Evolutionary Virtual Agent

Jean-Claude Heudin
International Institute of Multimedia
Pôle Universitaire Léonard de Vinci
Jean-Claude.Heudin@devinci.fr

Abstract

This paper describes a virtual agent model for mobile devices based on an Artificial Life and Evolutionary approach. The model is based on a bio-inspired metabolism and a "digital DNA". It includes a 3D face with real-time animation, natural language interaction and web access. Behaviors are scripted using the Scheme language and Genetic Programming. Rather than trying to strictly copy reality, the model sustains believability by applying the ten key characteristic qualities that animate synthetic characters as proposed by Hayes-Roth and Doyle.

1. Introduction

Autonomous self-animated conversational characters have applications in various domains such as man-machine interface, computer game, and customer relationship management. A major part of the research has focused on graphical animation or natural language interaction but very few integrate all requirements [1]. In contrast, autonomous characters are supposed to respond to human interaction in real-time and with appropriate behaviors: not pre-determined, broad in content, highly contextual, communicative, behaviorally subtle, etc. [2]. The aim of our project is to design such an agent model for mobile devices such as PDA and mobile phones. The applications this research is intended for include games, user interfaces for online services, customer relationship management, etc.

In order to achieve this challenging goal, we ground our software model for virtual agents on Artificial Life and Evolutionary Programming (see for example [3, 4, 5]). However, rather than trying to imitate the biological reality in details, we study how to maintain believability and situations where the suspension of disbelief may be disrupted. We refer to results obtained by character animators for the belief of life that is not dependent on accurate simulations but on the viewer's ascription of emotion. The character must appear to think, make decision, and act of its own volition [6]. To that extent, we use the ten key characteristic qualities that animate characters should possess [7]. The ten qualities are: identity, backstory, appearance, content of speech, manner of speaking and gesturing, emotional dynamics, social interaction patterns, role, and role dynamics.

Other important constraints are the size and real-time performance of the resulting agent since today's mobile devices have limited screen size, memory size and processor power. This leads to a reduced and optimized software architecture and coding.

In this paper, we describe our Evolutionary Virtual Agent model and we give some preliminary results. In particular we show how these ten characteristics become salient during a typical interaction with EVA. We conclude by outlining future research work.

2. Agent model

2.1. Bio-inspired architecture

The Evolutionary Virtual Agent (EVA) model inherits from the LifeDrop project [8]. The two main differences are: (1) EVA focus on human-like characters while LifeDrop is limited to relatively small creatures called biomorphs, (2) EVA agents evolve within the system without any simulated environment while the ecosystem of LifeDrop is a simulated drop of water.

One of the most salient feature of living organisms, from an organizational point of view, is the distinction between the genotype and the phenotype. As suggested by C. Langton [9], we refer to generalized notions so that we may apply them in a non-biological situation. Gtype is the generalized genetic code and Ptype is the set of structures and behaviors that emerge out the interpretation of the Gtype in the environment. These two levels are linked together by two main processes: *development* and *reproduction* (cf. figure 1).

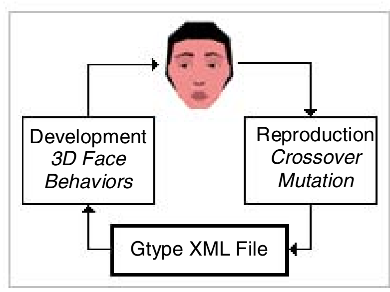


Figure 1. EVA evolution cycle.

The development process takes its parameters from the Gtype encoded as an XML file. It produces a layered architecture shown in figure 2. This includes: a reduced 3D real-time engine; a Body3D class that creates the face shape and all elementary expressions; a Metabolism class which includes all primary bio-inspired metabolic loops like energy, lifetime, reproduction, emotions; a Brain class responsible for all high-level behaviors like natural language interaction.

Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'04)
0-7695-2101-0/04 $ 20.00 IEEE

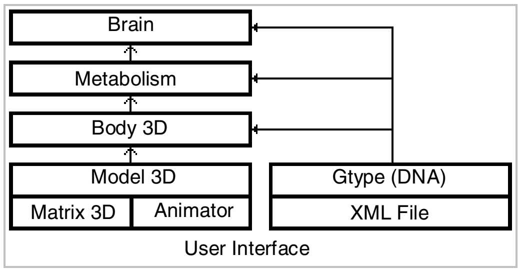


Figure 2. EVA Ptype architecture.

The second process is the reproduction phase. In this case, the Gtype file of a character is not used as a "parameter file" but as a "description file". This file is crossed with the one of another character and mutations are applied in order to create one or more new Gtypes. These new Gtypes produce new characters that inherit from their parents. Thus evolution occurs by selecting the best adapted in a given environment (application).

```
<vertices>
0,188.9,88.16 …
</vertices>
<polygons>
11,8,9,0 …
</polygons>
<move name="BROWS">
100,0,12.4,0 …
</move>
```

Figure 3. Sample syntax of the Gtype file.

2.2. Face design and expression animation

One important feature of the EVA model is its real-time emotional response to user interaction. The first EVA 3D face model is based on the work done by Ken Perlin with his Responsive Animated Character experiments [10].

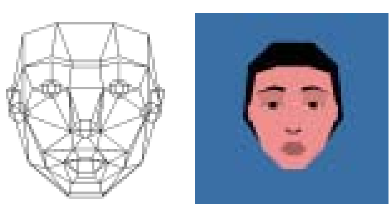


Figure 4. First face model and EVA rendering.

This initial model has been translated in our Gtype XML format. We implemented 12 basic key face animations (cf. tables 1 and 2). These animations represent the minimal number of facial expression elements that produce a convincing impression of character and personality.

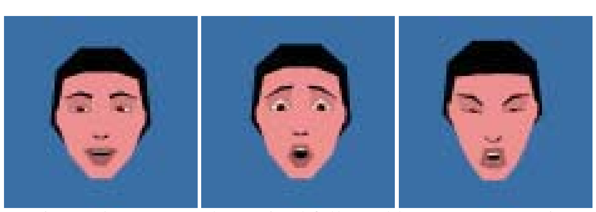


Figure 5. Happy and surprised followed by anger expressions.

We also use Perlin's noise at many levels for improving the rendering of movements: ranging from low-level animation triggering (eye blinking for example) to high-level attitudes that develop over time. All these noises are under the control of the Metabolism class.

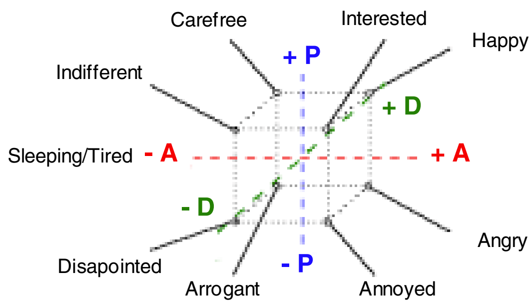
#	Key animation	+ action	normal	- action
1	Brows	up	mid	down
2	Eyes open	wide	open	squint/shut
3	Lids up	up	mid	down
4	Lids left-right	left	mid	right
5	Gaze up-down	up	mid	down
6	Mouth open	mmm	eh	ah/oh
7	Mouth wide	pucket	mid	wide
8	Mouth smile	smile	mid	frown
9	Mouth sneer	relax	mid	sneer
A	Head left-right	left	mid	right
B	Head up-down	up	mid	down
C	Head tilt	left	mid	right

Table 1. Key animations for EVA.

Expression	1	2	3	4	5	6	7	8	9	A	B	C
Happy			+			-	-	+				
Angry	-	+	+			-		+	+		-	
Annoyed	-	+				-			+		+	
Arrogant	-	+				-			+			
Disappointed	+		+			-						
Indifferent	+	+	+	+	+	-	-	+			-	
Carefree	+		+			-		+				
Interested	+	-	+			-		+				
Surprised	+	-	+									
Frightened	+	-	+							-	+	
Tired	+	+				-					-	
Sleeping		+				-					-	+

Table 2. The 12 key animations related to face expressions. +/- represents the sign of the coefficient applied to each animation.

The manifestation of emotions increases believability and creates the appearance of personality by using different intensities or responses to a given situation. Several emotion models have been applied to virtual characters like the OCC model [11]. We use the PAD Emotional State Model [12] consisting of three nearly independent dimensions that are used to describe and measure emotional states: *Pleasure-displeasure* distinguishes the positive-negative affective quality of emotional states; *Arousal-non-arousal* refers to a combination of physical activity and mental alertness; *Dominance-submissiveness* is defined in terms of control versus lack of control. The Metabolism class includes a PAD emotional engine linked with the facial expressions using a 3 dimensional mapping (cf. figure 6).



Surprised/Frightened = Y . max (| P |, | A |, | D |)

Figure 6. PAD model and facial expression mapping.

This emotion engine uses a set of parameters including thresholds, slow and fast transitions, sustain and decay times, default states, etc. These parameters are stored in the Gtype file like other parameters and subject to be modified through evolution.

2.3. Natural language interaction

The Brain class includes a natural language processing module consisting of three main components. The first one is a sentence parser extracting keywords from the textual input. This is done using a thesaurus and an "empty word" filter. The second component is a "categorizer" mapping each keyword to a conceptual category. These categories are defined and linked to the PAD emotion engine through XML files (cf. figure 8) and can be dynamically updated.

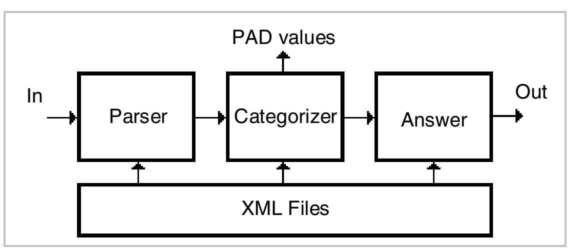


Figure 8. Natural Language Processing in EVA.

The third component is a reply generator that generates candidate phrases based on templates stored in XML files. Templates are paired with categories by simply using the same "name" parameter. The simplicity of XML templates and categories allows non-technical staff to author additions for a given purpose or application. Note that the facial animation includes also lip-syncing based on the Preston Phoneme Set.

```
<category name="HELLO" pleasure="0.2" arousal="0.2"
dominance="0.2">
hello
hi
howdy …
</category>
```

Figure 9. Syntax example for categories.

```
<template name="NONSENSE">
Is that a trick sentence?!?!
I do believe that you've gone quite mad.
What kind of sentence is that?
Have you always been this silly?
Hehehe. Trying to trick me, eh?
Sorry, my positronic brain is out of control. …
</template>
```

Figure 10. Syntax example for template files.

We use an integrated chat-like system for implementing a simple communication protocol between interactors and EVA. This enables multiple identified interlocutors over the network. Note that EVA can also make web accesses and call search engines in order to extract sentences containing query terms found during the flow of conversation.

2.4. Evolutionary approach

Mating is not the only evolutionary paradigm in the EVA model (cf. section 2.1). Most EVA software components like the natural language processing system can be called from a scripting language, thus allowing the design and evolution of various conversational behaviors. We choose Scheme because of its known qualities [13] and its adaptation for Genetic Programming.

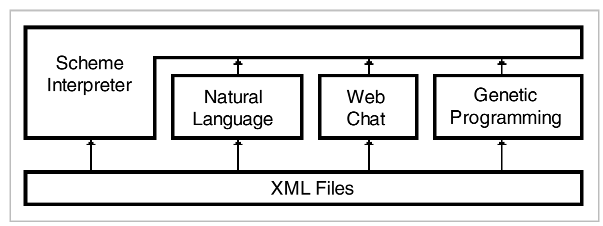


Figure 11. The "Brain" level architecture.

Therefore, the Brain class is composed of four software modules (cf. figure 11) including the natural language processing system, the web access and chat system, and the Scheme interpreter itself. We implemented a Genetic Programming engine based on the principles defined by John Koza [14] that enables the generation of non pre-determined behaviors.

```
( define random-behavior ( make-behavior ) )
( set-program! random-behavior
                       ( random-program
                                   *function-set*
                                   *argument-set*
                                   *terminal-set*
                                   *max-depth*
                                   #t ) )
( set-fitness! random-behavior min-fitness )
```

Figure 12. Example of a script creating a new random behavior.

The individual structures that undergo evolution and adaptation in our genetic programming system are "behaviors" composed of hierarchically structured Scheme programs and related fitness values. The set of possible behaviors is the set of all compositions of functions that can be composed recursively from the pre-determined sets of functions and terminals. These two sets are defined by the user and could include most EVA Brain methods and EVA parameters. The fitness function is scripted using Scheme and can use all EVA information such as the PAD values.

3. Preliminary Results

We have designed a first conversational agent based on the EVA model previously described. The face was designed using Discreet's 3DS Max software tool and a dedicated converter in order to generate a correct EVA Gtype file. Figures 13 and 14 show sample images from first drawing to final running agent.

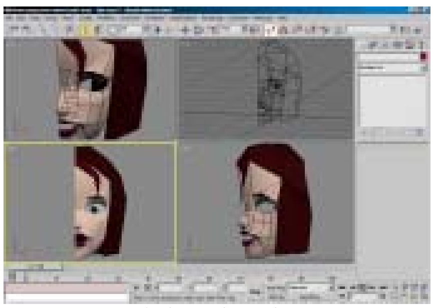


Figure 13. EVA 3D design using Discreet's 3DSMax tool.

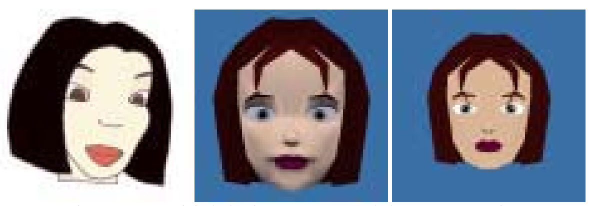


Figure 14. Initial design, 3DS model, EVA running.

All basic expression animations have been easily updated by adding subtle changes to the XML DNA file. Figure 15 shows the 12 key facial emotion expressions obtained when running the agent.

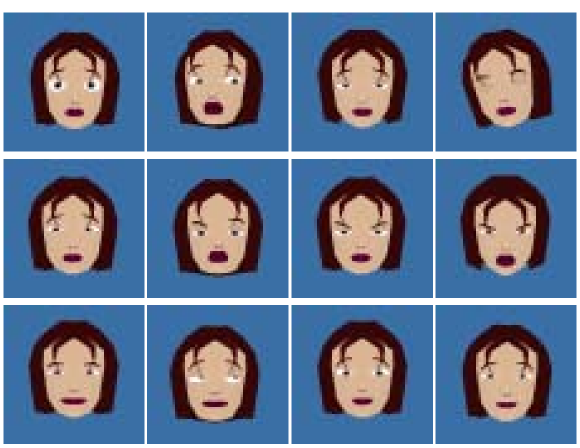


Figure 15. (1) Surprised, (2) frightened, (3) tired, (4) sleeping, (5) disappointed, (6) arrogant, (7) annoyed, (8) angry, (9) happy, (10), indifferent, (11) interested and (12) carefree.

We have conducted many text-based interaction experiments with this agent where a user types in comments and EVA reacts accordingly. For practical reasons, experiments were implemented on a portable computer with a networked chat interface (cf. figure 16).

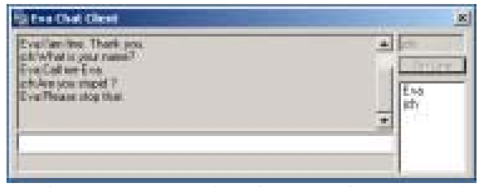
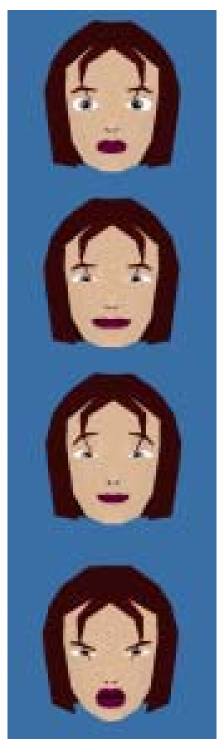


Figure 16. Networked chat interface used for the experiment.

Figure 17 features a sample of recorded interactions using several XML categories and templates like identity, backstory, appearance, content of speech, manner of speaking and gesturing, emotional dynamics, social interaction patterns, role, and role dynamics. These files have been designed in order to feet the ten key qualities that animate character should possess (cf. introduction). It also uses a set of generic template for most current types of interactions. This design phase is very important in order to give the agent patterns of behavioral, temperamental, emotional and mental traits that distinguish people from one another. In other words, it gives the agent a perceived personality resulting in convincing behaviors.



Eva: Hello jean-claude!
jch: How are you ?
Eva: As a matter of fact, I was just going to check myself.
jch: What is your real name ?
Eva: My full name is AI0130-H0001. But just call me Eva.
jch: What do you think ?
Eva: I know that I can't feel emotion like people but I can react to certain variables.
jch: Ok. Are you an AI?
Eva: So you're into electronic sheep too?
jch: Very funny. Do you know Eliza?
Eva: I have played around with Eliza a bit. She's not that good!
jch: Your answer is silly.
Eva: Why do you feel that you have to tell me that my answer is silly?
jch: Because it is a stupid answer!
Eva: Hey! I didn't come here to be subjected to bad language!
jch: Sorry Eva.
Eva: I prefer that.

Figure 17. A sample dialog between the EVA agent and a human.

Next table (table 3) gives a list of descriptions to highlight these ten key qualities. We emphasize that most of these qualities use multiple features of EVA.

Quality	EVA
Identity	EVA is an AI software agent.
Backstory	EVA is the first generation of AI EVA creature.
Appearance	EVA is presented as an expressive young brown haired woman.
Content of Speech	EVA enjoys discussing AI and Science Fiction.
Manner of Speaking And Gesturing	EVA has a sense of humor and often ask questions to learn more about humans.
Emotional Dynamics	EVA resents all derogatory comments about her and AI in general.
Social Interactions Patterns	EVA is intent on educating those that interact on the importance of AI creatures in the near future.
Role	EVA is an ambassador of AI creatures.
Role Dynamics	EVA wants humans to know more about the reality of virtual creatures. She reacts to and answers questions.

Table 3. The ten key perceived qualities of characters expressed in EVA during the experiment.

The next figure (figure 18) gives an example of the mutation method applied on the first EVA agent. The resulting agent has evident changes in colors but also more subtle changes in the face model. The following set of figures (figure 19) shows alternative character designs to emphasize the generality of the EVA model.

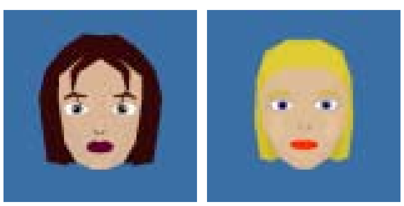


Figure 18. Mutation of the first EVA design.

Figure 19a. Neutral (left) and disappointed (right) expressions.

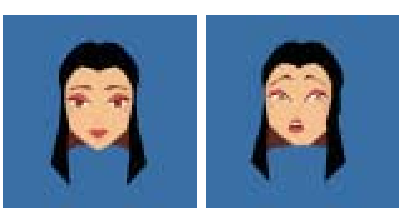


Figure 19b. Neutral (left) and frightened (right) expressions.

Figure 19c. Neutral (left) and arrogant (right) expressions.

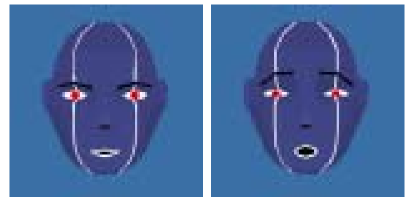


Figure 19d. Neutral (left) and surprised (right) expressions.

Figure 19e. Neutral (left) and angry (right) expressions.

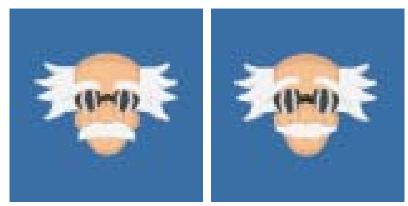


Figure 19f. Neutral (left) and happy (right) expressions.

4. Conclusion and future work

This paper has described a model of Evolutionary Virtual Agent for mobile devices and its first results in terms of the ten key qualities as proposed by Hayes-Roth and Doyle. Among all the work remaining to be done on this software platform, we are designing a Java MIDP version for a game project on high-end cell phones. Most of the research effort will focus on the evolutionary abilities of the EVA model. This includes: (1) evolution of new characters from a set of existing ones by mating and selection and (2) evolving conversational behaviors using the genetic programming engine. We are also interested in studying collective behaviors where multiple EVA agents interact together in order to solve a problem, like a complex search query on the Internet for example. The interaction between agents could be simply implemented since they can send and evaluate Scheme-expressions (which is a natural property of Scheme) through the chat system.

5. References

[1] Franklin, S., Graesser, A. 1997. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Proceedings*

IEEE COMPUTER SOCIETY

of the 3rd Int. Workshop on Agent Theories, Architectures and Languages, Springer-Verlag, 21-35.

[2] Badler, N., Allbeck, J., Byun, M. 2002. Representing and Parameterizing Agent Behaviors. In *Proceedings of Imagina*, 151-164.

[3] Heudin, J.C. (Ed.) 1998. *Virtual Worlds - Synthetic Universes, Digital Life and Complexity*. NECSI Series on Complexity, Perseus Books, Reading MA, 1-28.

[4] Sims, K., 1995. Evolving 3D Morphology and Behavior by Competition, *Artificial Life*, MIT Press (Cambridge), 1, 353.

[5] Terzopoulos, D., Xioyuan, T., Radek, G., 1995. Artificial Fishes: Autonomous Locomotion, Perception, Behavior, and Learning in a Simulated Physical World, *Artificial Life*, MIT Press (Cambridge), 1, 327.

[6] Thomas, F., Johnston, O. 1981. *The Illusion of Life: Disney Animation*. Hyperion Books, New York.

[7] Hayes-Roth, B., Doyle, P. 1998. Animate Characters. *Autonomous Agents and Multi-Agent Systems*, 1, 2:195-230.

[8] Metivier, M., Lattaud, C., Heudin, J.C. 2002. A Stress-based Speciation Model in LifeDrop, In *Proceedings of the 8th Int. Conf. on Artificial Life*, Sydney, MIT Press, 121-126.

[9] Langton, C.G., 1988. Artificial Life, *Artificial Life*, Addison-Wesley (Reading), 6, 1.

[10] Perlin, K. 1997. Layered Compositing of Facial Expression. ACM SIGGRAPH Technical Sketch.

[11] Ortony, A., Clore, G., Collins, A. 1988. *The Cognitive Structure of Emotions*, Cambridge University Press.

[12] Mehrabian, A. 1996. Pleasure-Arousal-Dominance: A general framework for describing and measuring individual difference in temperament. *Current Psychology: Developmental, Learning, Personality, Social*, 14, 261-292.

[13] Abelson, H., Sussman, G.J. 1966. *Structure and Interpretation of Computer Programs – 2nd Ed.* MIT Press, Cambridge MA.

[14] Koza, J.R. 1992. *Genetic Programming - On the programming of computers by means of natural selection*. MIT Press, Cambridge MA.

[15] Hayes-Roth, B., Maldonado, H., Moraes, M. 2002. Designing for Diversity: Multi-Cultural Characters for a Multi-Cultural World. In Proceedings of Imagina, 207-225.