

Qualité des règles et des opérateurs en découverte de connaissances floues

Maurice Bernadet

IRIN / Ecole Polytechnique de l'Université de Nantes
Rue Christian Pauc, La Chantrerie
BP 60601, 44306 Nantes Cedex 3, France
Maurice.Bernadet@polytech.univ-nantes.fr

Résumé. Nous décrivons ici l'évaluation de la qualité des règles et celle de la qualité des opérateurs dans un système de découverte de connaissances floues. Nous indiquons d'abord comment il est possible de définir des partitions floues pour remplacer les attributs classiques par des attributs flous. Nous généralisons ensuite à des règles floues trois indices de qualité utilisés pour des règles classiques et nous détaillons un algorithme d'extraction de règles floues. Etant donné que les logiques floues fournissent une infinité d'opérateurs logiques, nous proposons une méthode pour évaluer la qualité d'un jeu d'opérateurs flous quand les règles extraites doivent être exploitées dans des systèmes à base de connaissances à l'aide du Modus Ponens Généralisé. Nous appliquons cette évaluation sur un exemple simple et nous exposons quelques résultats de ces méthodes sur des bases de données réelles. Nous indiquons les jeux d'opérateurs flous qui s'avèrent de meilleure qualité et en quoi ces résultats correspondent à la théorie. En rappelant des résultats sur la réduction de règles floues, nous constatons que les jeux d'opérateurs dont la qualité est la meilleure pour utiliser le Modus Ponens Généralisé dans des systèmes à base de connaissances, ne sont pas les mieux adaptés à la réduction des règles.

1. Introduction

Quand on réalise un système de découverte de connaissances, des choix initiaux doivent être effectués : le processus peut être supervisé ou non et l'on peut représenter les connaissances à l'aide d'arbres de décision, de règles d'associations, de réseaux neuronaux... Pour le système que nous avons conçu, nous avons choisi, pour des raisons de simplicité, un processus supervisé et une représentation des connaissances utilisant des règles. Ensuite, considérant qu'il est souvent difficile d'interpréter des règles sur des intervalles d'attributs numériques et que des seuils stricts provoquent souvent des coupures trop brutales entre intervalles, nous avons décidé d'utiliser une représentation des connaissances basées sur les logiques floues. Les logiques floues peuvent être considérées comme des généralisations des logiques à valeurs multiples et autorisent des valeurs de vérité intermédiaires entre le vrai et faux (Zadeh 1974). Elles permettent d'exprimer les connaissances de manière plus naturelle que les logiques booléennes classiques, en utilisant des attributs gradués dans des formules comme "X est plutôt haut" (ou "X est haut est plutôt vrai"), "Y est un peu chaud" (ou "X est chaud est un peu vrai")... En outre, les logiques floues offrent le choix de nombreux opérateurs logiques (Klir et Yuan 1995), ce qui permet une bonne diversité dans l'expression des connaissances.