SNS COLLEGE OF ENGINEERING

Kurumbapalayam (Po), Coimbatore – 641 107

An Autonomous Institution

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A' Grade
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

COURSE NAME : 19CS732 INFORMATION RETRIEVAL TECHNIQUES

IVYEAR / VII SEMESTER

Unit 2- MODELING AND RETRIEVAL EVALUATION

Topic 7 : Retrieval Evaluation and Retrieval Metrics

Problem

- Bad input selection.
- Noisy data.
- Very big dataset.
- Unsuitable structure.
- Inadequate number of hidden neurons.
- Inadequate learning rate.
- Insufficient stop condition; and/or.
- Bad dataset segmentation.

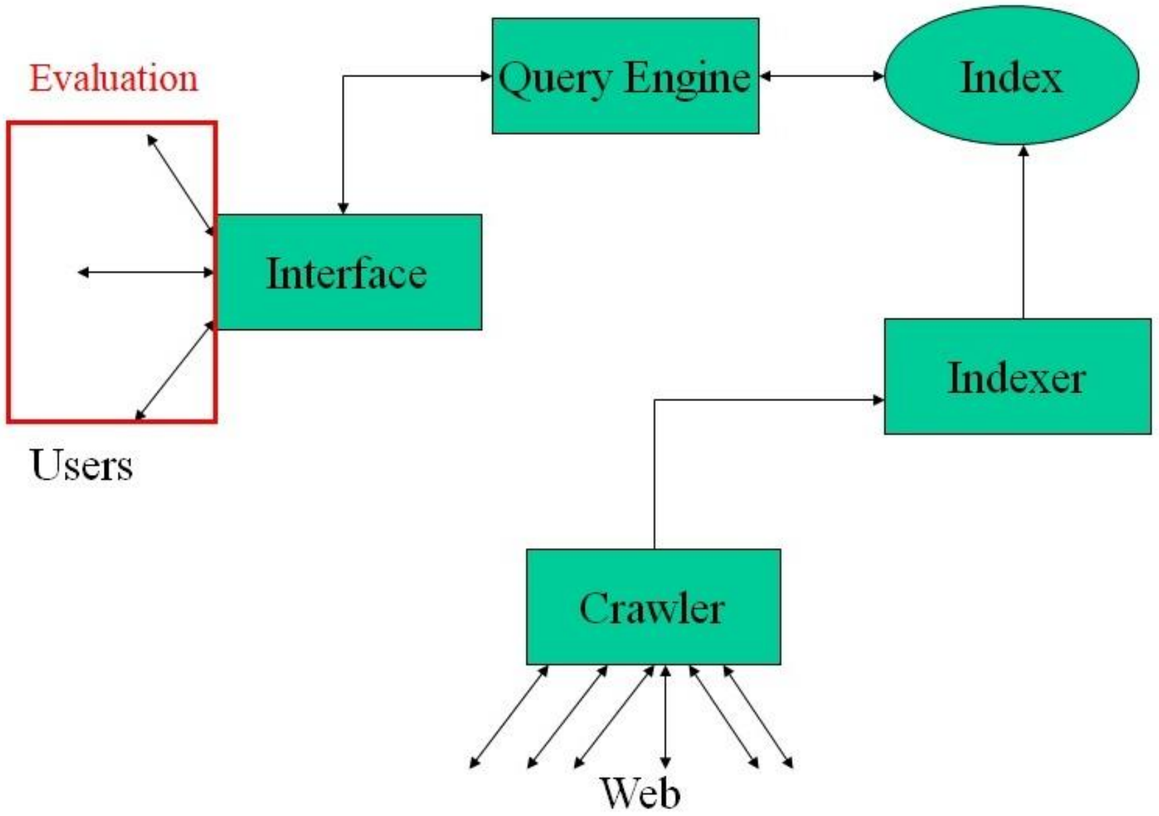
Retrieval Evaluation

- How do we know if our results are any good?
  - Evaluating a search engine
    - Benchmarks
    - Precision and recall
- How fast does it index
  - Number of documents/hour
  - (Average document size)
- How fast does it search
  - Latency as a function of index size
- Expressiveness of query language
  - Ability to express complex information needs
  - Speed on complex queries
- Uncluttered UI
- Is it free?

Measures for a search engine

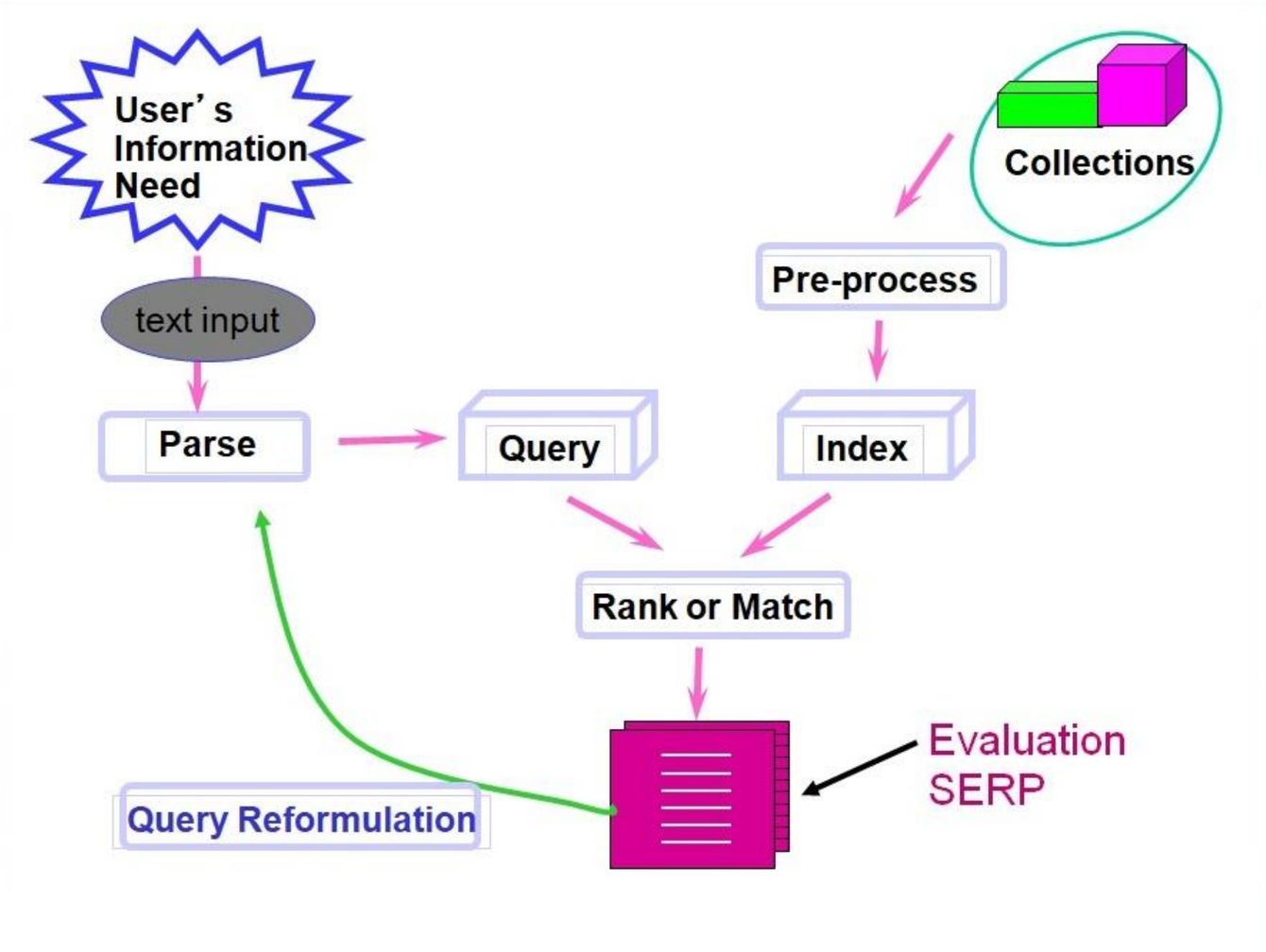
- All of the preceding criteria are *measurable*: we can quantify speed/size
  - we can make expressiveness precise
- The key measure: user happiness
  - What is this?
  - Speed of response/size of index are factors
  - But blindingly fast, useless answers won't make a user happy
- Need a way of quantifying user happiness

Evaluation –Cont..

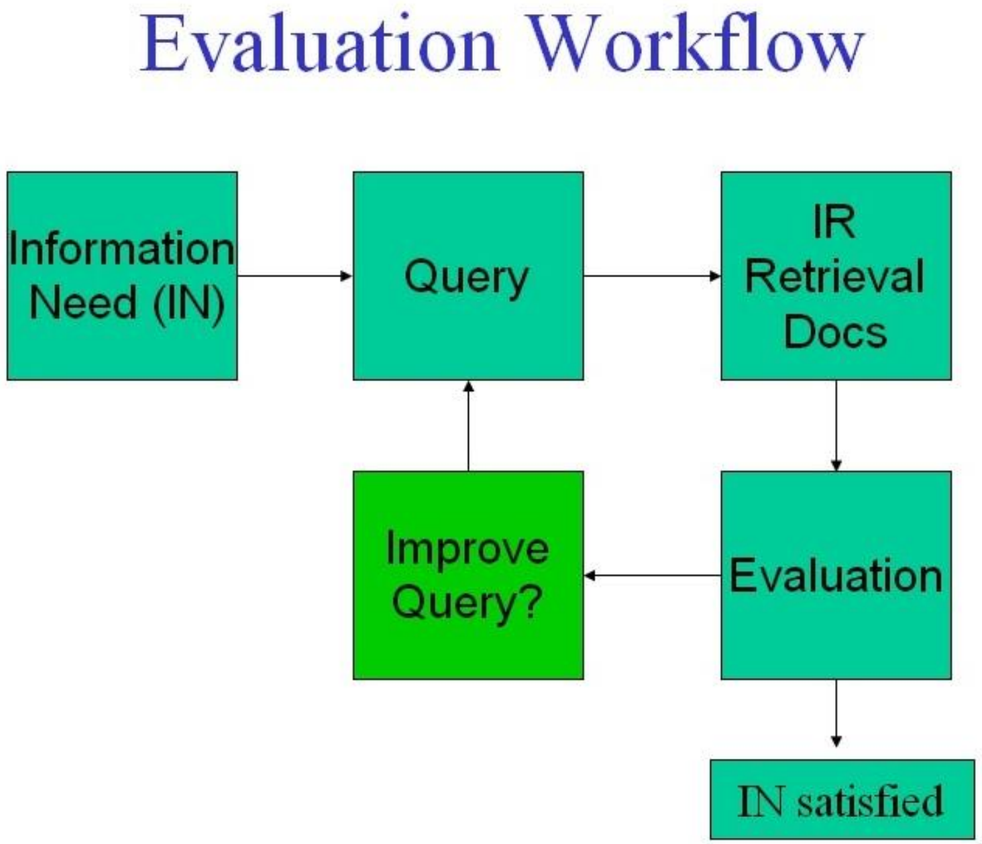


A Typical Web Search Engine

Evaluation –Cont..



Evaluation –Cont..



What does the user want? Restaurant case

- The user wants to find a restaurant serving sashimi. User uses 2 IR systems. How we can say which one is better?
- ✓Evaluate
- ✓Why Evaluate?
- ✓What to Evaluate?
- ✓How to Evaluate?

Why Evaluate?

- Determine if the system is useful
- Make comparative assessments with other methods/systems
  - Who's the best?
- Test and improve systems
- Marketing
- Others?

What to Evaluate?

- How much of the information need is satisfied.
- How much was learned about a topic.
- Incidental learning:
  - How much was learned about the collection.
  - How much was learned about other topics.
- How easy the system is to use.
- *Usually based on what documents we retrieve*

Relevance as a Measure

- *Relevance is everything!*
- How relevant is the document retrieved
  - for the user's information need.
- Subjective, but one assumes it's measurable
- Measurable to some extent
  - How often do people agree a document is relevant to a query
    - More often than expected
- How well does it answer the question?
  - Complete answer? Partial?
  - Background Information?
  - Hints for further exploration

Relevance

- In what ways can a document be relevant to a query?
- Simple - query word or phrase is in the document.
  - Problems?
- Answer precise question precisely.
- Partially answer question.
- Suggest a source for more information.
- Give background information.
- Remind the user of other knowledge.
- Others ...

How to Evaluate relevance document

- What can be measured that reflects users' ability to use a system? (Cleverdon 66)
  - Coverage of Information
  - Form of Presentation
  - Effort required/Ease of Use
  - Time and Space Efficiency
  - Effectiveness

Measures of relevance:

Binary measure

1 relevant

0 not relevant

N-ary measure

3 very relevant

2 relevant

1 barely relevant

0 not relevant

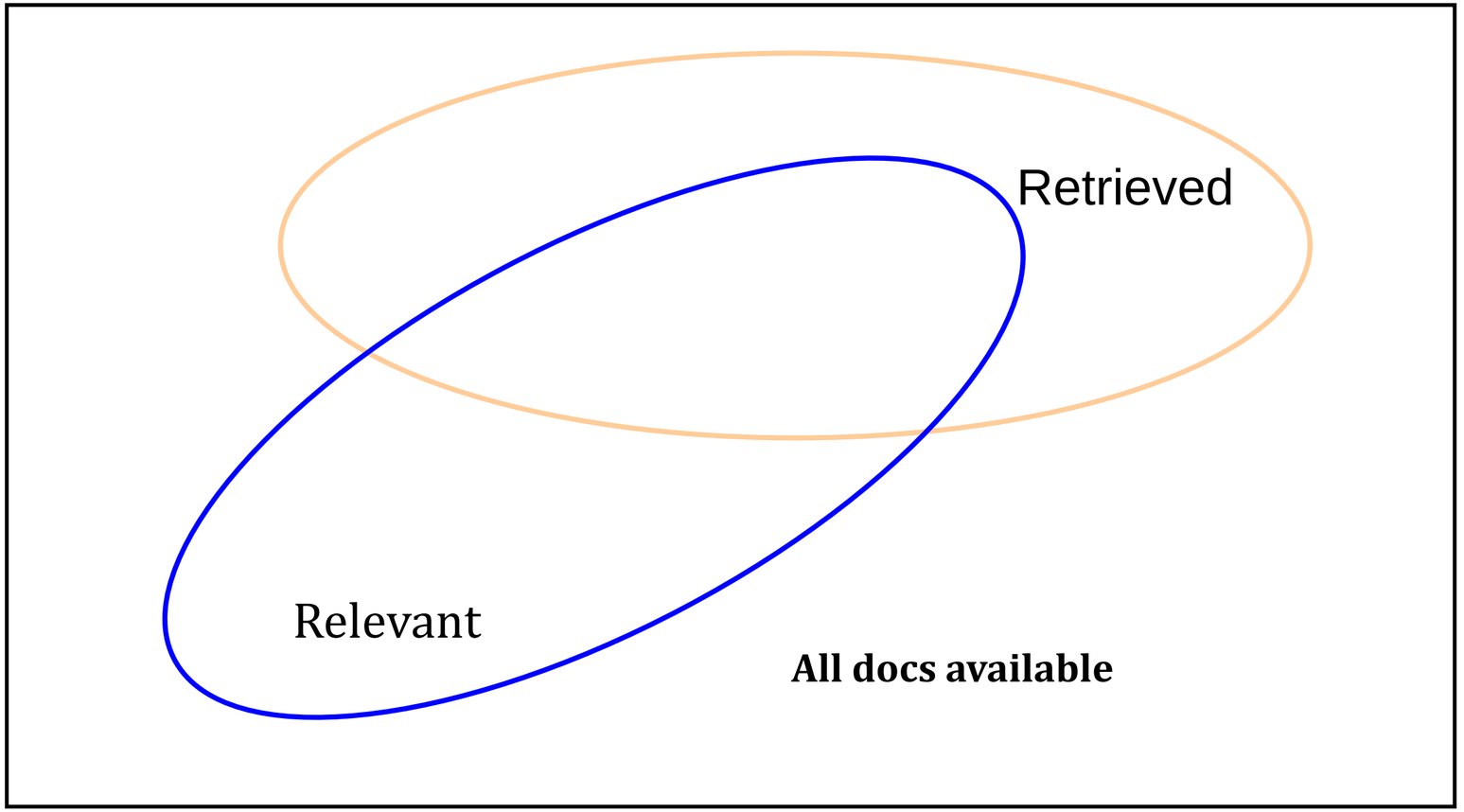
Negative values?

N=? consistency vs. expressiveness tradeoff

How to measure the effectiveness

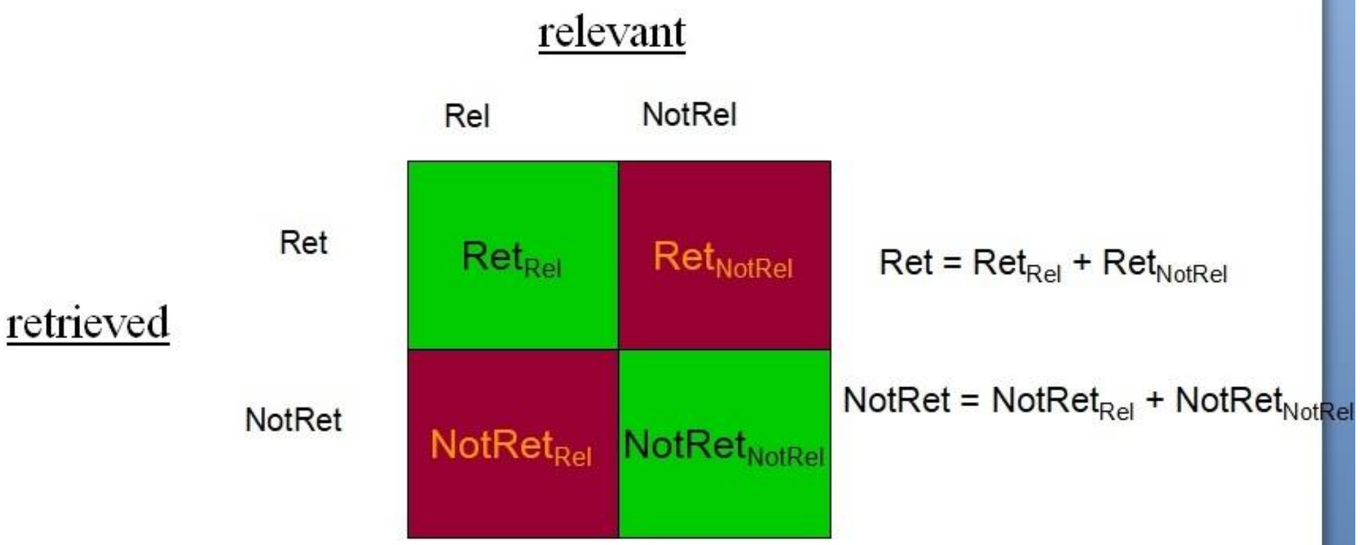
- Recall
  - proportion of **relevant** material actually retrieved
- Precision
  - proportion of retrieved material actually **relevant**

Relevant vs. Retrieved Documents



How to measure the effectiveness-Cont..

Contingency table of relevant and retrieved documents



$Relevant = Ret_{Rel} + NotRet_{Rel}$ $\quad$ $Not\ Relevant = Ret_{NotRel} + NotRet_{NotRel}$

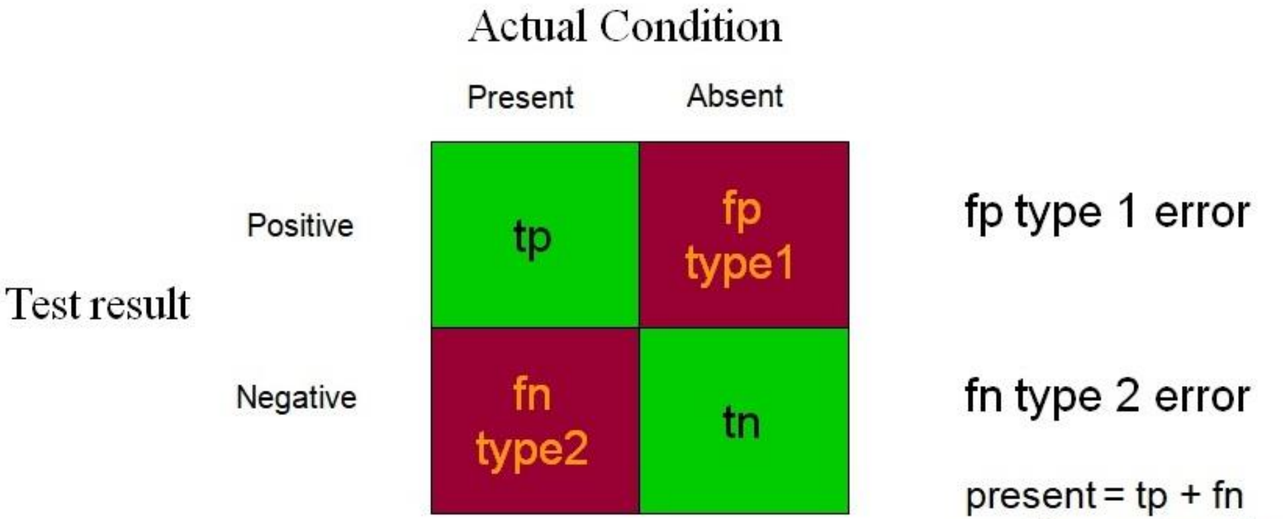
$Total\ \#\ of\ documents\ available\ N = Ret_{Rel} + NotRet_{Rel} + Ret_{NotRel} + NotRet_{NotRel}$

- Precision: $P = Ret_{Rel} / Retrieved$
- Recall: $R = Ret_{Rel} / Relevant$

$P = [0,1]$
$R = [0,1]$

How to measure the effectiveness-Cont..

Contingency table of classification of documents



fp type 1 error

fn type 2 error

present = tp + fn
positives = tp + fp
negatives = fn + tn

Total # of cases N = tp + fp + fn + tn

Precision: P= tp / tp + fp

- Recall: R = tp / tp + fn

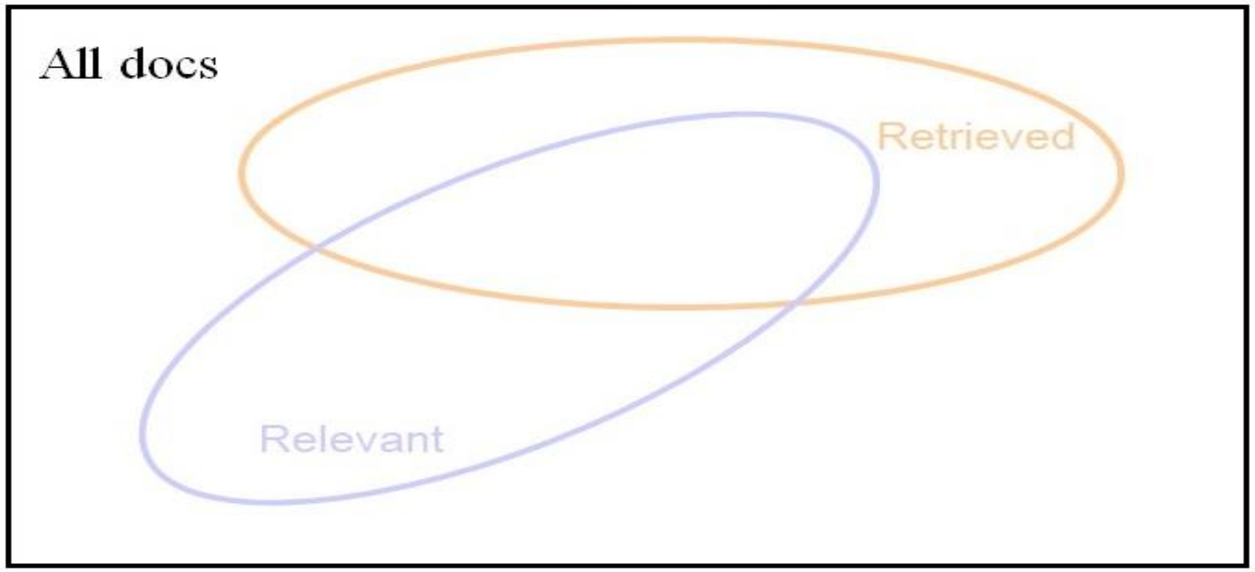
- False positive rate α = fp/(negatives)
- False negative rate β = fn/(positives)

How to measure the effectiveness-Cont..

Precision vs. Recall

$$\text{Precision} = \frac{|\text{RelRetrieved}|}{|\text{Retrieved}|}$$

$$\text{Recall} = \frac{|\text{RelRetrieved}|}{|\text{Rel in Collection}|}$$



How to measure the effectiveness-Cont..

Recall

$$R = \frac{\text{Number of relevant items retrieved}}{\text{Total number of relevant items in collection}}$$

Precision

$$P = \frac{\text{Number of relevant items retrieved}}{\text{Total number of items retrieved}}$$

Goal

high recall and high precision

Activity

Disadvantages

- Hardware dependence
- Unexplained behavior of the network
- Determination of proper network structure
- The duration of the network is unknown

Advantages

Performance evaluation

Speed

Space

Tradoff

Common for all systems. Not discussed here.

Retrieval performance (quality?) evaluation

= goodness of a retrieval strategy

A **test reference collection**: docs and queries.

The “correct” set (or ordering) provided by “experts”

A **similarity measure** to compare system output with the “correct” one.

Assessment 1

1. List out the Advantages of Retrieval Evaluation
   a)_____
   b)_____
   c)_____
   d)_____

2. Identify the disadvantages of Retrieval Evaluation
   a)_____
   b)_____
   c)_____
   d)_____

TEXT BOOKS:

1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, —Modern Information Retrieval: The Concepts and Technology behind Search, Second Edition, ACM Press Books, 2011.

2. Ricci, F, Rokach, L. Shapira, B.Kantor, —Recommender Systems Handbook||, First Edition, 2011.

REFERENCES:

1. C. Manning, P. Raghavan, and H. Schütze, —Introduction to Information Retrieval, Cambridge University Press, 2008.

2. Stefan Buettcher, Charles L. A. Clarke and Gordon V. Cormack, —Information Retrieval: Implementing and Evaluating Search Engines, The MIT Press, 2010.

THANK YOU