Extraction de motifs séquentiels à partir de traces d'utilisation pour la construction automatique de modèles de tâches dans les systèmes tutoriels intelligents

Philippe Fournier-Viger*, Engelbert Mephu Nguifo**, Roger Nkambou*

*Département d'informatique, Université du Québec à Montréal
201, avenue du Président-Kennedy, Montréal, Canada, H2X 3Y7
philippe.fournier-viger@courrier.uqam.ca, <http://www.dinfo.uqam.ca>

**LIMOS - CNRS UMR 6158

UFR Sciences et Technologies, Université Blaise Pascal - Clermont 2
Complexe scientifique des Cézeaux, BP 125, 63173 AUBIERE cedex, FRANCE.
mephu@isima.fr, <http://www.isima.fr/limos/>

1 Introduction

Pour construire un système tutoriel intelligent capable d'offrir une assistance personnalisée à des apprenants lors d'activités de résolution de problèmes, il faut donner au système un modèle de la tâche que les apprenants doivent accomplir. L'analyse cognitive de tâche, approche classique pour leur construction, consiste à observer plusieurs utilisateurs pour ensuite en modéliser à la main les différentes façons de la réaliser. Or, ce processus d'acquisition est une tâche souvent ardue et coûteuse en temps (Aleven et al., 2006). De plus, pour certains domaines dits « mal-définis », il peut-être tout simplement impensable de définir un modèle de tâche complet ou partiel à la main (Lynch et al., 2006). La solution que nous proposons à cette problématique est d'extraire automatiquement des modèles de tâche à partir d'un ensemble de résolutions de problèmes enregistrées par plusieurs utilisateurs de différents niveaux d'expertises. Le domaine d'application de notre travail est CanadarmTutor (Kabanza et al., 2005), un système tutoriel intelligent pour l'entraînement des astronautes à la manipulation du bras robotisé Canadarm2 déployé sur la station spatiale internationale. Canadarm2 est un bras télémanipulateur à 7 degrés de liberté. Son opération est une tâche complexe puisqu'un opérateur doit éviter les collisions et qu'il ne possède pas une vue directe du bras et de son environnement, mais doit plutôt en tout temps choisir trois caméras parmi une dizaine disponibles sur la station pour les visualiser. De plus, il existe un très grand nombre de possibilités pour déplacer le bras d'une position à une autre, et il n'y a pas de façon simple de formaliser comment choisir les meilleures manipulations à faire à partir d'une configuration donnée. Pour cette raison, l'application d'une approche d'extraction automatique de modèles de tâches à partir de traces d'utilisation du bras robotisé a été privilégiée pour modéliser la tâche dans CanadarmTutor.

2 Un algorithme pour l'extraction de modèles de tâche

Pour l'extraction automatique de modèles de tâches à partir de traces d'utilisation, nous proposons le processus suivant. Tout d'abord, un ensemble de solutions pour une tâche est enregistré sous forme de séquences d'actions dans une base de données séquentielle. Ensuite,

un algorithme de recherche de motifs séquentiels (Han et Kamber, 2006) est appliqué pour extraire des séquences fréquentes. Ces motifs séquentiels peuvent ensuite être exploités pour offrir des services tutoriels.

Les premiers essais que nous avons conduits avec des algorithmes classiques de recherche de motifs séquentiels ont montré que ce contexte particulier bénéficierait d'un algorithme adapté. Nous avons développé un algorithme hybride combinant les propriétés de plusieurs algorithmes de recherche de motifs séquentiels (Fournier-Viger et al., 2008)¹:

- **recherche de motifs séquentiels avec un support minimum par pseudo-projection**
- **recherche de motifs séquentiels avec dimensions** : pour pouvoir indiquer avec des dimensions le succès/échec de chaque trace de résolution de problème et le profil de la personne l'ayant effectué, et découvrir des motifs annotés avec ces informations.
- **acceptation de contraintes temporelles sur les séquences** : pour limiter la taille des « trous » entre itemsets dans les motifs séquentiels découverts.
- **découverte de motifs fermés avec dimensions** : pour éliminer une forme de redondance dans les motifs découverts et accélérer la recherche.

En plus de ces propriétés, l'algorithme offre la caractéristique originale d'accepter des séquences où les actions (items) sont annotées avec des nombres, et fait un regroupement des valeurs similaires lors de la découverte des motifs séquentiels. Ceci a le bénéfice de permettre d'annoter les actions effectuées par les apprenants avec des valeurs les quantifiant (ex. : nombre de degrés de rotation d'un joint du bras robotisé) et de grouper les valeurs similaires. L'algorithme a été appliqué avec succès dans le logiciel CanadarmTutor et a permis l'implémentation de services tutoriels comme la suggestion des prochaines actions à faire aux utilisateurs lors de leur pratique avec le simulateur (Fournier-Viger et al., 2008).

Des améliorations des algorithmes et des services tutoriels offerts sont en cours. De plus, nous effectuons actuellement des évaluations empiriques pour évaluer les bénéfices à l'apprentissage.

Références

- Aleven, V., McLaren, B.M., Sewall, J., et Koedinger, K. (2006). *The Cognitive Tutors Authoring tools (CTAT): Preliminary evaluation of efficiency gains*. Proc. ITS 2006, 61-60.
- Fournier-Viger, P., Nkambou, R. et Mephu Nguifo, E. (2008). *A Knowledge Discovery Framework for LearningTask Models from User Interactions in Intelligent Tutoring Systems*. Proc. of the 7th Mex. Intern. Conf. On Artif. Intel. (MICAI 2008), 765-778.
- Kabanza, F., Nkambou, R. et Belgith, K. (2005). *Path-Planning for Autonomous Training on Robot Manipulators in Space*. Proc. IJCAI 2005. 1729-1773.
- Han, J. et Kamber, M. (2006). *Data mining : concepts and techniques, 2nd ed.*, San Francisco :Morgan Kaufmann Publisher,.
- Lynch, C., Ashley, K., Aleven, V. et Pinkwart, N. (2006). *Proc. of the Intelligent Tutoring Systems for Ill-defined Domains Workshop (ITS2006), 2006*.

¹ Une implémentation de l'algorithme est disponible à l'adresse : <http://www.philippe-fournier-viger.com/spmf/>.