Submitted to Game-On'02, London, United Kingdom, November 29-30, 2002.

LEARNING BY IMITATION OF BEHAVIORS FOR AUTONOMOUS AGENTS

Cédric Buche Marc Parenthoën Jacques Tisseau

Laboratoire d'Informatique Industrielle, ENIB
EA 2215 / Université de Bretagne Occidentale
Parvis Blaise Pascal, BP 30815, F-29608 Brest Cedex,
France
E-mail:{buche,parenthoen,tisseau}@enib.fr

KEYWORDS
Learning of behaviors, Animat, Autonomy, Imitation.

ABSTRACT

The goal of this work is to provide more autonomy for virtual actors by endowing them with a learning ability by imitation. While acting in his virtual world, our virtual actor uses prototypic behaviors defined by Fuzzy Cognitive Maps (FCMs) to simulate other actors' behavior in his imaginary world. This simulation allows him to carry out predictions and choices of strategies. We propose a method allowing virtual actor to adapt a prototypic behavior of FCMs to a model by simple observation. Prototype adapts itself to its model and simulation of other actors' behavior in the imaginary world comes closer to reality. This method uses meta-knowledge about learning allowing to preserve a "personality" and emotions.

INTRODUCTION

Our study takes place in the framework of Interactive Fictions where autonomous entities improvise with avatars [Hayes-Roth 96]. The idea is to provide the ability for virtual actors to adapt his representation of other actors' behavior, and therefore to carry out accurate predictions by simulating.

Each *Animat* [Meyer 91] has its own behavioral culture implemented in a library of behavioral prototypes. This culture gives it self-perception and perception of others. While reacting in virtual world, it can simulates in its imaginary world its vision of other entities in order to choose a strategy according to the prediction of the simulation [Maffre 01]. We propose to endow these *Animats* with learning by imitation [Mataric 01]. By observing another entity or avatar, our *Animat* modifies one behavioral prototype from its library in order to imitate the observed model with more accuracy, increasing the relevance of its predictions. It can also imitate another agent preserving its own "personality".

Fuzzy Cognitive Map (FCM) [Kosko 86] can specify and control emotional and perceptive (not only sensitive) *Animats* behavior [Parenthoën 01]. FCM is declarative and explanatory, it can therefore be specified by a non-specialist in computer science. *Animats* behavioral culture consists in a library of prototypic FCMs allowing it to simulate and to anticipate agents' behavior in its imaginary world. We propose to adapt prototypic FCM by learning process in order to imitate an observed behavior.

Applications are implemented in the multi-agent environment *oRis* [Harrouet 02] showing a sheepdog gathering sheep. The learning mechanism allows the dog to adapt its prey prototype to a given sheep in real time.

Next section explicits the context in which our learning algorithm is situated . We will justify the choice of the FCMs as foundation of the behavioral library, explain the notion of imaginary world and explain how we envisage the learning mechanism. Next, we will present the learning algorithm. Finally, we will apply this algorithm on the example of the sheepdog and explicit the obtained results.

CONTEXT

FCMs are graphs of influences allowing to specify and to control an *Animat* behavior. FCM is a dynamic system constituted by nodes and links. Nodes represent concepts and links causal connexions between concepts. Every concept has semantics. Information relating to the perception of an *Animat* are fuzzyfied to activate sensor concepts, while activations of motor concepts are defuzzyfied to determine its effectors. FCM is not only sensory but also perceptive thanks to self-excitator links and to links from internal to perceptive concepts.

We consider that an *Animat* has sensors allowing it to perceive its environment, effectors to perform, and also a library of prototypic behaviors specified by FCMs. A FCM is not only declarative, it is an explanatory graph fitting to behavior specification. Thus, an expert in collaboration with an ergonome will be able to develop a library of prototypic behaviors. This library represents the behavioral culture of the *Animat*. For example the library of an animal can be constituted of a prototypic behavior of prey, and a prototypic behavior of predator.

In parallel to the virtual world, an *Animat* has also an imaginary world, where it can simulate its own behavior and also other actors' behavior. This imaginary world corresponds to an approximate representation of the environment from *Animat* perception and to a representation of other actors' behavior. In Fact, an *Animat* uses prototypic behaviors in order to simulate other actors' behavior. It imagines its behavior in simulating its own decisonal mechanism and imagines other actors' behavior with prototypic FCMs. It can use its imaginary world to choose a strategy between several possibilities, not by a logical reasoning but by a behavioral simulation. Thus, it will be able to make predictions on the future.

We want to provide the ability for an *Animat* to adapt its representation of other actors' behavior and consequently its predictions become more pertinent. Thus, we propose to endow an *Animat* with a learning ability by imitation. An *Animat* must be able to modify a behavior to mime an observed behavior of a model that could be another actor or an avatar controlled by a human operator [Stoffregen 99]. By simple observation of the imitated model, the virtual actor must adapt its representation of the model behavior. The mechanism used to control the model behavior to imitate is independent of learning. Thus, imitated model can be piloted by any decision-making mechanism. The idea here is to modify prototypic FCMs representing other actors' behavior in comparing the result of the simulation in the imaginary world and the result of the virtual world. Thus we incorporate a third level to an *Animat* that we name "adaptative mode" (learning), adding to the reactive mode (virtual world) and to the "predictive mode" (imaginary world). These three modes represent the three levels used in cognitive psychology [Morineau 02]. The three methods are in communication, but they evolve in parallel.

LEARNING THROUGH IMITATION

In this section, we present a method allowing an adaptation of prototypic behavior by imitation in real time. An *Animat* observes its environment (other agents), allowing it to simulate other entities' behavior in its imaginary world with prototypic FCMs. The idea is to provide a more pertinent simulation by adapting prototypic FCMs by imitation. The modification of prototypic FCMs reduces the difference between predictions of the imaginary world and reality. We made the asumption that an *Animat* has sensors to estimate the information relating to prototypic FCMs, means an estimation of sensors and effector values that will allow to fuzzyfy sensors values and to compare the result of defuzzyfication of motor concepts activations with the effector values of the model.

The learning mechanism consists in getting back the result of the simulation in the imaginary world, comparing it to what happened in the virtual world, and deducting an adaptation of prototypic FCMs. We will limit our study to the learning of the weights of the causal connections between concepts in a prototypic FCM in order to imitate a given behavior, by modifying neither the structure of the influence graph of a FCM, nor the fuzzyfication of the sensors, nor the defuzzyfication of the concepts motors. This modification of the causal connections between concepts uses meta-knowledge about learning (the expert certifies notably structures of FCMs and the sign of links).

Kosko has proposed two different Hebb type methods [Hebb 49] for an expert given limit cycle learning by FCM [Kosko 88]. One is based on the correlations between activations [Kosko 92], the other on a correlation of their variations (differential hebbian learning) [Dickerson 94]. The differential learning modifies only the associated links to correlated variations of the concepts activations, while the non differential correlations learning risk to modify in a non pertinent way all links. Kosko's differential learning is based on the knowledge of a limit cycle including all concepts and provided by an expert. However, we can't have such a limit cycle, because only estimated model sensors and effectors can be observed and FCM having generated them is not available. In addition, Kosko's differential learning makes the assumption that external activations are constant. However, the virtual world is a dynamic system and external activations evolve in time. Thus, we will modify Kosko's hebbien differential learning to our case.

The algorithm of adaptation that we propose is an iterative cycle in four stages:

1. *Model estimation:*
   by simple observation the *Animat* estimates model-sensors and model-effectors,

2. *Simulation of the prototypic behavior:*
   sensors are fuzzyfied into perceptive concept external activations, calculation of the FCM dynamics, then image-effectors are obtained by motor concept inner activation defuzzyfication,

3. *Calculation of calling into question:*
   comparison between image-effectors and model-effectors is performed, generatation of a set of desired pseudo-activations obtained by going up the influence graph from motor concepts towards perceptive concepts without modifying links and by using meta-knowledge about learning,

4. *Update causal links:*
   FCM causal links are updated by applying discrete differential hebbian learning to the sequence corresponding to the passage from FCM activations towards desired pseudo-activations.

More precisely :

1. In the first stage, imitator measures features about the model, which are necessary for model-sensor and model-effector estimations.

2. The second stage corresponds simply to the usual use of a FCM for the control of a virtual actor, and determines image-actor FCM activations at moment $t + \delta t \approx t$ in the imaginary world, according to model-sensor estimation and FCM dynamics with N iterations:

$$\begin{array}{c} a(t + \frac{I}{N}\delta t) = S\left(G(f(t), L^T \cdot a(t + \frac{I-1}{N}\delta t))\right) \\ \text{for } I = 1, \cdots, N \; ; \;\; \delta t << 1 \end{array} \qquad (1)$$

N equals the length of the longest acyclic path added to the length of the longest cycle in the influence graph, in order to make sure that sensor information is spread to all nodes; n being FCM concept number, $f = (f_i)_{i[\![1,n]\!]}$ external activations coming from sensor fuzzyfication, $a{=}(a_i)_{i[\![1,n]\!]}$ innner activations, $L{=}(L_{ij})_{(i,j)[\![1,n]\!]^2}$ link matrix, $G : (\Re^2)^n \rightarrow \Re^n$ a comparison operator and S a standardization function transforming each coordonate by the sigmoidal function: $\sigma(x) = \frac{1+\delta}{1+e^{-\rho(x-a_0)}} - \delta$, with parameters $(\delta, \rho, a_0) \in \{0, 1\} \times \Re^+_* \times \Re$. FCM motor concept defuzzyfication at moment $t + \delta t \approx t$ provides image-effectors. For more clearness, we note a the resulting inner activations $a(t + \delta t)$ in next paragraphs.

3. The third stage recursivly generates sets of pseudo-activations $(P_i)_{i\in[\![1,n]\!]}$ translating an orientation for FCM dynamics. The principle consists in going up the influence graph from motor concepts towards perceptive concepts proposing pseudo-activation values according to meta-knowledge about learning and bringing image-effectors closer to model-effectors estimation. We did not use the method of gradient backpropagation [Rumelhart 86]. FCM is a cyclic process and its topology is not organized in layers (recurrent links). In addition, the method of gradient backpropagation does not hold graph semantic and we wished to have the possibility to apply specific meta-knowledge to a specific node. Let's detail the recursive process:

   Initialisation $m = 0$*:* entering into the FCM from effectors. A set I_0 represents indices of concepts defuzzyfied onto image-effectors. For each $i \in I_0$, we apply the decision learning meta-knowledge: two potential pseudo-activations $p_i^{\pm} = \sigma(a_0 \pm \frac{2\alpha_i}{\rho})$ simulate an active/inactive concept C_i, $\alpha_i \geq 1$ translating choise radicality. With the a_i value, that makes 3 possible pseudo-activations $p_i = a_i, \; p_i^+$ or p_i^- for each C_i. The $3^{\mathrm{Card} I_0}$ combinations are defuzzyfied, compared to model-effector estimation and the best combination $(p_i^{0,\{\}})_{i\in I_0}$ is kept (the 0 deals with defuzzyfication and the $\{\}$ is a set of future labels). $\forall i \in I_0, \; P_i = \{p_i^{0,\{\}}\}$. The other pseudo-activations sets $(P_i)_{i\in([\![1,n]\!]\setminus I_0)}$ are empty.

   Progression from m *to* $m+1$*:* Let $I_m \subset [\![1,n]\!]$ be the index set of concepts whose desired pseudo-activation set is not empty. For $i \in I_m$, note a_i (reps. f_i) inner (resp. extern) activation of concept C_i, $P_i = \{p_i^{k_1,\{\cdots\}}, \cdots, p_i^{k_L,\{\cdots\}}\}$ its desired pseudo-activation set which cardinal equals L and $J \subset [\![1,n]\!]$ the index set of concepts which are causes for the concept C_i (*i.e.*:$L_{ji} \neq 0$) and such that the arc from C_j to C_i has not been studied: $\forall \lambda \in [\![1,L]\!], \; j \neq k_\lambda$. We will calculate pseudo-activations P_j for $j \in J$ as follows:

   - For each $j \in J$, we apply the decision learning meta-knowledge: two potential pseudo-activations p_j^+ and p_j^- are calculated (2) so that their influence on a_i causes a clear choise between an active C_i or an inactive C_i, taking into account external activations, with $\alpha \geq 1$ translating the choice radicality:

   $$p_j^{\pm} = \left(a_0 \pm \frac{2\alpha_j}{\rho} - f_i - \sum_{l \neq j} L_{li} a_l \right) / L_{ji} \qquad (2)$$

   - Then we randomly select a $\lambda \in [\![1,L]\!]$. That gives a $p_i^{k_\lambda,\{\cdots\}} \in P_i$ and we choose among the $3^{\mathrm{Card} J}$ possible combinations $p_j^i = a_j, \; p_j^+$ or p_j^- for $j \in J$, the one $p_j^{i,\{\cdots,k_\lambda\}}$ which gives a C_i activation $\sigma\left(G_i(f_i, \sum_j L_{ji} p_j^i)\right)$ the nearest to $p_i^{k_\lambda,\{\cdots\}}$,
   - Thus we obtain a new set of concept indices with a not empty desired pseudo-activation set: $I_{m+1} = I_m \cup J$ with $P_j = P_j \cup \{p_j^{i,\{\cdots,k_\lambda\}}\}$for $j \in J$.

   Termination: if for each $i \in I_m$, the corresponding J set is empty, every arc belonging to paths arriving into $(C_i)_{i\in I_0}$ has been studied.

   We use a discrete method by proposing three pseudo-activations. We choose a discrete method allowing us on one hand to limit the calculations and on the other hand to translate a radical choice. We argue that to learn semantic purpose, proposed modifications have to correspond to radical choices and not to light modifications.

4. The fourth and last stage modifies FCM link weights, in order to direct its dynamics towards a behavior approaching the model. Contrary to Kosko who uses a cycling cycle and a learning rate decreasing with time (see [Dickerson 94] page 186), we make only one stage from inner activations a to link corresponding desired pseudo-activations p for the weight modification without cycling and preserve a constant learning rate $r(t) = R$, in order to ensure a strong adaptivity for our virtual actor. Formally, noting $\mathcal{A} \subset [\![1,n]\!]^2$ the arc set of the FCM, $\beta \in]0; 1{+}\delta[$ a sensitivity level and $s : \Re \rightarrow \{-1, 0, 1\}$ the discrete function $s(x) = -1, \; 0$ or 1 if respectively

$x \leq -\beta,\ -\beta < x < \beta$ or $x \geq \beta$, the learning algorithm follows the equations:

$$\begin{array}{l}\forall (i,j) \in \mathcal{A},\ \text{if}\ \exists k \in [\![0,n]\!],\ p_j^{k,\{\cdots,i,\cdots\}} \in P_j, \\ \text{we take such a } k \text{ and :} \\ \begin{cases} \Delta_i = s(p_i^{j,\{\cdots\}} - a_i),\ \Delta_j = s(p_j^{k,\{\cdots,i,\cdots\}} - a_j) \\ \\ L_{ij\,(t+1)} = \left| \begin{array}{ll} L_{ij\,(t)} + R(\Delta_i \Delta_j - L_{ij\,(t)}) & ,\ \text{if} \quad \Delta_i \neq 0 \\ L_{ij\,(t)} & ,\ \text{if} \quad \Delta_i = 0 \end{array} \right. \end{cases} \\ \text{else } \mathcal{A}_{ij} \notin \{\text{path to effectors}\} :\ L_{ij\,(t+1)} = L_{ij\,(t)} \end{array} \tag{3}$$

It is to note that we preserve a coherence in our modification of links according to the initial prototype furnished by the expert. Thus, the following possibilities are forbidden: link emergence, link suppression, or modification of the sign of a link. We also keep some link weights inside given bondaries $\mathcal{B}_{ij} = [L_{ij}^{min}, L_{ij}^{max}]$ so that the adapted behavior remains believable according to the expert: if $L_{ij\,(t+1)} < L_{ij}^{min}$ then $L_{ij\,(t+1)} = L_{ij}^{min}$ and if $L_{ij\,(t+1)} > L_{ij}^{max}$ then $L_{ij\,(t+1)} = L_{ij}^{max}$. Moreover, the expert can decide to immobilize the weight of one or several links, therefore they will not be modified during the learning process. To immobilize links or to impose limits allows to adapt prototypic FCMs while preserving a "personality".

RESULTS

Our applications show a sheepdog gathering sheep. During the simulation one or several sheep can move away from the gathering zone. When approaching a sheep, the dog frightens it and obliges it to regain this zone. The dog simulates in its imaginary world several strategies to gather sheep. We have implemented three applications showing a sheepdog gathering sheep. First, the dog learns a way of gathering sheep by the imitation of a human operator or another dog. In that case, the prototypic FCMs used is its own FCMs. Second, an adaptation of dog's prey prototype to a given sheep occurs in real time. This application is described in this section. Third, a paranoiac sheep learns how to be surrounded by other sheep remains frightened but does not flee any more when viewing a dog. To immobilize paranoiac links allows to adapt sheep behavior while preserving a paranoid "personality".

To simulate sheep behavior, the dog uses prototypic FCMs of prey from its behavioral library. Actually, the dog represents each sheep behavior by prototypic FCMs of prey in its imaginary world. Each sheep is associated with its own prototype. Thus the dog can simulate sheep behavior and can do predictions. Prototype will be adapted to a sheep by imitation. A FCM controls the prototype's speed and another controls the prototype's angle.

The comparison between the result of the imaginary world and the virtual world allows an adaptation of prototypic FCMs in real time by learning. The figure (1) illustrates the modification by imitation of prototype's speed that defined the representation of one sheep's speed used the imaginary world. We imposed the learning period. Such a period allows the convergence of the process.

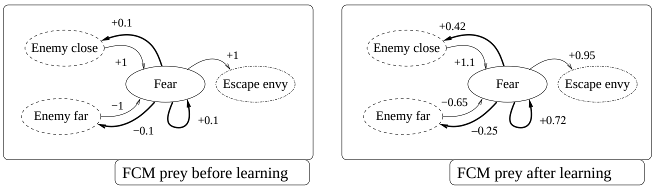

Figure 1: FCM of perceptive prey is coming from the library of prototypic FCMs and adapts itself by learning.

The dog observes the sheep to imitate. It adapts the prototypic behavior of prey allowing it to simulate the sheep's behavior in its imaginary world. By simple observation of the sheep to imitate, it estimates information necessary to the fuzzyfication for the prototype. The estimation of sensors values are fuzzyfied in activation of the concepts "Enemy close" and "Enemy far". The dynamic of the prototype occurs and by defuzzyfication of the activation of the effector motor "Escape envy" we get the image effector. Its corresponds to the representation that the dog has of prey's speed. This image effector from prototype is compared to an estimation of sheep's effectors. This comparison allows to calculate a set of pseudo activations that define desired modifications of FCM links. The prey prototype adapts itself to a sheep by reiterating the learning process. In pratice, the congercence occurs.

On figure (2), we compare the simulation of sheep behavior from prototype in the imaginary world and the sheep behavior in the virtual world, before and after learning while the dog performs the same trajectory. We note that the simulation is closest to reality after learning.

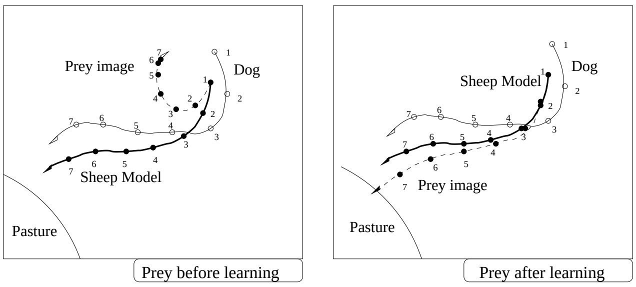

Figure 2: Learning by imitation allows to get more pertinent predictions from the imaginary world.

CONCLUSIONS AND FUTURE WORKS

Our *Animat* possesses a behavioral library composed by prototypic FCMs. While acting in the virtual world, the prototypic FCMs allows him to simulate other actors' behavior in its imaginary world. It simulates different strategies, allowing him to carry out predictions. We use FCMs because they represent an explicit knowledge and provide perception and emotions to the *Animat*. We have presented a learning algorithm allowing an adaptation of the prototypic FCMs to imitate a given actor. This adaptation provides a more pertinent imaginary world and therefore the *Animat* carries out predictions closest to the results of the virtual world. Our learning by imitation uses meta-knowledge from description of the prototypes by an expert, allowing to preserve the "personality" and the emotions of the prototype. In addition, our learning is based on a behavioral prototype allowing to simulate the model behavior to imitate. Moreover, we do not have to modify the structure of the influence graph of the FCM, the fuzzyfication of the sensors, and the defuzzyfication of the concept motors. Future works will try to set up a process that selects a prototype in the library by simple observation of the model behavior to imitate. Also, we work on the the adaption of the fuzzy transformations associated to the fuzzyfication and defuzzyfication.

REFERENCES

[Dickerson 94] Dickerson J.A., Kosko B., Virtual Worlds as Fuzzy Cognitive Maps, *Presence*, 3(2):173-189, MIT Press, 1994.

[Harrouet 02] Harrouet F., Tisseau J., Reignier P., Chevaillier P., oRis : un environnement de simulation interactive multi-agents, *Revue des sciences et technologie de l'information, série Technique et science informatiques (RSTI–TSI)* 21, no.4 :499-524, 2002.

[Hayes-Roth 96] Hayes-Roth B., Van Gent R., *Story-making with improvisational puppets and actors*, Technical Report KSL-96-05, Stanford University, 1996.

[Hebb 49] Hebb D.O., *The Organization of Behavior*, John Wiley & Sons (eds), New York, USA, 1949.

[Kosko 86] Kosko B., Fuzzy Cognitive Maps, *International Journal Man-Machine Studies*, 24:65-75, 1986.

[Kosko 88] Kosko B., Hidden patterns in combined and adaptative knowledge networks, *International Journal of Approximate Reasoning*, 2:337-393, 1988.

[Kosko 92] Kosko B., *Neural networks and fuzzy systems: A dynamical systems approach to machine intelligente*, Engelwood Cliffs, 1992.

[Maffre 01] Maffre E., Tisseau J., Parenthoën M., Virtual Agents Self-Perception in Virtual Story Telling, *ICVS'01 proceedings*, 155-158, Springer, 2001.

[Mataric 01] Mataric M.J., Sensory-Motor Primitives as a Basis for Learning by Imitation: Linking Perception to Action and Biology to Robotics, *Imitation in Animals and Artifacts*, Dautenhahn K. & Nehaniv C. (eds), MIT Press, 2001.

[Meyer 91] Meyer J.A., Guillot A., Simulation of adaptative behavior in animats: review and prospect, *from Animals to Animats*, 1:2-14, 1991.

[Morineau 02] Morineau T., Hoc J.M., Denecker P., *Cognitive Control Levels in Air Traffic Radar Controller*, To appear in Internal Journal of Aviation Psychology, 2002.

[Parenthoën 01] Parenthoën M., Reignier P., Tisseau J., Put Fuzzy Cognitive Maps to Work in Virtual Worlds, *Fuzz-IEEE'01 proceedings*, 1:P038, 2001.

[Rumelhart 86] Rumelhart D.E, Mc Clelland J.L. and the PDP research group, *Parallel Distribued Processing Exploration in the microstructure of cognition, Vol I, II and III.*, A bradford book, MIT press, Cambridge (MA), 1986.

[Stoffregen 99] Stoffregen T.A., Gorday K.M., Sheng Y-Y., Flynn S.B., Perceiving affordances for another person's actions, *Journal of Experimental Psychology: Human Perception and Performance*, 25:120-136, 1999.