Insights On
Relation Matching
And
Conceptual Graphs in IR

Presented By:

Dhiraj Bashyal(MSIISE05)

Outline

- Relation Matching
- Conceptual Graphs in IR

Information Retrieval

- **The processing of documents containing free text, so that they can be rapidly retrieved based on keywords specified in a user's query**
- The aim of information retrieval is **to provide the user with the “best possible" information from a database**
- A model of information retrieval predicts and explains what a user will find in relevance to the given query

Relation Matching[1]

- IR researchers in the late 1980s and early 1990s that the maximum attainable retrieval performance using keyword matching methods
- two ways in which relations can improve information retrieval effectiveness – through **relation matching** and through **query expansion**

Relation Matching[2]

- Relation matching
  - Both the **concepts** and the **relations** between the concepts as expressed in the **user's query are matched** with concepts and relations in documents
  - system first performs **concept or word matching**
- Query expansion
  - query concepts that are found in the document, the system further checks whether **the relations expressed between the query concepts match the relations expressed between the document concepts.**

Relation Matching[3] - Example

Consider the sentence Harry loves Sally.

If we use the sentence as a query in a keyword matching system, the system would look for documents containing the terms **Harry*, Sally* and love*** (where "*" is the truncation sign),

(1) Harry loves Sally.

(2) Sally loves Harry, but Harry hates Sally.

(3) Harry’s best friend loves Sally’s best friend.

(4) Harry and Sally loves pizza.

(5) Harry’s love for Sally is beyond doubt.

Relation Matching[4] - Example

- For example sentences from previous slide, sentences (1) and (5) should be ranked higher than the other sentences.
  - With relation matching, a document that not only has the keywords Harry, Sally and love but also expresses **the correct relation between the concepts** would be given a higher ranking in the retrieval results
- Relation matching improves **retrieval precision by reducing the number of non-relevant documents** retrieved
- System uses the additional criteria of relation matches to eliminate some non-relevant documents that would otherwise be retrieved by **keyword matching**

Main issues and research questions relating to relation matching[1]

- Comparison with keyword matching
  - To what extent does the use of relation matching improve retrieval effectiveness compared with keyword matching alone
  - For example : let's talk the word Bank
- Comparison with word proximity matching
  - To what extent does the use of relation matching improve retrieval results compared with using word proximity information?
  - Example Good: fine, excellent, great.

Main issues and research questions relating to relation matching[2]

- The difficulty of **identifying relations automatically**, especially when the **database is not limited** to a narrow subject area
  - Are simple methods of identifying relations (and the relatively low accuracy) good enough to yield a material improvement in retrieval effectiveness?
  - Will more accurate identification of relations yield better retrieval results than simple methods?

- The relation matching method
  - There are several ways of identifying relations automatically and several ways of performing relation matching.
  - There are also different types of relations and different sets of relations used by different researchers

Main issues and research questions relating to relation matching[3]

- The method of combining relation with keyword matching
  - What is the relative importance of keyword and relation matches in information retrieval?
  - How can relation matching be combined with keyword matching to estimate the likelihood that a document is relevant to the user?
  - Should different types of relations be weighted differently?
  - It is difficult to match the word by word

- The circumstances in which relation matching is important. Relation matching may not be helpful in all situations.
  - For what types of queries, documents, subject areas and applications is relation matching helpful in improving retrieval results?

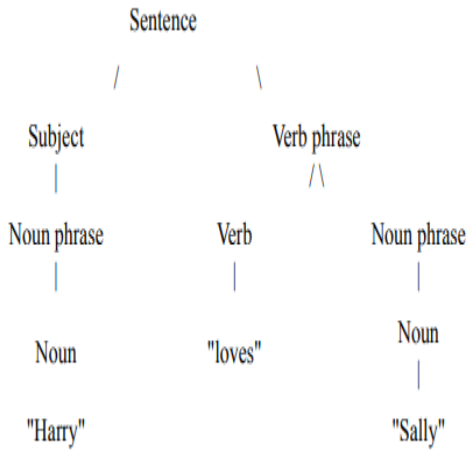
Matching With Manually Identified Relations

- Some **human indexer** has indicated that there is a relationship between two or more concepts in the content of the document
- Use of relations in manual indexing has not caught on probably because its effectiveness has not been clearly demonstrated
- **The effective use of relations requires training both for the indexer and the user**

Matching With Automatically Identified Relations

- Syntactic relations refer to relations that are derived from the syntactic structure of the sentence.
- Determining the syntactic structure of a sentence is one of the processing steps needed for determining the semantic relations and the meaning of the sentence

The following is a simple syntactic representation of the sentence Harry loves Sally:



Matching With Automatically Identified Relations

semantic relations - Examples

- For example, the same semantic relations between Harry, love and Sally as expressed in the sentence Harry loves Sally can be expressed in other ways:
  - **Harry's love for Sally is beyond doubt.**
  - **Harry declared his love for Sally.**
  - **Sally is Harry's one true love.**
  - **Sally is the object of Harry's love.**
  - **Sally is the love of Harry's life.**
  - **Sally, Harry's beloved, likes pizza.**

  Syntactic relation matching can fail to find a match between similar semantic relations if the relations are expressed using different syntactic structures
  - Information is retrieved from this knowledge store by comparing the information in the knowledge store with the knowledge representation of the user's query. Such a system is called a **conceptual information retrieval system**.

Factors that affect the usefulness of relation matching

- the accuracy of the automatic identification of relations
- the method used for **calculating the retrieval scores** (e.g. a tree matching method or term matching method) .
- the type of **documents** and the type of **queries**.
- the type of **relations** used and the set of relations used (i.e. syntactic or semantic, and which particular relations?)
- the degree of **relational ambiguity** between the concepts linked by a relation

Relational ambiguity

- For example, if the **words eat and apple** occur in the **same sentence**, we can quite confidently predict that apple has the patient (i.e. object) relation to the word eat without even reading the sentence. There is little relational ambiguity in this case

Relation matching with a wildcard

- Relation matching with a wildcard is helpful because it allows a match in the following cases:
  - when one member of the relation is specified in a different sentence in the document, as in the following two examples
  - (1a) The policeman surprised a burglar.
  - (1b) In the ensuing struggle, he killed the burglar.
  - (2a) The policeman surprised a burglar.
  - (2b) The burglar was killed in the ensuing struggle

  In both examples, there is a causal connection between the policeman and the burglar's death. In example (1), an anaphor is used in the second sentence to refer to the policeman in the first sentence.

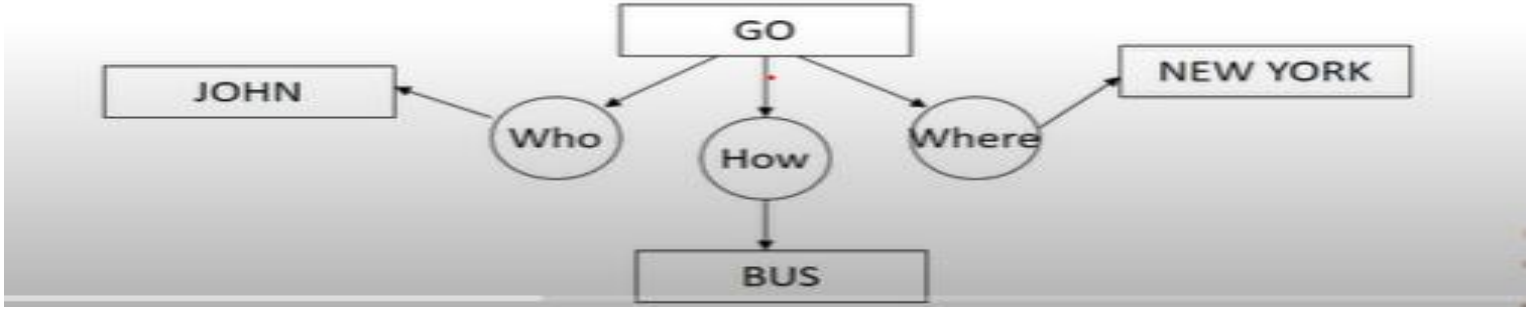
  In example (2), the policeman is not referred to in the second sentence but is implied by the context.

Future Work

- the **weighting for causal relation** matches had to be customized for individual queries
- it is not realistic to expect the user to give relevance judgements on a large sample of documents
- The effect of **one relation may be small and hard to detect**.
- Processing two or more relations in a query may give us a clearer idea of the effectiveness of relation matching

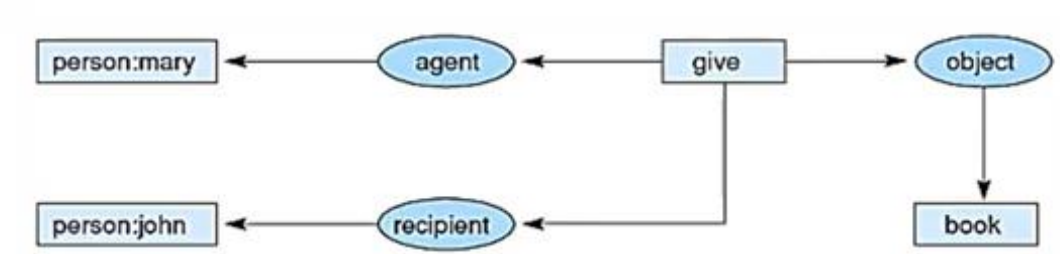
Conceptual Graph Matching

- conceptual graph are semantic nets representing the meaning of (simple)sentences in natural language
- Two types of nodes :
  - concept node: there are two types of concepts, individual concepts and generic concepts
  - The concept nodes represent entities, attributes, states and events in the world
  - The relation nodes show how the concepts are interconnected
  - Relation nodes(binary relations between concepts)



Conceptual Graph Matching - Examples

- Graph of “marry give john the book”



This graph uses conceptual relations to represent the case of the verb “to give” and indicates the way in which conceptual graphs are used to model the semantics of natural languages

Information Retrieval with Conceptual Graph Matching[3]
Comparison of conceptual graphs

- In general terms, our algorithm for the comparison of two conceptual graph representations of two texts consists of two main parts:
  - **Find the intersection of the two (set of) graphs,**
  - **Measure the similarity between the two (set of) graphs as the relative size of each one of their intersection graphs**

- In general, we can find more than one subgraph as the intersection of the initial graphs, but the measurement of similarity is applied to each one of them separately, and **only the highest value is kept**

Comparison of conceptual graphs[2]

- In the first step, we build the intersection G1 ∩ G2 = Gc of the two original conceptual graphs G1 and G2 . This intersection consists of the following elements:
  - All concept nodes that appear in both original conceptual graphs G1 and G2
  - All relation nodes that appear in both G1 and G2 and relate the same concept nodes.

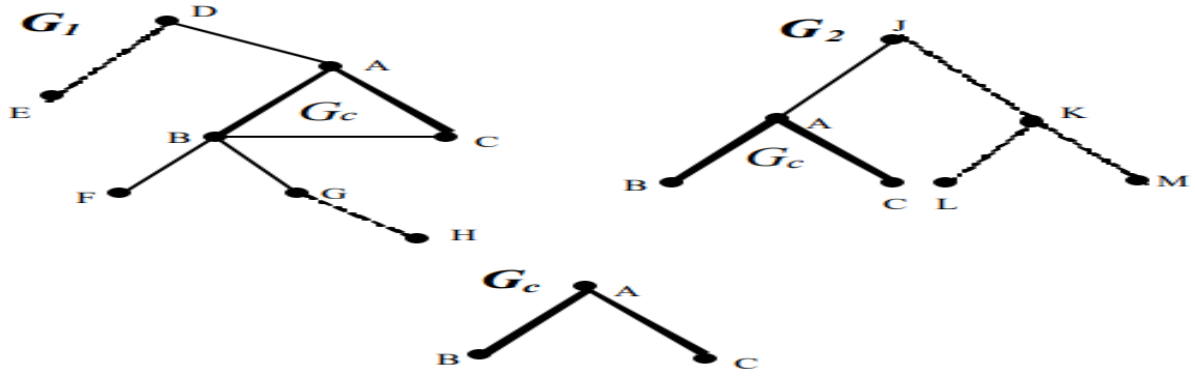


Fig.1. Intersection of two conceptual graphs

Information Retrieval with Conceptual Graph Matching[4]
Comparison of conceptual graphs[1]

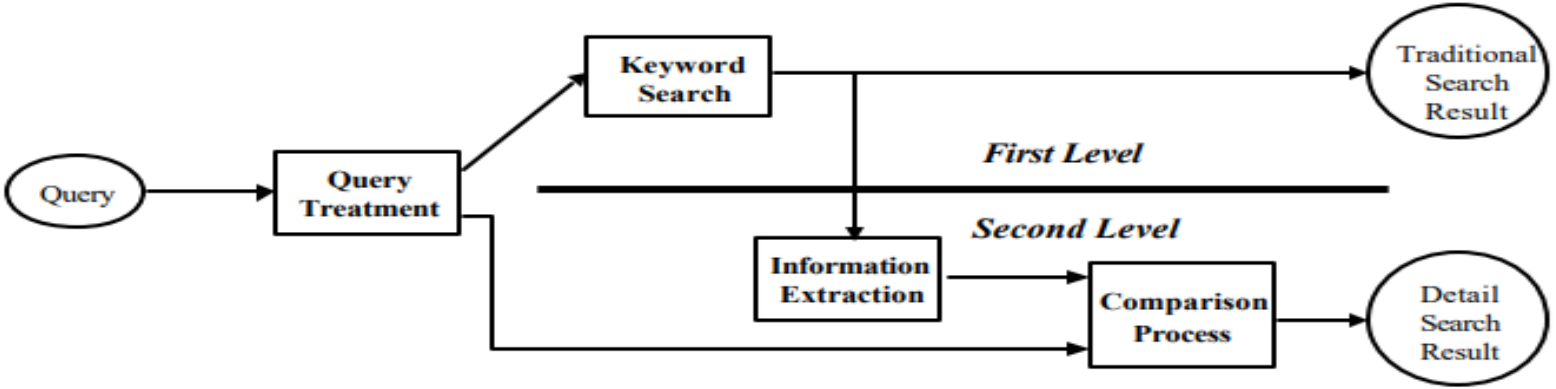


Fig. 3. Calculation of relational similarity.

References

[1] G. Salton. The SMART Retrieval System. Englewood Cliffs, N.J., 1971. Prentice Hall, Inc.

[2] C.J. van Rijsbergen. Information Retrieval. Butterworths, second edition, 1979.

[3] Chris H. Q. Ding, A Similarity-Based Probability Model for Latent Semantic Indexing, Proc. of SIGIR'99, Berkeley, August 1999

[4] Levow, G.-A., Oard, D. W., and Resnik, P., Dictionary-based techniques for cross-language information retrieval, Information Processing and Management, 41, 523–547, 2005.

[5] Khoo, Christopher Soo-Guan (1997). "The Use of Relation Matching in Information Retrieval". Electronic Journal ISSN 1058-6768, September 1997

Thank you !!!