category: **Artificial Life**

Coupling Morphology and Control in a Simulated Robot

Craig Mautner
Computer Science and Engineering Dept.
University of California, San Diego
La Jolla, CA 92109
cmautner@cs.ucsd.edu (619)453-4364

Richard K. Belew
Computer Science and Engineering Dept.
University of California, San Diego
La Jolla, CA 92109
rik@cs.ucsd.edu (619)534-2601

Abstract

The history of natural evolution displays an inseparable coupling of organic bodies and the nervous systems that control them. In contrast to this almost all research in Evolutionary Robotics to date begins with a robot body whose features are fixed and proceeds to evolve a control structure for this body. Our research program is focused on exploring the coupled evolution of both the body and the control structure in real robots. In this paper we take early steps toward this goal by exploring the space of sensor and effector selection and positioning coupled with a neural network linking them within a simulated environment. This space is explored using evolved grammars for generating both the body and neural network. Results from several problem worlds are presented and analyzed.

1 INTRODUCTION

Evolutionary Robotics is a very new field dating back only to the early 1990's. The vast majority of work in this field has explored evolving control structures for fixed robotic platforms. Taking this approach is justified when the purpose of the research is to explore learning algorithms solely, in which case the body can be considered to be part of the environment/task that is to be learned. This approach is also understandable when the availability of off the shelf robots with embedded controllers is contrasted with the difficulty of constructing multiple unique robots.

The research described below investigates the interaction between co-evolved body and control structures. Elsewhere we describe our model in more detail [13]. We have also began to apply these techniques to real robots [14].

In general, our approach uses Genetic Algorithms to evolve grammars which simultaneously define both the body plan and Neural Network (NNet) control structure of our agents. Grammars provide a framework in which structures can be easily encoded and reused as building blocks of larger, more complex structures. Grammars also offer a parallel to biological development in which start symbols equate to gametes; the execution of rules corresponds to cell divisions; and a derived string of the grammar matches a developed body. We represent and evolve these grammars using Kammeyer's methods [7].

Many aspects of this research have been investigated in isolation by others. Examples of evolved robots that have implemented NNet controllers include [17], [9], [5], [6], and [1]. Researchers who have investigated applying grammatical models to the construction of feed-forward NNets include [8], [17] and [11].

Very little prior work in evolving morphology exists. [15] investigated the nature of sensor usage by providing their agents with an evolvable NNet connected to sensors and effectors. [12] explored eye types and positioning. [16] demonstrated a simulation where the complete morphology of the individuals was involved. Sims created an artificial world in which each agent was grown from a genome that defined both the physical structure and the control structure. [4] has developed an evolutionary system that simulates the growth of a body based on differential gene expression. [3] and [10] have also worked on evolving both control structures and body plans.

Our work most closely resembles [10] in its exploration of body and control space. However our work differs in two significant ways. First of all, Lee use independent evolutionary pathways for the body (GA) and the control structure (GP). Our approach uses a com-

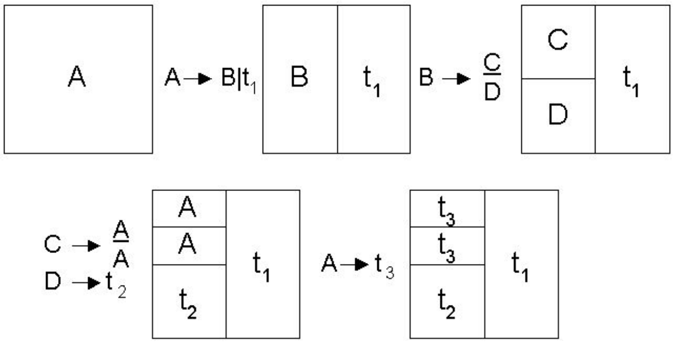

Figure 1: Development Process. Initially each cell begins as a gamete which is labeled with the starting non-terminal symbol of the grammar. Cell differentiation proceeds by selecting and applying the rules of the grammar. For each cell labeled with a non-terminal, a rule is found whose left side matches the non-terminal. The cell is then replaced with the one or two cells specified by the right side of the rule. Each cell is either labeled with a non-terminal symbol or is a terminal cell. The process continues replacing non-terminals with terminals and non-terminals until there are only terminals left. If there is no matching rule or the number of divisions exceed 6, the cell is replaced with a terminal cell with no weights or edges.

mon mechanism to explore the body space and control space as a unified whole. Secondly their bodies and control structures are directly specified by their genetic representation while our approach evolves an intermediate representation, the grammar.

2 Developmental Model

The grammar for an agent consists of a set of rules for rewriting non-terminal cells into terminal and non-terminal cells. Each agent starts as a single undifferentiated non-terminal cell and through repeated application of grammar rules is transformed into a body and NNet consisting only of several, tiled terminal cells. A derivation that takes four generations is shown in Figure 1. In this example the gamete is labeled with the starting symbol, A. The production rule $\mathrm{A} \rightarrow \mathrm{B}|\mathrm{t}_1$ indicates that the non-terminal A is converted into two cells. The first cell, a non-terminal B, is to be placed to the left of the second which is the terminal cell, t_1.

A production rule of the grammar specifies how to replace a non-terminal cell with one or two terminal and non-terminal cells. The production rule also contains the orientation of the cell (i.e. whether it is to be horizontally or vertically flipped or left in its normal orientation) and it specifies whether the non-terminal produces one or two cells, and if two, the relative position (i.e. above, below, to the left of, or to the right of) of each. In general we indicate non-terminal cells by labeling them with the capital letters A-Z.

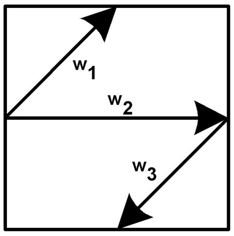

Figure 2: Terminal Cell Example

A terminal cell is a set of directed, weighted edges from the sides of the cell to other sides. A typical terminal cell is shown in Figure 2. There are sixteen potential edge weights for each terminal cell: eight edges from a side to an adjacent side, four edges from each side to the opposite side and four bias weights that can be associated with each side. Each weight can be one of 512 values (-8 to $7\frac{31}{32}$ in $\frac{1}{32}$ increments). The number of unique terminal cells is therefore $512^{16} = 2^{144}$.

2.1 NEURAL NETWORK INTERPRETATION

Once the cell division is complete, the body consists of a set of cells that have within them directed, weighted edges. The cells and edges are interpreted as sensors, effectors and the neural networks that connect them. Edges of a terminal cell are transformed into NNet edges through simple rules. These rules merge edges that point to the same terminal cell sides into a common NNet node that sums and squashes (using the tanh function) their output activation. When two cells abut one another, the nodes formed by the edges in one cell provide the input activation to the neighboring cells.[1]

Any directed edge that originates from the perimeter of the body becomes a sensor or input node. Sensor nodes detect signals of the environment. They provide the input that is propagated through the NNet of the body. Their activation is proportional to $\frac{1}{d^2}$ where d is the distance to the signal source.

Any directed edge that terminates on the perimeter of the body becomes an effector or output node. Effectors provide propulsion to the agent's body. The

[1]This connectivity permits many NNet architectures including recurrent connections which are handled by stochastic updating methods.

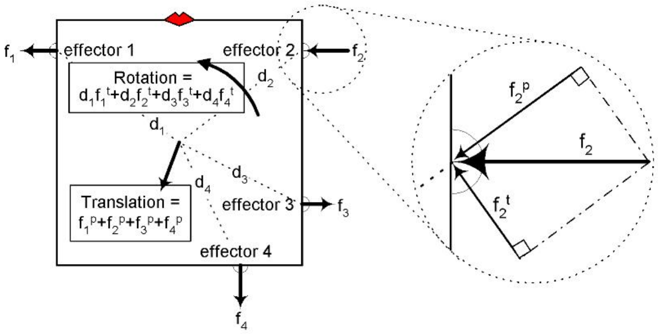

Figure 3: Conversion of Effector Outputs to Rotation and Translation. Effector outputs $\vec{f_i}$ are broken into two force components. Translational forces ($\vec{f_i^p}$) pass through the center of the agent and are vector-summed to produce a new location of the agent. Rotational forces ($\vec{f_i^t}$), perpendicular to the translational forces are multiplied by their distance to the center and then summed to produce a new angular orientation for the agent.

force of this propulsion is proportional to their output activation. The direction of propulsion of the agent is described in Figure 3.

3 Experimental Environments

Our experiments consisted of evolving agents in several simulated environments. Each evolution consisted of 1000 generations of populations of 200 individuals. Each agent is assigned a fitness score based on its performance in an environment. Those agents with higher score are preferentially selected via a random process for inclusion in the next population. The agents are generated from a grammar which is represented by a string of characters. Prior to evaluation of the population the character strings encoding the grammars are crossed with one another with a probability of 0.8 and each character is randomly mutated with a probability of 0.001. The environments consisted of either toroidal or walled worlds of dimension 500x500 units. The agents size within the world are 20 units on an edge.

The agents are placed at random locations in the world and allowed to wander freely (or to just sit in the case of many agents) for up to 30 time steps after which they are moved to a new location. All agents in a given generation are started from the same set of locations. The set of locations are changed at each generation. Each agent is evaluated for a total of 300 time steps.

3.1 A SIMPLE CENTER-SOURCE ENVIRONMENT

Our first experiments evolved agents that would approach a source of reward placed in the center of a toroidal world. The center was detectable by the agents' sensors whose input values fell off as the inverse-square of the distance from the sensor to the center of the world.

The fitness, f_i, at time step i is $\frac{1}{d^2}$ for $d \geq 28$ and 1 for $d < 28$ where d is the distance from the agent's mouth to the center of the world. In this environment, if an agent's mouth gets within 28 units of the center of the world then it is given a fitness of 1 for that time step and immediately moved to a new location.

A perfect agent would turn it's mouth toward the center of the world and approach the center in as few time steps as possible.

3.1.1 EVOLUTION FROM BRAITENBERG AGENTS

A classic design in the field of robotic control is the Braitenberg Vehicle 2b described in [2]. This agent has two sensors on the front and two effectors on the back. The agent's body is bi-laterally symmetric with each sensor connected via a positive weight to the effector on the opposite side. The effect of this connectivity is to steer the agent to the side with the stronger sensor. The effectiveness of this design has been demonstrated in a number of robots.

A simple grammar that generates a complete Braitenberg body is shown in Figure 4. The grammar consists of two rules. The first rule rewrites the undifferentiated cell body (start symbol A) into two nonterminal B cells one of which is horizontally flipped relative to the other. The second rule converts a B nonterminal cells into a terminal cell with two edges defined. The parsimonious nature of the grammar that generates the Braitenberg Vehicle under our system shows the representational adequacy of the grammar system.

Ten runs of populations seeded with an initial population of Braitenberg agents were performed. Figure 5 shows the body of the best agent in the last generation of the best run. Next to the body is shown the NNet and a trace of its behavior starting from five random positions. Compared with the original Braitenberg ancestors, we see that effectors were added to the front and sides of the body and sensors were added to the rear. The crossed pathways from the front sensors to the opposite side rear effectors were replaced with direct pathways from front to rear on the same

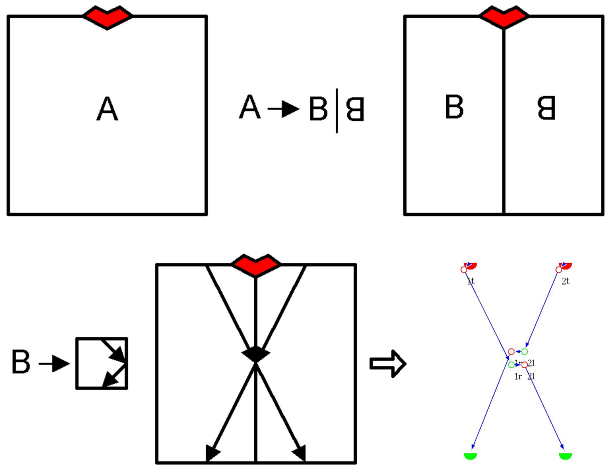

Figure 4: Grammar for Generating Braitenberg Vehicle

sides. The new front effectors are controlled by negative connections from the front and rear sensors on the opposite sides and by a positive weight from the rear sensor on the same side. The side effectors are controlled by a positive weight from the front sensor and a negative weight on the rear sensor of the same sides. Bias weights on the rear effectors provide constant propulsion which are negated by the front effectors as the agent approaches the center of the world.

Behavioral traces of this agent starting from five random positions and orientations are shown in the last figure. (The two long jumps shown are artifacts of the toroidal world). Note the accuracy of the final jumps into the center of the world. The trace beginning midway up on the left side actually falls short on its second jump but then takes a very small step to fall within the center ring.

Examining the best individuals of the final generation of each run we discovered the following:

- In all runs the original bi-lateral symmetry was retained.
- Four of the ten runs divided horizontally into two cells following the initial vertical division.
- Two of the ten runs lost the crossed pathways that characterized the original Braitenberg, replacing them with straight paths. Of the remaining eight runs that retained crossed pathways, none resembled the original pathways of the Braitenberg.

3.1.2 EVOLUTION FROM RANDOM GENOMES

We next initialized a population with completely random genomes as opposed to seeding the population with Braitenberg agents as above. The results achieved in ten runs are much less consistent than those achieved with the Braitenberg progenitors. Each run produced a solution however these solutions varied wildly in their implementations and quality of solution.

- Only two of the ten runs produced bi-laterally symmetric agents.
- Of the ten runs performed, nine of them produced agents that were inferior to those produced by the previous experiment.
- The tenth run produced a bilaterally symmetric agent that outperformed all of the Braitenberg descendants. This agent is shown in Figure 6.

This successful agent combined negative bias weights on the front with positive bias weights on the rear effectors to produce a strong jump. In addition the two effectors on the sides towards the front of the agent provided an accurate steering mechanism for pointing the front of the agent at the center of the world.

3.2 CENTER-SOURCE ENVIRONMENT WITH A VARIATION

The same fitness environment described above was modified slightly to speed up processing. In making this change a small opportunity for receiving exceptionally high fitness was allowed. This could only happen if an agent stopped at the edge of the reward ring ($d = 28$) with its mouth outside of the ring and then rotated its body without moving it so that its mouth fell within the reward ring. Ten runs of the Braitenberg seeded population were run in this environment. Seven out of the ten runs were able to discover the strategy for receiving high reward. This was surprising to us because we were not aware of the opportunity for receiving this high reward until the agents fitness began to skyrocket.

The best agent of the final generation of one of the best runs is shown in Figure 7. Note the behavioral trace in this figure. Note the two paths in the bottom left corner of the behavioral trace. Despite the fact that the second to the last step is nearly twice as far away in one trace than in the other the final jump takes the agent right to the edge of the ring in both cases. This demonstrates the extreme control that has evolved in the agent in order to exploit the opportunity for receiving high reward when landing at the edge. Compare

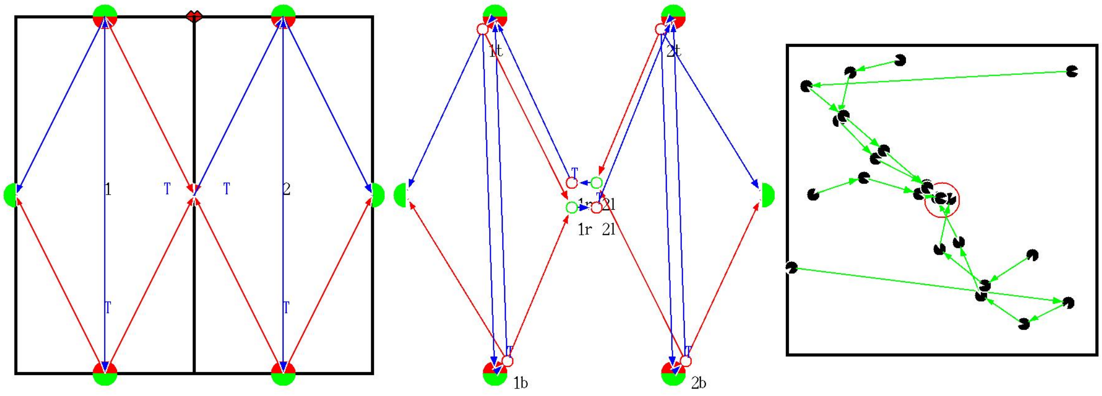

Figure 5: Center-Source Environment: Best Agent Evolved From Braitenberg. Inward facing semi-circles are sensors, outward facing semi-circles are effectors. T's indicate bias weights at edge. Darker edges are positive weights, lighter edges negative.

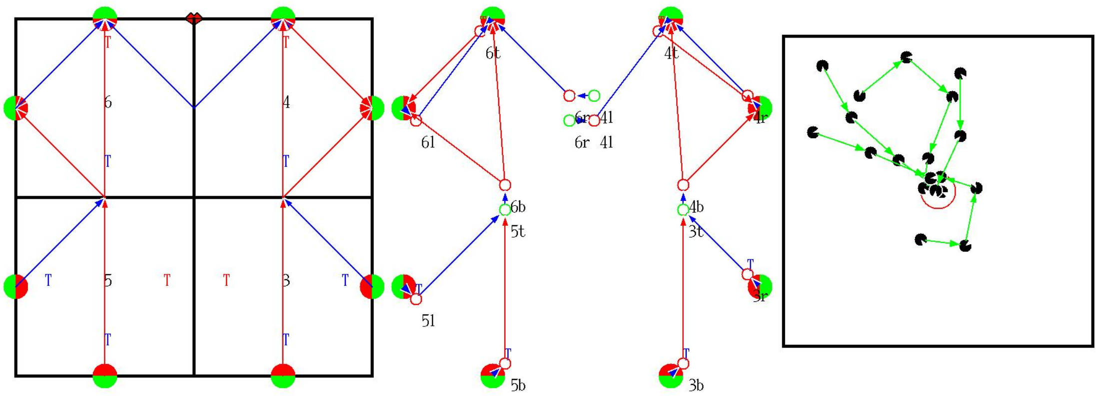

Figure 6: Simple Environment: Best Agent Evolved From Random

the location of all of the final jumps to the previous environments where the agents ended up very close to the center of the reward ring and far from its edges.

When the population was seeded with random agents rather than Braitenberg only one experimental run was able to discover and exploit the ring effect. The agent that displayed the behavior was similar to the one shown in Figure 7 in that it had positive bias weights at the bottom that serve to push it forward and positive weights from the bottom to the top that inhibit the bias weights when the agent approaches the center of the world. There were also weights that went from sensors on one side to effectors on the other but they were not bi-laterally symmetric like those found above.

3.3 DUAL-SOURCE ENVIRONMENTS

We next developed two more sophisticated environments. The purpose of these environments are twofold: first, to learn how our techniques scaled up with increased complexity; and secondly to explore evolution when multiple sensor types are available.

The new worlds are characterized by two sources of reward and two sensor types. One source of reward produces positive fitness while the other produces negative. Each source generates a fitness that is inversely proportional to the square of the distance. The total fitness is then the sum of the two component fitnesses. If the sum of all the fitnesses awarded over this period is less than 0, then the fitness for the agent is set to 0.

In the first world the sources are positioned in the center of opposite quadrants of the world as shown in Figure 8. Similarly to the earlier environment, if the agent moves itself within the rings shown by dashed lines, it receives a fitness of +1.0 for the upper left ring and -1.0 for the lower right ring. The agent is then immediately repositioned to a new starting point.

The two types of sensors available to the grammar are each capable of detecting exactly one of the sources of

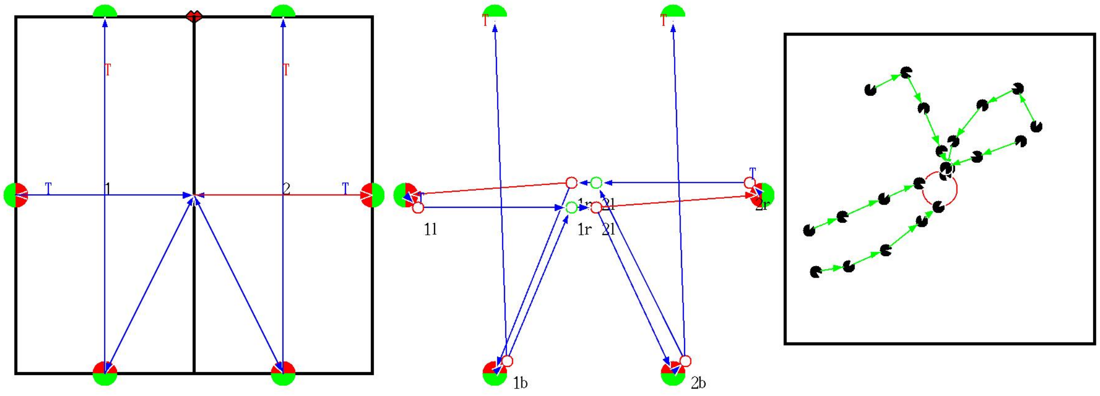

Figure 7: Environment with Edge Bonus: Best Agent that Evolved from Braitenberg

reward. The level of detection is proportional to the inverse square of the distance from the center of the rings.

In the single source worlds described above the source of sensor stimulation and reward was located in the center of the world. This centering meant that an agent was always closer to the reward point by going directly towards it rather than by crossing over the toroidal walls. This is not the case in the dual source worlds where it would be quicker to cross the boundary than to chart a course that avoids the negative reward source. We chose to disallow this strategy and removed the toroidal nature from the dual-source worlds. Agents who attempt to wander off the edge are stopped at the boundary.

Our expectations were that an ideal agent would follow a path to the positive source while avoiding the negative source. For this to occur the ideal agent must make use of both types of sensors. Agents that have only sensors that detect the positive source of reward would regularly achieve positive fitness scores but will occasionally receive negative fitness because they cannot detect the negative source and will occasionally stumble into the negative ring. Agents that detect only the negative source and avoid it will also score higher than an agent with random behavior. Such agents will occasionally accidentally pass through the positive ring although it is just as likely that they will be driven into a wall and receive very little reward for a given trial. As in prior experiments our agents surprised us.

All experimental runs were seeded with random genomes. Of the ten runs in this environment only one produced bi-lateral symmetry (Figure 8) and again the bi-laterally symmetric agent again performed the best of all runs. As can be seen from the behavioral trace this agent occasionally wandered into the negative ring. Only one run produced agents with sensors for both the positive and negative sources and that run performed the worst of all runs. All other runs produced agents that used positive reward sensors only.

It was clear from these results that there was not enough negative pressure to cause the agents to evolve behaviors that steered them away from the negative ring. In order to correct this we tripled the diameter of the negative ring and placed it in the center of the world (Figure 9).

Ten runs in this new environment produced the following results.

- Of the ten runs in this new environment three produced agents that had only negative source detecting sensors.

- Three runs produced agents that had only positive source detecting sensors.

- The remaining four runs produced agents that had both positive and negative source sensors.

The best agent of the last generation of the best run is shown in Figure 9. This agent was the only agent to exhibit bi-lateral symmetry. In addition it is one of the four that used both positive and negative sensors. The behavioral traces show that it is at first attracted to the positive ring but as it comes close to the negative ring it is diverted from its original path in order to avoid the negative ring.

In contrast to this predicted solution another interesting solution to the problem using only negative source sensors is shown in Figure 10. This solution was discovered by all three runs that produced only negative source sensors. In this solution the agent is drawn into a path that circles the negative source at a very precise fixed distance. This distance is the exact distance

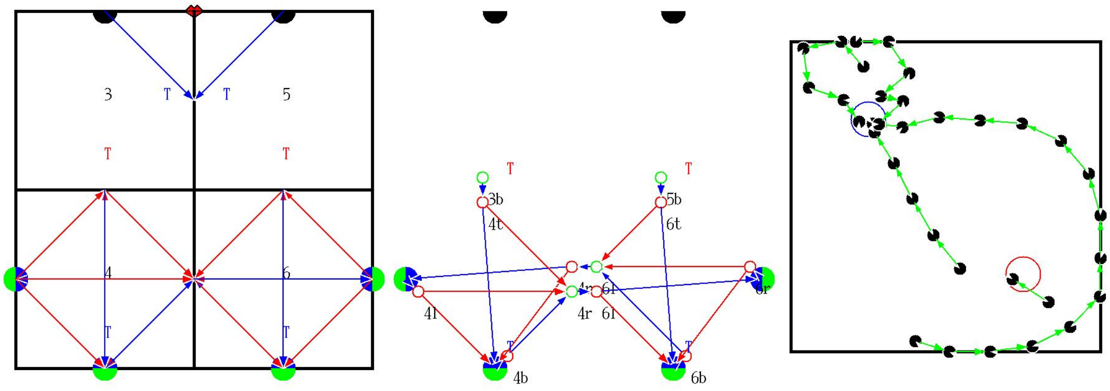

Figure 8: Dual-Source Environment: Best Agent. Points within upper left circle provides reward of +1.0, within lower left circle provides reward of -1.0, all other points provide reward of $\frac{1}{d_+^2} - \frac{1}{d_-^2}$ where d_+ and d_- are distances to the two centers.

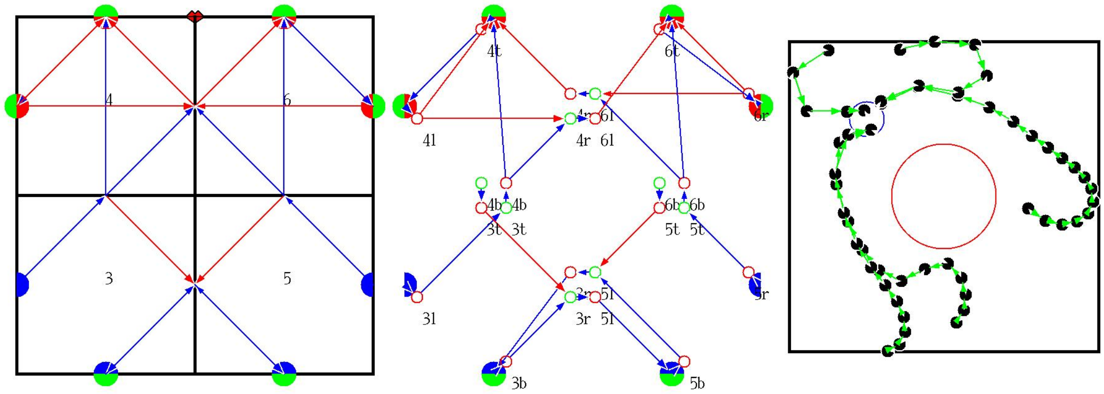

Figure 9: More Challenging Dual-Source Environment: Best Agent. Darker colored inward facing semi-circles are positive detecting sensors, lighter colored ones are negative detecting sensors.

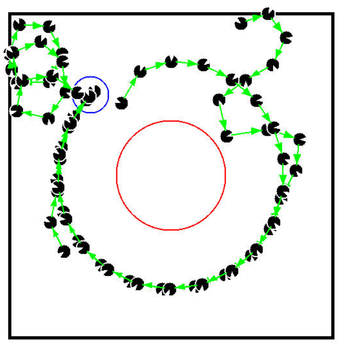

Figure 10: More Challenging Dual-Source Environment: Agent with Negative Sensors Only

from the center of the negative source to the center of the positive source. The effect thus causes the agent to fall into the positive source as it proceeds around the negative source.

4 CONCLUSION

We explored evolving grammars to generate strongly-coupled body plans and control structures. We applied this technique in several environments with varying levels of complexity.

We found that the strongest performers in each experiment demonstrated bi-lateral symmetry. In some cases the initial population was seeded with bi-laterally symmetric ancestors, but even in those experiments where this was not the case bi-lateral symmetry was occasionally discovered through the evolutionary process. We find this to be significant in its parallels with the course of biological evolution and intend to explore such bi-lateral symmetries further.

These experiments demonstrate the power and versatility of an evolved grammar to generate coupled body plans and control structures. Our current research is applying these techniques to real robots. We are doing this by mapping the existing sensors of a khepera robot to preset body positions in our agents. Following these experiments we will apply the grammars to evolving LEGO robots where both the physical body and control structure will be formed by our grammars.

References

[1] P. Angeline, G. Saunders, and J. Pollack. An evolutionary algorithm that constructs recurrent networks. *IEEE Trans. on Neural Networks*, 5:54–65, 1994.

[2] V. Braitenberg. *Vehicles: Experiments in Synthetic Psychology.* MIT Press, Cambridge, 1984.

[3] F. Dellaert and R. D. Beer. Toward an evolvable model of development for autonomous agent synthesis. In R. Brooks and P. Maes, editors, *Proceedings of the International Conference Artificial Life IV*, pages 246–257, Cambridge, MA, 1994. MIT Press.

[4] P. Eggenberger. Evolving morphologies of simulated 3d organisms based on differential gene expression. In P. Husbands and I. Harvey, editors, *Fourth European Conference on Artificial Life*, pages 205–213. MIT Press, 1997.

[5] D. Floreano and F. Mondada. Automatic creation of an autonomous agent: Genetic evolution of a neural network driven robot. In J.-A. M. D. Cliff, P. Husbands and S. Wilson, editors, *From Animals to Animats III*, Cambridge, MA, 1994. MIT Press.

[6] D. Floreano and S. Nolfi. Adaptive behavior in competing co-evolving species. In P. Husbands and I. Harvey, editors, *Fourth European Conference on Artificial Life.* MIT Press, 1997.

[7] T. Kammeyer. *Evolving Stochastic Grammars.* PhD thesis, Computer Science & Engineering, University of California, San Diego, 1998.

[8] H. Kitano. Designing neural networks using genetic algorithms with graph generation system. *Complex Systems*, 4(4), 1990.

[9] J. Kodjabachian and J. A. Meyer. Evolution and development of neural networks controlling locomotion, gradient following, and obstacle avoidance in artificial insects. http://www.biologie.ens.fr/fr/animatlab/perso/-kodjaba/kodjaba.html, 1997.

[10] W. Lee, J. Hallam, and H. Lund. Hybrid gp/ga approach for co-evolving controllers and robot bodies to achieve fitness-specified tasks. In *Proceeding of IEEE 3rd International Conference on Evolutionary Programming*, pages 384–389, New York, 1996. IEEE Press.

[11] S. M. Lucas. Growing adaptive neural networks with graph grammars. In *Proceedings of European Symposium on Artificial Neural Networks (ESANN '95)*, pages 235–240, 1995. http://esewww.essex.ac.uk/ sml/papers.html.

[12] A. Mark, D. Polani, and T. Uethmann. A framework for sensor evolution in a population of braitenberg vehicle-like agents. In C. Adami, R. K. Belew, H. Kitano, and C. E. Taylor, editors, *Artificial Life VI*, pages 428–432, Cambridge, Mass, 1998. MIT Press.

[13] C. Mautner and R. K. Belew. Evolving robot morphology and control. In M. Sugisaka and H. Tanaka, editors, *Proceedings of the 4th International Symposium on Artificial Life and Robotics (AROB 4th '99)*, pages OP22–OP27, 700 Dannoharu, Oita 870-1192, JAPAN, January 1999. AROB Secretariat c/o Sugisaka Laboratory.

[14] C. Mautner and R. K. Belew. Testing simulated controllers in real robots' bodies. In D. Floreano, editor, *Proceedings of the 1999 European Conference on Artificial Life*, Sept 1999. Submitted.

[15] F. Menczer and R. K. Belew. Evolving sensors in environments of controlled complexity. In R. Brooks, editor, *Proc. Fourth Conf. on Artificial Life*, 1994.

[16] K. Sims. Evolving 3d morphology and behavior by competition. In R. Brooks and P. Maes, editors, *Proceedings of the International Conference Artificial Life IV*, Cambridge MA, 1994. MIT Press.

[17] D. Whitley, F. Gruau, and L. Pyeatt. Cellular encoding applied to neurocontrol. In L. Eshelman, editor, *International Conference on Genetic Algorithms.* Morgan Kaufmann, 1995.