Ortus: an Emotion–Driven Approach to (artificial) Biological Intelligence

Andrew W.E. McDonald, Sean Grimes and David Breen

Drexel Unversity, Philadelphia, PA 19104
awm32@cs.drexel.edu

Abstract

Ortus is a simple virtual organism that also serves as a framework for developing biologically based artificial intelligence. Born from a goal to create complex virtual intelligence and an initial attempt to model C. elegans, Ortus implements a number of mechanisms observed in organic nervous systems, and attempts to fill in unknowns based upon plausible biological implementations, psychological observations. Implemented mechanisms include excitatory and inhibitory chemical synapses, bidirectional gap junctions, Hebbian learning with its Stentian extension. We present an initial experiment that showcases Ortus fundamental principles; specifically, a cyclic respiratory circuit, and emotionally–based associative learning with respect to an input stimulus. Finally, we discuss the implications and future directions for Ortus and similar systems.

Introduction

While much work has been done to develop artificial intelligence (AI) systems that borrow principles from organic nervous systems, far less has been done that specifically targets the intersection of biology and artificial intelligence such that biological principles—rather than a specific applicability of the technology—are of primary concern, with the main goal being a virtual system that exhibits biological intelligence (BI). As our understanding of organic nervous systems and access to computation power both increase, widespread interest in systems that do exaclty this is greatly increasing; as evidenced by DARPA's recent L2M project, in search of machines that learn throughout their lives (DARPA, 2017).

Researchers in the realm of computational biology and neuroscience have started making progress toward devloping systems that model specific organisms or neural circuits, such as C. elegans (Izquierdo and Beer, 2016), though these systems have the potential to require too much focus on organism–specific details, decreasing their potential for generalized creation of neurologically–inspired intelligent systems.

On the other hand more traditional (application focused) AI research has started taking more inspriation from human learning, as did Bowren et al. (2016), developing an auto–encoder augmented by Hebbian learning, by Bowren et al. (2016), decreasing the need for an initial superivised–like learning period. Further, Marblestone et al. (2016) discusses ways that artificial neural networks (ANNs) can more closely approximate neural functionality. In the context of biologially–inspired AI, the frameworks underlying these approaches may be too constraining for full exploration.

Recent work at the intersection of these two areas includes Sinapayen et al. (2016), which investigates the applicability and biological plausibility of spiking neural networks learning by "stimulation avoidance". Though perhaps the project most closely aligned to Ortus is a biologically inspired neural network modeled off of a honey bee's visual system, which merges biological mechanisms and neural networks Roper et al. (2017).

Ortus is an approach to virtual life that has the potential to approximate the intelligence of living organisms. Bourne from the study and analysis of C. elegans connectome and behavior,

One of Otus' main goals was to eliminate the need for backpropagation, which is costly, and doesn't appear to have a biological counterpart. Ortus' synaptic weights change according to Hebbian (and opposite, fix this) learning principles as part of it's operation. This means that there is no explicit training period, but rather that the system's initial behavior is based upon neural structure, and as it ages, modifies and refines the network—in much the same way as a baby.

This paper will outline Ortus' design,

GENERALLY:

- Why is Ortus/this project important? Who cares?
- What else is out there? Approaches to dynamically structuring neural networks based upon input data (NIT)
- What is Ortus, and why is Ortus different?

Ortus' creates its initial structure is based upon a set of specifications, similar to the way DNA controls organic development (though vastly simplified)

- Give overview of paper.

citet is just year

citep is author and year.

System Design

One of Ortus' goals was to only implement functions that either had a known analogous biological process, or which may have an analogous biological implementation along with a psychological argument that is currently unknown. Following each Ortus design element (ODE) underlying Ortus' functionality is its biological rationale (BR):

1)

It is well known that emotions are a driving force in motivation and behavior of intelligent organisms (cite). While at the human level, the emotional "part" of the brain is quite complex, it is not unreasonable to assume that as organismic complexity (and thereby intelligence) decreases, the complexity of emotions decreases. If one follows this line of thought, the possibility emerges that organisms like C. elegans may, in fact, be driven by "emotions" as well. For example, C. elegans is capable of toxin avoidance, a tap–withdrawal response, as well as learning that it found food at a certain temperature (citations?) – one must ask how this can be. There is nothing external that tells it what is good vs. bad, but it wants to avoid certain things, while being attracted to others. The premise behind Ortus, is that these behaviors are a result of a *very* simple emotional subsystem that forms the basis for C. elegans' behavior.

While Ortus' current implemention is a vastly simplified approach to a virtual organism, there are a few underlying principles. First, its core, least mutable circuit, is its respiratory circuit, which naturally cycles between inhalation of O_2, and exhalation of CO_2 (show graph). Second, as a part of its "life", a constant amount of O_2 is decreased, while a constant amount of CO_2 is introduced—this causes inhalation, which increases O_2. Third, all sensory inputs, are, after input consolitation, tied to all emotional centers that exist. The `FEAR` center, has a causal relationship between increased CO_2 and increased `FEAR` activation. This enables associative learning between an input, and an emotional state.

Synapic plasticity, or Hebbian learning, as well as (find out the opposite to it...) are implemented. Synapses strengthen when a presynaptic neuron and postsynatptic neuron or muscle activate (relative to the synapse's mutability index), and weaken gradually according to (insert equation) both over time, and when the presynapic neuron reaches its threshold, but the postsynaptic element (neuron or muscle) does not. Synaptic habituation

the ORT filetype is a specification of "rules", akin to very simplified "virtual DNA" that define elements (neurons and muscles), and relationships between them, currently: A causes B, A is correlated with B, A opposes B, and A dominates B. Each realtionship has a number of paramters, such as Age, Mutability, Threshold, and either a chemical synapse (CS) weight (unidirectional), or a gap junction (GJ) weight (bidirectional) (NOTE: equations!). Ortus' neurons are based upon C. elegans' neurons, and are non–spiking, but do not transmit any "activation" (footnote about why activation is used instead of voltage? or use 'potential?) below a given threshold, (see [1] below). The equations for CS and GJ synapses were derived (simplified) from Wicks' 1996 paper (cite), and appear to have been used in recent C. elegans' models (cite Beer's 2013).

Ortus bulids its connectome based upon the rules specified in a supplied ORT file, based upon guidelines that aim to approximate the way neural circuits form in organic brains (cite? expand upon this? e.g, how does it do that?).

Sensory stimulation is sensed by sensory neurons that are topographically arrainged, similar to those in organic systems. A recent study suggesting that the brain is organized based upon an n = 2something... and something about permutations lends credence to this approach (cite n something 2 permutation paper).

ODE: Emotional learning is achieved by strengthening synapses of interneurons that form junctions between consolidated sensory information, and interneurons that are tied to each emotional "center" by GJ connections. The GJ connection allows an emotion to propagate over the entire system once activated. The effect of this, is that if a certain sensory input is causing the activation of a given emotion, the introduction of another sensory input will cause the synapses at the junction of the newly introduced sensory input and elevated emotion to strengthen. (graph here)

BR: Specifically for the Dopamine circuit in the brain, there exists a small locus of dopaminergic cells, but the axons project all througout the brain. From a psychological perspective, it is clear that when one enters an emotional state, the emotion is encompassing, not localized. It is also clear that a stimulus or activity is likley to be "colored" by the emotional state one was in when the stimulus occurred.

ODE: Each interneuron that acts as an "emotion extension interneuron" (EEI) that branches away from the primary neuron for a given emotion, is connected to its primary emotion neuron via a GJ. It also has a CS connection back to the sensory consolidatory interneuron that has a CS into it This creates an emotional feedback circuit, allowing various emotional states to cause Ortus to "remember" other things that are related to that emotional state, by *slightly* activating the sensory consolidatory interneuron that would trigger the emotional response in the presence of the actual stimulus. In Ortus' current state, to "remember" means a given neuron or neural circuit has a measurably elevated activation, as a result of stimulation by an interneuron that would not have caused an increase in activation had associative learning not occured (diagram of S -¿ SEI -¿ SCI =¿ EEI —-— E thing).

BR: only psychological...

ODE: Sensory consolidatory interneurons have weights set initially such that the total incoming weight to any given SCI is the same; so the weight of any individual presynapse is $\frac{1}{\#inputs}$. This allows simple snesory, error–less sensory consolidation. However one current issue is if we have SCIs for sensory neurons `A`, and `B`, we will have `iA`, and `iB`, and `iAB`. If `iAB` is activated, that means that `iA` and `iB` will also be activated. A solution might be for more complex SCIs to inhibit less complex ones, though that would perhaps unneccesarily add to computational complexity.

BR: cite paper that talks about connection weight is usually inversely proportional to incoming connections.

Talk about the combinatorial explosion, but then discuss how that can be pruned away... similar to babies. The $2^n - 1$ paper suggested reducing the resolution of inputs, so, instead of each sensor being able to connect with every other sensor, funnel 3 sensors (for example) into 1, to create a "functional connectivity module" (FCM). That approach might not be bad, however, it could be severely limiting. Perhaps this can be cut down by looking at relationships between sensors, and getting rid of things that don't matter?

Choices are: 1) an array of 9 sensors, if 0,1,2 ¿ A, 2,3,4 ¿ B, 4,5,6 ¿ C, 6,7,8 ¿ D, then sum nCr(k,n) for n from 1 to k is $2^n - 1$. It is clear that to get a fully connected network, capable of expressing all possible combinations, even from a living organism perspective, things become unfeasible.

2) Rather, it is probably the case that our brains aren't quite as connected as they appear to be. (e.g., within a certain radius, two places on your skin feel like the same location. visual system either might be an exception, however it is well known that visual information gets interpolated.) There must be some sort of neural processing shortcuts ... approximation. e.g., if you are quickly reading, you might read "New York", when the only word there was "York"

floor so that learned stuff can't be unlearned beyond a certain point – e.g., as a relationship ages, it's mutability should tend toward 0.

¡¡IMPLEMENTATION¿¿

Implementation Details - C++, OpenCL - each neuron operates on its incoming connections, so it only does "postsynaptic" work, leaving its "presynaptic" work to neurons that it is presynaptic to (see "shortcomings") - sensory consolidation causes resolution loss, currently maximum resolution for any sensor array (e.g., not just O2) has a maximum resolution of 2 for intra–sensor consolidation, and 3 for inter–sensor consolidation. Single neuron sensors have no minimum for either. (NOTE: sensor array not implemented yet... take this out if it doesn't get implemented)

Discussion

While the implementation presented is quite simple, the initial results presented pave the way for quick advances to Ortus' capability with regard to "behaviors" resulting from emotional amalgamation as well as the introduction of more complex sensory input.

While Ortus' eventual goal is to develop complex "neural" functionality akin to that observed in complex systems, we

From an artificial intelligence standpoint, the

1. less connections
2. ability to specify initial behavior
3. don't need a massive traning set
4. 3 types of interactions: inhibitory CS, excitatory CS, and GJ.

Together, these afford artificial neural networks far greater flexibility and thereby power. It is possible that this will enable more nuanced behavior.

From an artificial life standpoint: people are modeling networks with less complexity than this to attempt to recreate C. elegans' behavior from its connectome (cite beer.. and newer). This work pivoted from that approach, because in attempting to model the entire connectome, it became clear that there were too many unknowns. However, it also became clear that there was a certain logic to the way the neurons were wired; a notion backed up by (cite c elegans wiring works... have some in mendeley).

A simple idea is, how can C. elegans ¡DO SOMETHING¿ – point here is to describe the requirement that certain things inhibit while others excite, in order to allow a certain behavior to happen.

Shortcomings and Issues. There is no temporal differentiation between CS and GJ transmission. On each timestep, neurons collect all incoming "activation" from.

There is minimal presynaptic processing – (NOTE:how true is this?)

No location–based axon growth, as is seen in, relative to hebbian plasticity (cite)

Conclusion

We have presented Ortus, an initial approach to a framework that exhibits basic emotionally–based learning (fear conditioning), self–sustaining inherent behavior (cyclic breathing), and a topographical approach to sensory processing that enables intersensory associative learning, with or without an emotional component.

Ortus also has a system to map sensory inputs such as "breathe" to the stimulation of O2 sensors.

[1] https://www.ncbi.nlm.nih.gov/pubmed/210264

Acknowledgements

This work was supported by NSF grant No. PHY-9723972.

References

Bowren, J., Pugh, J., and Stanley, K. (2016). Fully Autonomous Real-Time Autoencoder-Augmented Hebbian Learning through the Collection of Novel Experiences. *Proceedings of the Artificial Life Conference 2016*, (Alife 15):382–389.

DARPA (2017). Toward machines that improve with experience.

Izquierdo, E. J. and Beer, R. D. (2016). The whole worm: Brain-body-environment models of C. elegans. *Current Opinion in Neurobiology*, 40:23–30.

Marblestone, A. H., Wayne, G., and Kording, K. P. (2016). Towards an integration of deep learning and neuroscience. *bioRxiv*, 10(September):1–61.

Roper, M., Fernando, C., and Chittka, L. (2017). Insect Bio-inspired Neural Network Provides New Evidence on How Simple Feature Detectors Can Enable Complex Visual Generalization and Stimulus Location Invariance in the Miniature Brain of Honeybees. *PLOS Computational Biology*, 13(2):e1005333.

Sinapayen, L., Masumori, A., and Ikegami, T. (2016). Learning by Stimulation Avoidance: A Principle to Control Spiking Neural Networks Dynamics. page 17.