adaptive Behaviour*, 5(3-4):317–342, Jun. 1997.

[6] J.-B. Lamarck. *Philosophie zoologique*. 1809.

[7] S. Luke, C. Cioffi-Revilla, L. Panait, K. Sullivan, and G. Balan. MASON: A multiagent simulation environment. *Simulation*, 81(7):517–527, Jul. 2005. `http://www.cs.gmu.edu/~eclab/projects/mason/`.

[8] S. Luke, L. Panait, G. Balan, S. Paus, Z. Skolicki, J. Harrison, J. Bassett, R. Hubley, and A. Chircop. ECJ 16: A java-based evolutionary computation research system, Jan. 2008. `http://cs.gmu.edu/~eclab/projects/ecj/`.

[9] J. G. Mendel. Versuche über pflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, 4 Abhandlungen:3–47, 1866.

[10] D. E. Moriarty and R. P. Miikkulainen. Evolutionary networks for value ordering in constraint satisfaction problems. Technical Report AI94-218, University of Texas, Austin, TX, USA, May 1994.

[11] G. S. Nitschke, M. C. Schut, and A. E. Eiben. Collective specialization in multi-rover systems. In D. Thierens, H.-G. Beyer, M. Birattari, J. Bongard, J. Branke, J. A. Clark, D. Cliff, C. B. Congdon, K. Deb, B. Doerr, T. Kovacs, S. Kumar, J. F. Miller, J. Moore, F. Neumann, M. Pelikan, R. Poli, K. Sastry, K. O. Stanley, T. Stützle, R. A. Watson, and I. Wegener, editors, *GECCO'07: Proc. 9^{th} Genetic and Evolutionary Computation Conf.*, pages 342–342, London, England, UK, Jul. 2007. ACM Press.

[12] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. *Learning internal representations by error propagation*, volume 1 of *Parallel Distributed Processing*. MIT Press, Cambridge, USA, 1986.

[13] A. Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Methuen and Co., Ltd. 1904, London, United Kingdom, fifth edition, 9 Mar. 1776.

[14] P. J. Werbos. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, Cambridge, MA, USA, 1974.

[15] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann Publishers Inc., San Francisco, 2nd edition, 2005. `http://www.cs.waikato.ac.nz/ml/weka/`.

[16] C. H. Yong and R. P. Miikkulainen. Cooperative coevolution of multi-agent systems. Technical Report AI01-287, University of Texas, Austin, TX, USA, 2001.