

Nouveau modèle pour un passage à l'échelle de la θ -subsomption

Hippolyte Léger, Dominique Bouthinon, Mustapha Lebbah, Hanane Azzag

Universite Paris 13, Sorbonne Paris Cité, L.I.P.N
UMR-CNRS 7030 F-93430, Villetaneuse, France
{leger, bouthinon, lebbah, azzag}@lipn.univ-paris13.fr

Résumé. Le test de θ -subsomption, opération fondamentale en Programmation Logique Inductive (PLI) pour tester la validité d'une hypothèse sur les exemples, est particulièrement coûteux. Ainsi, les systèmes d'apprentissage de PLI les plus récents ne passent pas à l'échelle. Nous proposons donc un nouveau modèle de θ -subsomption fondé sur un réseau d'acteurs, dans le but de pouvoir décider la subsomption sur de très grandes clauses.

1 Introduction

La θ -subsomption est utilisée dans de nombreux systèmes de Programmation Logique Inductive (PLI) pour tester la validité d'une hypothèse sur les exemples. Une clause C θ -subsume une clause D si et seulement si il existe une substitution θ telle que $C\theta \subseteq D$. Malheureusement, la complexité temporelle de pire cas de la θ -subsomption est ($O(|D|^{|C|})$). De nombreuses recherches ont été menées pour créer des algorithmes de θ -subsomption efficaces (Ferilli et al. (2003); Kuzelka et Zelezný (2008); Santos et Muggleton (2010))¹. Cependant, le passage à l'échelle de la subsomption sur des plate-formes distribuées a reçu beaucoup moins d'attention et aucun système à notre connaissance ne se concentre sur la θ -subsomption. Notre but est de créer un modèle générique de θ -subsomption pouvant passer à l'échelle et être facilement intégré à des systèmes d'apprentissage relationnel en utilisant des plate-formes Big Data distribuées.

2 Préliminaires

Nous considérons ici la θ -subsomption entre deux clauses C et D où C et D sont des clauses de Horn définies sans fonctions, C contenant des variables et D des constantes. Une substitution est un ensemble fini $\{X_1/v_1, \dots, X_n/v_n\}$ où X_i est une variable et v_i est une constante (une variable apparaît une seule fois dans une substitution). Deux substitutions θ_1 et θ_2 ne sont pas compatibles si elles assignent deux valeurs distinctes à une même variable, par exemple $\theta_1 = \{Y/a\}$ et $\theta_2 = \{Y/b\}$. En revanche l'union de deux substitutions compatibles est toujours valide. Nous présentons ci-dessous, un exemple θ -subsomption utilisé tout au long de l'article.

1. Voir Ferilli et al. (2003) pour une étude plus approfondie de ces travaux