

Relational Learning as Search in a Critical Region

1 Contexte et problématique

Cet article s'intéresse à la calculabilité et la complexité de l'apprentissage relationnel. L'évaluation d'une hypothèse est NP-difficile au pire et peut être représentée sous forme d'une variable aléatoire dans un espace lié aux problèmes considérés. Trois classes de problèmes peuvent alors être distingués : les problèmes sous-contraints, les problèmes sur-contraints et ceux appartenant à une transition de phase abrupte entre les 2 ensembles où l'on peut observer un pic de complexité.

Les contributions apportées par l'article sont la mise en relation de la présence de la phase de transition avec la "réussite" de l'apprentissage relationnel, ainsi qu'une argumentation des résultats expérimentaux.

2 Démarche et résultats préliminaires

L'article cherche à évaluer la qualité de l'apprentissage pour un couple (m, L) en générant un ensemble de problèmes relationnels $\Pi_{m,L}$ composée d'une hypothèse h , conjonction de m prédicats, de 2 ensembles d'apprentissage et de test, E_L et E_T de 100 exemples positifs et de 100 exemples négatifs chacun sur L constantes, et en fournissant E_L à un algorithme d'apprentissage (FOIL, SMART, G-NET).

3 Résultats et Interprétations

L'évaluation repose sur 3 critères : la précision prédictive, l'identification du concept et le coût computationnel et permet de catégoriser les problèmes en trois classes de complexité : les problèmes faciles qui sont correctement résolus, les problèmes faisables qui sont approchés de manière efficace et les problèmes durs.

Par ailleurs, la phase de transition se comporte comme un attracteur de l'apprentissage, ce que l'on peut expliquer avec des arguments sur les critères d'arrêt des algorithmes basés sur la capacité discriminatoire des hypothèses qui est maximisée dans la transition de phase.

Les auteurs mettent également en lumière l'inadéquation des méthodes d'optimisation gloutonne de fonction de fitness pour les choix réalisés lors de l'apprentissage, dans les régions autres que la phase de transition. Cet état de fait rend compte de la difficulté d'apprendre sauf pour les problèmes sur-contraints.

4 Perspectives et Avis personnel

Les perspectives proposées sont l'utilisation d'algorithmes bottom-up, de connaissance à priori et enfin de reconsidérer les stratégies d'apprentissage afin de prendre en compte l'importance de la phase de transition.