NTCIR-8 Multilingual Opinion Analysis Task (MOAT)
First Round Table Meeting

Toyohashi University of Technology
Yohei Seki
June 22th, 2009
Outline

- NTCIR-8 Multilingual opinion analysis task
  - Task Design
  - New Challenge (1): Scenario-Based Evaluation
  - New Challenge (2): Cross-Lingual Opinion Analysis
  - New Challenge (3): Collaborative Annotation
  - Comments from participants

- Pilot Task

- Schedule
NTCIR-8 MOAT Task Design

- Continue four subtasks from NTCIR-6/7:
  A. opinion judgment, B. polarity judgment,
  C. opinion holder identification, and
  D. opinion target identification

- Data: Newspapers from 2002 to 2005 (Japanese, English, and (simplified/traditional) Chinese)

- New Challenge
  1. Scenario-based evaluation using opinion question
  2. Cross-lingual opinion analysis
     (Q: English -> Extraction multilingual opinion)
  3. Collaborative annotation with participants
New Challenge (1)
Scenario-based evaluation

- Provide topic description as follows:
  - Topic: swine flu
  - Question: what negative prospects were discussed about swine flu, which began in Mexico?
  - Polarity: negative
  - Opinion type: prospect discussion
  - Keyword: swine flu, Mexico

- Evaluate the accuracy of answer extraction for opinion question from relevant documents sets. (combine subtask evaluations)
New Challenge (2)  
Cross-lingual opinion analysis

- Provide English question and extract opinions from document sets in four languages (English, Japanese, Chinese (simplified/traditional))
- Clarify language-universal opinion property and language-dependent opinion property.
- Develop language adaptation technology by participants.
New Challenge (3)
Collaborative annotation

- Participants challenge opinion annotation.
  - Provide annotation site (tool) and manual by organizers.
  - Annotate the documents sets in three topics (known in previous before formal run).
  - While provide annotations by two assessors from organizers, provide annotations by one assessor from participants.
  - Check the agreements of annotation by discussing between organizers and participants.
Comments: NTCIR-8 Multilingual Opinion Analysis Task
Outline

- NTCIR-8 Multilingual opinion analysis task
- Pilot Task
  - Blog task: overview
  - Blog task: task design (plan)
  - Yahoo! answers pilot task (in Japanese)
  - Comments from participants
- Schedule
Blog task: overview

- Japanese copyright law was revised.
  - The revised copyright law was accepted in the national Diet in Japan on June 17th, 2009 and will be in effect from January 1st, 2010.
  - The copyrighted works could be recorded in the media as necessary for information analysis.

- ICWSM 2009 multilingual blog data
  - We obtained this data for NTCIR-8 MOAT use.
  - It contained 62 million multilingual blogs, and 2.6 million Japanese blogs, published between August 1st and October 1st, 2008.
  - Participants should obtain this data from ICWSM 2009 organizers.
Blog task: task design (plan)

1. Opinion analysis task
   - Opinion/polarity judgment, and opinion holder identification subtasks.
   - Overload for newspaper and blog annotation.
   - Implement transform script for sentence units.
   ⇒ Too much overload for organizers.

2. Recommendation task with Yahoo! answers.
   - With (discussion/consultation type) questions in Yahoo! Answers, recommend answer blogs.

3. Application-oriented task.

4. Cancel blog task.
Yahoo! Answers pilot task

- Best answer estimation
  - Best answer estimation using Yahoo! Answers second version data (in Japanese), which will be provided in NII soon.
Comments: Pilot task
Outline

- NTCIR-8 Multilingual opinion analysis task
- Pilot task
- Schedule
Schedule

- Close CFP: June 30th, 2009
- NTCIR8 Meeting: June, 2010.
You are most welcome for participation. We are also looking for organizers in pilot task.

ntcadm-opinion@nii.ac.jp

http://ntcir.kde.ics.tut.ac.jp/moat/index.php/MOAT