Multifaceted Analysis of News Articles by Using Semantic Annotated Information
Masaharu Yoshioka(Hokkaido University), Noriko Kando(National Institute of Informatics)

Objectsives
- Proposal of comparative multifaceted analysis interface for news articles from different countries
- Usage of semantic annotated text
  * Person, Place, Organization, Sentiment, News source, …
- Support interactive construction of multifaceted structured query for comparative analysis
- For example
  * Key persons related to Barack Obama about “Global Warming”
  * Comparison of negative issues related to Barack Obama in each country
  * Comparative trend analysis on positive and negative articles for “Economics” that contains different keywords (e.g., Euro, Japan, China; …)

Background: NSContrast: World News Surveillances System
- News surveillance system that focuses on the difference of each news resource.
- Comparison of characteristic features (keywords) among different countries

NSContrast

Semantic Annotation with Wikipedia Ontology

Wikipedia has a large number of entries about named entities
Usage of Category Information for estimating type of the entry
Classify named entities into several semantic categories by using Wikipedia category
Semantic category

Person: 日本の人物 (Japanese Person), …
Place: 北海道の地理 (Geography of Hokkaido), …
Organization: 国連機関 (Organization of United Nation), …

Varieties of description
e.g., Barack Obama and President Obama
Usage of Wikipedia redirect to identify synonym

Example of Link Discovery
'Narosura' populated place in Kenya
http://en.wikipedia.org/wiki/Narosura

Extract new link with higher precision and detect errors

Semantic Annotation with Geographic Information by using GeoNames and Wikipedia

Automatic Link Discovery between GeoNames and Wikipedia
Usage of Category Information for estimating location and type of the entry

Geography of Japan

Geography of Japan by prefecture

Geography of Tokyo

Mining in Japan

Mining Companies of Japan

Lake Okutama

Sumitomo Group

Japanese Alps

Landforms of Japan

Mountain ranges of Japan

Example of Link Discovery

‘Narosura’ populated place in Kenya
http://en.wikipedia.org/wiki/Narosura

Categories: Educational institutions established in 1910, Japanese national universities, National Seven Universities, Hokkaido University

Universities and colleges in Sapporo

Facet: Person

Input: バラク・オバマ (Barack Obama)

Template: Pos-Neg in each country (Time sequence)

Multifaceted Interface for Comparison

- Multiple windows that represents different information based on different structured queries
- Structured query
  * Topic keywords for selecting articles
  * Multiple windows for representing different characteristic information related to the topic keyword
    * Type of information
    * Representation style
    * Additional query keywords for selecting articles to compare
  * All windows are updated when topic keywords are modified
- Template for analysis
  * Support novice users to construct structured query for common analysis cases
    * Comparison among positive and negative articles trends in different country
      * Type of information: Sentiment (positive, negative and neutral)
      * Representation style: Time sequence graph for representing number of articles
    * Additional query keywords: News resource (Japan, Korea, USA and China)

Conclusion and future works
- Proposed a novel framework that enables multifaceted analysis of news articles.
- Evaluation is important, but it takes some time to use this interface in effective way