

Régler le processus d’exploration dans l’analyse relationnelle de concepts – Le cas de données hydroécologiques

Amirouche Ouzerdine*, Agnès Braud**, Xavier Dolques***,
Marianne Huchard*, Florence Le Ber**

* LIRMM, Univ Montpellier, CNRS, Montpellier, France
labib23dz@hotmail.com, huchard@lirmm.fr

** ICube, Université de Strasbourg, CNRS, ENGEES, France
agnes.braud@unistra.fr, florence.leber@engees.unistra.fr,

*** Movidone, Strasbourg, France
xavier.dolques@laposte.net

Résumé. Cet article s’intéresse à l’exploration de jeux de données multi-relationnelles, et aux différentes manières de les analyser en utilisant l’analyse relationnelle de concepts (ARC), une variante de l’analyse formelle de concepts. L’ARC utilise plusieurs quantifieurs d’échelonnage qui rendent le processus d’analyse finement réglable, permettant une grande flexibilité dans l’exploration et dans ses résultats. En contrepartie, l’analyste peut être submergé par l’ensemble des choix qu’il doit faire au cours de l’analyse. Pour traiter ce problème, nous proposons trois sur-couches qui aident l’analyste à anticiper et contrôler les résultats de ses choix. Notre proposition est appliquée à un jeu de données sur la qualité des eaux de rivières.

1 Introduction

Les jeux de données multi-relationnelles suivent un schéma (modèle de données), où des entités (objets) de plusieurs catégories sont décrites par leurs caractéristiques (attributs) et où des relations relient les objets des différentes catégories. Les experts des domaines concernés les exploitent dans le cadre de différentes tâches : consultation ou exploration, requête, extraction de motifs, ou encore classification, au sens d’organiser dans une hiérarchie de généralisation des ensembles d’objets similaires. Cet article se concentre sur la tâche d’*exploration* fondée sur une telle organisation hiérarchique.

L’analyse formelle de concepts, ou AFC (Ganter et Wille, 1999), et ses extensions apportent des méthodes qui contribuent à l’exploitation des données, notamment multi-relationnelles (Liquière et Sallantin, 1998; Kötters, 2013; Ferré, 2015; Ferré et al., 2005; Hacene et al., 2013). Parmi ces méthodes, l’analyse relationnelle de concepts (ARC) a été spécialement conçue pour les tâches sus-citées (Hacene et al., 2013). Cette méthode construit un ensemble de classifications interconnectées, qui peuvent être utilisées pour extraire des motifs et des règles d’implication portant sur les liens inter-objets ainsi que sur des abstractions de ces liens.

L’ARC a pour caractéristique principale de construire des abstractions des liens inter-objets en appliquant des opérations d’échelonnage sur des groupes d’objets. Ces abstractions per-