Automatic Evolution of Multimodal Behavior with Multi-Brain HyperNEAT

Jacob Schrum
Dept. of Mathematics and Computer Science
Southwestern University
Georgetown, TX 78626 USA
schrum2@southwestern.edu

Joel Lehman
Center for Computer Games Research
IT University of Copenhagen
Copenhagen, Denmark
jleh@itu.dk

Sebastian Risi
Center for Computer Games Research
IT University of Copenhagen
Copenhagen, Denmark
sebr@itu.dk

ABSTRACT

An important challenge in neuroevolution is to evolve multimodal behavior. Indirect network encodings can potentially answer this challenge. Yet in practice, indirect encodings do not yield effective multimodal controllers. This paper introduces novel multimodal extensions to HyperNEAT, a popular indirect encoding. A previous multimodal approach called situational policy geometry assumes that multiple brains benefit from being embedded within an explicit geometric space. However, this paper introduces HyperNEAT extensions for evolving many brains without assuming geometric relationships between them. The resulting Multi-Brain HyperNEAT can exploit human-specified task divisions, or can automatically discover when brains should be used, and how many to use. Experiments show that multi-brain approaches are more effective than HyperNEAT without multimodal extensions, and that brains without a geometric relation to each other are superior.

Keywords

Indirect Encoding; Modularity; Multimodal Behavior

1. INTRODUCTION

Success in many domains requires agents capable of complex multimodal behavior, i.e. agents able to switch between distinct policies based on environmental context. This paper uses HyperNEAT [4] to evolve neural network brains for agents. However, HyperNEAT is extended to produce Multi-Brain HyperNEAT (MB-HyperNEAT)[1], which allows a single agent to have multiple brains.

2. BACKGROUND AND EXTENSIONS

HyperNEAT evolves Compositional Pattern Producing Networks (CPPNs) to specify connectivity patterns across indirectly-encoded *substrate* networks. The substrate must be embedded within a geometric space (Figure 1a). HyperNEAT can generate arbitrarily large networks, theoretically capable of generating multimodal behavior. However, an easier way to realize multimodal behavior is with several smaller networks, rather than with a single large one.

[1]Download at southwestern.edu/~schrum2/re/mb-hyperneat.html

GECCO'16 Companion, July 20–24, 2016, Denver, CO, USA.

ACM ISBN 978-1-4503-4323-7/16/07...$15.00

DOI: http://dx.doi.org/10.1145/2908961.2908965

Situational policy geometry [1] is an existing approach that generates multiple brains. Agents have distinct brains for different situations, but they share an underlying geometric relationship due to an additional *situation* input in the CPPN (Figure 1b). A human must decide which brain to use in each situation. These restrictions limit the applicability and effectiveness of the method.

This paper presents three main extensions to HyperNEAT to produce MB-HyperNEAT. Each idea is inspired by the direct-encoded Modular Multiobjective NEAT [3]: (1) The network structure from Multitask Learning is adapted to create multitask CPPNs (Figure 1c). In this case, a human must still specify when each policy is used. (2) Preference neurons [3] are added that allow evolution to discover when an agent should use each brain (Figure 1d). In this way, a different brain can be active on each time step, allowing evolution to autonomously discover an effective task division. (3) Module mutation operators [3], enable the automatic creation of new modules. Module mutation is a structural mutation operator that adds a new output module to a network. If brain substrates have preference neurons, then each application of module mutation to a CPPN creates a new brain substrate that the agent can arbitrate between using preference neurons.

3. EXPERIMENTS

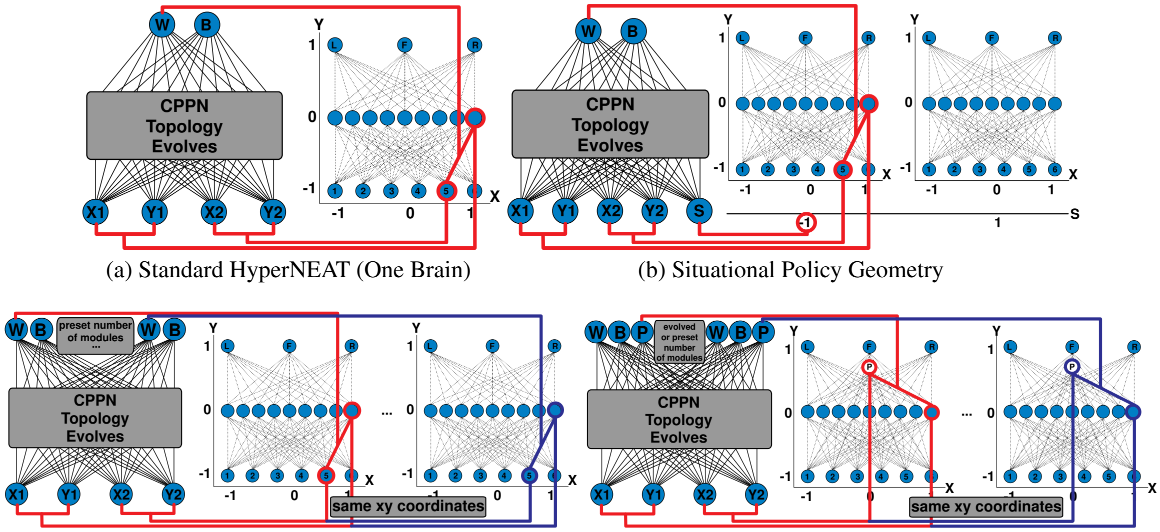
Experiments are conducted in two previous multimodal domains (team patrol and dual task) and two new ones (lone patrol and two rooms). These domains all use simulated Khepera robots.

Team patrol was originally used to demonstrate the effectiveness of situational policy geometry [1]. A team of robots must advance into a maze, and then return to their starting point. The new lone patrol domain uses the same environment, but requires a single robot to explore each part of the maze instead of a team. Dual task [2] consists of two isolated tasks, hallway navigation and foraging, with no clear geometric relation between the behaviors required in these tasks. The new two rooms domain requires hallway navigation and foraging comingled in a single environment: Two foraging rooms are separated by a convoluted hallway.

Standard HyperNEAT, which has only one module (`1M`), provided a performance baseline. Situational policy geometry (`SPG`) and multitask (`MT`) approaches had multiple modules using a human-specified task division. Approaches not depending on such task divisions include CPPNs with two (`2M`) and three (`3M`) preference modules, and CPPNs using different forms of module mutation: `MM(P)`, `MM(R)`, or `MM(D)`.

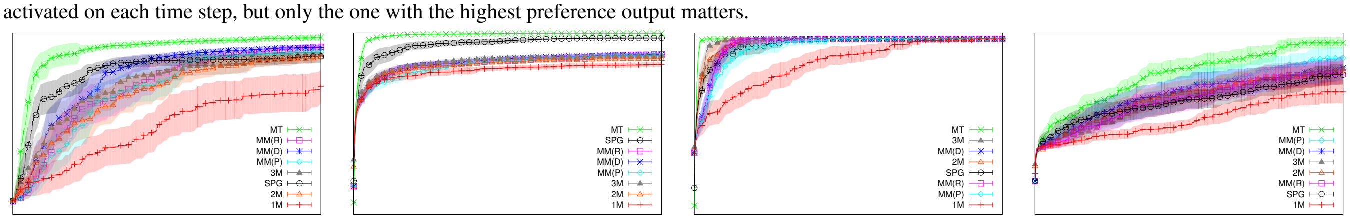
4. RESULTS

The results show that multimodal approaches discover better behavior faster than `1M`, and that multitask and preference neuron ap-



(a) Standard HyperNEAT (One Brain)
(b) Situational Policy Geometry
(c) Multitask
(d) Preference Neurons

Figure 1: Methods for Generating Substrate Brains. (a) A standard HyperNEAT CPPN creates a single-brain substrate [4]. For each possible connection, the xy-coordinates of neurons are input into the CPPN. The W output determines its weight (B determines a fixed bias). (b) Situational policy geometry [1] gives CPPNs an extra S input. There is a separate brain for each value of S the CPPN is queried with. The next two approaches are new to this work: (c) A multitask CPPN has a group of outputs for each brain (no S dimension). When the CPPN is queried, each output module supplies the corresponding connection weight for a different brain. (d) Preference neuron CPPNs add a P neuron to each output module that is used only when the postsynaptic neuron is a preference neuron. The preference neuron (white) of each brain has potential connections to all neurons of the hidden and input layers. All brains are activated on each time step, but only the one with the highest preference output matters.



(a) Team Patrol Results
(b) Lone Patrol Results
(c) Dual Task Results
(d) Two Rooms Results

Figure 2: Experimental Results Across Domains. Average champion fitness by generation (2,000 generations) across 30 runs of evolution for each approach in each domain. Transparent regions show 95% confidence intervals. Having multiple brains leads to better levels of performance faster.

proaches can evolve skilled multimodal behavior without any notion of situational policy geometry.

In all domains, the Kruskal-Wallis test indicates a significant difference between approaches ($p < 0.01$). In the dual task, time to reach the maximum score is compared rather than final score because all approaches reach the maximum. Except in two rooms, post hoc testing indicates that compared to all other methods, `1M` is significantly worse, and `MT` significantly better ($p < 0.01$). In two rooms, `MT` is significantly better than all but `MM(P)` ($p < 0.05$), and only `SPG`, `MM(P)`, and `MT` are significantly better than `1M` ($p < 0.05$). Though at least one preference neuron approach is always better than `SPG`, it is never *significantly* better. Fitness scores across generations are in Figure 2.

In each domain, observation of evolved behaviors reveals why certain approaches are superior, and how multiple modules are used by the modular approaches. Videos of representative behaviors are available at southwestern.edu/~schrum2/re/mb-hyperneat.html. A full-length version of this paper with further details, analysis, and discussion is also available.

5. CONCLUSION

This paper combines the HyperNEAT indirect encoding and the MM-NEAT approach to evolving modular networks, thereby realizing the strengths of both approaches. Results show that the resulting MB-HyperNEAT approach outperforms a previous attempt to merge HyperNEAT with multimodal extensions, and that it is possible to evolve modular ANNs when a human-specified task division is unavailable. The conclusion is that MB-HyperNEAT is a promising toolkit for evolving complex multimodal behavior that can reduce the need for specialized domain knowledge.

6. REFERENCES

[1] D. B. D'Ambrosio, J. Lehman, S. Risi, and K. O. Stanley. Task Switching in Multirobot Learning Through Indirect Encoding. In *International Conference on Intelligent Robots and Systems*, 2011.

[2] S. Risi and K. O. Stanley. An Enhanced Hypercube-Based Encoding for Evolving the Placement, Density and Connectivity of Neurons. *Artificial Life*, 18(4):331–363, 2012.

[3] J. Schrum and R. Miikkulainen. Discovering Multimodal Behavior in Ms. Pac-Man through Evolution of Modular Neural Networks. *TCIAIG*, 8(1):67–81, 2016.

[4] K. O. Stanley, D. B. D'Ambrosio, and J. Gauci. A Hypercube-based Encoding for Evolving Large-scale Neural Networks. *Artificial Life*, 15(2):185–212, 2009.