

Information Retrieval

CSC 9010: Special Topics. Natural Language Processing.

Paula Matuszek, Mary-Angela Papalaskari

Spring, 2005

Finding Out About

- There are many large corpora of information that people use. The web is the obvious example. Others include:
 - scientific journals
 - patent databases
 - Medline
 - Usenet groups
- People interact with all that information because they want to **KNOW** something; there is a question they are trying to answer or a piece of information they want.
- Information Retrieval, or IR, is the process of answering that information need.
- Simplest approach:
 - Knowledge is organized into chunks (pages or documents)
 - Goal is to return appropriate chunks

Information Retrieval Systems

- Goal of an information retrieval system is to return appropriate chunks
- Steps involve include
 - asking a question
 - finding answers
 - evaluating answers
 - presenting answers
- Value of an IR tool depends on how well it does on all of these.
- Web search engines are the IR tools most familiar to most people.