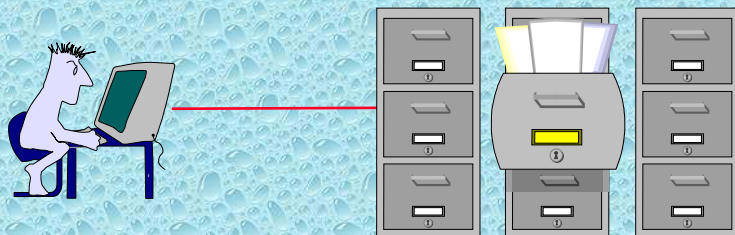


Databases and computerized information retrieval

Introduction



What is a database?

A database is a collection of similar data records stored in a common file (or collection of files).

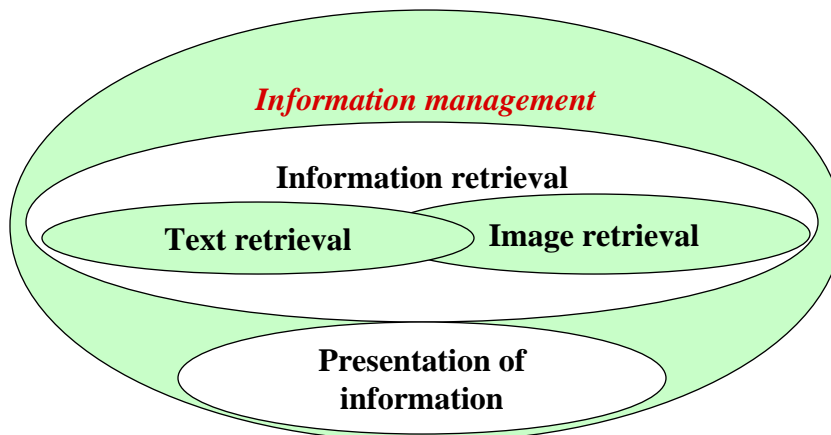


Types of databases: examples

Examples: The databases that form the basis for

- » catalogues of books or other types of documents
- » computerized bibliographies
- » address directories
- » a full text newspaper, newsletter, magazine, journal
+ collections of these
- » WWW and Internet search engines
- » intranet search engines
- » ...

Information retrieval and related activities: figure



Information retrieval and related activities: explanation

- **“Text retrieval” can be considered as a part of the larger concept “information management”.**
- **There is a great overlap:
“text retrieval” - “image retrieval”
because image retrieval is in most cases based on text retrieval:
in most cases retrieval of images is not based on computerized investigation of the images themselves, but on searches in the text that accompanies each image.**

Information retrieval: the terminology

Several words are used with similar or related meanings:

- » database / databank / corpus / collection / catalog / site / archive / file / web / ...
- » contents of a database / records / documents / items / (web) pages / ...
- » search / query / filter / ...
- » thesaurus / controlled vocabulary / dictionary / lexicon / term bank / ontology / ...
- » results / selection / retrieved documents / retrieved items /

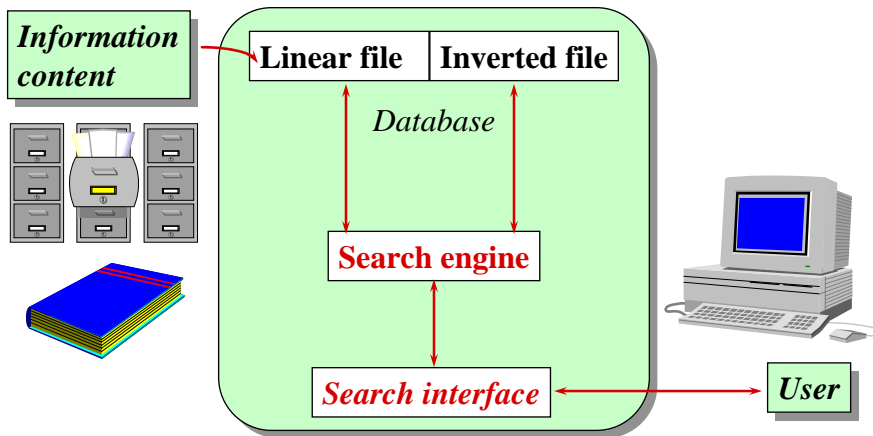
...

Information retrieval software: a particular type of DBMS

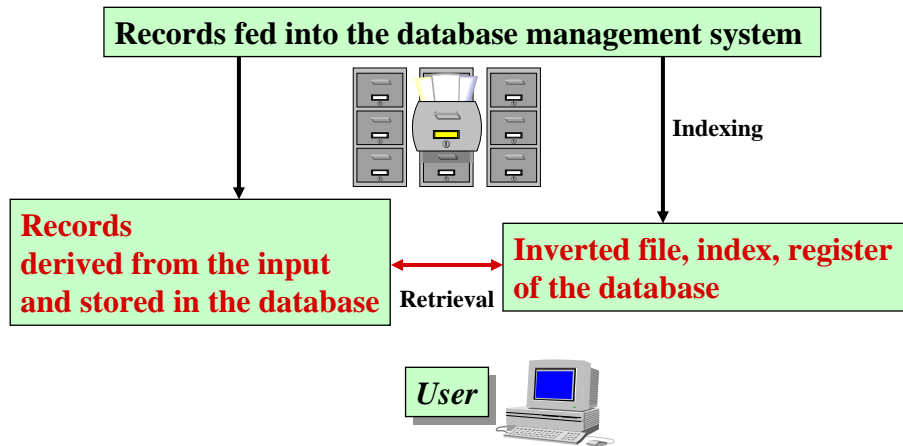
- **Software for information storage and retrieval**
(ISR software)
- **Text(-oriented) database management systems**
(Text-DBMS)
- **Text information management systems**
(TIMS)
- **Document retrieval systems**
- **Document management systems**



Information retrieval: via a database to the user



Information retrieval: building a database

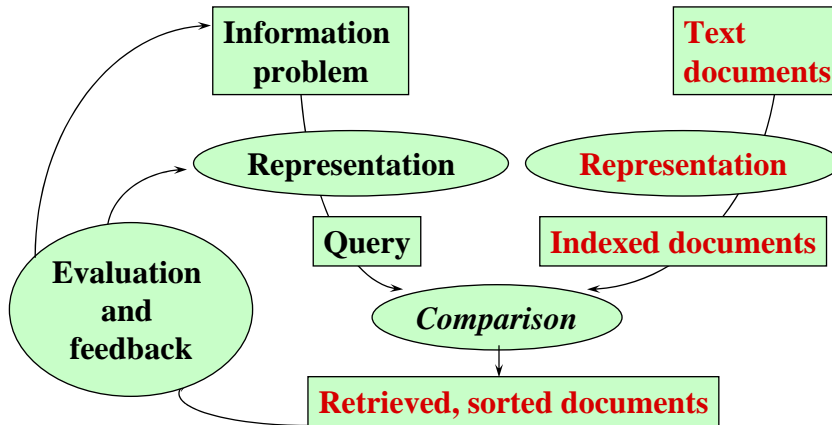


?? Question ??

The records input in a database system to be indexed do not necessarily appear completely in the output phase; that is: they are not shown completely to the user of the system in the results of a query. Can you illustrate this?



Information retrieval: the basic processes in search systems



Information retrieval systems: many components make up a system

- Any retrieval system is built up of many more or less independent components.
- To increase the quality of the results, these components can be modified more or less independent of each other.

Information retrieval systems: important components

the information content

system to describe formal aspects of information items

system to describe the subjects of information items

concrete descriptions of information items
= application of the used information description systems

information storage and retrieval computer program(s)

computer system used for retrieval

type of medium or information carrier used for distribution

Information retrieval systems: the information content

- The information content is the information that is created or gathered by the producer.
- The information content is independent of software and of distribution media.
- The information content is input into the retrieval system using
 - » a system (rules) to describe the formal aspects
 - » a system (rules) to describe the contents (classification, thesaurus,...)

Information retrieval systems: media used for distribution

- **Hard copy**
(for information retrieval systems only in the broad sense)
 - » Print
 - » Microfiche
- **For computers:**
(for information retrieval systems *strictu sensu*)
 - » Magnetic tape
 - » Floppy disk; optical disk (CD-ROM, Photo-CD, DVD...)
 - » Online

Information retrieval systems: the computer program

The information retrieval program consists of several modules, including:

- The module that allows the creation of the inverted file(s) = index file(s) = dictionary file(s).
- The search engine provides the search features and power that allow the inverted file(s) to be searched.
- The interface between the system and the user determines how they (can) interact to search the database (using menus and/or icons and/or templates and/or commands).

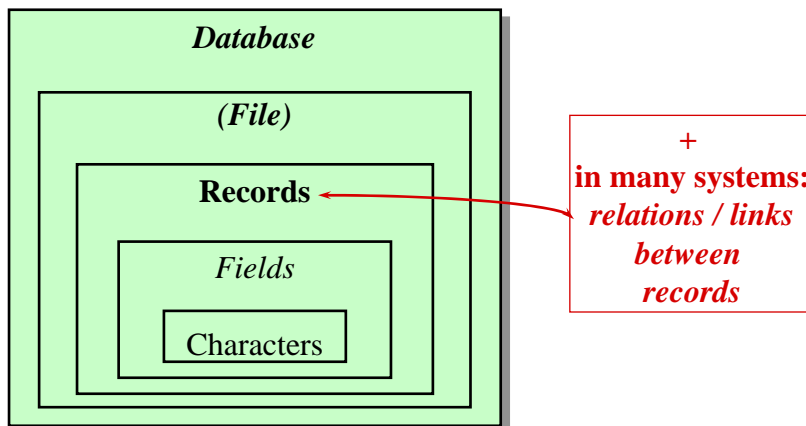
What determines the results of a search in a retrieval system?

1. the information retrieval system
(= contents + system)



2. the user of the retrieval system
and the search strategy applied to the system

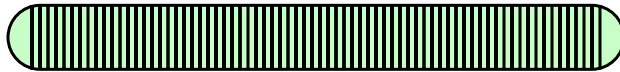
Layered structure of a database



A simple database architecture: all records together form a database

The ‘salami architecture’ = ‘sliced bread architecture’

- » the salami or the bread is a “database”
- » each slice of salami or bread is a “database record”
- » there are no relations between slices / records
- » the retrieval system tries to offer the appropriate slices / records to the user



!! Question !!

**The database architecture described here is simple,
but which factors make retrieval
nevertheless a complex procedure
in many real databases with this architecture?**

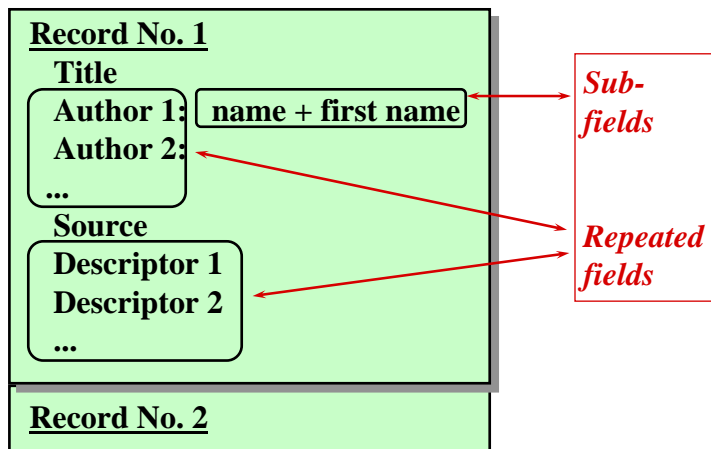


Characteristics / definition of structured text-information

- The text information is structured.
(files, records, fields, sub-fields, links/relations among records...)
- The length of records and fields can be “long”.
- Some fields are multi-valued = they occur more than once = repeated or repeatable fields



Structure of a bibliographic file



Databases and computerized information retrieval



Text retrieval and language

Text retrieval and language: an overview

Problems/difficulties related to language / terminology occur

- in the case of “multi-linguality”:
“cross-language information retrieval”;
that is when more than 1 language is used
 - » in the contents of the searched database(s)
and/or in the subject descriptors of the searched
database(s) OR
 - » in the search terms used in a query
- even when only 1 language is applied
throughout the system



Text retrieval and language: enhancing retrieval



- **Retrieval can be enhanced by coping with the problems caused by the use of natural language.**
- **Contributions to this enhancement of retrieval can be made by**
 - » the database producer
 - » the computerized retrieval system
 - » the searcher/user
- **(The distinction between these is not very sharp and clear in all cases.)**

!! Task - Assignment !!

Read about
Language and information retrieval
by Large, Andrew, Tedd, Lucy A., and Hartley, R.J.
Chapter 4 in: Information seeking in the online age:
principles and practice.
London : Bowker-Saur, 1999, 308 pp.



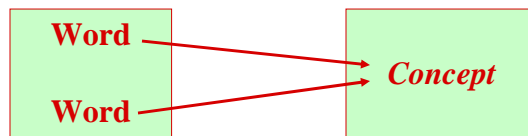
!! Task - Assignment !!

Read about
Information organization.
By Large, Andrew, Tedd, Lucy A., and Hartley, R.J.
Chapter 5 in: Information seeking in the online age:
principles and practice.
London : Bowker-Saur, 1999, 308 pp.



Text retrieval and language: a word is not a concept (a)

Problem:
A word or phrase or term is ***not*** the same as a concept or
subject or topic.



Text retrieval and language: a word is not a concept (a')

So, to 'cover' a concept in a search,
to increase the recall of a search,
the user of a retrieval system should consider an
expansion of the query;
that is:
the user should also include other words in the query to
'cover' the concept.



Text retrieval and language: a word is not a concept (a'')

» synonyms!
(such as :
Latin names of species in biology besides the common
names,
scientific names besides common names of substances in
chemistry...)



Text retrieval and language: a word is not a concept (a''')

- » narrower terms, more specific terms
(such as particular brand names);
including terms with prefixes
(for instance: *viruses*, *retroviruses*, *rotaviruses*...)
- » spelling variations
(such as UK English versus US English);
possible variations after transliteration



Text retrieval and language: a word is not a concept (a''')

- » singular or plural forms of a noun
(when this is used as a search term)
- » (relevant) related terms
- » various forms of a verb
(when this is used in the query)
- » broader terms (perhaps)



Text retrieval and language: a word is not a concept (b)



- **Method to solve the problem at the time of database production:**
 - » adding to each database record those codes from a classification system or terms from a thesaurus system that are relevant,
and providing the user with knowledge about the system used;
in some cases, this process is computerized
(with intellectual intervention or completely automatic)

Text retrieval and language: a word is not a concept (b')

- » However, this solution is not perfect:
 - Addition of terms by humans from a controlled vocabulary / from a thesaurus is not easy and time consuming.
Consequences:
 - the added value lags behind the availability of the document
 - the process can delay access to the document
 - the process is expensive
 - Moreover, in practice, most users of the resulting database do not exploit this method offered.

Text retrieval and language: a word is not a concept (c)



- **Method to solve the problem, provided by the computerized retrieval system:**
 - » offering to the user a partly computerized access to the particular subject description system used by the database producer, and then linking to the database for searching
 - » computerized, automatic, analysis of the ‘free text’ search terms applied in a query by the user, for transparent ‘mapping’ to the corresponding particular classification codes, categories, or thesaurus terms used by the database producer

Text retrieval and language: a word is not a concept (c')



- » offering the searching user access to a (general) thesaurus system, even when the database producer has not categorised the database contents; in this way, the user can refine his/her query
- » better, and more generally: computerized, automatic expansion of the query terms introduced by the user, based on a general thesaurus! (however, not many retrieval systems offer this feature)

Text retrieval and language: a word is not a concept (c'')



- » to avoid the problems of possible variations at the end of search terms:
 - offering the possibility to the user to truncate a search term explicitly
 - computerized, automatic, transparent truncation without explicit user action

Text retrieval and language: a word is not a concept (c''')



- » to avoid the problems of possible prefixes and suffixes:
 - computerized, automatic, transparent, intelligent morphological analysis of the query terms: 'stemming' of the 'free text' search terms used by the user;
however, this does not work perfectly and has not (yet) been implemented in most retrieval systems;
for languages that have a richer morphology than English, this can offer even a larger pay-off

?? Question ??

Which problems in text retrieval are illustrated by the following sentences?



**Time flies like an arrow.
Fruit flies like a banana.**

?

Time flies like an arrow.
Fruit flies like a banana.



Time flies like an arrow.
Fruit flies like a banana.

OK!

Text retrieval and language: ambiguity of meaning (a)

- **Problem:**
A word or phrase can have more than 1 meaning, because natural languages have evolved spontaneously, not strictly controlled.
- **Ambiguity of the meaning = polysemy.**
- **The meaning can depend on the context.**
- **The meaning may depend on the region where the term is used.**
- **This is a problem for retrieval.**
- **This decreases the precision of many searches.**

Text retrieval and language: ambiguity of meaning (a')

- **An example is the word “pascal”, which can have several meanings:**
 - » the philosopher Blaise Pascal,
 - » the programming language Pascal,
 - » the physical unit of pressure, and
 - » the name of many persons...
- **Another example:**
 - » Turkey, the country
 - » Turkey, the animal



Text retrieval and language: ambiguity of meaning (a'')

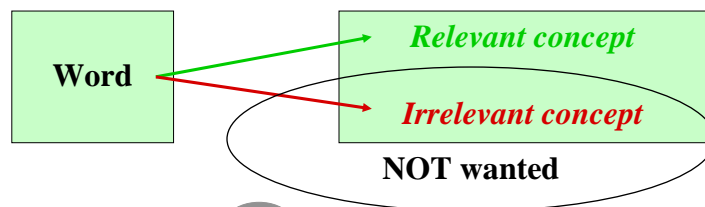
- **Example of sentences:**

- » The **banks** of New Zealand *flooded* our mailboxes with free *account* proposals.
- » The **banks** of New Zealand *flooded* with heavy rains *account* for the economic loss.



Text retrieval and language: ambiguity of meaning (a''')

Problem:
Ambiguity of meaning
may be the cause of low precision.



Text retrieval and language: ambiguity of meaning (b)



- **Method to solve the problem at the time of database production:**
 - » adding to each database record codes from a classification system or terms from a thesaurus system, and providing the user with knowledge about the system used;
in some cases, this process is computerized (completely automatic or with intellectual intervention);

Text retrieval and language: ambiguity of meaning (b')



- **Method to solve the problem, provided by the computerized retrieval system:**
 - » offering to the user a partly computerized access to the subject description system and then linking to the database for searching

Text retrieval and language: ambiguity of meaning (b'')



- » searching normally (without added value), but adding value by categorizing the retrieved items in the presentation phase to assist in the ‘disambiguation’; this feature is offered for instance by
 - the public access module of the book catalogue of the library automation system *VUBIS* at VUB, Belgium, when a searching items that were assigned a particular keyword

!! Task - Assignment !!

Search *Clusty* or *Vivisimo* or *Wisnut* as an example of a system that applies automatic, computerized subject categorization of database records.



Text retrieval and language: ambiguity of meaning (b'')



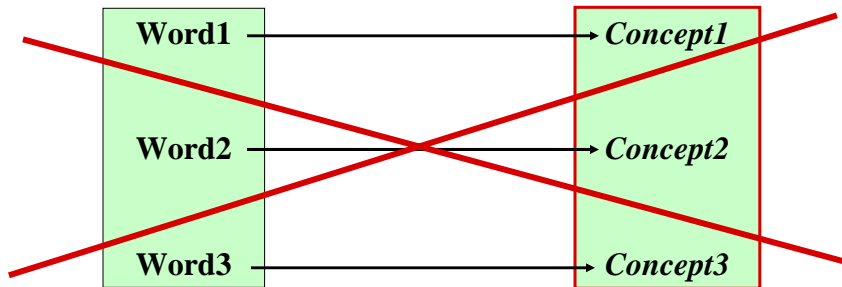
- » Natural language processing of the queries:
linguistic analysis to determine possible meanings of the query, which includes disambiguation of words in their context:
- “lexical” analysis = at the level of the word
“semantic” analysis = at the level of the sentence
- However, most queries are short and therefore it is difficult to apply semantic analysis for disambiguation.

Text retrieval and language: ambiguity of meaning (b'')



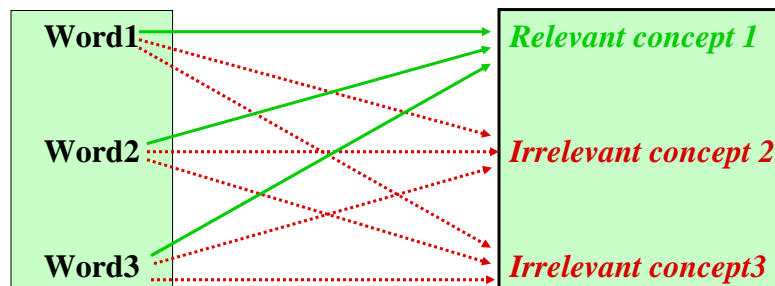
- » Natural language processing of the documents:
linguistic analysis to determine possible meanings of a sentence, which includes disambiguation of words in their context:
- “lexical” analysis = at the level of the word
“semantic” analysis = at the level of the sentence
- However, most retrieval systems do not apply this complicated method.

**A word is not a concept
A concept is not a word**



**The most simple relation
between words and concepts is NOT valid.**

**A word is not a concept
A concept is not a word**



- A concept cannot be “covered” by only 1 word or term;
this may be the cause of low recall of a search.
- The meaning of many words is ambiguous;
this may be the cause of low precision of a search.

Text retrieval and language: relation with recall and precision

Recapitulating the two problems discussed, we can say that

- Expansion of the query allows to increase the *recall*.
- Disambiguation of the query allows to increase the *precision*.



Text retrieval and language: evolution of meaning (a)

- **Difficulty:**
The meaning of a word or phrase can change over time.



Text retrieval and language: evolution of meaning (b)



- **Method to solve the problem at the time of database production:**
 - » using a categorization system and also adapting this continuously to the changing reality and meanings of terms

Text retrieval and language: phrases composed of words (a)

- **Problem:**
Most retrieval systems can search for words, but they do not directly recognize or 'know' phrases / terms composed of more than 1 word.



Text retrieval and language: phrases composed of words (b)



- **Methods to solve the problem, provided by the computerized retrieval system:**
 - » the user can and should indicate explicitly that a few words should be considered together by the retrieval system as forming a phrase/term (for instance in many Internet search engines by putting the phrase in quotes like “three word phrase”)

Text retrieval and language: phrases composed of words (b')



- » better:
 - the retrieval system automatically recognizes a phrase/term relying on a term bank that has been created in advance;
 - examples:
 - the Internet search engines *AltaVista* and *Scirus* work in this way

Text retrieval and language: searching more than 1 database (a)

- **Problem:**
Searching various databases at the same time,
or merging databases for searching,
suffers from the problem that these databases may use
categorization systems to make the problem of
terminology and language smaller, but in most cases these
systems are different and incompatible.



Text retrieval and language: searching more than 1 database (b)

- **Method to solve the problem,**
provided by the computerized retrieval system:
 - » mapping of the search term chosen by the user to the
various thesaurus terms used by the various databases;
only a few retrieval systems try to accomplish this



Text retrieval and language: relations among concepts (a)

- **Difficulty:**

In many cases, when the user combines several concepts in 1 search, the searching user cannot well communicate the intended relations among these concepts to the retrieval system.



Text retrieval and language: relations among concepts (a')

» Example:

concept 1 = children/sons/daughters/...

concept 2 = parents/fathers/mothers/...

concept 3 = beating/violence/...

How to find documents on

“children beating their parents”

while avoiding documents on

“parents beating their children”?



Text retrieval and language: relations among concepts (a'')

» Example:

concept 1 = computers

concept 2 = architecture

How to find documents on

*“the application/role/importance of
computers in architecture”*,

while avoiding documents on

“the architecture of computers”?



Text retrieval and language: relations among concepts (b)



- **Method to solve the problem,
provided by the database producer:**

» offering facilities to the user for disambiguation,
like in the more simple case of singular terms without
combinations with other terms

Text retrieval and language: relations among concepts (b')



- **Method to solve the problem, provided by the computerized retrieval system:**
 - » natural language analysis of both the documents and the natural language query to interpret their structure and meaning

Text retrieval and language: expressing the purpose of a search

- **Difficulty:**
Classical queries and retrieval systems work with terms to match the subject, the “aboutness” expressed in the query with the documents, but do not try to express and to understand the purpose, aim and context of the search.



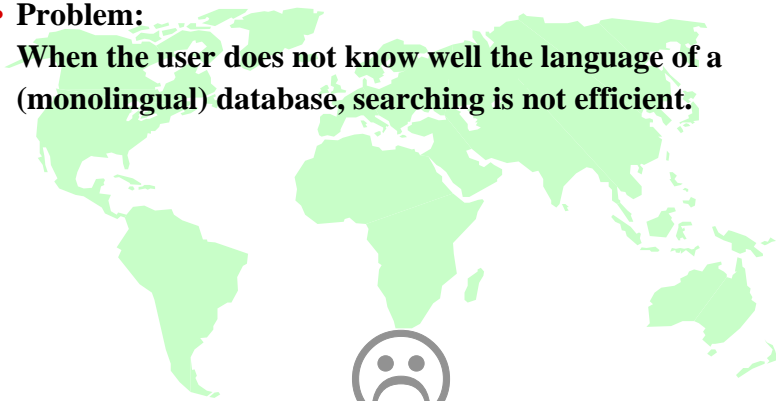
?? Question ??

Which are some of the problems caused by the use of language in information retrieval?



Text retrieval and multi-linguality (1a)

- **Problem:**
When the user does not know well the language of a (monolingual) database, searching is not efficient.



Text retrieval and multi-linguality (1b)



- **Methods to solve the problem, at the time of database production:**

- » adding subject descriptors in various languages (for instance in *Pascal* and *Francis* made by *INIST*)
- » adding abstracts in various languages (for instance the abstracts in English in *INSPEC*)
- » translation of the complete contents of the database

These processes can be partly computerized, but they are still time consuming and expensive.

Text retrieval and multi-linguality (1c)



- **Method to solve the problem, provided by the computerized retrieval system:**

- » translating the query of the user, by using a general multilingual thesaurus; however, most free text queries are quite short, which makes it difficult to use the context to limit possible ambiguity; disambiguation by user-computer interaction offered by the query interface, can increase the effectiveness here.

Text retrieval and multi-linguality (2a)

- **Problem:**
When documents in a database are written in more than 1 language, searching that database in a single language may not be sufficient to retrieve all interesting, relevant documents.



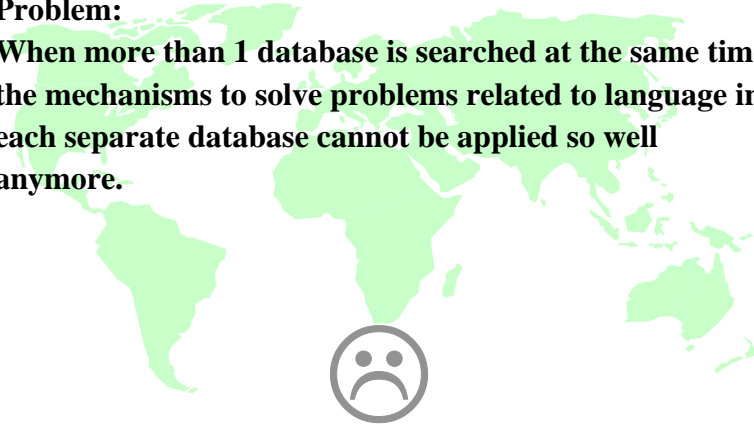
Text retrieval and multi-linguality (2b)

- **Method to solve the problem:**
 - » extensions of the methods when only 1 language is used in the documents



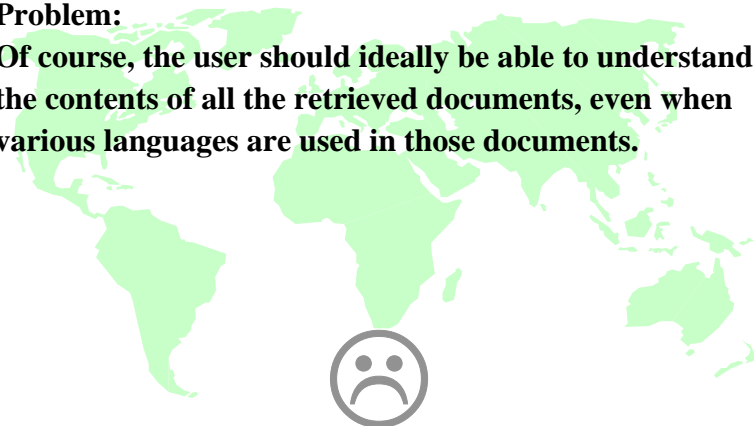
Text retrieval and multi-linguality (3)

- **Problem:**
When more than 1 database is searched at the same time, the mechanisms to solve problems related to language in each separate database cannot be applied so well anymore.



Text retrieval and multi-linguality (4a)

- **Problem:**
Of course, the user should ideally be able to understand the contents of all the retrieved documents, even when various languages are used in those documents.



Text retrieval and multi-linguality (4b)



- **Methods to solve the problem,
at the time of database production:**

- » adding abstracts in various languages
(for instance the abstracts in English in *INSPEC*)
- » translation of the complete contents of the database

These processes can be partly computerized,
but they are still time consuming and expensive.

Text retrieval and multi-linguality (4c)



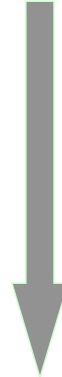
- **Methods to solve the problem,
provided by the computerized retrieval system:**

- » rapid automated translation
 - of the titles of retrieved records/documents
(for instance offered by the Internet search engine
AltaVista)
 - of the abstracts of retrieved records/documents
(for instance offered by the Internet search engine
AltaVista)
 - of the complete retrieved records/documents

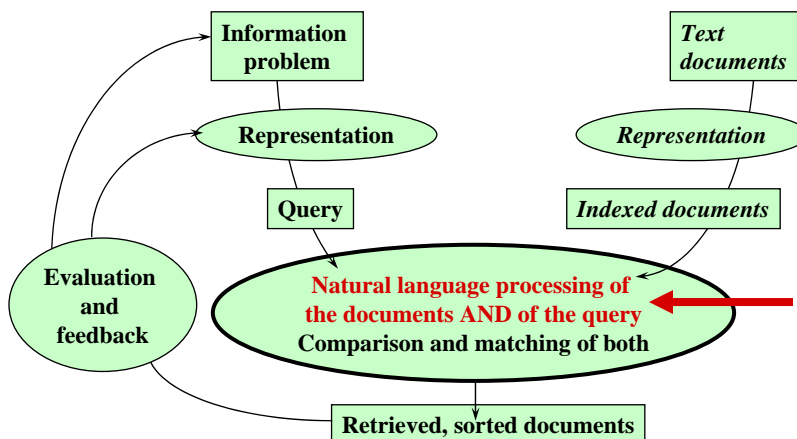
A good text retrieval system solves some problems due to language



- accepts words / terms / phrases in the query of the user
- maps the words to corresponding concepts
- presents these concepts to the user who can then select the appropriate, relevant concept (“disambiguation”)
- searches for this concept, even in documents written in another language
- presents the resulting, retrieved documents in the language preferred by the user



Enhanced text retrieval using natural language processing



Text retrieval and language: conclusions

- The use of terms and language to retrieve information from databases/collections/corpora causes many problems.
- These problems are not recognized or underestimated by many users of search/retrieval systems
= The power of retrieval systems is overestimated by many users.
- Much research and development is still needed to enhance text retrieval.



!! Task - Assignment !!

Recommended reading:

Veal, D.C.

Progress in documentation:

Techniques of document management:
a review of text retrieval and related technologies.

J. Doc., Vol. 57, No. 2, March 2001, pp. 192-217.



!! Task - Assignment !!

Recommended reading:

Chowdhury, G. G., and Chowdhury, Sudatta
Information retrieval in digital libraries.

In: Introduction to digital libraries.

London : Facet Publishing, 2003, 354 pp.



?? Question ??

Explain the basic relations/similarities in

- speech recognition (speech to text)
- translation of a text (text to text)
- summarizing texts (text to summary)
- text retrieval (query to texts)
- cross-language text retrieval (*combination*)



Databases and computerized information retrieval



Hints on how to use information sources

Hints on how to use information sources: overview (Part 1)

- **Know the purpose and motivation for each search.**
- **Do not be lazy: search on your own, before bothering experts with requests for advice.**
- **Plan your search in advance.**
- **Choose the best source(s) for each search.**
- **Use the available tools for subject searching well.**
- **Try to cope with the language problems;
avoid spelling errors in your search query;
use spelling variations in your search query**



Hints on how to use information sources: overview (Part 2)

- Match your search strategy with the type of source.
- Work cost-effectively.
- Use special care when searching for names.
- Be specific.
Avoid broad searches.
Limit your search to a specific country or region if required.
- Work iteratively.
- Keep a record of your work.




Hints on how to use information sources: overview (Part 3)


- Do not only focus on a single source.
- Consider citation indexes besides subject-oriented databases, as useful secondary information sources.
- Stop searching when “enough is enough”
- Give up if necessary... (Not all questions have an answer.)
- Be critical: not all information is correct or useful.



Hints on how to use information sources: overview (Part 4)

- **In computer-based retrieval systems, consider applying**
 - » truncation of search terms (using a symbol like * or ?)
 - » combine search terms, using 
 - Boolean operators:
OR **AND / +** **NOT / AND NOT / -**
 - proximity operators
 (for instance “**NEAR**”)
 - phrase searching (“**word1 word2**”)
 - » searching limited to a field (for instance URL, title...)


Hints on how to use information sources: subject searching

- **When you search for information on a particular topic/subject: investigate if the database producer offers** 
 - » a subject classification scheme and/or
 - » a controlled/approved/accepted subject terms, and/or
 - » a subject thesaurus

Exploit these, if they are available.
- **In most cases you should find and use synonyms and narrower terms**
- **Use broader and /or related terms, if appropriate.**

Hints on how to use information sources: language problems...

The problem of search terms with more than one meaning: solutions

- » Select the most specific, appropriate database. 
- » Limit to a specific, appropriate section of the database.
- » Find first synonyms or narrower terms using a vocabulary or thesaurus, and use these as search terms.
- » Limit the search to one (or several) fields.
- » ...

Hints on how to use information sources: Boolean combinations

Most text search systems understand the basic Boolean operators:



OR

= obtain records that contain one or both search terms



AND

= obtain records that contain both search terms

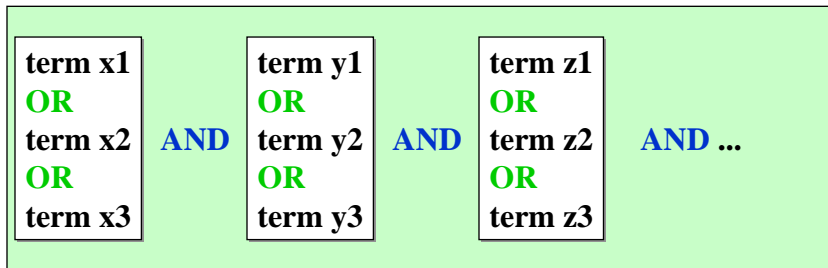


NOT or ANDNOT or AND NOT

= exclude records that contain a search term

Hints on how to use information sources: Boolean combinations

In the case of computer-based information sources, use Boolean combinations of search terms when appropriate and when possible.



?? Question ??

Suppose that you want to search for a topic that has several synonyms (for example, young people, adolescents, teenagers, teens). Then which one of the following operators would you use in your query?

ADJ AND NEAR NOT OR



Hints on how to use information sources: Boolean queries

- Most text search systems understand the basic Boolean operators typed in capital characters:

OR

AND

- So this leads us to queries like for instance
(word1 OR word2 OR word3 OR word4) AND (wordA OR wordB OR wordC)

Hints on how to use information sources: default Boolean operator

- Find out if there is a default implicit Boolean operator working in the search system that you use.
- This works even when no operator is used explicitly among words.
- This can be **OR**, **AND**, **NEAR**...
- So this leads us to queries like for instance
(word1 OR word2 OR word3 OR word4) (wordA OR wordB OR wordC)

?? Question ??

Why is it important to know the default Boolean operator in the search system that you use?
You can also explain this with an example.



!! Task - Assignment !!

You can read
Cohen, Laura
Boolean searching on the Internet. [online]
Available from:
<http://library.albany.edu/internet/boolean.html>
University Libraries, University at Albany, USA.
[cited 2006]



?? Question ??

You want to search a database for a low-fat recipe for pasta with either shrimp or chicken.

Which query demonstrates the proper use of nesting to get many search results that are very relevant?

1. **noodles or (pasta and shrimp) or chicken and low-fat**
2. **(noodles or pasta) and (shrimp or chicken) and low-fat**
3. **noodles or pasta and (shrimp or chicken) and low-fat**
4. **(noodles or pasta) and shrimp or (chicken and low-fat)**
5. **noodles or pasta and shrimp or chicken and low-fat**



?? Question ??

You need information on the communication strategies applied by the popular star Madonna.

Which query will probably be the most efficient one in some particular database, (of course in the case that the database understands the operators applied)

1. **Communication AND strategies**
2. **Madonna AND communication AND strategies**
3. **Madonna OR communication OR strategies**
4. **Strategies OR communication**
5. **Madonna**



?? Question ??

How many (and which) concepts/facets do you see in a search for “general reviews about monitoring seawater pollution that is due to effluents in Tanzania”?



!! Task - Assignment !!

Prepare off-line, on paper, a suitable search query in a generic format, to find “general reviews about monitoring seawater pollution that is due to effluents” as the basis for later, concrete searches in databases. (Limit yourself to 1 of the concepts.)



Hints on how to use information sources: example of a search query

Example: Searching for the concept “sea” can or should involve for instance the following words in a

Boolean OR-combination:

baltic OR bay OR bays OR coast OR coastal OR coastline OR coasts OR cove OR coves OR gulf OR mangrove OR mangroves OR **marine** OR mediterranean OR noordzee OR noordzeekust OR noordzeekusten OR **ocean** OR oceanic OR **oceans** OR pacific OR reef OR reefs OR “saline-freshwater interface” OR **sea** OR **seas** OR seashore OR **seawater** OR seawaters OR shore OR shores

?? Question ??

What did you learn from the exercise on the formulation of a query?



!! Task - Assignment !!

Prepare off-line, on paper, a suitable search query in a generic format, to find documents about
“how to evaluate the ability to find scientific information of starting university students up to professional scientists”
as the basis for later, concrete searches in databases.
(Limit yourself to 1 of the concepts.)



?? Question ??

How can we exploit in some searches the fact that many bibliographic databases (in particular the commercial, expensive ones) offer records with a field structure?



!! Task - Assignment !!

Read

Luther, Judy, Kelly, Maureen, and Beagle, Donald

Visualize this

(Visualization software may become a powerful new way to search
– or a footnote in technology history).

Library Journal, March 1, 2005, pp. 34-37.

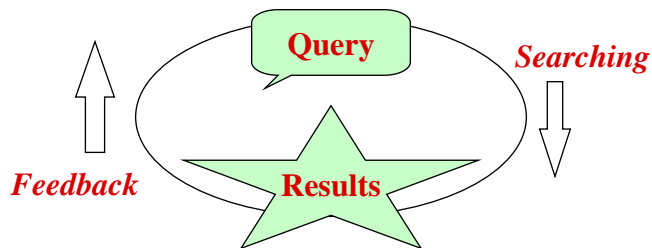


Hints on how to use information sources: work iteratively

Work iteratively =

search, investigate your results, refine your search, search again, and so on;

do not try to find everything in 1 step, with 1 search.



Hints on how to use information sources: work iteratively: example

When you search a database with subject keywords from a controlled list, added to each record:

1. Search with search terms that you know
2. Investigate the results and select good, relevant items
3. Look for the keywords added to these items
4. Select the good, relevant keywords
5. Formulate a new search with these keywords added
6. Execute the new search
7. Repeat the procedure



!! Task - Assignment !!

Search in the freely accessible *ERIC* database for documents on “courses offered through the web in the field of architecture, or history, or computer applications”.

This is not easy, because words like web, architecture, history, and computers, can have other meanings than titles of courses. Therefore, find and use the controlled subject terms that are added by the database producer and see that the results are better.



*“The ability to ask the right question
is more than half the battle of finding the answer.”*

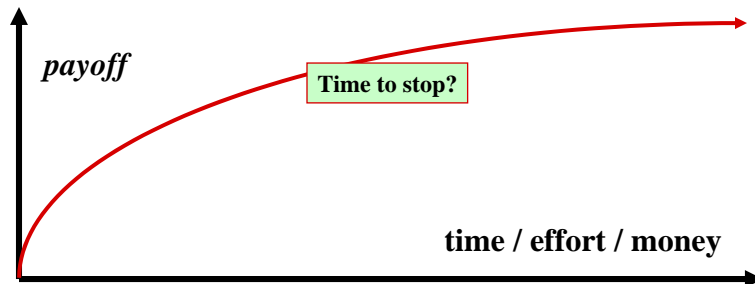
Thomas J. Watson



Hints on how to use information sources: when to stop searching?

Develop a feel for the “curve of diminishing returns”:

If you spend too much time, effort, and/or money with too few benefits, you should stop.



- **You are free to copy, distribute, display this work under the following conditions:**



» Attribution:

You must mention the author.



» Noncommercial:

You may not use this work for commercial purposes.



» No Derivative Works:

You may not change, modify, alter, transform, or build upon this work.

- **For any reuse or distribution, you must make clear to others the license terms of this work.**