

# Bio-inspired Cognitive Architecture for Adaptive Agents based on an Evolutionary Approach

Oscar Javier Romero López  
Universidad Politécnica de Madrid  
Boadilla del Monte 28660, Madrid  
+34 648777757  
ojrlopez@hotmail.com

Angélica de Antonio Jiménez  
Universidad Politécnica de Madrid  
Boadilla del Monte 28660, Madrid  
+34 913366925  
angelica@fi.upm.es

## ABSTRACT

In this work, an hybrid, self-configurable, multilayered and evolutionary subsumption architecture for cognitive agents is developed. Each layer of the multilayered architecture is modeled by one different Reinforcement Machine Learning System (RMLS) based on bio-inspired techniques. In this research an evolutionary mechanism based on Gene Expression Programming to self-configure the behaviour arbitration between layers is suggested. In addition, a co-evolutionary mechanism to evolve behaviours in an independent and parallel fashion is used too. The proposed approach was tested in an animat environment (artificial life) using a multi-agent platform and it exhibited several learning capabilities and emergent properties for self-configuring internal agent's architecture.

## Categories and Subject Descriptors

I.2.6 [Artificial Intelligence]: Learning – *Connectionism and neural nets, Induction, Knowledge acquisition, Parameter learning.*

## General Terms

Algorithms

## Keywords

Gene Expression Programming, Artificial Immune Systems, Extended Classifier Systems, Connectionist Q-Learning, Subsumption Architecture, Hybrid Behaviour Co-evolution, Cognitive Science.

## 1. INTRODUCTION

Recently, Cognitive Architectures have been an area of study that collects disciplines as artificial intelligence, cognitive science, psychology and more, to determine necessary, sufficient and optimal distribution of resources for the development of agents exhibiting emergent intelligence. One of the most referenced is the Subsumption Architecture proposed by Brooks [1].

According to Brooks [1], the Subsumption Architecture is built in layers. Each layer gives the system a set of pre-wired behaviours,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*Conference '04*, Month 1–2, 2008, City, State, Country.  
Copyright 2008 ACM 1-58113-000-0/00/0004...\$5.00.

where the higher levels build upon the lower levels to create more complex behaviours: The behaviour of the system as a whole is the result of many interacting simple behaviours. Another characteristic is its lack of a world model, which means that its responses are always and only reflexive as proposed by Brooks.

However, Subsumption Architecture results in a tight coupling of perception and action, producing high reactivity. Poor adaptability to new environments, no learning capabilities, no internal representation and the need of all patterns of behaviours must be pre-wired, are some weaknesses of the Subsumption theory.

Several extensions have attempted to add representation and behaviour arbitration to Subsumption like Behavior-Based Control Architecture [2] and Hormonal Activation Systems [3], but pre-wired behaviours and non-learning characteristics still remain becoming the architecture applicable and restricted only for a specific pre-configured environments.

The present research focuses on developing an Hybrid Multilayered Architecture for Cognitive Agents based on Subsumption theory. Additionally this work proposes an Evolutionary Model which allows the Agent to self-configure and evolve its arbitration of processing layers through the definition of processes (like inhibition, suppression and aggregation), kind of behaviours and number of layers. That means each agent instead of having a pre-configured structure of layers and processes it will have an Artificial Evolutionary Process which is responsible for defining the multilayered structure. On the other hand, instead of using an Augmented Finite Machine System as Subsumption theory states in [3] where no internal representation is done, in this paper we propose that each behaviour layer is driven by a different bio-inspired reinforcement machine learning system RMLS (chosen from a repertoire where behaviour co-evolution occurs) which learns from the environment and generates an internal world-model by means of an unsupervised and reinforced learning. The RMLSs used in the approach are: Extended Classifier System XCS [5], Learning Classifier System LCS [6], Artificial Immune System AIS [7], [8] and Neuro Connectionist Q-Learnig System NQL [9], [10].

The remainder of the paper is organized as follows. The description of the approach proposed is detailed in Section 2. Section 3 outlines and discusses the experimental results and emergent properties obtained. Finally concluding remarks are shown in Section 4.

## 2. PROPOSED HYBRID, SELF-CONFIGURABLE AND EVOLUTIONARY MODEL

In order to design an hybrid, self-configurable, scalable, and evolutionary architecture for cognitive systems which exhibits emergent behaviours and learning capabilities, the proposed work is exposed as follows.

Consider a virtual environment where there are several cognitive agents interacting with each others using a typical Subsumption Architecture. Some mayor constraints arise:

- o Environmental conditions changing continuously.
- o The number of behaviours inside each agent is variable.

- o Arbitration of behaviours is pre-wired and not depends on agent’s motivational states.

- o The cognitive agent can inhibit or suppress behaviours “only” if an applicability predicate is preestablished and new environment changes are not considered.

- o Agent’s behaviours do not generate a model of the world, do not couple with the environment via the agent’s sensors and actuators, do not learn about its own interaction with the environment and do not evolve the internal state of the behaviour.

These constraints address the following proposed approach of an hybrid, self-configurable and bio-inspired architecture for cognitive agents, depicted in Figure 1.

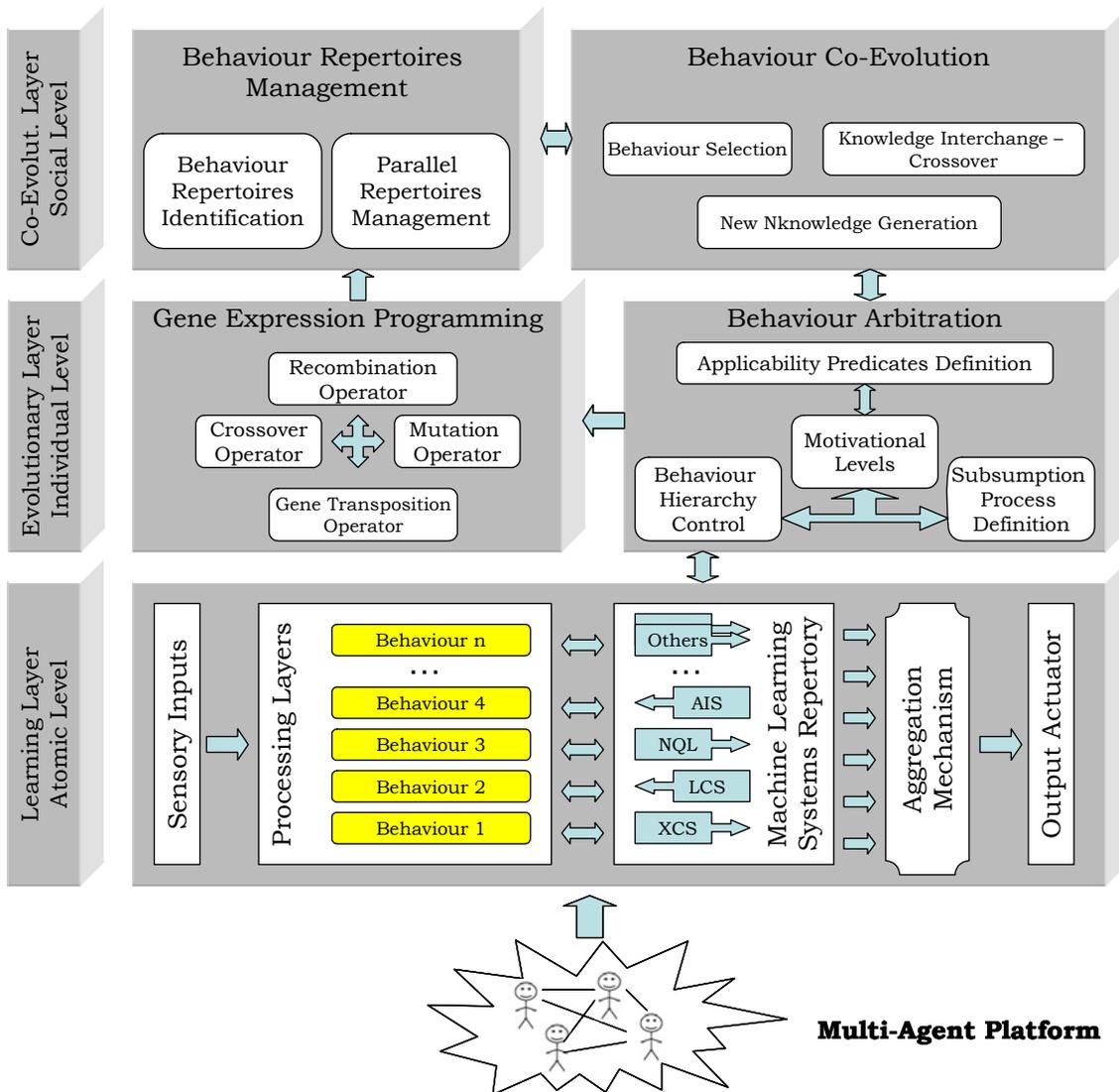


Figure 1. Hybrid and Evolutionary Architecture for Cognitive Agents.

The Figure 1 shows an hybrid architecture from which all the constraints mentioned before can be solved. An internal architecture based on subsumption principles but with few variations can be observed in every agent:

- Each processing layer is connected randomly with a different learning machine system (XCS, LCS, AIS, NQL, and scalable to others) which replaces the typical AFSMs proposed by Brook’s architecture in [1].
- After being trained, each agent’s behaviour is sent to a behaviour repertoire according to its type, where a co-evolutionary mechanism is applied so that every behaviour not only will learn in a local way inside of each agent but also will evolve in a global way, to be selected afterwards by another agent in the next generation.
- There is an evolutionary process driven by a Gene Expression Programming Algorithm GEP [12], which is in charge of self-configuring the agent (defining the number of layers, the behaviours that the agent will use, the connections and hierarchies between them -inhibit, suppress, aggregate-, the applicability predicates that determine which behaviour is activated at a certain situation and an activation time controlled by a timer).

## 2.1 Hybrid Learning Layer: Behaviours driven by different Machine Learning Systems

Every behaviour layer in the multilayered architecture will be associated to a Reinforcement Machine Learning System RMLS, that allows the architecture being hybrid and not only reactive since each behaviour will be able to exert deliberative processes using the acquired knowledge. Besides, this mechanism gives plasticity to the architecture because every behaviour “learns” in an unsupervised, independent and parallel way through its interaction with the environment, generating internal representations, rules and both specific and generalized knowledge. This mechanism is favored by the RMLSs characteristics: robustness, fault tolerance, use of bio-inspired techniques, adaptability and they do not require a previous definition of knowledge (unsupervised learning).

There are two principles formulated by Stone [13] that have motivated the proposed layered learning approach:

- Layered learning is designed for domains that are too complex for learning a mapping directly from an agent’s sensory inputs to its actuator outputs. Instead the layered learning approach consists of breaking a problem down into several behavioral layers and using RMLSs at each level. Layered learning uses a bottom up incremental approach to hierarchical task decomposition.
- RMLS is used as a central part of layered learning to exploit data in order to train and or adapt the overall system. RMLS is useful for training behaviors that are difficult to fine-tune manually.

The sensory inputs of each RMLS read the objects sensed around the agent while the actuator outputs indicate actions that the agent must to execute on the environment.

Accordingly, a common interface for all RMLSs (XCS, AIS, NQL, LCS, etc.) is proposed so although each RMLS has a

different internal process, they all have a similar structure that lets the system to be scalable introducing new RMLSs if is required and connecting them in an easy way with each behaviour layer in the agent’s multilayered architecture, as depicted in Figure 2.

Each RMLS has its advantages and disadvantages. However, no one RMLS is always better than others, so it is difficult to determine a good RMLS to drive each behaviour. On the other hand, the Cognitive Agent determines which behaviors must be inhibited or suppressed in a specific situation, but not always just one behaviour remains activated after inhibition, sometimes the Agent can require several behaviours to be activated concurrently. In order to merge the outputs of these behaviours activated in just one output, we propose an Aggregation Mechanism similar to proposed by Jiang in [11] but using different RMLS’s. This mechanism is based on Borda Counting Method. Aggregation of RMLS can improve the learning qualities as a whole because they can share some knowledge and utilize the strengths of the others to alleviate individual weaknesses.

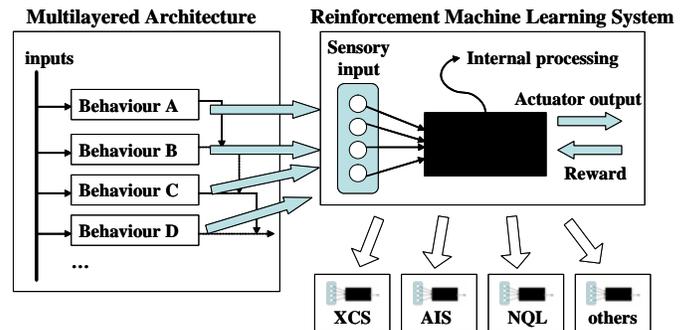


Figure 2. Multilayered Architecture connecting with RMLS interface.

## 2.2 Evolutionary Layer: Behaviour Arbitration

If each agent has an arbitrary behaviour set, how to determine: the interaction between them, the hierarchy levels, the Subsumption process (inhibition and suppression) and the necessary layers to do an adequate processing? These questions are solved next.

The internal multilayered structure of each agent is decomposed in atomic components which can be estimated and used to find the optimal organization of behaviors during the agent’s lifetime [4]. The main goal is that the agent in an automatic way self-configures its own behaviours structure. The model proposed by Ferreira in [12] called Gene Expression Programming GEP is used to evolve internal structures of each agent and generate a valid arbitration of behaviours.

GEP uses two sets: a function set and a terminal set. The proposed function set is: AND, OR, NOT, IFMATCH, INHIBIT, SUPPRESS and AGGREGATE. The AND, OR and NOT functions are logic operators used to group and exclude subsets of elements. The conditional function IFMATCH is an applicability predicate that matches with a specific problem situation. This function has five arguments; the first four

arguments belong to the rule's antecedent: they all indicate motivational levels in the agent (internal states, moods, etc.), for instance: energy level, bravery/cowardice level, hunger/thirstiness level, etc. If the first four arguments are applicable then the fifth argument, the rule's consequent, is executed. The fifth argument should be a INHIBIT/SUPPRESS/AGGREGATE function, or maybe and AND/OR function if more elements are necessary. The INHIBIT, SUPPRESS and AGGREGATE functions have two arguments (behaviourA, behaviourB) and indicate that behaviourA inhibits/suppresses/aggregate behaviourB.

On the other hand, the terminal set is composed by the behaviour set and the motivational levels set. Additionally "do not care" elements are included so whichever behaviour or motivational levels can be referenced. Behaviour Arbitration is driven by Agent's Motivational Levels which try to simulate moods or humor states in the Cognitive Agent. These moods are changing continuously whereas the Agent interacts with the environment.

Each agent has a chromosome with information about its self structure, e.g. Agent A can have a chromosome as: [{IFMATCH}, {ml1}, {ml2}, {ml3}, {ml4}, {INHIBIT}, {behaviour1}, {AND}, {behaviour2}, {behaviour3}], and this chromosome is a valid rule because both the antecedent and the consequent of IFMATCH function match to each required argument type, where {ml} is the abbreviation for motivational level. The above chromosome translates in the following rule:

IFMATCH:

ml1, ml2, ml3, ml4

THEN:

behaviour1 INHIBIT behaviour2 AND behaviour3.

Analyzing this rule we can infer that the agent has three behaviour layers: behaviour1, behaviour2, and behaviour3, and the two last ones are inhibited by the first one when agent has the motivational levels ml1, ml2, ml3, ml4. However, these chromosomes (applicability predicates) do not have always a valid syntax, so the GEP mechanism is used to evolve the chromosome until it becomes in a valid syntactic rule.

Each individual (agent) has a multigenic chromosome, that means, each chromosome has a gene set where each gene is an applicability predicate like the example, so the agent has several rules (genes) as part of its genotype and each one is applied according to the situation that matching the rule antecedent. Each gene is become to a tree representation and then a genetic operator set is applied between genes of the same agent and genes of other agents as in [12]: selection, mutation, root transposition, gene transposition, two-point recombination and gene recombination, in order to evolve chromosomal information.

After certain number of evolutionary generations, valid and better adapted agent's configurations are generated. A roulette-wheel method is used to select individuals with most selection probability derived from its own fitness. Fitness represents how good interaction with environment during agent's lifetime was.

## 2.3 Behaviour Co-evolution Layer: evolving globally

A co-evolutionary mechanism is proposed to evolve each type of behavior separately in its own genetic pool. Most evolutionary approaches use a single population where evolution is performed; instead, the behaviours are discriminated in categories and make them evolve in separate behaviour pools without any interaction, as proposed in [14].

First, each agent defines a specific set of behaviours that builds its own multilayered structure. For each required agent's behaviour, a behaviour instance is chosen from the pool (this instance is connected with one RMLS). Subsequently each agent will interact with the environment and each agent's behaviour will learn a set of rules and generate an own knowledge base.

After certain period of time a co-evolutionary mechanism is activated. For each behaviour pool is applied a probabilistic selection method of behaviours where those behaviours that had the best performance (fitness) will have more probability to reproduce. Then, a crossover genetic operator is applied between each pair of selected behaviours: a portion of knowledge acquired by each agent's behaviour (through its RMLS) is selected and interchanged with the other one; this is like Heritage of Knowledge.

Finally, new random rules are generated until complete the maximum size of rules that behaviours can have in their own knowledge base, so a new pair of behaviors is created and left in the corresponding behaviour pool to be selected by an agent in the next generation.

## 2.4 Emergent Properties of the Architecture

Brooks postulates in his paper [3] the possibility that intelligence can emerge out of a set of simple, loosely coupled behaviours, and emergent properties arise (if at all) due to the complex dynamics of interactions among the simple behaviours and that this emergence is to a large extent accidental.

The proposed architecture articulates a behaviour set that learns about environmental conditions in an independent and parallel fashion, and on the other hand evolve inside a categorized pool. Each simple behavior can be applied to a subset of specific situations but not to the whole problem space, however the individual level interaction between behaviours (inside each agent) allows covering multiple subsets of problem states and some characteristics are generated: robustness, redundancy in acquired knowledge: fault tolerance and a big plasticity level, so emergent properties in the individual and inside of the society (Multiagent systems) appear. So, the emergent properties arise from three points of view in a bottom-up approach:

- Atomic Level: in each behaviour of the multilayered architecture, when the associated RMLS learns from the environment how to associate sensory inputs and actuator outputs, in an automate way.
- Individual Level: when the agent self-configures its internal structure (chromosome), hierarchy and arbitration of behaviours through an evolutionary process driven by GEP.

- o Social Level: when an hybrid behaviour co-evolution mechanism is applied to all agent's behaviours, so behaviours learn not only themselves via the RMLS associated but also cooperating with other agents and communicating the acquired knowledge between them.

It is important to notice that emergence in different levels, from atomic to social point of view, provokes an overall emergence of the system, where some kind of intelligence we hope to arise. The experimentation focused on discovering some emergent characteristics in the agents. Nevertheless, expected emergent properties can vary according to the environment and the behaviour set.

### 3. EXPERIMENTATION

In order to evaluate the proposed architecture, following aspects were considered in each level:

- o Learning convergence rate of each proposed systems: XCS, AIS, LCS and NQL.
- o Learning and evolution convergence rate of each behaviour pool.
- o Variation of success rate vs. number of genes in GEP
- o Syntactically well-formed gene convergence rate

About overall System:

- o Subsumption architectures obtained on individuals after n iterations and emergent properties identified.

An artificial life environment called Animat (animal + robot) described in [6] is proposed to test the experiments. The environment simulates virtual agents (prey-depredator model) competing for getting food and water, avoiding obstacles, hunting, escaping from depredators, etc. Each animat driven by an agent in the environment disposes a set of 10 proximity sensors (see Figure 3) simulating a limited sight sense. 8 sensors read a safe zone and 2 sensors read a danger zone (to avoid collisions) as proposed by Romero [8].

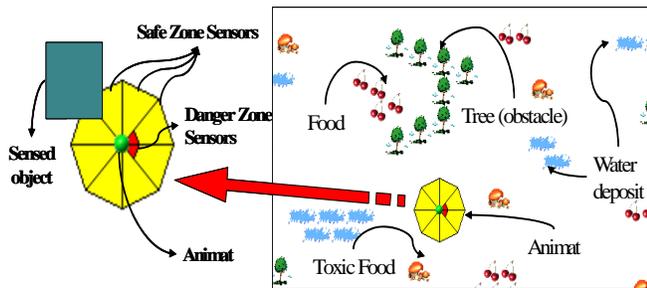


Figure 3. Animat Sensor distribution.

Thus, some experiments designed to evaluate the performance aspects mentioned above are described next.

#### 3.1 Learning convergence of each RMLS

In this experiment we chose an environment where the animat has to interact with using one different RMLS on a time. Table I shows the learning parameters used.

Table 1. Learning parameters of each RMLS

Parameter	XCS	AIS	NQL	LCS
Life Tax	-	0.005	-	0.005
Bid Tax	-	0.003	-	0.003
Cloning Rate x rule	1	4	-	1
Mutation Rate x rule	2	2	-	1
Similarity Threshold	-	0.8	-	-
Alpha $\alpha$	0.1	-	0.1	-
Beta $\beta$	0.2	-	-	-
Delta $\delta$	0.1	-	0.02	-
Gamma $\gamma$	0.95	-	0.8	-
Lamda $\lambda$	-	-	0.8	-
Layers in NN	-	-	5	-
Number of Epochs	50			
Nº runs x epoch	20			

Figure 4 shows a chart of the learning curve of the RMLSs: XCS, AIS, LCS, simple NQL and multilayered NQL.

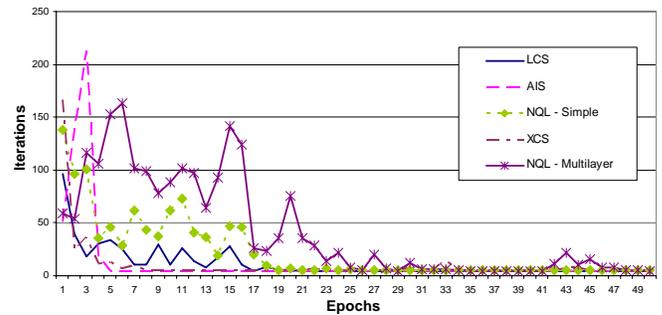


Figure 4. Learning Curve of each RMLS.

Figure 4 illustrates that AIS and NQL are more adaptive and robust than the others converging more quickly when changes in the learned environmental pattern are introduced. The peaks were registered because of changing patterns, but each RMLS adapted to new conditions quickly.

#### 3.2 Learning and evolution convergence of each behaviour pool.

The goal of this experiment is to examine if the fitness of every separate behaviour pool increments gradually until reaches a convergence point while evolution takes place. The experiment was carried out with the parameters on Table 2

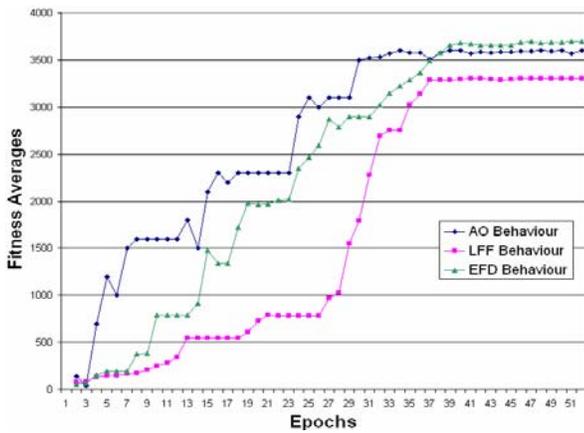
Three behaviour pools were selected for the experiment: Avoiding-obstacles, Looking-for-food and Escaping-From-Depredators, the results are depicted in Figure 5.

Figure 5 depicted some differences in each learning curve, due to environmental conditions, however the pools always tried to converge and reach certain stability in the same number of

epochs (approximately after 30 epochs), that means the evolution has been effective and each behaviour pool has established a coherent knowledge base getting a consensus between its own behaviour instances, about what the “behaviour category” should do.

**Table 2. Co-evolution Learning Parameters**

Parameters	Value
Epochs	50
N° runs x epoch	50
Crossover Prob.	0.7
Mutation Prob.	0.3
Mutation Rate $\eta$	0.85
Mutation Rate $\theta$	0.25
Mutation Rate $\kappa$	1.03
Mutation Rate $\gamma$	0.01

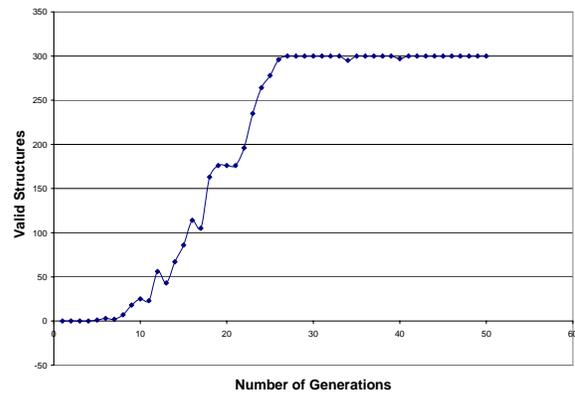


**Figure 5. Evolution convergence rate in 3 behaviour pools.**

### 3.3 Syntactically well-formed gene convergence

In this experiment, the progression of the number of syntactically well-formed structure (multigenic chromosomes) of each individual was analyzed. Figure 6 shows how the number of valid chromosomes increments whereas generations evolve through the time. The experiment was executed with a population of 300 individuals.

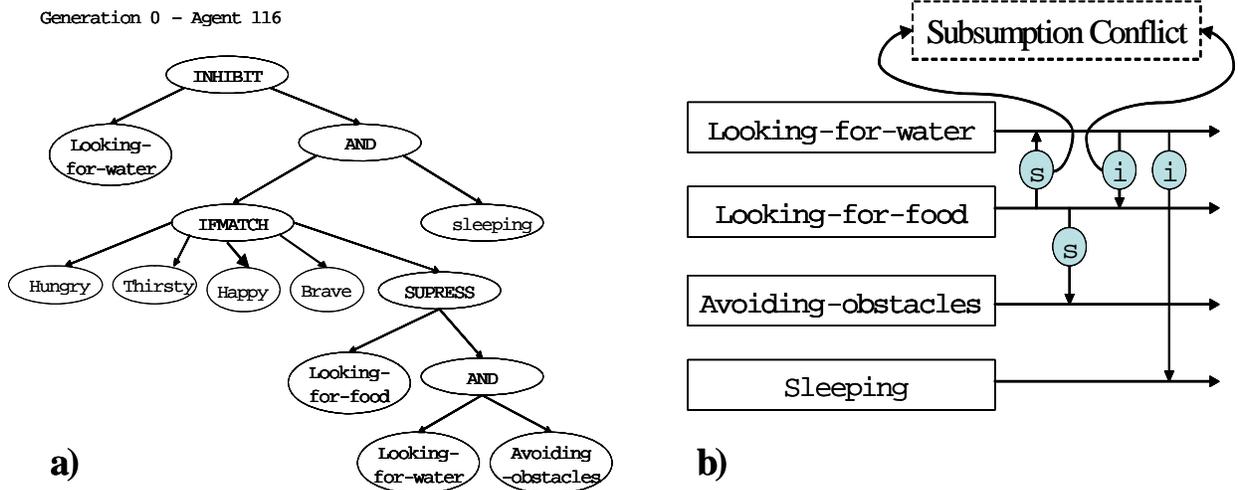
Figure 6 shows that a point of convergence (that means all chromosomes in population are valid) is given in the generation 27 approximately. Then, the system will need between 25 and 30 generations to evolve all individuals in the population.



**Figure 6. Valid Structures (chromosomes) through several Generations.**

### 3.4 Analysis of evolved architectures

Finally, after the whole system has evolved during a specific number of generations, we have analyzed the final structures of the best adapted agents where emergent properties arose.

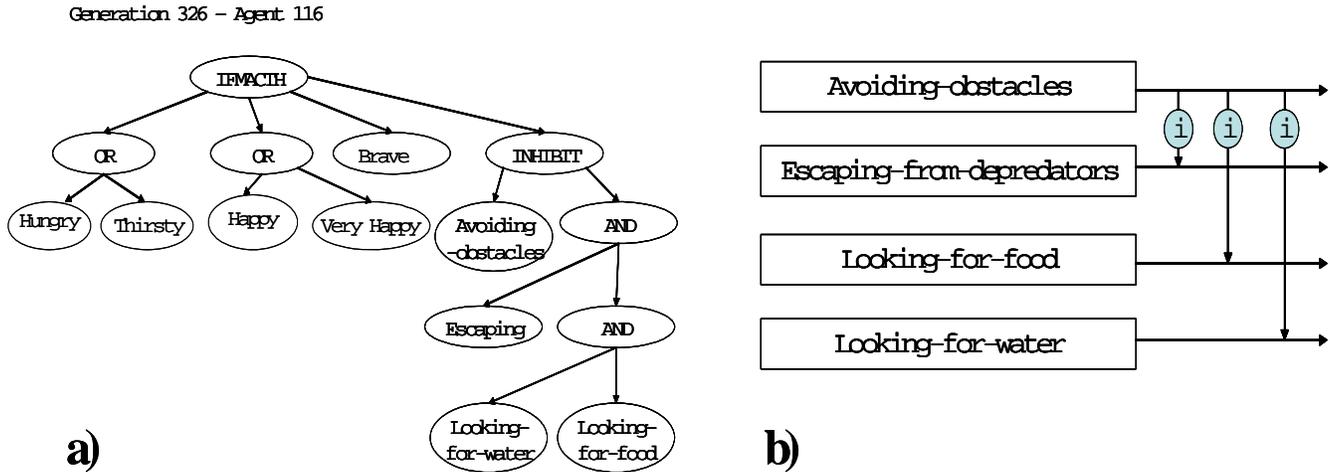


**Figure 7. Genotype and Phenotype of an initial Agent's Architecture.**

Figure 7 shows the genotype (Expression Trees ETs) and phenotype respectively of an initial architecture of a random agent without any evolutionary phase; in contrast, Figure 8 shows the genotype and phenotype respectively of the evolved architecture of the same agent.

In Figure 7 the chromosome represents four behaviours: looking-for-water LFW, looking-for-food LFF, avoiding-

obstacles AO and sleeping SL, where LFW inhibits LFF and SL and LFW suppresses AO, but there is a contradictory process when LFF tries to suppress LFW and LFF has been inhibited by LFW already. This is solved with the evolved architecture in Figure 8, which proposes a new structure adding escaping-from-predators EFD behaviour and excluding sleeping behaviour.



**Figure 8. Genotype and Phenotype of the Agent's Architecture after 326 evolutionary generations.**

As depicted in Figure 8, the initial contradictory inhibitory/suppressor processes in the agent's architecture are solved, and only hierarchical inhibitory processes are proposed by the evolved architecture. Furthermore, we can deduce too that evolved architecture has collected a set of specific behaviours becoming the agent to an animat with a prey identity.

It is important to notice in evolved architecture that EFD behaviour inhibits both LFF and LFW behaviours, but if the animat is escaping and its sensors read a "wall" or a "tree", then EFD behaviour is inhibited by AO behaviour until the obstacle is not in front of the animat anymore, and after that the animat continues its getaway, so we can say that emergent behaviour arises.

Finally, the experimentation demonstrate that specific parameter configurations in RMLSs, GEP and Co-evolutionary mechanism are required to reach certain robustness, adaptability and learning capacities in the overall system. Nevertheless, emergent properties did not arise every time or in a quick way, in several experiments animats died quickly and they could not learn to survive.

#### 4. CONCLUSIONS

The integration of multiple Reinforcement Machine Learning Systems in controlling the behaviours layers of an hybrid Subsumption Architecture approach, instead of using the typical Augmented Finite State Machines, have demonstrated important advantages in learning about the world of the agent, making

internal knowledge representations and adapting to environmental changes.

The evolutionary mechanisms used in this work, provided a plasticity feature allowing the agent to self-configure its own multilayered behaviour-based architecture; thus it can avoid creating exhaustive and extensive knowledge bases, pre-wired behaviour-based multilayered structures and pre-constrained environments. Instead of this, a cognitive agent using our architecture only needs to interact with an arbitrary environment to adapt to it and take decisions in a reactive and deliberative fashion.

In the experimentation, the emergent properties were difficult to discover because it takes a lot of time to evolve the overall system despite of using a multiagent platform in a distributed configuration. Maybe, it can be similar to the natural evolution where adaptation occurs slowly and sometimes produces poor adapted creatures.

In our future work we expect to continue working on designing more adaptive and self-configurable architectures, using fuzzy techniques in the RMLSs to improve the sensors readings and to manipulate motivational levels (moods). One concrete application of this research will be the development of a Cognitive Module for Emotive Pedagogical Agents where the agent will be able to self-learn about its own perspectives, believes, desires, intentions, emotions and perceptions.

## 5. ACKNOWLEDGMENTS

Supported by the Programme Alban, the European Union Programme of High Level Scholarships for Latin America, scholarship No. E05D056455CO”.

Diego Romero of Mechanical Engineering Department at National University of Colombia, who made numerous contributions to this research.

## 6. REFERENCES

- [1] R.A. Brooks, A Robust Layered Control System For A Mobile Robot, *IEEE Journal Of Robotics And Automation*, RA-2, 1986, 14-23.
- [2] M.J. Mataric, Behavior-based control: Main properties and implications, *Proceedings of the IEEE International Conference on Robotics and Automation*, Nice, Francia, 1992, 2-8.
- [3] R.A. Brooks, How to build complete creatures rather than isolated cognitive simulators, *Architectures for Intelligence*, 1991, 225-239.
- [4] J. R. Koza, Evolution of subsumption using genetic programming, *Proceedings of the First European Conference on Artificial Life*, Paris, 1992, 110-119.
- [5] M.V. Butz, T. Kovacs, S. Wilson, How {XCS} evolve accurate classifiers, *Proceedings of the Genetic and Evolutionary Computation Conference*, San Francisco, 2001, 927-934.
- [6] J.H. Holland, *Induction, Processes of Inference, Learning and Discovery*, (Mich:Addison-Wesley, 1953).
- [7] L. N. de Castro, J. Timmis, *Artificial Immune Systems: A New Computational Intelligence Approach*, (Ed. Springer, 2002.)
- [8] D. Romero, L. Niño, An Immune-based Multilayered Cognitive Model for Autonomous Navigation, *IEEE Congress on Evolutionary Computation*, Vancouver, 2006, 1115-1122.
- [9] C. Watkins, *Q-learning*, *Machine Learning* 8, Boston, 1992 – pp. 279-292.
- [10] V. Kuzmin, Connectionist Q-learning in Robot Control Task, *Proceedings of Riga Technical University*, 2002, 112-121.
- [11] J. Jiang, M.S. Kamel, Aggregation of Reinforcement Learning Algorithms, *IEEE Congress on Evolutionary Computation*, Vancouver, 2006, 68-72.
- [12] C. Ferreira, *Gene Expression Programming: A new adaptive algorithm for solving problems*, *Complex Systems*, forthcoming, 2001.
- [13] P.Stone, *Layered Learning in Multiagent Systems*, (Doctor Thesis CMU-CS-98-187, 1998)
- [14] A. Farahmand, Hybrid Behavior Co-evolution and Structure Learning in Behavior-based Systems, *IEEE Congress on Evolutionary Computation*, Vancouver, 2006, 979-986.