Use and Adequacy of Computer Paradigms to Simulate Bioinspired Synthetic Landscape Ecologies

Jean Le Fur1, Pape Adama Mboup2 and Moussa Sall3

1Institut de Recherche pour le Développement (IRD), Centre de Biologie pour la Gestion des Populations (CBGP), Campus Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez, France
2Lab. Informatique, Faculté des Sciences et Techniques, Univ. Cheikh Anta Diop (UCAD), Dakar, Senegal
3Dépt. Informatique, Univ. G. Berger/Saint-Louis Sénégal and Lab. IRD-BIOPASS, Campus Bel-Air, Dakar, Senegal

Keywords: Bio-inspiration, Ecoinformatics, Natural Computation, Synthetic Ecology, Robustness, Object-oriented Programming, Agent-based Model, Rodent.

Abstract: In the field of ecoinformatics, synthetic ecologies approach intends to reproduce in silico the architecture and functioning of a real domain 'as it works'. When the systems to be represented are complex, bioinspiration is proposed as a relevant approach to develop robust modelling. This study aims to develop, using basic object-oriented paradigms, and in collaboration with biologists, a comprehensive synthetic ecology about a given application domain: wild rodents' population dynamics. To address the complexity of the field, the architecture is gradually grown and shaped from integration of successive and diverse case studies modelling in the application domain. Developed for more than ten years, the same model has made it possible to represent historical, cellular and/or ecological processes at the scales of a country, a region, a city or a laboratory as well as a diversity of interacting living beings. Results present how principles such as composition, aggregation, inheritance, generalization have been used to elaborate a synthetic ecology. These paradigms altogether constitute a rich, and improvable, toolbox offering a varied set of possible uses to formalize bioinspired landscape or ecologies.

1 INTRODUCTION

Bioinspiration in the more common sense is considered as using natural solutions to solve engineering questions. Conversely, in the field of bio-socio-ecological modelling, one's aim is to use computer or formal engineering to question natural processes. The latter models often are abstraction of the processes under investigation. For example, differential equation sets can be used to model epidemic process such as in the classical susceptible-infected-recovered (SIR) model (e.g., Chen et al., 2020). In some cases, bioinspired algorithms are also used to formalize natural contexts. Several approaches belonging to so-called natural computing (Kari and Rosenberg, 2008) or ecoinformatics (Yao, 2006) are developed in that direction such as cellular automata, genetic programming, swarm intelligence, artificial immune systems... In this field of modelling too, this bioinspired approach may be not related to the biological processes studied. For example, neural networks may be used to describe a hierarchically structured ecosystem (Olden et al., 2006), or a bioinspired cell model (P-system) be developed to study a trophic web (Colomer et al., 2011) ... In any of these directions there is generally an assumed gap between the concrete problem to solve and the model formalism (strategic models, Holling 1966 in Evans et al., 2013). In the case of bio-ecological modelling this gap however may restrain each model application to the particular use case it is fitted to or trained on (Svoray and Benenson, 2009). This rules out an explicative description that could be generalized in other contexts or produce relevant forecasting even in a changing context. Model’s robustness may there be questioned.

On the other hand, tactical models (Holling 1966 in Evans et al., 2013) are focused on prediction and robustness. For this purpose, their aim is to capture the mechanisms that govern real-world dynamics with a virtual copy working the same way (Svoray and Benenson, 2009). Such approach can be found in synthetic biology, a field dominated by research at the microbiological level (Dunham, 2007, Shou et al., 2007). At the scale of individuals, the so-called Functional-Structural Plant Modelling or FSPM...
(Buck-Sorlin, 2013) integrates biology and physiology to reproduce plant growth (Dejong et al., 2011). At a larger scale the emerging field of synthetic ecology proposes bioinspired solutions in a large range of research domains such as city planning, sedimentation modelling (Cantrell and Holzman, 2014). In Ecology or natural landscape issues, DeAngelis and Mooij (2003) also proposed a so-called “mechanistically rich” approach in which a maximum of the entangled factors producing the natural dynamics studied could be integrated.

Developing such promising approaches supposes a focus on robustness. Indeed, it implies accounting for a diversity of items with highly heterogeneous distinct behaviours and need tough computation to manage this complexity.

On the other side, Nature appears as the one and only robust ‘model’ where drivers are universal, where processes and components function altogether and resist to change. Bioinspired computation may be then one possible direction to elaborate robust ecological or landscape models. Since Nature is indeed the only robust functioning system, the more one uses the same processes for computation, the more one can produce robust analogies and hence simulations, of the worlds investigated.

In this study we propose to use computer paradigms, in collaboration with bio-ecologists, to mimic the known behaviour of an application domain in Ecology with the objective to produce (i) the more robust possible model and (ii) a comprehensive tool for interaction between biologists and modelling.

Following Wirth (2006), Object-Oriented Programming was selected as the privileged approach to develop a model as closed as possible to what is known of the functioning of Nature.

We first present our approach to develop a synthetic ecology that could be as bioinspired as possible. The structure of the resulting architecture is then described by indicating the computational paradigms, derived mainly from object-oriented formalism, that have been shown as the most relevant for mimicking the known functioning of ecosystems. The advantages and disadvantages of object-oriented programming from this point of view are discussed.

2 MATERIAL AND METHOD
2.1 Model Purpose

The objective of this work is the long term elaboration of a generic model of wild rodent ecology. The aim is primarily to provide dynamic simulations where knowledge coming from multidisciplinary thematic like Bio-Eco-Sociology, Geography, etc. can be articulated. The expected outcome would be a relevant tool with which specialists could compare simulations with most indicators and knowledge they are accustomed.

Two specifications found the approach:

A) The model must be as comprehensive as possible to take into account the various sources of fluctuations at stake in the ecosystems or landscapes under investigation.

B) The model must hence be robust to multiple contexts. Robustness would permit also to instantiate a model that could be queried from multiple multidisciplinary points of view (i.e., provide the variety of indicators with which bio-ecologist are used to).

To achieve (or rather go in the direction of) this objective, we have retained as a principle that Nature is the unique example of architecture and function that fulfils robustness and exhaustiveness. Hence, the greater the fidelity to the Natural ‘model’, the greater the pledge of robustness. We therefore sought to reproduce the available bio-ecological knowledge with the most bioinspired modelling schemes possible.

2.2 Context: Research in Rodents’ Bio-ecology and Epidemiology

This long-term project was carried out in a laboratory of biology and ecology within a team of scientists specialised in wild rodents, mainly in West Africa (Granjon and Duplantier, 2009). These populations are studied as pests (crops or dwellings) or as reservoirs of pathogens involved in numerous epidemics.

A preliminary investigation was first carried out by mean of individual interviews with biologists belonging to or associated with this laboratory. The interviews aimed to identify the characteristics of the knowledge domain (rodents) to be represented. The content of five interviews was then reified and integrated to identify the different components, processes, variables, indicators implemented in rodent ecology in multiple contexts.

The result of this work (Le Fur, 2014) highlighted a rich field of knowledge with a great diversity of:
- constituents with, particularly, numerous species interacting,
- processes and scales, especially spatial ones, to be considered,
- approaches such as for example Agroecology, Biogeography, community Ecology, Eco-
Immunology, landscape or population Genetics, Phylogenetics, Physiology, - indicators used and the way they are obtained to observe the ecosystems concerned.

2.3 Approach: Growing the Model as a Heuristic

Given this complexity it was not possible to design ‘from scratch’ a simulator capable of representing most aspects of a synthetic ecology. We therefore chose, as a heuristic, to make grow the model by successively formalizing concrete case studies. Each new case study brings new facets and new constraints to be resolved in order to account for field knowledge. The iterative integration of various cases then makes it possible to gradually evolve model structures, function as well as model parameters and progress in the development of a synthetic model. This approach was also used as a testbed for the robustness of the model as it was developed (see below).

Each case study was chosen in the first place according to interest of biologists for a model of this type in their field. Beyond that, in the modelling project, we tried to select the most different studies possible, however still in the field of bio-ecology of wild rodents, so as to test the robustness of the model to various contexts: each particular situation or use case, once integrated, constitutes in turn a constraint for the preceding ones and has to be accounted within the code.

The model was developed in Java using the agent-based Repast-Simphony platform (North et al., 2013) and using Eclipse features for refactoring operations in particular.

Table 1 and the text below provide a description of the main contribution to the model for six distinct use cases out of the 11 developed throughout the whole project.

- A first study in a dynamic agricultural landscape made it possible to set up the notion of nested spaces in which similar space cells can be grouped to compose a higher level entity (a field, a road, etc.).
- The second case study looked at representing the genetic process of reproduction, leading to the production of one being from two others. The model represents the genetic process including gene, chromosome, ..., meiosis, fusion and crossing of chromosomes, rejection of aberrations, transfer of genes. The model was validated by reproducing the results of an animal facility experiment (Comte, 2012) and was generalized to all case studies.
- These first stages of the model made it possible to implement and refine a study on a rodent trapping experiment where we could carry out ‘reality-inspired’ formalization of rodent traps scientifically scattered in the simulated landscape. Bio-inspiration permitted here to reproduce straightforwardly the trapping protocol and produce the necessary ecological and genetic

Table 1: Main characteristics of the most distinctive case studies successively formalized to elaborate the model.
outputs that were necessary to reproduce and compare the composite estimators familiar to environmentalists (Diakhate et al., 2014).

- The following studies were interested in the application and the adaptation of the model to much larger scales, it was thus possible to represent the historical evolution of commercial transport on the scale of a century and a whole country and study the probability of rodents (black rats) boarding commercial vehicles (boats, trains, trucks) to gradually colonize the entire country. The model has been refined here by taking into account human agents as well as various adaptations such as graph routes (Mboup et al., 2015a, 2015b) to represent their moves.

- Another implementation has been developed to account for the behaviour of rodents in a Sahelian savannah in Africa. The hierarchy of behaviours has been reorganized here using knowledge from biologists. Here we have taken into account predation (owls that eat gerbils) and circadian activity (cycles of human and animal activity different between night, day, dawn and dusk).

- Finally, a recent study (Sall et al., In prep.) looked at the spread of epidemics (zoonoses) linked to urban mice and their proximity to humans (commensal animals). The epidemic chain has been formalized by integrating biting ticks and pathogenic bacteria, cats and human activity at high resolution. This choice has led to an interest in particular phenomena such as the ability of mice to walk along residential walls to move around (Sall et al., 2019).

2.4 Use of Computer Paradigms

With each addition of a new case study, the model has been reworked and reshaped according to new features and constraints introduced. This work mainly consisted in generalization that is factorizing attributes or methods shared between several studies and move them up to an upper parent class (owl nest and rodent burrow generalized into animal home for example).

One important step in this reworking has been also the repeated use of refactoring. As we progressed in understanding the emerging architecture, the 'natural' meaning of variables and methods was refined to fit as much as possible to biological knowledge. Refactoring hence gradually improved consistency, readability, genericity and, consequently, maintenance of the growing model and the backward compatibility with all formalized use cases.

Finally, each behaviour formalized in the model was questioned, with the help of biologists, as possibly generic. This led to enrichment of the class hierarchy with encapsulation of natural functioning at the specific level at which they occur.

Depending on the specific needs, encapsulation, delegation, aggregation, polymorphism, interfacing and other paradigms of the sort have been used to reproduce by analogy the structures known as occurring in Nature and elaborate a clean architecture. In this process, apart from generalization, three particular object-oriented programming paradigms have been focused on to elaborate the model: composition, aggregation and inheritance (see results section).

At the end of this compilation of the case studies, 1,143 minor or major reconfigurations (version commits) of the model architecture and function were carried out. The final source code represents 22,248 lines (including data input, data retrieval, and display output) of which 9,066 lines of code (41%) for the business model which constitutes the synthetic ecology itself. These 9,066 lines bring together 98 classes, 149 attributes, 181 relations and 761 methods. Regarding the latter, they contain 168 (22%) calls to super and 175 (23%) overridden methods.

3 RESULTS

The architecture obtained at the end of the study is presented in more detail (UML diagrams) in Le Fur et al. (2017). In this article we focus on programming paradigms that have been useful in advancing towards a bioinspired synthetic ecology. To develop and refine the code in this direction, three programming paradigms were found particularly relevant: composition, aggregation, inheritance.

3.1 Composition

Composition was used in this model for two generic structures of ecosystems and landscapes.

The first is bio-inspired and straightforward to implement: genomes are known and described as composition and they were formalized in this way. Thus, during mating of any two reproducing agents, we obtain the following functional structure:

\[
\text{genome} \rightarrow \text{diades (paired chromosomes)} \rightarrow \text{chromosomes} \rightarrow \text{genes}.
\]
The composition principle has made it possible to computationally reproduce the functioning of a genome in a manner analogous to known reality.

The second use of composition concerns the formalization of space into successive entities such as "cell → region → landscape" or "cell → cage → animal facility"... However, this choice is here an arbitrary classification linked to observation and which does not correspond to a genuine architecture of Nature.

3.2 Aggregation and Recursion

The principal use of aggregation in a bio-inspired perspective has been developed to account for a first fundamental organization of Nature. Nature is universally a nested system ranging from the smallest quantum particle to the Universe. Within this system each level arises from the emergence of interactions between components of the lower level: atom - molecule - cell - organ - organism - population... These different levels are hence embedded within each other. At the scale of an ecosystem or a landscape each component is always simultaneously a container for other components and itself a content. To reflect this architecture, all agents have gradually inherited from an abstract class ('Container') implementing an interface of the same name (Figure 1).

Figure 1: The minimal interface finally obtained for any object, included agent, sharing roles of container and content. Methods underlined recursively run through the cascade of containers contained (see Figure 3).

Upon completion of the case studies integration, the Container class becomes a founding class, close to the root of the business model, and from which all objects and agents inherit. This thus made it possible to `naturally' formalize a rodent in a culture (in its meal), a prey in the talons of an owl, a bacterium in a tick or an embryo inside an animal (Figure 2).

Figure 2: Selected examples of embedded chains of containers currently observable in the model’s simulations.

3.3 Inheritance

A second constitutive architecture of Nature was used to structure the model which results from the theory of evolution (Darwin, 1859). Living species diversify over time by inheriting the fundamental characteristics of their ancestors while having new functions or structures that are specific to them. Organisms are therefore organized according to a single phylogeny unique to the living world and within which each individual of any species has genetically inherited all the specific properties of its ancestors. The object oriented principle of inheritance is a paradigm working like a phylogeny. Taxonomic hierarchy (Marcos and Cavero, 2002) was thus chosen as a robust approach to account in a sparing way of biodiversity as it is naturally constructed. This approach has led to the definition of a tree structure of genomes delegated to that of the modelled agents (see below). Each genome provides specific characteristics to agents that bare them.

Each genome is thus analogous to a knowledge base associated with any agent and allowing it to achieve its growth, reproduction, life cycle... with the proper parameters (speed, litter size, weaning age...) of its species. Over the course of the case studies, and without limitation, it has been possible to characterize plants, mammals, animals, birds, arthropods, bacteria (Figure 4), without calling into question this architecture. These characteristics are transmitted at each act of reproduction occurring in a simulation, whether it is a rodent, bird, etc.

Coupled with the preceding genome hierarchy, and as a founding principle in agent-based modelling, the inheritance paradigm is also the core of the functional model within which agents’ behaviours and interaction are coded. It was elaborated here in...
4 DISCUSSION

4.1 Genericity, Robustness and Reusability

Developed for ten+ years, the model has thus made it possible to represent a laboratory experiment, historical processes at the scale of a country, intracellular or ecological processes at the scales of a landscape, country or city. By trying to systematically represent Nature as it is known, and thanks to the process of generalization, this approach makes it possible to reach a robust modelling allowing to formalize and make interact in the same shared schema a great diversity of living beings (humans, rodents, birds, plants, parasites, predators).

The resulting architecture remains consistent and extensible. It permits, almost straightforwardly, to put a tree in a landscape, then a nest in the tree and an owl in the nest. One can also address complex issues such as the occupation of an organism by a parasite entering and leaving its organic container. Moreover, the scheme improves continually. As an example, the very first use case started with a reduced set of possible desires (forage, reproduce) used in the deliberation scheme of rodent agents; the model then gradually enriched with, among others, new desires such as resting, wandering, dispersing, flee, hide, suckle... Over time, the relevance of the simulations for each and every specific case studies thus continuously increases.

4.2 Paradigms Adequacy

The functional hierarchy that has been established uses nothing more than the classical paradigms of the object-oriented approach. But this tree structure also revealed itself, with the collaboration between biologists and modellers, biologically relevant. The robustness obtained makes it possible, for example, to...
integrate new types of organisms without calling into question the model. It is also likely that, when retaining only the uppermost abstract classes and interfaces, the model could be dedicated to simulate natural systems entirely distinct from rodent biology.

The computational paradigms that have been found efficient to copy natural mechanisms (composition, recursion, aggregation, etc.) come from a wide range of possible implementations that could be exploited for bioinspired modelling. Each is not used systematically but whenever necessary and appropriate. They thus constitute a combination of ‘ad hoc’ formalisms which are available as in a toolbox to represent in a bioinspired way particular aspects of Nature.

However, certain modes of operation of Nature seem difficult to take into account. This is the case, for example, of mechanisms operating simultaneously on several spatial and temporal scales. A study with this same model (Mboup et al., 2017) showed that it was possible to develop algorithms formalizing this mode of operation. But addressing this issue leads to overly sophisticated algorithms that move away from the bioinspired issue.

Finally, the question arises of the relevance or the need to use multiple inheritance to develop a synthetic ecology. For example, in this study, rodents can be colonial, fossorial, domestic, commensal, circadian (Major functions, Figure 5) with different combinations depending on the species. For a bioinspired model, to this “logic” of Nature should correspond an adapted computer formalism. In the context of the Java language that was chosen at the start of the project, the interface paradigm is not sufficient to account for the multiple inheritance of behaviours because it requires code duplications. However, implementations do exist with the so-called interface default methods of Java 8 (Mohnen, 2002) that could be used.

Achieving a synthetic ecology using a language dedicated to the management of multiple inheritance (e.g., Malayeri, 2009) could thus constitute an approach to be favoured for the development of bioinspired simulations of ecologies or landscapes. However, multiple inheritance gives rise to inconsistencies such as the diamond problem (Truyen et al., 2004). These problems may question its relevance to mimic a real functioning of Nature which, for its part, does not present any inconsistencies.

5 CONCLUSION

Computer paradigms and in particular those related to object-oriented formalism all together constitute an
Use and Adequacy of Computer Paradigms to Simulate Bioinspired Synthetic Landscape Ecologies

The authors would like to thank the members of the “rodent group” of the Centre for Biology for Population Management (CBGP) as well as the members of the BioPASS laboratory for their essential support in the development of this study.

REFERENCES


les transports humains sur un siècle.


