

Un algorithme multi-agent de classification pour la construction d'ontologies dynamiques

Kévin Ottens, Nathalie Aussenac-Gilles

IRIT, Université Paul Sabatier,
118 Route de Narbonne,
F-31062 TOULOUSE CEDEX 9
{ottens,aussenac}@irit.fr,
<http://www.irit.fr/>

Résumé. La construction d'ontologies à partir de textes reste une tâche coûteuse en temps qui justifie l'émergence de l'*Ontology Learning*. Notre système, *Dynamo*, s'inscrit dans cette mouvance, en apportant une approche originale basée sur une architecture multi-agent adaptative. En particulier, l'article présente le cœur de notre approche, un algorithme distribué de classification hiérarchique qui s'applique sur les résultats d'un analyseur syntaxique. Cet algorithme est évalué et comparé à un algorithme centralisé plus conventionnel. Forts de ces résultats, nous discutons ses limites et dressons en perspective les aménagements à effectuer pour aller vers une solution complète de construction d'ontologies.

1 Introduction

L'analyse syntaxique et terminologique de textes organisés en corpus est envisagée depuis une quinzaine d'années comme une issue possible au problème de la construction d'ontologies ou de terminologies. Non seulement la référence à l'utilisation de la langue est le moyen d'assurer une certaine pertinence aux connaissances ainsi identifiées, mais surtout, la disponibilité de textes sous format électronique autorise le recours à leur analyse par des logiciels de Traitement Automatique des Langues (TAL). Enfin, dans le cas particulier où l'ontologie à construire doit servir à indexer ou annoter des documents par des méta-données, le recours aux textes fournit une terminologie riche et au plus près de celle utilisée dans ces documents.

Plusieurs méthodes ont été définies pour guider ce processus et le rendre systématique, comme Archonte (Bachimont, 2004) ou Terminae (Aussenac-Gilles et al., 2003). Toutefois, ces méthodes se heurtent à la lourdeur du dépouillement des résultats des logiciels d'analyse syntaxique, lexicale ou sémantique des textes auxquels elles ont recours. De plus, toutes s'accordent à souligner le rôle majeur de l'analyste (que nous appellerons aussi par la suite ontologue) dans l'interprétation de ces résultats. Ainsi, lorsqu'un outil d'aide au repérage de relations selon une approche par patron présente des phrases contenant des termes en relation, il n'est pas sûr que cette relation doive être intégrée dans l'ontologie, ou encore qu'elle mette en relation exactement les concepts auxquels renvoient ces termes (Séguéla, 2001).

Aussi, depuis quelques années, plusieurs tentatives sont effectuées pour réduire la charge de l'analyste. Elles profitent des avancées de l'apprentissage automatique ou de méthodes de clas-