

Open-Ended Dynamics of Prediction Games

Dissertation Proposal
Thomas Willkens

Background

Much of the artificial life community is currently focused on the problem of understanding *open-ended evolutionary systems*, which can be broadly defined as evolving systems that never settle into a single stable equilibrium state. The Earth’s biosphere provides the most obvious example of an open-ended system, as new and ever more complex forms of life evolve over time. However, the dynamics of these systems are impossible to capture in closed form, and they have proven difficult to model (Thurner et al., 2018). Open areas of research in the field include the establishment of metrics for measuring open-ended behavior, the identification of systems believed capable of open-ended growth, and the characterization of such systems in terms of activity and complexity (Packard et al., 2019; Stepney, 2021).

Pioneering work in the field of open-ended evolution includes the identification of punctuated equilibria within a variation of the iterated prisoner’s dilemma in Lindgren and Nordahl (1994), the self-replicating machine code of Tierra and Avida (Ray, 1992; Ofria and Wilke, 2004) and the evolutionary activity statistics of Channon (2001). While more recent work has produced systems that exhibit features of open-endedness by certain metrics (Adams et al., 2017; Soros and Stanley, 2014), an unambiguously open-ended system has yet to be developed and is considered to be one of the “grand challenges” of artificial life. The dynamics of these systems in general are agreed to be poorly understood, but novel analytical frameworks such as the *Measurements of Open-Ended Dynamics in Evolutionary Systems* (MODES) Toolbox have been proposed to in an effort to better characterize them (Dolson et al., 2019).

The DEMO Lab of Jordan Pollack has produced much work in the related area of *coevolution* and *coevolutionary algorithms* (CEAs). CEAs have found practical use on a number of difficult problems, such as the sorting networks of Hillis (1990) and the game playing strategies of Rosin and Belew (1995). It has been argued in Arulkumaran et al. (2019) that part of the success of some deep-reinforcement learning systems such as AlphaStar can be attributed to the use of competitive coevolutionary principles. However, experimental work by the DEMO Lab showed that coevolutionary systems are prone to unique pathologies that may stall adaptive progress in certain domains (Watson and Pollack, 2001). Theoretical work by the lab led to the development of algorithms specifically to

address these pathologies (Noble and Watson, 2001; De Jong and Pollack, 2004; De Jong, 2004).

More recently, the lab has begun to explore the potential of coevolutionary dynamics to create systems capable of open-ended complexity growth (Harrington and Pollack, 2019; Moran and Pollack, 2019). This work was cited by Ken Stanley, a leader in the artificial life community, as of special interest in a survey paper of the field (Stanley, 2019). Work by this author (Willkens and Pollack, 2022) proved that previous hypotheses regarding conditions necessary for open-ended growth in these systems, such as the presence of both cooperative and competitive interactions, in fact do not hold in the general case. Moreover, certain systems were found to be sensitive to the selection method employed, with more sophisticated coevolutionary algorithms producing dramatically greater rates of complexity growth.

Chapter 1: Introduction

The introduction will outline the general aims of the dissertation, summarize the prior literature on the subject, and provide an overview of the necessary terminology.

The intent of this work is help pave the way for the creation of truly open-ended systems through a careful statistical analysis of a set of novel yet simple systems known to exhibit open-ended complexity growth. We identify these systems as *coevolutionary prediction games*. They are some of the simplest systems known to produce such behavior and are therefore ideal for statistical analysis. The intent is to better describe the dynamics of these systems and identify general principles that may apply to more complex applications such as generative modeling and multi-agent reinforcement learning problems.

The substantive contributions of this work include the following:

1. The first analytical treatment of prediction games considered as a general class of coevolutionary systems, employing the MODES Toolbox as the primary framework;
2. The introduction of two new prediction game domains, the *Continuous Prediction Game* and the *Harmonic Prediction Game*
3. The introduction of a new artificial organism model based on the *Genetic Program Parse Tree* (GP-Tree) of Koza (1992).
4. The introduction of a novel analytical technique, *Qualitative Dynamical Evolutionary Stable State Analysis* (Q-DESSA)

In our experiments, we study the coevolutionary dynamics of populations of *artificial organisms* within *coevolutionary ecosystems*. Organism genotypes are represented as *heterogenous graphs* corresponding to phenotypes that take the form of *abstract machines*. A coevolutionary system comprises the populations of organisms and the interactions that take place between within a

given *interactive domain*. These interactions often have either a “cooperative” or “competitive” nature, in which two organisms will have aligned or opposing interests, respectively.

Each interaction, which is performed in pairwise fashion between all members of two interacting populations, yields some measure of *fitness* to each individual involved. Using the results of these interactions, a *coevolutionary algorithm* produces members of the next generation. A *selection function* chooses which members have the greatest likelihood to reproduce. For each class of artificial organisms, we define a set of *mutation operators* to be stochastically applied to each parent to produce a child with a modified genotype; in all our experiments reproduction is performed asexually.

We provide here a summary of the artificial organisms under consideration in this dissertation:

1. The *Deterministic Finite State Machine* (DFSM) introduced in [Moran and Pollack \(2019\)](#)
2. The *GNARL Recurrent Neural Network* with operators defined in [Willkens and Pollack \(2022\)](#)
3. The *Genetic Program Parse Tree* (GP-Tree) based on the Artificial Ant of [Koza \(1992\)](#) with modifications novel to this dissertation

We consider four interactive domains:

1. The *Linguistic Prediction Game* (LPG) introduced in [Moran and Pollack \(2019\)](#)
2. The *Collision Game* introduced in [Willkens and Pollack \(2022\)](#)
3. The *Continuous Prediction Game* novel to this dissertation
4. The *Harmonic Prediction Game* also novel to this dissertation

We consider two selection functions:

1. *Fitness Proportionate Selection* (FPS) where the probability of selection is proportional to the fitness of the organism
2. *Discovery of Search Objectives* (DISCO), a new selection function based on work by the DEMO Lab specifically to address coevolutionary pathologies ([Liskowski and Krawiec, 2016](#))

We employ two analytical paradigms:

1. The *Measurements of Open-Ended Dynamics in Evolutionary Systems* (MODES) Toolbox introduced by [Dolson et al. \(2019\)](#)
2. The *Q-DESSA* framework novel to this dissertation

CoEvo

It is necessary to perform these experiments under the same conditions and with the same analytical framework. For this purpose, the CoEvo framework was developed. Comprising over 8,000 lines of Julia code, the framework is designed to be modular and extensible and is equipped with an extensive test suite. Genotypes are stored on disk in stable JLD2 dataformat along with random seeds and other metadata. It is possible to “revive” organisms at any point in the evolutionary timeline to evaluate their complexity, study their interactions, and continue their evolutionary progress in a reproducible way.

The framework is the first of its kind to provide a unified interface for high-performance simulation and analysis of arbitrary coevolutionary systems. It will be made freely available to researchers and practitioners on GitHub.

Open Science

The disseration shall be undertaken in the spirit of open science movement ([Vicente-Saez and Martinez-Fuentes, 2018](#)). All code and data will be made public to members of the scientific community as it is produced. Progress will be performed in the open-notebook style, with regular status reports accessible via the Internet.

Chapter 2: Analytical Methods and Complexity Growth

Chapter 2 shall outline the statistical methods and analytical approaches to be applied to the experimental results.

The MODES Toolbox

The *MODES Toolbox* is a set of metrics and analytical techniques introduced in [Dolson et al. \(2019\)](#) to better understand the open-ended potential of evolutionary systems in terms of change, novelty, complexity, and ecology. The toolbox is shown to produce results supporting current intuitions and understanding of systems such as rugged NK landscapes and the Avida digital evolution platform ([Lenski et al., 2003](#); [Kauffman and Levin, 1987](#)).

Before applying the metrics to data resulting from an evolutionary system, it is necessary to first screen out deleterious and neutral genotypes so that the focus is placed on truly adaptive mutations. The primary tool for doing so is the application of a *persistence filter*, which has its roots in *coalescence theory*, an area of theoretical population genetics ([Fu and Li, 1999](#)). This is done by identifying the *persistent lineages* according to a sliding window of length t generations. Coalescence theory predicts that in the absence of selective pressure, a single individual is expected to become the sole common ancestor of the entire population in a median time of $2N$ generations, where N is the population size;

this value typically becomes much smaller when selective pressure is present. At each generation $A - t$, each reproductive individual passes on to its offspring a unique identifier, which is passed on their offspring in turn if they are able to reproduce. This continues for t generations until A is reached, whereupon the identifiers are collected from members of the current population such that the persistent individuals from generation $A - t$ can be retrospectively identified.

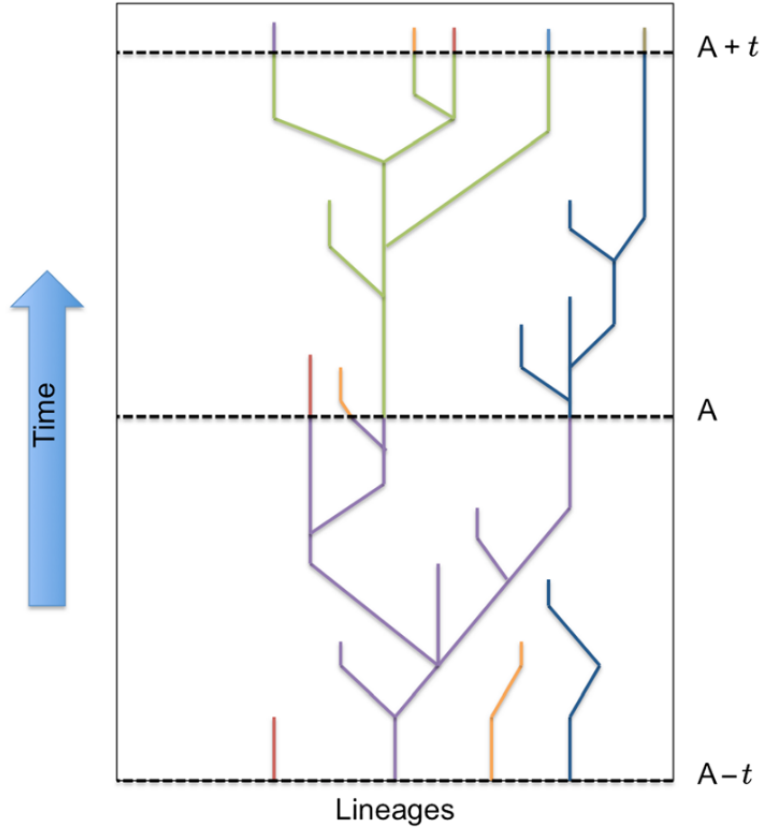


Figure 1: Illustration of persistence filtering taken from [Dolson et al. \(2019\)](#). At timepoint A , the purple lineage has proven to be persistent and therefore the original component from $A - t$ will be considered meaningful. Similarly, the green and blue lineages persist to timepoint $A + t$, and so the original green and blue components will be considered meaningful as they existed at timepoint A .

This step alone dramatically reduce the volume of data for processing, at least by a factor of N , but there is more that can be done. It is recommended to identify the *informative sites* of these persistent genotypes to prune noncoding genetic material. A simple means of doing so is the *knockout* technique: For

each site in the genotype, remove that site and observe the fitness effect for that individual in its original environment. If the fitness either remains the same or increases, this is evidence that the site is nonadaptive or not meaningful and can therefore be pruned. There are downsides to this approach as it fails to capture interactions between sites. This work will explore the possible use of alternative heuristics; preliminary experiments using an incremental fitness-preserving pruning method using information relating to the age of the genes have been promising.

Once the persistent individuals and meaningful genotypic sites have been identified, the core MODES metrics can be applied at each generation interval A given a window size of t :

1. **Change:** A count of the number of unique persistent genotypes at generation A that are different than those at generation $A - t$;
2. **Novelty:** A count of the number of persistent genotypes in generation A that have not been seen at all in the intervals preceding it;
3. **Complexity:** The greatest count of informative sites among all persistent genotypes at generation A ;
4. **Ecology:** The Shannon entropy obtained from measuring the proportion of the persistent population occupied by each unique genotype at generation A . This metric provides an intuitive notion of diversity or evenness, as its value grows higher as a greater number of unique genotypes occupy more equal proportions of the population.

Q-DESSA

The MODES toolbox provides an excellent set of analytical tools for explicating core dynamics of evolutionary systems. However, while we get an overall sense of the growth and behavior of the system, we do not gain a sense of how two systems that exhibit similar MODES dynamics might differ. For example, two systems might both exhibit cycling behavior, evidenced by a negative slope of the novelty metric. But we do not get a sense of the frequency of the cycles, the geometry of the genotype space, or the regions in which those cycles occur.

Q-DESSA (Qualitative Dynamical Evolutionary Stable State Analysis) is a proposed method for analyzing these aspects of potentially open-ended systems. The method employs a graph neural network component trained in an unsupervised fashion over the dataset of persistent organisms filtered using MODES. An advantage of all the models considered in our prediction game is that they all have a natural graph representation, and so this method can be applied to all of them in a similar fashion. We collect together the set of all persistent individuals across all trials of all ecosystems for a given artificial organism type. The *UGraphEmb framework* (Ding et al., 2019) may be employed to obtain a graph-level embedding that reflects the graph edit distance between the genotypes of the organisms. In such a fashion we can derive a vector space in which the embeddings of the two graphs preserve their graph-graph proximity. The resulting

dataset of graph embeddings can be reduced in dimensionality using a method such as *t-distributed Stochastic Neighbor Embedding* (t-SNE) and visualized in a 2D space (Van der Maaten and Hinton, 2008).

Preliminary experiments have shown that the resulting embeddings capture intuitive notions such as that members of one population will tend to be clustered together in the embedding space, distinctly from members of other populations.

We then turn to dynamical systems theory of Strogatz (2000), considering each 2D reduced embedding as a *phase point* in a dynamical system. In this sense, each point taken from a lineage at generation A , $A + t \dots$ forms a *trajectory* in the phase space. The path the trajectory takes is a reflection of the evolutionary dynamics of the system, and by considering the trajectories taken by different species within each ecosystem, we can better understand the evolutionary relationship between each population. It is possible then to identify the presence of evolutionary stable states as well as to classify them according to the terminology of Watson and Pollack (2001):

1. *Disengagement*: When the trajectories of two or more population do not leave some *fixed-point region* of the space. As these systems are out of equilibrium, and there always exists the mutational possibility of components being added or removed from the genome and we do not expect a true fixed point to emerge. But orbits constrained within a region reflect that the populations have disengaged evolutionarily and that no more complex adaptations are expected to occur.
2. *Convention chasing*: When the trajectories of two or more populations orbit between two or more *semi-stable regions*. This is a situation where populations cycle through a limited number of adaptive solutions and better solutions either do not exist or are not reachable through mutation.
3. *Open-ended behavior*: When the transient phase of the trajectory continues indefinitely, and there are no fixed points or semi-stable regions. Open-ended growth necessarily implies open-ended behavior, as a lineage with genotypes that tend to grow over time without bound will not remain within any region.

The combination of MODES and Q-DESSA will give a richer picture of the dynamics of these systems than in the previous literature. Moreover, it allows the measurement of differences between evolutionary trajectories and between trials with different random seeds.

Chapter 3: The Linguistic Prediction Game and the Complexity of DFSMs

This chapter will finish work begun in Moran and Pollack (2019) and continued in a paper currently under review for publication in the 2023 Conference on Artificial Life titled “MODES Analysis of Prediction Games.”

The *linguistic prediction game* is a simple two-player coevolutionary domain. Each player simultaneously emits a bit. If the game is *cooperative*, then players share the same goal: either to match (or mismatch) the bit provided by the other. In a *competitive* interaction, however, only one player is rewarded for matching (or mismatching) bits. Neither player is aware of the nature of their relationship upon starting the game, and their decisions must be made based solely on the patterns of bits produced.

Organisms in the game take the form of *deterministic finite state machines* (DFSMs). Each state has either a 0 or 1 label, corresponding to the bit emitted at that state, and two transition links. The bit emitted by the other player determines which of the links to choose to transition to the next state. Due to the discrete nature of the machines, a loop eventually must occur. The simulation ends once such a loop is detected, and the score for each player is determined by taking the average of all their scores in the loop.

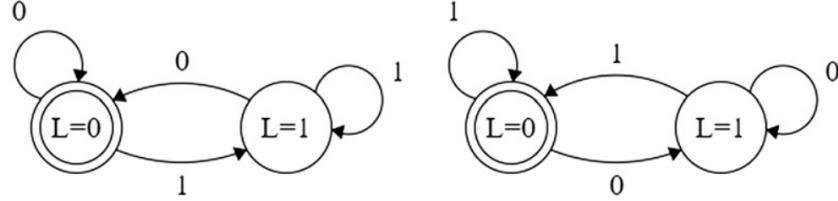


Figure 2: Two examples of DFSM organisms. Node labels indicate the symbol to be emitted at that state, while transition labels indicate the symbol emitted by the other that causes the transition to occur.

Evolution begins with identical populations consisting of a single *start state* with two links to itself. There is an equal probability that one of four mutations will occur: (1) Add state, (2) remove state, (3) flip state label, (4) reassign link.

The ecosystems to be studied include the following:

1. In the *Control* ecosystem, populations evolve without the presence of selective pressure. The results are used to interpret the open-ended behavior of the other ecosystems. In general, genetic drift will lead the size of the organisms will increase gradually over time due to the fact that the number of states is bounded at one.
2. The *Two-Species Competitive* (Comp) ecosystem involves two populations: one whose members seek to match bits, and another whose members seek to mismatch.
3. In the *Two-Species Cooperative* (Coop) variety, both seek either to match or to mismatch.

4. In the *Three-Species Mixed* (3-Mixed) variety, one “host” species is in a cooperative relationship with one population and a competitive relationship with a “parasitic” other.

It was argued that ecosystems possessing only cooperative or competitive interactions tend to plateau in terms of complexity, there defined as the count of the number of states once a DFSM is pruned using *Hopcroft’s algorithm* (Hopcroft, 1971). Meanwhile, *mixed* ecosystems were believed to hold the potential for *unbounded growth*, there determined as when the slope of the line of best fit to the mean complexity growth of a population exceeds that of a control or “shadow” ecosystem that lacks selection pressure.

A number of hypotheses were proposed to explain the behavior of the systems. These include that primarily cooperative species are more likely to exhibit growth; that *convention chasing* may be responsible for cycles between competitive species, leading to *evolutionary stable states*; and that clusters of competitive interactions appear to choke growth in larger configurations, leading to *degenerate ecosystems* (Ficici and Pollack, 1998; Watson and Pollack, 2001). However, more comprehensive theoretical and statistical justification was left for future work, as the volume and apparent intricacy of the machines and their interactions made manual inspection and analysis intractable over evolutionary timescales.

Goals

This chapter will employ MODES Toolbox, the Q-DESSA method, and the DISCO selection method to answer the following open scientific questions:

1. Is the complexity growth observed in mixed ecosystems adaptive or non-adaptive?
2. Can competitive ecosystems be statistically shown to demonstrate convention-chasing or cycling behavior?
3. Preliminary experiments suggest that the original work by Moran and Pollack (2019) failed to adequately characterize the behavior of various ecosystems. Can the underlying dynamics be better characterized and understood?

Chapter 4: The Collision Game and the Growth of GNARL Networks

The Collision Game

The *collision game* is a parameterized finite-horizon two-player game introduced in Willkens and Pollack (2022) that takes place on a real-valued one-dimensional number line. On timestep zero, each agent is assigned a position on the line such that they are some distance apart. At every timestep, each agent emits two

	<i>Collision</i>	<i>Evasion</i>
Affinitive	(1, 1)	(0, 0)
Avoidant	(0, 0)	(1, 1)
Adversarial	(1, 0)	(0, 1)

Table 1: Payoffs for the collision game

real-valued outputs between -1.0 and 1.0. The first is a *movement* value, which is summed with the agent’s current position value to determine its position at the next timestep, while the second is a *communication* value.

At the start of the next timestep, each player is provided with the distance between their two respective positions along with the communication value of the other agent. Each then determines new action and communication values, and so on. The episode ends upon one of two conditions: If the position of the agent that began with the lesser value becomes greater than or equal to the position of the other, the game is over and the interaction is classified as a *collision*; however, if the agents fail to collide by some specified timestep, it is an *evasion*.

The linguistic prediction game and the collision game differ in various ways: The former is infinite-horizon, with discrete actions and a one-dimensional action space; meanwhile, the latter has finite-horizon episodes, real-valued actions, and a two-dimensional action space. Moreover, the collision game features three classes of interaction:

- *Affinitive*: If the episode terminates with a collision, both agents are rewarded with a single point; otherwise neither agent is rewarded. We classify this as a *cooperative* relationship as both interests are aligned.
- *Avoidant*: Both agents are rewarded if no collision occurs and are given nothing otherwise. This also is a *cooperative* relationship, but one incentivizing a different pattern of behavior.
- *Adversarial*: One species is assigned the role of predator, the other the prey. If the episode ends with a collision, the predator is awarded a point while the prey receives nothing, and vice-versa. This is a *competitive* relationship due to the misalignment of their goals.

We note that the reward in the linguistic prediction game is real-valued, while in the collision game it is discrete. But there are strong similarities between the two games: They can be easily mapped to various ecosystem topologies, agents begin with ignorance of their partner’s identity, different flavors of competitive and cooperative interactions are involved, and there is room for diverse strategies to emerge.

GNARL Networks

The abstract machine chosen to represent the organism model in this experiment is the recurrent GNARL network as described in [Angeline et al. \(1994\)](#). GNARL networks evolve both their weights and topologies starting from a minimal state. Our networks here have two input nodes corresponding to the distance and communication value of the other agent, one bias node, and two output nodes corresponding to the movement and communication actions. The first generation has zero hidden nodes and zero connections.

Each node output is the weighted sum of all incoming connections passed through a nonlinear function. We use the *tanh* activation function for all connections so that output is bound in the range $[-1.0, 1.0]$. Each node retains this output value as a hidden state; in the case of a self-loop, the state value from the prior activation is weighted and likewise included in the sum. The movement action of the *greater* (or rightmost) agent is negated in simulation, so that an output of 1.0 always corresponds with moving towards the other agent and -1.0 with retreat.

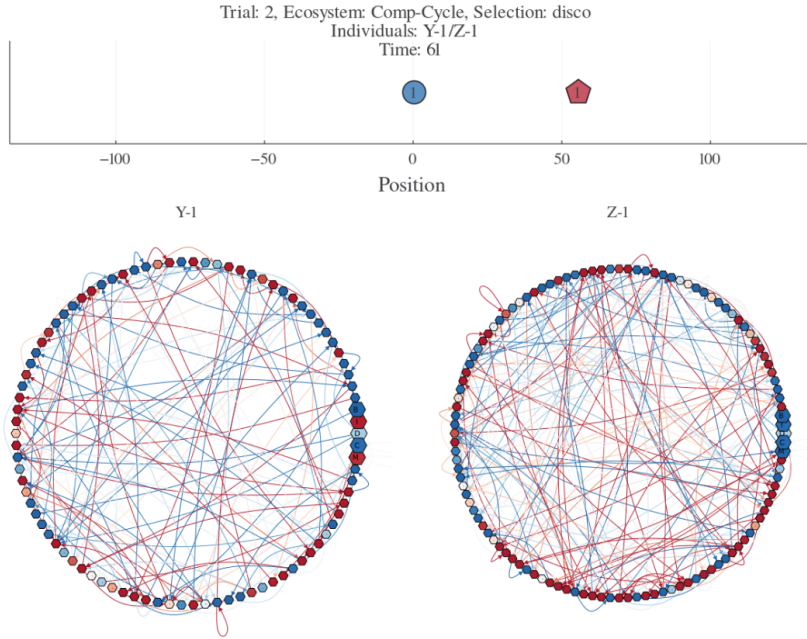


Figure 3: Complex GNARL Networks at one timestep of an interaction in the collision game ([Willkens and Pollack, 2022](#)). The state of each node and its weighted outputs is represented by color, with red being values close to -1.0 and blue being values close to 1.0.

Goals

In [Willkens and Pollack \(2022\)](#), it was found that rates of complexity growth followed completely different patterns than those predicted by [Moran and Pollack \(2019\)](#). Moreover, for all the ecosystem configurations considered, fitness-proportionate selection failed to produce unbounded complexity growth, while the DISCO selection method succeeded in some. This chapter will investigate the following questions using the MODES Toolbox and the Q-DESSA method:

1. How does the DISCO selection method affect evolutionary dynamics such that unbounded complexity growth is produced in the collision game?
2. Is the complexity growth adaptive or nonadaptive by the standards of the MODES Toolbox?
3. Will open-ended transients be observed upon application of the Q-DESSA method?

Chapter 5: The Continuous and Harmonic Prediction Games with Genetic Programming

Here we introduce two new domains predicted to exhibit open-ended characteristics.

The first, the *continuous prediction game*, may be considered a variant of the linguistic prediction game. The game can be envisioned as a pursuit about the unit circle. Each agent has a position corresponding to a point on the unit circle, beginning at the same point. At each timestep, each agent is given a scalar observation corresponding to the arclength distance between the two agents. Each agent then produces a scalar action value, which is applied to produce a rotation of that point on the unit circle. The distance between the two points is measured and subjected to a scoring function and the process repeats for a given number of timesteps. The overall score is the average over all timestep scores in the episode and is illustrated in Figure 4.

In this game, a “cooperative” interaction is defined as one in which each agent is rewarded for minimizing the distance between the two points on each timestep, while in a “competitive” interaction, one will seek to minimize the distance while the other is rewarded for maximizing it.

The second, the *harmonic prediction game*, springs from the observation that the “distance” function in such prediction games is ultimately arbitrary, and that multidimensional outputs and inputs may be interpreted in creative ways. If a domain provides sufficient adaptive pressure and coevolutionary tension, we could expect to observe various forms of open-ended complexity across a wide range of representations and media.

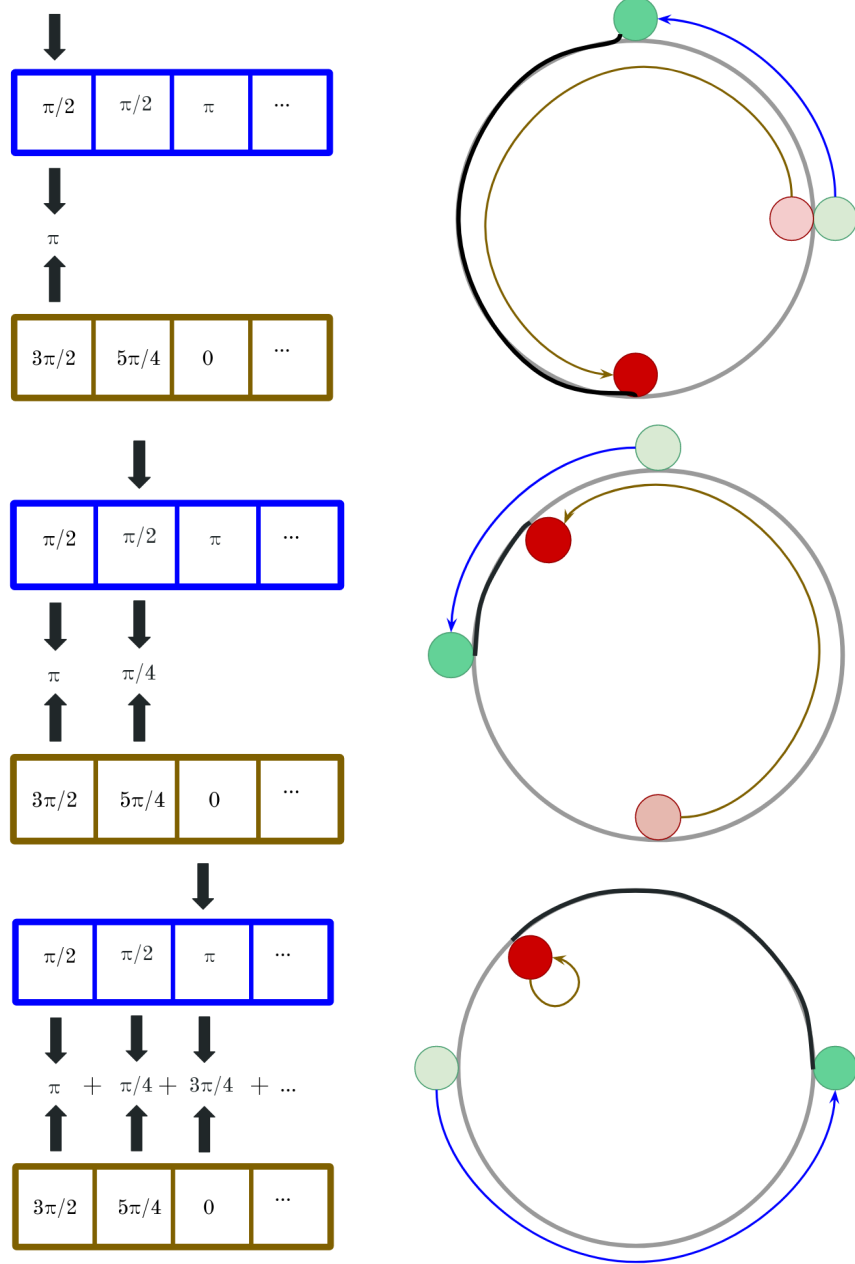


Figure 4: Example of scoring in the continuous prediction game.

The harmonic prediction game expands the scalar observation and action spaces of the continuous prediction game to vectors. Each value in the vector

corresponds to the strength of each harmonic within the overtone series of a synthesized tone. The first value in the vector then corresponds to the frequency of the fundamental, the second to the strength of the second overtone, the third to the strength of the third overtone, and so on. As an alternative “distance” metric, we may instead use a *dissonance model* such as that of Sethares (2005) as presented in Cook (2017) and shown in Figure 5.

The interval created by the harmony of the two tones may be then measured using the dissonance model; a “cooperative” interaction is one in which both are rewarded for producing harmony, while a “competitive” interaction corresponds to when one agent is rewarded for producing harmony, while the other is rewarded for producing dissonance.

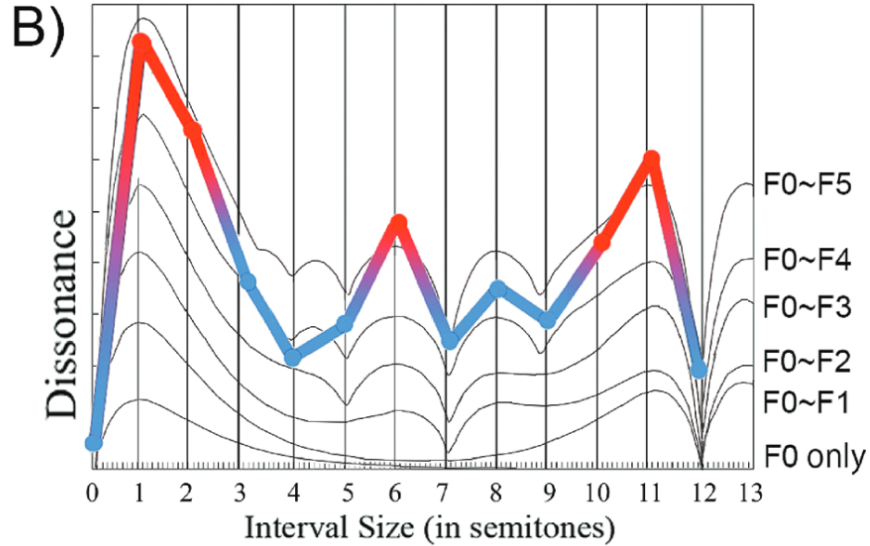


Figure 5: Dissonance model shown in Cook (2017). The colored line represents the empirical curve of perceived dissonance reported by human subjects, while the black lines represents the addition of overtones to the theoretical model of Sethares (2005).

Genetic Programming

As a third substrate, we will use constructs from the *genetic programming* paradigm to represent our artificial organisms. We use the Artificial Ant automata as presented in Koza and Poli (2005) as a basis. The difference is that in addition to protected arithmetic operations, we also include operations for writing values to an output tape and reading values produced by the other agent from its own respective output tape. These tapes will begin with a single entry of real value and grow in length over the course of the episode. They may

be read in a circular fashion, where each read operation moves the read head forward one position, and reaching the end returns the tape to the beginning; see Figure 6 for details. Additional computational power could be provided by equipping each organism with a stack, along with operations for pushing and popping values resulting from intermediate calculations.

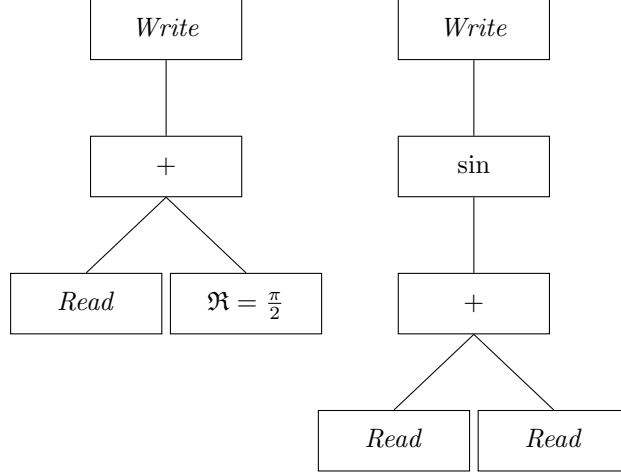


Figure 6: Two simple GP organisms. In this example, the set of functions is $\{Write, +, \sin\}$, while the set of terminals is $\{Read, \mathfrak{R}\}$, where \mathfrak{R} denotes the generation of a random constant. An organism must have a minimum of two nodes: a *Write* node and a *Read* node. In an interaction, each organism begins with an output tape of $[0]$ and operations are performed simultaneously starting with the leftmost terminal and then looping once the root is reached. So the operations performed by the first organism would be $[Read, \frac{\pi}{2}, +, Write, Read, \dots]$. In the event that *Write* and *Read* operations are performed simultaneously, the *Write* operation is performed first. Sixteen timesteps of interaction between the two results in an output tape for the left organism of $[0, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, 1 + \frac{\pi}{2}]$ and an output tape for the right of $[0, 0, 1, 0]$.

Goals

The goals of this chapter are to answer the following questions:

1. How do the continuous and harmonic prediction games compare to other games in terms of open-ended complexity when performed using analogous ecosystems and the GP representation?
2. Will the results be of “auditory interest” in human terms?
3. How will the interactions compare with human music in terms of statistical complexity?

Chapter 6: Conclusion and Future Work

This section will summarize the results of the previous analysis and discuss various avenues for future work. Depending on the pace of progress, some of this may be included in the final dissertation as experiments are easy to perform using the CoEvo framework. Future work may include:

1. Testing the effects of different organism representations across different interactive domains. For example, we may test the effects of using the GNARL representation in the continuous prediction game, or the effects of using the GP representation in the collision game. It may be that a more powerful computation model is required to produce open-ended complexity in certain ecosystems of the linguistic prediction game.
2. Testing the effects of different selection methods across different interactive domains. It is possible that the DISCO method could unlock complexity growth in the linguistic or continuous prediction games.
3. Testing the effects of varying other hyperparameters, such as the population size. Larger populations may be required to produce open-ended complexity in certain ecosystems.

With a better understanding of how selection methods impact complexity growth and its adaptive nature or nonadaptive nature, we may develop a more principled approach to the design of more powerful coevolutionary systems. It is the hope of the author that this dissertation will contribute to the ultimate effort of developing truly open-ended artificial systems.

References

- Adams, A., Zenil, H., Davies, P. C., and Walker, S. I. (2017). Formal definitions of unbounded evolution and innovation reveal universal mechanisms for open-ended evolution in dynamical systems. *Scientific reports*, 7(1):1–15.
- Angeline, P., Saunders, G., and Pollack, J. (1994). An evolutionary algorithm that constructs recurrent neural networks. *IEEE Transactions on Neural Networks*, 5(1):54–65.
- Arulkumaran, K., Cully, A., and Togelius, J. (2019). Alphastar: An evolutionary computation perspective. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO ’19, page 314–315, New York, NY, USA. Association for Computing Machinery.
- Channon, A. (2001). Passing the alife test: Activity statistics classify evolution in geb as unbounded. In *European Conference on Artificial Life*.
- Cook, N. (2017). Calculation of the acoustical properties of triadic harmonies. *The Journal of the Acoustical Society of America*, 142:3748–3755.

- De Jong, E. and Pollack, J. B. (2004). Ideal evaluation from coevolution. *Evolutionary computation*, 12(2):159–192.
- De Jong, E. D. (2004). The incremental pareto-coevolution archive. In *Genetic and Evolutionary Computation–GECCO 2004: Genetic and Evolutionary Computation Conference, Seattle, WA, USA, June 26–30, 2004. Proceedings, Part I*, pages 525–536. Springer.
- Ding, B., Proximity, P. G., Bai, Y., Ding, H., Marinovic, A., Chen, T., and Wang, W. (2019). Unsupervised inductive whole-graph embed-.
- Dolson, E. L., Vostinar, A. E., Wiser, M. J., and Ofria, C. (2019). The MODES Toolbox: Measurements of Open-Ended Dynamics in Evolving Systems. *Artificial Life*, 25(1):50–73.
- Ficici, S. G. and Pollack, J. B. (1998). Challenges in coevolutionary learning: arms-race dynamics, open-endedness, and medicocre stable states.
- Fu, Y. X. and Li, W. H. (1999). Coalescing into the 21st century: An overview and prospects of coalescent theory. *Theoretical population biology*, 56 1:1–10.
- Harrington, K. and Pollack, J. (2019). Escalation of memory length in finite populations. *Artificial life*, 25(1):22–32.
- Hillis, W. D. (1990). Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D: Nonlinear Phenomena*, 42(1-3):228–234.
- Hopcroft, J. E. (1971). An $n \log n$ algorithm for minimizing states in a finite automaton.
- Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128(1):11–45.
- Koza, J. and Poli, R. (2005). *Genetic Programming*, pages 127–164.
- Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA, USA.
- Lenski, R. E., Ofria, C., Pennock, R. T., and Adami, C. (2003). The evolutionary origin of complex features. *Nature*, 423:139–144.
- Lindgren, K. and Nordahl, M. G. (1994). Evolutionary dynamics of spatial games. *Physica D: Nonlinear Phenomena*, 75(1-3):292–309.
- Liskowski, P. and Krawiec, K. (2016). Online discovery of search objectives for test-based problems. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion*, pages 163–164.
- Moran, N. and Pollack, J. (2019). Evolving complexity in prediction games. *Artificial Life*, 25(1):74–91.

- Noble, J. and Watson, R. A. (2001). Pareto coevolution: Using performance against coevolved opponents in a game as dimensions for pareto selection.
- Ofria, C. and Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial life*, 10(2):191–229.
- Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K. O., and Taylor, T. (2019). An Overview of Open-Ended Evolution: Editorial Introduction to the Open-Ended Evolution II Special Issue. *Artificial Life*, 25(2):93–103.
- Ray, T. S. (1992). Evolution, ecology and optimization of digital organisms. *Santa Fe*.
- Rosin, C. D. and Belew, R. K. (1995). Methods for competitive co-evolution: finding opponents worth beating. In *ICGA*, pages 373–381. Citeseer.
- Sethares, W. (2005). *Tuning, Timbre, Spectrum, Scale*.
- Soros, L. and Stanley, K. (2014). Identifying necessary conditions for open-ended evolution through the artificial life world of chromaria. In *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 793–800. MIT Press.
- Stanley, K. O. (2019). Why Open-Endedness Matters. *Artificial Life*, 25(3):232–235.
- Stepney, S. (2021). Modelling and measuring open-endedness. *Artificial Life*, 25(1):9.
- Strogatz, S. H. (2000). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press.
- Thurner, S., Klimek, P., and Hanel, R. (2018). *Introduction to the Theory of Complex Systems*. Oxford University Press.
- Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).
- Vicente-Saez, R. and Martinez-Fuentes, C. (2018). Open science now: A systematic literature review for an integrated definition. *Journal of business research*, 88:428–436.
- Watson, R. A. and Pollack, J. B. (2001). Coevolutionary dynamics in a minimal substrate. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2001)*, pages 702–709.
- Willkens, T. and Pollack, J. (2022). Evolving Unbounded Neural Complexity in Pursuit-Evasion Games. volume *ALIFE 2022: The 2022 Conference on Artificial Life of ALIFE 2022: The 2022 Conference on Artificial Life*. 9.