

# Dictionaries merger for text expansion in question answering

Bernard JACQUEMIN

b\_jacquemin@yahoo.fr

COLING 2004

Workshop "Enhancing and using electronic dictionaries"

Genève 29th August 2004

# Introduction

- Current Question Answering approaches :
  - query expansion
  - search engine using significant words of the query
  - particular refinement techniques
- Needs : expansion elements
  - synonyms
  - derivatives
  - taxonomy
  - ...
- Lexical resources are needed
- Variable characteristics according to the QA method

# Introduction – Example

Question : *A quel chef Domitien succède-t-il ?*

(Which leader does Domitian succeed ?)

Query expansion :

	cuisinier		remplacer	
	général		suivre	
A quel	chef	Domitien	succède	– t – il ?
	tête		relayer	
	commandant		successeur	

Precise needs of QA vary with the methodology

# Plan

- My QA method and its particular needs
  - semantic disambiguation
  - consistent dictionaries
- How to make dictionaries compatible
  - synonyms
  - derivatives
  - taxonomy
- Experienced problems
- Conclusion

# "Meaning-oriented" QA

Classic query expansion

	<i>cuisinier</i>		<i>remplacer</i>
	<i>général</i>		<i>suivre</i>
A quel	chef	Domitien	succède – t – il ?
	<i>tête</i>		<i>relayer</i>
	<i>commandant</i>		<i>successeur</i>

# "Meaning-oriented" QA

Word Sense Disambiguation applied before expansion

	<i>cuisinier</i>		<i>remplacer</i>
	<i>général</i>		<i>suivre</i>
A quel	chef	Domitien	succède – t – il ?
	<i>tête</i>		<i>relayer</i>
	<i>commandant</i>		<i>successeur</i>

A more precise expansion reduces the noise

The applied WSD rule

- indicates a sense number in a dictionary
- selects expansion elements corresponding to the sense

# Dictionary needs of QA

- Relating to WSD
    - sharing out of the lexical data by SENSE
    - as many differential data as possible
  - Relating to the utterance expansion
    - as many expansion elements as possible
    - expansion elements are synonyms, derivative and taxonomy
    - expansion elements shared out by sense following WSD
- ⇒ More than one dictionary is necessary
- reference dictionary for WSD : *Dubois' French dictionary*
  - expansion dictionaries : *EWN, Memodata, Bailly*

# (In)compatible dictionaries

- Synonyms
  - sharing out of the synonyms has to follow the WSD senses
  - dictionaries have varied structures

⇒ deletion of proposed synonyms semantically different from the current sense (Dubois' semantic classes)

*Entry :*

succéder (1)

law

change (fig)

*Proposed synonyms :*

relayer

sport/sociology

change

**remplacer**

**law/sociology**

**change**



# (In)compatible dictionaries

- Derivatives

- Dubois' indicates derivation kind by sense
- some tools generate derivatives from a word

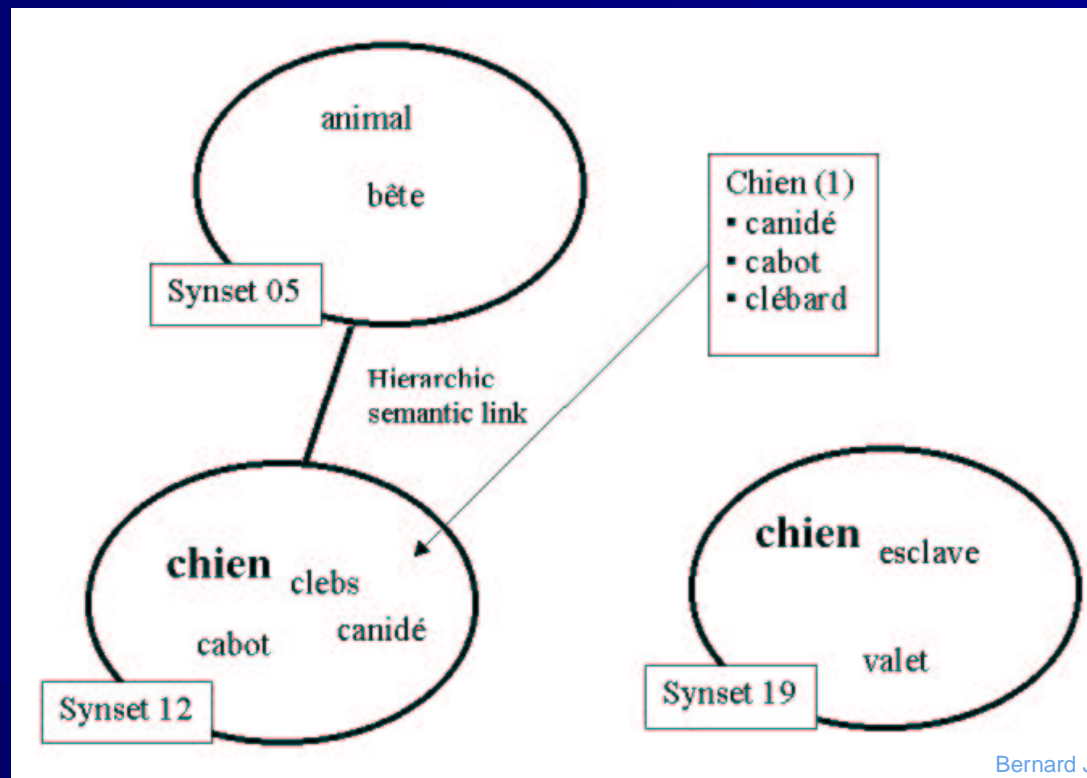
⇒ filtering and sharing out of proposed derivative by Dubois' data

For "succéder" (to succeed sb)

Proposed derivative	Dubois' instructions	corresponding sense
successeur	name -sseur	law sense (nb 1)
succession	name -ssion	all senses
<b>succion</b> (sucking)	—	—

# (In)compatible dictionaries

- Taxonomy
  - EWN hierarchies : hyperonymy and holonymy between synsets
  - a word in a given sense belongs to a synset if its synonyms in this sense belong to the synset



# Experienced problems

- Synonyms
  - no sense sharing out for multiword expressions
  - no evaluation
- Derivatives
  - the derivation tool cannot generate short words
  - the derivation tool uses only suffixes (no prefixes)
- Taxonomy
  - not yet implemented
  - EWN covers only a small part of the lexicon

# Conclusion

- The dictionaries are imperfect and unfinished
- QA has considerable needs of synonyms and derivatives for text expansion

⇒ Use of many dictionaries

This article presents an original method to make several dictionaries semantically consistent in order to expand text for question answering

Thank you for your attention  
Any questions ?