

*Individual Project for Problem Seminar (CSC 400),
University of Rochester, Fall 2007
sAdvisor: James F Allen*

Semantic Search on Help Files

Naushad UzZaman
University of Rochester
17 December 2007



Introduction

- Searching technology improved a lot in last several years
- Searching in help file?
- What about a semantic search?

Problem

- Test application: Browser.
 - Firefox and Internet Explorer
- Query: open a web page
- Internet Explorer:
 - "To disable a browser add-on"
 - In description it has, "Open an Internet Explorer".
 - In notes it has, "Some Web pages, or Internet Explorer...".
- Firefox:
 - No result!

Example

- i. opening a bank account
 - ii. open a bank account
 - iii. create a bank account
 - iv. modify loan information
-
- Titles:
 - *"Opening a bank account"*
 - *"Edit loan information"*

Solution?

- Semantic search as solution!
 - Need understanding of texts
 - Create and open can be used for same word
 - Not always feasible for web: extra overhead of representing the documents in semantic form

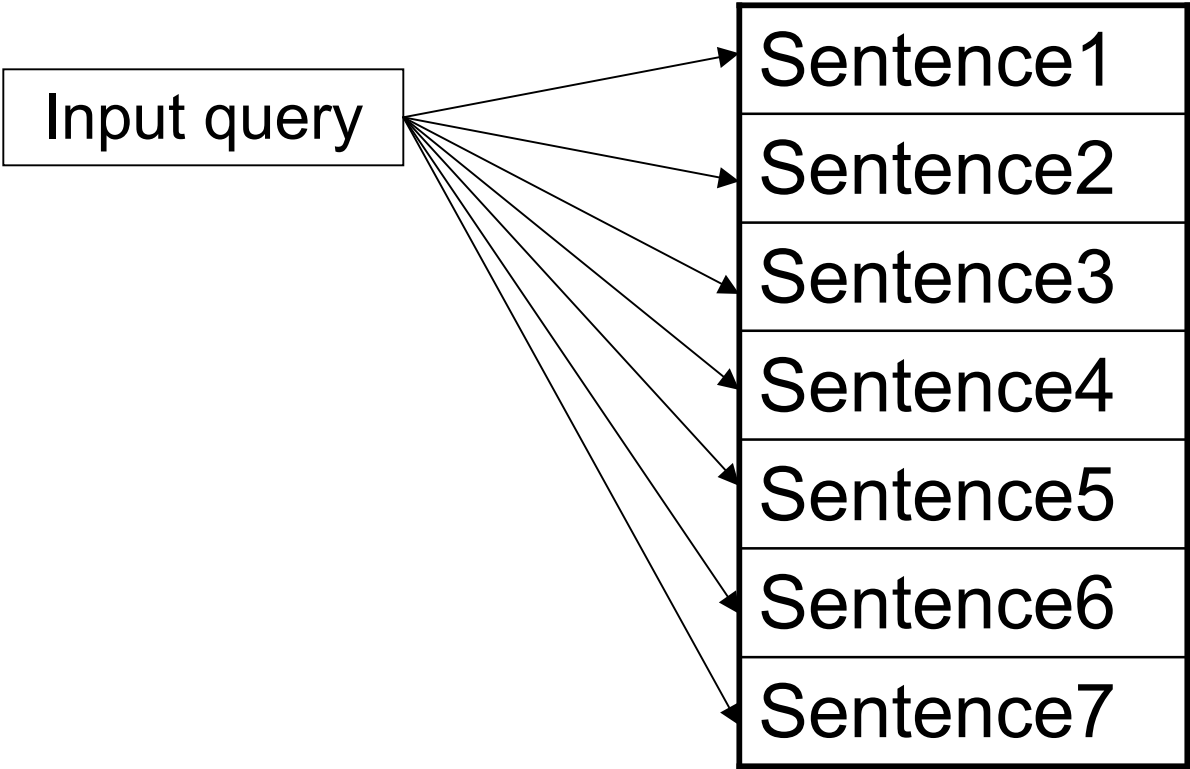
Semantic concepts

- Logical Form (LF): is the description of semantic content that can be mapped to a meaning via contextual interpretation, but not the meaning.

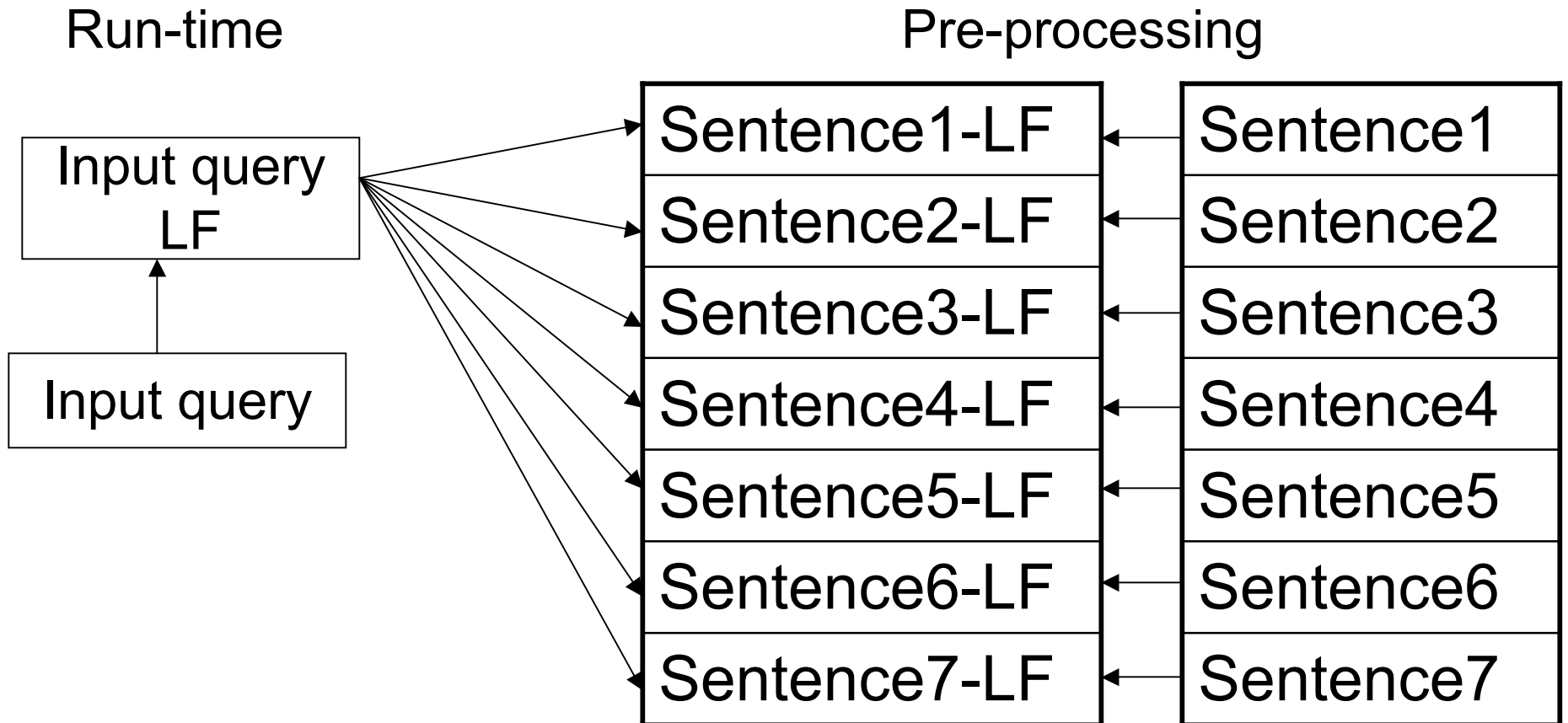
But the man wants to eat it

((LF::SPEECHACT V107605 W::SA_TELL :CONTENT
V107499 :MODS (V107426))
(LF::F V107426 (:* LF::CONJUNCT W::BUT) :OF V107605)
(LF::F V107499 (:* LF::WANT W::WANT) :ACTION V107537
:EXPERIENCER
V107492 :TMA ((W::TENSE W::PRES)))
(LF::THE V107492 (:* LF::MALE W::MAN))
(LF::F V107537 (:* LF::CONSUME W::EAT) :THEME
V107545 :AGENT V107492)
(LF::PRO V107545 (:* LF::REFERENTIAL-SEM W::IT)
:CONTEXT-REL W::IT))

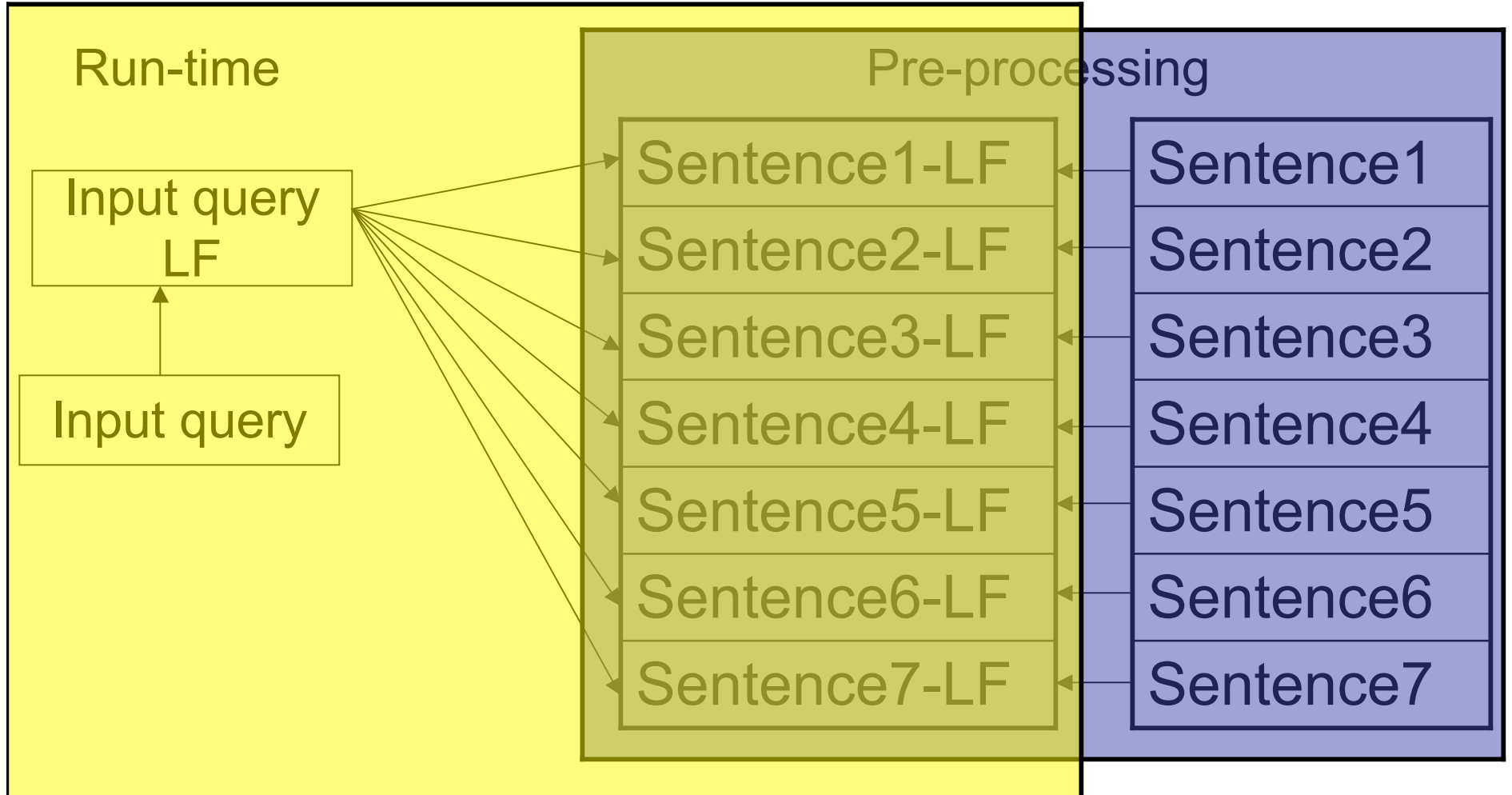
Normal Search



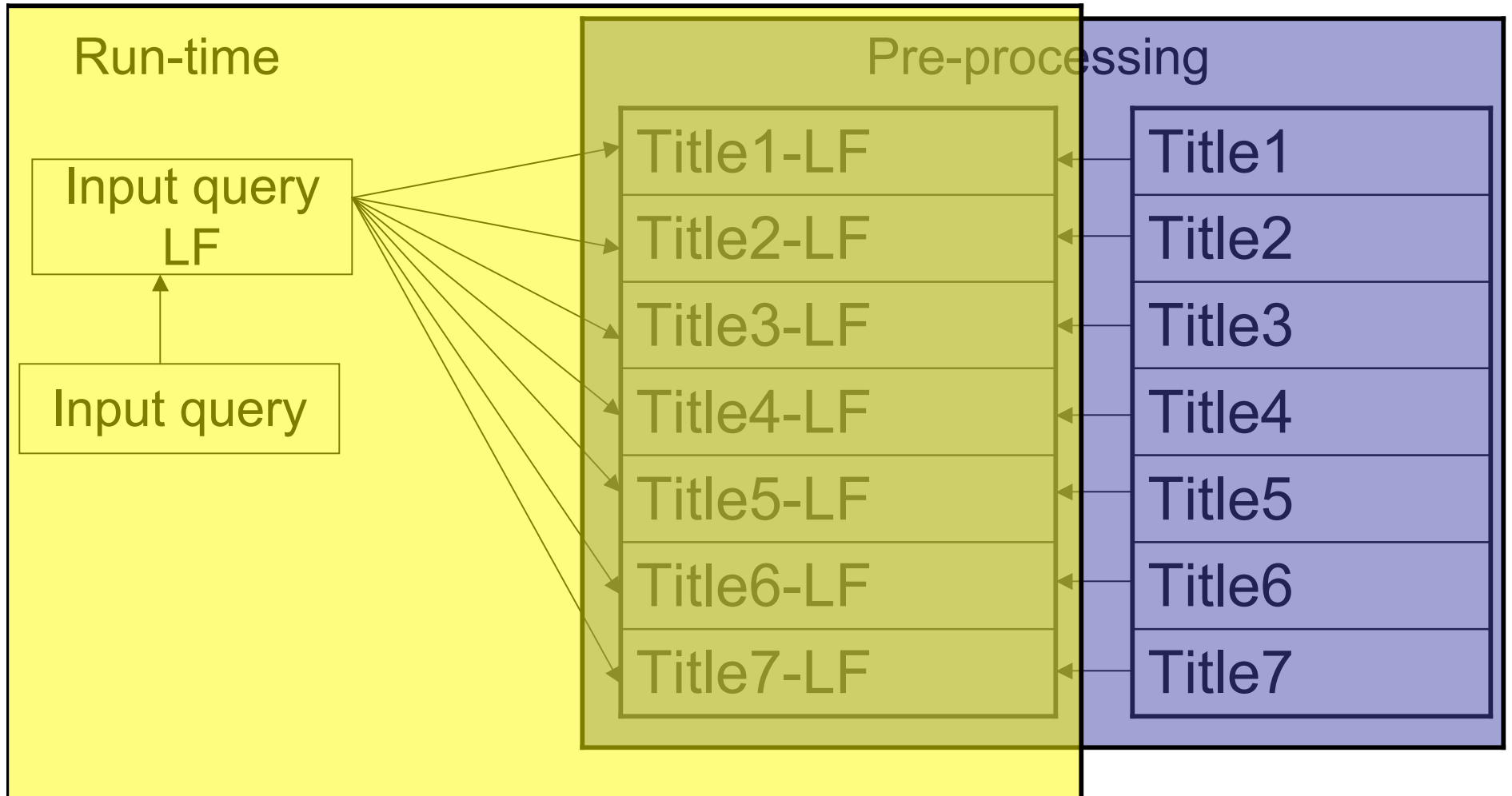
Semantic Search



Semantic Search



Semantic Search on Help File



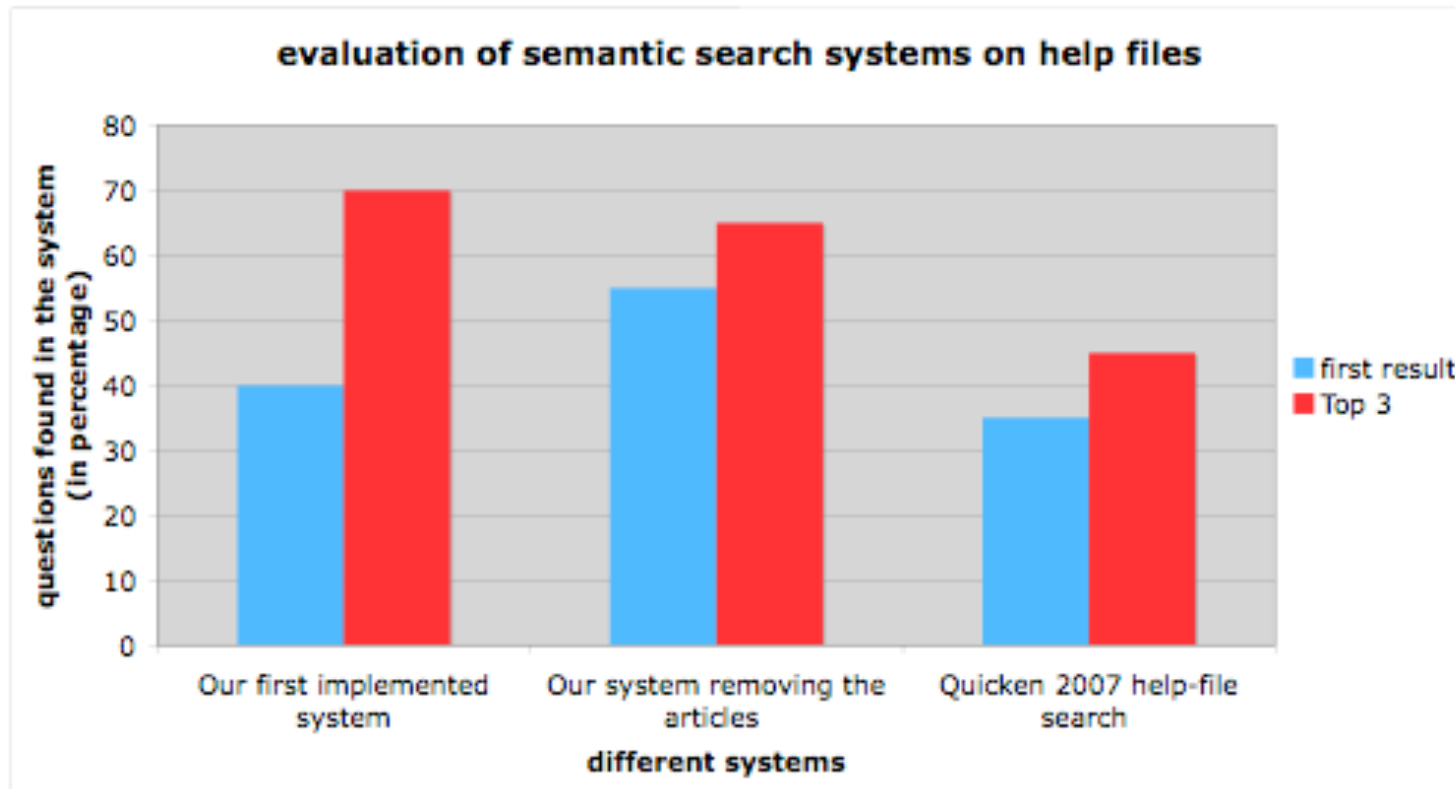
Experiment

- Three systems
 - Our first implemented system
 - Our system removing the articles
 - Quicken 2007 help-file search

Performance

System	First result	Top 3
Our first implemented system	8 (40%)	14 (70%)
Our system removing the articles	11 (55%)	13 (65%)
Quicken 2007 help-file search	7 (35%)	9 (45%)

Performance



Still Problems!!

- **Cannot handle substring:**
 - *“edit category”*
 - *“Editing a category or subcategory”*
- **Different ordering:**
 - *“set up web connect account”*
 - *“Setting up an account for Web Connect”*

Future Work

- Hybrid solution
 - Current Semantic Search
 - Keyword matching

Conclusion

- Easy, but neglected, area of software, search engine for help file
- Prototype on Quicken
- Experimented on 20 queries
- Our system perform better than Quicken
- More opportunity to improve by a hybrid solution

Thank you