Extracting and Utilizing Event-Context Relationships in Blogsphere
Yukino Baba (The University of Tokyo, Japan), Fuyuki Ishikawa (National Institute of Informatics, Japan) and Shinichi Honiden (The University of Tokyo/National Institute of Informatics, Japan)

SUMMARY
We propose a method for extracting relationships between an event and the time/location it occurred from blog entries. This method targets the upcoming Ubiquitous era when user experiences are recorded in real time with spatiotemporal metadata. To extract the [Event]-[Time, Location] relationships, we analyze the spatiotemporal metadata from blog entries describing the same event.

PURPOSE
- Supporting real time creation of blog entries annotated with spatiotemporal metadata on smart devices
- Popularization of entries annotated with such metadata
- Similarity between blog-creation time/location and event-occurrence time/location

Limitation of estimating event-occurrence time and location only by text mining:
- Difficult to know "which word represents the actual time and location"
- Examples:
  "The World Cup soccer finals held in Germany, a match between France and Italy"
  "The Academy Award ceremony in Los Angeles, "Philadelphia", featuring Denzel Washington, won the Oscar for best original song"

Current Scope
A topic associated with specific time and location
Event occurrence time
Event occurrence location
Event Database

Event Database

Ubiquitous Computing
- Supporting real time creation of blog entries annotated with spatiotemporal metadata on smart devices
- Popularization of entries annotated with such metadata
- Similarity between blog-creation time/location and event-occurrence time/location

Limitation of estimating event-occurrence time and location only by text mining:
- Difficult to know "which word represents the actual time and location"
- Examples:
  "The World Cup soccer finals held in Germany, a match between France and Italy"
  "The Academy Award ceremony in Los Angeles, "Philadelphia", featuring Denzel Washington, won the Oscar for best original song"

Approach

Blog entries
Blog entries having the same topic

Event occurrence time
e.g., "2007-11-13 17:21:35"

Event occurrence location
e.g., "lat: 35.15 lng: 129.05"

Method for selecting blog entries by topic:
1. Estimate the probability of a word belonging to a topic by probabilistic topic modeling (*)
2. Estimate the probability of an entry belonging to a topic using the result of step 1
3. Select blog entries having high probability of belonging to a topic
(*) e.g., CPLSA [Mei & Zhai 06]

CONTACT
Yukino Baba / The University of Tokyo
Phone: +81-3-4212-2678 FAX: +81-3-4212-2697
E-mail: ybaba@nii.ac.jp