

Une mesure de similarité contextuelle pour l'aide à la navigation dans un treillis

Saoussen Sakji*, Marie-Aude Aufaure*, **
Géraldine Polaiillon*, Bénédicte Le Grand***, Michel Soto***

*Supélec – Computer Science department – plateau du Moulon – 3, rue Joliot Curie
F-91192 Gif sur Yvette Cedex

{Saoussen.Sakji, Marie-Aude.Aufaure, Geraldine.Polaiillon}@supelec.fr

** INRIA Paris-Rocquencourt – Domaine de Voluceau – Rocquencourt

B.P. 105 F-78153 Le Chesnay Cedex

Marie-Aude.Aufaure@inria.fr

***Laboratoire d'Informatique de Paris 6 – 104, av du P^{dt} Kennedy – F-75016 Paris
{Benedicte.Le-Grand, Michel.Soto}@lip6.fr

Résumé. La recherche d'information et la navigation dans les pages web s'avèrent complexes du fait du volume croissant des données et de leur manque de structure. La formalisation conceptuelle d'un *contexte* associé à une ontologie rend possible l'amélioration de ce processus.

Nous définissons un *contexte conceptuel* comme étant l'association d'un treillis de concepts construit à partir de pages web avec des ontologies. La recherche et la navigation peuvent alors s'effectuer à plusieurs niveaux d'abstraction : le niveau des données, le niveau conceptuel et le niveau sémantique. Cet article s'intéresse essentiellement au niveau conceptuel grâce à une représentation par les treillis de concepts des documents selon les termes qu'ils ont en commun. Notre objectif est de proposer une mesure de similarité permettant à l'utilisateur de mieux naviguer dans le treillis. En effet, une bonne interprétation du treillis devrait passer par un choix rigoureux des concepts, objets, relations et propriétés les plus intéressants. Pour faciliter la navigation, il faut pouvoir indiquer à l'utilisateur les concepts les plus pertinents par rapport au concept correspondant à sa requête ou pouvoir lui proposer un point de départ. L'originalité de notre proposition réside dans le fait de considérer un lien sémantique entre les concepts du treillis, basé sur une extension des mesures de similarité utilisées dans le cadre des ontologies, afin de permettre une meilleure exploitation de ce treillis. Nous présentons les résultats expérimentaux de l'application de cette mesure sur des treillis construits à partir de pages web dans le domaine du tourisme.

1 Introduction

L'objectif de nos travaux est de faciliter la recherche d'information dans des pages Web par l'utilisation conjointe de treillis de Galois et d'ontologies, qui constitue ce que nous appelons un « contexte conceptuel ». Les regroupements conceptuels fournis par les treillis, associés aux liens sémantiques de l'ontologie, permettent d'améliorer la recherche d'information en fournissant des niveaux de navigation plus abstraits et complémentaires.