# AN EVOLUTIONARY LEARNING STRATEGY FOR MULTI-AGENT SYSTEMS

*Marcio Lobo Netto, Eder Antonio Rangel Marques and Henrique Schützer Del Nero*

University of São Paulo
São Paulo - Brazil

*Abstract*

In this paper we describe the incorporation of an adaptive learning strategy to the framework we have designed to describe artificial life characters. This strategy is based on the usage of language structures to represent actions and strategies developed by these artificial characters to solve their problems. A set of problems related to how characters evolve and learn may be studied here, ranging from basic survival in their environment to the emergence of knowledge exchange supported by the usage of language to communicate ideas.

## 1. INTRODUCTION

Artificial Life characters have been employed in a large diversity of situations, ranging from those where analytical models as cellular automata [1-3] or logic coupled maps [4] are well applied, to other cases where a more close representation of true characters are used [5-8]. This last case comprehends many computer animation purposes, where movie directors want just to coordinate sequences of actions played by virtual actors.

While in the first case analytic functions can be used to represent each element and its relation (dependence) with other elements in the virtual environment, in the second case structured and hierarchical models need to be used. We need then to consider some kind of structural and functional hierarchy to describe these characters.

Our group has been working on some of the many aspects related to artificial life, and therefore we conceived a framework to describe the parts that compose the multi functional structure of interesting virtual characters. This framework allows us to add new features to those modules already conceived, or even to add new modules. This paper describes the inclusion of two new modules responsible respectively for the visual perception and cognition of the actor. However emphasis is given to the second aspect. Our framework has been initially conceived to support an interface between the actor and its environment based on sensors, simulating what is performed by the human senses. We have focused our implementation first in the simulation of the vision system, including an image capturing process and its consecutive analysis and classification. The resulted perception is passed to a cognitive module, responsible for the decision taking, which leads to commands given to actuators and communicators.

The following sections present the previous work (section 2), the character framework (section 3), a proposed language model (section 4) and the characters learning skills (section 5). Finally, we present the conclusion with further work (section 6).

## 2. PREVIOUS WORK

Many authors are currently working with artificial life, which is a rich field of research. In this paper we focus our attention to those that have been applying cognitive skills to control the behavior of their virtual characters, providing these actors with some kind of personality.

### 2.1. Work from other Groups

Some researchers have been involved with this field looking for models that would describe how real life began and evolved. In fact they were looking for a universal life concept, which should be independent of the media on which they exist [1]. Other scientists were looking for physical models to give natural appearance to their characters. Very interesting results have been achieved here [5-7]. Although very different in nature, these works have been related to evolution and natural selection concepts. In many of them, evolutionary computing schemes such as genetic algorithms have been an important tool to assist the conduction of transformations in the genotype of virtual creatures, allowing them to change from one generation to another. Mutation and combination (by reproduction) provide efficient ways to modify characteristics of a creature, as in real life. Selection plays a role, choosing those that, by some criteria, are the best suited, and therefore allowed to survive and reproduce themselves.

Terzopoulos [5] presented some papers showing how to develop artificial creatures supporting strategies to simulate natural behavior and cognition. Herewith he showed the possibility to train these characters to perform a certain class of actions, or even to let them learn how to perform sophisticated actions.

Sims [7] proposed an evolutionary model to evolve creatures, where both morphology and behavior adaptation is considered. The results are very impressive. After performing some adjustments on initially simple models, leads to characters well adapted to their environment.

### 2.2. Our Previous Work

A general and flexible framework has been proposed in our first project allowing us to establish an open and extensible model to support different types of implementation. This first model – WOXBOT [9] – consists of three main classes of modules [Fig. 1]: a set of sensors, responsible for gathering environmental information and translating it into an internal representation at each character. This information passes through different types of processing stages, which are intended to classify it. For instance, the visual module transforms a pictorial (bitmap) representation of an image, gathered through a visual process simulating photography, into a set of logical symbols, which are in turn given to the cognitive module. The cognitive module is able to conduct decision processes based on this input and on its own internal state. In this case, we do not consider knowledge acquirement neither handling. No learning skill is present, and the actors perform their action instinctively. The decisions taken here are given to actuators (not represented in the figure), which in turn implement the command, as determined by the cognition.

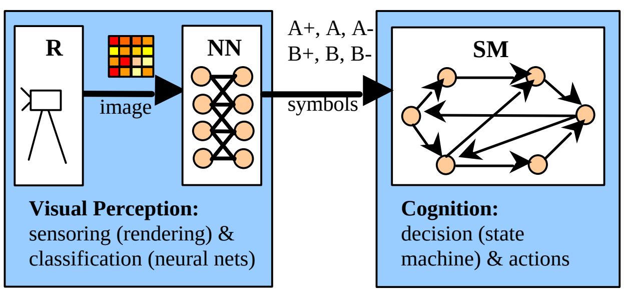


**Figure 1. WOXBOT Character Framework**

In the WOXBOT project the input and output modules were designed to perform their tasks, and no adjustment has been carried out after any creature has become alive. The vision system was implemented by a virtual camera followed by a neural network, trained in advance to correctly classify the scene elements, in order to provide right information about it to the cognitive module of each virtual creature, a robot.

The cognitive module has been conceived as a state machine, represented as a bit string that as such could evolve through generations. Therefore the state machine can be adaptively adjusted to perform actions in a more convenient way. The state change in the machine was controlled by the inputs passed by the sensors. Each state could be associated with an action.

The initial state machines, corresponding to those of the first generation of characters from specie, were randomly produced. Therefore the behavior of these characters was not expected to be appropriate. But a statistical dispersion of behaviors made some more consequent then others. Periodically some of these characters have been selected to reproduce based on their measured fitness. This process of reproduction, combining crossover and mutation, allowed an improvement in collective behavior after some generations [Fig. 2]. But no learning ability was present in these virtual beings. Our proposal in this paper is to introduce this skill.

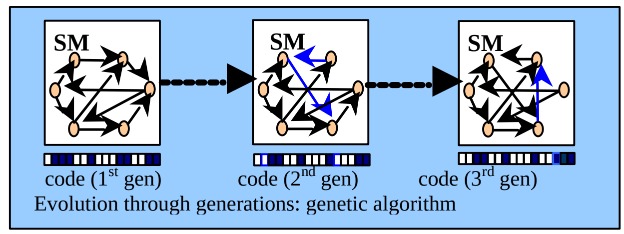


**Figure 2. WOXBOT Evolutionary Model.**

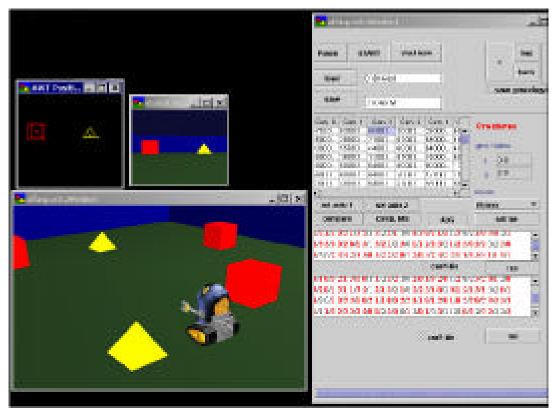

**Figure 3. WOXBOT Simulation**

An overview of the WOXBOT simulation process, showing the environment and some results, is presented below [Fig. 3]. It presents the robot in the arena, and shows also its own view of the scene, which is used to take the decisions. Different genotypes have been produced after some generations, showing a diversity of strategies that can be considered to be well adjusted to the survival proposal.

## 3. CHARACTER FRAMEWORK

The new character has two main differences if compared with WOXBOT. Its visual perception consists of an adaptively adjusted classifier based on fuzzy logic, replacing the original neural network. Its cognitive model is represented by a simple language structure, instead of a state machine. This paper focuses on the second aspect.

The cognition of each character is implemented as a simple language interpreter, which periodically selects one sentence to be executed. This selection is controlled by the inputs from the visual system, which passes symbols describing the recognition and positioning of scene elements. The selection of the sentence to be executed considers also the history of success or failure when executing that sentence in the same case (input).

Before presenting this model in details we describe its structure, considering components and how they are combined in more elaborated structures.

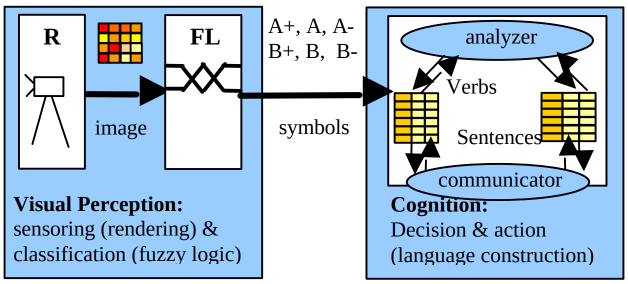


**Figure 4. ALGA Character Framework.**

### 3.1. Perception

The perception module is based on a fuzzy logic classifier that takes as input a bitmap image of the scene, as seen by the character, and classifies its objects. We intend to add an adaptive module to control the features of this fuzzy logic classifier, in order to observe if the character can improve its ability to classify scene objects.

### 3.2. Cognition

The cognition is based on an analyzer that takes a set of words (representing single actions) and combines them in small sentences. The level of success or failure associated to the execution each sentence is kept internally, and represents the characters own knowledge. This information is used to assist the selection process for new sentences.

A communicator is present in this framework to show the possibility of knowledge exchange between characters, allowing some characters to teach others. The language model, described ahead, gives more details on this structure and functionality.

#### *3.2.1. Sentences Constructor*

The sentences constructor is responsible for analyzing the repertoire of sentences in order to propose the construction of new ones, using its vocabulary. The vocabulary is in part instinctive (the character is born with a small set of words) and in part may be extended during its existence. In order to construct new sentences the constructor combines existing sentences, considering their level of success.

Every character is born with a small vocabulary, which corresponds also to all known sentences at the beginning of its life. New vocabularies and sentences may be acquired when talking to other characters.

New sentences may be acquired in two forms: a) listening to speeches from colleagues and considering their suggestions (reflecting how appropriate they are), or b) acquired by an internal inspection conducted by the constructor, which in some cases proposes new sentences, combining other already existent. This concept emulates a type of reasoning or self-reflection by the character.

#### *3.2.2. Sentences Executor*

The sentences executor is periodically required to take one sentence and to execute it, leading to a character action. The executor makes some measurements on internal character variables (states) in order to observe how they change, and based on them modifies the history of success or failure associated to that sentence in conjunction with the circumstance where it happened. A typical internal variable is the stored energy of the character. The execution of a sentence can pertain to one of three classes, listed bellow:

*Action Sentences*: These sentences lead to an action as gathering, touching, bringing, etc… When executing bad actions (inside a context) the character is punished, while good actions rewards it.

*Speech Sentences*: Speech sentences determine the production of a small talk (normally just a sentence accompanied by its context). For instance a valid sentence here could be:

Context: If ball is close    Sentence: then catch the ball.

*Movement Sentences*: Movement sentences control the character motor system with actions as: step ahead, step back, turn right, turn left or just stay.

## 4. LANGUAGE MODEL

The language model presented here has some components (tables), composition rules and analyzers / constructors to create new sentences and to execute them. These parts are presented bellow.

### 4.1. Language Components

Two tables are used to keep the known vocabulary and repertoire of sentences of each character.

#### *4.1.1. Dictionary / Vocabulary*

Every character has a small vocabulary, which basically consists of actions that it can perform. This vocabulary is a subset of the complete vocabulary that comprehends all words known by all individuals.

A character is born with a fixed small set of words, representing those actions that can be associated with an instinctive procedure. During its life it can learn new words, based on its own personal experience. An example of a dictionary is presented below [Table 1]. Two possible situations are considered here: learning by doing and learning by talking, which are better described ahead. Every word has two indices: one showing how frequently it appears in sentences, and another saying how important they are. The importance is calculated analyzing the effectiveness of the correspondent action.

**Table 1. Dictionary**

| Word | Frequency of use | Importance |
|---|---|---|
| A | 10% | High |
| B | 30% | Low |
| C | 40% | Medium |
| C | 20% | Medium |
| D | 50% | High |
| instinctive vocabulary (genetically coded) – basic actions | | |
| vocabulary acquired by experience (not genetically coded) | | |

*4.1.2. Sentences Book*

Every character has also a table where it keeps its acquired knowledge, expressed in simple language structures, composed by a set of few words. One example could be: *step ahead – turn right – catch something*, where each of these 3 actions corresponds to one word of its vocabulary. The sentences book [Table 2] keeps all sentences from one character, from single to multiple word sentences. Although one sentence relates to just one situation (input), the same situation can be associated with multiple sentences. In this case a Monte Carlo method selects the sentence, based on the distribution of their probability of occurrence and on their importance (fitness).

**Table 2. Sentences Book**

| Sentences | Inputs | Probability | Fitness |
|---|---|---|---|
| A | x | 100% | High |
| B | y | 100% | Medium |
| C | z | 100% | High |
| AB | X + y | 50% | High |
| BCA | z + w + | 70% | Low |
| CA | X * w | 20% | Low |
| ADC | x | 80% | Medium |
| DDB | x*y + z | 10% | Low |
| instinctive sentences (genetically coded) – basic actions | | | |
| sentences acquired by experience, on self-reflexive analysis (a type of reasoning) or on talks (not genetically coded) | | | |

When a new generation of characters is computed, the genetically coded part of these tables are combined, producing a new one. Herewith we allow the perpetuation of basic (instinctive) skills through generations. The non-genetically coded lines remain empty, and will be filled by experience, acquired during life process.

### 4.2. Language Analysis and Composition

The process of analysis and composition is responsible for continuously evaluating the history of all valid sentences, in order to classify them according to its convenience of being selected for each input. Furthermore, the composition is responsible for the proposition of new sentences, based on expected results, which can be foreseen considering the history of other known sentences.

These processes consider both, the local vocabulary and sentence book. The usage of the vocabulary allows trials with those words that have not yet been inserted into sentences. But the analysis of existing sentences, and their combination into new ones allows the exploitation of more complex sequences of actions. The system acts in response its own experimentations. Therefore if a new sentence seems to be inappropriate, it will probably receive a low importance score, and may be either rarely used, or even excluded from the sentence book. By the other side good sentences tend to receive high importance scores, and therefore will probably be selected more frequently than other with lower scores.

### 4.3. Execution of Language Statements

The execution of language statements, or sentences, is performed by a process that considers the importance conferred to them in different situations, as well as their probability of occurrence, based on their history. Therefore both aspects are relevant to the decision. Those sentences associated to higher frequency of occurrence represent a more conservative character behavior, while those with higher importance may represent more dared character behavior.

The framework permits the usage of a variable to describe the predominant behavior, or the character personality.

### 4.4. Evolving Strategies from Combination of Sentences

Our goal is to implement a system, formed by the analyzer and composer, able to propose sentences where strategies are present. Since we have implemented this system as an open framework, more sophisticated models based on artificial intelligence may be added to it. We believe that AI may provide fruitful results for the achievement of intelligent strategies, but we are also considering the usage of simple schemas based on genetic algorithms to evolve these sentences.

## 5. LEARNING SKILLS

The learning skills of our character are two fold: communication and self-analysis abilities.

### 5.1. Communicating Knowledge

The first learning skill results from the possibility of information exchange between characters, which have been implemented with an inter-character communication mechanism. This mechanism allows the exchange of symbols accompanied by statistics associated to it. Herewith we allow a character, wanting to cooperate with another one, to tell some sentences to this other one, which are from its point of view, convenient at this certain circumstance. As a consequence we can identify learning by talking skill.

*Learning by Talking*: The proximity of two characters may induce their conversation, which in fact is the transfer of a single word or sentence involved by its semantics, from one to the other. For instance if both recognize that something is good for them one can teach the other one what to do in that case, based on its own previous experience. Two situations are foreseen here. The character may want to cooperate, teaching the right action (word or sentence), or it can defect, telling the other wrong words or sentences.

This pre-disposition to tell the truth always, sometimes or never may be represented in some intern personality feature of the character, and may be part of those coded genetically or not.

### 5.2. Self-Analysis (Adaptive Learning)

The cognitive module has been conceived as a set of tables containing language information (words, sentences and the history of success / failure associated to them) and an analyzer, which selects one sentence to be executed. For this selection all sentences related to the current input are considered, but weighted by their importance, which in turn is obtained from its history and rate of success.

As presented in section 4, new sentences may be proposed, or removed from the sentences book. Here we can identify the learning by reasoning skill:

*Learning by Doing:* In this case some actions may lead to a situation where the repertoire of the character is increased, or the relative importance of its elements change, by the assumption of a new word or sentence that has a special meaning at that circumstance.

## 6. CONCLUSION AND FURTHER WORK

Up to now we have conceived the ALGA framework and are working on its implementation. This new model has been proposed based on our previous experience with the WOXBOT model. In ALGA we are adding learning skills to our characters, based on their ability to compose new sentences and verify how appropriate they are to conduct more powerful actions (those that increases some scores, stored energy for instance, for the character) more frequently. Furthermore we have proposed a communication structure that will allow characters to exchange part of their own knowledge (words or sentences associated with the character perception and its importance in certain circumstances). The scenario used for the current project is an aquarium, inhabited by fishes of different species. We expect to be able to identify both competition between different species and cooperation between fishes of the same family.

Further work will first analyze a large set of experiments conducted in a scene populated by these characters. The following steps will be the use of genetic algorithms to produce new sentences, and to evolve the fuzzy logic classifier. We intend to use this model to simulate diverse social behaviors. Members of our cognitive sciences group are interested to use this framework to study global economical and sociological phenomena.

## ACKNOWLEDGMENTS

Our thanks to CNPq for the partial support granted to this project including a master scholarship.

## REFERENCES

[1] Adami C., “Introduction to Artificial Life”, Springer Verlag, 1998

[2] Bonabeau E. W. and Therulaz G., “Why do we need artificial life?” in Artificial Life an Overview, Langton C. G. (editor), MIT Press, 1995

[3] Fogel D. B., “Evolutionary Computation, Toward new philosophy of machine intelligence”, IEEE Press, 1995

[4] Kaneko, K. “The Coupled Map Lattice: Introduction, Phenomenology, Lyapunov Analysis, Thermo-dynamics, in Theory and Applications”, John Wiley Sons, 1993.

[5] Terzopoulos D. (org.), “Artificial Life for Graphics in Animation, Multimedia and Virtual Reality”, in ACM/SIGGRAPH 98 Course Notes 22, 1998

[6] Phillips C. and Badler N. I., “Jack: A toolkit for manipulating articulated figures” in ACM/SIGGRAPH Symposium on User Interface Software, Banff, Canada, 1988

[7] Sims K., “Evolving 3D Morphology and Behavior by Competition”, Artificial Life, 1994

[8] Netto M. L. and Kogler Jr., J. E. (eds.) “Artificial life: Towards New Generation of Computer Animation”. Computers & Graphics. 25/6 Amsterdam, Holland, Elsevier Science 2001

[9] Miranda, F. R. et al “An Artificial Life Approach for the Animation of Cognitive Characters”. Computers & Graphics. 25/6 Amsterdam, Holland, Elsevier Science 2001

Evolutionary multi-agent systems. Published online by Cambridge University Press: 25 March 2015.Â  Solving optimisation problems with evolutionary algorithms requires choosing the encoding of solutions according to the search space definition, crossover and mutation operators appropriate for the encoding, as well as configuring a selection mechanism, or other components of specialised techniques like migration strategies for the island model of PEA (BÃ¤ck et al., Reference BÃ¤ck, Fogel and Michalewicz1997).Â  3.2 Evolutionary multi-agent system with immunological selection. The main idea of applying immunological inspirations to speeding up the process of selection in EMAS is based on the assumption that â€˜badâ€™ phenotypes come from â€˜badâ€™ genotypes. We've discovered that evolution strategies (ES), an optimization technique that's been known for decades, rivals the performance of standard reinforcement learning (RL) techniques on modern RL benchmarks (e.g. Atari/MuJoCo), while overcoming many of RL's inconveniences. In particular, ES is simpler to implement (there is no need for backpropagation), it is easier to scale in a distributed setting, it does not suffer in settings with sparse rewards, and has fewer hyperparameters.Â  To describe the behavior of the agent, we define a policy function (the brain of the agent), which computes how the agent should act in any given situation. Accordingly, practical negotiation systems should be empowered by effective learning mechanisms to acquire dynamic domain knowledge from the possibly changing negotiation contexts. This article illustrates our adaptive negotiation agents, which are underpinned by robust evolutionary learning mechanisms to deal with complex and dynamic negotiation contexts. Our experimental results show that GAâ€based adaptive negotiation agents outperform a theoretically optimal negotiation mechanism that guarantees Pareto optimal.Â  Nirmal Choudhary, K. K. Bharadwaj, Evolutionary learning approach to multi-agent negotiation for group recommender systems, Multimedia Tools and Applications, 10.1007/s11042-018-6984-3, (2018). Crossref. View Evolutionary Strategy Research Papers on Academia.edu for free.Â  Learning to cooperate in multi-agent systems by combining Q-learning and evolutionary strategy. Save to Library. Download.Â  We present an evolutionary strategy based algorithm which generates a topology for a non-slicing floorplan, represented as a sequence pair. Our search procedure uses large population of solutions to simultaneously explore multiple more. We present an evolutionary strategy based algorithm which generates a topology for a non-slicing floorplan, represented as a sequence pair. Our search procedure uses large population of solutions to simultaneously explore multiple solution regions. Non-cooperative Multi-agent Systems with Exploring Agents. Multi-agent learning is a challenging problem in machine learning that h 05/25/2020 âˆ™ by Jalal Etesami, et al. âˆ™ 0 âˆ™ share. read it. Detection under One-Bit Messaging over Adaptive Networks.Â  . Of particular interests are systems where multiple agents learn concurrently and independently by interacting with each other. This multiâ€“agent learning problem has attracted a great deal of attention due to number of important applications. Among existing approaches, multiâ€“agent reinforcement learning (MARL) algorithms have become increasingly popular due to their generality. Hu98multiagentreinforcement ; Bowling02multiagentlearning ; Claus1998 ; Peshkin00learningto .