



Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>
Eprints ID: 7999

To link to this article: DOI: 10.1162/artl.2009.Gras.012
URL : <http://dx.doi.org/10.1162/artl.2009.Gras.012>

To cite this version:

Gras, Robin and Devaurs, Didier and Wozniak, Adrianna and Aspinall, Adam
An individual-based evolving predator-prey ecosystem simulation using a fuzzy cognitive map as the behavior model. (2009) Artificial Life, vol. 15 (n° 4). pp. 423-463. ISSN 1064-5462

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes.diff.inp-toulouse.fr

An individual-based evolving predator-prey ecosystem simulation using Fuzzy Cognitive Map as behavior model

Gras R.*, Devaurs D.*, Wozniak A.*, Aspinall A.*

* University of Windsor, School of Computer Science

Correspondence to: R. Gras, University of Windsor, 401 Sunset Avenue, N9B 3P4 Windsor, Ontario, Canada. rgras@uwindsor.ca, +1 519 253 3000 ext. 2994

Abstract

This paper presents an individual-based predator-prey model with, for the first time, each agent behavior being modeled by a Fuzzy Cognitive Map (FCM), allowing the evolution of the agent behavior through the epochs of the simulation. The FCM enables the agent to evaluate its environment (e.g., distance to predator/prey, distance to potential breeding partner, distance to food, energy level), its internal state (e.g., fear, hunger, curiosity) with memory and choosing several possible actions such as evasion, eating or breeding. The FCM of each individual is unique and is the outcome of the evolution process throughout the simulation. The notion of species is also implemented in a way that species emerge from the evolving population of agents. To our knowledge, our system is the only one that allows modeling the links between behavior patterns and speciation. The simulation produces a lot of data including: number of individuals, level of energy by individual, choice of action, age of the individuals, average FCM associated to each species, number of species. This study investigates patterns of macroevolutionary processes such as the emergence of species in a simulated ecosystem and proposes a general framework for the study of specific ecological problems such as invasive species and species diversity patterns. We present promising results showing coherent behaviors of the whole simulation with the emergence of strong correlation patterns also observed in existing ecosystems.

Keywords

Evolving ecosystem, predator-prey model, individual-based model, Fuzzy Cognitive Map, speciation

1 Introduction

Individual-based modeling is a bottom-up approach to simulating ecosystems that allows for the consideration of the traits and behavior of individual organisms. Whereas classical approaches to modeling ecology often ignore individual behavior and instead consider an entire ecosystem as a whole, individual-based models aim to "treat individuals as unique and discrete entities" [10]. 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 classical approaches, namely that the assumptions made regarding individual behavior (such as the desire for fitness and shelter) provide for a more detailed simulation than using a state-variable model that may begin by calculating birth and death rates.

It has been suggested that because models are not well categorized, it is difficult to isolate any one model as being a specific type, such as individual-based [38]. Critics of this approach suggest that individual-based models are merely a tool for simulating very specific environments. However, advocates who favor the use of individual-based models are driven by paradigmatic motivation [10] 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 Individual-based Modeling and Ecology reserve several sections to discuss the generality of individual-based models [11]. They describe the difficulty of creating generic ecological models by comparing ecology to physics. "Individuals [of ecology] are not atoms but living organisms" and because "individual organisms have properties an atom does not have", such as the variation between them and their adaptive behavior, aiming for generality in ecological models is much more difficult. Despite this, there continues to be a rise in the use of individual-based models [16].

While the use of individual behavior has been included in many models during recent decades [15], the individual-based modeling approach is exponentially increasing as the cost to purchase and operate a machine capable of running time consuming simulations reduces. The contributions of individual-based models are discussed in [4] which examines, among others, how forest ecology [32], a fish-recruitment model [29], and models depicting spatial heterogeneity [18] have all benefited from this approach. Few attempts have been made to model a complete ecosystem. A pioneer in this domain is J. Holland with his platform Echo [12, 13] which includes an evolutionary mechanism. However, the organisms in Holland's simulation are very simple and do not involve any behavioral model. A predator-prey model has also been proposed by Ward et al. [39], with more complex agent models. Nevertheless, the agent model is dedicated to represent schooling behaviors and the evolution is an offline mechanism using a genetic algorithm. More recently, Ronkko [30] has proposed a high scale simulation based on a particle system approach. There is, however, no evolution mechanism in this artificial ecosystem.

As the agent behavioral model is crucial to creating complex interacting agent, we have chosen a sophisticated but efficient model called Fuzzy Cognitive Map (FCM) [17]

to model the agents' behavior. In our simulation, FCM is not only the base for describing and computing the agent behaviors, but also the platform for modeling the evolutionary mechanism and the speciation events. Additionally, we have implemented a speciation mechanism based on gene pool and, to our knowledge, for the first time in such simulation, linked behavioral patterns to speciation. To date, there is also no large scale individual-based ecosystem simulation that integrates a complex behavioral model for the agents, an evolutionary mechanism and a speciation mechanism. In particular, there is no use of FCM or equivalent model in such a large scale simulation and in the context of evolution. Our study includes important ecological and evolutionary concepts at a computationally acceptable cost. As we include in the same timescale of the simulation speciation events and individual behaviors, we have chosen to only represent tendency of behavior for our individuals. Therefore, a time step in the simulation represent a relatively long time period. The individuals perform multiple actions during this period but with a specific tendency corresponding to the action represented in our simulation (section 3.3). We show that such complex adaptive systems lead to a generic ecosystem with behaviors similar to those found in existing ecosystems. These are the key components needed in order to show that this kind of approach can be used to understand existing ecosystems and make some interesting and valid predictions.

The rest of the article is organized as follows: In section 2, we present and define the FCM model. In section 3, we describe the agents, the speciation concepts, the evolutionary mechanism and all the other components of our simulation. In section 4, we show the results we obtained for one run of the simulation and discuss about the pertinence of these results considering existing ecosystem behaviors. Finally in section 5 we conclude about this work and propose several possible extensions and dedicated applications to enhance our method.

2 Fuzzy Cognitive Maps

The Fuzzy Cognitive Maps rely on a concept derived from cognitive maps that was originally introduced by psychologists to model complex behaviors [36]. Recently, these FCMs have been extended in several steps: formalization as an oriented graph [1], association with fuzzy logic (Fuzzy Cognitive Map) [17], dynamic integration of external information [34] and learning [35]. FCMs aim to represent the causal relationship between concepts and to analyze inference patterns (the final state of the system after convergence). They are able to handle temporal information and fuzzy activation levels for each concept. They have been used in a wide variety of fields involving economic system modeling [33], machine learning [9], freeway modeling [37], autonomous agent modeling [34], etc. FCMs have also been used to represent complex biological systems such as ecosystems [26] and regulatory networks [41, 5]. FCMs have been used to model individual agent behaviors [34, 35] but only for few none evolving individuals. This last application led to very promising results that demonstrate the ability to represent complex internal concepts as emotions and desires, and to build agents that are able to perceive, make decision and act. Nevertheless, to our knowledge, a FCM has never been used in a large scale individual-based modeling of an ecosystem and has never been used in an evolutionary context.

We used mostly the definition of the FCM coming from [34]. FCMs are graphs which contain a set of nodes C , each node C_i being a concept, and a set of edges I , each edge I_{ij} representing the influence from a concept C_i to a concept C_j . A positive value of I_{ij} corresponds to an excitation of the concept C_j from the concept C_i whereas a negative value corresponds to an inhibition (a value of 0 meaning that there is no influence of C_i on C_j). An activation level a_i is also associated to each concept. The FCM allows the computation of the value of the concepts of an agent based on its perception and on the current activation level of its concepts. This computation is called the dynamic of the map and is a normalized matrix product (see section 2.1)

The FCM is used to model the agent behaviors (structure of the graph) and to compute the next action of the agent (dynamic of the map). A map contains three kinds of concepts: sensitive, internal and motor. The activation level of a sensitive concept is computed by a *fuzzification* of the information coming from the environment. The activation level of the motor concept is used to determine what the next action of the agent will be and a *defuzzification* of its value can be used to determine the amplitude of the action. Finally, the internal concepts' activation level corresponds to the level of intensity of the internal states of the agent and affects the computation of the dynamic of the map.

2.1 A formal definition of FCM

A FCM F is a quadruplet (C, L, A, R) where:

- $C = \{C_1, \dots, C_n\}$ is the set of n concepts
- L is a matrix $n \times n$ with $L_{ij} \in \mathbb{R}$. L_{ij} is the influence of concept C_i on concept C_j . If $L_{ij} = 0$, there is no edge between C_i and C_j .
- $A = \begin{cases} C \rightarrow [0,1]^{\mathbb{N}} \\ C_i \rightarrow \{a_i(0), a_i(1), \dots\} \end{cases}$ is a function that associates the series of all its successive activation levels to each concept C_i such as for $t \in \mathbb{N}$, $a_i(t) \in [0,1]$ is its activation level at time t .
- R is a recursive relation between $a_i(t+1)$ and $a_i(t)$ with $1 \leq i \leq n$ which describes the dynamic of the map F .

$$\forall i \ 1 \leq i \leq n, a_i(t_0) = 0; \forall i \ 1 \leq i \leq n, \forall t \geq t_0, a_i(t+1) = \sigma \circ g(a_i(t), \sum_{1 \leq j \leq n} L_{ji} a_j(t))$$

where $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a function such as: $\min(x,y)$ or $\max(x,y)$ or $\alpha x + \beta y$, and where $\sigma : \mathbb{R} \rightarrow [0,1]$ is a normalization function with two possible modes:

(a) continuous mode, where σ is the sigmoid function $\sigma_{(\delta, a_0, k)}$ centered in $(a_0, (1-\delta)/2)$ with a slope of $k \cdot (1+\delta)/2$ in a_0 and with limits in $\pm\infty$ respectively of 1 and 0:

$$\sigma_{(\delta, a_0, k)} : \begin{cases} \mathbb{R} \rightarrow [0,1] \\ a \rightarrow \frac{1}{1 + e^{-k(a-a_0)}} \end{cases}$$

(b) ternary mode:

$$\sigma : a \rightarrow \begin{cases} 0 & \text{if } a < s_1 \\ \frac{a - s_1}{s_2 - s_1} & \text{if } s_1 \leq a \leq s_2 \\ 1 & \text{if } a > s_2 \end{cases}$$

2.2 A simple example

A very simple map can be defined to model an agent perceiving and reacting to its distance to a foe. The closer the foe, the more frightened the agent. Depending on this distance and then on the fright level the agent will decide whether or not it will evade. The more frightened the agent, the faster the evasion. An FCM corresponding to this example is given in figure 1. In this example there are two sensitive concepts: *foeClose* and *foeFar*, one internal: *fear* and one motor: *evasion*. There are also three influence edges: closeness to a foe excites fear, distance to a foe inhibits fear and fear causes evasion. Activations of the concepts *foeClose* and *foeFar* are computed by fuzzyfication of the real value of the distance to the foe, and the defuzzyfication of the activation of *evasion* tells us about the speed of the evasion.

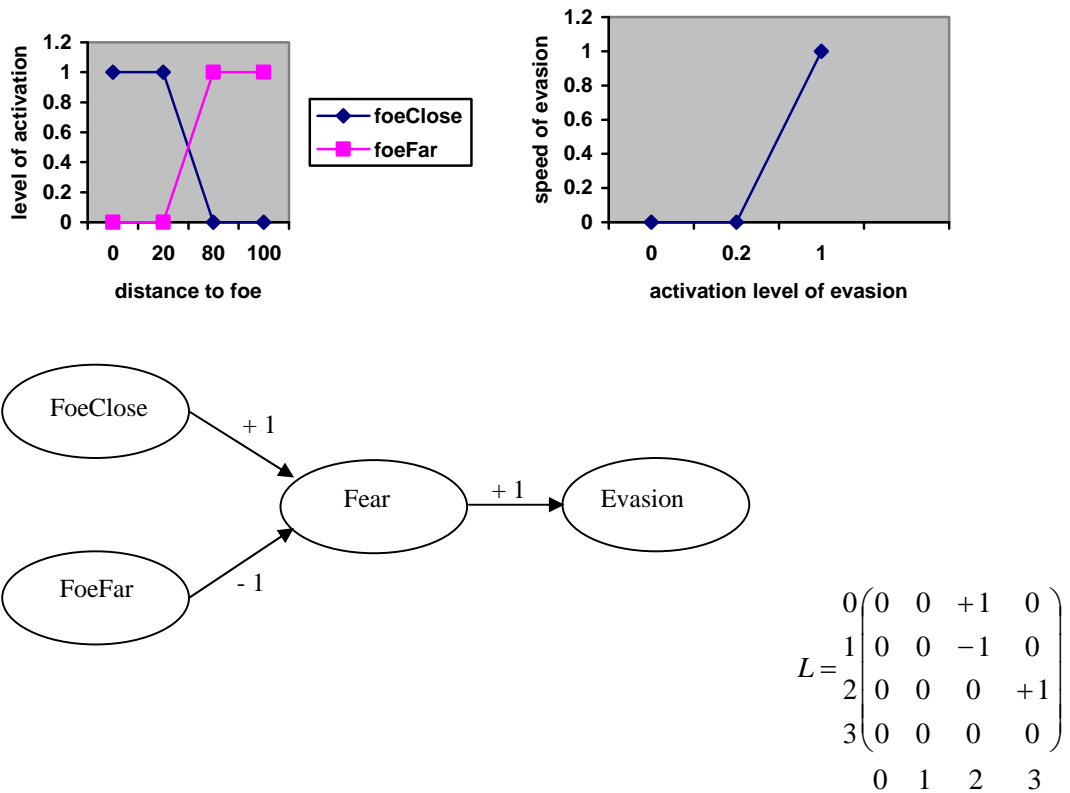


Figure 1. A simple fuzzy cognitive map for detection of foe and decision to evade with its corresponding matrix L and 0 for “Foe close”, 1 for “Foe far”, 2 for “Fright” and 3 for “Evasion” and the fuzzyfication and defuzzyfication functions.

With the FCM model it is possible to distinguish the perception from the sensation: the sensation is the real value coming from the environment and the perception is the sensation modified by the internal states. For example, it is possible to add three edges to the previous map (figure 2): one auto-excitatory edge from the concept *fear* to itself, an excitatory edge from *fear* to *foeClose* and one inhibitory edge from *fear* to *foeFar*. A given real distance to the foe seems higher or lower to the agent depending on the activation level of *fear*. Also the fact that the agent is frightened at time t influences the level of fright of the agent at time $t + 1$. This kind of mechanisms gives the possibility to model a degree of paranoia and a degree of stress for the agent. It also allows to memorize information from previous time steps: fear maintains fear. If the dynamic of the map is computed several times using the same sensation value (several applications of R updating each time the values of the concepts) it allows the sensitive information to go through each level of the map, even if there are loops, and influences the motor concepts before any action is undertaken. As a decision making model that can be understood as: take time to think before acting. It is therefore possible to build very complex dynamic systems involving feedback and memory using FCM, which is what is needed to model complex behaviors and abilities to learn from evolution.

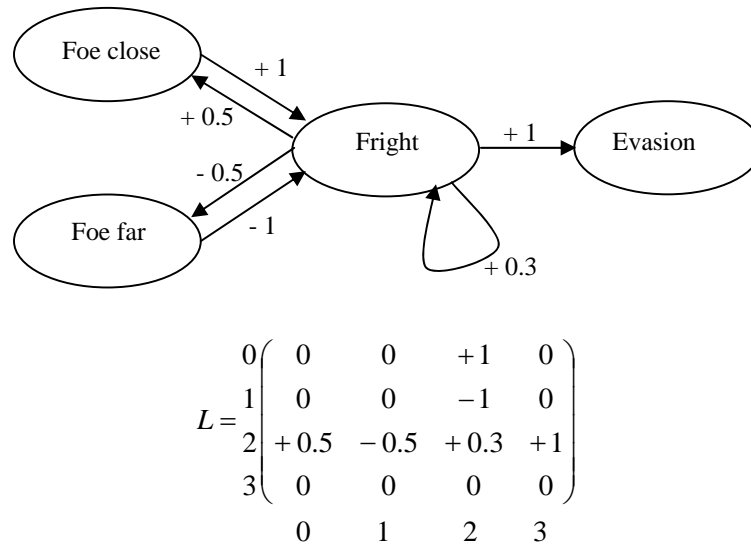


Figure 2. A simple fuzzy cognitive map for detection of foe and decision to evade with its corresponding matrix L and 0 for “Foe close”, 1 for “Foe far”, 2 for “Fright” and 3 for “Evasion”.

3 An evolving ecosystem

We have chosen an individual-based approach for our simulation of an evolving ecosystem. We aimed to develop a generic platform able to simulate complex ecosystems with “intelligent” agents interacting and evolving in a large and dynamic environment. An important property that we wanted to integrate was the fact that the agents have to develop efficient behaviors to be able to survive in this environment. We have therefore chosen a predator-prey model in which behaviors of preys and predators have to evolve

simultaneously to give them abilities to survive. Our ecosystem is composed of individuals belonging to two trophic levels: prey and predator. We also handle two resources: grass and meat which are respectively the food for the preys and the predators (see section 3.3). This concept could be easily applied to a more complex food chain by adding more resources and creating a higher hierarchy of predator/prey. Each agent possesses its own genome (the matrix L of its FCM, see section 3.1), can interbreed with other genetically similar individuals and produce offspring with a modified combination of the genomes of its parents (see section 3.5). We also represent species (see section 3.2). New species can emerge from the evolution of individuals and get extinct if all their members die.

3.1 Agents

Each agent has several properties that determine its physical capabilities and its behaviors. The Behaviors are determined by the interaction between the FCM and the environment. Each agent possesses its own FCM that represents its genome. This FCM contains sensitive concepts: (1) *foeClose* (prey only), (2) *foeFar* (prey only), (3) *preyClose* (predator only), (4) *preyFar* (predator only), (5) *foodClose*, (6) *foodFar*, (7) *mateClose*, (8) *mateFar*, (9) *energyLow*, (10) *energyHigh*, (11) *quantityOfLocalFoodHigh*, (12) *quantityOfLocalFoodLow*, (13) *quantityOfLocalMateHigh*, (14) *quantityOfLocalMateLow*; internal concepts: (15) *hunting* (predator only), (16) *fear*, (17) *hunger*, (18) *sexualNeeds*, (19) *curiosity*, (20) *sedentarity*, (21) *satisfaction*, (22) *annoyance*; and motors concepts: (23) *evasion* (prey only), (24) *searchForPreys* (predator only), (25) *searchForFood*, (26) *socialization*, (27) *exploration*, (28) *resting*, (29) *eating*, (30) *breeding*. It also contains links and weights representing the mutual influences of these concepts. Concepts (1) to (8) are computed by the fuzzyfication (using ternary mode (b) from section 2.1) of the distance of the closest corresponding feature (foe, prey, food and mate). Concepts (11) and (12) are computed by the fuzzyfication (using ternary mode (b) from section 2.1) of the number of food units currently available in the cell of the agent. Concepts (13) and (14) are computed by the fuzzyfication (using ternary mode (b) from section 2.1) of the number of possible mates currently present in the cell of the agent. The FCM of an agent is transmitted to its offspring after being combined with the one of the other parent and after the possible addition of some mutations. The behavior model of each agent is therefore unique¹. Links between concepts can appear or disappear during this process so the structure and complexity of the map can also change during the evolutionary process.

The values of the FCM used to initialize all first preys and predators are given in tables 1 and 2. It is important to notice that such a behavioral model allows the representation of very complex phenomena. For example, looking at table 1, it appears that the concept of *evasion* is excited by the concepts of *fear* and *annoyance* and inhibited by the concepts of *hunger*, *sexualNeeds*, *curiosity* and *satisfaction*. These concepts in turn are excited or inhibited by all the sensitive concepts. That means that the activation level of the motor concept of *evasion* depends on a complex and non-linear combination of all the sensitive concepts and of 6 internal concepts. This is true for all the motor concepts. Another important thing to notice is that the activation levels of all the concepts of an

¹ In fact, the uniqueness is not guaranteed but the probability that two identical FCM appears during the simulation is very close to zero.

agent are never reset during its life. As the previous time step activation level of a concept is involved in the computation of its next activation level, it means that all the previous states of an agent during its life participate in the computation of its current states. It means therefore that an agent has a memory of its own past that will influence its future states. As the action undertaken by an agent at a given time step depends on the current activation level of all its motor concepts, the global behavior of an agent dynamically depends on a complex combination of all the information it currently receives from its environment, all its current internal states and all the past states it went through during its life.

	16	17	18	19	20	21	22	23	25	26	27	28	29	30
1	4	0	0	.1	0	-1	1	0	0	0	0	0	0	0
2	-4	0	0	0	0	.5	-.5	0	0	0	0	0	0	0
5	0	.5	0	-.1	.1	.5	-.5	0	0	0	0	0	0	0
6	0	0	-.4	.2	-.2	-.7	.7	0	0	0	0	0	0	0
7	0	0	.5	-.1	.1	.5	-.5	0	0	0	0	0	0	0
8	0	0	-.4	.2	-.2	-.5	.5	0	0	0	0	0	0	0
9	.4	4	-1.5	0	0	-2.2	2.2	0	0	0	0	0	0	0
10	0	-1	1.5	.2	-.2	1.5	-1.5	0	0	0	0	0	0	0
11	0	-.2	0	-.3	.3	1.1	-1.1	0	0	0	0	0	2.6	0
12	0	.2	0	1	-1	-1.1	1.1	0	0	0	0	0	-4	0
13	0	0	0	-.4	.4	.5	-.5	0	0	0	0	0	0	1.5
14	0	0	.5	.3	-.3	-.8	.8	0	0	0	0	0	0	-4
16	.5	0	0	0	0	0	0	3.5	-.8	-1	.3	-1	-1	-1
17	0	.3	0	0	0	0	0	-.8	2.1	-.7	.7	-.5	4	-1.8
18	0	0	.2	0	0	0	0	-.2	0	1.5	.5	-.3	-.4	3
19	0	0	0	.1	0	0	0	-.1	.5	.3	1.5	-.2	-.3	-.2
20	0	0	0	0	.1	0	0	0	-.5	-.3	-1.2	.2	.3	.2
21	0	0	0	0	0	0	0	-.1	-.8	-.2	-2	1.5	.8	.7
22	0	0	0	0	0	0	0	.4	1	.2	2	-1.2	-.7	-.7
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	.2	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 1. Initial matrix L for preys

	15	17	18	19	20	21	22	24	25	26	27	28	29	30
3	.7	0	0	-.1	0	.5	-.5	0	0	0	0	0	0	0
4	-.5	.7	.1	.4	-.4	-.5	.5	0	0	0	0	0	0	0
5	-.5	.7	0	-.1	.1	.5	-.5	0	0	0	0	0	0	0
6	.8	-.2	.1	.2	-.2	-.6	.6	0	0	0	0	0	0	0
7	0	0	.7	0	0	.4	-.4	0	0	0	0	0	0	0
8	0	0	-.5	.3	-.3	-.4	.4	0	0	0	0	0	0	0
9	3.5	5	-1.2	0	.2	-1.5	1.5	0	0	0	0	0	0	0
10	-2	-3	1.4	.3	-.3	1	-1	0	0	0	0	0	0	0
11	-1.5	.3	-.2	-.3	.3	1	-1	0	0	0	0	0	4	0
12	1.7	0	.2	1	-1	-1	1	0	0	0	0	0	-5	0
13	-.3	0	0	-.4	.4	.8	-.8	0	0	0	0	0	0	2
14	.3	0	.5	.3	-.3	-.8	.8	0	0	0	0	0	0	-.5
15	.2	0	0	0	0	0	0	1.5	-.2	-.4	.3	-.4	0	-.4
17	0	.3	0	0	0	0	0	1.5	2.5	-1.2	.3	-.4	3.5	-.8
18	0	0	.2	0	0	0	0	-.8	-.8	1.5	.3	-.5	-.6	3
19	0	0	0	.1	0	0	0	.3	.3	.3	1.5	-.4	-.3	-.2
20	0	0	0	0	.1	0	0	-.3	-.3	-.3	-1.5	.4	.3	.2
21	0	0	0	0	0	0	0	-.8	-.8	-.2	-1.8	1	.8	.8
22	0	0	0	0	0	0	0	1	.8	.2	2	-1	-.6	-.8
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	.2	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2. Initial matrix L for predators

The physical capabilities are:

- Maximum and current level of energy. At each time step, each agent spends energy depending on its action (breeding, eating, running...) and on the complexity of its behavior model (number of nodes and edges in its FCM). The more complex its model is, the more energy the agent spends at each time step. The maximum level of energy (*maxEnergyPrey* and *maxEnergyPredator*) is associated with the type of agent (predator or prey).
- Maximum and current age. The maximum age (*maxAgePrey* and *maxAgePredator*) of an individual is determined randomly at birth from a distribution centered at a value associated with the type of agent. At each time step the age of each agent is incremented by one. When the current age of an agent is equal to its maximum age it dies.
- Minimum age for interbreeding. The minimum age at which an individual can begin to interbreed is associated with the type of agent (*ageInterbreedPrey* and *ageInterbreedPredator*).
- Maximum and current speed. The current speed of an agent is calculated when it undergoes a moving action. The speed value corresponds to the defuzzification of the activation level of the corresponding motor concept. The maximum speed is associated with the type of agent (*maxSpeedPrey* and *maxSpeedPredator*).

- Vision distance. This parameter determines how far (in number of cells) an agent can perceive things (food, foe...). The vision distance is associated with the type of agent (*distanceVisionPrey* and *distanceVisionPredator*). At most each individual can view and memorize the 5 closest individuals and resources of each type within its vision range.
- The energy transmitted to offspring. It determines the minimum percentage of energy that is transmitted to the (unique) offspring from its parents (see section 3.6). The maximum percentage is *birthEnergyPreyMax* for the preys and *birthEnergyPredatorMax* for the predators. The level of energy of the offspring is uniformly selected between these minimum and maximum values. Each parent loses half of this value. The amount of energy transmitted from the parents to their offspring is also submitted to evolution. The energy value for the offspring is the value (possibly mutated) of one of the two parents. The energy transmitted to the offspring is initially associated with the type of agent (*birthEnergyPrey* and *birthEnergyPredator*)².

3.2 Species

To the best of our knowledge, there were two models embodying mechanism of speciation: species were migrating and getting extinct but did not originate and the number of species was a predefined parameter. We have built a model implementing speciation mechanism that is related to the genotypic cluster definition proposed by Mallet [16]. The speciation mechanism we implemented accounts also for the gradualism and fuzziness of the speciation process. Traditionally, good species are populations that do not exchange genes with other populations so that there is no blurring of the species border: “*species level is reached when the process of speciation has become irreversible, even if some of the (component) isolating mechanisms have not yet reached perfection* [23]³. Yet, our model accounts for the fact that isolating barriers operating between populations and delimiting species boundaries undergo evolution, so that their appearance in itself is a part of speciation [21]. What is more, these boundaries are not permanent over time. For example, in the case of young related species we cannot exclude the possibility that even species considered as good will backcross, i.e. fuse back into one via hybridization. Indeed, recent studies have shown that hybridization is frequent (Cichlids in African lakes, common in plants). Our model accounts for the fact that speciation is not always a sharp and clear-cut process and that there are numerous groups showing substantial reproductive isolation but also exchange genes with sympatric relatives to some degree [2]. Thus, with these assumptions we can ask ourselves a few questions, such as A) how exactly intergradation takes place, implicit in the gradual nature of evolution, which presupposes the presence of intermediating forms, and B) what about the origin of speciation in sympatry, driven by behaviour (such as mating preferences).

² These two parameters are used to initialise the populations of preys and predators at the first time step. As these parameters are subjected to evolution, they are specific to each agent born during the simulation.

³ Similarly according to some of those who plead for sympatric speciation, the gene flow should approach zero in order to consider speciation as complete [6].

A) Indeed, one of the problems the simulation allows us to tackle with, is *the* problem of speciation, i.e. the origin of discrete groups of organisms [3, 7] or, in other words, the origin of organic diversity and the level at which the evolutionary process of differentiation is concerned. Yet, Darwinian evolution, synonymous to speciation, is considered a gradual process. Thus, a number of biologists have argued that gradual nature of the process of evolution implies gradual character of speciation [3, 22, 25]. We propose a model that accounts for speciation as a quasi-continuous process that yields intermediate stages.

B) Indeed, the innovation of our ecosystem simulation lies in the fact that it encompasses a complex behavioural model for the agents altogether with a speciation mechanism. We can then also examine the potential role of non-genetically driven individual variation, such as behaviour, or learning, in generating local selective pressure. This is crucial to determine whether reproductive isolation must be genetic or can have ontogenetic, and particularly behavioural basis, e.g. in sympatric speciation.

In our simulation a species is a set of individuals associated with the average of the genetic characteristics of its members. The average map is computed on the basis of the FCM matrixes of all individuals that are members of a species. It is considered that an individual belongs to a species if the difference between its matrix and the average matrix of the species is below a speciation threshold; the speciation threshold is the same for all species. Interbreeding can take place if the distance between individual matrixes is below the reproductive threshold. When a newborn appears, the distance between its matrix and the average matrixes of all existing species is computed. If the distance with respect to the closest average (i.e. the more similar species) is greater than the speciation threshold, then the individual forms a new species S . If in subsequent time-steps matrixes of some individuals turn out to be closer to S rather than to the average of their original species, the membership of these individuals will be switched to that new species.

More formally, we define a species as a set of individuals S and a centre $C(S)$ that represents the average genome of its members. We then define a metric D , such that $D(x,y)$ is the distance between the genomes of two individuals x and y :

$$D(x, y) = \sum_{1 \leq i, j \leq n} |L_{ij}^x - L_{ij}^y|$$

with L_{ij}^x (resp. L_{ij}^y) the edge between concepts i and j in the matrix L of x (resp. y). The center $C(S)$ is a matrix $L^c(S)$ such as:

$$\forall i, j \ 1 \leq i \leq n, 1 \leq j \leq n, L_{ij}^c(S) = \frac{\sum_{x \in S} L_{ij}^x}{|S|}$$

Then, it is also possible to compute the distance between an individual x and a species S :

$$D(x, S) = \sum_{1 \leq i, j \leq n} |L_{ij}^x - L_{ij}^c(S)|$$

Considering the current set Σ of existing species, an individual x is a member of a species S if:

$$D(x, S) = \min_{S' \in \Sigma} (D(x, S'))$$

Using the metric D and a speciation threshold T , a speciation event appears when a new offspring x is born such that: $\forall S \in \Sigma \ D(x, S) > 2 * T$. Considering two individuals possibly

from two different species), we define the probability $P(x,y)$ that these individuals can interbreed by:

$$P(x, y) = \begin{cases} f(D(x, y)) & \text{if } D(x, y) \leq T \\ 0 & \text{otherwise} \end{cases}$$

with $f: [0, T] \rightarrow [0, 1]$ a decreasing function of $D(x, y)$.

3.3 The world

Our simulation takes place in a toric virtual world composed of 1000 cells in both dimensions. Each cell can contain resources (grass and meat) and an unlimited number of individuals of both kinds. Because we want to focus on evolution of populations, we have chosen to make a coarse grain simulation. Even if every individual is simulated independently with a complex behavior, the world is not considered in high details. Therefore, a cell represents a large space that can contain an unlimited number of individuals. There is however a limit in the amount of resource available in each cell. This allows a competition for resource between individuals to occur.

We have also chosen an almost (see below) synchronous mode with discrete time. A time step corresponds to: the computation of all the agents' sensitive concepts, the computation of several dynamics of the map for each agent, the execution of one action by agents and the update of the parameters of the world. A time step also represents a relatively long time period. So, an action undertaken by an agent can in fact be viewed as a tendency. The agent performs a lot of small actions during a time step but the whole behavior is directed toward the realization of the given action. As a consequence, the total number of actions performed by each agent during its life is relatively small (a few dozens). This allows us to obtain a high level of population renewal which is an important criterion for studying an evolutionary process.

The maximum numbers of unit of each resource (*maxGrass* and *maxMeat*) by cell is a parameter of the simulation. At the initialization time there is no meat in the world and the number of grass units is randomly determined for each cell. For each cell, there is a probability, *probaGrass*, that the initial number of units is strictly greater than 0. In this case, the initial number is generated uniformly between 1 and *maxGrass*. Each unit provides a fixed amount of energy to the agent that eats it. The preys can only eat the grass and the predators have two modes of predation: hunting and scavenging. When a predator hunting action succeeds, new meat units are added in the corresponding cell. When a predator eating action succeeds (which can be viewed as a scavenging action), one unit of meat is removed in the corresponding cell. The amount of energy is *energyGrass* for one grass unit when eaten by a prey and is *energyMeat* for one meat unit eaten by a predator. The number of grass units grows at each time step (see section 3.4) and when a prey dies in a cell the number of meat units in this cell increases by 2. The number of grass units in a cell decreases by 1 when a prey eats and the number of meat decreases by 1 also when a predator eats. The number of meat units in a cell also decreases at each time step even if no meat has been eaten in this cell.

The initial position of the individuals is generated non-uniformly to form clusters of individuals. The idea is to model a realistic initial world state by having the individuals grouped in clusters. The parameter *sizeClusterPrey* (resp. *sizeClusterPredator*) sets how many preys (resp. predators) are members of each initial cluster. For the first member of the cluster, its position is uniformly generated in the whole world. Then, for all other

members of the cluster, their positions are uniformly generated among all cells that are in a *sizeCluster*-radius from the position of the first member. The initial number of preys (resp. predators) is determined by the simulation parameter *initNbPrey* (resp. *initNbPredator*) which is a multiple of *sizeCluster*. At the first time step all preys (resp. the predators) are members of the same species with a center corresponding to the initial FCM of preys (resp. predators) since all the individuals of the same type initially have the same FCM.

The preys and the predators are stored in two different lists in an age ascending order. We use this order to determine who acts before whom. For example, if in a given cell there is only one food unit and two agents that have chosen the action of eating, the youngest will act first and so it will be the only one that can eat (in this cell) at this time step. The action of the other one fails and it does nothing at this time step (except losing some energy). So even if every agent looks at its environment simultaneously, then makes a decision of action simultaneously, the simulation is not completely synchronic because there is an ordering of the actions based on the age of the agents. With this system the younger ones are advantaged compared to the older ones. It is a way to simulate the fact that the young can act faster than the old.

3.4 Update

At each time step we need to update the value of the state of all the parameters of our model. Here is the overview of the successive phases of the update process:

```

For every prey: Perception of the environment (1)
For every prey: Computation of all concepts (2)
For every prey: Application of their action and update of the energy level (3)
Updating the list of prey (4)
For every predator: Perception of the environment (1)
For every predator: Computation of all concepts (2)
For every predator: Application of their action and update of the energy level (3)
Updating the list of predators (4)
Updating the list of preys (5)
Updating the prey species (6)
Updating the predator species (6)
For every cell in the world {
    Updating the grass level (7)
    Updating the meat level (8)
}
Updating of the age of the agents (9)

```

The steps (1) to (9) are detailed here (for the predator, steps (1) to (4) and (6) are similar to those of preys):

- (1) For every prey, computation of the five closest foes; cells with food units and mates within the vision range of the prey; its current level of energy; the quantity of grass units in its cell; the number of possible mates in its cell are performed. The possible mates of a prey that is a member of a species *S* are the preys that are

- members of a species S' in which $D(S, S') < 2 * T$. With this mechanism we model the fact that an individual can evaluate its similarity with other individuals and then estimate if it can interbreed. This estimation is not precise because only the distance between their corresponding species is taken into account and with a threshold twice higher than the threshold for interbreeding. So individuals can try to interbreed even if mating will fails.
- (2) For every prey, computation of: the value of its sensitive concepts by fuzzyfication of the previous values; three dynamics of the map by applying the recursive formula R given in section 2.1 three consecutive times. Function g is $x + y$. Function σ is the continuous mode (a) for the internal and motor concepts and the ternary mode (b) for computing the initial value of the sensitive concepts.
 - (3) For every prey, in the age ascending order, application of the action corresponding to the motor concept that has the highest activation level⁴ and computation of the corresponding speed of the prey. Then computation of its new energy level by applying the formula:

$$energy = energy - nbconcepts - \frac{nbedges}{10} - speed^{1.4}$$

- with $nbconcepts$ the sum of the number of sensitive, internal and the motors concepts, $nbedges$ the number of edges in the prey FCM that have a value different from 0 and $speed$ the distance traveled by the prey during this time step.
- (4) For every prey, removing it from the list and adding two meat units in its cell if its energy is lower than or equal to 0 or if its age is greater than its maximum age. Adding all new prey's offspring to the beginning of the list of preys.
 - (5) Removing from the list every prey that has been killed by the predators.
 - (6) Removing every member of the current prey species (the species are reset before reallocating the preys to their closest species). Then, for all prey p , in the age ascending order, applying the algorithm:

```

dmin = 0
Smin = ∅
For all species S in Σ
    d = D(p, S)
    if d < dmin
        dmin = d
        Smin = S
If (dmin < 2 * T)
    S = S ∪ p
Else
    Create a new empty species S'
    S' = S' ∪ p
    Σ = Σ ∪ S'

```

⁴ If the highest motor concept is *breeding* then it is also required that the age of the prey (resp. predator) is greater than *ageInterbreedPrey* (resp. *ageInterbreedPredator*) otherwise it is the action corresponding to the second highest motor concept that is chosen.

- Then, removing every prey species that does not have anymore members, for every prey species S computing its new center $C(S)$ and for every prey species S computing its distance $D(S, S')$ with all other species.
- (7) For every cell of the world: if its number of grass units is greater than zero adding *growGrass* units of grass else if one of its 8 adjacent cells has a level of grass greater than zero adding *growGrass* units of grass with a probability of *probaGrowGrass*. With this mechanism if the agents eat all the grass in one cell the grass cannot grow anymore unless there is still grass in an adjacent cell. That prevents agents from staying in one place waiting for the grass to grow and models the problem of overexploitation of resource. That also models the mechanism of diffusion of resources through the world changing and renewing the interest of regions of the world. After this process if the number of grass units in the cell is greater than *maxGrass*, it is set to *maxGrass*.
 - (8) For every cell of the world: if its number of meat units is greater than zero subtracting *decreaseMeat* meat units. With this mechanism we model the fact that meat is perishable.
 - (9) Incrementation of the age of all agents.

The possible actions for the agents are:

- 1) Evasion (for preys only). The evasion direction is the direction opposite to the direction of the closest foe within the vision range of the prey compared to the current position of the prey. If no predator is within the vision range of the prey the direction is chosen randomly. Then the new position of the prey is computed using the speed of the prey (see below) and the direction. The current activation level of *fear* is divided by two.
- 2) Search for food. The direction toward the closest food (grass or meat) within the vision range is computed. If the speed of the agent is high enough to reach the food, the agent is placed on the cell containing this food otherwise the agent moves at its speed towards this food.
- 3) Socialization. The direction toward the closest possible mate within the vision range is computed. If the speed of the agent is high enough to reach the mate, the agent is placed on the cell containing this mate and the current activation level of *sexualNeeds* is divided by three otherwise the agent moves at its speed towards this mate. If no possible mate is within the vision range of the agent the direction is chosen randomly.
- 4) Exploration. The direction is computed randomly. The agent moves at its speed in this direction. The activation level of *curiosity* is divided by 1.5.
- 5) Resting. Nothing happens.
- 6) Eating. If the current number of grass (resp. meat) units is greater than 1, then this number is decreased by one and the prey's (resp. predator) energy level is increased by *energyGrass* (resp. *energyMeat*). Its activation level for *hunger* is divided by 4. Otherwise nothing happens.
- 7) Breeding. The following algorithm is applied to the agent A^5 :
 If $A.energyLevel > 0.125 * maxEnergyPrey$ then
 For all A' of the same type in the same cell

⁵ This algorithm is given for preys. The algorithm is almost identical for predators.

If $A'.energyLevel > 0.125 * maxEnergyPrey$ and
 $D(A,A') < T$ and
 A' has not acted at this time step yet and
 A' choice of action is also breeding

Then

interbreeding(A,A')

$A.sexualNeeds \leftarrow 0$

$A'.sexualNeeds \leftarrow 0$

if A' satisfies all the criteria the loop is cancelled

If none of the A' agents satisfies all the criteria the breeding action of A fails.

The interbreed() function is explained in the next section about evolution.

For every action requiring that the agent moves, its speed is computed by the formula:

$Speed = C_a * maxSpeedPrey$ for the preys

$Speed = C_a * maxSpeedPredator$ for the predators

with C_a the current activation level of the motor concept associated with this action.

3.5 Evolution

The evolution in this simulation comes from several mechanisms: interbreeding, mutation and speciation. The process of speciation is described in section 3.2 and it is linked to the notion of distance between FCMs. With this notion, depending on the FCMs of new offspring and of individuals that die, species can emerge or disappear at any time step. It allows us to model the evolution of populations of individuals that share important genetic properties. It will be a very important tool to study concepts such as the controversy between allopatric and sympatric speciation, diffusion of an invasive species in an existing ecosystem, species-abundance distribution.

Due to our species model, evolution of species is derived directly from the evolution of individuals. Evolution of individuals occurs when there is an interbreeding event. In this case, one unique offspring is conceived by two parent agents. The offspring inherits a combination of the genomic information of its parents with possible mutations. The genome of an agent is all the information that is transmitted from the parents to the child and submitted to possible mutations. In our current implementation, the elements that correspond to these criteria are the edge weight values of the matrix L of the agents' FCM and the parameters $birthEnergyPrey$ and $birthEnergyPredator$. These values are also used to compute the genetic distance D^6 . The process of generation of a new offspring corresponds to the function `interbreeding()` mentioned in the section 3.4. First the value of $birthEnergyPrey$ is transmitted with possible mutations (1) from one parent to the offspring. Second the edge's values are transmitted with possible mutations and the initial energy of the offspring is computed (2). To model the crossover mechanism, the edges are transmitted by block from one parent to the offspring (3). For each concept, all its incident edges are transmitted together from the same parent. Third, the maximum age

⁶ In the current implementation only the value of the matrix L is taken into account for the computing of the distance D .

of the offspring is computed (4). Finally, the energy level of the two parents is updated (5). Here is the algorithm⁷ of the interbreeding function:

Interbreeding(A_1, A_2)

Select uniformly A, one of the two parents to transmit *birthEnergyPrey* to its offspring O
 If randomNumber(0,1) < *probaMut* (1)

$r \leftarrow$ generate uniformly a number in $[-highMut, highMut]$
 $O.birthEnergyPrey \leftarrow (1 + r/100) * A.birthEnergyPrey$

Else $O.birthEnergyPrey = A.birthEnergyPrey$

If $O.birthEnergyPrey > birthEnergyPreyMax$

$O.birthEnergyPrey \leftarrow birthEnergyPreyMax$

$diff \leftarrow birthEnergyPreyMax - O.birthEnergyPrey$

$p \leftarrow$ generate uniformly a random number between 0 and diff

$p \leftarrow p + O.birthEnergyPrey$

$O.energy \leftarrow maxEnergyPrey * p/100$ (2)

For all i

Select uniformly A, one of the two parents, to transmit edge weights issue from concept C_i (3)

For all j

If $A.L_{ij} \neq 0$

If randomNumber(0,1) < *probaMut*

$r \leftarrow$ generate uniformly a number in $[-Mut, Mut]$

$O.L_{ij} \leftarrow A.L_{ij} + r$

If $|O.L_{ij}| < minEdge$

$O.L_{ij} \leftarrow 0$ // if the weight is too small it is set to 0

Else $O.L_{ij} \leftarrow A.L_{ij}$

Else // an edge that does not exist in the parents' FCM can emerge

If randomNumber(0,1) < *SmallProbaMut*

$r \leftarrow$ generate uniformly a number in $[-highMut, highMut]$

$O.L_{ij} \leftarrow r$

If $|O.L_{ij}| < minEdge$

$O.L_{ij} \leftarrow 0$

$r \leftarrow$ generate uniformly a number in $[-25, 25]$

$O.maxAge \leftarrow maxAgePrey + maxAgePrey * r/100$ (4)

$r \leftarrow .05 + O.birthEnergyPrey$ (5)

$A_1.energy \leftarrow A_1.energy - maxEnergyPrey * r/200$

$A_2.energy \leftarrow A_2.energy - maxEnergyPrey * r/200$

This mechanism allows the apparition of new edges, the disappearance of old ones and the variation of the weights associated with other edges. The apparition of new edges is very important in the sense that new influences between concepts could emerge during the evolutionary process. It leads to more complex and adaptive behaviors, and the inherent natural selection process, coming from the interaction of the individuals with

⁷ To make it simple we present only the interbreeding algorithm for the preys. The algorithm for the predators is almost identical.

their environment, will allow the preservation and the transmission of such behaviors if they have a selective advantage. As a counterpart, the possibility that edges disappear is also fundamental. When the complexity (number of existing edges) of the FCM grows, the agent needs more energy to survive, and then also needs a more efficient behavioral model to be able to obtain this energy. The possibility for the edges to disappear allows the evolutionary process to test the interest of some influence links, to remove them if they are not helpful enough, to react to the changes into the environment and to balance the interest of a complex behavioral model with its energy cost.

Most of the modifications consist in fact in small differences in the values of a few edges. By this mechanism, the concept of the neutral theory of evolution is integrated in the evolutionary model of the simulation. One mutation is almost neutral considering the behavioral model. Therefore, a unique breeding event, that generates a mutated offspring, has low probability to result in a new behavior model. It is the accumulation of neutral mutations during several generations that allows the apparition of new individual behaviors and then new species.

3.6 Complexity of the algorithm

The simulation algorithm is mostly linear with the number of agents N . More precisely an important part of the complexity comes from the computation of the dynamic of the map for the FCM of each agent. If we consider that there are N_1 preys and N_2 predators ($N=N_1+N_2$), that the matrix L size is n_1*m_1 for preys and n_2*m_2 for predators, the complexity of this part is $O(N_1n_1m_1+N_2n_2m_2)$, but as the size of L during the whole process is constant the complexity is in fact $O(N_1+N_2)$. Another computationally expensive part is the resolution of the breeding action. In this case, considering that an agent executing the breeding action has to compute its genomic distance D with p other agents⁸, and that $N_1' < N_1$ preys and $N_2' < N_2$ predators in the whole world execute the breeding action at a given time step, the complexity of this part is $O(N_1'pn_1m_1+N_2'pn_2m_2)$ or also $O(N_1'p+N_2'p)$. Another time consuming part, corresponding to updates 9) and 11) of section 3.4, is the computation of new species. For all preys and all predators, the distance D between their matrix L and the centre of each species has to be computed, and then the new center of all species has to be computed as well. If we consider the previous number of prey species is S_1 , the new number of prey species is S_1' , the previous number of predator species is S_2 , the new number of predator species is S_2' , the maximum number of agent members for all prey species is M_1 and the maximum of agent members for all predator species is M_2 then the complexity of this part is $O(N_1S_1n_1m_1+N_2S_2n_2m_2+S_1M_1n_1m_1+S_2M_2n_2m_2)$ or $O(N_1S_1+N_2S_2+S_1M_1+S_2M_2)$. The only non linear part corresponds to the computation of the distance D between all species. The complexity of this part is $O(S_1^2n_1m_1+S_2^2n_2m_2)$ or $O(S_1^2+S_2^2)$.

4 Running the simulation

4.1 The parameters

⁸ Practically, the number p is small (2 or 3) because during the simulation agents spread in the whole world. So the number of agents sharing a cell is small and only a fraction of them choose the breeding action. Moreover, the current level of energy of the agents could be not high enough to allow interbreeding. In this situation the distance D is not computed.

Even if the complexity of this algorithm is not high, in practice this simulation is intensively computationally expensive. As, in general, S is at least three orders of magnitude smaller than N , the dominant part is the computation of the new species. In our first experiments, the simulation manages up to 400,000 individuals and up to 500 species at a given time step, leading to a computational time of over forty minutes for one time step. Therefore our longest simulation until now has been running for two and a half months, corresponding to 7,112 time steps⁹. As it is the most complete run that we have currently obtained, we consider this run for the discussion in this section. The parameters used for this run are given in table 3. The initial FCM for preys uses table 1 and the initial FCM for predators uses table 2.

The first important thing to notice is that, even if this simulation is a very complex and large adaptive system, the whole behavior of the ecosystem is relatively stable and present interesting correlation patterns. Moreover, having tested numerous different sets of initial parameters, we have noticed that the overall behavior of the simulation is stable, which means that the same phenomenon of epochs of correlated inflation and deflation of the number of individuals, of species, of resources, appears systematically (see section 4.2). As it is a complex dynamic system, even if there are such regularities, the simulation is far from being easily predictable¹⁰. The amplitude and time of inflation and deflation vary considerably, but their mutual correlation is conserved.

<i>maxGrass</i>	10	<i>maxAgePrey</i>	39
<i>maxMeat</i>	8	<i>maxAgePredator</i>	34
<i>probaGrass</i>	0.07	<i>ageInterbreedPrey</i>	6
<i>energyGrass</i>	250	<i>ageInterbreedPredator</i>	8
<i>energyMeat</i>	500	<i>maxSpeedPrey</i>	7
<i>sizeClusterPrey</i>	10	<i>maxSpeedPredator</i>	12
<i>sizeClusterPredator</i>	3	<i>distanceVisionPrey</i>	40
<i>sizeCluster</i>	5	<i>distanceVisionPredator</i>	50
<i>initNbPrey</i>	12000	<i>birthEnergyPreyMax</i>	60
<i>initNbPredator</i>	1200	<i>birthEnergyPredatorMax</i>	75
<i>maxEnergyPrey</i>	650	<i>birthEnergyPrey</i>	30
<i>maxEnergyPredator</i>	1000	<i>birthEnergyPredator</i>	50
<i>probaMut</i>	0.005	<i>minEdge</i>	0.075
<i>Mut</i>	0.15	<i>highMut</i>	0.2
<i>SmallProbaMut</i>	0.001	<i>growGrass</i>	0.8
<i>probaGrowGrass</i>	0.0085	<i>decreaseMeat</i>	1
<i>TPrey</i>	0.75	<i>TPredator</i>	0.7

Table 3. Initial parameters of the simulation

4.2 Overall analysis of the simulation

⁹ The current implementation of the simulation is written in C# and has been running on an AMD Athlon 64 X2 Dual 4200+ processor with 4GB of memory.

¹⁰ There is no way, excluding the simulation itself, to predict the state of the system at time step t knowing the state of the system at time step $t-1$.

What is very useful for a biological interpretation with such detailed simulation is that all the parameters of all components remain accessible at any time in the evolving process. We have, for example, access to general parameters describing our population such as the current number of individuals and food units, the number of agents doing each type of action, the average energy level of individuals, the average current age of individuals, the average age of death of individuals, the average and maximum number of interbreeding by individual, etc. We also have access to the average value of the activation level of all concepts of the individuals, to the current number of species and for each species we have access to its average matrix. We could even have access to the speciation events and then construct the complete exact phylogeny of the evolving predator and prey species. To illustrate the behavior of the simulation and to see if it has properties that are known to exist in ecosystems, we have extracted several of these parameters and we have also computed the cross correlations between them. As these correlations may not be in phase (for example the number of predators at a given time step will have an influence on the number of preys several time steps later) and that this difference of phase is unknown and can differ for every couple of parameters, we have computed the maximum cross correlation value by shifted one time series against the other using the Pearson formula:

$$r(d) = \frac{\sum_i [(x(i) - m_x) \cdot (y(i-d) - m_y)]}{\sqrt{\sum_i (x(i) - m_x)^2} \sqrt{\sum_i (y(i-d) - m_y)^2}}$$

with $x(i)$ the value of the time series x at time step i , d the shift value, $y(i-d)$ the value of the time series y at time step $i-d$ and m_x (resp. m_y) the average value of the time series x (resp. y). We present the results in table 4 and 5. Several of the cross correlation coefficients are very high such as between the number of preys and the number of eat actions for prey (0.98), the number of preys and the number of breed actions for preys (0.99) or the number of socialize actions and breed actions for predators (0.94). For these cases, it seems that there is a direct correlation between the number of individuals and the number of individuals choosing an action, which means that an almost constant proportion of individuals choose this action during the whole simulation. For others, even if the correlation coefficient is quite high, the phenomenon is more complex. We selected several of them and presented and discussed their correlated evolutions¹¹. In figure 4, 5, 8, 9, 16 and 18 it should be notice that there exists cycles in the correlations between several parameters. This phenomenon clearly illustrates the fact that most of the parameters follow an oscillatory pattern. Long term correlations (high values of d) thus represent correlations between two different cycles of the oscillatory patterns of two parameters.

¹¹ We use TeeChart library from Steema Software for the visualisation of the graphs.

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	1	-0.57	0.45	-0.34	-0.67	0.86	-0.51	0.68	0.86	0.87	0.83	-0.41	0.86
1	-0.36	1	-0.81	0.87	0.68	0.5	-0.52	-0.52	-0.32	0.44	0.42	0.78	-0.51
2	-0.39	0.66	1	0.84	-0.48	-0.38	0.43	-0.15	-0.3	-0.36	-0.41	0.71	-0.5
3	-0.43	0.84	0.79	1	0.76	-0.58	0.55	-0.47	-0.56	-0.56	-0.58	0.81	-0.47
4	0.62	-0.39	0.46	-0.48	1	0.81	-0.61	0.52	0.67	0.81	0.76	-0.53	0.49
5	0.82	-0.52	0.29	-0.53	0.75	1	-0.58	0.68	0.86	0.94	0.92	-0.6	0.55
6	-0.5	0.3	-0.46	0.48	-0.58	-0.63	1	-0.45	-0.48	-0.64	-0.57	0.34	-0.53
7	0.66	-0.48	0.16	-0.43	-0.75	0.69	0.39	1	0.76	0.68	0.7	-0.49	0.4
8	0.82	-0.52	0.26	0.43	-0.76	0.79	0.47	0.78	1	0.71	0.69	-0.36	0.42
9	0.87	-0.51	0.25	-0.53	0.76	0.99	-0.66	0.68	0.86	1	0.98	-0.58	0.69
10	0.81	-0.49	-0.29	-0.54	0.73	0.95	-0.57	0.72	0.87	0.95	1	-0.61	0.61
11	-0.28	0.83	0.71	0.84	0.7	-0.68	-0.46	-0.52	-0.54	-0.64	-0.64	1	-0.45
12	0.86	-0.61	0.56	-0.45	0.4	0.65	-0.55	0.44	0.65	0.69	0.64	-0.49	1
13	0.98	-0.61	0.51	-0.36	-0.68	0.79	0.46	0.67	0.89	0.8	0.77	-0.44	0.8
14	0.99	-0.56	0.44	-0.34	-0.65	0.87	-0.55	0.66	0.84	0.89	0.86	-0.4	0.9
15	-0.31	0.99	-0.77	0.84	0.71	0.52	-0.54	-0.57	-0.37	0.47	0.45	0.78	-0.47
16	-0.42	0.71	-0.64	0.77	0.82	-0.55	-0.52	-0.65	-0.56	-0.53	-0.53	0.61	-0.39
17	-0.55	0.94	-0.88	0.87	0.59	0.32	-0.41	-0.32	-0.34	0.25	0.25	-0.78	-0.67
18	0.45	-0.41	0.41	-0.31	0.26	0.3	0.21	0.37	0.38	0.35	0.35	-0.34	0.6
19	-0.67	-0.54	-0.75	0.64	0.41	-0.23	-0.55	-0.41	-0.56	-0.24	-0.19	-0.65	-0.66
20	-0.34	1	-0.8	-0.86	0.66	0.52	-0.53	-0.51	-0.3	0.46	0.44	0.78	-0.5
21	-0.38	0.99	-0.82	-0.88	0.59	0.5	-0.51	-0.45	-0.23	0.44	0.42	-0.76	-0.55
22	-0.75	0.6	-0.4	0.7	0.87	-0.93	0.74	-0.69	-0.82	-0.9	-0.88	0.7	-0.47
23	0.73	-0.58	0.51	0.51	0.78	0.87	-0.47	0.6	0.82	0.86	0.83	-0.65	0.45
24	-0.72	0.57	0.25	0.53	0.81	-0.81	-0.47	-0.82	-0.8	-0.8	-0.8	0.6	-0.43
25	-0.39	-0.51	-0.33	0.31	-0.43	-0.39	0.33	0.28	0.31	-0.41	-0.37	-0.34	-0.5

	13	14	15	16	17	18	19	20	21	22	23	24	25
0	0.98	0.99	-0.6	-0.46	-0.43	0.38	0.6	-0.58	-0.56	-0.69	0.7	-0.7	0.54
1	-0.43	-0.36	0.99	-0.76	0.94	-0.53	-0.62	0.99	0.99	-0.65	0.64	0.51	-0.71
2	-0.42	-0.4	0.62	0.43	0.64	-0.57	0.4	0.66	0.66	0.44	-0.34	0.25	0.57
3	-0.47	-0.41	0.81	0.77	0.81	-0.56	0.6	0.82	0.8	0.7	-0.51	0.53	-0.54
4	0.6	0.61	-0.38	-0.78	-0.5	0.36	-0.47	-0.37	-0.36	-0.82	0.88	-0.57	0.14
5	0.78	0.8	-0.55	-0.62	-0.5	-0.24	0.59	-0.51	-0.5	-0.85	0.86	-0.76	0.39
6	-0.39	-0.54	0.3	0.47	0.35	-0.27	0.33	0.28	0.28	0.65	-0.5	0.42	-0.32
7	0.65	0.66	-0.52	0.62	-0.3	0.38	0.54	-0.48	-0.42	-0.64	-0.64	-0.78	0.57
8	0.87	0.74	-0.55	0.69	0.47	-0.26	0.84	-0.53	-0.5	-0.68	0.75	-0.77	0.51
9	0.8	0.89	-0.54	-0.61	-0.48	-0.24	0.52	-0.5	-0.49	-0.88	0.85	-0.76	0.39
10	0.75	0.82	-0.52	-0.57	-0.46	0.27	0.48	-0.49	-0.47	-0.85	0.84	-0.75	0.39
11	-0.31	-0.29	0.83	0.68	0.8	0.5	-0.47	0.82	0.8	0.73	-0.66	0.6	-0.65
12	0.8	0.9	-0.61	-0.36	-0.51	0.57	-0.21	-0.61	-0.6	-0.43	0.49	-0.49	0.5
13	1	0.95	-0.63	0.46	-0.44	0.39	0.66	-0.62	-0.6	-0.64	0.69	-0.69	0.59
14	0.95	1	-0.59	-0.45	-0.43	0.41	0.51	-0.57	-0.55	-0.7	0.69	-0.69	0.53
15	-0.38	-0.3	1	-0.8	0.91	0.53	-0.66	0.99	0.98	-0.67	0.68	0.54	-0.74
16	-0.46	-0.4	0.7	1	0.72	-0.44	0.69	0.68	0.65	0.69	-0.57	0.65	-0.6
17	-0.61	-0.55	0.91	0.72	1	-0.66	0.6	0.92	0.94	-0.5	0.45	0.39	-0.49
18	0.45	0.48	-0.42	-0.21	0.35	1	-0.12	-0.41	-0.4	-0.23	0.29	-0.28	-0.79
19	-0.71	-0.66	-0.5	0.72	-0.64	-0.76	1	-0.53	-0.56	-0.41	-0.31	0.21	-0.49
20	-0.41	-0.34	0.99	-0.77	0.93	0.54	-0.61	1	0.99	-0.67	0.65	0.5	-0.72
21	-0.45	-0.38	0.98	-0.74	0.94	0.57	-0.6	0.99	1	-0.65	0.59	0.44	-0.69
22	-0.71	-0.74	0.62	0.76	0.58	-0.21	-0.67	0.57	0.55	1	-0.87	0.79	-0.39
23	0.71	0.7	-0.59	-0.78	-0.61	0.28	0.7	-0.56	-0.54	-0.87	1	-0.76	0.32
24	-0.69	-0.71	0.6	-0.65	0.49	0.25	-0.64	0.56	0.54	0.79	-0.75	1	-0.49
25	-0.37	-0.39	-0.52	0.41	-0.43	-0.78	0.3	-0.52	-0.53	0.4	-0.4	0.35	1

Table 4. Value of the maximum cross correlation for 26 parameters with a shift $-2500 < d < 2500$. A cell of coordinate (x,y) corresponds to the highest cross correlation between x and y with $0 \leq d \leq 2500$ if $x < y$ and with $-2500 \leq d \leq 0$ if $x > y$. The parameters are: (0) number of preys, (1) number of predators, (2) number of prey species, (3) number of predator species, (4) grass level, (5) meat level, (6) prey average energy, (7) predator average energy, (8) escape, (9) search food for prey, (10) socialize for prey, (11) explore for prey, (12) wait for prey, (13) eat for prey, (14) breed for prey, (15) hunt, (16) search food for predator, (17) socialize for predator, (18) explore for predator, (19) wait for predator, (20) eat for predator, (21) breed for predator, (22) foe close, (23) satisfaction for prey, (24) prey close, (25) satisfaction for predator.

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	1424	280	1518	2475	25	2	432	246	0	1	1330	0
1	1638	0	1721	172	2285	1524	1417	41	74	1438	1818	0	1651
2	2076	0	0	147	258	301	236	2499	839	271	267	0	2074
3	870	0	0	0	2499	162	22	552	757	130	138	0	1630
4	207	1049	2499	0	0	214	12	2499	435	167	162	0	2499
5	0	772	2499	0	0	0	2	241	232	0	1	71	0
6	1	360	1840	42	0	35	0	2468	540	2	1	94	0
7	73	0	395	0	1707	110	2469	0	128	73	82	0	63
8	0	1031	225	2499	2229	0	2499	199	0	0	1	0	0
9	7	827	2499	0	0	32	4	10	504	0	1	100	0
10	17	1060	0	0	22	40	17	3	502	7	0	0	1
11	2097	281	0	153	2499	160	1569	32	339	117	100	0	2101
12	6	1776	307	1820	0	36	3	678	529	5	1	1502	0
13	1	1533	265	1529	2427	22	2499	266	33	0	1	1331	0
14	7	1551	300	1530	2492	34	3	432	262	3	1	1339	0
15	1643	8	1747	182	2192	1548	1418	56	64	1456	1826	0	1814
16	950	0	1640	0	2090	42	1742	927	73	130	1	0	1628
17	1650	6	1734	236	2487	1522	1428	33	1641	1258	1603	1650	1649
18	352	1913	408	2031	268	125	0	0	758	98	347	2043	95
19	1314	1623	1581	214	1523	188	1750	1372	1394	477	144	1378	1321
20	1640	7	1724	1945	2231	1533	1417	55	76	1437	1825	0	1655
21	1643	2	1718	1938	2322	1527	1419	29	75	1433	1818	1649	1652
22	47	454	1950	0	2499	68	19	498	305	34	27	0	31
23	88	866	2499	2499	19	148	57	1	147	114	103	0	2499
24	65	689	0	0	2246	103	2428	24	127	78	74	0	55
25	957	143	433	2139	546	1403	1387	146	0	899	1758	72	88

	13	14	15	16	17	18	19	20	21	22	23	24	25
0	0	0	1409	1026	1242	63	2499	1511	1513	0	110	368	1494
1	1626	1637	0	2367	0	1426	2292	0	0	1642	1685	2	52
2	2061	2073	0	477	0	1402	195	0	0	275	588	0	1627
3	1458	1471	0	65	0	1311	0	0	0	0	3	0	2499
4	217	186	1033	1158	964	2157	1053	1051	1041	202	291	540	2499
5	0	0	740	756	748	657	2499	782	779	0	96	339	766
6	1	0	573	756	400	1658	89	363	352	0	486	595	2499
7	67	72	0	2498	3	2	2499	1	701	1	2496	2	867
8	0	0	1010	2499	2499	1008	2499	1066	1127	0	40	79	1649
9	11	0	799	790	786	684	2499	1040	1038	0	140	371	810
10	23	6	826	835	830	5	2499	1056	1058	3	190	367	815
11	2087	2095	275	326	305	283	2499	263	280	235	4	27	129
12	6	0	1766	1542	1763	75	1572	1764	1769	0	494	601	1501
13	0	0	1542	2499	1401	49	2499	1528	1532	0	103	369	1501
14	7	0	1423	1056	1250	65	2499	1540	1545	0	131	368	1501
15	1631	1641	0	2366	2	37	2351	1	3	1646	1663	1	58
16	1451	1037	1	0	3	1338	3	1	0	0	4	38	2497
17	1642	1647	0	7	0	1433	131	1	4	1650	1688	0	60
18	345	93	1905	1930	139	0	2239	1906	1888	1385	748	1013	1
19	1333	1314	1603	226	1643	1342	0	1605	1606	1744	5	602	2499
20	1626	1641	0	2366	1	36	2286	0	2	1637	1662	0	55
21	1633	1641	1	2367	5	34	2274	2	0	1521	1659	5	48
22	51	36	521	570	453	1938	2499	522	578	0	4	0	83
23	79	88	711	887	701	2268	2499	948	866	80	0	28	340
24	62	62	652	2499	675	631	2498	679	686	0	0	0	334
25	953	899	135	2120	139	0	2465	131	133	1383	1040	1661	0

Table 5. Value of d for which the cross correlation is maximum for 26 parameters with a shift $-2500 < d < 2500$. The parameters are the same as in table 4.

In figure 3, the correlated evolution of the number of preys, the number of predators and the number of grass units are presented for the whole simulation (7,112 time steps). As we should expect with a predator-prey system, it is clear that there is a dependency between the number of preys and the number of predators. The evolution of the number of predators follows the one of preys and vice-versa. When the number of preys grows, the number of predators also grows few time steps later. But when the number of predators grows too much, the number of preys decreases few time steps later, leading to a new later decrease in the number of predators. The maximum cross correlation value between the number of preys and predators is -0.57 for $d = 1424$. But since there is a third entity, the number of grass units involved in this interacting system, its level influences the other two. The maximum cross correlation between the number of preys and the level of grass is -0.67 for $d = 2475$. The figures 4 and 5 show the variation of these two cross correlations for all values of d . It appears that the correlation between the number of preys and the number of predators is much more complex than the one between the number of preys and the level of grass. By studying figure 3 we can see that after the time period from time step 1 to 250, in which the number of preys is very high, the number of grass units decreases very fast until time step 1,725. During this period there is also a high decrease in the number of preys. Then it takes a long time to recover from this situation and the number of preys and grass units do not reach such a high level anymore. It is the union of the phenomena of high number of predators and low number of preys that allows a fast growth of grass between time steps 3,550 and 4,750. Then, as the number of predators decreases and the number of grass units is high, the population of preys enters a period of very fast growth, which in turn leads to an increase in the number of predators and a decrease in the number of grass units.

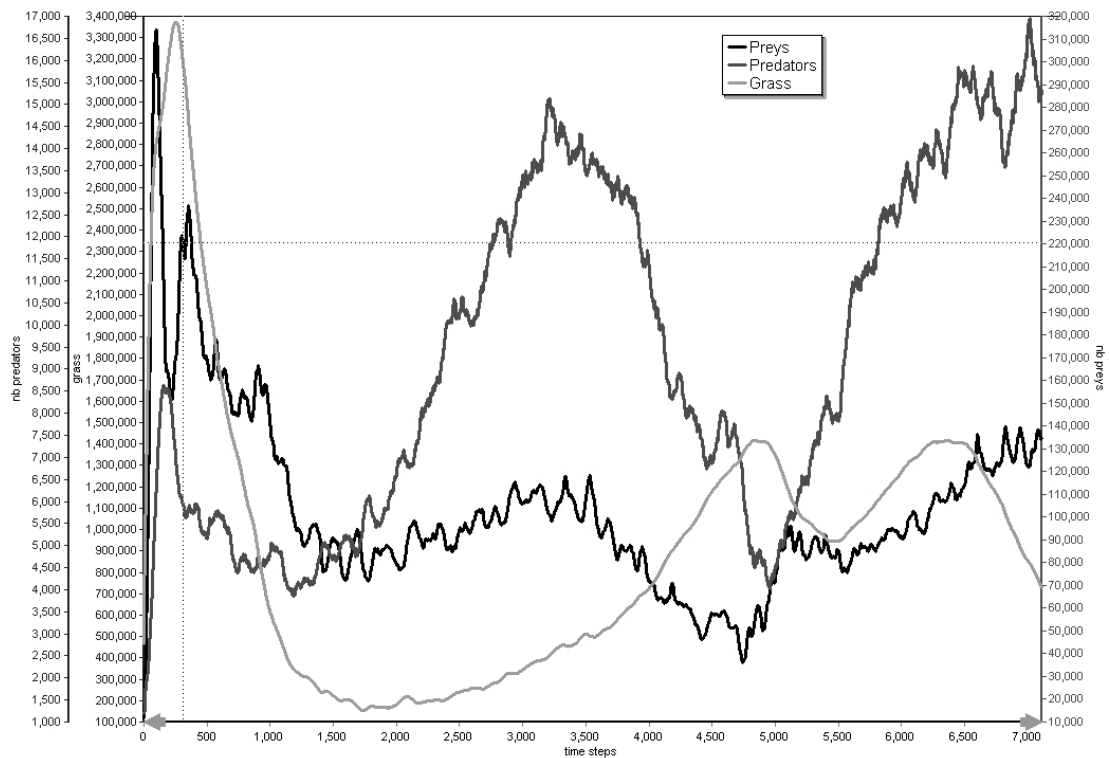


Figure 3. Evolution of the number of individual preys, the number of individual predators and the number of grass units

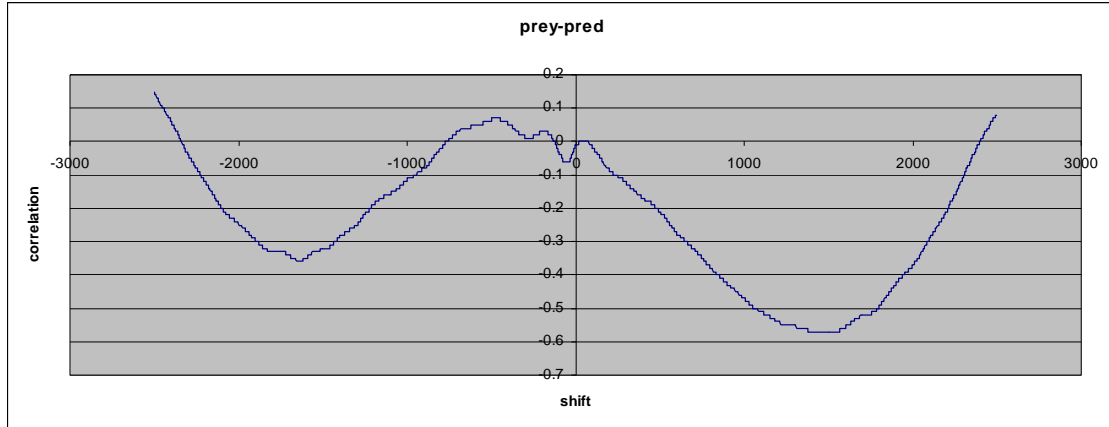


Figure 4. Cross correlation between the number of preys and the number of predators for $-2500 \leq d \leq 2500$.

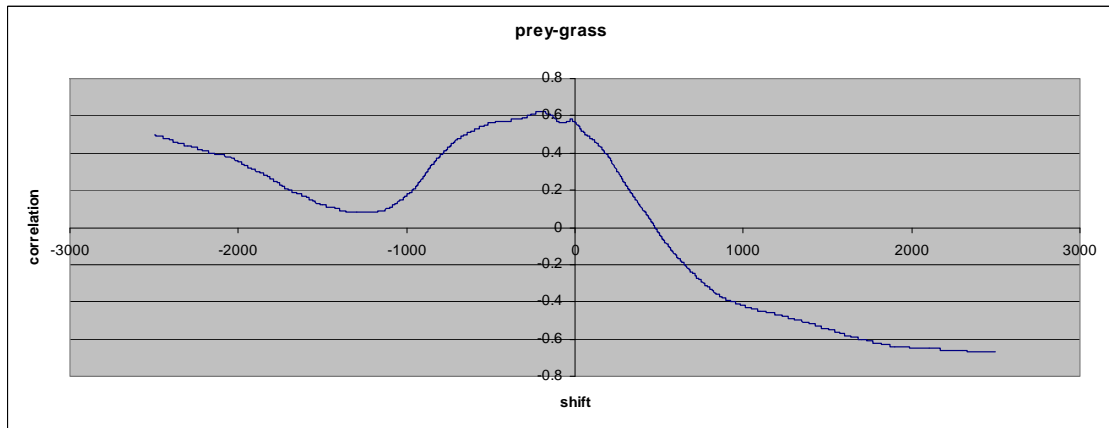


Figure 5. Cross correlation between the number of preys and the level of grass for $-2500 \leq d \leq 2500$.

Figures 6 and 7 illustrate the correlation between the number of individuals and the number of species for each type of individual. In figure 7, it appears that¹² the number of predator species is closely correlated to the number of predator individuals. The maximum cross correlation is 0.87 with $d = 172$. This correlation is also strong for the number of prey species and the number of prey individuals but with a higher difference in the amplitude of the fluctuations. The maximum cross correlation is 0.45 with $d = 280$. Figure 8 and 9 show that the correlation is much stronger between the number of predators and the number of predator species than between the number of preys and the number of prey species. For each of them the number of individuals increases before the number of species increases. This kind of correlation is what should be expected in an

¹² There is in fact a latency delay, between time step 1 and time step 128, needed for some offspring to have evolved enough in comparison to the initial matrix and then for the first species to emerge.

ecosystem. Several publications on existing organisms' populations show correlation patterns between the number of individuals and the number of species [31]. The difference for these publications is that the data come from different spatial locations since it is very difficult to collect data on the number of individuals and the number of species during a long period of time.

The species-area scaling relation is a classical ecological pattern. Its underlying intuitive idea is: if individuals are collected in different zones, the bigger the sampled area, the more species we find. This relation is used for example in conservation biology, in order to estimate the effects of the size of a reserve on species diversity. More significantly, species-area relations are the fundamentals of the theory of island biogeography [19]. Actually, islands in an archipelago provide ecologists with natural sampling habitats of varying sizes, but with similar environments. When species' richness is calculated for habitats of increasing size (such as islands), the following scaling relation holds: $S = cA^z$, where S is the total number of species found, A is the size of the sampled area, and c and z are regression constants¹³. This relation is empirically well supported [8, 19, 27], and the value of z is often around 0.25 for small scale ecological communities. However, no satisfying explanation of the underlying mechanisms of this phenomenon has been proposed yet [8]. Note also that exceptions to the species-area relation have been presented, e.g. in [28], and that the value of z can be greater than 0.25 when sampling large scale and complex areas characterized by greater habitat heterogeneity [19]. In figures 10 and 11¹⁴, we present graphs between the log of the number of individuals and the log of the number of species but for our evolving populations (different time steps). The linear dependency between the log of the number of predators and the log of the number of predator species is particularly clear with a slope of 0.99, which is higher than the 0.25 expected but controversial. Even if it is less convincing for preys, the phenomenon is still visible on figure 11.

¹³ On the logarithmic scale, the following equivalent linear relation holds: $\log(S) = z\log(A) + \log(c)$.

¹⁴ As the speciation process takes time to get stabilized, the data presented in these figures correspond only to the time steps between 1200 and 7112 for the preys and to the time steps between 1,600 and 7,112 for the predators.



Figure 6. Evolution of the number of individual preys and the number of prey species



Figure 7. Evolution of the number of individual predators and the number of predator species

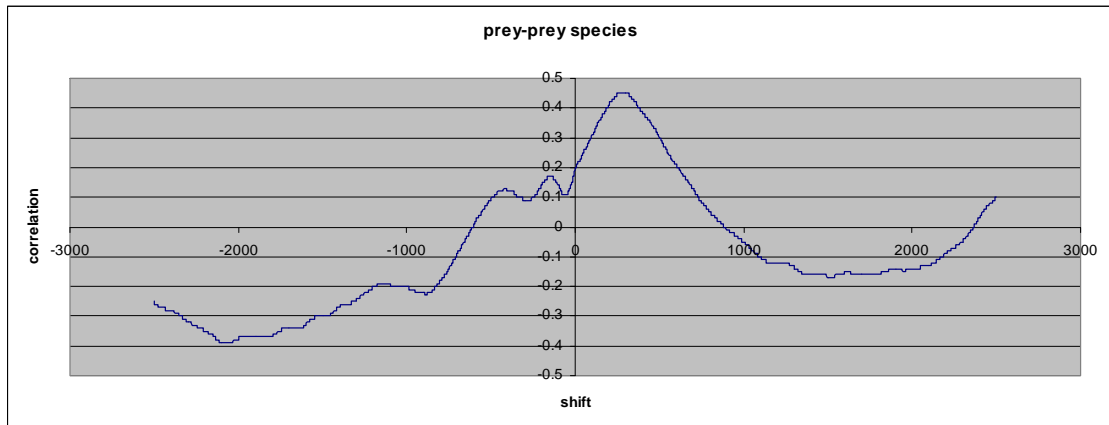


Figure 8. Cross correlation between the number of preys and the number of prey species for $-2500 \leq d \leq 2500$.

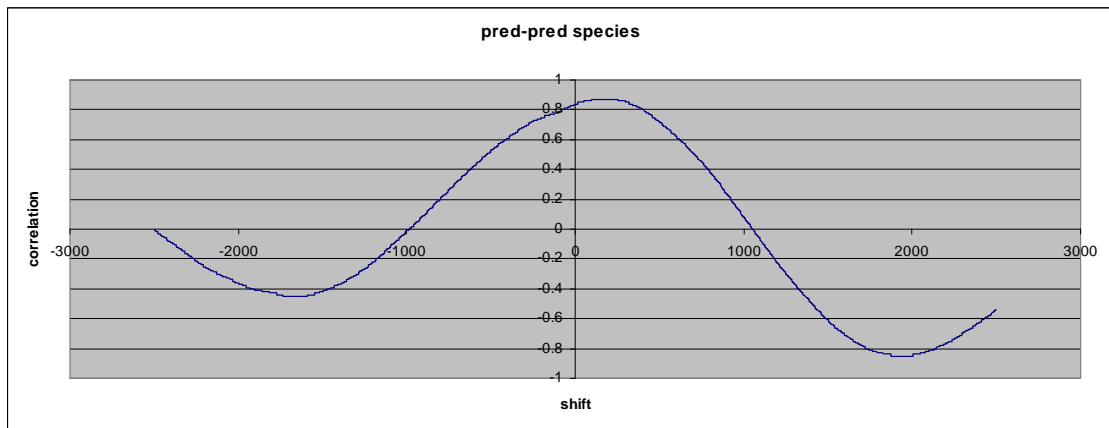


Figure 9. Cross correlation between the number of predators and the number of predator species for $-2500 \leq d \leq 2500$.

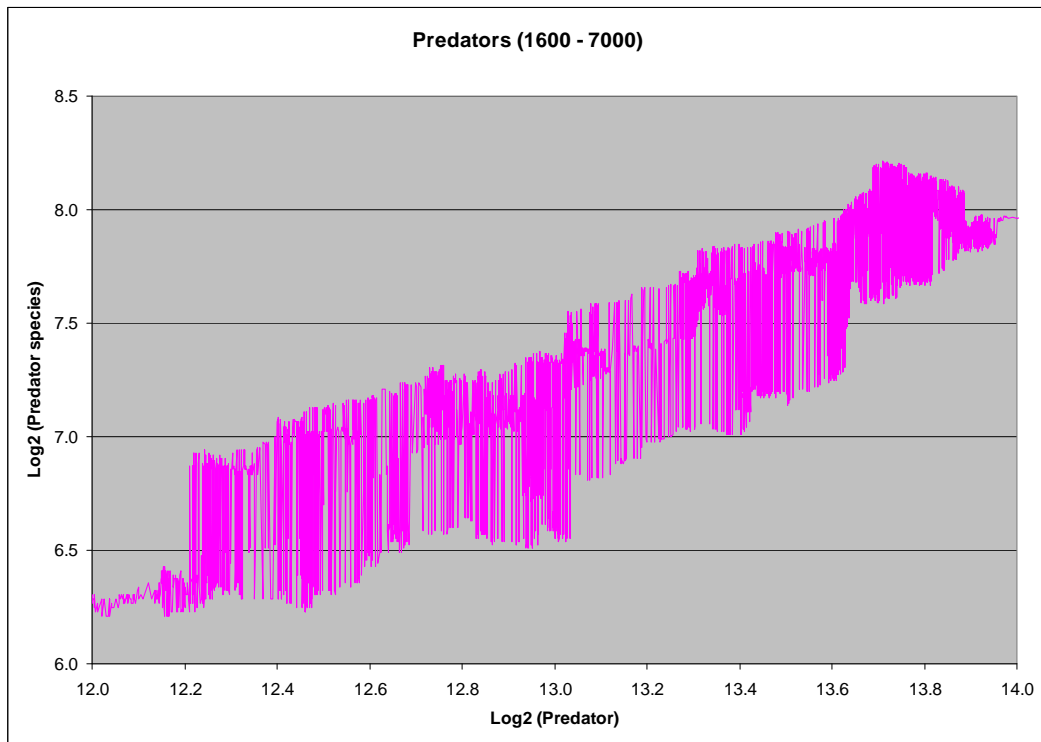


Figure 10. Correlation between the number of individual predators and the number of predator species. Slope 0.99

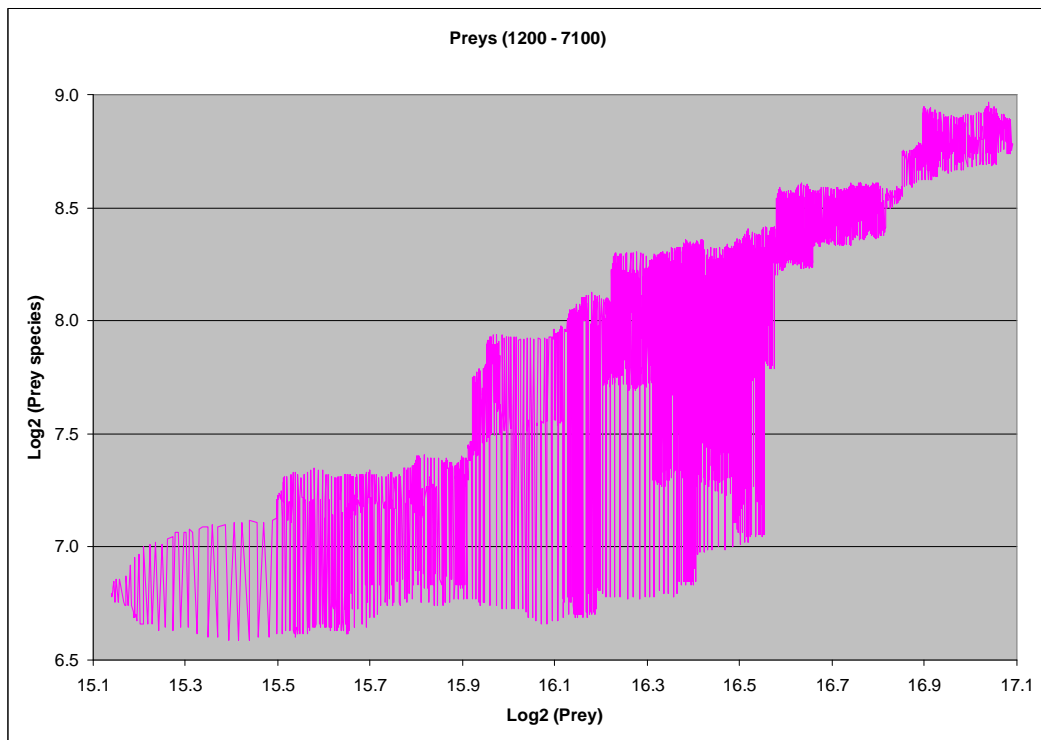


Figure 11. Correlation between the number of individual preys and the number of prey species. Slope 1.22

Figure 12 presents the overall evolution of the populations of preys and of predators in term of complexity of the behavioral model. The average number of edges in the Matrix L of both populations of preys and predators is computed at each time step. This average value grows almost monotonically for both populations, with a higher slope for predators. As a higher number of edges in a matrix also increases the energy used by the individuals at each time step, it appears that there is enough interest in adding new influence edges between concepts and then to have a more complex behavioral model to compensate for the loss of energy. It is a very interesting result since it shows the interest of the FCM as a behavioral model. The FCM behavioral model is sophisticated and useful enough for the agents (it provides the agents with an efficient way to survive and to propagate their genomic information through generations) to such an extent that a gain in behavior complexity is enough to compensate for the loss of energy. It shows also the capability of this simulation to test some evolutionary hypothesis such as showing how more complex behaviors, even with the associated drawback of an increase in energy needs, could lead to organisms with better abilities to survive and to transmit their genetic information. The number of preys is also plotted in this figure to show the correlation with the number of edges. From time steps 4,750 to 5,100 the prey population grows very fast. The average number of edges for preys grows much faster also during this period. The acceleration in the increase in the number of edges for preys begins around time step 4,450, which is before the acceleration in growth of the population of preys. This could be explained by the fact that mutations in some individuals allow one or more well adapted species to emerge and, after few generations, trigger the growth acceleration of the population of preys event. This hypothesis is strengthened by the fact that this phenomenon is also correlated to a decrease in the number of prey species from time steps 4,550 to 4,800 (the most efficient species dominate) and then to an increase in this number from time steps 4,800 to the end of the simulation (the most efficient species lead to the emergence of new efficient species).

Another analysis that can be performed is the study of the evolution of the average activation level of the concepts of a population. For example, in figure 13, we focus on the activation levels of the action concepts *explore* and *wait* of preys and we correlate their evolution with the total number of preys. The maximum cross correlation between the number of preys and the activation level of *explore* for prey is -0.41 for $d = 1330$ and the maximum cross correlation between the number of preys and the activation level of *wait* for prey is -0.41 for $d = 0$ and $d = -6$. These two coefficients are not very high and thus show that a more complex interaction should be involved. The activation level of *explore* grows very fast at the beginning of the simulation and then remains almost constant until the end. For the activation level of the *wait* concept, there is a low and constant decline from the beginning of the simulation to the time step 1,550. Globally, the activation level of *explore* is much higher than the one of *wait* during the whole simulation. The period of fast growth of the population of preys from time steps 4,750 to 5,100 corresponds to a growth of the activation level of *wait* and *explore* as well, but after time step 5,100 the activation level of *wait* tends to grow slowly and the activation level of *explore* to decrease slowly. It seems that the explanation given for figure 12 about important mutations that change the efficiency of prey species after this time step can also explain the change in behaviour observed with the change in the level of activation

of the *wait* and *explore* concepts. Noticing that the number of grass units and the overall number of preys are relatively high during this period, the exploration behaviour could be less interesting and the wait behaviour, avoiding too much energy consumption, could be more attractive.

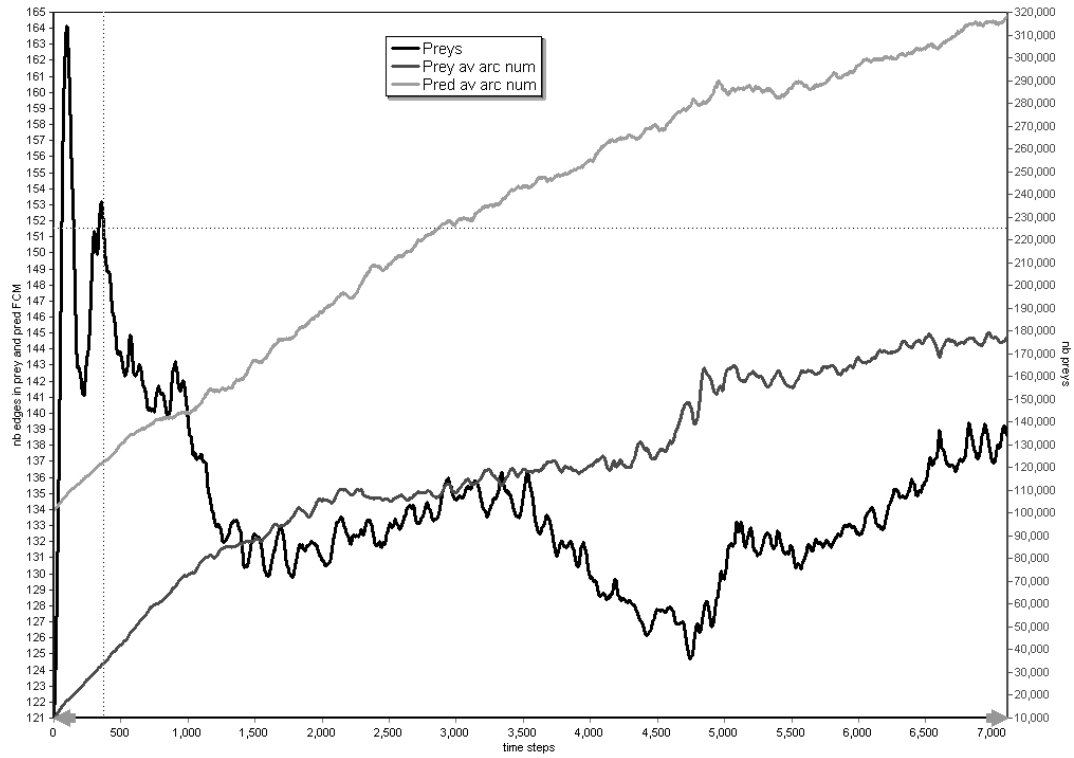


Figure 12. Evolution of the number of individual preys, of the number of edges in the prey FCM and of the number of edges in the predator FCM.



Figure 13. Evolution of the number of individual preys, of the number of preys choosing the *explore* action and of the number of preys choosing the *wait* action.

Figure 14 shows that very different patterns of evolution of behaviors can emerge. The number of predators choosing the *hunt* action follows almost exactly the number of predators (the maximal cross correlation is 0.99 for $d = 0$ and $d = 8$) whereas the evolution pattern of the number of predators choosing the *searchForFood* action vary considerably from the one of the number of predators. The maximal cross correlation is -0.76 for $d = 2367$ but also 0.71 for $d = 0$ showing that there is an overall quite good correlation between the *searchForFood* action and the number of predators. It seems that the action of hunting is constantly important during the whole simulation. Approximately the same fraction of the whole predator population chooses the *hunt* action at each time step. For the *searchForFood* action, this is not the case. At the beginning of the simulation a very small fraction of the predators choose this action. Then the number of predators choosing the *searchForFood* action grows almost constantly until the time step 3,650, even when the overall number of predators highly decreases. This can be explained by the fact that the evolutionary process allows the predators to discover the importance of searching for food to take advantage of the available food units (generated by the large amount of preys dying of lack of energy or of old age) in the world. After this maximum, the number of predators choosing to search for food decreases with the total number of predators but with a much lower slope, and then stabilizes even when the number of predators grows very fast from time step 4,950 to the end of the simulation. This final phenomenon does not seem to be correlated to the number of meat units available since this number tends to grow at the end of the simulation (data not shown).

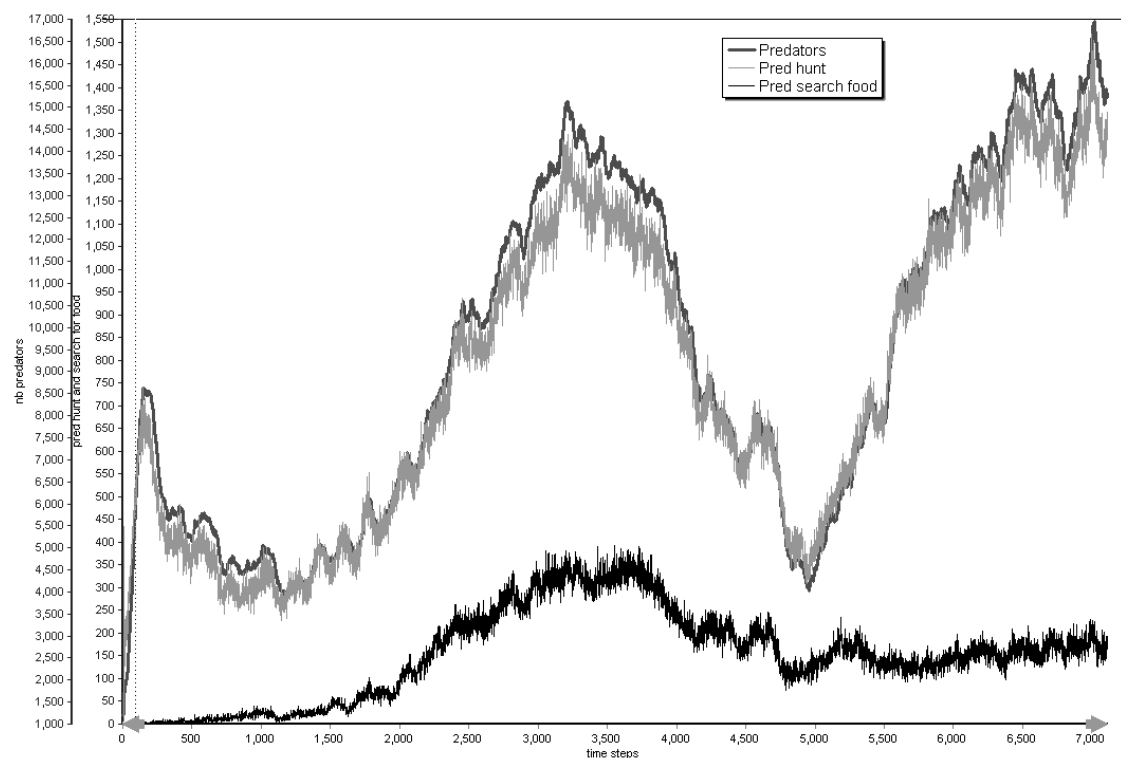


Figure 14. Evolution of the number of predators, of the number of predators choosing the *hunt* action and of the number of predators choosing the *searchForFood* action.

Figure 15 shows another non-linear phenomenon. This figure presents the evolution of the average activation level of the sensitive *foeClose* concept of preys correlated with the number of preys and the number of predators. As we could have guessed, the preys' perception of close predators depends on the overall density of both populations. The maximum cross correlation is -0.75 for $d = -47$ between the number of preys and the *foeClose* concept and is -0.65 for $d = 1642$ between the number of predators and *foeClose*. Figure 16 shows also that the number of foes close to the prey is inversely correlated in a near future with the number of preys (the preys are killed by their foes). But this phenomenon is more complex, as it is associated with an evolution of the behavior models of the individuals and with their relative positions. The dominant pattern of the beginning of the simulation (until time step 2000) is a fast and constant growth in the preys' perception of their proximity to their foes. This phenomenon is even conjugated with a simultaneous decrease in both populations of preys and predators. It should then be associated with an important modification of the relative positions of preys and predators in the world; predators should have learned to be closer to their preys. From time step 3,600 to 5,850, the pattern of the activation level of the *foeClose* concept follows almost slavishly the evolution pattern of the number of predators. Then, the preys' perception of their proximity to their foes slightly decreases and stabilizes, whereas both populations of preys and predators continue to increase. This situation is more complex to explain but could also arise from new evolved behaviors resulting in new relative positions of preys and predators.

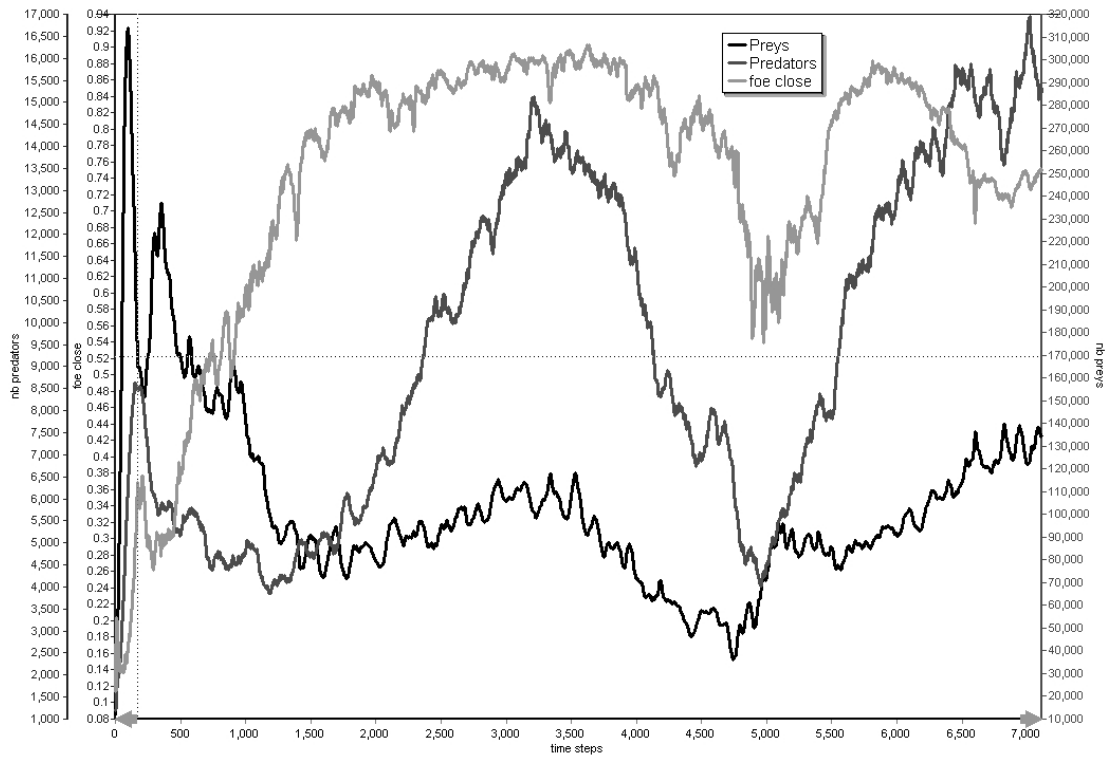


Figure 15. Evolution of the number of preys, of the number of predators and of the value of the preys' sensitive *foeClose* concepts.

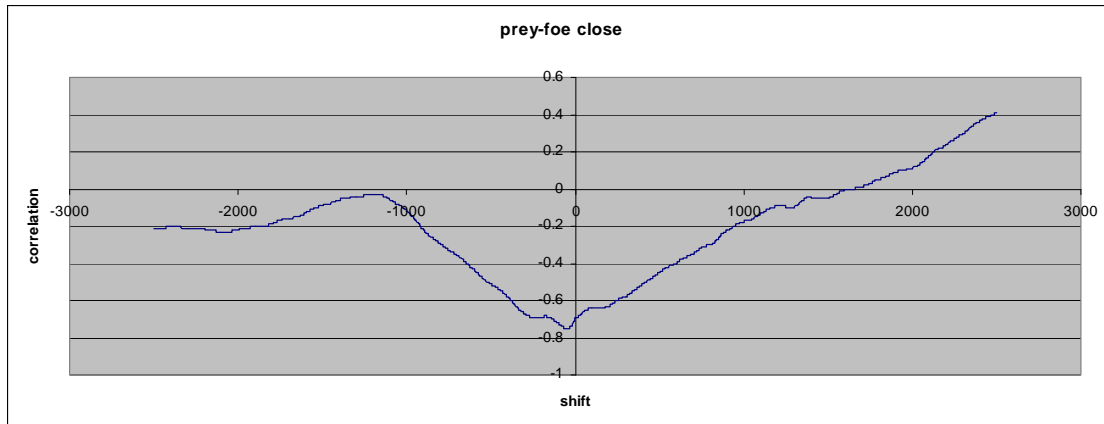


Figure 16. Cross correlation between the number of preys and the value of the sensitive *foeClose* concepts for $-2500 \leq d \leq 2500$.

It is also possible to study internal concept activation levels. Figure 17 shows the evolution of the average preys' *satisfaction* activation level and correlates it with the number of preys and the number of predators. The maximum correlation between the number of preys and their satisfaction level is 0.73 for $d = -88$ and between the number of predators and the satisfaction level of preys is 0.64 for $d = 110$, showing an almost direct influence between these parameters. It appears that after a very fast and short increase in the satisfaction level a long decrease follows, until time step 1,650, corresponding mostly to the important decrease in the population of predators. Then the preys' satisfaction level

increases very slowly almost continuously until time step 4,750, and then has a peak corresponding to the minimum number of predators and finally a very slow decrease corresponding to the important growth of the predator population. Even if the preys' level of satisfaction seems to be closely related to the number of predators, it is not the only factor. For example from time step 1,650 to 3,200, the preys' satisfaction level grows during a phase of large expansion of the predator population. The figure 18 also clearly shows that the variation in the preys' level of satisfaction follows the variation in the number of preys. Other phenomenon, such as the evolution of the number of grass units, the relative distance between preys and predators, the quantity and distance of potential partners for interbreeding should also intervene. It is likely too that the evolution of the preys' behavior models allows them to find strategies to improve their level of satisfaction.



Figure 17. Evolution of the number of preys, of the number of predators and of the value of the preys' internal *satisfaction* concept.

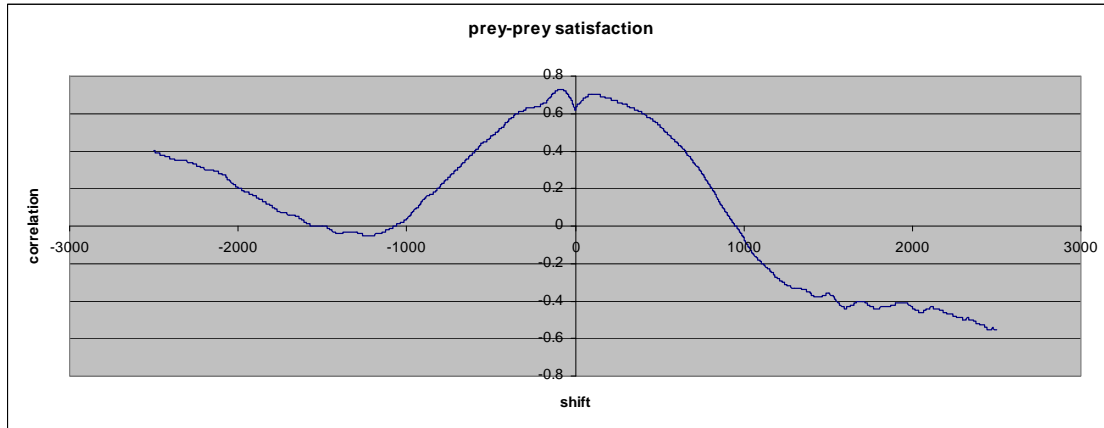


Figure 18. Cross correlation between the number of preys and the value of the sensitive *satisfaction* concepts for preys for $-2500 \leq d \leq 2500$.

It is important to notice, based on these figures and the measures of cross correlation, that the correlations between some parameters of the simulation are quite high, but perhaps not so high that what should be expected a-priori. In fact, most of the correlations in this system are complex and involve several mutually dependent parameters (for example there are clear mutual dependencies between the number of preys and predators, the level of grass, the average level of energy of preys and predators...). Moreover there is a difference in phase between the times series due to the delay between a cause and its effects. It is therefore normal that cross correlation, measuring only binary dependencies, cannot reveal all the complex network of dependencies. A more complex statistical measure coupled with a machine learning approach should be used to extract and model all these important information.

4.3 Evolution of the behavioral models

An important specificity of this simulation is that the evolution is governed by modifications in the genome of the simulated organisms and that these modifications have a direct effect on the behavioral model of the organisms. As all initial individuals (of the same type) at the first time step have the same genome, it is easy to compare the matrix of a current individual with the initial one at any time of the simulation. As each species is represented by the average matrix L of its members, it is also easy to compare a species, in term of behavioral models of its members, with the initial model. It is therefore possible to analyze what the evolutionary process has conserved from the beginning, what new influence links between concepts have emerged and what influence links have disappeared. It is also possible to see how a particular modification has spread among all the species, and therefore to have an indirect information about the importance of this influence link for the survival and the capability for the individuals of a species to transmit their genome.

As an example, shown in table 6, of such possible study, we have chosen, at the time step 7,100 of our longest run, the matrix corresponding to the 8,785th prey species that has appeared during the run. This species contains 164 members at this time step. If we compare table 6 with table 1, we can see that there are variations in most of the initial

weights (more than 0.1 in absolute value for 25 of them), there are 31 new links and 5 old links have disappeared. The 5 links that have disappeared are from (1) to (19), (5) to (19), (7) to (19), (19) to (19) and from (7) to (20), and all of them are low weight links in the initial matrix. It can be supposed that their influence on the whole behavior is negligible and, as less links means less energy spend by the individual, it is an advantage for the individual to lose these links. The large increase in the number of edges confirms the importance of a more complex behavioral model.

	16	17	18	19	20	21	22	23	25	26	27	28	29	30
1	3.95	0	0	0	0	<u>-97</u>	.9	0	<u>-13</u>	0	0	0	0	0
2	-4	0	0	0	0	<u>.27</u>	<u>-.59</u>	0	0	<u>-.11</u>	0	0	<u>.33</u>	0
5	0	<u>.56</u>	0	0	<u>.22</u>	<u>.47</u>	<u>-.48</u>	0	0	0	0	0	0	0
6	0	<u>.08</u>	<u>-.42</u>	<u>.08</u>	<u>-.35</u>	<u>-.66</u>	<u>.69</u>	0	0	0	0	0	0	0
7	0	0	.6	0	0	<u>.41</u>	<u>-.67</u>	0	0	0	0	0	0	0
8	<u>.13</u>	0	<u>-.33</u>	<u>.11</u>	<u>-.12</u>	<u>-.45</u>	<u>.36</u>	0	<u>.1</u>	0	0	<u>-.01</u>	<u>-.02</u>	0
9	<u>.31</u>	<u>3.94</u>	<u>-1.45</u>	0	0	<u>-2.24</u>	<u>2.13</u>	0	0	0	0	0	0	0
10	0	<u>-1.28</u>	<u>1.44</u>	<u>.2</u>	<u>-.43</u>	<u>1.57</u>	<u>-1.42</u>	0	0	0	0	<u>.01</u>	<u>.01</u>	0
11	<u>-.18</u>	<u>-.23</u>	0	<u>-.3</u>	<u>.41</u>	<u>1.19</u>	<u>-1.05</u>	0	0	0	0	0	<u>2.55</u>	0
12	0	<u>.2</u>	0	<u>.97</u>	<u>-1</u>	<u>-1.15</u>	<u>1.1</u>	0	0	0	0	0	<u>-3.95</u>	0
13	0	0	0	<u>-.37</u>	<u>.4</u>	<u>.56</u>	<u>-.67</u>	0	0	0	0	0	0	<u>1.39</u>
14	0	0	<u>.42</u>	<u>.3</u>	<u>-.35</u>	<u>-.86</u>	<u>.83</u>	0	0	0	0	0	0	<u>-3.72</u>
16	<u>.52</u>	0	0	0	0	0	0	<u>3.46</u>	<u>-.64</u>	<u>-1.17</u>	<u>.2</u>	<u>-1.09</u>	<u>-1.02</u>	<u>-.88</u>
17	0	<u>.33</u>	0	0	0	0	0	<u>-.74</u>	<u>2.12</u>	<u>-.64</u>	<u>.78</u>	<u>-.67</u>	<u>4.07</u>	<u>-1.74</u>
18	0	0	<u>.44</u>	0	<u>-.17</u>	<u>-.08</u>	0	<u>-.19</u>	0	<u>1.34</u>	<u>.44</u>	<u>-.2</u>	<u>-.48</u>	<u>3</u>
19	0	<u>-.14</u>	0	0	0	0	<u>-.03</u>	<u>-.26</u>	<u>.38</u>	<u>.26</u>	<u>1.57</u>	<u>-.09</u>	<u>-.3</u>	<u>-.2</u>
20	0	<u>-.3</u>	0	0	<u>.1</u>	0	0	0	<u>-.55</u>	<u>-.3</u>	<u>-1.24</u>	<u>.16</u>	<u>.54</u>	<u>.3</u>
21	<u>-.09</u>	0	<u>.02</u>	0	0	<u>-.01</u>	0	<u>-.24</u>	<u>-.84</u>	<u>-.2</u>	<u>-1.87</u>	<u>1.45</u>	<u>.88</u>	<u>.88</u>
22	0	<u>-.27</u>	0	0	0	0	0	<u>.35</u>	<u>.95</u>	<u>.2</u>	<u>1.97</u>	<u>-1.1</u>	<u>-.71</u>	<u>-.61</u>
23	<u>-.01</u>	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	<u>.03</u>	0	0	0	0	0	0	0	<u>-.17</u>	0	0	0	0
27	0	<u>-.01</u>	0	0	0	0	0	0	0	<u>.04</u>	0	0	0	0
28	<u>-.01</u>	0	0	0	<u>.12</u>	0	0	0	0	0	0	<u>.38</u>	0	<u>-.13</u>
29	<u>.09</u>	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	<u>-.23</u>	<u>.29</u>	0	0	0	0	0	0

Table 6. Example of a prey species' centre. It corresponds to the 8,785th prey species that has emerged. These values correspond to the average of the matrix L of the 164 preys, members of this species at the time step 7,100. In italic the edges that have disappeared compared to the initial matrix (table 1), in bold the edges that have appeared and underlined the weights that have significantly changed (more that 0.1 in absolute value).

Among the new edges, several have relatively high values (an absolute value greater than 0.1) and therefore can have no negligible effects on the overall behavior. For example, the edge from (2) to (29) is particularly high. This is an influence that makes sense: when there is no nearby predator the prey can stay and eat. The negative edge from (11) to (16) means that when the level of food is locally high the level of fear of the prey decreases. This is a very interesting notion that was not included into the initial model. In

the initial model, the level of fear is only influenced by itself, the distance to predators and the prey's energy level. With this edge, the evolutionary process has discovered something that is directly linked to the semantic of the internal concept of *fear*. Because *fear* influences the actions of the prey, the evolutionary process has used this concept with its right semantic to increase the fullness of the behavioral model: when there is enough food, the prey is less frightened. Similarly, the new positive edge from (8) to (16) creates a new semantically correct notion: when there is no possible mate close to it the prey's level of fear increases. The new negative edge from (18) to (20) is also very logical: when the prey wants to find a potential sexual partner, its desire of sedentarity decreases. Other very interesting edges appear as well that create feedback from action concepts to internal concepts. For example the negative edge from (30) to (22) means that the action of breeding will decrease the feeling of annoyance. This is also a very semantically important notion that has emerged from the evolutionary process. The new negative edge from (30) to (23) is very interesting and complex too. It means that the action of breeding will influence the next action by causing the evasion of the prey. To understand the importance of this new notion, it is important to remember that the offspring will act before their parents due to the ordering of the list of individuals. The parents have also lost an important amount of energy during the interbreeding action, so it is highly probable that their energy level is low and that they need to eat. But since the offspring will act before them and since it needs food too, a good next choice for the parents is to leave the current cell to increase their chance to find some available food. As both parents should be interested to leave the cell, it is important for them as well to end their moves on different cells to increase their probability to find food. Therefore, they should not try to go towards the closest food (by increasing *searchForFood*) because there is a high probability that they will reach the same cell. So a random move is better, and that is what the *evasion* action does. With this last example, it appears that some very complex and efficient behaviors can emerge from the evolution of the FCM model. It is also important to notice that all these concepts interact thanks to the dynamic of the map process. Therefore, much more complex behaviors, difficult to discover by just looking at the matrix, can emerge and drastically change the strategy of the agent.

From the species' point of view, it is interesting to study how modifications spread among the species. We do not want to develop this study here but we still want to make a short remark. We have observed, from our example matrix of table 4, that the important notions like the one corresponding to the edge (2) to (29) or (30) to (23) and to a smaller extent (8) to (16) or (30) to (22), but also many others that do not appear in the matrix of table 4, are conserved by the evolutionary process and spread among a large number of species. That confirms that the simulated evolutionary ecosystem allows the discovery, the conservation and the propagation of new important behaviors and that the species is an efficient vector of this propagation. It should be particularly interesting to extend this study in the framework of the underlying phylogeny deriving from the successive speciation events to see how important evolutionary discoveries influence the speciation process.

5 Conclusion

We have conceived an individual-based evolving predator-prey ecosystem simulation. The main characteristics of our approach are: the use of a complex model, the Fuzzy Cognitive Maps, as a model for the behavior of our individuals, the definition and implementation of a speciation concept and of a species model based on genomic distance measurement, an evolutionary mechanism allowing the combination of genomic information of the two parents and mutations and a direct link between the behavior model and the evolution mechanism as the behavior model is coded in the individual genome. The results obtained are very promising as, even with a very complex multi-levels dynamic adaptive system (involving behavioral model allowing feedback effects and short term memory, large number of interacting individuals, a multi-level food resource system, emergence and evolution of species...), the overall dynamic of the system at any level seems to be coherent and presents strong correlation patterns between its components. Agents make pertinent decisions of action considering several independent sensitive pieces of information, the evolutionary system generates and selects behavioral models with new innovative notions, the correlation between the number of species and the number of individuals is in accord with such existing correlation.

One major and unique contribution of this simulation is that it combines a behavioral, an evolutionary and a speciation mechanism. This is the only simulation modeling the fact that individual behaviors affect evolution and speciation. That promises future interesting studies on theoretical biology. For example, we will be able to study 1) whether the heterogeneity of population increases speciation events by increasing the probability for individuals of that population to find compatible reproductive partners; 2) whether a relation between individual distance values and the probability of interbreeding evolves (evolution of mating strategies); 3) hybridization phenomenon. For the latter, while computing individuals' membership, we consider individuals from the same species as well as from different species. Such model accounts for the facts that speciation can arise not only from the fission of existing species (allopatric, sympatric) but also from the hybridization of existing species. Since we compute the distance between the genotype of a newborn and the average "genotype" of all existing species (i.e. we consider individuals from different species), we keep the possibility for hybridization opened. Speciation can be reversible, contrary to Mayr [23], according to whom speciation is not complete until irreversible. We keep the possibility that future reproductive isolation will be undone; reproductive isolation is not irreversible, so there is no claim for the permanence of isolating barriers and two "good" species can be fused back into one.

An important limitation of such kind of large scale simulations is that it produces a huge amount of data with complex correlation patterns that are not easy to analyze and interpret. We have computed cross correlation coefficients between several parameters of the simulation and showed that several of them are highly correlated. We have also discussed more complex situations in which more than two parameters are correlated. In these cases, the measurement of cross correlation coefficients is not enough to reveal the dependencies. More complex statistical analyzes have to be applied to these data to discover the complex interdependent pattern produced by the simulation. We plan to use a machine learning approach based on dynamic Bayesian network to analyze our data.

This complex statistical model has the capacities to reveal multiple dependencies, network of dependencies and shifts between analyzed time series.

There are a lot of other possible improvements and perspectives for this project. The more practical and pressing things are to rewrite the application to allow the use of both a parallel shared memory approach using openMP and a distributed approach using MPI and to create a visual interface to display the states of the world online¹⁵. These two improvements are essential to be able to run a large scale and long term simulation, to have a better understanding of the local behavior of the agents and a global view of the distribution of species. Then several very interesting properties could be added such as: more types of individuals and more resources to increase the food resource hierarchy; capacity of life time learning by experimentation and by observation for the agents, which is very easy to include in this simulation (comparing the level of satisfaction of an agent before and after an action and reinforcing the edges that allow to choose positive actions and decreasing the ones that allow the choice of negative actions); using a non-linear (Gaussian) distribution for the probability of interbreeding according to the genomic distance; conceiving a diploid genome for the agents with dominant and recessive alleles; constructing the phylogenetic tree of the species, etc. Finally, this simulation will be the framework for the study of some specific ecological questions in collaboration with biologists. For example, this approach can be used to study complex ecological and evolutionary processes such as the species abundance distribution, patterns and rates of speciation, the evolution of sexual and asexual populations, the interaction and diffusion of an invasive species into an existing ecosystem, etc.

Acknowledgments

The authors would like to thank Melania Cristescu (Great Lakes Institute for Environmental Research, Windsor, Ontario) for her support and for her instructive and detailed comments.

References

1. Axelrod R., Structure of decision, Princeton University Press, Princeton, 1976
2. Clarke, C.A., Sheppard P.M., Interactions between major genes and polygenes in the determination of the mimetic patterns of *Papilio dardanus*. *Evolution* 17:404-413, 1963
3. Coyne J, Orr H.A. Speciation. Sinauer Associates, Sunderland, Massachusetts, 2004
4. DeAngelis D. and Mooij W.M., Individual-based modeling of ecological and evolutionary processes. *Annu. Rev. Ecol. Evol. Syst.*, August 2005
5. Dickerson J.A., Cox Z., Wurtele E.S., Fulmer A.W., Creating metabolic and regulatory Network models using fuzzy cognitive maps, *Proc. of IFSA World Congress Vol. 4*, 2171-2176, 2001
6. Dieckmann, U. & Doebli, M., On the origin of species by sympatric speciation. *Nature* 400: 354-357, 1999

¹⁵ The current version is only a text version

7. Dobzhansky T. Genetics and the origin of species. 3d ed. Columbia University Press, New York, 1951
8. Durrett R. and Levin S.A., Spatial models for species-area curves. *Journal of Theoretical Biology*, 179(2):119–127, 1996
9. Georgiou D.A., Makry D., A Learner's Style and profile Recognition via Fuzzy Cognitive Map, *Proceedings of the IEEE International Conference on Advanced Learning Technologies*, 2004
10. Grimm V., Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling*, 1999
11. Grimm V. and Railsback S.F., *Individual-based Modeling and Ecology*. Princeton University Press, 2005
12. Holland J.H., *Hidden Order: How Adaptation Builds Complexity*, Addison-Wesley 1995
13. Hrabar, P.T., Jones T. and Forrest S., The ecology of Echo. *Artificial Life* 3(3):165-190 1997
14. Hubbell S.P., *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press, 2001
15. Huston M., DeAngelis D., and Post W., New computer models unify ecological theory. *BioScience*, 38(10):682-691, 1988
16. Judson O.P., The rise of the individual-based model in ecology. *Trends in Ecology and Evolution*, 1996
17. Kosko B. Fuzzy Cognitive Maps, *International Journal Man-Machine Studies* 24, 65-75, 1986
18. Levin S.A., Patchiness in marine and terrestrial systems: from individuals to populations. *Philos. Trans. R. Soc. London Ser. B*, 1994
19. MacArthur R.H., and Wilson E.O., *The Theory of Island Biogeography*. Princeton University Press, 1967
20. Mallet J., A species definition for the Modern Synthesis, *Trends in Ecology and Evolution* 10(7), 294-299, 1995
21. Mallet, J., McMillan, W.O. & Jiggins, C.D., Mimicry and warning color at the boundary between races and species. In: *Endless Forms: Species and Speciation*. Eds: Howard, DJ & Berlocher, SH, Oxford Univ. Press, New York, 390-403, 1998
22. Mallet, J., Species, concepts of. In: *Encyclopedia of Biodiversity* (S. A. Levin, ed.), vol. 5, pp. 427±440. Academic Press, San Diego, CA, 2001
23. Mayr, E., *Animal species and evolution*. Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1963

24. Medina, R. F. The Role of Host-Plant Species in the Differentiation of Sympatric Populations of Hymenopteran Parasitoids, PhD thesis, University of Maryland, 2006
25. Mishler, B. D. and Donoghue M.J., Species concepts: a case for pluralism. *Syst. Zool.* 31:491-503, 1982
26. Ozesmi's U., Ecosystems in the mind: Fuzzy Cognitive Maps of the Kizilirmak Delta Wetlands in Turkey, Ph.D. Dissertation, University of Minnesota, 1999
27. Preston F.W., The canonical distribution of commonness and rarity: Part I. *Ecology*, 43(2):185–215, 1962
28. Preston F.W., Noncanonical distributions of commonness and rarity. *Ecology*, 61(1):88–97, 1980
29. Rice J.A., Miller T.J., Rose K.A., Crowder L.B., and Marschall E.A., Growth-rate variation and larval survival: inferences from an individual-based size-dependent predation model. *Can. J. Fish. Aquat. Sci.*, 1993
30. Ronkko M. An Artificial Ecosystem: Emergent Dynamics and Lifelike Properties, *Artificial Life* 13, 159-187, 2007
31. Rosenzweig M.L., Species diversity in space and time, Cambridge University Press, 1995.
32. Shugart H. H., A theory of forest dynamics: The ecological implications of forest succession models. Springer-Verlag, 1984
33. Smith E., Eloff J.H.P., Transaction based Risk Analysis – Using Cognitive Fuzzy techniques, in Eloff J.H.P. et al. (Eds.), *Advances in Information Security Management & Small Systems Security*, Kluwer Academic Publishers, 2001
34. Tisseau J., *Réalité virtuelle – autonomie in virtuo*. Habilitation à Diriger les Recherches, Université de Rennes, France, 2001
35. Tisseau J., Parenthoën M., Buche C., and Reignier P., Comportements perceptifs d'acteurs virtuels autonomes. une application aux cartes cognitives floues. *Revue des Sciences et Technologies de l'Information, série Techniques et Sciences Informatiques (RSTI-TSI)*, 24(10):1259-1293, 2006
36. Tolman E.C. Cognitive maps in rats and men, *Psychological Review* 42(55):189-208, 1948
37. Tsadiras A.K., Margaritis K.G., On Transition and Convergence Properties of Extended Fuzzy Cognitive Maps, Proc. 2nd Hellenic European Conf. on Mathematics and Informatics, Athens, 1994
38. Uchmanski J. and Grimm V., Individual-based modeling in ecology: what makes the difference? *Trends in Ecology and Evolution*, 1996
39. Ward C.R., Gobet F., Kendall G., Evolving Collective Behavior in an Artificial Ecology, *Artificial Life* 7, 191-209, 2001

40. Williamson M. and Gaston K.J., The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. *Journal of Animal Ecology*, 74, 409-422, 2005
41. Wurtele E.S., Li J., Diao L., Zhang H., Foster C.M., Dickerson J., Brown A., Cox Z., Cook D., Lee E.K., Hofmann H. (2003), *Comparative and Functional Genomics* 4(2), 239-245