

## University of Windsor Scholarship at UWindsor

---

Electronic Theses and Dissertations

---

2010

# K-Means Clustering as a Speciation Mechanism within an Individual-Based Evolving Predator-Prey Ecosystem Simulation

Adam Aspinall  
*University of Windsor*

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

---

### Recommended Citation

Aspinall, Adam, "K-Means Clustering as a Speciation Mechanism within an Individual-Based Evolving Predator-Prey Ecosystem Simulation" (2010). *Electronic Theses and Dissertations*. Paper 315.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

K-Means Clustering as a Speciation Mechanism within an Individual-Based Evolving  
Predator-Prey Ecosystem Simulation

by

Adam Aspinall

A Thesis  
Submitted to the Faculty of Graduate Studies  
through Computer Science  
in Partial Fulfillment of the Requirements for  
the Degree of Master of Science at the  
University of Windsor

Windsor, Ontario, Canada

2010

© 2010 Adam Aspinall

K-Means Clustering as a Speciation Mechanism within an Individual-Based Evolving  
Predator-Prey Ecosystem Simulation

by

Adam Aspinall

APPROVED BY:

---

Dr. Melania E. Cristescu  
GLIER

---

Dr. Alioune Ngom  
School of Computer Science

---

Dr. Robin Gras, Advisor  
School of Computer Science

---

Dr. Luis Rueda, Chair of Defense  
School of Computer Science

September 14, 2010

# DECLARATION OF PREVIOUS PUBLICATION

---

This thesis includes 1 original conference paper that has been previously published by Springer in the Lecture Notes in Computer Science, as follows:

Thesis Chapter	Publication title/full citation	Publication status
<i>Chapter 3, “Classical Speciation” and “K-Means Clustering for Speciation”</i>	<i>K-Means Clustering as a Speciation Mechanism within an Individual-Based Evolving Predator-Prey Ecosystem Simulation</i>	<i>Published</i>

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted

material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

# ABSTRACT

---

Although presence of individual-based modeling in ecology continues to rise, to this date, there has been little to no studies of speciation in an evolving ecosystem simulation. This thesis presents a new method for modeling speciation within a previously created individual-based evolving predator-prey ecosystem simulation. As an alternative to the classical speciation mechanism originally implemented, k-means clustering provides a more realistic method for modeling speciation that, among other things, allows for species splitting, the recreation of the species tree of life, and more in-depth analysis of speciation. This thesis introduces the predator-prey ecosystem simulation with specific emphasis on the speciation mechanism. Moreover, the k-means speciation mechanism is presented, and the improvements it provides, including improved runtime performance and better modeling of biological theories, are provided.

# **DEDICATION**

---

For my beautiful wife, Alice, whose love and support made this thesis possible.

# ACKNOWLEDGEMENTS

---

To my supervisor, Dr. Robin Gras, I owe great appreciation for the help and guidance provided to me since the beginning of my studies in winter 2008. My acknowledgements also extend to Dr. Melania Cristescu, from the Great Lakes Institute of Environmental Research, from whom I received advice and direction in the area of biological speciation.



# TABLE OF CONTENTS

---

DECLARATION OF PREVIOUS PUBLICATION.....	iii
ABSTRACT.....	v
DEDICATION.....	vi
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
<b>CHAPTER</b>	
<b>I. INTRODUCTION</b>	
Individual-Based Models.....	1
Modeling and Ecology.....	2
Classical Modeling.....	3
Pattern-Oriented Modeling.....	4
<b>II. REVIEW OF LITERATURE</b>	
Individual-Based Models.....	6
A Sample of Individual-Based Models.....	10
Biological Concept of Species.....	11
Individual-Based Models with Speciation.....	12
<b>III. DESIGN AND METHODOLOGY</b>	
Evolving Predator-Prey Simulation.....	14
Individuals and the Behaviour Model.....	14
FCM as an Agent's Genome.....	15
Measuring Genetic Similarity.....	18
Classical Speciation Mechanism.....	19
K-Means Clustering for Speciation.....	20
<b>IV. ANALYSIS OF RESULTS</b>	
Performance.....	25
Simulation Runs.....	27

Continuity of Evolution and Discontinuity of Species .....	28
Numbers of Individuals and Species .....	29
Species Spatial Distribution.....	34
Species' Sizes .....	36
Species Splitting .....	38
Physical vs. Genetic Distance.....	41
Frequency of Interbreeding.....	45
Intra- and Inter-Cluster Distances.....	48
Genetic Drift .....	50
<b>v. CONCLUSIONS AND RECOMMENDATIONS</b>	
Conclusions.....	52
Recommendations.....	53
<b>APPENDICES</b>	
Complete Prey Species Tree of Life at Generation 6300 .....	55
Simulation Parameters .....	56
Initial Fuzzy Cognitive Map for Prey .....	57
Initial Fuzzy Cognitive Map for Predators .....	58
<b>REFERENCES.....</b>	<b>59</b>
<b>VITA AUCTORIS .....</b>	<b>64</b>

# LIST OF TABLES

---

TABLE 1. SUMMARY OF PUBLICATIONS ON INDIVIDUAL-BASED MODELS .....	9
TABLE 2. MATRIX OF EDGES, L, FOR THE SAMPLE FUZZY COGNITIVE MAP (FIGURE 1). ..	17
TABLE 3. THE LENGTH OF TIME NEEDED TO COMPLETE PREY SPECIATION.....	26
TABLE 4. RUNS OF THE SIMULATION USING BOTH SPECIATION METHODS .....	27
TABLE 5. AVERAGE NUMBERS OF INDIVIDUALS AND SPECIES USING BOTH SPECIATION METHODS WITH STANDARD DEVIATIONS SHOWN IN PARENTHESES.....	31
TABLE 6. BASIC STATISTICS OF BOTH SPECIATION METHODS. ....	33
TABLE 7. GENETIC DISTANCE BETWEEN PREY SPECIES 15 AND 37 DURING RUN NUMBER 3 USING 2-MEAN SPECIATION. ....	51

# LIST OF FIGURES

---

FIGURE 1. SIMPLE EXAMPLE OF A FUZZY COGNITIVE MAP. ....	16
FIGURE 2. THE AVERAGE NUMBER OF PREY AND PREDATORS FOR RUNS USING 2-MEAN SPECIATION. ....	29
FIGURE 3. THE AVERAGE NUMBER OF PREY AND PREDATORS FOR RUNS USING CLASSICAL SPECIATION. ....	29
FIGURE 4. THE AVERAGE NUMBER OF PREY AND PREDATOR SPECIES FOR RUNS USING 2-MEAN SPECIATION. ....	30
FIGURE 5. THE AVERAGE NUMBER OF PREY AND PREDATOR SPECIES FOR RUNS USING CLASSICAL SPECIATION. ....	31
<i>FIGURE 6.</i> CROSS-CORRELATION BETWEEN THE NUMBER OF PREY AND PREY SPECIES. ....	34
<i>FIGURE 7.</i> VISUALIZING THE LOCATIONS OF INDIVIDUALS (2-MEAN SPECIATION). ....	35
<i>FIGURE 8.</i> VISUALIZING THE LOCATIONS OF INDIVIDUALS (CLASSICAL SPECIATION). ....	35
<i>FIGURE 9.</i> LOCATIONS OF INDIVIDUALS IN TIME STEP 4500 (LEFT) AND 4599 (RIGHT). ....	36
<i>FIGURE 10.</i> AVERAGE SPATIAL SIZE OF PREY SPECIES DURING RUN NUMBER 8 (CLASSICAL SPECIATION). ....	37
<i>FIGURE 11.</i> AVERAGE SPATIAL SIZE OF PREY SPECIES DURING RUN NUMBER 2 (2-MEAN SPECIATION). ....	38
<i>FIGURE 12.</i> COMPLETE PREY SPECIES TREE OF LIFE AT TIME STEP 6300 OF RUN NUMBER 2. ....	39

<i>FIGURE 13.</i> SUBSET OF PREY SPECIES TREE OF LIFE OF RUN NUMBER 6 USING 2-MEAN SPECIATION.....	39
<i>FIGURE 14.</i> SPATIAL SIZES OF PREY SPECIES 26, 81, AND 105 DURING RUN NUMBER 6.....	40
<i>FIGURE 15.</i> PHYSICAL AND GENETIC DISTANCE BETWEEN INDIVIDUALS IN SPECIES 141. ..	41
<i>FIGURE 16.</i> PHYSICAL AND GENETIC DISTANCE BETWEEN INDIVIDUALS IN SPECIES 92. ....	42
<i>FIGURE 17.</i> PHYSICAL AND GENETIC DISTANCE BETWEEN INDIVIDUALS IN SPECIES 11. ....	43
<i>FIGURE 18.</i> LOCATIONS OF INDIVIDUALS IN PREY SPECIES 286 DURING TIME STEP 4546 OF RUN NUMBER 3.....	44
<i>FIGURE 19.</i> LOCATIONS OF INDIVIDUALS IN PREY SPECIES 286 (BLACK) AND 425 (GREY) DURING TIME STEP 4547 OF RUN NUMBER 3. ....	44
<i>FIGURE 20.</i> FREQUENCY OF INTERBREEDING FOR BOTH SPECIATION MECHANISMS DURING RUN NUMBER 3.....	46
<i>FIGURE 21.</i> FREQUENCY OF INTERBREEDING BETWEEN SPECIES DURING TIME STEP 6300 OF RUN NUMBER 3.....	47
<i>FIGURE 22.</i> SUBSET OF THE PREY SPECIES TREE OF LIFE AT TIME STEP 6300 OF RUN NUMBER 3.....	48
<i>FIGURE 23.</i> SUBSET OF THE PREY SPECIES TREE OF LIFE AT TIME STEP 6300 OF RUN NUMBER 3.....	48
<i>FIGURE 24.</i> THE COMPACTNESS OF PREY SPECIES 286 DURING TIME STEP 4546 OF RUN NUMBER 2. ....	49
<i>FIGURE 25.</i> THE COMPACTNESS OF PREY SPECIES 286 (LEFT) AND 425 (RIGHT) DURING TIME STEP 4547 AFTER SPLITTING OF RUN NUMBER 2. ....	49

## INTRODUCTION

---

### Individual-Based Models

Individual-based modeling is a bottom-up approach to simulating the interactions among individuals or groups of individuals, which intends to show a strong influence of the behaviour of individuals on the entire system. For ecological modeling, individual-based models (IBMs) allow for the consideration of the traits and behaviour of individual organisms.

Whereas classical approaches to modeling ecology often ignore individual behaviour and instead consider an entire ecosystem as a whole, individual-based models aim to “treat individuals as unique and discrete entities” (Grimm, 1999). By modeling individuals with varying ages, social ranks, and adaptability, for example, the properties of the system that the individuals represent can begin to emerge. This has a distinct advantage over the classical approach, namely that the assumptions made regarding individual behaviour (such as the desire for fitness and shelter) provide for a more realistic simulation than using a state-variable model that may begin by calculating birth and death rates.

The lack of explicit criteria for differentiating between classical modeling approaches and individual-based models and the computational cost associated with individual-based modeling are frequently described as reasons why individual based models may not provide a new method for modeling ecology (Uchmanski and Grimm, 1996). Those against this approach may feel that individual-based models are merely a

tool for simulating very specific environments. Advocates who favour the use of individual-based models are driven by paradigmatic motivation (Grimm, 1999) where such models may be used to formulate general theories of ecology.

The generality of individual-based modeling is an important area of consideration. As beneficial as a specific model may be, it is often more worthwhile to formulate general theories. The authors of the book *Individual-based Modeling and Ecology* (Grimm and Railsback, 2005) reserve several sections for a discussion about the generality of individual-based models. They describe the difficulty of creating generic ecological models by comparing ecology to physics. "Individuals [of ecology] are not atoms but living organisms" and that because "individual organisms have properties an atom does not have", such as the variation between them and their adaptive behaviour, aiming for generality in ecological models is much more difficult. Despite these reservations, there continues to be a rise in the use of individual-based models (Judson, 1994).

For this thesis, I introduce our evolving predator-prey ecosystem simulation with a focus on the definition of our behavioural model and how it is used to cluster individuals into species. Subsequent to this introduction, I emphasize a focus on the method in which speciation occurs in the simulation as my objective is to present a new method for speciation which is not only computationally less expensive than a previous method but is also strongly similar to what is known regarding biological speciation

### Modeling and Ecology

The goal of modeling is to solve problems or answer questions, according to Grimm and Railsback (2005), who describe three key points for ecological modeling.

Identifying a problem or question that needs to be addressed or answered should be one of the very first steps executed during the modeling process. Merely aiming for “realism”, according to Grimm, is not a strong enough guideline for modeling and should not be the reason why a modeller sets out to represent a system or environment. Modeling an ecological system requires constraints and rules that restrict our attention to the problem. Starfield et al. (1990) provide a comprehensive description on what they believe to be an appropriate process for modeling. They suggest that a modeller begins by phrasing the problem to be solved. Grimm and Railsback agree by saying, "Good science requires good questions" (Grimm and Railsback, 2005). Drawing a simple diagram of the system to be modeled is identified as the second phase. Starfield et al. (1990) suggest that imagining yourself inside the system should be the third step. "What is going on around me?" and "What affects me, and what do I affect?" are two questions that Grimm and Railsback suggest a modeller asks him or herself (Grimm and Railsback, 2005). Identifying the essential variables, outlining any simplifying assumptions, and attacking the problem through the use of many small steps (versus addressing it head-on) are suggested as the three last high-level phases in the modeling process (Starfield et al., 1990).

### Classical Modeling

Uchmanski and Grimm (1996) describe basic models of classical ecology as ones that focus on an "average individual." Uncomplicated life cycles are routinely simulated and the development, metabolism and the aging of individuals is scarcely considered in classical modeling. "Classical models cannot take into account discrete individuals, which create local population non-uniformity that can affect population dynamics and



ecosystem function" according to DeAngelis and Mooij (2005) who describe characteristics of simulations in the context of both classical modeling approaches and individual-based models. Few system characteristics can be simulated in classical models, according to DeAngelis and Mooij, such as some form of implicit learning where, for example, "a predator population could increase its preference for certain prey relative to other prey types" (DeAngelis and Mooij, 2005). However, phenotypic characteristics occur at the individual level and classical models cannot replicate this behaviour to the same degree that occurs in nature.

Traditional modeling techniques often use state variables, such as population density, to describe an environment. Recognizing this drawback, Judson (1994) alleges that models of this type "sometimes produce dynamics that are not realistic." Even as modellers began to add age, size, and organism classes to classical models, this traditional approach does not produce simulations that accurately portray real environments.

### Pattern-Oriented Modeling

It is the belief of Grimm et al. (1996) that "ecological modelling should take its orientation more from real patterns observed in nature." Pattern-oriented modeling forces a relationship between spatial and temporal scales, according to Grimm et al., and that a pattern-oriented model is a tool that assists modellers to create predictions that are more easily tested than the predictions formulated by other modeling techniques.

Pattern-oriented modeling is described as bottom-up by Grimm et al. (2005) because it begins by collecting relevant information about individuals, then proceeds to formulate theories regarding the individuals' behaviour, and finally tests the theories in a

computer simulation that allows the modeller to observe the environmental properties that emerge. This technique forces a modeller to use real patterns observed in nature as an aid during the design of a model or simulation. Moreover, the pattern-oriented modeling process described by Grimm et al. (2005) unifies the concepts of individual-based models with the idea of modeling based on patterns. This method creates a model structure that is optimal, rigorous, and realistic, while sculpting the model into one that has an ideal complexity.

Testing alternative theories about the behaviour of individuals is easier with pattern-oriented modeling. By comparing a model's output with the data retrieved from real patterns observed in nature, modellers are quickly and easily able to draw conclusions about the accuracy of the hypothesis that was tested.

DeAngelis and Mooij (2005) provide a small discussion on how individual-based models can be used to show the emergence of patterns such as the formation of swarms, flocks, schools, herds, and other groups.

For all of these reasons, pattern-oriented modeling has been found to be a large stepping stone in the direction from classical ecological modeling to an individual-based approach.

## CHAPTER II

# REVIEW OF LITERATURE

---

### Individual-Based Models

As one of the original milestone papers to discuss the arrival of a new modeling technique, Huston et al. (1988) isolate two key reasons why classical models "violate" theoretical ecology. Firstly, classical models attempt to describe all individuals in an environment with one variable. This tactic assumes, for example, that individuals do not vary in their behaviour or physiology, which is a fundamental flaw. Secondly, Huston et al. (1988) criticize classical models for not spatially isolating organisms within a system. Individuals are simulated in a way that causes each of them to have an equal effect on each and every other individual. However, it is a generally accepted principle that interactions between individuals occur only between organisms that come into contact with one another.

Huston et al. (1998) elaborate on the benefits of an individual-based approach by describing the effect of the degree to which individuals initially vary. When trees, for example, are at relatively similar heights to begin with, they will grow naturally at the same rate as the competition for light is an equal fight. However, when the initial variance of height is high, the result will be very few large plants and many smaller strained individuals.

The rise of the individual-based model is discussed by Judson (1994) who provides a short analysis of the properties and problems of individual-based models. The degree to which an individual's life cycle will be simulated, whether or not resource

dynamics are taken into account, how the size of the population is represented, and the extent of variability among individuals of the same age are described as several classification criteria for evaluating the effectiveness of an individual-based model (Uchmanski and Grimm, 1996).

Although it is predominantly accepted that individual-based models are providing a new outlook on ecological modeling, an examination of how significant the contributions are reveals that there is little common motivation behind the movement (Grimm, 1999). Grimm describes the use of individual-based models simply as a tool as having "pragmatic motivation." On the contrary, individual-based models that are designed to support theoretical ecology are driven by "paradigmatic motivation" which he describes as the pathway to developing generic IBMs.

In more recent years, as a means of providing a basis for the development of individual-based models, many frameworks are being developed. Railsback et al. (2001) draw the concepts from complex adaptive systems as guidelines that will help "make the design of IBMs less ad hoc." Identifying what behaviours should emerge from the model, outlining what adaptive behaviours are to be simulated, deciding on what measures will be used to test fitness, and determining to what extent individuals are able to predict the outcome of their behaviour are all steps that Railsback (2001) suggests should be executed during the individual-based modeling process.

Modeling tools such as ECOTALK by Baveco and Lindeman (1992) and Baveco and Smeulders (1994), HOBOT by Lhotka (1994), ECOSIM by Lorek and Sonnenschein (1998) and MOAB by Carter and Finn (1999) are all examples of tools that were designed to help develop individual-based models.

In more recent years, there has been some focus on how the environment in an IBM is represented (Bian, 2003). How the environment is represented in a model, says Bian, "is a critical part of individual-based models." Two traditional approaches for simulating the environment are prevalent - the grid model and the patch model. Bian (2003) analyses the implications of using both of these techniques and concludes that how the environment is represented in an individual-based model will have an effect on the data that is produced. Similarly, different scheduling methods for individual-based models will produce varying results (Caron-Lormier et al., 2008).

One of the most significant contributions to the study of ecological modeling is by Grimm and Railsback (2005). Many compositions on the subject cite *Individual-based Modeling and Ecology* by Grimm and Railsback as a book that covers a broad spectrum of topics: a generic modeling process, pattern-oriented modeling, and individual-based modeling. Moreover, it discusses the goals of IBMs, what makes a model an IBM, and many examples of individual-based models. It presents a framework for the design and development of individual-based models and it reserves chapters for the examination of how individual-based models should be analyzed and how the model and the data produced by the model should be communicated and presented.

Table 1.

*Summary of Publications on Individual-Based Models.*

Author(s)	Journal	Title	Contribution
Michael Houston, Donald DeAngelis, and Wilfred Post	BioScience, 1988.	New Computer Models Unify Ecological Theory.	One of the first milestone papers to discuss the emerging technique of IBMs.
Olivia P. Judson.	Trends in Ecology and Evolution, 1994.	The rise of the individual-based model in ecology	Presents an analysis of generality in ecology and its implications for individual-based models.
Peter T. Hraber, Terry Jones, and Stephanie Forrest.	Artificial Life, 1997.	The Ecology of Echo.	Presents a generic individual-based simulation model.
Adam Lomnicki.	Ecological Modelling, 1999.	Individual-based models and the individual-based approach to population ecology.	Presents four factors for describing relations between individuals.
Ling Bian.	Ecological Modelling, 2003.	The representation of the environment in the context of individual- based modeling.	Analyses two approaches to representing the environments in IBMs.
Broder Breckling, Ulrike Middelhoff, and Hauke Reuter.	Ecological Modelling, 2006.	Individual-based models as tools for ecological theory and application: Understanding the emergence of organisational properties in ecological systems.	Analyzes the potential of IBMs and presents a generic framework for IBMs.
Geoffrey Caron- Lormier, Roger W. Humphry, David A. Bohan, Cathy Hawes, and Pernille Thorbek.	Ecological Modelling, 2008.	Asynchronous and synchronous updating in individual-based models.	Investigates two approaches for scheduling and updating IBMs.

### A Sample of Individual-Based Models

Listing and describing all existing individual-based models is not possible in a relatively short survey such as this. Nevertheless, a short summary of several IBMs is presented here with the goal of demonstrating the wide range of applications of this modeling technique.

Fahse et al. (1998) demonstrate their protocol for extracting population parameters from individual-based models with the use of an IBM that simulates "nomadic birds in a heterogeneous landscape," similar to some living in parts of South Africa. Habitat selection by stream salmonids is simulated in an individual-based model by Railsback and Harvey (2002) and Bian (2003) uses an IBM that simulates salmon growth to support her theory that how the environment is represented in an IBM will affect the model's results.

Individual-based models of vegetation are also available. An individual-based model is used in by Breckling et al. (2006) to conduct a risk-analysis of genetically modified plants.

Upwards of 27 individual-based models are given as examples by Grimm and Railsback (2005). They cover an extensive array of topics such as simulating the grouping behaviour of birds and fish, the population dynamics of social animals, the movement and dispersal of trout, the dynamics of plant populations, and the evolving traits of marine fish. However, in each of these models, speciation is not a critical component.

### Biological Concept of Species

The existence of species is the well-discussed introduction in *Speciation* (Coyne and Orr, 2004) in which the authors aim to answer three questions: “Are species real? If so, what are they? Finally, why do they exist?” The authors consider, among other things, *The Origin of Species* by Charles Darwin which describes the possibility that species are neither real nor distinct but are instead “theoretical constructs of the human mind” designed for categorization of individuals. Unarguably, there is a clear distinction between an eagle and a crow, suggests Coyne and Orr, but whether or not this discontinuity constitutes distinct species of birds is up for debate.

In fact, Ridley (1996) suggests that “The fact that independently observing humans see much the same species in nature does not show that species are real rather than nominal categories. The most it shows is that all human brains are wired up with the similar perceptual cluster statistic.” On the contrary, a number of biologists show studies designed to statistically identify the existence of discrete clusters of organisms. Among them are Neff and Smith (1978) who aimed to distinguish between hybrids and parental species of the sunfish. Mayr (1992) conducted an extensive analysis of the discontinuity of plants in Concord Woods, Massachusetts in attempt to demonstrate the existence of discrete clusters of organisms.

The decision of the existence of species may not be concluded simply by experiment. Instead, it seems that the “species problem” (Brookfield, 2002) is not a scientific problem that can be solved conclusively. Brookfield suggests that the notion of species is “about choosing and consistently applying a convention about how we use a word.” For us, we accept and implement the genotypic cluster definition of species



described by Mallet (1995) which is the belief that a “good species” is one that shows genetic isolation from populations of other species.

### Individual-Based Models with Speciation

While the presence of individual-based models continues to rise, to our knowledge there has been very little detailed study on the simulation of various speciation methods within an evolving individual-based ecosystem. Among the few such simulations, J. H. Holland (1995) presented Echo, a platform for modeling complex adaptive agents that are able to collect resources and move to neighbouring sites. However, both the organisms and the speciation methods in Holland’s platform are quite simple, and Hraber et. al (1995) have shown that Echo did not match “exactly with quantitative predictions” when they compared the output data on species diversity with real data observed in nature.

Mamedov and Udalov (2002) recognized the fact that many of the individual-based modeling frameworks demand that ecologists encompass some set of programming skills, and consequently they developed the CENOCON system. Alleging that it is flexible and requires no programming skill at all, Mamedov and Udalov promote CENOCON as a framework that "generates a virtual space, creates and populates the space with individuals" and "manages these virtual entities to act as real components of real ecological communities." CENOCON limits the number of species to 256, all of which must be predefined in an external text file, and provides no means of representing any kind of learning or evolution.

Another artificial life system is Avida (Adami and Brown, 1994) which, within a 2D geometry, models cells, the interactions between them, the breeding of cells, and their

ability to adapt. In Avida, a genome (which the authors refer to as a “string”) is represented as an entirely separate piece of code running on its own virtual computer. During self-replication (or reproduction), a string may be subject to mutations either during or immediately after the copy method is performed, resulting in a new string to be placed in a nearby cell. This, the authors note, “is the driving force of evolutionary change and diversity” and, in fact, is similar to the evolutionary mechanism we implement in our evolving predator-prey ecosystem. However, there is no explicit speciation mechanism implemented in this simulation. Moreover, not only do individuals in Avida not move but each of them also possesses the same predefined fitness function which is a large limitation of an ecological individual-based model. Instead, this restriction suggests that Avida is an optimization problem. For us, however, individual behaviour and adaptation is governed by the variability of the environment in which an individual lives. Because of the constant changing environment, as opposed to an optimization problem, in our simulation, an *ultimately fittest individual* cannot exist.

## CHAPTER III

# DESIGN AND METHODOLOGY

---

### Evolving Predator-Prey Simulation

Introduced by Dr. Gras et. al. (2009), we developed and continue to study the results of an individual-based evolving predator-prey ecosystem simulation that includes a behaviour model using a “fuzzy cognitive map” (FCM) (Kosko, 1986). In our simulation, complex adaptive agents (or, simply, individuals) are either a prey or a predator, inhabiting a world implemented as a  $1000 \times 1000$  matrix of cells. To remove any bias about the edges of our world, we allow both the left and right sides to wrap and the top and bottom to wrap. This causes our world to take the shape of a torus, in which an unlimited number of prey and predators may survive. In addition to prey and predators, every cell in our world may contain some value of grass and meat (both are limited by an upper bound) which can be eaten by the prey and predators, respectively.

At the initial time step, grass is randomly distributed throughout the world and there is no meat available for the predators to eat. When a predator kills a prey, meat is added to the corresponding cell and is made available to eat by the same or surrounding predators.

### Individuals and the Behaviour Model

Individuals in our simulation make decisions based on their behavioural model which is represented by an FCM. Our implementation of a fuzzy cognitive map combines the concepts of a directed graph (Axelrod, 1976), fuzzy logic (Kosko, 1986), external information (Tisseau, 2001), and learning (Tisseau et al., 2006). Although

FCMs have been used to model the behaviour of individuals, to our knowledge they have never been used in an evolutionary system as complex as ours. In fact, in addition to being a mechanism for decision making, the FCM is the basis for our evolutionary platform, and is also the object we use to cluster individuals into species.

In our simulation, an FCM is a directed graph that contains a set of nodes,  $C$ , where each node,  $C_i$ , is a concept and a set of edges,  $L$ , where each edge,  $L_{ij}$  represents the influence of concept  $C_i$  on concept  $C_j$ . Every edge in  $L$  has a weight,  $w$ , such that a positive value corresponds to an excitation caused by one concept onto another, and a negative value corresponds to an inhibition. An edge,  $L_{ij}$ , may exist with weight 0 which represents the lack of influence of concept  $C_i$  on  $C_j$ . Moreover, an activation level,  $a_i$ , is associated with every concept  $C_i$ . Thus, with this implementation, our FCM allows for the representation of concepts that may be updated by an individual's perception about the world around it, such as the distance to nearby friends, foe, and food, and allows for the computation of a decision of action for the agent depending on its perceptions and its internal states. The matrix of all the weights,  $L_{ij}$ , which describes unambiguously the behavioural model of an agent, is considered in our simulation to be the agent's genome.

### FCM as an Agent's Genome

Recall that every individual in the simulation has its own FCM and it is within every FCM that we define an individual's concepts, of which there are three kinds: sensitive (perceptions about their environment), internal (levels of emotion), and motor concepts (physical actions). Using fuzzy logic, an individual's perceptions about the world in which it lives influences its internal sensitive concepts. Afterwards, the activation levels of the individual's motor concepts are the driving force behind deciding

on an action to perform. The degree to which the action is performed is dependent upon the level of activation following some defuzzification of the value itself.

Consider, for example, a very simple FCM which could be used to model an individual's perception about its distance to a foe. Naturally, being close to a foe could largely increase the individual's internal concept of fear which, in turn, may affect that individual's desire to evade from its current location. A diagram of this FCM is below (figure 1) where it can be seen that as the sensitive concept "foeClose" increases, so does the level of "fear" by a value of 1. As one might expect, the sensitive concept "foeFar" has a negative influence on the internal level of "fear" and an increase in the individual's level of "fear" will also increase the individual's desire to "escape".

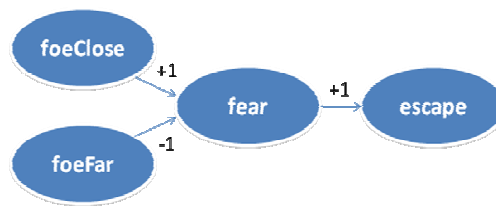


Figure 1. Simple example of a Fuzzy Cognitive Map.

This simple example could be made more complicated by adding a positive self-loop from the concept "fear" to itself. This adds to the model a degree of paranoia or stress such that an individual's level of "fear" at time step  $t_1$  will influence the individual's level of fear at time step  $t_2$ . The corresponding matrix of edges,  $L$ , for the FCM in figure 1 is given below.

Table 2.

*Matrix of Edges,  $L$ , for the Sample Fuzzy Cognitive Map (Figure 1).*

	0	1	2	3
0	0	0	+1	0
1	0	0	-1	0
2	0	0	0	+1
3	0	0	0	0

Recall that  $L$  represents the set of edges present in an individual's FCM and that the edges in  $L$  describe the influence of one concept on another. It is on this set that we implement evolution, in that we allow for the emergence, disappearance, and variation of weights of existing edges to occur at the point at which an individual is born. The learning process for us, then, occurs during the birth of a prey or predator.

A newly created individual inherits a recombination of the genomic material of its parents and with this recombination also exists the possibility of mutation. To model the crossover mechanism, edges are transmitted block by block from one parent to its offspring. More precisely, for each concept, the child inherits all incident edges on this concept from one parent.

With some small probability, a new edge may appear between concept  $C_i$  and  $C_j$  in the FCM of the newly created individual. If this edge has some positive influence on the individual's behaviour, it is a representation of a "fitter" individual. We can observe during the course of the simulation new edges appearing, existing edges disappearing, and existing edges varying in weights within the FCMs of individuals. In fact, because we begin the simulation with every prey having the same FCM and every predator having the same FCM, we can study the evolutionary distance of individuals at time step 10,000,

for example, from the beginning of the simulation and observe how similar or dissimilar the FCMs are among the prey population. After all, this is, indeed, the goal here – we aim to measure the genetic similarity and dissimilarity among individuals and among species during the course of the simulation.

### Measuring Genetic Similarity

Suppose, for example, that there are two prey individuals,  $P_1$  and  $P_2$ , living in the world at time  $t$ . Recall that when each of these prey were created, they inherited a recombination of the genome of their parents – with some possibility of mutation – and that this genome represents the sensitive, internal, and motor concepts for the individual.

We define a distance function,  $D(FCM_1, FCM_2)$ , which computes and returns the numerical distance between two FCMs,  $F_1$  and  $F_2$ , which is a sum of the distances between the weights of matching edges in  $L_1$  and  $L_2$ , the edge matrices for  $F_1$  and  $F_2$ , respectively. For us, this arithmetic distance between two FCMs, or two genomes, is a kind of representation of genetic distance. We use this computation of *genetic distance* between two individuals as a method of clustering individuals into *similar* groups which represent species. Thus, our simulation embodies species as a set of individuals sharing *similar genomes* and how we cluster individuals into groups is what we define as our mechanism for speciation.

We also define an FCM for each species – the average FCM of the individuals contained within that species. Referring to this as a “species centre”, like any other FCM, we are able to apply our distance function,  $D(FCM_1, FCM_2)$ , and use this genetic distance to compare the genetic similarity and dissimilarity of species.

### Classical Speciation Mechanism

In the original implementation, our simulation used a basic mechanism for speciation, called our classical speciation model. For every newborn prey or predator, the numerical distance (a representation of a genetic distance) between the newborn's FCM and the FCMs of every existing species was calculated. If the distance to the closest existing species was below a predefined threshold in the simulation's parameters file, the newborn was assigned to that existing species. Otherwise, a new species,  $S$ , was created with the newborn as its only member. In subsequent time steps, existing individuals may switch to species  $S$  if the genetic distance between the individual and  $S$  is smaller than the genetic distance between the individual and its current species. This implementation caused several limitations.

Our classical speciation created a rigid rule that every new species had an initial size (number of members) of 1. It does not model the basic principle of speciation – that every new species is the result of a splitting of an existing species. Indeed, “most biologists agree that discrete clusters [of organisms] exist” (Coyne and Orr, 2004) and that these clusters form discrete, or near-discrete, species. This phenomenon is observed in our 2-means speciation method.

In addition to this limitation, we were not able to recreate a well-structured tree of life for species. Because our simulation allows for the rare occurrence of interbreeding, a newborn,  $N$ , may have parents from two different species,  $S_1$  and  $S_2$ . If the newborn individual forms a new species,  $S_3$ , then species  $S_3$  will have two parents. With this design, we were not able to extract from our data a tree of life that could represent any kind of species splitting.



Our classical method for speciation was one of the most computationally expensive parts of the simulation. For all individuals, the distance,  $D$ , between their FCM and the FCMs of every existing species was calculated. Moreover, the new map of every species was recalculated. Suppose there are  $N_1$  prey and  $S_1$  prey species during time step  $t_1$ . Then, in our classical speciation mechanism, the complexity of determining the existence of new prey species is  $O(N_1 S_1)$ . This is repeated for checking the emergence of new predator species, resulting in a combined complexity of  $O(N_1 S_1 + N_2 S_2)$ , where  $N_2$  is the number of predators during time  $t_1$  and  $S_2$  is the number of existing predator species during time step  $t_1$ .

#### K-Means Clustering for Speciation

As an alternative mechanism for clustering individuals into *similar* groups, we implemented a k-means clustering technique designed to allow for (1) the splitting of an existing species  $S$  into  $S_1$  and  $S_2$ , and (2) the clustering of individuals that initially belonged to  $S$  into one of either  $S$  or  $S_1$  (thus, more specifically, a 2-means clustering algorithm).

In this new implementation, every newborn individual,  $I_3$ , which is created as the result of reproduction between individuals  $I_1$  and  $I_2$ , is added to the closer of  $S_1$  or  $S_2$ , the species of  $I_1$  and  $I_2$ , respectively. This is accomplished by calculating and comparing the distance  $D(I_3, S_1)$  and  $D(I_3, S_2)$ . The species that is more genetically similar to  $I_3$  gains the addition of  $I_3$  as a member. Speciation then occurs later as a separate method executed within the same time step of the simulation.

Our speciation method begins by finding the individual in a species  $S$  with the greatest distance from the species' FCM. If this distance is greater than a predefined

threshold for speciation, 2-means clustering is performed. Otherwise, species  $S$  remains unchanged. If clustering is to be performed, two new species are created – one centered around a random individual in  $S$ , denoted  $I_r$ , and another centered around the individual in  $S$  that is farthest from  $I_r$ , denoted  $I_f$ . Subsequently, all remaining individuals in  $S$  are added to one of the two new species – whichever species the individual is more genetically similar. After recalculating the new FCMs for the two new species, the process of clustering is repeated for convergence.

After the 2-means clustering is completed, two new species exist,  $S_1$  and  $S_2$ , whose members are a subset of the original members of  $S$ . It is at this point in the speciation method that the closer of  $S_1$  or  $S_2$  to the original species  $S$  inherits the properties of species  $S$ , such as the species ID and the ID of its parent species. Thus, one of the new species will continue to represent the original species while the other will represent a split off of the original species. For example, species number 15 (with 3,000 members) may split into species number 15 (now with 2,500 members) and species number 20 (new species with 500 members).

The first part of our 2-means prey speciation is to determine whether or not clustering should take place. For each prey species,  $S$ , and for every individual,  $I$ , within  $S$ , the distance  $D(I.FCM, S.FCM)$  is calculated. Clearly, this iterates over the number of prey species and the number of prey in each species – which is a complete subset of the total number of prey in the entire world. Thus, this part of our 2-means prey speciation takes  $O(N_I)$  time, where  $N_I$  is the total number of existing prey.

If the largest distance between an individual in  $S_i$  and the centre of  $S_1$  is greater than a threshold, an individual in  $S_i$  is randomly selected - this is  $O(k)$  time, where  $k$  is a

constant. Immediately afterward, the furthest individual from the randomly selected individual is selected. This takes  $O(N_{S_i})$  time, where  $N_{S_i}$  is the size of species  $S_i$ . The creation of two new species,  $S_{i1}$  and  $S_{i2}$ , centered on the two chosen individuals in  $S_i$  takes constant time. 2-mean speciation iterates a predefined number of times (8 was used for the results discussed in this paper). Subsequent to this, all remaining individuals in the current species,  $S_i$ , are grouped into one of the two new species,  $S_{i1}$  or  $S_{i2}$ . This grouping takes  $O(|S_i| - 2)$ , or more simply,  $O(|S_i|)$ .

If we consider that there are  $P_{i1}$  prey in species  $S_{i1}$  and that the size of the matrix  $L$  in each prey is  $n_1 \times m_1$ , then the recalculation of the FCM for species  $S_{i1}$  has complexity  $O(P_{i1}n_1m_1)$ . This, combined with the recalculation of the map for species  $S_{i2}$  creates a total complexity of  $O(P_{i1}n_1m_1 + P_{i2}n_2m_2)$ , where  $P_{i2}$  is the number of prey in species  $S_{i2}$  and  $n_2 \times m_2$  is the size of the matrix  $L$  for each of the prey in species  $S_{i2}$ . However, because the size of  $L$  is constant throughout the simulation, this complexity can be reduced to  $O(P_{i1} + P_{i2})$  or, more simply,  $O(N_{S_i})$ . This is the most computationally expensive part of our 2-means speciation.

The remainder of our 2-means speciation mechanism, lines 19-32, is responsible for assigning a new species ID to either  $S_{i1}$  or  $S_{i2}$  – whichever species is further from the original species  $S_i$ . This takes  $O(k)$  time, where  $k$  is a constant.

The overall complexity is then:

$$O(N_1 + N_2) + \sum_{i=1}^{k_1} P_{i1} + P_{i2} + \sum_{i=1}^{k_2} Pd_{i1} + Pd_{i2}$$

Such that  $N_1$  is the total number of prey,  $N_2$  is the total number of predators,  $k_1$  is the number of prey species for which splitting occurs,  $k_2$  is the number of predator

species for which splitting occurs,  $P_{i1}$  and  $P_{i2}$  is the number of prey in the splitting species and  $Pd_{i1}$  and  $Pd_{i2}$  is the number of predators in the splitting predator species.

The above equation is smaller than  $O(N_1S_1 + N_2S_2)$ , which is the complexity of our classical speciation mechanism. As the two sums are smaller than  $O(N_1 + N_2)$ , simply because the process of speciation is only applied to a subset of the existing species, the total complexity can be reduced to  $O(N_1 + N_2)$ .

The algorithm for our 2-mean prey speciation method can be found below.<sup>1</sup>

1. For each prey species S,
  2. Find the individual, I, with the greatest distance from S.FCM
  3.  $grDst \leftarrow D(I.FCM, S.FCM)$
  4. If ( $grDst > Ts$ ) Then // Where Ts is a predefined threshold for speciation
  5.  $Ir \leftarrow$  random individual in S
  6.  $If \leftarrow$  furthest individual in S from Ir
  7. S1  $\leftarrow$  new species centered around Ir
  8. S2  $\leftarrow$  new species centered around If
  9. For i = 0 to Tc // Repeat clustering for convergence
  10. For each prey, P, in S  $\neq$  Ir or If
  11.  $dst1 \leftarrow D(P.FCM, S1.FCM)$
  12.  $dst2 \leftarrow D(P.FCM, S2.FCM)$
- 

<sup>1</sup> The concepts are easily applied to predator speciation.

```

13.           If (dst1 < dst2) Then
14.               InsertMember(P, S1)
15.           Else
16.               InsertMember(p, S2)
17.           End If
18.       End For
19.       Recalculate the FCMs of S1 and S2
20.   End For
22.   dst1 <- D(S.FCM, S1.FMC)
23.   dst2 <- D(S.FCM, S2.FCM)
23.   If (dst1 < dst2) Then
24.       S1.id <- S.id
25.       S2.id <- next available prey species id
26.       S1.parent <- S.parent
27.       S2.parent <- S.id
28.   Else
29.       S1.id <- next available prey species id
30.       S2.id <- S.id
31.       S1.parent <- S.id
32.       S2.parent <- S.parent
33.   End If
34.   Remove S from the list of prey species
35.   Add S1 and S2 to the list of prey species
36. End If
37. End For

```

## CHAPTER IV

# ANALYSIS OF RESULTS

---

### Performance

Recall that every execution of our simulation takes as input a set of parameters. Among other things, the parameters are designed to allow the system to reach a level of stabilization between the fluctuation numbers of prey, predators, grass, and meat. Stabilization is not achieved immediately; in fact, it is not unusual for the simulation to spend the first 3,000 time steps achieving stabilization. Keeping this in mind, and keeping in mind the fact that our simulation has no definitive end, once the system has reached stabilisation, the length of each execution is entirely dependent upon our needs. It is natural to say, “The more data the better.” However, this may be unrealistic as a single time step in our simulation could require as much as 100 MB of storage space, resulting in a single run of 20,000 time steps requiring a total of 2 TB.

A random sample of 5 time steps for classical speciation and 5 time steps for 2-mean speciation in which 1, 5, 10, 15, 20, and 25 prey species existed are taken. Rounded to the nearest second, the table below depicts the length of time needed to complete the prey speciation method.

Table 3.

*The Length of Time Needed to Complete Prey Speciation*

		Number of Prey Species						
		Sample #	1	5	10	15	20	25
Classical Speciation	1	1s.	4s.	9s.	10s.	50s.	65s.	
	2	1s.	4s.	9s.	13s.	48s.	65s.	
	3	1s.	3s.	10s.	15s.	49s.	64s.	
	4	1s.	3s.	10s.	11s.	49s.	65s.	
	5	1s.	3s.	11s.	11s.	50s.	62s.	
2-Mean Speciation	1	0s.	3s.	2s.	3s.	4s.	3s.	
	2	0s.	1s.	2s.	3s.	4s.	4s.	
	3	0s.	1s.	2s.	3s.	4s.	9s.	
	4	1s.	1s.	3s.	3s.	5s.	4s.	
	5	1s.	1s.	2s.	4s.	3s,	3s.	

Clearly, very little time savings occurs when there are just a few prey species. In fact, it seems that the largest time savings is achieved when there are 10 or more living prey species – which is the norm for any given run of the simulation.

Although table 3 depicts large time savings for running the prey speciation method, because of the variability in runs of the simulation, it would be inaccurate to conclude that a run of the simulation using the 2-mean speciation mechanism costs 1/10<sup>th</sup> of the time needed for a run using the classical speciation mechanism, simply because speciation is just one computation among many that occurs during a single time step of the simulation.

Every run of the simulation discussed in this thesis was executed on the Shared Hierarchical Academic Research Computer Network (SHARCNET: [www.sharcnet.ca](http://www.sharcnet.ca)) which provides to Canadian academic institutions a high-performance computing infrastructure. More specifically, these results were produced using the Narwhal cluster

hosted by the University of Guelph, which provides users with a 4-core CPU and 8.0 GB of memory.

### Simulation Runs

Reference is made throughout this thesis to several runs of the simulation. These runs are illustrated in the following table, listed newest to oldest by speciation mechanism.

Table 4.

#### *Runs of the Simulation Using Both Speciation Methods.*

	Speciation Mechanism	Time Steps	Length of the Run (days)
Run 1	2-Mean	3,238	6
Run 2	2-Mean	9,541	15
Run 3	2-Mean	8,558	16
Run 4	2-Mean	8,245	13
Run 5	2-Mean	12,615	22
Run 6	2-Mean	13,762	26
Run 7	2-Mean	33,641	32
Run 8	Classical	10,323	34
Run 9	Classical	5,632	24
Run 10	Classical	5,189	25
Run 11	Classical	8,168	28
Run 12	Classical	4,292	18
Run 13	Classical	4,276	15

Computing the average number of time steps completed for both speciation methods reveals that when using the 2-mean speciation mechanism, the simulation is able to produce approximately 643 time steps per day while a run with our classical speciation mechanism is able to produce just 258 time steps per day. This represents a runtime performance improvement of 248.8%.



Run number 4 produced an abnormally large population size for both prey and predators (as many as 347,472 prey, for example, compared with 166,943 prey during run number 2 – the run which produced the second highest amount of individuals). This run of the simulation was discarded and not considered in any of the data – averages or otherwise – presented in this thesis.

### Continuity of Evolution and Discontinuity of Species

It is widely understood that, with the exception of a sudden environmental or mutational change, evolution is a continuous process and that “a species is a single lineage of ancestral descendent populations or organisms, which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate” (Wiley, 1978). Wiley continues to promote this belief by saying that “Separate evolutionary lineages (species) must be reproductively isolated from one another.” Although we simulate evolution in a continuous way, a limitation of our simulation is that we are unable to represent the creation of species as being part of the continuity of evolution. The reason for this is simple; at some time step, an evaluation of the genetic similarity and dissimilarity of individuals in a species must be performed and if the degree of dissimilarity exceeds a predefined threshold for speciation, we make the decision to create a new species at that exact moment in time. Nevertheless, the whole process is still continuous and all of the intermediate steps, including partial speciation and sister species with hybridization events, can be observed through the process of our simulation.

### Numbers of Individuals and Species

The average number of prey and predator individuals for the first 10,000 time steps of six runs of the simulation using 2-mean speciation is shown in figure 2. The same averages but rather from runs using classical speciation is shown in figure 3.

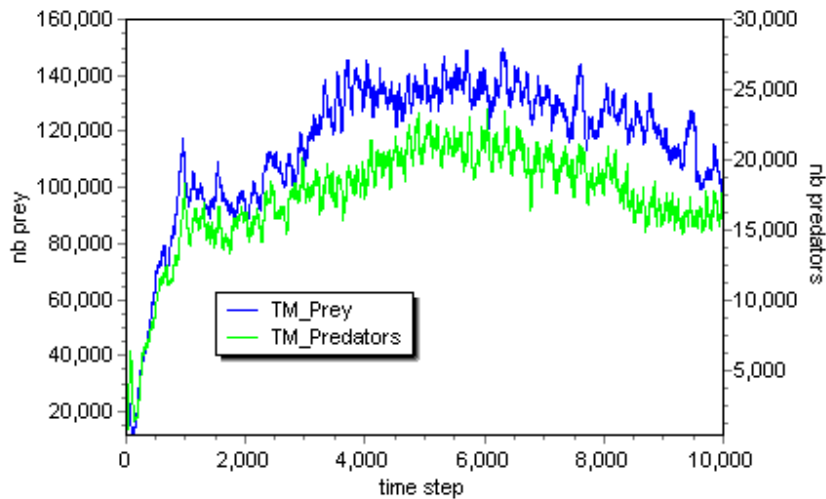


Figure 2. The average number of prey and predators for runs using 2-mean speciation.

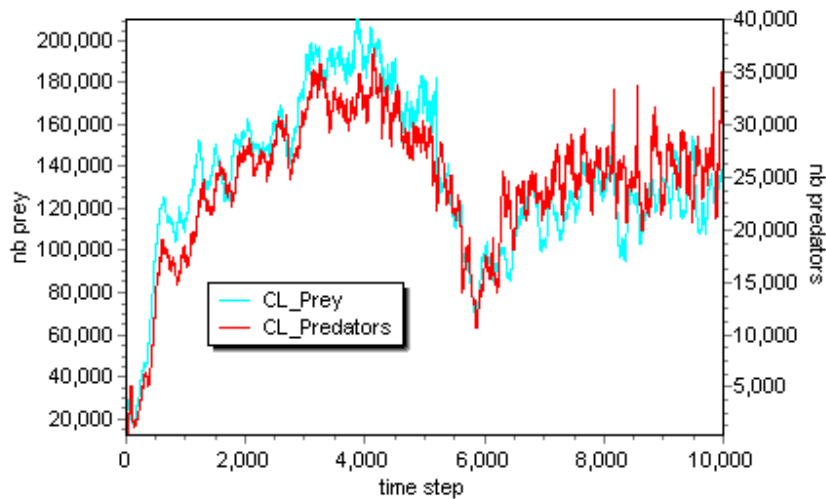


Figure 3. The average number of prey and predators for runs using classical speciation.

On average, runs of the simulation using the classical method for speciation produced larger population sizes for both prey and predators. Consequently, classical

speciation produced more prey and predator species. Also, it appears that there is less fluctuation in the number of individuals when using the 2-mean speciation mechanism, which leads to a more stable system.

It is largely understood (Devaurs and Gras, 2010) and described by fisher's log series (Fisher et al., 1943) that there exists a strong correlation between the size of a population and the number of existing species. This phenomena is predominant in our simulation as the cross-correlation between the number of prey and the number of prey species can be as high as 0.66.

Figures 4 and 5 depict the average number of prey and predator species for 2-mean speciation and classical speciation, respectively.

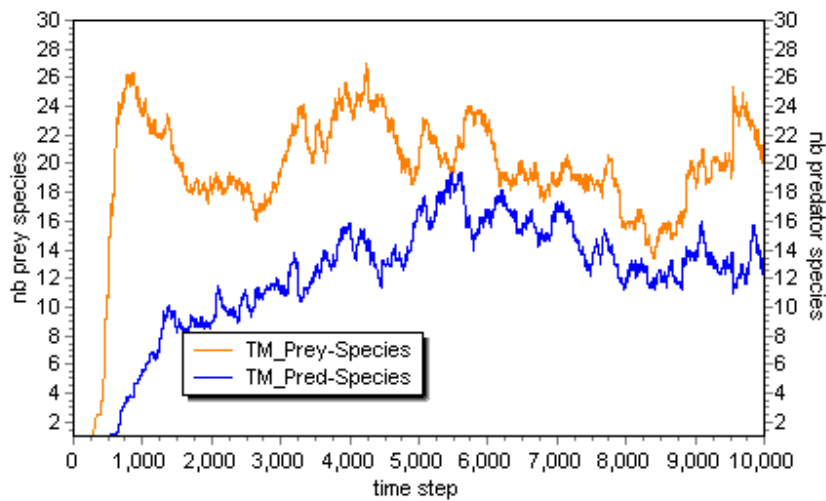


Figure 4. The average number of prey and predator species for runs using 2-mean speciation.

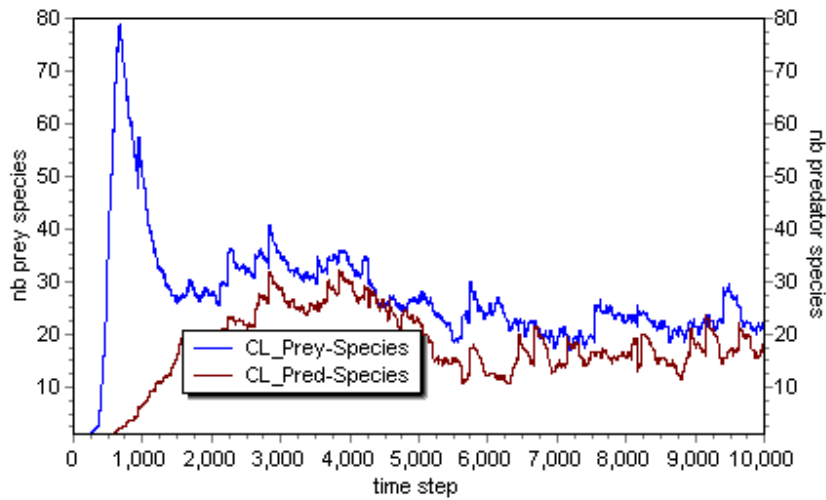


Figure 5. The average number of prey and predator species for runs using classical speciation.

These averages are summarized in table 5.

Table 5.

*Average Numbers of Individuals and Species Using Both Speciation Methods with Standard Deviations shown in Parentheses.*

	2-Mean Speciation	Classical Speciation
Avg. No. of Prey	114,857 (SD = 17,760)	144,629 (SD = 46,404)
Avg. No. of Predators	11,683 (SD = 5,628)	24,148 (SD = 8,890)
Avg. No. of Prey Species	23 (SD = 5)	33 (SD = 14)
Avg. No. of Predator Species	15 (SD = 4)	18 (SD = 10)

The simulation appears to spend approximately the first 3,000 time steps achieving equilibrium. Following this stage, our results show stabled curves for the number of prey and the number of prey species (figure 2). Although the simulation is a large, complex, and evolving system, and although many of the data series show

oscillations with high amplitude, there is a strong correlation between many of the dependent properties (figure 6).

In our classical speciation method, on average, the simulation produced a second prey species during time step 357. The 2-mean speciation method created a second prey species much later, indeed, on average, during time step 438. It can be observed in figure 5 that classical speciation for prey species occurs, on average, very frequently near the beginning of the simulation before stabilizing around 33 prey species. This behaviour is not seen when using our 2-mean speciation. Instead, 2-mean prey speciation occurs gradually towards stabilization around 23 prey species.

When speaking about the size of a species (the quantity of members), it is more useful to use relative sizes by comparing the average size of a species per time step,  $T$ , to the total quantity of individuals during  $T$ . The average prey species size during 2-mean speciation was just 5.85%. The average prey species size during classical speciation was 9.19% of the total population size of the prey, which is 157% larger than that of 2-mean speciation. This degree of difference demonstrates that the 2-means speciation mechanism produced, on average, prey species of smaller size relative to the quantity of prey individuals. According to Devaurs and Gras (2010), who discusses species abundance, we would expect to observe this exact phenomenon. Indeed, it is “widely observed by ecologists that species are far from being equally abundant” (Fisher et al., 1943). Instead, more species are represented with fewer individuals.

Classical speciation produced a maximum of 247 living prey species (time step 688 during run number 12) and 78 living predator species (at time step 2860 of run number 13). 2-mean speciation produced a maximum of 63 living prey species (at time

step 773 of run number 5), which is quite less than that of our classical speciation.

Similarly, 2-means speciation produced a maximum of just 48 living predator species (at time step 5,572 of run number 2), which is also much smaller than the number of predator species created by our classical speciation method. These measurements are summarized in table 6.

Table 6.

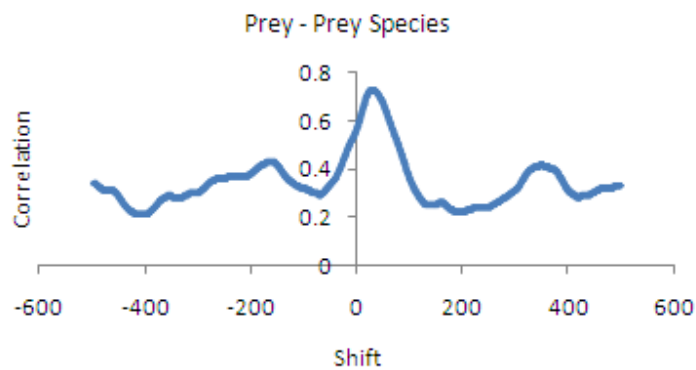
*Basic Statistics of Both Speciation Methods*

	Classical Speciation	2-Mean Speciation
Avg. time step for first prey speciation event	357	438
Avg. time step for first predator speciation event	406	535
Max. number of living prey species	247	63
Max. number of living predator species	78	48

Figure 6 demonstrates the dependency between the prey and prey species data series presented in figure 2. This dependency has been widely discussed and is the basic principle of Fisher’s log series (Fisher et al., 1943) – a species abundance distribution model which describes the dependent relationship between the size of a community and the total number of species within the community.

Computing the cross-correlation function between two data series is a method for measuring the similarity of two waveforms which takes into consideration temporal differences between the two series. The result is a function that returns values between -1 and +1, such that a high positive cross-correlation between two data series at a shift of distance  $d$  suggests that as one series increases or decreases, so does the other  $d$  time steps later. We use this measurement for cross-correlation to show the dependency

between the number of prey and the number of prey species at possible distance shifts of -500 to +500. The resulting function describes for us at what shift the two data series – prey and prey species – are most correlated. The results are revealed in figure 6 which shows a strong positive correlation between the number of prey and the number of prey species, with the highest correlation value of 0.73 occurring at time step 25. This suggests that, as the quantity of prey individuals increases, so does the quantity of prey species approximately 25 time steps later.



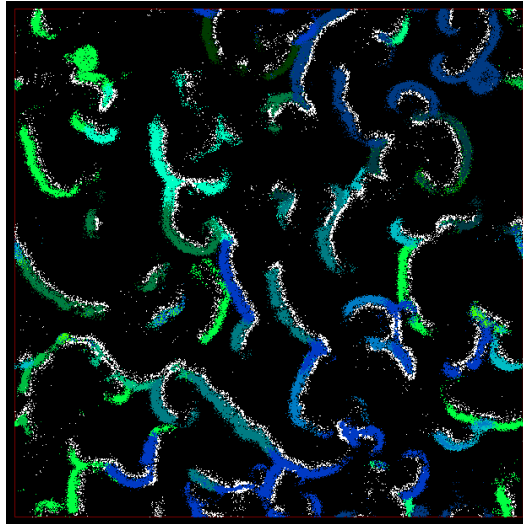
*Figure 6.* Cross-correlation between the number of prey and prey species.

### Species Spatial Distribution

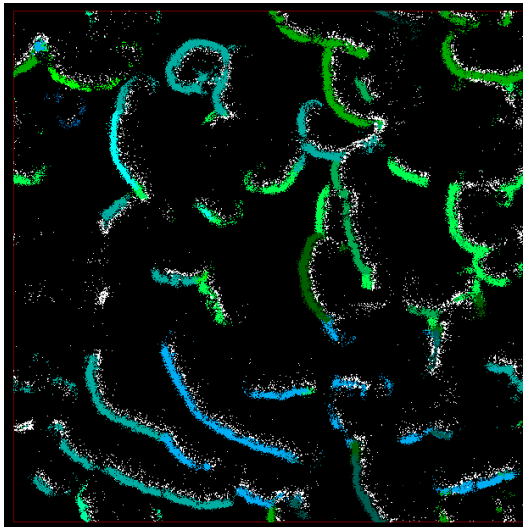
By design, the speciation mechanism used in the simulation has no direct consequence on the spatial patterns and spatial distribution of species in the environment. As a consequence, regardless of the speciation mechanism used – classical or 2-mean – the shapes and positions of species is similar.

Figure 7 depicts the locations of predator individuals (white dots) and prey individuals (coloured dots, such that all prey in the same species are represented using the same colour) for run number 2 of the simulation using 2-mean speciation. Using the same colour scheme (white dots for predators and coloured dots for prey), figure 8

depicts the locations of prey and predators for run number 8, a run which used our classical method for speciation.



*Figure 7.* Visualizing the locations of individuals (2-mean speciation).



*Figure 8.* Visualizing the locations of individuals (classical speciation).

It can be seen that the predators strongly border the prey and influence the direction in which they move. What are interesting are the strong spiral wave patterns that form from the initially tightly packed groups of prey and prey species. Consider, for



example, prey species 46 (from run number 2; using 2-mean speciation) which, during time step 4500, demonstrates a small spiral shape forming. Ninety-nine time steps later, the same prey species has now taken the shape of a much larger and prominent spiral wave (figure 9).

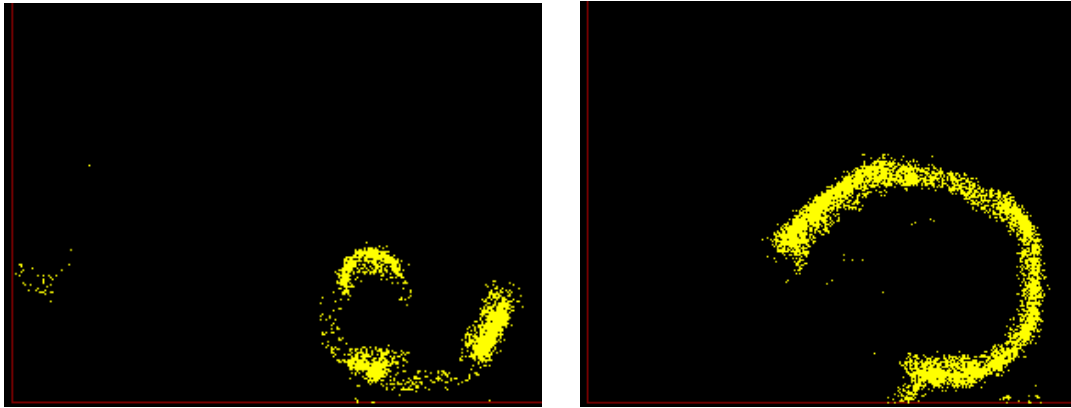


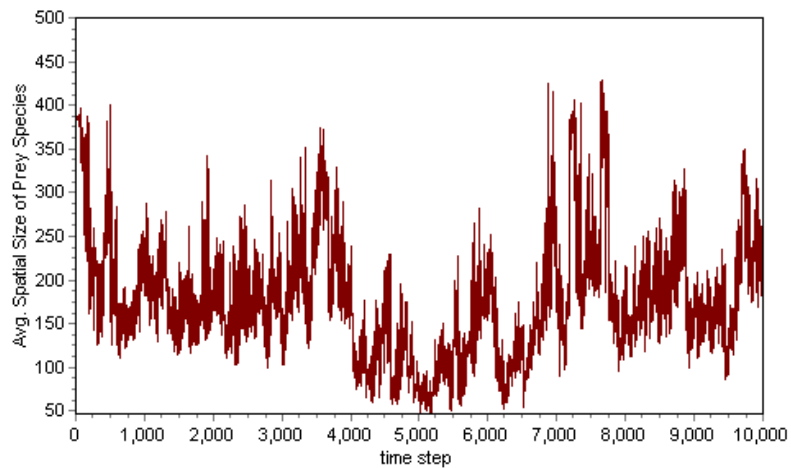
Figure 9. Locations of individuals in time step 4500 (left) and 4599 (right).

In fact, strong and robust spiral waves are a common phenomena among complex and dynamic biological systems (Rohani et al., 1997). Self-organized spiral patterns have been seen not only within chemical reactions but also among populations of bacteria (Rohani et al., 1997) and snowshoe hares in Northern Canada (Bascompte et al., 1997). The phenomenon is predominant throughout the simulation.

### Species' Sizes

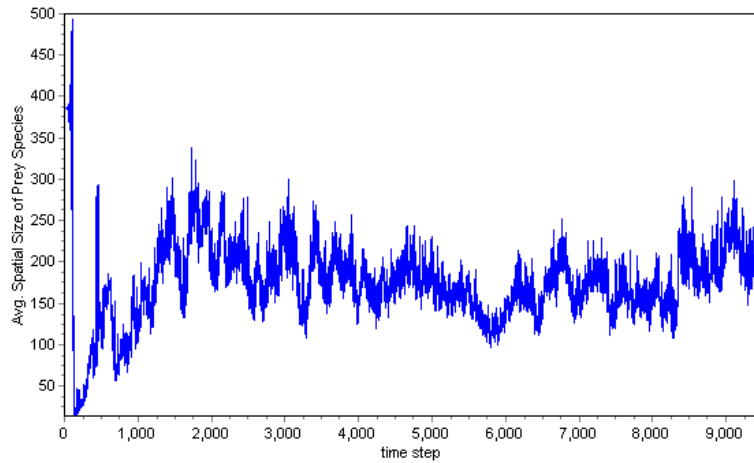
We refer to *size* here in two ways: (1) the *size* of a species  $S$  is the number of individuals in  $S$ , and (2) the *spatial size* of a species  $S$  as the average of the pairwise physical distances in the world between the individuals in  $S$ . Our definition of a species' spatial size allows us to comparatively measure the amount of space in the world that a species occupies.

Consider, for example, run number 8 – the most recent execution of the simulation using classical speciation. The spatial sizes of prey species for the first 10,000 time steps of this run are depicted in figure 10. There is a great deal of variability among the spatial sizes of prey species during the course of this run. In fact, the standard deviation of the average spatial size of prey species during this run is 68.323.



*Figure 10.* Average spatial size of prey species during run number 8 (classical speciation).

The average spatial size of prey species during run number 2 – a run using 2-mean speciation – is shown in figure 11. Clearly, there is a large reduction in the amount of variability among the spatial sizes of prey species during this run. This suggests that 2-mean speciation outperforms classical speciation as a clustering mechanism. The standard deviation of the average spatial size of prey species during this run was just 37.77.



*Figure 11.* Average spatial size of prey species during run number 2 (2-mean speciation).

### Species Splitting

Recall that unlike our classical speciation mechanism, our 2-means speciation allows for the splitting of an existing species  $S$  into  $S$  and  $S_I$ , such that the individuals contained within  $S_I$  are a subset of the individuals originally contained in  $S$ .

Dendroscope, bioinformatics software that is used to visualize phylogenetic trees, was used to produce figure 12 – the recreation of the prey species tree of life at time step 6300 run number 2 – a run using our 2-mean speciation mechanism. The length of a branch represents a distance in time proportional to the generation number in which the speciation event occurred. As a result, it can be seen that the frequency of speciation events is higher near the beginning of the simulation, prior to the environment reaching a level of stabilization. The complete, enlarged version of the prey species tree of life for run number 2 can be seen in Appendix A.

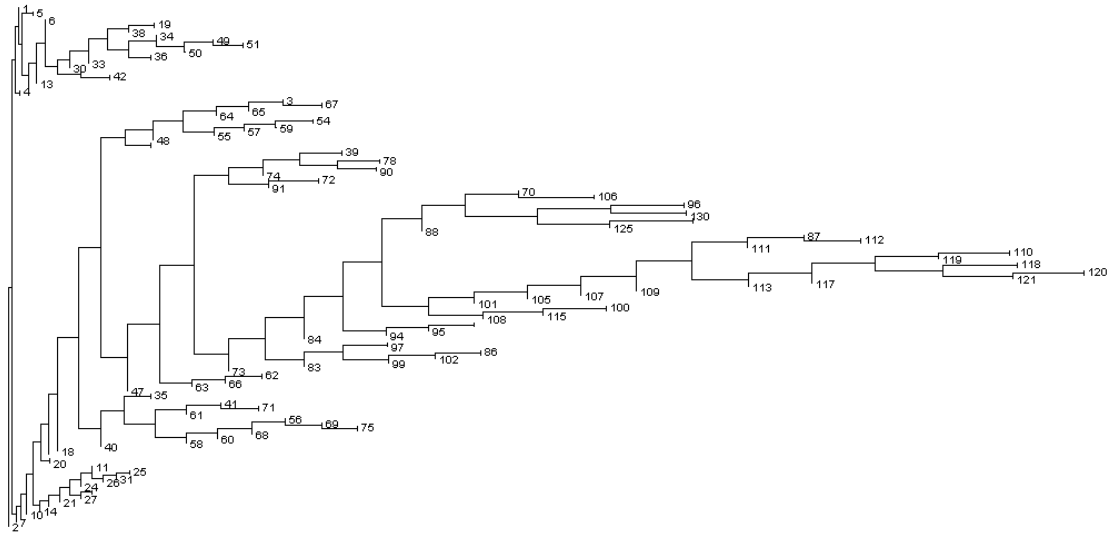


Figure 12. Complete prey species tree of life at time step 6300 of run number 2.

Figure 13 demonstrates a very small subset of the entire prey species tree of life produced from run number 6. The prey species number can be seen at the end of every branch while the number in brackets represents at which time step the species splitting occurred.

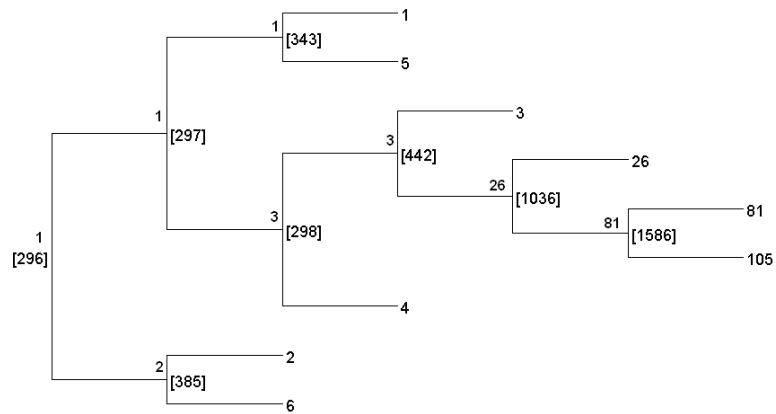
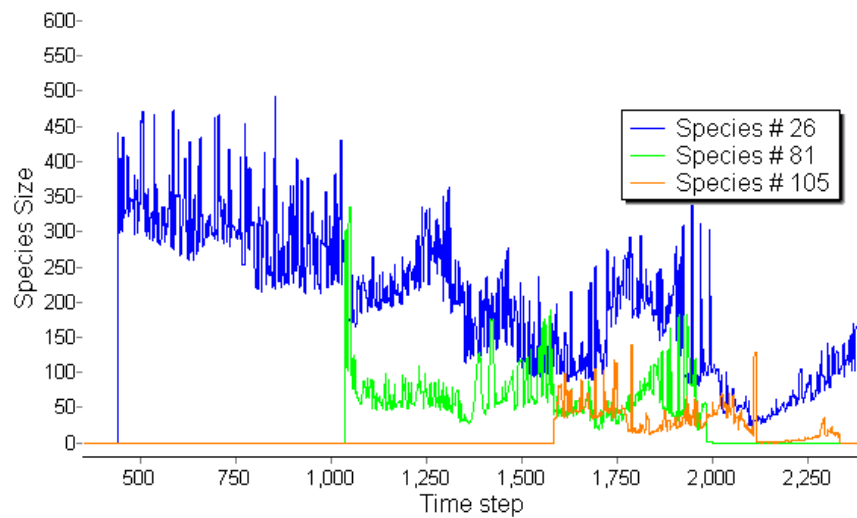


Figure 13. Subset of prey species tree of life of run number 6 using 2-mean speciation.

We use this sample tree to demonstrate the splitting of prey species 26 into species 26 and species 81 (at time step 1036). During time step 1586, prey species 81 splits into species 81 and 105. This splitting can be seen in figure 14, which displays

graphs for the spatial sizes of prey species 26, 81, and 105. It can be seen that prey species 26 emerges during time step 442 (figure 14). This prey species experiences a drop in its spatial size during time step 1036 when it splits into prey species 81 (figure 14). Similarly, prey species 81 experiences a similar drop in its spatial size when it splits into prey species 105 during time step 1586 (figure 14).



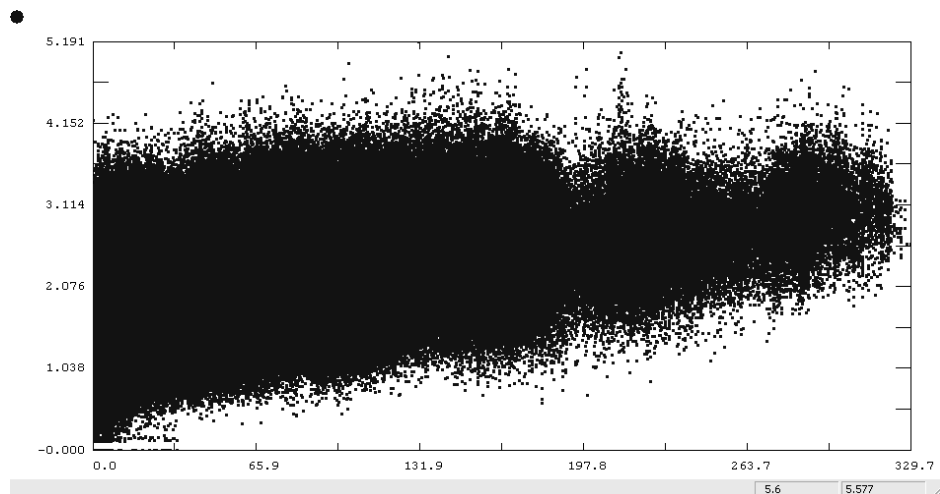
*Figure 14.* Spatial sizes of prey species 26, 81, and 105 during run number 6.

Therefore, with the new 2-means speciation mechanism, we have the ability to make a much more in-depth analysis of the history of the speciation events and correlate them with useful large scale qualitative and quantitative measures associated with each species. For example, we may deduce from figure 14, which illustrates the spatial sizes of species 26, species 28 (which split from species 26), and species 105 (which split from species 28), that there is a correlation between the spatial distances and genetic distances between individuals in a species. This gives us other criteria with which we may compare our speciation methods.

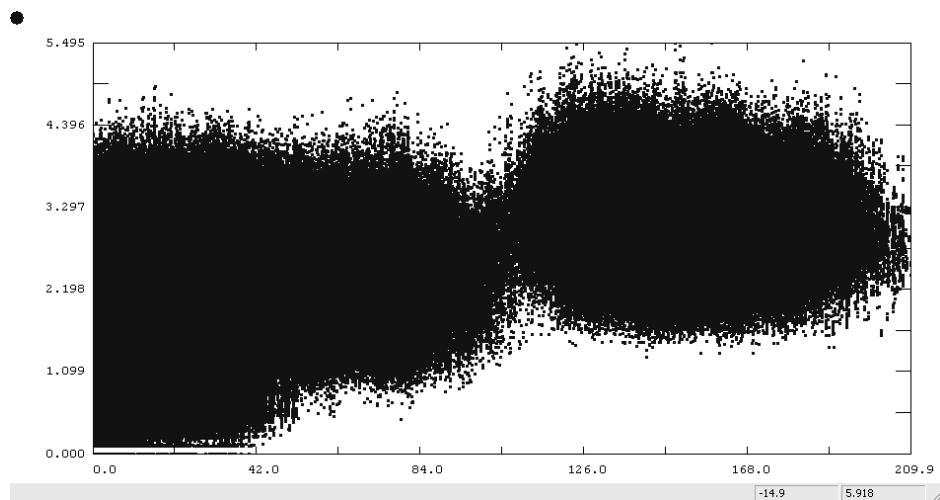
### Physical vs. Genetic Distance

Introduced in 1943 by Sewall Wright, “Isolation by distance” is a biological theory that suggests a positive correlation between physical distances and genetic differences. Subsequent authors, including Kimura and Weiss (1964), Nagylaki (1976), and Slatkin (2007) have continued to study this phenomenon, the last of which demonstrated that on samples of genes from two populations, it is possible to identify isolation by distance.

For every pair of individuals in a species, ( $I_1$ ,  $I_2$ ), measuring the physical distance and genetic distance between  $I_1$  and  $I_2$  demonstrates some evidence of isolation by distance. Depicted in figures 15 and 16, for some prey species, it can be seen that as the physical distance between two individuals,  $I_1$  and  $I_2$ , increases (the x-axis), so does the genetic distance between  $I_1$  and  $I_2$  (the y-axis). For others, however, this correlation is not as evident (figure 17).



*Figure 15.* Physical and genetic distance between individuals in species 141.

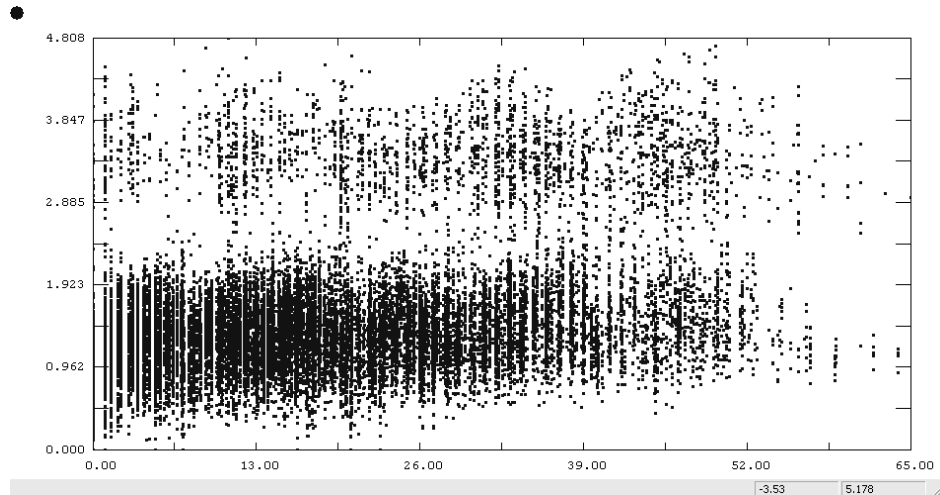


*Figure 16.* Physical and genetic distance between individuals in species 92.

Prey species 141, which contained 1,692 individuals, demonstrates a strong relationship between the physical distance between a pair of individuals and the genetic distance between the same pair of individuals. As the physical distance increases along the x-axis, the genetic distance appears to increase along the y-axis. This suggests that two individuals within this species which are physically separated by a large distance are also likely to be genetically separated by a large distance.

Prey species 92, which contained 2,104 individuals at this time, shows a similar pattern. Although there is little evidence of isolation by distance among all pairs of individuals in the species, there is evidence of genetic differences among the two clusters of individuals in species 92 which are physically isolated from each other (figure 16). Particularly interesting, prey species 11 (the oldest living prey species during time step 3075) demonstrates some genetic isolation among individuals in the species (figure 17). Measuring the correlation coefficient between the physical distance and genetic distance

between individuals in prey species 11 reveals a coefficient of 0.2768 – small, yet still a positive correlation.

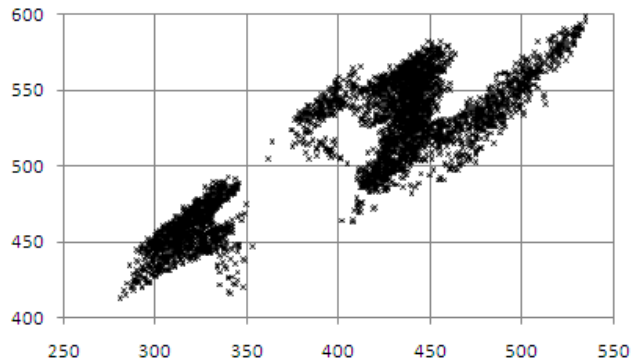


*Figure 17.* Physical and genetic distance between individuals in species 11.

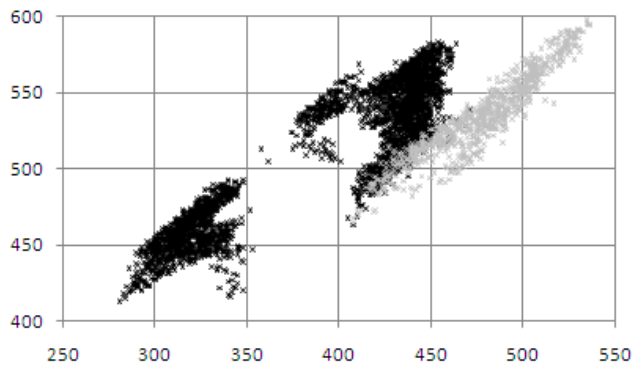
Similar to measuring the correlation between the number of prey and prey species, cross-correlation can be used to quantify the dependency between physical distance and genetic distance among pairs of individuals in a species. The calculations for such a statistic for the three species discussed above is as follows: for species 141, the correlation coefficient measured at 0.4551, for species 92 it was 0.5745, and for species 11, it was 0.2768 which, although small, shows a positive correlation between physical and genetic distance.

Visualizing the physical location of individuals within the world helps us to further identify a relationship between the physical location of individuals and their genetic similarity.





*Figure 18.* Locations of individuals in prey species 286 during time step 4546 of run number 3.



*Figure 19.* Locations of individuals in prey species 286 (black) and 425 (grey) during time step 4547 of run number 3.

Figure 18 depicts the locations of individuals, within the world of  $1000 \times 1000$  cells, belonging to prey species 286 during time step 4546 of run number 3. Immediately following this time step, prey species 286 splits into a new prey species, 425. It can be seen in figure 19 that the new clusters of genetically similar individuals, which form the new prey species, are also physically located near each other. These diagrams further illustrate the belief that there exists a strong relationship between the physical location of individuals and the genetic distance between them. Recall that in order for a set of

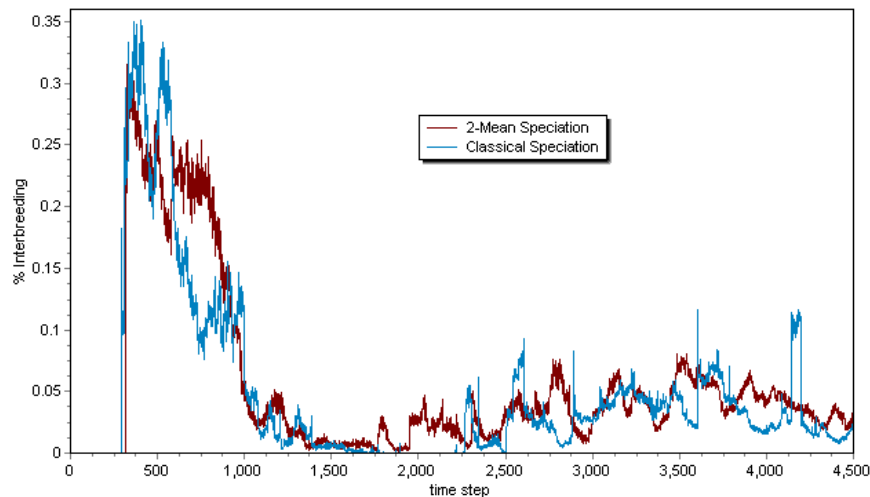
individuals to be clustered together into a new species, they must be genetically similar to each other, and to an individual in the parent species that was genetically dissimilar from the parent species. For this reason, it is known that the FCMs of the individuals forming the new species must be *close* to each other in the world of all possible FCMs – after all, the individuals are genetically similar. Furthermore, it can be clearly seen in figure 19 that the individuals forming the new species are physically located near each other. Thus, it may be concluded that there exists a relationship between an individual's physical distance to all other members in the species and the individual's genetic distance to all other members in the species.

#### Frequency of Interbreeding

Recall that a “good species” is one that shows genetic isolation. However, due to design, the boundaries of our species are not distinct and interbreeding is neither forced nor forbidden. Interbreeding between two individuals of two different species is possible provided that all requirements for reproduction are met (these same requirements, including being physically near each other, wanting to reproduce with each other, and being genetically similar, apply to any reproduction event – not just to interbreeding).

Figure 20 depicts the frequency of interbreeding events for both speciation mechanisms during run number 3 (for 2-mean speciation), and run number 8 (for classical speciation). The y-axis is the ratio of interbreeding events to the total number of reproduction events that occurred during that time step. For example, near the beginning of the simulation when there are just a few new species emerging and when there has been very little genetic diversity, there is a high degree of interbreeding (as much as 35% for classical speciation, for example). There appears to be a stabilization reached for both

speciation mechanisms after time step 1000 when the variability of the frequency of interbreeding converges to an average of approximately 5.40% for 2-mean speciation and 5.35% for classical speciation.



*Figure 20.* Frequency of interbreeding for both speciation mechanisms during run number 3.

Figure 21 displays the frequency of interbreeding events between a pair of individuals in two different species during time step 6300 of run number 3. For example, during this time step, there were 620 interbreeding events between an individual in species 96 and an individual in species 106 (the leftmost bar in figure 21), which are, in fact, direct descendent species from the same original parent species – prey species number 70. Moreover, prey species 96 and prey species 106 are relatively close to each other on the prey species tree of life (a subset of which is shown in figure 22).

This phenomenon is also visible between prey species 110 and 111, which not only demonstrate the second highest amount of interbreeding events, but are also direct descendents of the same original parent species – prey species number 87. A different

subset of the entire prey species tree of life demonstrates the relationship between these descendent species – number 87, 110, and 111 (figure 23).

It is clear that the highest frequencies of interbreeding occur between species sharing a common ancestor. However, because of spatial separation, it is not always guaranteed that related species will experience a high degree of interbreeding. Consider, for example, species 117 and 118 which contain 12,822 and 14,552 prey, respectively (or 7.02% and 7.97% of the entire prey population, respectively). These two species are among the largest living prey species during this time step and yet, due to physical separation in the world, they exhibit the lowest frequency of interbreeding.

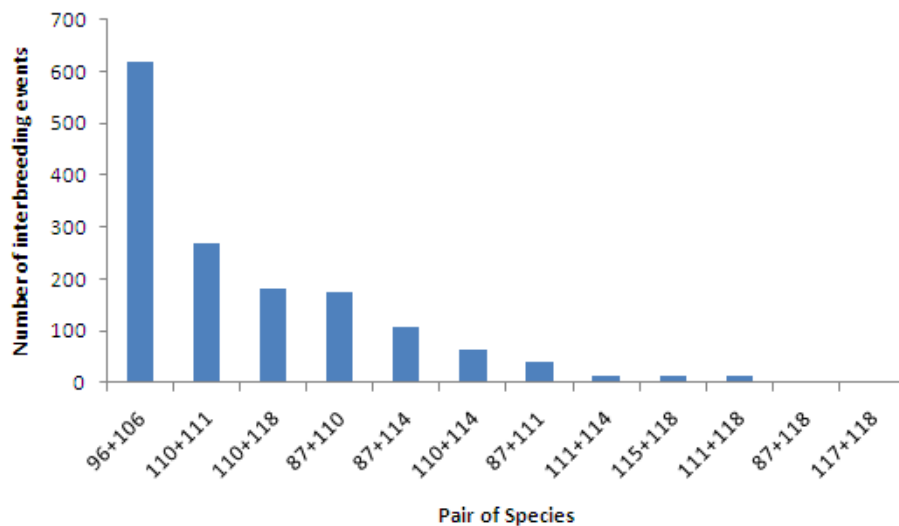


Figure 21. Frequency of interbreeding between species during time step 6300 of run number 3.

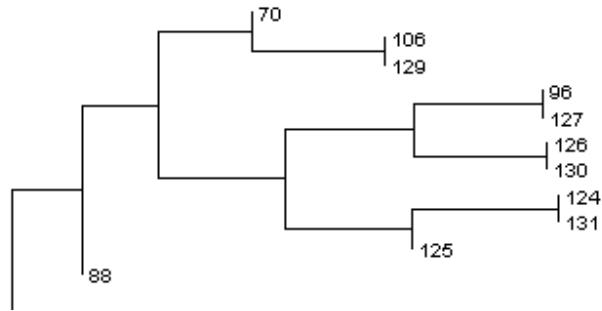


Figure 22. Subset of the prey species tree of life at time step 6300 of run number 3.

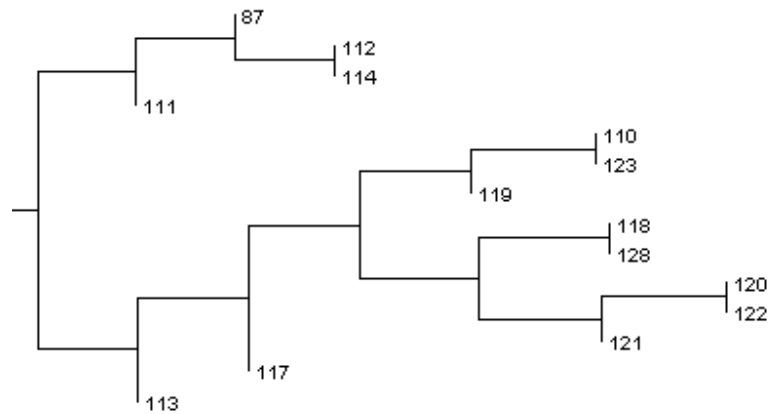


Figure 23. Subset of the prey species tree of life at time step 6300 of run number 3.

### Intra- and Inter-Cluster Distances

Calculating intra- and inter-cluster distances is one method of illustrating and measuring “compactness” of cluster of FCMs. Thus, these distances can be used to reveal how genetically similar individuals are in a single species and also how genetically dissimilar two species are in the world. However, as previously explained, because our classical speciation method is not designed to allow for species splitting, this measurement of cluster compactness before and after species’ splitting can only be used with data from our 2-mean speciation method.

Immediately before the splitting of a species  $S$  (such as species 286 in run number 2; figure 24), there is a high value for intra-cluster distance. This reflects the fact that there is at least one pair of individuals within  $S$ ,  $I_1$  and  $I_2$ , such that the genetic distance between them exceeds our predefined threshold for speciation. For prey species 286, the largest genetic distance between every individual and the center of the species is 2.91041. Moreover, there exists a pair of individuals in species 286 such that the genetic distance between the two individuals is 6.12 – the greatest of every pair of individuals in the species.

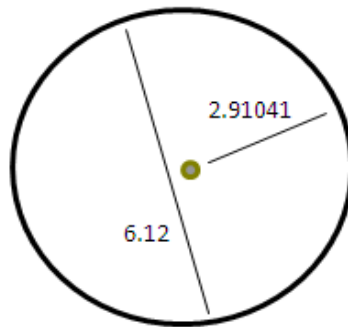


Figure 24. The compactness of prey species 286 during time step 4546 of run number 2.

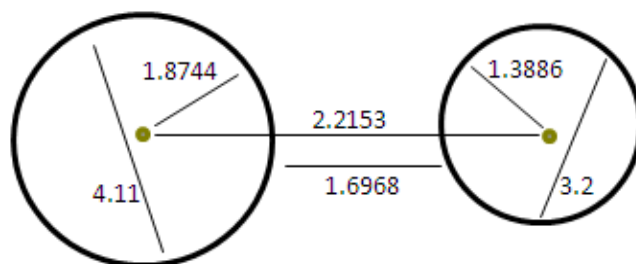


Figure 25. The compactness of prey species 286 (left) and 425 (right) during time step 4547 after splitting of run number 2.

Subsequent to prey species 286 splitting, as seen in figure 25, the largest genetic distance from every individual and the center of the species has reduced to 1.8744. In addition, the greatest distance between every pair of individuals has reduced to 4.11. The new species, species number 425, is even more compact. The greatest distance from an individual in species 425 to the centre of species 425 is just 1.3886 and the largest distance between every pair of individuals in species 425 is 3.2. These results suggest that immediately after a species,  $S_t$ , splits into  $S_{t+1}$  and  $S_{2t+1}$ , the two new species are more compact than the predecessor parent species. Moreover, the genetic distance between the centres of species 286 and species 425 is 2.2153 and the genetic distance between the closest pair of individuals, such that one individual is from species 286 and the other is from species 425, is 1.6968. This reveals genetic separation among the two species post-split.

### Genetic Drift

Genetic drift among species can be demonstrated using the measurement of inter- and intra-cluster distances. Consider, for example, prey species 15 which, during time step 3010 of run number 3, split into prey species 37. At this time, the distance between the centers of species 15 and 37 is 1.93201 – a relatively small value which indicates that these two species are genetically very similar (a result we would expect as species 37 had just split from species 15). Moreover, there exists a pair of individuals – one from species 15 and the other from species 37 – which are genetically quite similar (a genetic distance of 1.14725 between them). One thousand time steps later, the distance between the centres of prey species 15 and prey species 37 has increased to 4.27654 – an increase of more than 220%. In addition, the pair of individuals from both species which are more

genetically similar than any other pair of individuals is separated by a genetic distance of 3.57179. This is strong evidence of genetic drift between two species which were once genetically very similar. These genetic distances are summarized in the following table.

Table 7.

*Genetic distance between prey species 15 and 37 during run number 3 using 2-mean speciation.*

	Time step	
	3010	4081
Genetic distance between centers	1.93201	4.27654
Genetic distance between furthest pair of individuals	1.14725	3.57179



## CONCLUSIONS AND RECOMMENDATIONS

---

### Conclusions

This thesis presents a new mechanism for speciation implemented within an individual-based evolving predator-prey ecosystem simulation. Not only does the 2-mean speciation mechanism have a reduced complexity (when compared to the classical speciation mechanism which, in turn, reduces the runtime performance), it also more accurately models what we understand regarding biological speciation. That is, rather than create a new species centered initially around a single individual – as seen in classical speciation – 2-mean speciation allows for the splitting of an existing species,  $S$ , such that a subset of the original individuals in  $S$  now belong to the newly created species,  $S_1$ . This significant modification is more biologically relevant and also allows us to perform further rigorous studies on speciation, including studying the effect of species splitting on spatial sizes of species, intra- and inter-cluster genetic distances, and genetic drift between a parent and a child species.

It was concluded that the chosen mechanism of speciation has little to no effect on the population sizes of prey and predators, and the number of prey species and predator species. The degree of variability among the population size is due to some amount of randomness in the simulation and its complex and chaotic behaviour (Farahani et al., 2010), and the number of existing species at any given time during the simulation is largely dependent on the size of the prey and predator populations.

Species distribution and spatial patterns were studied and the phenomenon of spiral waves, as seen previously in other systems such as chemical reactions, populations of bacteria, and populations of the Canadian snowshoe hare, were evident. These spatial patterns emerged as a natural result of our system and were not forced by design.

Further analysis revealed a strong correlation in our simulation between the physical location of individuals within a species and their genetic distance to other individuals in the same species.

The frequency of interbreeding was examined and it revealed that species sharing a common ancestor, when able to (i.e. they are physically located near each other), demonstrate a much higher frequency of interbreeding events than observed between species which are genetically distant from each other.

### Recommendations

There are great deals of other studies on speciation that can be performed. Understanding population sizes, species splitting, the relationship between physical and genetic distances, the frequencies of interbreeding, intra- and inter-cluster genetic distances, and genetic drift will be even more necessary if it is decided to study invasive species and the effect on speciation when obstacles are added into the environment, the latter of which is currently being studied by a colleague.

Further analysis on the genetic similarity and dissimilarity between closely related species can be performed. Specific emphasis may be focused on comparing the actual FCM of two sister species, studying specifically the emergence and disappearance of edges in their respective FCMs, and understanding how influential each edge in their respective FCM is on the length of time that each species exists. This study, then, would

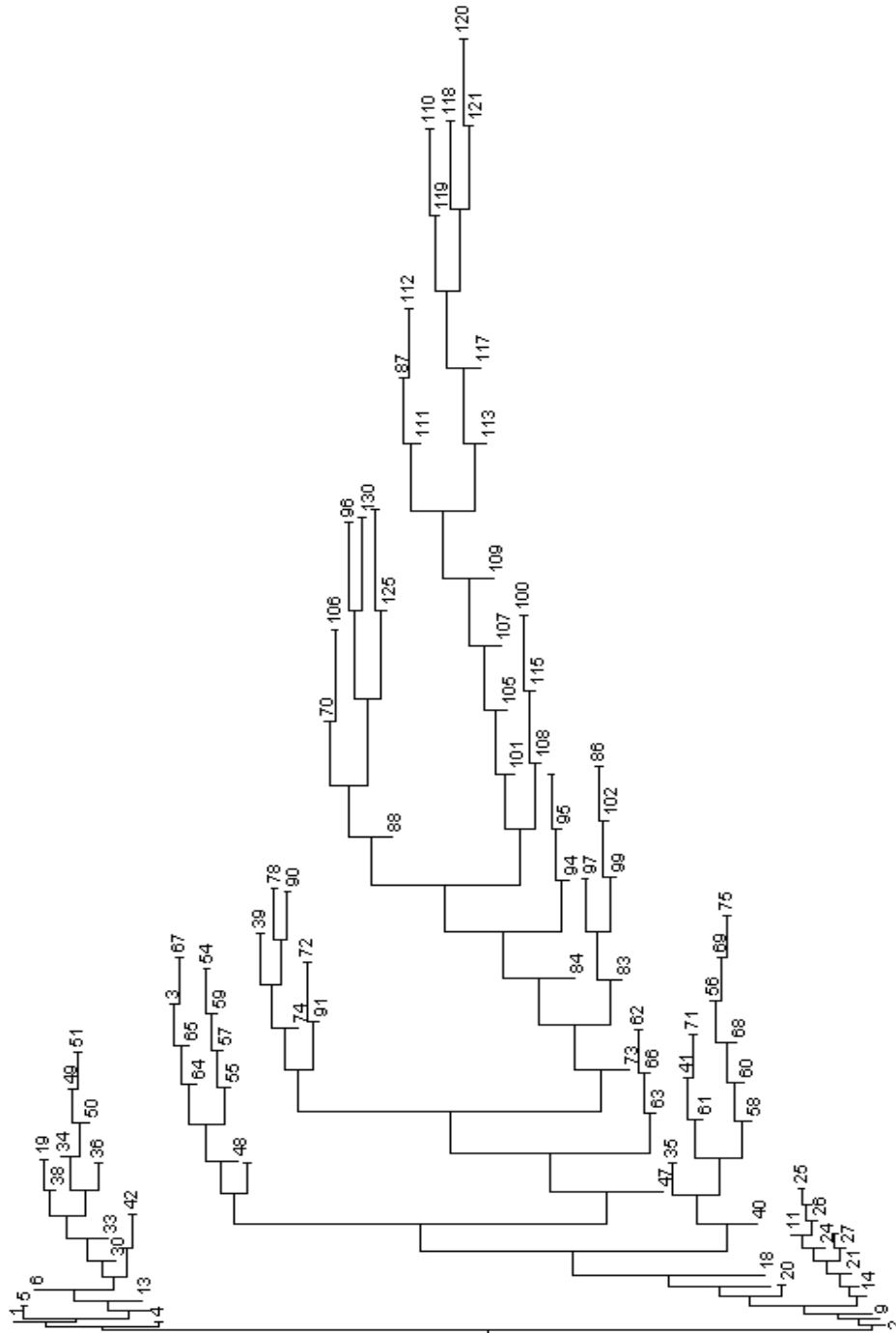
focus a great deal on the details within each FCM – not only on the edges themselves but also on the variation of weights within the FCMs – and aim to understand in more detail how the FCMs of individuals in a species leads to higher-level species behaviour.

Finally, additional studies may be performed to identify or dismiss the existence of co-evolution between prey and predator species. This examination would require identifying a prey species, for example, which is on the evolutionary path towards speciation and following it during the course of the simulation. Questions that could be answered include: how long after a prey species splits does a nearby predator species split? What is the genetic drift between the parent and newly created prey species? Does a nearby predator species which also undergoes speciation demonstrate similar genetic drift? And, what is the relationship between the FCM of a prey species and the FCM of a nearby predator species?

APPENDICES

APPENDIX A

Complete Prey Species Tree of Life at Generation 6300



## APPENDIX B

### Simulation Parameters

Below lists the input simulation parameters which were used to produce the data discussed within this thesis.

Parameter	Value	Parameter	Value	Parameter	Value
Width	1000	MinArc	0.075	EnergyPred	1000
Height	1000	InitNbPrey	13000	SpeedPrey	6
ProbaGrass	0.187	InitNbPredator	350	SpeedPred	11
ProbaGrowGrass	0.0028	DistanceSpeciesPrey	1.5	VisionPrey	20
ValueGrass	250	DistanceSpeciesPred	1.3	VisionPred	25
ValuePrey	500	DistanceMin	0.1	StateBirthPrey	30
MaxGrass	8	AgeMaxPrey	46	StateBirthPred	40
SpeedGrowGrass	0.5	AgeMaxPred	42	nbSensProie	12
MaxMeat	8	AgeReprodPrey	6	nbConceptsProie	7
ProbaMut	0.005	ClusterPrey	10	nbMoteurFixProie	3
ProbaMutLow	0.001	ClusterPredator	3	nbSensPredateur	12
PercentMut	0.15	RayonCluster	5	nbConceptsPredateur	7
PercentMutHigh	0.2	EnergyPrey	650	nbMoteurDepPredateur	4
				nbMoteurFixPredateur	3

## APPENDIX C

### Initial Fuzzy Cognitive Map for Prey

	fear	hunger	searchLeave	curiosity	sedentary	satisfaction	nuisance	escape	searchFood	socialize	exploration	wait	eat	reproduce	move
PredClose	4	0	0	0.1	0	-1	1	0	0	0	0	0	0	0	0
PredFar	-4	0	0	0	0	0.5	-0.5	0	0	0	0	0	0	0	0
FoodClose	0	0.5	0	-0.1	0.1	0.5	-0.5	0	0	0	0	0	0	0	0
FoodFar	0	0	-0.4	0.2	-0.2	-0.7	0.7	0	0	0	0	0	0	0	0
FriendClose	0	0	0.5	-0.1	0.1	0.5	-0.5	0	0	0	0	0	0	0	0
FriendFar	0	0	-0.4	0.2	-0.2	-0.5	0.5	0	0	0	0	0	0	0	0
EnergyLow	0.4	4	-1.5	0	0	-2.2	2.2	0	0	0	0	0	0	0	0
EnergyHigh	0	-1	1.5	0.2	-0.2	1.5	-1.5	0	0	0	0	0	0	0	0
FoodLocalHigh	0	-0.2	0	-0.3	0.3	1.1	-1.1	0	0	0	0	0	2.6	0	0
FoodLocalLow	0	0.2	0	1	-1	-1.1	1.1	0	0	0	0	0	-4	0	0
PartnerLocalYes	0	0	0	-0.4	0.4	0.5	-0.5	0	0	0	0	0	0	1.5	0
PartnerLocalNo	0	0	0.5	0.3	-0.3	-0.8	0.8	0	0	0	0	0	0	-4	0
Fear	0.5	0	0	0	0	0	0	3.5	-0.8	-1	0.3	-1	-1	-1	0
Hunger	0	0.3	0	0	0	0	0	-0.8	2.1	-0.7	0.7	-0.5	4	-1.8	0
SearchLeave	0	0	0.2	0	0	0	0	-0.2	0	1.5	0.5	-0.3	-0.4	3	0
CuriosityStrong	0	0	0	0.1	0	0	0	-0.1	0.5	0.3	1.5	-0.2	-0.3	-0.2	0
Sedentary	0	0	0	0	0.1	0	0	0	-0.5	-0.3	-1.2	0.2	0.3	0.2	0
Satisfaction	0	0	0	0	0	0	0	-0.1	-0.8	-0.2	-2	1.5	0.8	0.7	0
Nuisance	0	0	0	0	0	0	0	0.4	1	0.2	2	-1.2	-0.7	-0.7	0
Escape	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9
SearchFood	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6
Socialize	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
Exploration	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3
Wait	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0	-0.3
Eat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.2
Reproduce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.2

## APPENDIX D

### Initial Fuzzy Cognitive Map for Predators

	chaseAway	hunger	searchLeave	curiosity	sedentary	satisfaction	nuisance	searchPrey	searchFood	socialize	exploration	wait	eat	reproduce	move
PreyClose	0.7	0	0	-0.1	0	0.5	-0.5	0	0	0	0	0	0	0	0
PreyFar	-0.5	0.7	0.1	0.4	-0.4	-0.5	0.5	0	0	0	0	0	0	0	0
FoodClose	-0.5	0.7	0	-0.1	0.1	0.5	-0.5	0	0	0	0	0	0	0	0
FoodFar	0.8	-0.2	0.1	0.2	-0.2	-0.6	0.6	0	0	0	0	0	0	0	0
FriendClose	0	0	0.7	0	0	0.4	-0.4	0	0	0	0	0	0	0	0
FriendFar	0	0	-0.5	0.3	-0.3	-0.4	0.4	0	0	0	0	0	0	0	0
EnergyLow	3.5	5	-1.2	0	0.2	-1.5	1.5	0	0	0	0	0	0	0	0
EnergyHigh	-2	-3	1.4	0.3	-0.3	1	-1	0	0	0	0	0	0	0	0
FoodLocalHigh	-1.5	0.3	-0.2	-0.3	0.3	1	-1	0	0	0	0	0	4	0	0
FoodLocalLow	1.7	0	0.2	1	-1	-1	1	0	0	0	0	0	-5	0	0
PartnerLocalYes	-0.3	0	0	-0.4	0.4	0.8	-0.8	0	0	0	0	0	0	2	0
PartnerLocalNo	0.3	0	0.5	0.3	-0.3	-0.8	0.8	0	0	0	0	0	0	-5	0
ChaseAway	0.2	0	0	0	0	0	0	1.5	-0.2	-0.4	0.3	-0.4	0	-0.4	0
Hunger	0	0.3	0	0	0	0	0	1.5	2.5	-1.2	0.3	-0.4	3.5	-0.8	0
SearchLeave	0	0	0.2	0	0	0	0	-0.8	-0.8	1.5	0.3	-0.5	-0.6	3	0
Curiosity	0	0	0	0.1	0	0	0	0.3	0.3	0.3	1.5	-0.4	-0.3	-0.2	0
Sedentary	0	0	0	0	0.1	0	0	-0.3	-0.3	-0.3	-1.5	0.4	0.3	0.2	0
Satisfaction	0	0	0	0	0	0	0	-0.8	-0.8	-0.2	-1.8	1	0.8	0.8	0
Nuisance	0	0	0	0	0	0	0	1	0.8	0.2	2	-1	-0.6	-0.8	0
SearchPrey	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9
SearchFood	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9
Socialize	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7
Exploration	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
Wait	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0	-0.5
Eat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.6
Reproduce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1.5

## REFERENCES

- Adami, C., Brown, C.T. (1994). Evolutionary Learning in the 2D Artificial Life System, *Avida. Artificial Life*, 377–381.
- Axelrod, R. (1976). *Structure of decision*. Princeton, NJ: Princeton University Press.
- Baveco, J. M. and Lindeman, R. (1992). An Object-oriented Tool for Individual Oriented Simulation: Host-parasitoid System Application. *Ecol. Model.*, 61, 267-286.
- Baveco, J. M. and Smeulders, A. M. W. (1994). Objects for Simulation: Smaltalk and Ecology. *Simulation*, 62, 42-57.
- Bian, L. (2003). The Representation of the Environment in the Context of Individual-based Modeling. *Ecological Modelling*, 159, 279-296.
- Breckling, B., Middelho, U. and Reuter, H. (2006). Individual-based Models as Tools for Ecological Theory and Application: Understanding the Emergence of Organizational Properties in Ecological Systems. *Ecological Modelling*, 194, 102-113.
- Brookfield, J. (2002). Review of Genes, Categories, and Species by Jody Hey. *Genet. Res.*, 79, 107-108.
- Caron-Lormier, G., Humphry, R. W., Bohan, D. A., Hawes, C. and Thorbek, P. (2008). Asynchronous and Synchronous Updating in Individual-based Models. *Ecological Modelling*, 212, 522-527.
- Carter, J. and Finn, J. T. (1999). Moab: a Spatially Explicit, Individual-based System for Creating Animal Foraging Models. *Ecol. Model.*, 119, 29-41.
- Coyne, J.A. and H.A. Orr. (2004). *Speciation*. Sinauer Associates Inc., Sunderland, MA.



- DeAngelis, D. and Mooij, W. M. (2005). Individual-based Modeling of Ecological and Evolutionary Processes. *Annu. Rev. Ecol. Evol. Syst.*, 36, 147-168.
- Devaurs, D. and R. Gras. (2010). Species Abundance Patterns in an Ecosystem Simulation Studied through Fisher's Logseries. *Simulation Modelling Practice and Theory*, 100–123.
- Fahse, L., Wissel, C. and Grimm, V. (1998). Reconciling Classical and Individual-based Approaches in Theoretical Population Ecology: A Protocol for Extracting Population Parameters from Individual-based Models. *The American Naturalist*, 152, 838-852.
- Farahani, Y. Majdabadi, A. Golestani, and R. Gras. (2010). Complexity and Chaos Analysis of a Predator-Prey Ecosystem Simulation. *The Second International Conference on Advanced Cognitive Technologies and Applications*.
- Fisher, R. A., Corbet, A. S., Williams, C. B. (1943). The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. *The Journal of Animal Ecology*, 42–58.
- Grimm, V. (1999). Ten Years of Individual-based Modeling in Ecology: What Have We Learned and What Could We Learn in the Future? *Ecological Modelling*, 115, 129-148.
- Grimm, V., Frank, K., Jeltsch, F., Brandl, R., Uchmanski, J. and Wissel, C. (1996). Pattern-oriented Modelling in Population Ecology. *The Science of the Total Environment*, 186, 151-166.
- Grimm, V. and Railsback, S. F. (2005). *Individual-based Modeling and Ecology*. Princeton University Press.

- Holland, J.H. (1995). *Hidden order: How adaptation builds complexity*. Addison-Wesley, Reading, MA.
- Hraber, P. T., Jones, T. and Forrest, S. (1997). The Ecology of Echo. *Artificial Life*, 3, 165-190.
- Huston, M., DeAngelis, D. and Post, W. (1988). New Computer Models Unify Ecological Theory. *BioScience*, 38(10), 682-691.
- Judson, O. P. (1994). The Rise of the Individual-based Model in Ecology. *Trends in Ecology and Evolution*, 9(1), 9-14.
- Kosko, B. (1986). Fuzzy Cognitive Maps. *International J. of Man-Machine Studies.*, 65–75.
- Lhotka, L. (1994). Implementation of Individual-oriented Models in Aquatic Ecology. *Ecol. Model.*, 74, 47-62.
- Lorek, H. and Sonnenschein, M. (1998). Object Oriented Support for Modeling and Simulation of Individual-oriented Ecological Models. *Ecol. Model.*, 108, 77-96.
- Mallet, J. (1995). A Species Definition for the Modern Synthesis. *Trends in Ecology and Evolution*, 10(7), 294–299.
- Mamedov, A. and Udalov, S. (2002). A Computer Tool to Develop Individual-based Models for Simulation of Population Interactions. *Ecological Modelling*, 147, 53-68.
- Mayr, E. (1992). A Local Flora and the Biological Species Concept. *Am. J. Bot.*, 79, 222-238.
- Neff, N. A. and G. R. Smith. (1978). Multivariate Analysis of Hybrid Fishes. *Syste. Zool.*, 28, 176-196.

- Pejman Rohani, Timothy J. Lewis, Daniel Grunbaum, Graeme D. Ruxton (1997). Spatial Self-organization in Ecology: Pretty Patterns or Robust Reality? *Trends Ecol. Evol.*, 12, 70-72.
- Railsback, S. F. and Harvey, B. C. (2002). Analysis of Habitat-selection Rules Using an Individual-based Model. *Ecology*, 83(7), 1817-1830.
- Railsback, S. F., Lamberson, R. H., Harvey, B. C. and Duy, W. E. (1999). Movement Rules for Spatially Explicit Individual-based Models of Stream Fish. *Ecological Modelling*, 123, 73-89.
- Ridley, M. (1996). *Evolution*. Second Edition. Blackwell Scientific Publications, Oxford, UK.
- Slatkin, M. (2007). Isolation by Distance in Equilibrium and Non-Equilibrium Population. *Evolution*, 54, 1606–1613.
- Starfield, A., Smith, K. A. and Bleloch, A. L. (1990). How to Model It: Problem Solving for the Computer Age, McGraw-Hill.
- Tisseau, J. (2001). Réalité virtuelle—Autonomie in virtuo. *Unpublished habilitation à diriger les recherches dissertation*, University of Rennes, France.
- Tisseau, J., Parenthoën, M., Buche, C., & Reignier, P. (2006). *Comportements perceptifs d'acteurs virtuels autonomes. Une application aux cartes cognitives floues*. *Revue des Sciences et Technologies de l'Information, Série Techniques et Sciences Informatique*, 24(10), 1259–1293.
- Uchmanski, J. and Grimm, V. (1996). Individual-based Modelling in Ecology: What Makes the Difference? *Trends in Ecology and Evolution*, 11(10), 437-441.

Wiley, E. O. (1978). The Evolutionary Species Concept Reconsidered. *Syst. Zool.*, 27,  
17-26.

## VITA AUCTORIS

Mr. Adam Aspinall was born in Sudbury, Ontario in 1985. He attended Harrow District High School from 1999 to 2003 where, during the course of his last year, he was voted Valedictorian of his graduating class.

In 2003, Adam attended the University of Windsor and in 2008 he graduated with distinction from the H.B.Sc. Computer Information Systems program. He co-authored a publication in *Artificial Life* in 2009 and was recently published by Springer in the *Lecture Notes in Computer Science* for his contribution to the International Conference on Active Media Technology 2010.

Adam is currently a candidate for the Master's degree in Computer Science at the University of Windsor.