

Extraction de Règles en Incertain par la Méthode Implicative

Régis Gras*, Raphaël Couturier**, Fabrice Guillet*, Filippo Spagnolo***

** LINA– Ecole Polytechnique de l'Université de Nantes - BP 60601 44306 Nantes
regisgra@club-internet.fr et Fabrice.Guillet@polytech.univ-nantes.fr

** Institut Universitaire de Technologie de Belfort, BP 527, rue E. Gros, 90016 Belfort
Raphael.Couturier@iut-bm.univ.fcomte.fr

*** G.R.I.M., Department of Mathematics, University of Palermo
spagnolo@math.unipa.it.

Résumé. En relation avec des approches classiques de l'incertain, l'analyse statistique implicative (A.S.I.) peut apparaître innovante, particulièrement pour l'opérateur d'implication. L'article montre en effet que la notion de variables à valeurs intervalles et celle de variables-intervalles sont efficaces dans la détermination de leur distribution et dans la recherche de règles entre variables floues. De plus, elles apportent de riches informations sur la qualité de ces règles, tout en permettant d'étudier le rôle des variables supplémentaires dans l'existence de ces règles. Cette nouvelle perspective épistémologique de l'incertain ouvre d'intéressantes perspectives d'application.

1 Introduction

Partant du cadre défini et formalisé par (Lofti et al. 2001), (Dubois et al. 1987), ce texte vise à étudier les proximités formelle et sémantique des cadres de l'incertain et de l'analyse statistique implicative (A.S.I.) entre variables à valeurs intervalles et variables-intervalles (Gras et al. 2001). On s'intéresse particulièrement à l'opérateur « implication » avec lequel on extrait des règles d'association. Ce texte s'inscrit dans le cadre initié par (Gras 1979) sur l'analyse de données, A.S.I., qui vise à extraire et représenter, des règles d'association entre variables ou conjonctions de variables, du type $a \Rightarrow b$. Nous considérons celles qui croisent des sujets et des variables, présentant des modalités nettes ou floues. La qualité de la règle sera d'autant plus grande que son nombre de contre-exemples sera invraisemblablement petit sous l'hypothèse d'indépendance a priori, eu égard aux occurrences. .

2 problématique

Bien que les applications de la logique floue soient nombreuses en intelligence artificielle (par exemple en matière de diagnostic médical ou de reconnaissance des formes), plusieurs questions restent bien souvent latentes : comment obtient-on des distributions des degrés d'appartenance dans le cas de variables numériques ? Sur quelles connaissances sont-elles établies ? Sont-elles données a priori et mises à l'épreuve de la réalité ou bien sont-elles des construits ? S'il s'agit de ce dernier cas, quel processus d'extraction de connaissances à partir de données peut y conduire et quel type de règle peut-on alors extraire dans ce cadre ? Quelle signification peut-on donner à une règle associant deux sous-ensembles ou deux attributs flous ? On rejoint alors une des problématiques du data mining et de la qualité des règles.