

Apprentissage par Imitation pour l'Autonomie des Acteurs Virtuels

Marc Parenthoën*, Cédric Buche*, Jacques Tisseau*

* Laboratoire d'Informatique Industrielle (*Li*², EA 2215 UBO/ENIB)
ENIB - Parvis Blaise Pascal - 29280 Plouzané - BP 30815 - F-29608 Brest Cedex FRANCE
Tél : +33 (0)298 05 66 31 Fax : +33 (0)298 05 66 29 Web : [{parenthoen, tisseau, buche}@enib.fr](http://www.enib.fr/chercheur/LI2)

Mots-Clés: *apprentissage, autonomie, acteur virtuel, carte cognitive floue, imitation*

Résumé étendu

Le but de ce travail est de donner plus d'autonomie à des acteurs virtuels en les dotant d'une capacité d'imitation. Un acteur virtuel est un agent autonome ayant sa propre personnalité, ses propres émotions [Hayes-Roth 96]. Il possède des capteurs et des effecteurs contrôlés par des Graphes Cognitifs Flous (GCFs, en anglais: Fuzzy Cognitive Maps ou FCMs [Kosko 86]) qui spécifient le comportement [Parenthoën 01]. Un expert en collaboration avec un ergonome peut fournir les prototypes de tels GCFs [Parenthoën 02]. La connexion entre l'agent autonome et un GCF se traduit par une fuzzification des capteurs pour déterminer l'activation forcée des concepts perceptifs de ce GCF et par une défuzzification des activations des concepts moteurs fixant les effecteurs. Les concepts internes peuvent traduire des émotions et sont utilisés selon le graphe d'influence lors de la dynamique du GCF convergeant vers son attracteur. Un acteur virtuel possède également un monde imaginaire, dans lequel il peut simuler son propre comportement mais aussi celui d'autres acteurs [Maffre 01]. Il imagine son comportement en utilisant ses propres GCFs, et celui des autres acteurs avec des GCFs prototypiques. Par exemple, un chien de berger voulant garder un troupeau de moutons possèdera ses GCFs lui permettant de regrouper des moutons, mais aussi des GCFs prototypiques d'une proie lui permettant de simuler dans son monde imaginaire le comportement de moutons et des GCFs prototypiques de chiens de berger pour simuler les autres chiens. Cette perception de lui-même et des autres implémentée par une simulation dans la simulation est l'une des clés pour l'autonomie d'apprentissage. Dans le cadre d'un apprentissage par imitation, l'acteur virtuel doit pouvoir modifier son comportement pour mimer le comportement observé d'un modèle qui peut être un autre acteur ou un avatar contrôlé par un opérateur humain [Stoffregen 99]. Par exemple, un chien de berger peut être amené à apprendre le comportement proposé par le berger, par un autre chien ou par un mouton. Nous allons voir comment un acteur virtuel peut adapter en temps réel un comportement dans son monde imaginaire pour imiter un comportement observé. Nous limiterons notre étude à l'apprentissage des poids des connexions causales entre concepts dans un GCF prototypique afin d'imiter un comportement donné, sans modifier ni la structure du graphe d'influence du GCF, ni la fuzzification des capteurs, ni la défuzzification des concepts moteurs. Cette modification des connexions causales entre concepts utilise des méta-connaissances sur l'apprentissage.

Ce paragraphe présente l'algorithme d'apprentissage par imitation pour des acteurs virtuels contrôlés par des GCFs. Nous souhaitons mettre en place une méthode d'apprentissage par imitation permettant d'adapter des GCFs prototypiques existants, issus d'un travail d'expertise. L'acteur observe le modèle et simule le comportement d'un acteur-image devant suivre ce modèle. Cette image d'acteur peut être sa propre image, ou celle d'un autre acteur dont il possède un GCF prototypique. L'acteur devra s'efforcer d'imiter dans son monde imaginaire le comportement d'un modèle en adaptant ses GCFs. Ce modèle piloté par ses effecteurs peut être un acteur virtuel ou un avatar contrôlé par un opérateur humain. L'imitateur n'a accès qu'à l'observation des effecteurs et à une estimation des capteurs du modèle. L'apprenant doit posséder lui-même de tels effecteurs et des capteurs lui permettant cette estimation. Il fait l'hypothèse que le comportement du modèle est contrôlable par un GCF dont il connaît une structure prototypique. Il lui suffit alors d'identifier la personnalité du modèle en modifiant la matrice des liens de son GCF prototypique. Un GCF adaptatif doit modifier les poids des liens de causalité entre ses concepts en fonction de données acquises du monde extérieur, de telle sorte que ses bassins d'attraction génèrent le comportement souhaité. Kosko a proposé deux méthodes différentes de type Hebb [Hebb 49] pour l'apprentissage par un GCF d'un cycle limite déterminé par un expert [Kosko 88]. L'une est basée sur les corrélations entre activations [Kosko 92], l'autre sur une corrélation de leurs variations (apprentissage hebbien différentiel) [Dickerson 94]. L'avantage de l'apprentissage des variations est que ne sont modifiés que les liens associés à des variations corrélées des activations des concepts, alors que les corrélations non différentielles risquent de modifier de

manière non pertinente l'ensemble de tous les liens. Autrement dit, l'apprentissage hebbien différentiel ne tient compte que des moments où il y a changement significatif entre le comportement dynamique du GCF et le comportement dynamique souhaité. Nous allons décrire une méthode permettant une adaptation en temps réel. L'algorithme d'apprentissage hebbien différentiel de Kosko d'une part se base sur la connaissance d'un cycle limite des activations de l'ensemble de tous les concepts fourni par un expert et d'autre part fait l'hypothèse que les activations forcées externes sont constantes. Or les activations externes évoluent dans le temps et nous ne pouvons disposer d'un tel cycle limite, car nous ne pouvons observer que les effecteurs du modèle et ne disposons pas du GCF les ayant générés: un opérateur humain peut conduire ce modèle. L'acteur va simuler le comportement du modèle dans un espace imaginaire avec un GCF adaptatif, et comparer les effecteurs de l'image avec les effecteurs du modèle pour mettre à jour ce GCF. Dans un premier temps, c'est le modèle qui décide la période d'apprentissage. A la fin de cette période, le GCF adapté à l'imitation du modèle remplace le prototype initial. L'algorithme d'adaptation que nous proposons est un cycle itératif en trois étapes entre t et $t + 1$:

1. fuzzification des capteurs en activations externes des concepts perceptifs, calcul de la dynamique du GCF convergeant vers son attracteur, puis des effecteurs-images par défuzzification des activations internes des concepts moteurs,
2. comparaison des effecteurs-images avec les effecteurs du modèle, et génération d'un vecteur de pseudo-activations souhaitées obtenu en remontant le graphe d'influence depuis les concepts moteurs vers les concepts perceptifs sans modifier les liens et en utilisant des méta-connaissances sur l'apprentissage des concepts,
3. mise à jour des liens causaux du CGF en appliquant l'apprentissage hebbien différentiel sur la séquence des 2 termes correspondant au passage des activations du GCF vers les pseudo-activations souhaitées.

La première étape correspond simplement à l'utilisation usuelle d'un GCF pour le contrôle d'un acteur virtuel, et détermine dans le monde imaginaire les activations à l'instant t de l'acteur-image.

$$a(t + \frac{1}{N}\delta t) = \sigma \circ G(f_a(t), L^T \cdot a(t + \frac{1}{N}\delta t))$$

pour $I = 1, \dots, N$; $\delta t \ll 1$

f_a étant les activations externes provenant de la fuzzification des capteurs, a les activations internes, L la matrice des liens, G un opérateur de comparaison et σ une fonction de normalisation transformant chaque coordonnée par une sigmoïde: $x \mapsto \frac{1+\delta}{1+e^{-k(x-a_0)}} - \delta$. La défuzzification des concepts moteurs de ce GCF fournit à l'instant $t + \delta t \approx t$ les effecteurs-images.

La seconde étape génère un vecteur de pseudo-activations traduisant une orientation pour la dynamique du GCF. Le principe consiste à remonter le graphe d'influence depuis les concepts moteurs vers les concepts perceptifs en proposant des valeurs des concepts d'après des méta-connaissances sur l'apprentissage et allant dans le sens d'une décision rapprochant les effecteurs de l'image de ceux du modèle à l'instant t . Notons $I_k \subset \llbracket 1; n \rrbracket$ l'ensemble des indices des concepts dont la pseudo-activation souhaitée est déterminée. Prenons $i \in I_k$, notons a_i (reps. f_{a_i}) l'activation interne (resp. externe) de C_i , p_{a_i} sa pseudo-activation souhaitée et $J \subset (\llbracket 1; n \rrbracket \setminus I_k)$ l'ensemble des indices des concepts non encore pseudo-évalués et qui sont des causes pour le concept C_i , i.e: $L_{ji} \neq 0$ pour $j \in J$. Nous allons calculer les pseudo-activations p_{a_j} pour $j \in J$ comme suit:

- Pour chaque $j \in J$, on applique la méta-connaissance d'apprentissage de décision: deux pseudo-activations potentielles $p_{a_j}^+$ et $p_{a_j}^-$ sont calculées pour que leur influence sur a_i provoque un choix net entre un C_i actif ou un C_i inactif compte tenu des activations externes, i.e avec $\alpha_j \geq 1$ traduisant la radicalité du choix:

$$p_{a_j}^\pm = \left(a_0 \pm \frac{\alpha_j}{2k} - f_{a_j} - \sum_{l \neq j} L_{li} a_l \right) / L_{ji}$$

- Ensuite, on choisit parmi les $3^{\text{Card}J}$ combinaisons possibles $p_{a_j} = a_j$, $p_{a_j}^+$ ou $p_{a_j}^-$ pour $j \in J$, celle qui donne le résultat le plus proche de p_{a_i} au sens d'une distance,
- On obtient ainsi un nouvel ensemble d'indices de concepts possédant une pseudo-activation:

$$I_{k+1} = I_k \cup J$$

Si pour chaque $i \in I_k$, l'ensemble J correspondant est vide, alors on termine la construction des pseudo-activations en donnant aux éventuels concepts non pseudo-évalués leur activation de l'instant $t + \delta t \approx t$, i.e: $\forall i \in \llbracket 1; n \rrbracket \setminus I_k, p_{a_i} = a_i$. L'initialisation I_0 est obtenu en déterminant la meilleure défuzzification (distance minimale entre effecteurs observés et calculés) toujours selon le méta-principe d'une prise de décision sur les concepts moteurs.

La troisième étape permet de modifier les liens du GCF afin d'orienter sa dynamique vers un comportement s'approchant de celui du modèle. Contrairement à Kosko qui utilise un cycle, nous ne faisons qu'une seule étape lors de la modification des poids et conservons un taux d'apprentissage constant, afin d'assurer une adaptivité forte pour notre acteur virtuel. Dans la pratique, les expériences menées sur le chien de berger montrent que la convergence a bien lieu. On peut bien-sûr modifier le taux d'apprentissage au cours du temps, en lui faisant suivre une suite décroissante tendant vers zéro et dont la série associée diverge vers l'infini. Cela assurerait une convergence théorique des poids du GCF, mais l'adaptivité serait de moins en moins forte avec l'âge de l'acteur.

La complexité de cet algorithme est polynomiale par rapport au nombre n de concepts intervenant dans l'expertise, et même en $\mathcal{O}(n)$. Sans faire attention, on pourrait croire que cet algorithme présente une complexité exponentielle or il n'en est rien. En effet, pour un expert, les causes d'un concept sont toujours en nombre très limité (rarement plus de cinq), donc sur chaque concept n'arrivent qu'au maximum 5 arcs. Le $3^{\text{Card}J}$ est donc majoré dans la pratique, quel que soit le nombre de concepts impliqués dans le GCF. Il en va de même pour le calcul de la dynamique du GCF qui est en $\mathcal{O}(n)$ alors qu'il pourrait apparaître comme un $\mathcal{O}(n^2)$, ce grâce au grand nombre de zéros dans la matrice des liens (le nombre de liens non nul dans une colonne de L n'exédant pas 5, quel que soit n). Cet algorithme peut donc être implémenté pour une utilisation en temps réel.

Nous avons doté des acteurs virtuels de la capacité d'imitation en temps réel. C'est une des clés de l'autonomie. Les applications implémentées dans l'environnement *oRis* [Harrouet 00] montrent un chien de berger réglant ses GCFs afin d'imiter un autre chien, un modèle piloté par un opérateur humain ou afin d'adapter son prototype de proie à un monton donné. Cependant, les périodes d'apprentissage ne sont pas choisies par l'acteur lui-même: c'est un opérateur humain qui les impose. Nous travaillons à inclure cette capacité d'imitation dans un processus sensorimoteur plus complexe permettant une véritable autonomie d'apprentissage par l'action.

References

- [Dickerson 94] Dickerson J.A., Kosko B., Virtual Worlds as Fuzzy Cognitive Maps, *Presence*, 3(2):173-189, MIT Press, 1994.
- [Harrouet 00] Harrouet F., *oRis: s'immerger par le langage pour le prototypage d'univers virtuels à base d'entités autonomes*, Thèse de Doctorat, Université de Bretagne Occidentale, Brest, France, 8 Decembre, 2000.
- [Hayes-Roth 96] Hayes-Roth B., Van Gent R., Story-making with improvisational puppets and actors *Technical Report KSL-96-05*, Knowledge Systems Laboratory, Stanford University, 1996.
- [Hebb 49] Hebb D.O., *The Organization of Behaviour*, John Wiley & Sons, 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 adaptive 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 intelligence*, Engelwood Cliffs, 1992.
- [Maffre 01] Maffre E., Tisseau J., Parenthoën M., Virtual Agents Self-Perception in Virtual Story Telling, *in ICVS 2001 proceedings*, 155-158, Springer LNCS 2197, Avignon, France, September, 2001.
- [Meyer 91] Meyer J.A., Guillot A., Simulation of adaptative behavior in animats: review and prospect, *in form Animals to Animats'91 Proceedings*, 1:2-14, 1991.
- [Parenthoën 01] Parenthoën M., Reignier P., Tisseau J., Put Fuzzy Cognitive Maps to Work in Virtual Worlds, *in FUZZ-IEEE 2001 proceedings*, 1:56-60, Melbourne, Australia, December, 2001.
- [Parenthoën 02] Parenthoën M., Tisseau J., Morineau T., Proactive Perception for Virtual Actors, *to appear in SCI 2002 proceedings*, Orlando, Florida, USA, October, 2002.
- [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.