

Chapitre 8 : Graphe de règles d'implication statistique pour le raisonnement courant. Comparaison avec les réseaux bayésiens et les treillis de Galois.

Martine Cadot

Université Henri Poincaré / LORIA, Nancy, France
Martine.Cadot@loria.fr
<http://www.loria.fr/~cadot>

Résumé. Les règles d'implication statistique ressemblent aux règles du raisonnement mathématique. Ce qui permet de les utiliser facilement pour raisonner sur les données. Toutefois, le modèle sous-jacent aux règles d'implication statistique n'est pas le modèle de la logique formelle utilisé en mathématique, mais un modèle statistique aboutissant à des relations approximatives. Contrairement au raisonnement mathématique, le raisonnement courant se satisfait de règles approximatives. Mais il a besoin d'un graphe pour savoir quels enchaînements de règles sont possibles car en faisant se succéder des approximations, on finit par arriver à des incohérences. On montrera dans ce chapitre comment fonctionne l'enchaînement de ces règles, notamment à travers la construction du graphe des règles d'implication tel que proposé dans les différentes versions de CHIC et on comparera ce modèle statistique des données à deux autres modèles proches : un modèle algébrique, les treillis de Galois, et un modèle probabiliste, les réseaux bayésiens. Pour permettre des comparaisons aisées, le fonctionnement des trois modèles sera illustré à l'aide d'un même jeu de données médicales librement disponible sur Internet.

1 Introduction

Les règles d'implication statistique ont été conçues à l'origine par Régis Gras dans sa thèse de mathématique (1979). Une règle d'implication statistique entre deux « notions » A et B s'exprime sous la forme « si A alors B » (ou $A \rightarrow B$) et indique que si l'élève a acquis la notion A alors il a acquis la notion B. La connaissance de telles règles aide l'enseignant de mathématique à organiser son cours pour faire assimiler un certain nombre de notions à ses élèves. En suivant le modèle de Régis Gras, l'enseignant peut établir un réseau de règles entre des notions (ou plutôt leur assimilation) sans avoir besoin de définir théoriquement ce qu'est chaque notion, du moment qu'il peut déterminer pratiquement si un élève donné l'a acquise ou non. En effet, les règles sont obtenues de façon statistique, par calcul à partir des résultats d'observations d'un ensemble d'élèves, à travers les résultats d'une interrogation écrite par exemple. La méthodologie de construction de règles d'implication statistique n'est pas restée cantonnée à la didactique car elle permet d'établir des liens de « cause à effet »