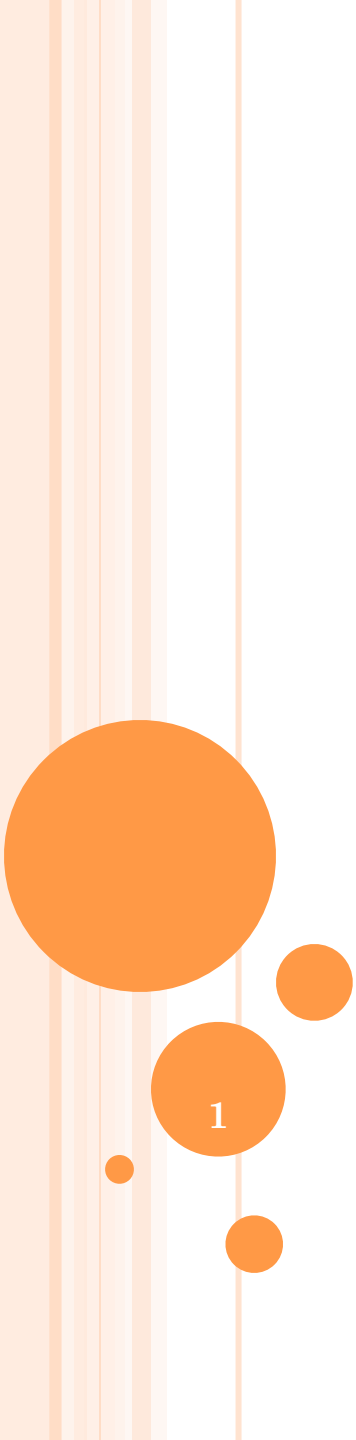




CHIHEB TRABELSI

SUJET DE PRÉSENTATION : Ontology Matching

Jérôme Euznat

Pavel Shvaiko

PLAN

- Introduction
- Le problème d'alignement
- Classification des techniques d'alignement d'ontologie
- Les techniques de bases
- Méthodes d'apprentissage
- Conclusion

INTRODUCTION

- **ontologie** : ensemble structuré des concepts d'un champ d'informations.
- Dans un système évolutif «Web Sémantique »



Contrainte : combiner différentes ontologies.

- Exemple : une bibliothèque universitaire et un site commercial de vente de livre



Hétérogénéité et différence

INTRODUCTION

- Intention d'interagir : Passer une commande de livres au site internet ou numériser les livres de la bibliothèque



Nécessité de faire la correspondance

- Un *livre* correspond à un *Volume*.
- Le *prix* est multiplié par le taux du tax.
- Trouver les correspondances  Alignement d'ontologie.

INTRODUCTION

- L'alignement d'ontologie : solution à l'hétérogénéité.
- Les types de correspondances : équivalence, inclusion, exclusion... entre les entités.
- L'alignement d'ontologie permet l'interopérabilité des entités alignées.
- Comment est-ce que l'alignement est effectué?
- Quelles sont les techniques utilisées?

LE PROBLÈME D'ALIGNEMENT

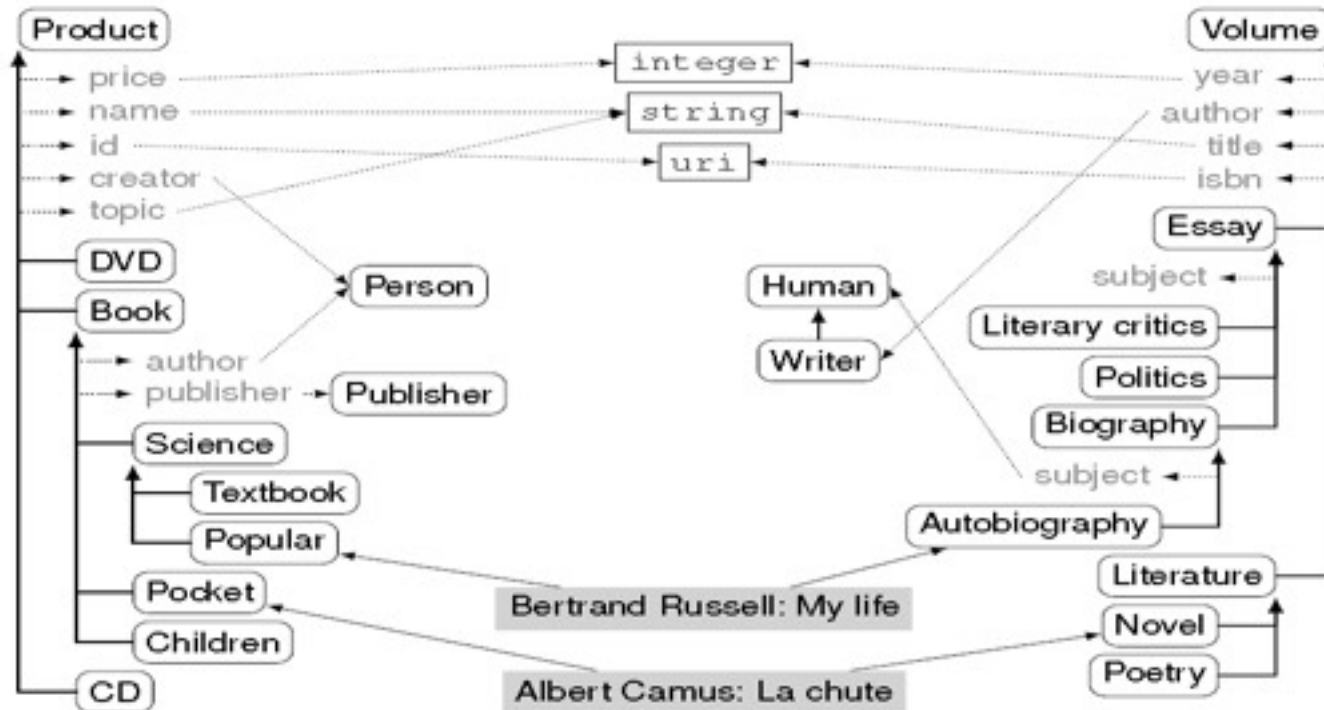


Figure 1 : Fragments de deux ontologies

- Classes : « *Book* », « *Person* ».
- Individus : « *Bertrand Russel : My life* ».
- Relations : « *creator* », « *topic* ».
- Types de données : « *String* », « *Integer* ».
- Valeurs : « *'My life'* ».

LE PROBLÈME D'ALIGNEMENT

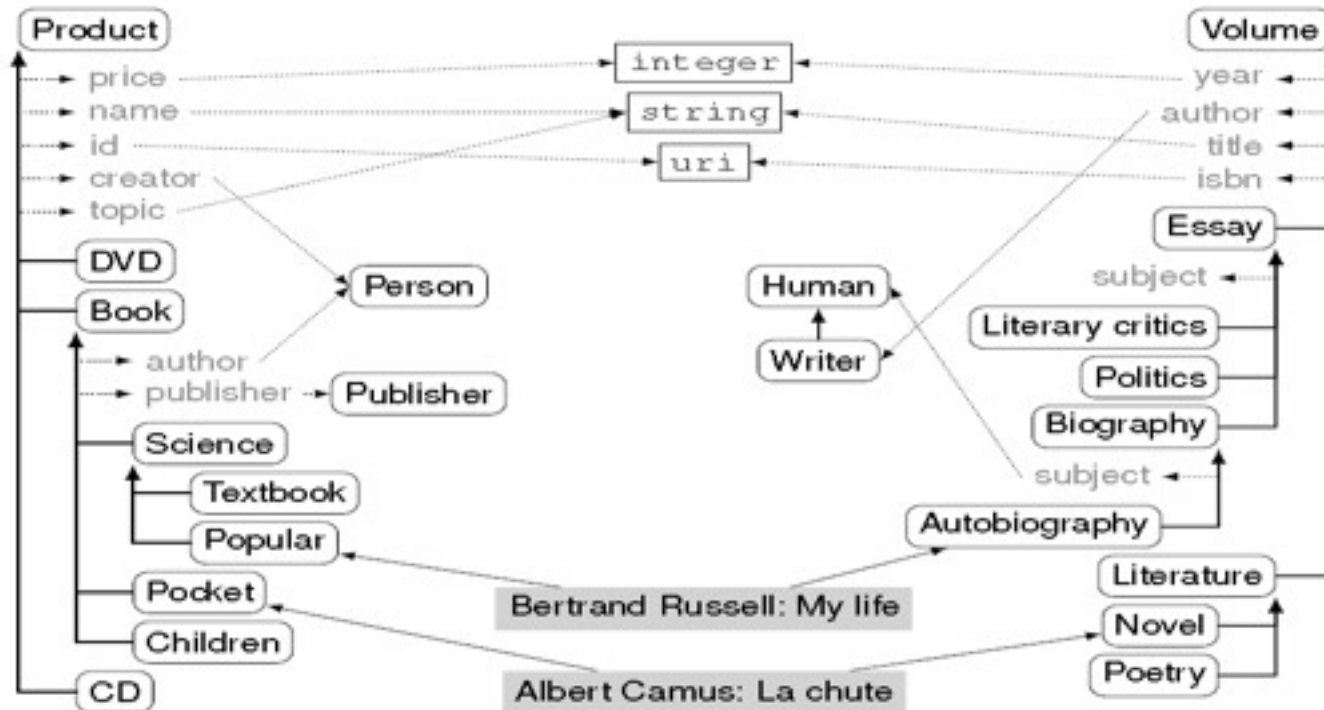


Figure 1 : Fragments de deux ontologies

- La classe « *Book* » est une **spécialisation** de la classe « *Product* ».
- La classe « *Product* » peut être déclarée comme **exclusive** de la classe « *Person* ».
- « *Bertrand Russel : My life* » est une **instance** de la classe « *Popular* ».

LE PROBLÈME D'ALIGNEMENT

- But de l'alignement d'ontologie : Réduire l'hétérogénéité.
- Types d'hétérogénéité :
 - L'hétérogénéité syntaxique : Langage d'ontologies différents.
 - L'hétérogénéité terminologique : « Paper » vs « article ».
 - L'hétérogénéité conceptuelle : Axiomes de définitions différents.

LE PROBLÈME D'ALIGNEMENT

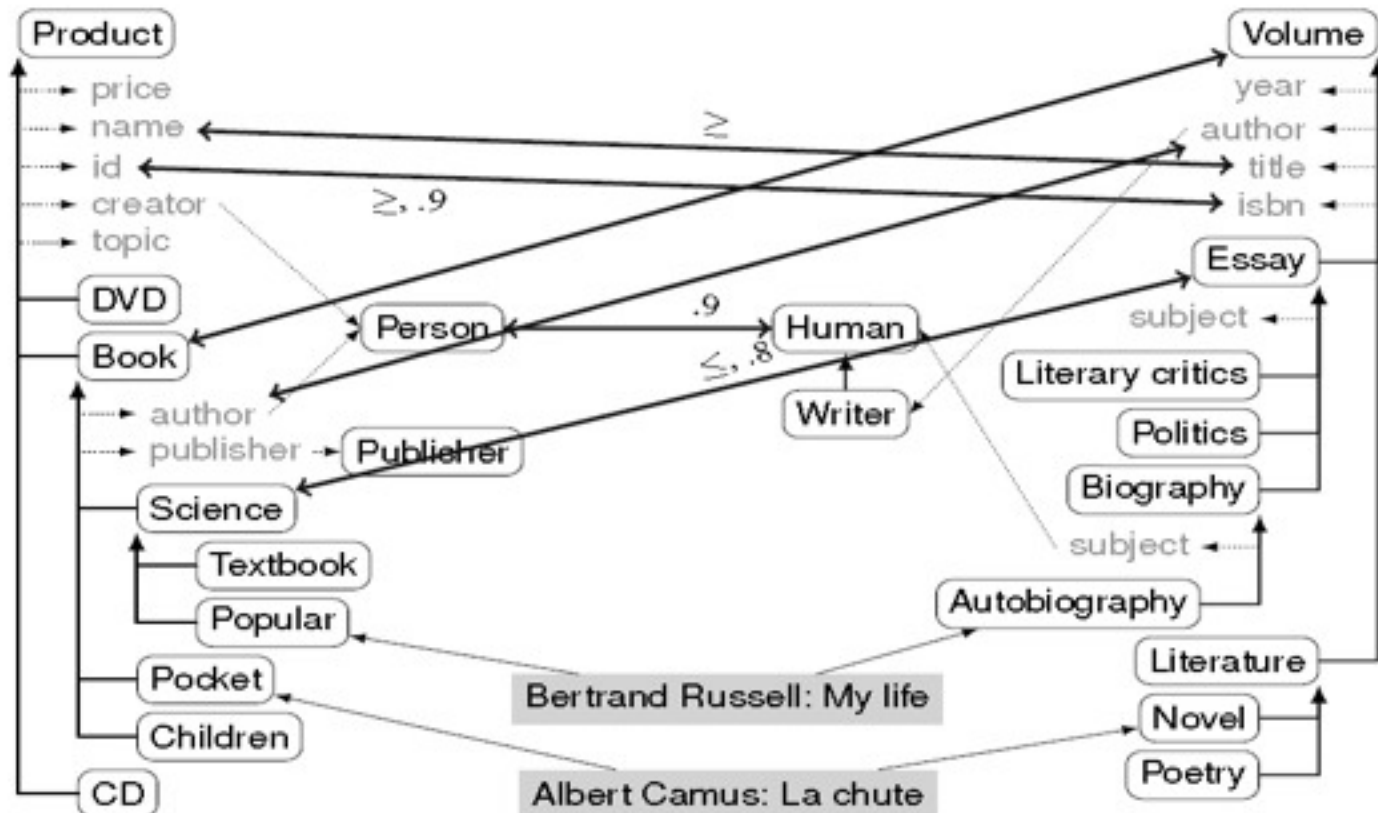


Figure 2 : L'alignement des deux ontologies

Book $=_{1.0}$ Volume
 id $\geq .9$ isbn
 Person $=_{.9}$ Human

name ≥ 1.0 title
 author $=_{1.0}$ author
 Science $\leq .8$ Essay

CLASSIFICATION DES TECHNIQUES D'ALIGNEMENT D'ONTOLOGIE

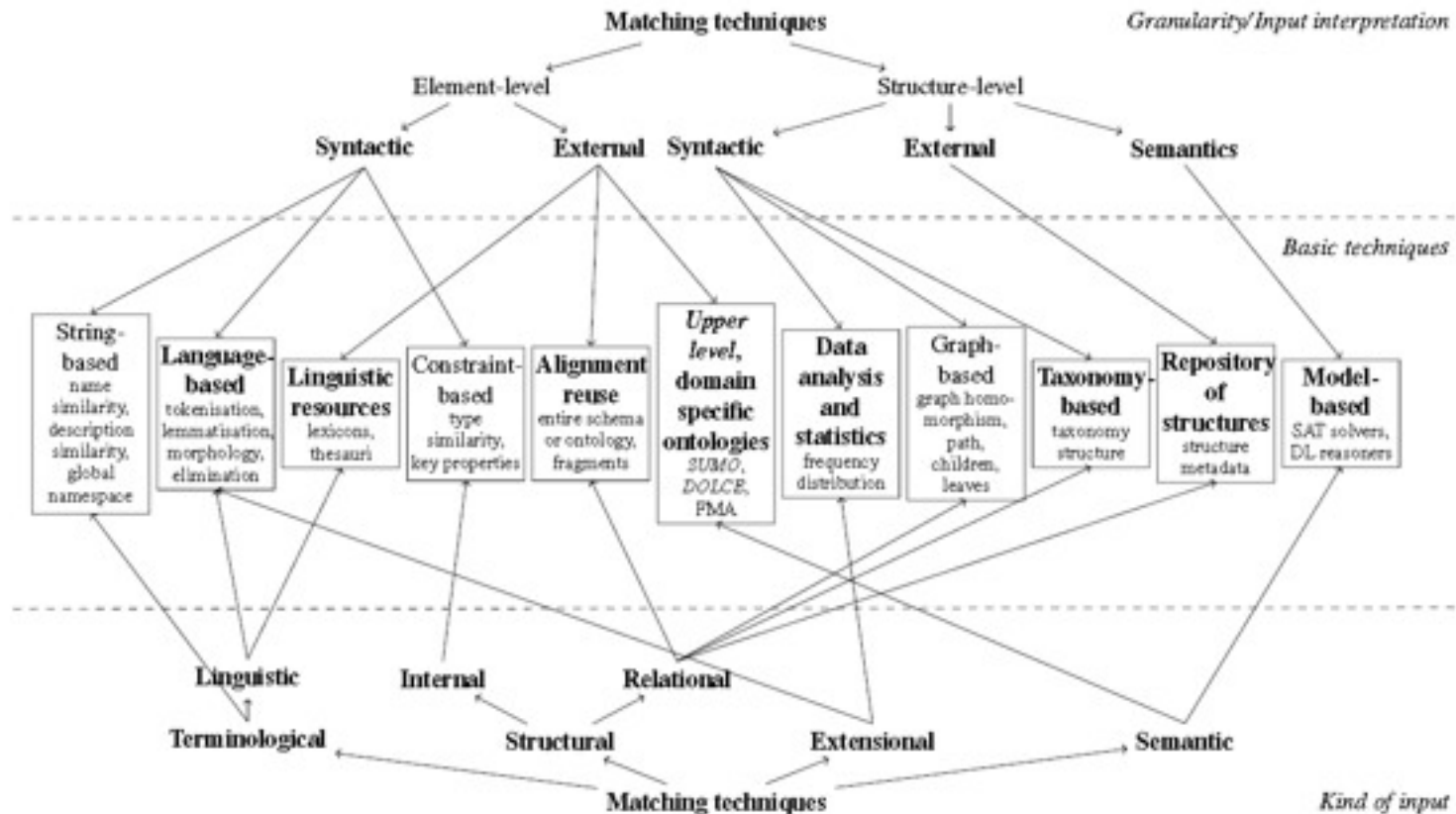


Figure 3 : Les types de techniques d'alignement

LES TECHNIQUES DE BASE

- Les techniques de base  trouver les relations entre entités.
- Les relations d'équivalence  mesure de la similarité.

$$\forall x, y \in o, \sigma(x, y) \geq 0$$

- Similarité $\forall x \in o, \forall y, z \in o, \sigma(x, x) \geq \sigma(y, z)$
 $\forall x, y \in o, \sigma(x, y) = \sigma(y, x)$

- Dissimilarité normalisée $\bar{\delta} = 1 - \bar{\sigma}$

- La similarité est une mesure métrique.

LES TECHNIQUES DE BASE

- Méthodes basées sur les chaînes de caractères :
 - La comparaison entre les chaînes est utile si on utilise des chaînes similaires pour représenter le même concept.
 - Si on utilise des synonymes mais que les structures des ontologies est différentes, on obtiendra une faible similarité entre les entités.
 - On risque d'obtenir de faux positives si deux chaînes sont similaires et qu'ils ne dénotent pas le même concept.

LES TECHNIQUES DE BASE

- Méthodes linguistiques :

- Ressources disponibles  Correspondances possibles .
 - Plusieurs correspondances potentielles pour une entité.
-  Prendre en compte la structure des entités de l'ontologie.

LES TECHNIQUES DE BASE

- Techniques basées sur les structure interne :
 - Fournissent des techniques de comparaisons efficaces et faciles à implémenter.
 - Pas beaucoup d'information sur l'entité à comparer.
-  Doivent être utilisées avec d'autres techniques pour remédier au problème.

LES TECHNIQUES DE BASE

- Techniques basées sur les structures relationnelle :
 - Permet de prendre en considérations toutes les relations entre les entités.



- Structures relationnelle puissante.

LES TECHNIQUES DE BASE

- Techniques extensionnelles :
 - Fournissent une information indépendante de la partie conceptuelle de l'ontologie.
 - Ne fonctionne pas dans le cas où les instances ne sont pas disponibles.
- Techniques Sémantiques :
 - Ne sont pas encore bien développés.
 - Ne permettent pas à elles seules d'identifier les correspondances.

LES MÉTHODES D'APPRENTISSAGE

- Le réseau de neurones :

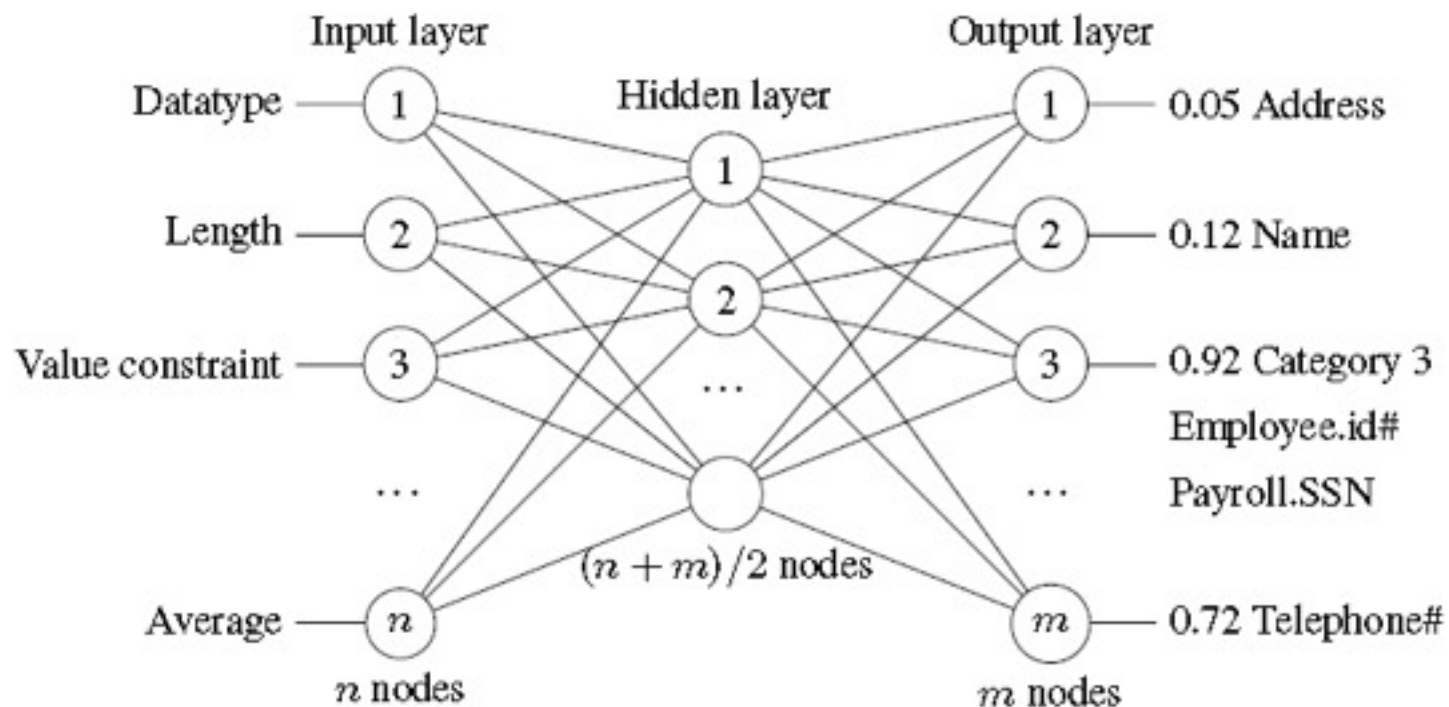


Figure 4 : Réseau de neurones multicouches pour l'alignement d'ontologies

LES MÉTHODES D'APPRENTISSAGE

- Arbre de décision :

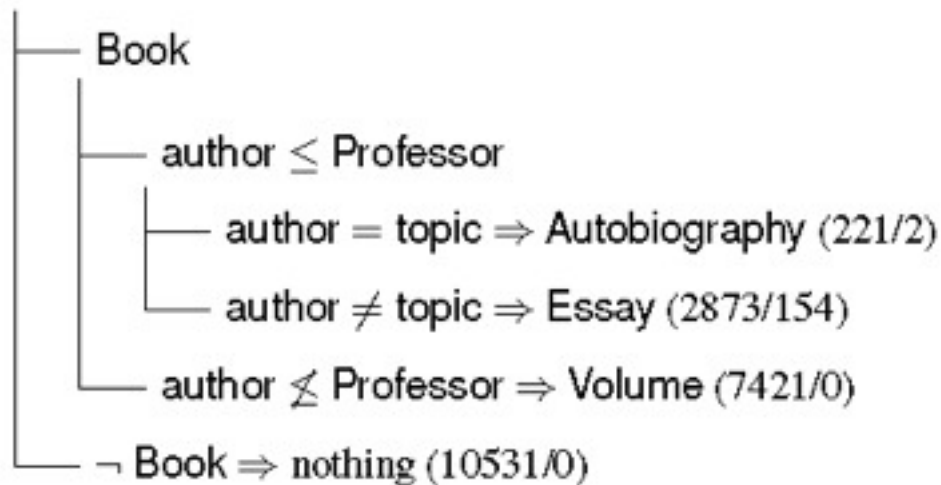


Figure 5 : Fragment d'un arbre de décision binaire pour l'alignement d'ontologies

- Générer des règles pour identifier de nouvelles instances.
- Alignement de livres de la première ontologie à la deuxième.

CONCLUSION

- Malgré le progrès satisfaisant, l'alignement d'ontologies semble être impossible.
- Il faut comprendre la véritable signification des concepts.
- La véritable signification est dans la tête des développeurs des ontologies.

QUESTIONS



MERCI



Questions?