Voice-based Information Retrieval

Lin-shan Lee

National Taiwan University

Why Is Text-based Information Retrieval Useful and Attractive?
How about Voice-based Information Retrieval?

|	Text-based	Voice-based
Resources	• Rich resources—huge quantities of text documents available over the Internet • Quantity continues to increase exponentially due to convenient access	• Spoken/multimedia documents are the new trend • Can be realized even sooner given mature technologies
Accuracy	• Retrieval accuracy acceptable to users • Retrieved documents properly ranked and filtered	• Problems with speech recognition errors, especially for spontaneous speech under adverse environments • More reliable retrieval/ranking technologies needed
User Interfaces	• Retrieved documents easily summarized on-screen, thus easily scanned and selected by user • Users may easily select query terms suggested for next iteration retrieval in an interactive process	• Spoken/multimedia documents not easily summarized on-screen, thus difficult to scan and select • Lacks effective output presentation and efficient user-system interactions • "Content Analysis": based on automatically generated summaries, titles, key terms, and semantic structures

More Reliable Retrieval Techniques

- Problems
  - Short queries, short segments
  - Lower recognition accuracies for spontaneous speech under adverse environments
- Possible approaches
  - More reliable spoken term detection techniques
  - Lattices, confusion networks, term positions
  - Subword units (covering OOV words, across different languages and using less space)
  - Other techniques useful in text-based retrieval: query expansion, concept matching, etc.
 - Methods for reducing computation/memory requirements
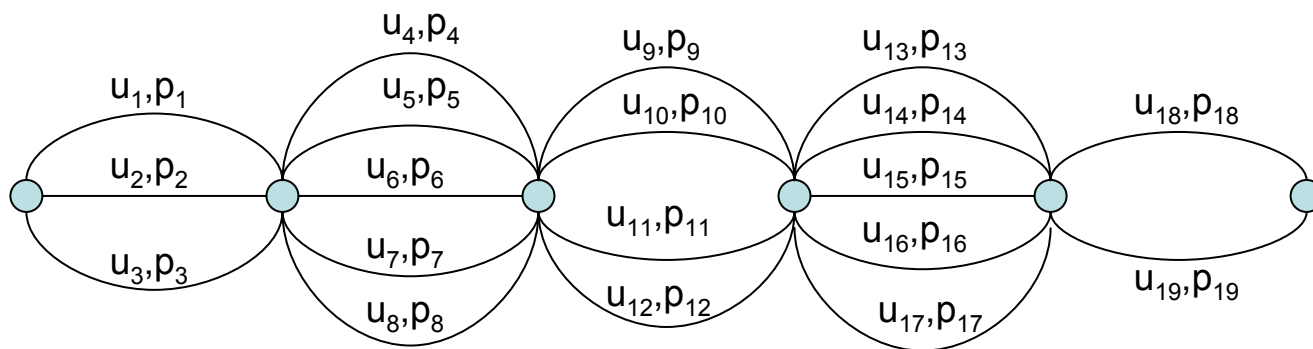
Output Presentation and User-System Interactions

- Problems
  - Spoken/multimedia documents not easily summarized on-screen, thus difficult to scan and select
  - Lacks effective output presentation and efficient user-system interactions

- Possible Approaches
  - Automatic summary/title generation for spoken/multimedia documents
  - Topic hierarchy construction for retrieved documents, with nodes labeled by key terms
  - Multi-modal user-system dialogue
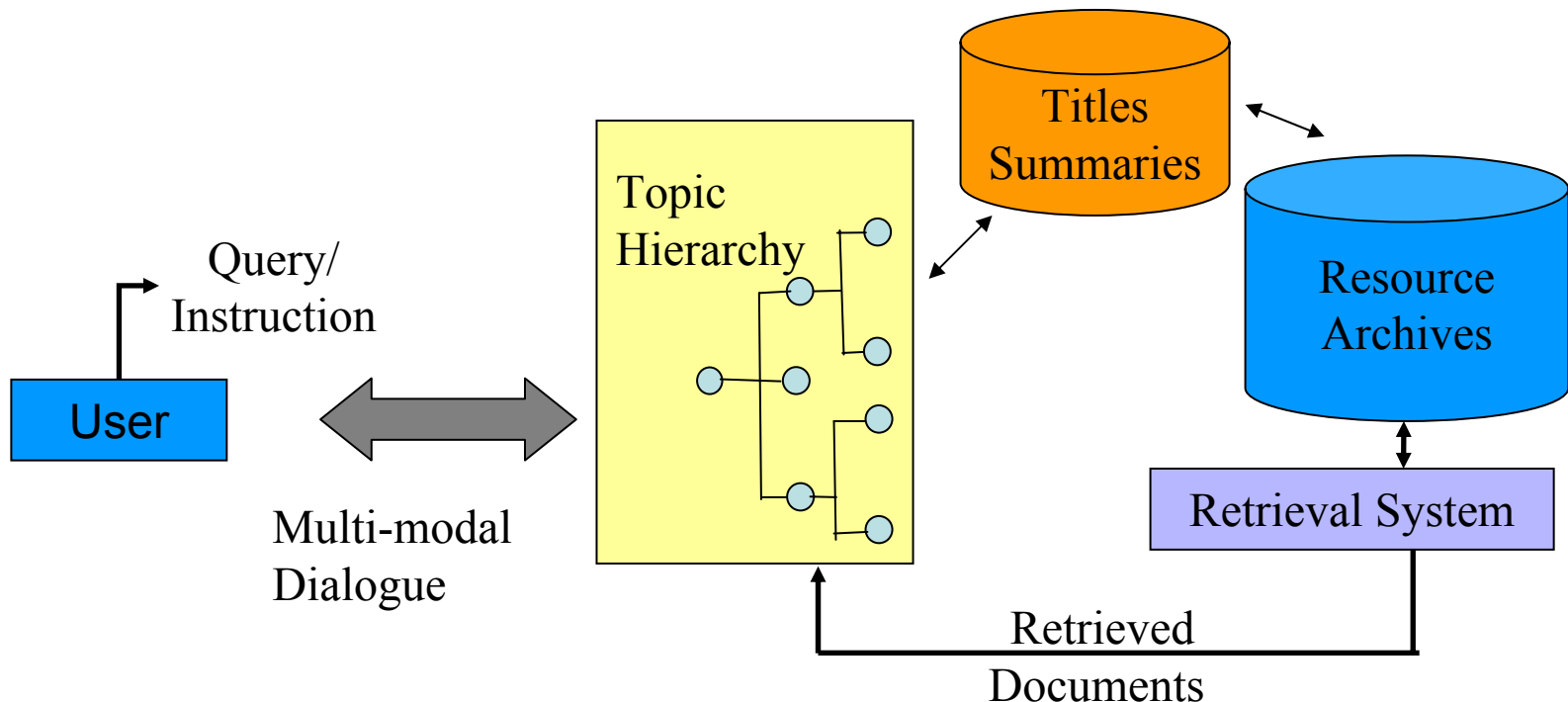
Spoken Language Processing on the Internet
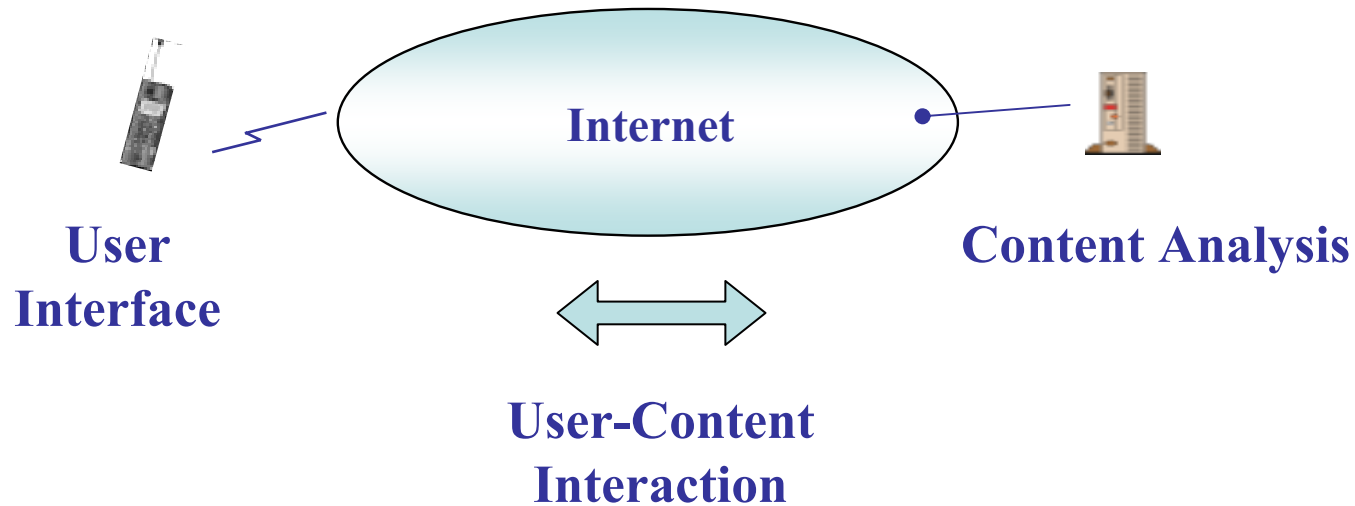
- User Interface
  - Difficult since users always expect technology to replace human beings
- Content Analysis/User-Content Interaction
  - Technology cannot perform as well as human beings
  - Can handle massive quantities of content, unlike human beings
- Voice-based Information Retrieval
  - Integrates user interface with content analysis/user-content interaction
  - Could be killer application for spoken language processing technologies some day