2nd class, March 17th
Webware main topics

- Search engines
- Server-side programming
- Client-side programming
- Mark-up languages (especially XML)
- Security and cryptography
- Multimedia
- Application layer

Precision/Recall

- There is a tradeoff between Precision and Recall
Precision/Recall Curves

- There is a tradeoff between Precision and Recall
- Therefore, we measure Precision at different levels of Recall (e.g., 10%, 20%, ...)

Precision/Recall Curves - Example

- Assume: we know all relevant documents in the collection:
  - \{d3,d5,d9,d25,d39,d44,d56,d71,d89,d123\}
    - 10 documents
  - Obtained by the query (ranked):
    - \{d123,d84,d56,d6,d8,d9,d511,d129,d187,d25,d38,d48,d250,d113,d3\}
  - Compute:
    - Precision at 10%, 20%, ... recall.
    - Build a point graph
    - Interpolate to have a line graph
Example

Precision

100 %

50 %

0 %

10 % 20 % 30 % 40 % 50 % 60 % Recall

Precision/Recall Curves
### Boolean Model

- **Advantages**
  - Simplicity
  - Semantic precision
- **Drawbacks**
  - Binary decision criterion
  - Difficult to formulate
- **Example:**
  
  \[ q = k_a \land (k_b \lor \neg k_c) \]

- **Disjunctive Normal Form (DNF)**
  
  \[ q_{dnf} = (1,1,1) \lor (1,1,0) \lor (1,0,0) \]

- **Conjunctive components**