

Overview of the NTCIR-11 Cooking Recipe Search Task

Michiko Yasukawa (Gunma University, Japan) Fernando Diaz (Microsoft, USA)
Gregory Druck (Yummlly, USA) Nobu Tsukada (Kagawa Nutrition University, Japan)

Introduction

- About 90% of males and females in all age groups eat meals prepared at home for breakfast and dinner in Japan.
- Only 30% of males and 50% of females have sufficient/some knowledge and skills for food choice and preparation.
- Efficient access to recipe information is necessary.

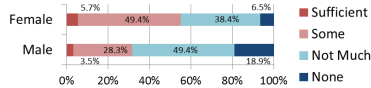


Figure: Knowledge and skills for food choice and preparation
Source: National Health and Nutrition Survey, Ministry of Health, Labour and Welfare (1999)

IR technologies for recipe search

- Users have complex information needs.
 - dietary constraints, course/dish, ingredient requirements, etc.
- Complicated constraints in realistic situation.
 - conditions for food allergy.
 - minimum/maximum intake of nutrients for medical treatments.

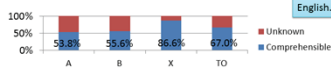


Two types of search users

- Non-professional users:
 - enjoy cooking for fun or for their daily living.
 - specific product names, abstract food names, etc.
 - Professional users:
 - experts in nutrition (dietitians).
 - recipe ingredients and standard food names.
- e.g., Standard Tables of Food Composition in Japan (2010), Dietary Reference Intakes for Japanese (2010).
- Semantic interpretation of food names and culinary words is needed. [queries → recipes]

A study of English queries

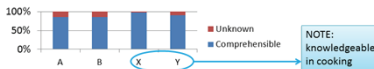
- Annotations of 500 English query samples were obtained.
- Query samples were categorized as comprehensible or not. (NOTE: the query samples were obtained from real user query log.)
- Annotators: two student assessors, a family member of a student assessor, one of the task organizers.
- Student A: 269/500 = 53.8% of queries are comprehensible. Others contain unknown words.
- Student B: 278/500 = 55.6% are comprehensible.
- Assessor X: 433/500 = 86.6% are comprehensible.
- Task Organizer: 335/500 = 67.0% are comprehensible.



NOTE: Assessor X is knowledgeable in cooking and English.

A study of Japanese food names

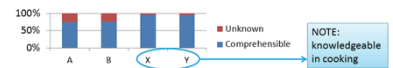
- Annotations on 110 pseudo queries were obtained and categorized as comprehensible or not. (NOTE: The pseudo queries were made from food names in recipes.)
- Annotators: two student assessors and two family members of one of the student assessors.
- Student A: 94/110 = 85.5% of queries are comprehensible. Others contain unknown words.
- Student B: 94/110 = 85.5% of queries are comprehensible.
- Assessor X: 108/110 = 98.2% of queries are comprehensible.
- Assessor Y: 100/110 = 90.9% of queries are comprehensible.



NOTE: knowledgeable in cooking

A study of Japanese food names (cont.)

- 794 Japanese standard food names appeared in Japanese recipes. (without spelling-variant normalization)
- Annotations were obtained by the four annotators:
 - Student A: 598/794 = 75.3% of queries are comprehensible. Others contain unknown words.
 - Student B: 608/794 = 76.6% of queries are comprehensible.
 - Assessor X: 764/794 = 96.2% of queries are comprehensible.
 - Assessor Y: 760/794 = 95.7% of queries are comprehensible.



NOTE: knowledgeable in cooking

Subtask 1: Ad-hoc recipe search

- search task with simple keywords.
- Similar to previous NTCIR tasks, such as NTCIR-3/4/5 WEB, NTCIR-7/8 IR4QA, NTCIR-8/9 GeoTime.
- NTCIR-11 RecipeSearch ad-hoc search is more technically challenging:
 - complex Boolean (AND/OR/NOT) search. e.g., "carrot cake with no butter"
 - semantic interpretation to process ambiguous words.
 - progressive query formulation using synonyms, paraphrases, spelling variants, abstract nouns, categorical words, etc. e.g., "fresh juice cocktails vodka" → what do "fresh juice" and "cocktails" mean? e.g., "fruit pie" → fruit could be "apple" or "blueberry"

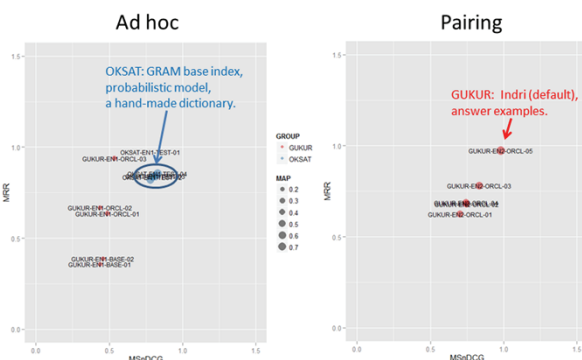
Subtask 1: Ad-hoc recipe search (cont.)

- Synonyms of negation words:
 - e.g., "eggless banana bread"
 - eggs, egg whites, egg yolks → must not appear
 - eggless, no-egg, egg-free → may appear
- English subtask 1: real user queries.
- Japanese subtask 1: pseudo queries.
- To give a guideline for relevancy, answer recipes for the queries were provided for the English/Japanese ad hoc search.

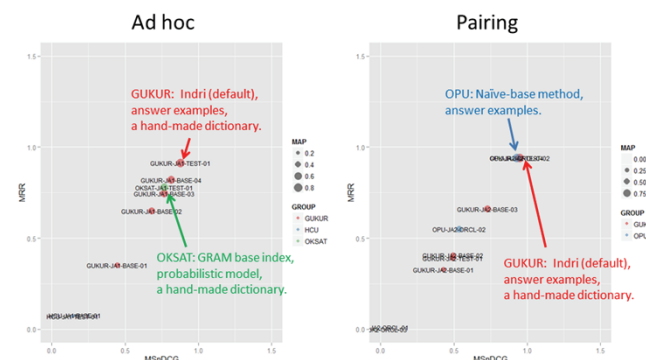
Subtask 2: Recipe pairing

- A simple form of menu planning: a user submit a recipe, and receives a complementary dish.
- Scenario: A user has selected a main dish and would like to find a complementary side dish.
- "Sweet and Sour Chicken" → "Fried Rice" (Chinese course)
- "Pineapple Chicken" → "Hawaiian Dessert" (Hawaiian cuisine)
- TOs provided answer-pairs of main dish and side dish.
- English answer pairs: often viewed in the same search session.
- Japanese answer pairs: manually chosen from a popular recipe author's repertoire.

Official results for English subtasks



Official results for Japanese subtasks



Official evaluation measures

Abbr.	Description
MAP	Mean average precision
MRR	Mean reciprocal rank
MSnDCG	Microsoft version of nDCG (normalized discounted cumulative gain)

Run type description

Abbr.	Description
BASE	Vanilla search. System input consists of only queries.
ORCL	Vanilla search. System input contains answer examples. (queries may be also included.)
TEST	Any input, any system for testing. (In addition to the testing data, queries and answer examples may be included.)

Conclusions

- Two subtasks: Ad hoc and Pairing
- Sufficiently challenging and moderately limited.
- Interesting problems of semantic analysis in recipes.
 - synonyms and negation of synonyms of culinary words.
- Major outcomes:
 - findings from the user study.
 - task-specific methodologies of query development and relevance judgment.
 - various approaches and experiments by participants.