

# Chunking-based Question Type Identification for Multi-sentence Queries

Mineki Takechi(Fujitsu Ltd.)  
 Takenobu Tokunaga(TITECH)  
 Yuji Matsumoto(NAIST)  
 July 27th, 2007 Amsterdam.  
 SIGIR 2007 Workshop Focused Retrieval

- Open domain question-answering: e.g., TREC,NTCIR

Query type in TREC and NTCIR

- Mainly factoid and single sentence questions  
 e.g., “When did the Jurassic Period end?”
- Non-factoid and multi-sentence queries(MSQs)  
 studied not enough

- The Web question-answering service: e.g., FMWORLD, MSN, Oshiete! goo

Queries in Web question-answering services

- Multiple sentences and multiple questions
- Factoid and Non-factoid questions mixed in a query

## Example Query in Web Q&A services

s1: Even when I sleep enough every night, I'm very tired all day.

Question A

s2: My friends tell me that these symptoms resemble depression, but what is the definition of depression?

s3: In my office, I have no time to relax because of my post.

Question B

s4: My wife is concerned about my recent condition and recommends that I see the doctor.

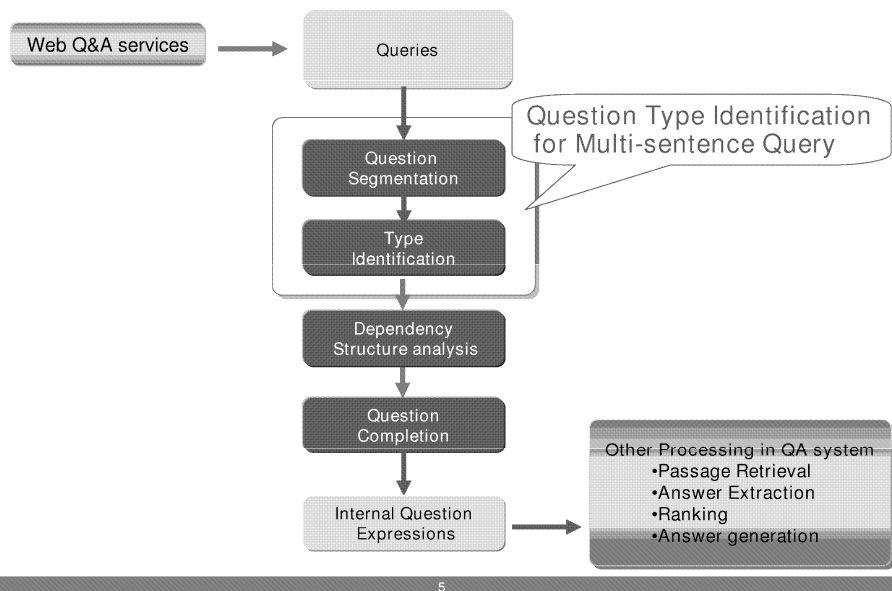
s5: I'm not sure whether I should see the doctor.

s5': Should I see the doctor?

s6: Please tell me your advice if you have same experience.

## Contents

- Query processing and annotation
- Proposed technique
- Evaluations
- Summary



- Identifying all questions in MSQs
  - Extract questions described in natural languages and identify their (semantic) question types.
- Various question types proposed
  - Maily address single sentence queries
  - Depending on tasks and based on answer types
  - TREC/NTCIR : Name, Location, Organization...
  - QA services[Kurohshi2000][Matsui2003]: How-to, Request, ...
- Not many studies of question types in the Japanese Web Q&A services

- “Oshiete goo!”: Web Q&A service, 2234 queries(in 2001,2006)
- Assign question types to multi-sentence queries; four factoid and five non-factoid types

Yes-No	Yes-No questions	Do your browser show “save” tab?
How-to	Ask methods to solve problems	How is the software installed?
Reason	Ask a reason	Why does it require the optional packages?
Evaluation	Ask sentiments or opinions	How do you feel new digital camera released by A Limited?

- Check inter-annotator agreement of question type annotation

- Queries including multiple questions: 56%(1252/2234 articles)
  - Dominantly exist in real queries

Yes-No question	Yes-No	43%
Non-factoid	How-to, Evaluation, Reason	19%
Mix. of factoid and non-factoid	Description, Consultation	19%
factoid	Name, Location, Time	19%

} Related to non-factoid: 38%

- Inter-annotator agreement on the F-measure(MUC98) without sentences including multiple question items
  - High agreement : Yes-No, Location, How-to : 0.8
  - Low agreement : Description : 0.5

- Many non-question sentences: greetings and apologies
- Ellipses and reference expressions often exist in question

“After catching a cold, many small, white dots appeared on the downside of my tongue. What do that seem to be?”

- To know the antecedent of “that” and semantic class of “what”, the first sentence is needed.

- One question described in multiple sentences often appear

“Plants that can grow in indoor or small beranda. Additionally, I wants them to be easy to care for beginner like me, so do anyone know such plants?”

- The sentence including multiple questions is existed but few and mostly stereotyped
  - Excluding these cases in the current stage of study

- Required new unit of question type identification
  - Questions in MSQs often described by multiple sentences

### Question segment

- A minimal set of sentences to identify a question type

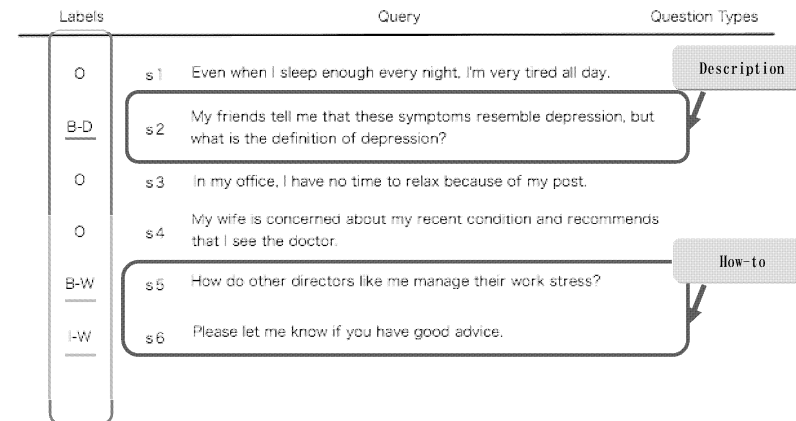
### Question type identification for MSQs

- Extract all question segments in a query and identify their question types

- Mainly pattern matching approaches for single sentence queries but some machine learning based methods[Li02,Zhang03,Suzuki05,Tamura05]
- QTI for MSQs in Japanese[Tamura et.al. 2005,2006]
  - Two pass processing; question extraction and the type identification
  - Accuracy of core sentence extraction: F-measure 0.897. But not clear in type identification for MSQs
- High computational cost using Support Vector Machines [Vapnik95]
  - Required reduce the computational cost

- Chunking-based question type identification
  - Chunking : identification of semantic “chunks”  
e.g., noun phrase, named entity,...
  - A chunk in our task equals to a question segment
  - Sentence chunking : assign a chunk tag to each sentence
  - Chunk tags : represented by a pair of the portion in the question segment and the question type

ex.1 B:the first sentence of the question segment  
and D:Description type ⇒ B-D  
ex.2 I :the second or more sentence and How-to  
⇒ I-W



- Perform question segmentation and the type identification concurrently; more efficient than the method in previous work
- Conditional Random Fields(CRFs): single exponential model for the joint probability of the entire sequence of labels given the observation sequence
  - Advantageous in computational cost and the flexibility to design the parameters of the model
- Using CRF++: the CRF implementation by [Kudo06]

## Evaluations

- Features : the uni-gram and bi-gram of words
  - Comparing the performance when using the feature of functional words alone with one when using all words
- Baseline : n words from the head and tail of sentence
- Traditional representation of chunk
  - IO/IOB1/IOB2/IOE1/IOE2/IOBES[Kudo2000]
- Varying the number of sentence exploited in the features of CRF

- Rows of binary feature vector based on words occur in each sentence

	Group A : m frequent POS				Group B : n POS at end of sentence			chunk tags
	feature1	feature2	...	feature m	feature m+1	...	feature m+n	
s1 Even when ...	w1	w2	...	wm	w1,m+1	...	w1,m+n	O
s2 My friend ...	w1	w2	...	wm	w2,m+1	...	w2,m+n	B-D
s3 In my office ...	nil	nil	...	nil	w3,m+1	...	w3,m+n	O
s4 My wife ...	nil	nil	...	nil	w4,m+1	...	w4,m+n	O
s5 ..	w1	nil	...	wm	w5,m+1	...	w5,m+n	<b>B-W</b>
s6 Please let me ..	w1	nil	...	wm	w6,m+1	...	w6,m+n	I-W
s7 I need some ...	w1	w2	...	wm	w7,m+1	...	w7,m+n	O

- Segmentation: F-measure 0.60(Bi-gram, window size=five)
  - When adding the words and the contexts, improved the accuracy
- Type identification: F-measures(window size=one)
  - When adding the contexts of words, incline the accuracy

How-to	0.50	Uni-gram
Location	0.34	Uni-gram
Reason	0.35	Uni+Bi-gram
Description	0.34	Uni+Bi-gram

- Different tendency with the contexts of words by question types
- Error Analysis : errors often occur when anaphora resolution and word sense disambiguation are required
  - e.g., “Could you give me any advice”; no antecedent in adjacent sentence

- Question type identification for multi-sentence queries characterized low computational cost
  - Handling the query including multiple questions of single sentence query in a same frame work
  - Disappointed accuracies in some experiments yet

- Required anaphora resolutions and word sense disambiguation for queries required

- Consider different tendency between question segmentation and type identification for same features

Thank you