Speaker: Dr. Min-Yen Kan  
Dept of Computer Science, Columbia University

Date: Feb. 25, 2002  
Time: 3:00pm - 4:00pm  
Place: Olsen Hall 311  
Refreshments are served at 2:30pm

Automatic Text Summarization In Information Retrieval

Automatic text summarization is the task of condensing a document to produce a shorter version using a computer. A natural application of this is in information retrieval, where summaries of the search results could assist users in making full-text retrieval judgments. In the Centrifuser project, we explore how grouping documents by their topical similarity and differences can produce effective multidocument summaries for this application. The Centrifuser system benefits from a number of library resources, such as MARC cataloguing records and annotated bibliography entries. These resources shape the resulting summaries in both form and content. This talk examines design motivations for Centrifuser, initial evaluation results and current developments.