

# Answering Contextual Questions Based on the Cohesion with the Knowledge

Tatsunori MORI  
Shinpei KAWAGUCHI  
Yokohama National University  
mori@forest.eis.ynu.ac.jp

# Outline

- Introduction
- An approach: reference resolution by the cohesion with the knowledge
- Algorithm
- Experimental results in NTCIR-5 QAC3
- Discussion
- Conclusion

# Introduction

- Contextual question answering
  - Answering questions by taking into account the context, i.e., previously asked questions and their answers
- Two types of approaches
  - Taking account of context in the document/passage retrieval. (e.g. Takaki[13])
  - Completing reference expressions and ellipses (i.e. zero pronouns) in questions by using contextual information
    - A completed question is submitted to a non-contextual QA system. (e.g. Fukumoto et al.[1])
    - Zero pronouns: Ellipses of obligatory cases in Japanese.

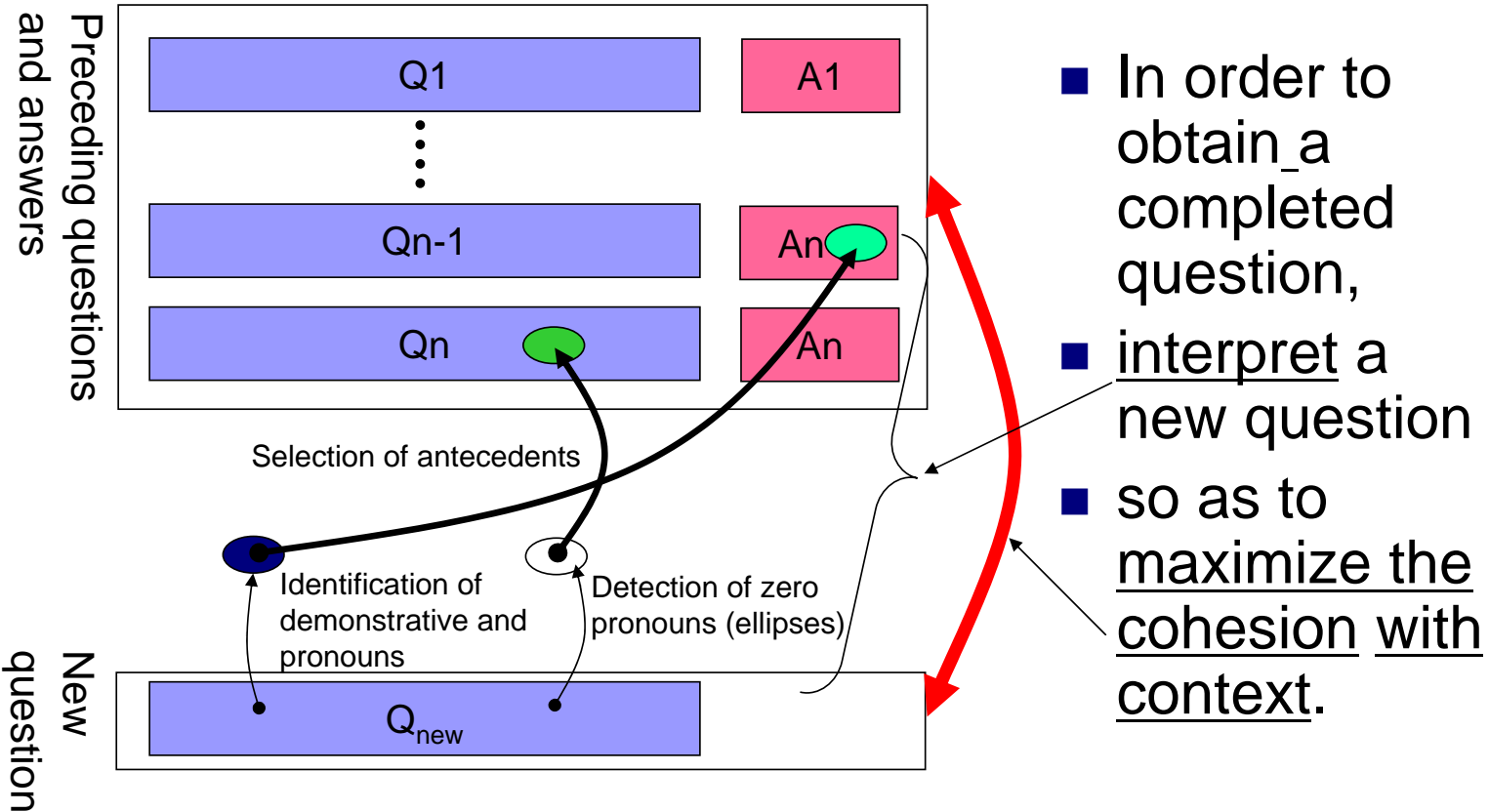
[13] Takaki. NTT DATA Question-Answering Experiment at the NTCIR-4 QAC2. NTCIR-4 (2004)

[1] Fukumoto et al. Rits-QA: List answer detection and Context task with ellipses handling. NTCIR-4 (2004)

# Our approach

- 2<sup>nd</sup> type: **completing questions by using contextual information.**
  - We expect that completed questions have sufficient information for answering.
  - Pro: Modularity in the question answering mechanism. We can use a conventional non-contextual QA system.
  - Con: Indirect use of context via references. There is room to introduce contextual information more directly. (e.g. document/passage retrieval phase).
- Completion of question---Reference resolution
  - Previous works: **the cohesion with the context**
  - Our approach: **the cohesion with the knowledge** as well as the cohesion with the context

# Reference resolution of question by the cohesion with the context



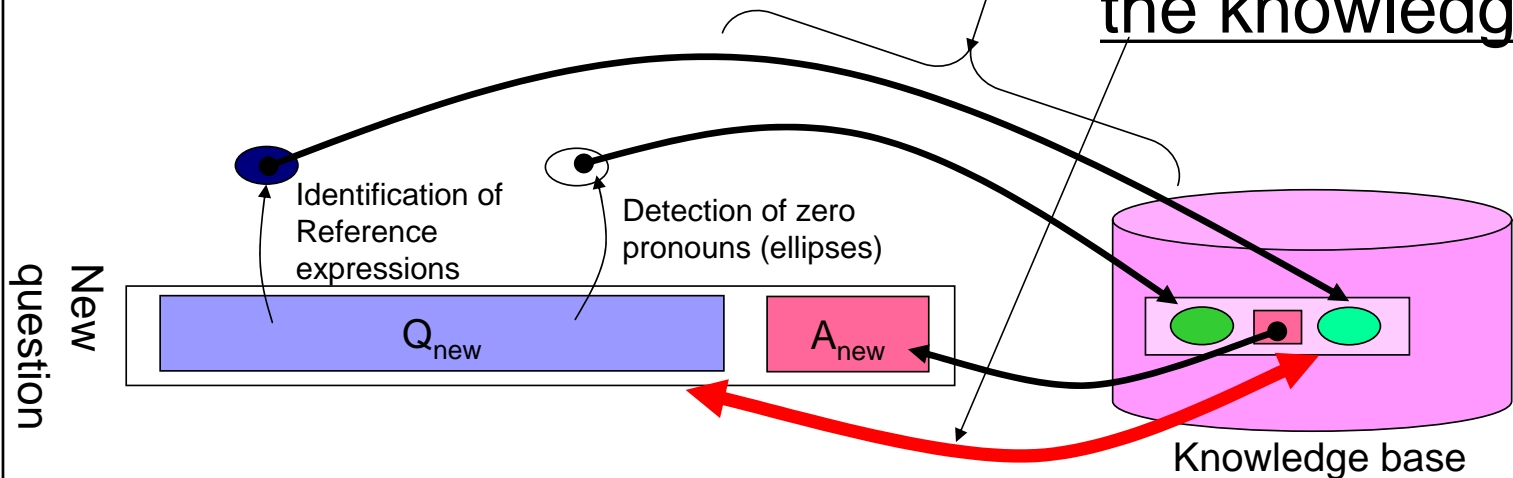
- In order to obtain\_a completed question,
- interpret a new question
- so as to maximize the cohesion with context.

However, in QA, we have another resource, i.e., the knowledge base.

# Reference resolution by the cohesion with the knowledge **only**

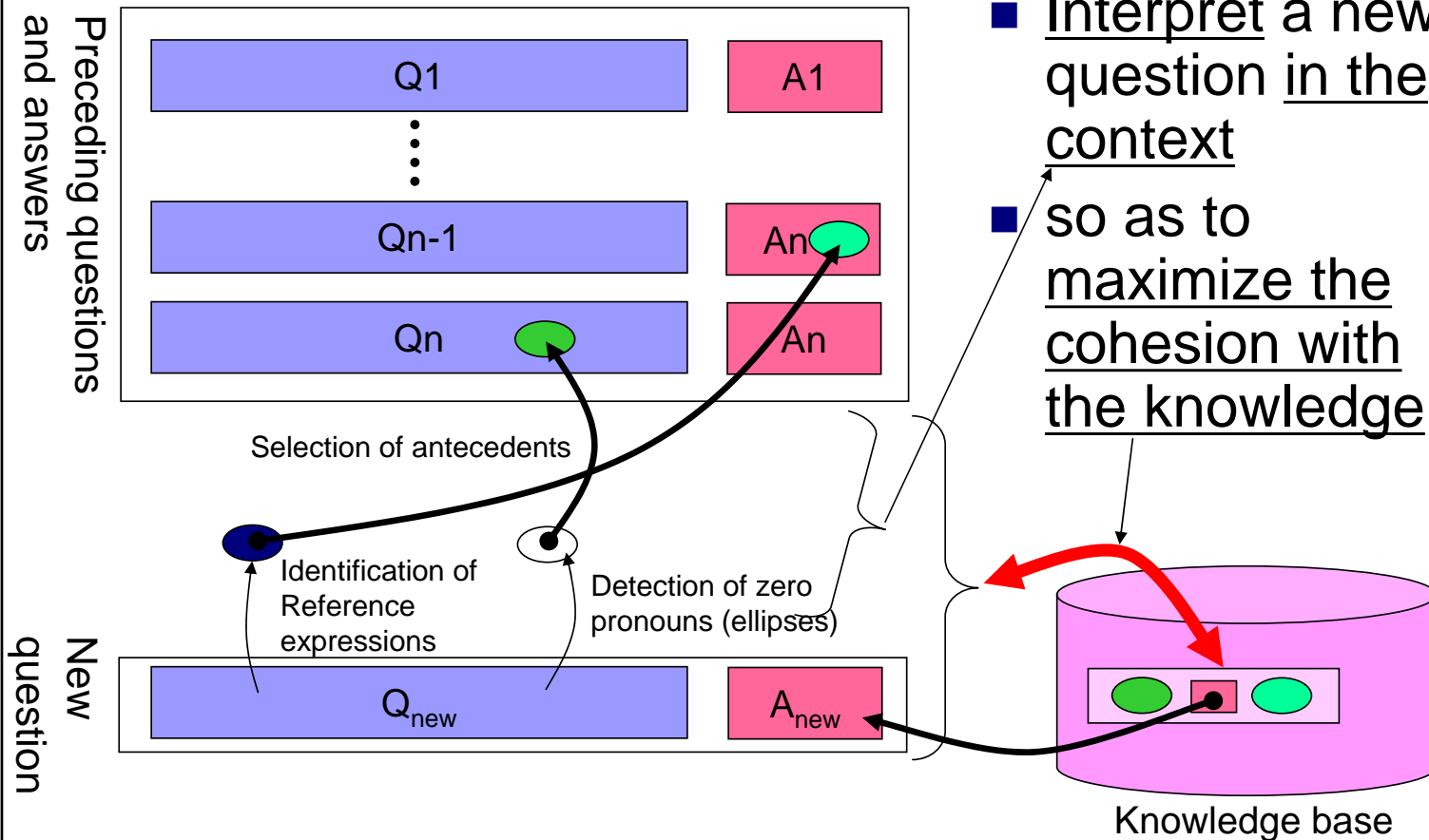
The context is totally abandoned.

- Interpret a new question
- so as to maximize the cohesion with the knowledge.



# Reference resolution of question by the cohesion with the knowledge

- Interpret a new question in the context
- so as to maximize the cohesion with the knowledge.



# Measuring the degree of cohesion with the knowledge (1)

- Our current implementation: **the score of the answer** (list) calculated by a non-contextual QA system
  - The score represents the goodness of the answer for a completed question candidate.
  - The goodness is a combination of the following factors.
    1. **The consistency of the question sentence with the context of an answer candidate.**
    2. The consistency between the question type and the type of the answer candidate.



# Measuring the degree of cohesion with the knowledge (2)

- The score of the current implementation is a linear combination of four sub-scores for an answer candidate  $AC$  in the  $i$ -th sentence  $L_i$  with respect to a question sentence  $L_q$  :

$$S(AC, L_i, L_q) = Sb(AC, L_i, L_q) + Sk(AC, L_i, L_q) \\ + Sd(AC, L_i, L_q) + St(AC, L_i, L_q)$$

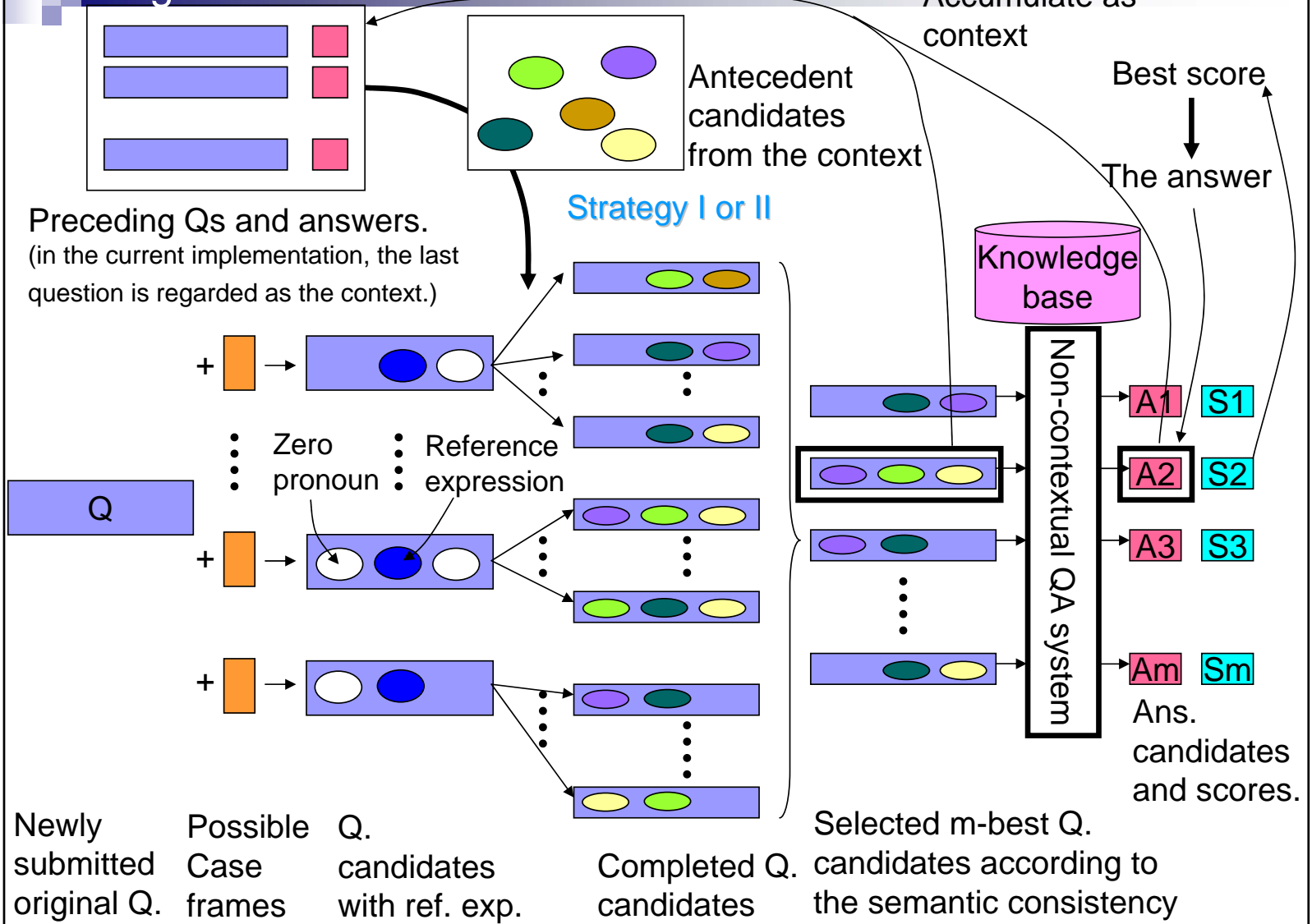
$Sb()$  : Matching score in terms of 2 - grams.

$Sk()$  : Matching score in terms of keywords.

$Sd()$  : Matching score in terms of dependency relation  
between an answer candidate and keywords

$St()$  : Matching score in terms of the question type

# Algorithm



# Gathering candidates of antecedents

- Strategy I: **simply gathering all possible nouns and NPs**
  - NPs with the Japanese topic marker “WA” and all nouns in the preceding completed question
  - All phrases in the answer list of the preceding question
  - All nouns and NPs with the topic marker “WA” **in the first question**
- Strategy II: **a method based on Nariyama’s SRL-based centering theory** [11]
  - SRL (salient referent list) pools all over arguments (i.e. **NPs with case markers or topic markers**) which have appeared up to the current question.
  - A zero pronoun is resolved by selecting the most salient argument in the SRL. The order of salience is defined as follows
    - Topic (marked by the case marker “WA”) > Nominative (“GA”) > Dative (“NI”) > Accusative (“O”) > Others.
  - Our modification (current implementation)
    - SRL is obtained from the completed preceding question.
    - Demonstratives and pronouns in a new questions are resolved before zero pronouns.
    - The interrogative in the completed preceding question is replaced with each of answers in the answer list.

[11] S. Nariyama. Grammar for ellipsis resolution in Japanese, In Proc. of the 9<sup>th</sup> Int’l conf. on Theoretical and Methodological Issues in MT (2002)

## Narrowing down antecedent candidates using the selectional restriction

- For each reference expressions (and ellipses), **candidates of antecedent are narrowed down** using a selectional restriction.
- **The selectional restriction is based on the similarity  $sim(a,r)$**  in a thesaurus between the categories of an antecedent  $a$  and the reference expressions  $r$ . The candidates that have the similarity under a threshold  $Th_{sim}$  are discarded.
- With regard to zero pronouns, the category information is obtained from case frames.

$$sim(a, r) = \begin{cases} \frac{2L_{ar}}{l_a + l_r} & \text{if } a \notin r \\ 1 & \text{if } a \in r \end{cases}$$

$l_a, l_r$  : the depths of the categories of  $a$  and  $r$  in the thesaurus

$L_{ar}$  : the depth of the lowest common ancestor of the categories

# Narrowing down completed question candidates

- The process so far may generate a lot of question candidates, and the non-contextual QA may take a very long time to process them.
- We introduce a measure  $C(S)$  for a completed sentence  $S$  as the degree of consistency in reference resolution, and narrow down the question candidates by selecting the m-best candidates.
- Some extra point is added to the value if the antecedent candidate is a named entity because a named entity tends to be an antecedent.

$$C(S) = \sum_{\langle r_i, a_i \rangle \in \text{resolv}(S)} c_1(r_i, a_i)$$

$$c_1(r, a) = \begin{cases} 1 & \text{if } a \in r \wedge a \text{ is not an NE} \\ 1.5 & \text{if } a \in r \wedge a \text{ is an NE} \\ \text{sim}(a, r) & \text{if } a \notin r \end{cases}$$

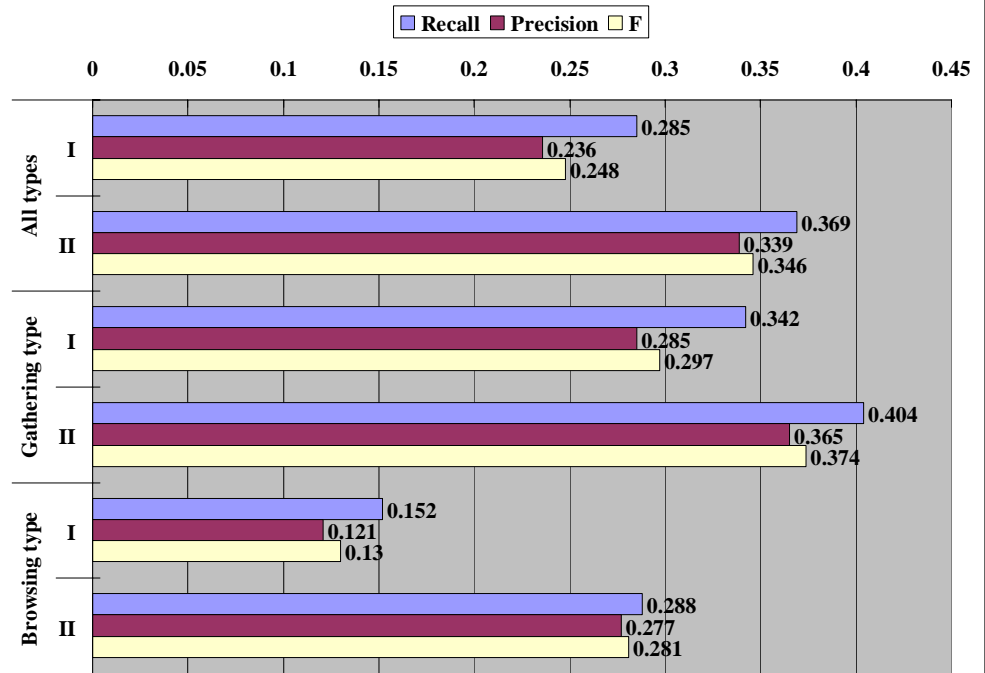
$\text{resolv}(S)$ : Set of pairs of reference expressions and its antecedent candidate in Sentence  $S$ .

# NTCIR-5 QAC3 Formal runs

- We evaluated the two strategies in terms of the accuracy of reference resolution and question answering.
  - Reference resolution: Recall, precision and F-measure
  - Question answering: Recall, precision and MMF1
- Setting of experiment
  - Non-contextual QA system: a Japanese real-time QA system by Mori[9]
  - Case-frame dic.: “Nihon-go goi taikei” (a Japanese lexicon)
  - Thesaurus for selectional restriction: “Nihon-go goi taikei”
  - # of completed Q. candidates to be selected (m): 20

# Reference resolution

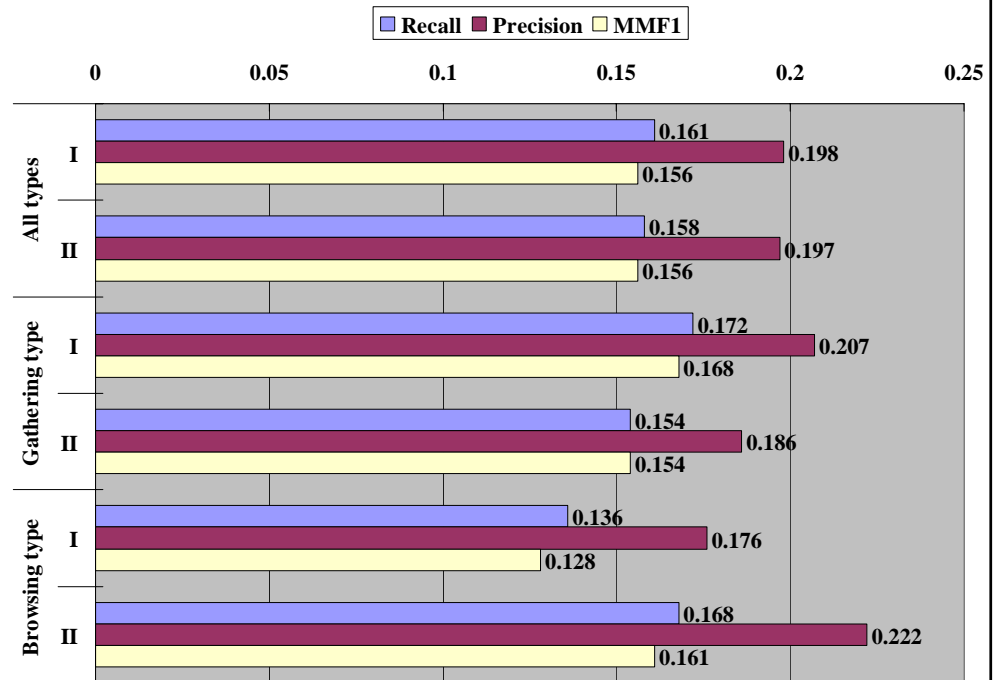
- Centering-theory-based method (Strategy II) is more accurate.
  - The centering theory is a method with an established reputation and works well in many cases.
  - Difference in unit of antecedents
    - In Strategy II, only noun phrases with case markers or topic markers can be antecedents.
    - In Strategy I, all nouns can be antecedents.



Strategy I : All possible nouns  
 Strategy II: Centering-theory based method  
 Gathering type: Q series of gathering type  
 Browsing type: Q series of browsing type

# Question answering

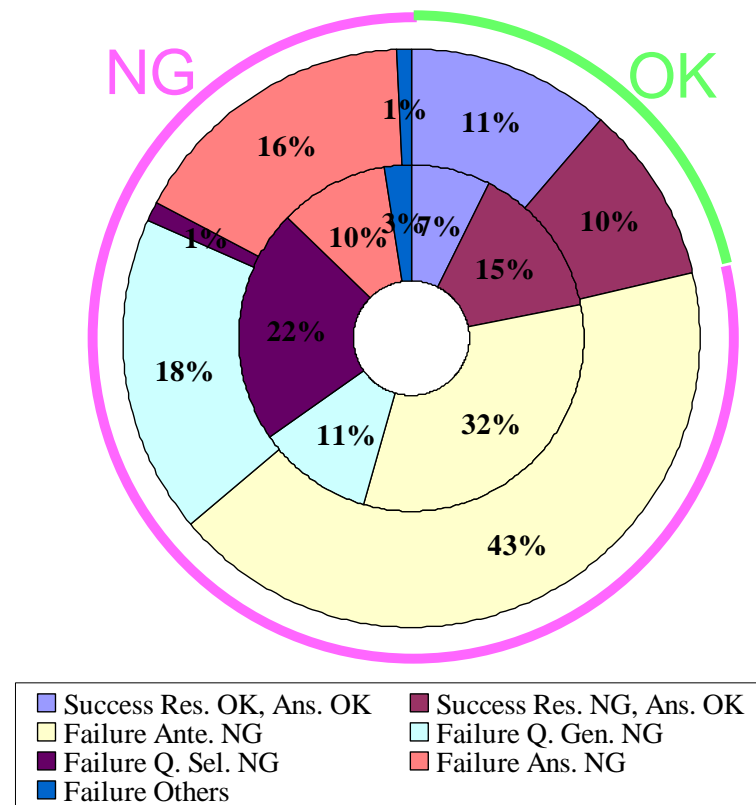
- Strategy I has the almost same accuracy as Strategy II in spite of its insufficient performance in reference resolution. It has much better performance for series of the gathering type.
- Strategy II is well-balanced. It works for the gathering type as well as browsing type with almost same accuracy.





# Failure analysis

- There are many cases that reference resolution is failed but the system successfully finds the answers. Strategy I has stronger tendency to succeed in such cases.
  - Newly introduced expressions may work well in the early stages of question answering, e.g. document/passage retrieval. The non-contextual QA system is robust to non-grammatical questions.
- The main reason of failure lies in the case that appropriate antecedents do not appear in either the completed preceding question or its answer list.
  - In the current implementation, the last (completed) question is regarded as the context.
  - The system failed to find correct answers for some previous questions, or failed to find appropriate antecedents in completing some previous questions.



Inner ring: Strategy I  
Outer ring: Strategy II

# Conclusion

- We introduced the notion “the cohesion with the knowledge,” and based on it, proposed a contextual QA system using a non-contextual QA system.
- Experimental results in NTCIR-5 QAC3
  - Strategy I has much better performance for series of the gathering type than the browsing type.
  - Strategy II is well-balanced.
  - The main reason of failure is that the appropriate antecedents in the current question do not appear in either the completed preceding question or its answer list.
- We need some device to maintain antecedent candidates in the more broad context like Nariyama’s SRL.