Instinct and Learning Synergy in Simulated Foraging Using a Neural Network

Thomas E. Portegys
School of Information Technology, Illinois State University
portegys@ilstu.edu

Abstract

Instinct and experience are shown to form a potent combination to achieve effective foraging in a simulated environment. A neural network capable of evolving instinct-related neurons and learning from experience is used as the brain of a simple foraging creature that must find food and water in a 3D block world. Instincts provide basic tactics for unsupervised exploration of the world, allowing pathways to food and water to be learned. The combination of both instinct and experience was found to be more effective than either alone. As a comparison, neural network learning also proved superior to Q-Learning on the foraging task.

1. Introduction

Foraging is an essential activity for many species, including some human societies. It thus also provides a valuable test bed for behavioral simulation with an aim toward artificial animal intelligence. In some organisms foraging consists of a combination of instinctive and learned behaviors. For example, honey bees will search their environment for nectar sources, learning their locations through visual cues and communicating this information to other bees [11]. Ants also forage and use pheromone signals to mark the location of food sources in the environment for further exploitation. The computational field of Ant Colony Optimization (ACO) [2,3] is largely based on this phenomenon.

Due to their similarity to natural nervous systems, artificial neural networks seem the most fruitful means of achieving generalized systems from the solutions of specific problems like foraging. Thus a neural network was chosen for the foraging task. Over the past 15 years a number of other systems have studied foraging and related problems with neural networks. Zhou and Shen [12] constructed a system that allows foraging "bugs" to learn an environment containing food, obstacles, and competing bugs. Their network learned from exposure to two-epoch trial cases that shaped an abstract force field gradient which in turn allowed test cases to be categorized with similar trials. Erdur and Güngör [4] follow a theme similar to this project in utilizing a combination of evolution with a genetic algorithm and experiential Hebbian learning to modify the neural network configuration to produce effective foraging as well as other behaviors. Nolfi and Parisi [5] pointed out that although evolution is a good way to get reasonable initial behavior, learning is indispensable to adapt to specific and changing conditions. Mazes have also been employed as an environment to investigate goal-seeking learning in artificial neural networks [1,10].

This study is believed to be novel in the way it combines instinctive behavior as a means of training experiential learning. This is a plausible counterpart to the way simple animals learn, and is therefore a useful approach to simulating them. The way this works in the foraging task is as follows: instincts guide the creature to effectively explore its environment, producing a stream of stimuli and responses which are then incorporated into new neurons recording pathways in the environment. These learned neurons are reinforced by consistent repetition as well as by association with the acquisition of food and water goals. Over trials the learned network often overrides instincts to guide the creature directly along paths to food and water.

The Mona goal-seeking neural network was used for this task. Mona has been shown to be capable of supporting instinct evolution to solve the Monkey and Bananas Problem [7], as well as effectively learning mazes requiring the retention of context information over time [6].

Q-Learning [9], a well-known reinforcement learning technique that is amenable to stimulus-response search space tasks, was used as a comparison to the neural network.

1.1. A brief overview of Mona

Mona is based on the rationale that brains are goal-seeking entities. It has a simple interface with the environment: all knowledge of the state of the environment is absorbed through senses. Responses are expressed to the environment with the goal of eliciting sensory inputs which are internally associated with the reduction of needs.

Events can be drawn from sensors, responses, or the states of internal neurons, calling for three types of neurons. Neurons attuned to sensors are receptors, those associated with responses are motors, and those mediating other neurons are mediators. Mediators can be structured in hierarchies representing environmental contexts. A mediator neuron controls the transmission of need through and the enablement of its component neurons.

To elucidate by example, consider this somewhat whimsical task: let Mona be a mouse that has been out foraging in a house and now wishes to return back to her mouse-hole in a certain room. For the sake of keeping peace with her fellow mice, she must not make the mistake of going into a hole in another room. Figure 1 shows her neural network at this juncture.

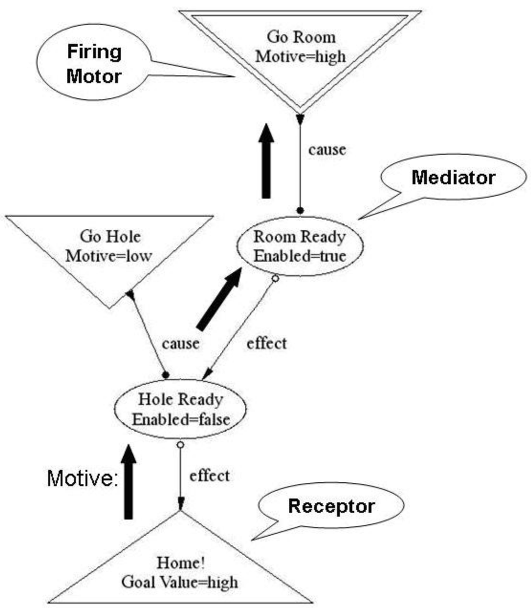

Figure 1. Initial mouse network

The triangle-shaped object at the bottom is the receptor neuron that fires once she has reached her hole; the inverted triangles are motor neurons that accomplish the responses of going to the correct room (Go Room), and going into the hole (Go Hole). The ellipses are mediator neurons. Each is linked up to a *cause* and *effect* event neuron. The “Hole Ready” mediator is not *enabled*, reflecting the importance of not going into a hole in the wrong room. The “Room Ready” mediator is enabled, signifying an expectation that if its cause event fires, its effect will also fire.

The “Home!” receptor neuron has a high goal value, indicating that it is associated with a need. Because of this, *motive* influence propagates into the network, flowing into motor neurons whose firings will navigate to the goal. Since the “Hole Ready” neuron is not enabled, the motive bypasses the “Go Hole” motor neuron in search of a mediator whose firing will enable “Go Hole”. Since “Hole Ready” is an effect of “Room Ready”, it flows into the “Go Room” motor via the enabled “Room Ready” mediator and causes it to fire (double outline).The flow of motive illustrates how mediators representing contexts work together. The appropriate context for “Hole Ready” is “Room Ready”, which means that the latter should necessarily contribute something to the former in order to enable it. This something is called a wager. A wager temporarily modifies the enablement of a mediator that is the effect event of another mediator. It is called a wager because the base-level enablement of the wagering mediator will be evaluated based on subsequent firing of the effect neuron.

In Figure 2 the “Go Room” cause firing can be understood as a conditional probability event: given that Mona is in the correct room (“Room Ready”), she is quite certain that she can go into her own hole. This accomplished by a wager from “Room Ready”, triggered by “Go Room” that boosts the enablement of “Hole Ready”. After this enablement occurs, motive flows into the “Go Hole” motor neuron, causing it to fire. Subsequently the Mona senses that she is home in her hole.

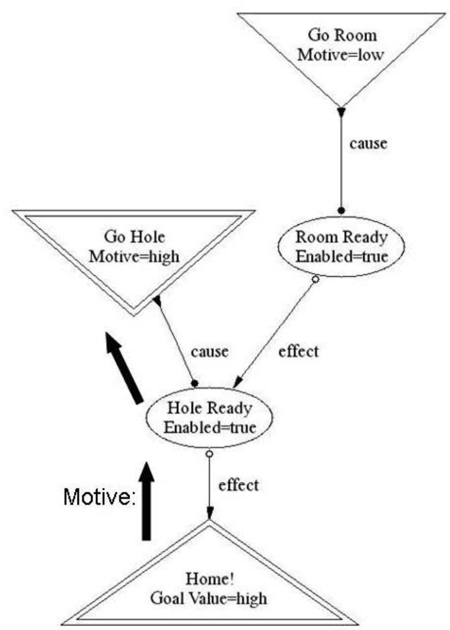

Figure 2. Final mouse network

2. Description

A muzz is a creature that lives in a 3D block world such as that shown in Figure 3. The right panel of the display shows a top view of the world, and the left panel shows the muzz in the upper left corner of the world as viewed by another muzz facing it. Blocks are randomly marked with letters of the English alphabet. Striped ramps also may lead to platforms of various heights. A mushroom (a small circular shape from the top view) and a pool (a larger circular shape) also appear somewhere in the world as a food and water source for the muzz respectively.

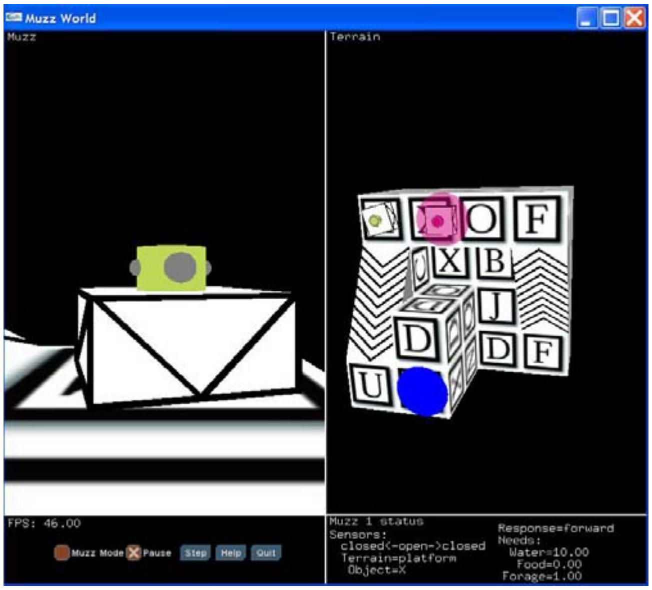

Figure 3. A muzz world

A muzz must forage for mushrooms and water in this world. A muzz has the following sensory capabilities: 3 sensors for detecting if the way is open to move in the forward, right, and left directions; a sensor to detect the terrain in the forward direction: { platform, wall, drop off, ramp up, ramp down }; and an object sensor for detecting objects in the forward direction: { mushroom, pool, muzz, empty, <block letter> }. Its response repertoire consists of: wait, move forward, turn right or left, eat, and drink.

A muzz also has 3 needs: food, water, and foraging. The forage need is based on a fraction of the maximum food or water need, which means when they are satisfied the muzz has no need of foraging. Initially, all of the needs are positive, meaning that they may compete to drive the muzz's responses. In other words, a learned path to a pool may "vie" with a different path to a mushroom. By attenuating need-derived motives as they drive through the network, the path to the closest goal will be preferred. This is assuming that the need for water and food are equal, which is the case in this study. Once a need has been satisfied, e.g. by drinking water, only motives associated with other positive needs will drive the network toward goals satisfying those needs.

2.1. Instincts

Receptor neurons for sensing mushrooms and water were initially placed into the neural network and given goal values associated with the reduction of hunger and thirst respectively. These were terminal goals for learned mediators (see learning section). Upon sensing a mushroom a muzz will automatically eat it if it is hungry. The same goes for a pool and drinking.

Three mediator neurons were also "hard wired" into the muzz to implement foraging instincts. One of them associates the "forward open" receptor neuron with the "move forward" response. The others associate receptors indicating openings to the right and left with turning right and left respectively. The goal values of these instinct mediators determine the probability of expressing the movement responses. These are critical values, since it is possible to set these such that foraging fails completely. For example, if the move forward mediator always dominates, the muzz will never turn down a side pathway that may lead to a goal, or will always follow walls and never explore an open space. If the turn right mediator dominates, on the other hand, the muzz will rotate endlessly in an open area. To determine effective settings, an evolutionary selection procedure (see procedure section) was used to select the instinct mediator goal values to produce effective foraging. These values were evolved in the presence of learning to achieve synergistic behavior.

2.2. Learning

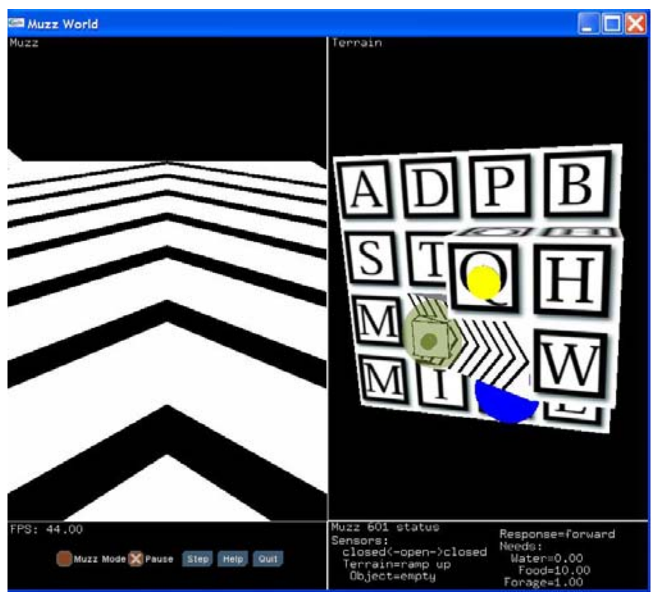

Figure 4. Muzz approaching mushroom

As foraging proceeds, a stream of sensory inputs and responses is generated. The neural network creates new receptor and mediator neurons to record these streams. The Mona neural network prefers to retain mediators that excel at being reliable/repeatable or lead to need-reducing goals, which in this task are the mushroom and pool sensing receptors. In this study mediators were capped at a maximum of 200, which, coupled with the exploratory nature of foraging, meant that most learned mediators were eventually destroyed.

As an example, Figure 4 shows a muzz ascending a ramp toward a mushroom on the platform above. Figure 5 is an annotated snapshot of the mediator controlling this activity, showing the sequence of stimuli and responses involved.

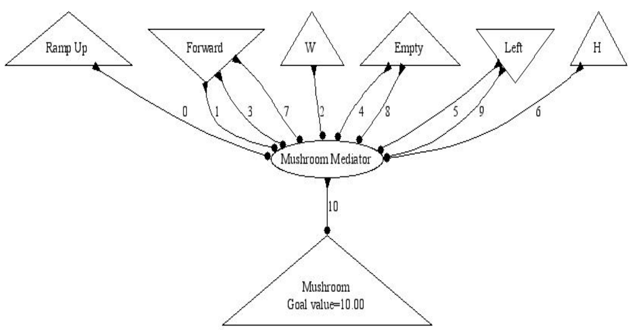

Figure 5. Mushroom seeking mediator neuron

2.3. Procedure

An initial population of 40 muzzes was generated and given random foraging instinct values. For each trial, a single muzz, mushroom and pool were placed in the world, and the muzz allowed to forage for 500 response steps. The fitness of a muzz was a function of whether it found food and/or water, and by how quickly it did so. The fittest 20 muzzes were used to create the next generation through mutation and mating. Mutation consisted of copying learned neurons into the offspring and probabilistically (10%) randomizing instinct goal values. Mating consisted of randomly choosing instinct goal values from a parent and randomly copying the strongest neurons from either parent into the offspring until the maximum of 200 was reached. Since each neuron is uniquely identified by a recursively computed MD5 hash, duplicating neurons was prevented. For an individual evolution run, the world configuration, consisting of the topography and object locations, was the same; these varied for different runs. Each evolution run proceeded for 30 generations. A set of 25 runs was done for 3 world dimensions: 4x4, 8x8, and 12x12.

2.4. Q-Learning

The block world presents a search space in which a stimulus-response stream can take a muzz from an initial point to a foraging goal. Consequently it was chosen as a comparison to the neural network. Just as for neural network experiential learning, Q-Learning was initially guided by foraging instincts. To tune them to work together, several Q-Learning parameters, shown in Table 1, were evolved in conjunction with instincts. This was done along the lines of the instinct evolution; hence the Q-Learning parameters of a mutant muzz were set to randomized values within minimum and maximum values. Also, since there were two goals, water and mushrooms, there were actually two concurrent Q-Learning processes, each sensitive to one of the goals. Each contributed to response selection as long as its respective goal was unsatisfied, which is a mechanism also incorporated in the neural network. Thus, combined with instincts, there were possibly three influences on response selection.

Table 1. Q-Learning parameters

Name	Initial	Minimum	Maximum
Reward	1.0	.001	5.0
Q value	.001	.001	1.0
Learning rate	.9	.1	.9
Rate attenuation	.9	.1	.9
Discount	.9	.1	.9

3. Results

For each world dimension setting, the fittest 10 muzzes for each of the 25 runs were tested, scored, and averaged under a variety of conditions to create the graphs shown below. The score was how many response steps out of a maximum of 500 were needed to get both food and water. Table 2 provides the legend for the graph symbols.

Table 2. Graph symbol legend

FI, ~FI	Foraging instincts enabled/not enabled
LC, ~LC	Learning capability enabled/not enabled
LE, ~LE	Learning experience used/not used
QLE,~QLE	Q-Learning experience used/not used

Figures 6, 7, and 8 show the 4x4, 8x8, and 12x12 world performances respectively. As observed, scaling the world for the most part seems to scale the results accordingly. In the first base case experiment (~FI,~LC), the muzzes were "lobotomized" by disabling both foraging instincts and learning capability. In most configurations, the muzzes were simply unable to locate food and water within the 500 step limit. Some configurations placed the muzz, mushroom, and pool close enough to allow success by making random responses. In the second (~FI,LC) experiment, only the learning capability was enabled. This resulted in performance as poor as the lobotomized muzzes, which is a stark testimony to the importance of having some tactics available to engage the environment. The next experiment (FI,~FC) indicates what a powerful effect the few simple instincts alone had on task success.

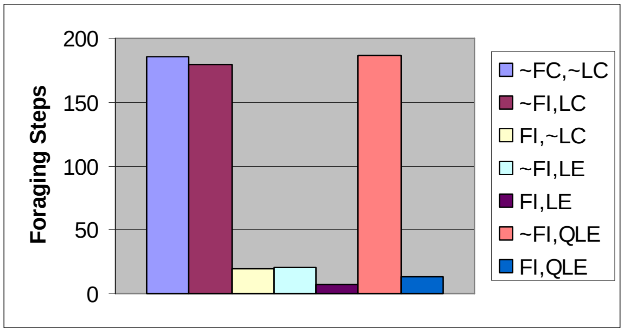

Figure 6. 4x4 World performance

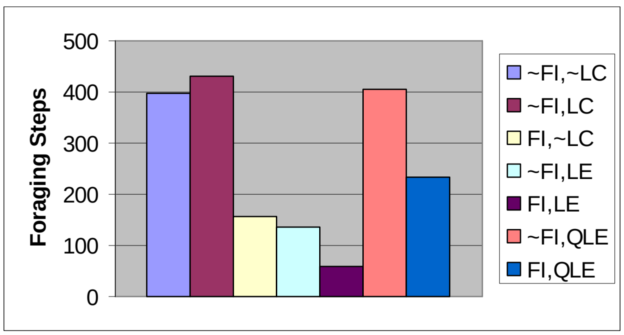

Figure 7. 8x8 World performance

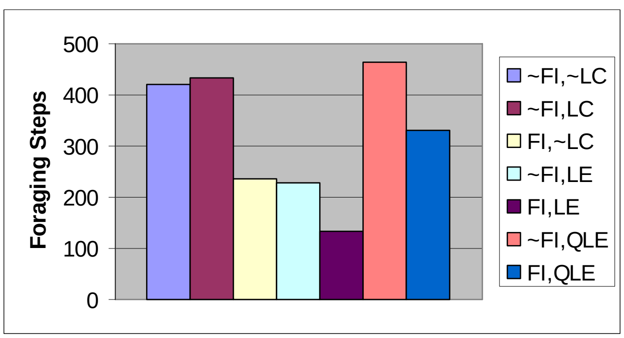

Figure 8. 12x12 World performance

The next experiment (~FI,LE) was interesting and somewhat unexpected. Here instincts and learning were enabled and the world foraged. For the test, foraging instincts were disabled and learning experience alone enabled. The result is comparable performance to foraging instincts alone. On closer observation, it appears that not only were a number of environmental paths learned, but that foraging itself was learned: the muzzes moved about with exploratory movement patterns. In the last experiment, the synergistic benefit of instinct and experiential learning was striking, cutting the time to find food and water approximately in half relative to either alone. Looking closer at a number of trials, especially in the 12x12 world, foraging appears to serve to get the muzz on to a learned pathway, whereupon learned behavior can activate to take the muzz directly to a goal.

The Q-Learning performance was unexpectedly poor. Not only was learning experience running without foraging instincts highly ineffective in all three world dimensions, but in the 8x8 and 12x12 worlds it actually hindered the effectiveness of foraging instincts. While it was expected that Q-Learning would in some instances be confounded by redundant sensory states within a goal path and by the three-way vying for control between instincts and the two goal-specific Q-Learning processes, the extent of the degradation was surprising.

4. Conclusion

The use of a few basic hard-wired neurons, tuned by evolution, has been shown to radically improve foraging performance. Moreover, the superiority of the instinct/learning synergy suggests that more ambitious studies are warranted. For example:

- In order to more closely mimic nature, an environment might be constructed that contains generalities related to foraging, such as a certain type of fruit that grows in proximity to environmental cues, such as odors or terrain markings. Then creatures might learn more generalized patterns related to resource acquisition.
- The addition of manipulable objects in the environment could be used to study such behaviors as nest-building.
- The addition of other creatures could be used to study social behaviors such as predator/prey strategies.
- The embodiment of the creatures in simple physical robots would create an opportunity to mesh other fields such as pattern recognition and kinematics with the neural network.

As a final note, the utility of uniquely identifying each neuron with an MD5 hash to prevent duplication during mating should be underscored. What it means is that any two neurons in different networks having the same id are recursively structurally identical. One of Mona's design goals is to address the critical problem of non-modularity in classical feed-forward networks [8] by being able to configure neurons that do specific jobs, something that biological neurons are also capable of. Imagine the possibilities of exchanging and even sharing neurons between networks, something that nature does not design for.

The C++/OpenGL source code for Mona and the muzz world are available at:

www.itk.ilstu.edu/faculty/portegys/research/muzz/muzz.zip (zip) or muzz.tgz (tarball).
It can be compiled with either gcc/make or Microsoft Visual Studio .NET.

5. References

[1] J. Blynel, and D. Floreano, "Exploring the T-Maze: Evolving Learning-Like Robot Behaviors using CTRNNs", In Raidl, G. et al. (Eds.) *Applications of Evolutionary Computing*, Heidelberg: Springer Verlag, 2003.

[2] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. ISBN 0-19-513159-2, 1999.

[3] M. Dorigo and T. Stützle, *Ant Colony Optimization*, MIT Press, 2004.

[4] S. Erdur and T. Güngör, "An Investigation of Artificial Neural Network Architectures in Artificial Life Implementations", *International 13th Turkish Symposium on Artificial Intelligence and Neural Networks* (TAINN 2004), 191-199, İzmir, 2004.

[5] S. Nolfi and D. Parisi, "Neural networks in an artificial life perspective", In: W. Gerstner, A. Germond, M. Hasler and J.-D. Nicoud (Eds.), *Lecture Notes in Computer Science 1327*. Berlin: Springer-Verlag, pp.733-738, 1997.

[6] T. Portegys, "An Application of Context-Learning in a Goal-Seeking Neural Network", *The IASTED International Conference on Computational Intelligence* (CI 2005), Calgary, Canada, 2005.

[7] T. Portegys, "Instinct Evolution in a Goal-Seeking Neural Network", *The IASTED International Conference on Computational Intelligence* (CI 2006), San Francisco, USA, 2006.

[8] J. Tan and S. Nolfi, "Learning to perceive the world as articulated: An approach for hierarchical learning in sensory-motor systems", *Neural Networks*, 12(7–8):1131–1141, 1999.

[9] C. Watkins, *Learning from Delayed Rewards*, Thesis, University of Cambidge, England, 1989.

[10] Y. Yamauchi and R. Beer, "Sequential behavior and learning in evolved dynamical neural networks", *Adaptive Behavior*, 2(3):219-246, 1995.

[11] S. Zhang, F. Bock, A. Si, J. Tautz, and M.V. Srinivasan, "Visual working memory in decision making by honey bees", *Proceedings of the National Academy of Sciences of the United States of America,* 102 (14): 5250-5, 2005.

[12] Z-H Zhou and X-H Shen, "Virtual Creatures Controlled by Developmental and Evolutionary CPM Neural Networks**",** *Intelligent Automation and Soft Computing*, 9(1): 23-30, 2003.