

Extraction de connaissances d'adaptation par analyse de la base de cas

Fadi Badra*, Jean Lieber*, Amedeo Napoli*

*LORIA (UMR 7503 CNRS–INPL–INRIA–Nancy 2–UHP)
BP 239, 54506 Vandœuvre-lès-Nancy, FRANCE
{badra,lieber,napoli}@loria.fr

Résumé. En raisonnement à partir de cas, l'adaptation d'un cas source pour résoudre un problème cible est une étape à la fois cruciale et difficile à réaliser. Une des raisons de cette difficulté tient au fait que les connaissances d'adaptation sont généralement dépendantes du domaine d'application. C'est ce qui motive la recherche sur l'acquisition de connaissances d'adaptation (ACA). Cet article propose une approche originale de l'ACA fondée sur des techniques d'extraction de connaissances dans des bases de données (ECBD). Nous présentons CABAMA-KA, une application qui réalise l'ACA par analyse de la base de cas, en utilisant comme technique d'apprentissage l'extraction de motifs fermés fréquents. L'ensemble du processus d'extraction des connaissances est détaillé, puis nous examinons comment organiser les résultats obtenus de façon à faciliter la validation des connaissances extraites par l'analyste.

1 Introduction

Raisonner à partir de cas consiste à résoudre un problème à l'aide d'une base de cas, dans laquelle un cas représente un problème déjà résolu accompagné de sa solution (Riesbeck et Schank (1989)). Un système de raisonnement à partir de cas (RÀPC) sélectionne un cas dans la base de cas, puis adapte la solution associée. L'adaptation nécessite des connaissances spécifiques au domaine d'application. L'acquisition de connaissances d'adaptation a pour but d'extraire ces connaissances, ce qui peut être réalisé soit directement auprès d'un expert du domaine (d'Aquin et al. (2006)), ou encore par analyse de la base de cas (voir par exemple Hanney et Keane (1996), McSherry (1998), Craw et al. (2006)).

Un cas est généralement représenté par un couple $(pb, Sol(pb))$ dans lequel pb représente un énoncé de problème et $Sol(pb)$ une solution de pb . L'ensemble des *cas sources* $(srce, Sol(srce))$ d'un système de RÀPC constitue la *base de cas* BC . Lors d'une session particulière de RÀPC, le problème à résoudre est appelé *problème cible*, dénoté par *cible*. Une inférence à partir de cas associe à *cible* une solution $Sol(cible)$, compte tenu de la base de cas BC et de bases de connaissances additionnelles, en particulier O , l'*ontologie du domaine*, qui introduit les concepts et les termes utilisés pour représenter les cas.

Le processus de RÀPC est principalement composé d'une étape de remémoration et d'une étape d'adaptation. La *remémoration* sélectionne $(srce, Sol(srce)) \in BC$ tel que $srce$ est