

Fictions et connaissances

Serge Abiteboul

INRIA & ENS Cachan

Conseil national du numérique & Académie des sciences



La vérité sur le Web

- Erreurs
- Exagérations, mensonges, omissions, manipulations
- Fiction
- Délires, folies
- Quelle vérité ?
 - Qui est le gouvernement légitime en Syrie ?

La fiction sur le Web

- Exemple : la fan fiction
 - Une fanfic, est un récit que des fans écrivent pour prolonger, amender ou même totalement transformer un roman, un manga, une série télévisée, un film, un jeu vidéo...
- D'autres références que dans le monde réel
 - Personnages, lieux....
- Extraire les connaissances du Web, c'est aussi retrouver ces références

Démarrons plus modestement

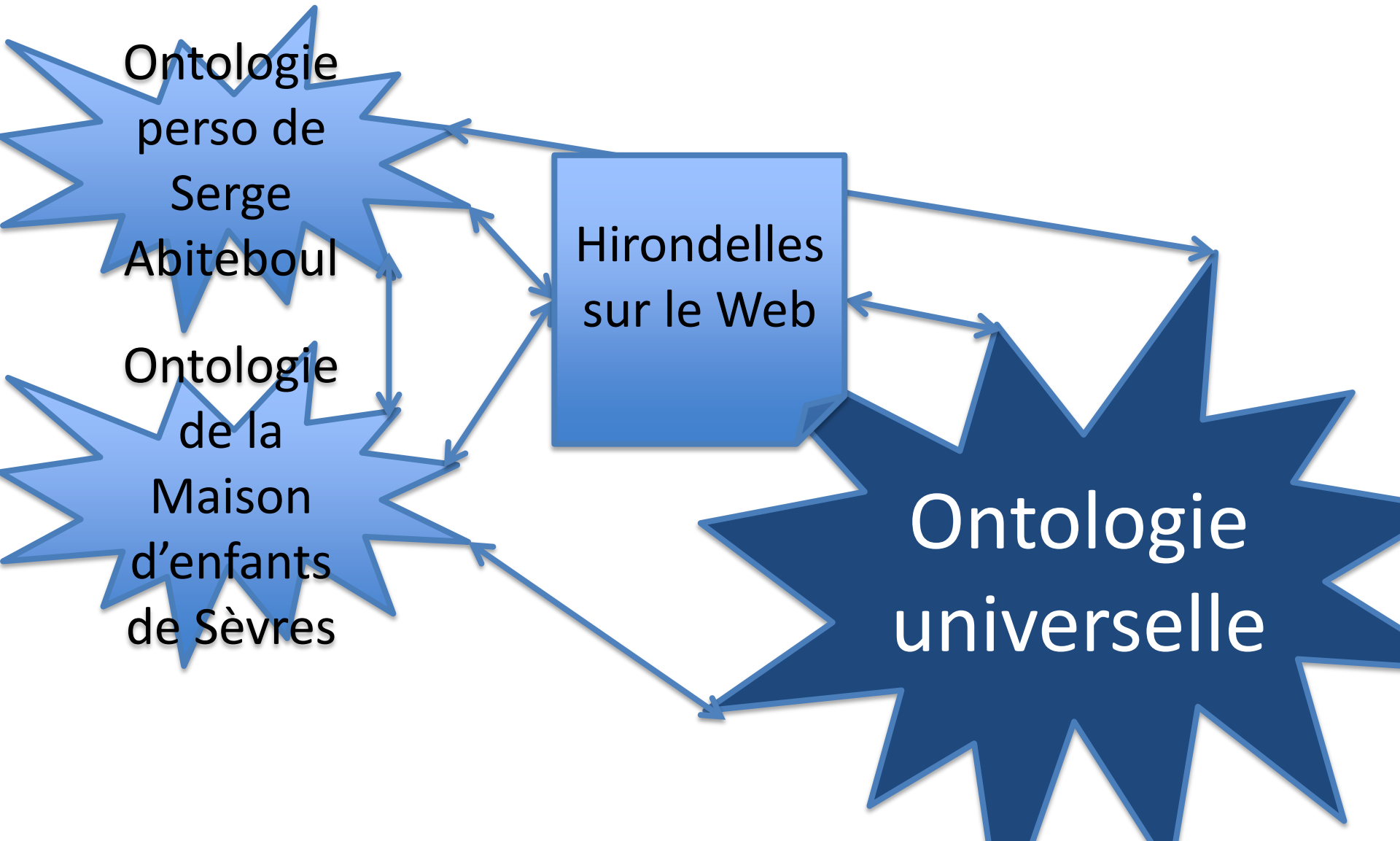
Extraire les connaissances d'un roman

Le mélange banal : réalité et fiction

Exemple : Hironnelles sur le Web

- Des lieux, des personnages, des événements vrais
 - Autour de la ville de Sèvres
 - Autour de l'histoire oubliée de Simone Hagnauer – Hironnelle
- La fiction
 - Une énigme policière inventée
 - Personnages, événements, lieux fictifs
 - La fin du roman se passe dans le futur

Contexte du roman: plusieurs ontologies



Crowdsourcing par les lecteurs

- Personnage fictif
- Personnage réel – voir la thèse de ...
- Probablement Place du Général de Gaule...
- Pourrait être Yann F. de l'ontologie de SA
- ...

Construction automatique

- Aligner les différentes ontologies
- Aligner les entités/rerelations du texte avec ces ontologies
 - La ville de Sèvres
 - Le directeur d'Espace dans le roman, dans la réalité, dans les contacts de Serge A.
- Résoudre la localisation
- Résoudre la temporalité
- Séparer le réel de la fiction
 - Personnages réels ou pas
 - Evènements réels ou pas
 - Lieux réels ou pas

Difficulté: précision et rappel (recall)

- Exemple « trouver toutes les personnage de romans »
- Trop de précision : on rate un animal qui joue un rôle essentiel
- Trop de témérité : on obtient l'arme du crime, puis les armes en général, les téléphones portables, etc.

Difficultés

- Il n'est pas simple de construire/vérifier une nouvelle logique en monde ouvert
- En monde ouvert, tout nouveau fait positif est accepté dans la théorie
- Les gens donnent peu de fait négatifs :
aujourd'hui je ne suis pas à la plage, je ne vais pas chez le coiffeur, je ne rencontre pas Alice, etc.

Difficultés

- Gérer les contradictions, imprécisions, l'incomplétude
- Probabilités : ce personnage pourrait être Yann F. du monde réel avec une assez forte probabilité
- Mondes possibles : le coupable du deuxième meurtre pourrait être M. X ou M. Y – le roman laisse le doute
- Mondes parallèles : dans le roman, Tadeo Ando a construit un musée sur l'Île Seguin
 - Ce n'est pas juste faux, c'est un autre monde

Séparer le vrai du faux

- Contradictions
 - Elvis Presley est toujours vivant
 - Vote
 - Tenir compte de l'expertise
- Vraisemblance
 1. Obama vient d'arriver à Namur
 2. Obama était à Namur hier
 3. Obama va assister au prochain cours
- Tenir compte du contexte
 - Dans le monde de Harry Potter (J. K. Rowling), on peut voler avec un balai

Dépend de tout un monde logique

Séparer le vrai du faux - fiction

Roman

- On sait que c'est un roman
- Donc par défaut les personnages sont inventés
- Mais ils sont aussi « inspirés »
 - Essayer de trouver des proximités avec des personnes vraies



Sur la toile : on aimerait pouvoir flécher les sites « fiction »

Des dernières difficultés pour la route

- L'ontologie peut évoluer
 - Les identités peuvent changer
 - Des pays fusionnent (Allemagne), se découpent (Yougoslavie)
 - Les lois changent
 - En France, un mariage unit un homme et une femme
- Des ontologies peuvent elles-mêmes contenir des erreurs
 - Les personnes sont des véhicules

