

Correction et complétude d'un algorithme de recherche d'information par treillis de concepts

Nizar Messai, Marie-Dominique Devignes, Amedeo Napoli, Malika Smaïl-Tabbone

UMR 7503 LORIA, BP 239, 54506 Vandœuvre-lès-Nancy, FRANCE
{messai,devignes,napoli,malika}@loria.fr
<http://www.loria.fr/~messai>

Résumé. Dans cet article nous présentons BR-Explorer, un algorithme de recherche d'information correct et complet qui s'appuie sur la classification par treillis de concepts. L'algorithme BR-Explorer a pour objectif la recherche d'objets pertinents pour une requête donnée. Initialement, nous disposons d'un contexte formel représentant la relation entre un ensemble d'objets décrits par un ensemble d'attributs et du treillis de concepts correspondant à ce contexte. Étant donné une requête, l'algorithme BR-Explorer commence par générer un concept formel représentant la requête puis l'insère dans le treillis de concepts. Ensuite, BR-Explorer localise, dans le treillis résultant, un "concept pivot" à partir duquel il construit la réponse à la requête étape par étape en effectuant un parcours en largeur des concepts subsumants le concept pivot jusqu'au concept le plus général du treillis. Finalement BR-Explorer retourne en réponse l'ensemble des objets pertinents ordonnés selon leur degré de pertinence pour la requête.

1 Introduction

La notion de treillis de Galois (ou treillis de concepts) d'une relation est à la base d'une famille de méthodes de classification conceptuelle (Godin et al., 1995b). Introduite par Barbut et Monjardet (1970), cette approche a été à la base de l'analyse formelle de concepts (Ganter et Wille, 1999) où on propose de considérer chaque élément du treillis comme un concept formel et le graphe (diagramme de Hasse) comme une relation de généralisation/spécialisation entre les concepts. Cette forme de classification constitue l'une des motivations de l'application des treillis de concepts pour la recherche d'information dès l'apparition de l'analyse formelle de concepts. La recherche d'information a par la suite été explicitement mentionnée comme étant l'une des applications possibles des treillis de concepts (Godin et al., 1995b). Plus tard, une approche pour la recherche d'information en utilisant les treillis de concepts, la recherche d'information par treillis, a été proposée dans (Carpineto et Romano, 2000, 2004). La motivation principale qui a conduit à l'utilisation des treillis de concepts pour la recherche d'information vient du fait que la classification en treillis permet de combiner la recherche par requête et la recherche par navigation (Godin et al., 1995a). En effet, les concepts formels qui s'organisent en un treillis de concepts (voir les définitions dans la section 2) sont vus comme des classes de documents pertinents pour un ensemble de contraintes données (représenté par une requête).