Improving Ranking of the PLIR System By Local and Global Approaches

Elham Ashoori, Farhad Orumchian, and Babak Nadjar Araabi
Control and Intelligent Processing Center of Excellence
Department of Electrical and Computer Engineering
University of Tehran
IRAN
ashoori@ut.ac.ir foroumchian@acm.org araabi@ut.ac.ir

Abstract: The PLIR system is an experimental high precision retrieval system based on the theory of the plausible reasoning of Collins and Michalsky. The PLIR system retrieves documents through plausible inferences. These inferences could be considered as sources of evidence of relevance of a document to a query or query term. A series of experiments were conducted to improve the quality of ranking of the PLIR system. For these experiments the application of Dempster-Shafer theory(DS) of evidence was considered for combining the evidence gathered through plausible inferences. The experiments were conducted with different assumptions and settings using the titles and abstracts of the CACM text collection. Two approaches in combining evidences of plausible inferences were discussed and it was observed that the application of DS theory and weighting inferences based on their overall usefulness in retrieval has improved the quality of ranking of the system.

Keywords: Plausible Reasoning, Dempster-Shafer Theory of Evidence, Plausible Inferences, Information Retrieval and High Precision Information Retrieval.

1 Introduction

The PLIR system is an experimental high precision retrieval system that attempts to simulate the reasoning aspect of a reference librarian when trying to reformulate a query and find other terms or references of interest to the user. The theory of plausible reasoning developed by Collins and Michalski [1] has been utilized for this purpose. They developed the theory for question-answering situations where information is incomplete or uncertain or dynamically changing. It consists of a set of inferences modeled after inferences used by human beings faced with similar situations. Many well-known logics such as predicate logic are subsumed by this theory. Therefore it seemed reasonable to formulate and investigate research questions such as “Is it possible to represent document contents using primitives of the theory of plausible reasoning?”, “Is it possible to represent user queries by primitives of the theory?”, “Is it possible to retrieve references by using plausible inferences of the theory”, "Does expressive power of plausible reasoning subsume other logics and inferences proposed for information retrieval?”, and “Does plausible reasoning perform as well as other simpler but powerful models such as the vector space model?”. These questions are addressed in [5], [6] but in this paper the main research question is “Which approach in combining evidences of plausible inferences is more effective?” In the following sections, first a brief introduction to the theory of plausible reasoning is presented and then an attempt is made to answer the above research question.

2 Basics of The Theory of Plausible Reasoning

For approximately 15 years, Collins and his colleagues have been collecting and organizing a wide variety of human plausible inferences made from incomplete and inconsistent information. These observations led to the development of a descriptive theory of human plausible inferences that categorizes plausible inferences in terms of a set of frequently recurring inference patterns and a set of transformations on those patterns. According to the theory, a specific inference combines an inference pattern with a transformation that relates the available knowledge to the questions based on some relationship (i.e. generalization, specialization, similarity or dissimilarity) between them. The primitives of the theory consist of basic expressions, operators and certainty parameters. In the formal notation of the theory, the statement "coffee grows in the Llanos" might be written:

$$\text{GROWS-IN (Lilanos) = Coffee, } \gamma = 1.0$$

This statement has the *descriptor* GROWS-IN applied to the argument Lianos and the *referent* Coffee. The certainty of the statement (γ) is 1.0 since it declares a fact about the Lianos. The pair descriptor and argument is called a *term*. Expressions are terms associated with one or more referents. All descriptors, arguments and referents are nodes in (several) semantic hierarchies. Any node in the semantic network can be used as a descriptor, argument or referent when appropriate. Figure 1 shows the basic elements of the core theory.

There are many parameters for handling uncertainty in this theory. However there is no complete agreement on their computational definitions and different computer models of the theory have implemented them in different ways. The definitions of the most important ones according to [1] are:

1. γ The degree of certainty or belief that an expression is true.
2. ϕ Frequency of the referent in the domain of the descriptor (e.g. a large percentage of birds fly).
3. τ Degree of typicality of a subset within a set (e.g. robin is a typical bird and ostrich is not a typical bird).
4. δ Dominance of a subset in a set (e.g. chickens are not a large percentage of birds but are a large percentage of barnyard fowl).
5. σ Degree of similarity of one set to another set.
6. α Conditional likelihood that the right-hand side of a dependency or implication has a particular value (referent) given that the left-hand side has a particular value.
7. β Conditional likelihood that the left-hand side of a dependency or implication has a particular value (referent) given that the right-hand side has a particular value.
8. μ_r Multiplicity of the referent (e.g., many minerals are produced by a country like Venezuela).
9. μ_a Multiplicity of the argument (e.g., many countries produce a mineral like oil).

The theory has a rich set of transforms and inferences, which provides a means of converting one statement to another and inferring unknown concepts from the known ones. Interested readers are referred to the references [1] and [2] for an in depth explanation or reference [4] for an implementation of the theory.

referents r1, r2, r3, { r2 … }
e.g., collie, brown and white,
{brown…} (translation: brown plus other colors)
arguments a1, a2, F(a1)
e.g., Fifo, Collie, Fido's master
descriptor d1,d2
e.g., breed, color
terms d1(a1), d2(a2), d2(d1(a1))
e.g., breed(Fido), color(Collie), color(breed (Fido))
statement d1(a1) = {r1} : γ , φ
e.g. , means-of-locomotion (bird) = {flying} : certain, high frequency
(translation: I am certain almost all birds fly
dependencies between terms
d1(a1) <----> d2(a1) : α , β , γ
e.g., latitude(place) <----> average-temperature(place): moderate, moderate, certain (translation: I am certain that latitude constrains average temperature constrains the latitude with moderate reliability)
implications between statements
d1(a1) = r1 <===> d2(a1) = r2 : α , β , γ
e.g., grain (place) = { rice,… } <====> rainfall (place) = heavy :
high , low , certain (translation : I am certain that if a place produces rice, it implies the place has heavy rainfalls with high reliability , but that if a place has heavy rainfall it only implies the place produces rice with low reliability)

Fig.1 Elements of Expressions in the Core Plausible Reasoning Theory

3 Knowledge Representation

This section explains how the text of documents is scanned and relations extracted, and how these relations are transformed into logical terms and statements. The calculation of certainty parameters for one type of relations is also given here. The knowledge (documents, index terms and phrases and their relationships) has been represented in the form of a hierarchical semantic network. Four different relationships connect index terms and phrases to each other and to the documents. These are: Broader-Narrower relationship (BN), X, Y and Reference (REF) relationships. Here, two phrases are considered to have the *Broader-Narrower (BN)* relationship if one ends with the other. X is a relationship between parts of a phrase with the phrase itself. The Y relationship is very similar to the X relationship but it is extracted from the text using some linguistic clues. The Documents in a collection are connected to their phrases and index words by REF relations. Figure 2 shows some examples of these relations.

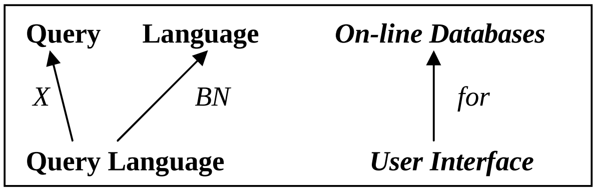

FIG.2 Examples of Relation X, BN and Y in two phrases

If we consider the example of the "Query" and "Query Language" in figure 2, a "query" is an instance of what can be expressed by the "query language". Therefore, the concept query could be a member of the set of what is represented by the concept query language. In example of the phrase "User Interface for On-line Databases", the preposition *for* is a clue indicating some kind of relationship between the concept "User Interface" and the concept of "On-line Database".

BN relationships have intuitive meanings but X and Y relationships manifest their existence but either we are unable to define their meaning exactly or it is not necessary in this experiments to do so. The uncertainty of a relationship is described by four parameters. These are the Frequency (F), Confidence(C), Dominance (D) and Acceptability (A) parameters. For instance, the Value of certainty parameters For Broader-Narrower Relationship is proposed in figure 3.

4 Information Retrieval By Plausible Inferences

There are four elements in a logic based IR system.

Those are the description of documents, the representation of queries, a knowledge base containing domain knowledge and a set of inference rules. This study also acknowledges that retrieval is inference but relevance is not material implication [11]. Here, a document is retrieved only if its partial description can be inferred from a query description. Thus the retrieval process is expanding a query description by applying a set of inference rules continuously on the description of the query and inferring other related concepts, logical terms and statements until locating a document or documents

Frequency F= Max(f, C1)

f = frequency of co-occurrence of a phrase with its parent in the same document.

C1= constant.

Confidence C= Max ($\sum$ws. fs - $\sum$wr . fr, C2)

ws= weight of supportive evidence or clue.

fs = frequency of supportive evidence.

wr= weight of rejecting evidence or clue.

fr = frequency of rejecting evidence.

Dominance $D = \dfrac{F_{child}\, C_{child}}{\sum F_{children} C_{children}}$

Acceptability A= C3, constant

C2 = constant

Fig . 3 The Value Of Certainty Parameters For Broader-Narrower Relationship

which are described partially by these concepts or logical terms or statements.

4.1 Document Representation

In this model, documents are represented in possible worlds by a partial set of phrases, logical terms and logical statements, i.e., the representation of a document is not limited to the set of its representative phrases or logical terms and statements. Any concept that can be inferred from representation, by plausible reasoning using the given knowledge base, is also a representative of the document content. A possible world is the finite set of all phrases and logical terms and statements that can be inferred from the partial representation of a document in a snapshot of the knowledge base.

Since the knowledge base is dynamically changing, so are the possible worlds.

In its simplest form, a typical document such as Van Rijsbergen’s 1986 article entitled “A non-classical logic for information retrieval” can be represented as follows:

1. REF (Information Retrieval) = {doc#l }
2. REF (Non-classical Logic)= {doc#l }
3. REF (Non-classical Logic (Information Retrieval))= {doc#l }

The first statement indicates the concept Information Retrieval is a reference for doc#l. The second statement states that the concept Non-classical Logic is a reference for doc#l. The third statement expresses that the term Non-classical Logic (Information retrieval) is a reference for doc#l.

4.2 Representing a query as an incomplete statement

A query can be represented as an incomplete logical statement in which the descriptor is the keyword REF (reference) and its argument is the subject in which the user is interested. The referents of this statement i.e. the desired documents, are unknown. A typical query in logical notation will have the form:

REF (A-Subject)={?}

Therefore the retrieval process can be viewed as the process of finding referents and completing this incomplete sentence.

Queries in the CACM database contain phrases such as time Sharing System or intermediate languages or sentence fragments like communication mechanisms for disjoint process.

A query with a single phrase, such as "Time Sharing System", can be formulated as:

REF (Phrase) = (?)

For example:

REF (Time Sharing System)={?}

A query consisting of a sentence fragment can be treated as a regular text. Therefore it can be scanned for extracting its logical terms. For example, consider the query number 32 from the CACM collection:

“I’m especially interested in any heuristic algorithms for graph coloring and ,...”

This query contains the sentence fragment “heuristic algorithms for graph coloring”. That can be converted into a logical term, which is revealed by the proposition *for*. The query can be represented as:

REF(heuristic algorithm (graph coloring))= {?}

Queries with more than one concept or term can be represented as a set of simple queries and the system can retrieve a set of references for each one separately and then reexamine the sets by combining the confidence on references, which are members of more than one set. Then the sets can be joined and the resulting set can be sorted according to the confidence value.

4.3 Document Retrieval by completing an incomplete query statement

Since the query is always represented as an incomplete statement the retrieval process can be seen as finding referents to complete the query statement. The first step is similar to any other retrieval system that is to look for documents, which are indexed by the query terms. Figure 4 shows this process in logical form. There are two situations, since a document could be indexed by either a concept or a logical term. In the example, user’s interest in "Automatic Translation of Machine Language Programs", is represented as:

REF(Automatic Translation (Machine Language programs))

which could be read as ‘Inferences for Automatic Translation of Machine Language Programs”. Document number 167 in the CACM collection is indexed with this term; therefore it is a referent for the query. This is a case of direct indexing where the document is indexed by the term.

REF (Phrasel) = { ? }
REF (Phrasel) = {doc# } : δ, A

REF (Phrasel) = {doc#} : γ = F1(δ, A)
or
REF (Phrasel (Phrase2)) = { ? }
REF (Phrasel (Phrase2)) = {doc# } : δ, A

REF(Phrasel (Phrase2))= {doc#} γ = F1(δ, A)

Example:
Query:
REF (Automatic Translation (Machine Language Program)) = {?}
REF(Automatic Translation(Machine Language Program)) ={ Doc#167 } δ = 0.02, A= 0.5

REF(Automatic Translation (Machine Language program)) = {Doc#167 } δ = 0.22

FIG.4 Finding References By Completing Incomplete Query Statement, Direct Approach

There is another situation where a document is not directly indexed by the query term, yet the document still can be retrieved if it is indexed by a concept, which is the referent of the query term. This can be seen as indirect indexing and it is demonstrated in Figure 5. In this example, it is assumed that there is no document indexed by the query term "language (programming)". However, there are documents which are indexed by the concept "Fortran". This concept is a referent of the query term. So in the inference, the query term is replaced by its referent making it possible for the document to be retrieved as a referent of the query. This referent is associated with the document 1150 in the CACM collection. Therefore this document can be retrieved although it is not directly indexed by the original query term.

REF(Phrasel (Phrase2)) = { ? }
Phrasel (Phrase2) ={ Phrase3) : γ 2 = F1(γ 1 , A1)
REF(Phrase3) = {doc# } : δ 1, A2

REF(Phrasel (Phrase2)) = { doc # } :
γ = F2 (γ 2, δ 1, A2)

Example:
Query:
REF(Language (Programming g))={?}
Language (Programming)= { Fortran } : γ = 0.82
REF(Fortran) = { Doc#l150 }: δ = 0.17, A= 0.5

REF(Language (Programming g)) = { Doc#l150 }:
γ = 0.24

FIG. 5 Finding References By Completing Incomplete Query Statement, Indirect Approach

The certainty of the relevance of a document depends on two factors, first the dominance of that document among other related documents and second acceptability of the document as a viable reference for the term in the query. The dominance is computed for each document by the system by propagating the dominance of all the terms/ statements involved in the inference process. On the other hand, the acceptability is computed from the feedback of the users on conclusions drawn by the system.
The methods of computation of certainty parameters are not discussed here since they are more experimental than theoretical.
The value of certainty (γ) ranges from 0 to 1, where 1 indicates 100% belief in the correctness of a statement and 0 means that there is no information about the truth of the statement. The acceptability (A) ranges from 0 to 1 where one indicates that 100% of users accept the statement or believe in its truth, while 0 only expresses that there is no special information about how users perceive the truth of a statement.

4.4 Document Retrieval Inferences Using Referent Transforms

In this application of the theory, there are several referent and argument -based transforms (generalization, specialization, similarity, dissimilarity) [5]. For reasons of brevity only the specialization (SPEC-) based referent transform is described in detail here.
The SPEC-based referent transform in the core theory is a strategy to utilize the part-of and kind-of relationships to find other referents for a given statement. In the IR situation this strategy could be applied to the concepts found relevant in earlier stages of retrieval in order to find other relevant concepts and their associated references. As an example, let’s consider the query:

REF(algorithm (distributed)) = { ?}

which expresses that there is an interest in references for distributed algorithms. Let's assume that the phrase "concurrent-program" has been already established as a referent for the term "algorithm (distributed)". By applying the SPEC-based transform, all the children of the node "concurrent-program" can be examined for retrieval, In this example, the phrase "concurrent process" is more specific than the "concurrent-program" and is associated with document 3128 in the collection. So, this document can be presented to the user as a reference. Figure 6 illustrates this transformation.
The example in Figure 6 is based on the query number 7 from the CACM collection, where a user requests references for "distributed algorithms" and expresses interest in "synchronization by using message passing" among other things. Document number 3128 is about "synchronization of concurrent processes". The inference starts by identifying "concurrent programs" as an example of distributed algorithms. Concurrent processes are specialization of "concurrent programs" in the context (CX) of "synchronization", which is of interest to the user. The dependency between "synchronization algorithms" and "distributed algorithms" can be established by looking at their co-occurrence in the collection. In this case they have appeared in the same documents in 10% of cases. Line 4 adds no new information and it is only for consistency between the example and the symbolic inference. Line 5 concludes that concurrent process is also an example of distributed algorithm. Line 6 identifies document number 3128 as a reference for concurrent process, and therefore it is a reference for distributed algorithm.

REF(d(a))={?}
1- d(**a**) ={**r**} : $\gamma 1$: premise
2- **r'** SPEC **r** in **CX** (**d**, D(**d**)) : $\delta 1$, $A1$: premise
3- D(**d**) <----> A(**d**) : $\alpha 1$, $\gamma 2$: premise
4- **a** SPEC **A** : $\delta 2$, $A2$: premise
5- d(**a**)= {**r'**} **:** $\gamma 3$ **= F3 (**$\gamma 1, \delta 1, A1, \alpha 1, \gamma 2, \delta 2, A2$**)** : From 1,2,3,4 by **SPEC-based Referent Transform**
6- **REF(r')**={**x**} **:** $\delta 3$, $A3$: premise

REF(d(**a**))= {**x**} **:** γ **= F2 (**$\gamma 3, \delta 3, A3$**)**
From 5, 6 by Indirect Approach

Example :
Query: REF(algorithm (distributed))={ ? }
1- Algorithm (**distributed**) = {**concurrent_program**} : $\gamma 1 = 0.30$
2- **Concurrent-process** SPEC **concurrent_program** in **CX** (**algorithm**, synchroronization (**algorithm**))
: $\delta 1 = 0.40$, $A1 = 0.5$
3- Syncbronization (**algorithm**) <----> distributed (**algorithm**)
: $\alpha 1 = 0.10$, $\gamma 2 = 0.6$
4 - **Distributed** SPEC **distributed** : $\delta 2 = 1.0$, $A2 = 0$
5 - Algorithm (**distributed**) = (**concurrent-process**)
: $\gamma 3 = 0.03$
6 - **REF** (**concurrent-process**) = {**Doc#3128**}
: $\delta 3 = 0.01$, $A3 = 0.5$

REF (algorithm (**distributed**))= { **Doc#3128** } γ **= 0.002**

FIG.6 Finding References by SPEC-based Referent Transform

5 Retrieval Algorithm

The PLIR system executes the following steps in order to find the references for a typical query such as REF(a(b)) = {?}.

1- Simple Retrieval
Find all references for

- all the terms such as a(b) that are in query
- all the referents such as c for the term a(b) where a(b) = {c}

2- Use Relationships & Inferences
Find all references for

- all the terms such as d(e) with mutual dependency relationship term a(b) where a (b) <---> d(e)
- all the referents such as f with SPEC , GEN and SIM relationship with referent C where a(b) = { f} and f SPEC , GEN , SIM c
- all the referents such as g for the term d(e) where d(e) = {g}
- all the terms such as i(j) with mutual dependency relationship with d(e) Where d(e) <---> i(j)

3- Repeat Step Two
As long as user seeks new references repeat step 2 for newly found terms and referents.

Since a term, referent or document could be reached through several different relationships or inference, therefore a method for combining the certainty of different certainty values coming from different inference should be devised.

Different certainty values calculated by different inferences or relationships have been treated as different sources of evidence. Two different approaches namely global and local have been tried with different weightings and theories for combining these values.

Figure 7 depicts the local combination method .In the next section different variations of each approach that has been investigated is explained.

6 Experiments

Several experiments have been conducted using the CACM collection to investigate the effectiveness of plausible inferences. 48 queries out of 64 standard CACM queries were used in these experiments. The effectiveness is measured by Precision and Recall.

After processing the CACM collection a knowledge base was built which contains the documents, phrases and logical terms, X, Y, REF and BN relationships.

Documents are represented by phrases and logical terms which occur in them. Two methods for computing importance of a phrase or term have been compared in these experiments. One weight is called Dominance and implemented as:

$$\text{Dominance D} = \frac{F(child)\ C\ (child)}{\sum F(children)\ C\ (children)}$$

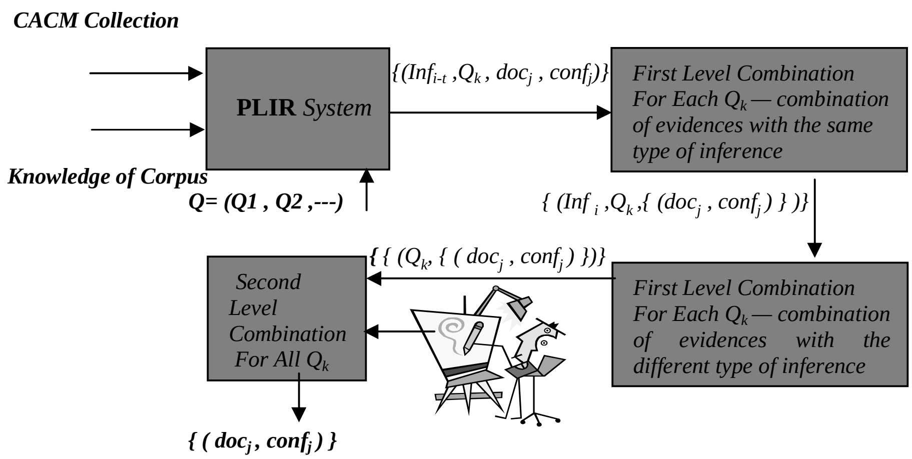

Fig.7 Multi Level Local Approach to combine different Evidences in Information Retrieval

where F(child) is the frequency of a child node among other children of the parent node, C(child) is the confidence on a particular child and the sum is over all children of the parent node. The other weighting is the traditional tf.idf weight. Two plausible retrieval systems were built, one based on dominance and the other based on tf.idf.

The vector space model with tf.idf weighting also was implemented and used as benchmark in these experiments. The vector space model represents queries and documents as vectors in multidimensional space. Two different version of the vector space also was implemented.One of them only used vectors of single words, and the other utilized the vectors of single words and phrases. The phrases were produced by PLIR system. Also the query vectors were extended with additional single words and phrases that were related to query in the PLIR knowledge Base. So the only difference between the PLIR systems and Vector Space systems were on their reasoning power.

Since plausible reasoning, the way it is implemented here, retrieves many fewer documents than vector space model, the output of the plausible reasoning systems were padded with that of vector models after eliminating redundant documents.
All the queries submitted to the plausible retrieval systems are manually generated after reading the text of the original queries and they are expressed as logical terms with REF as descriptor and another logical term as argument e.g. REF (A (B)). Only a subset of plausible inferences as listed below have been implemented: [5], [6]

1- Finding referents and completing incomplete query statements, direct approach by using only terms.
2- Finding referents and completing incomplete query statements,
3- Indirect approach.
4- SPEC-based Referent Transform.
5- SPEC-based Argument Transform.
6- A special case of Argument-based Mutual Dependency where both descriptor and argument of a term are specifications of the descriptor and the argument of another term. In Other words:

IF a SPEC A AND b SPEC B THEN
a(b) <------> A(B)

For global combination approach, two different methods were used. In global-pessimistic approach, the highest value of different certainty values was considered. On the other hand, the global-optimistic approach considered the inferences to be independent of each other and combined the confidences as independent probabilities. The local combination approaches combined the confidence values in two or three stages. The use of Dempster-Shafer (DS) theory of evidence [3], weighted addition and smets rule in combining the evidences provided by plausible inferences was investigated in first level. In the second level, we combined all these degrees of relevance to estimate the overall degree of relevance to the whole query by averaging. Table 1 shows different experiments for local approach.
Experiments differ in 4 below aspects:
1- The method of computing confidence on one inference that could be simple or weighted.

2- The combination of evidences of relevance from different inferences for each query term that could be Dempster's rule of combination or addition.
3- The method of computing unassigned certainty
4- The combination of evidences of relevance from

Table 1 Different Experiment for Local Approach

Exp#	$m_i(doc_j)$	$m_i(T)$	$m_i(\phi)$	m (doc_j)
Exp#1	c_i	$1-c_i$	0*	$m1 \otimes m2 \otimes m3 \otimes ...m5$
Exp#2	$w_i * c_i$	$w_i * (1-c_i)$	< >0*	$m1 \otimes m2 \otimes m3 \otimes ...m5$
Exp#3	$w_i * c_i$	1- $w_i * c_i$	0*	$m1 \otimes m2 \otimes m3 \otimes ...m5$
Exp#4	$w_i * c_i$	$w_i * (1-w_i * c_i)$	< >0*	$m1 \otimes m2 \otimes m3 \otimes ...m5$
Exp#5	$w_i * c_i$			$m1 + m2 + ...m5$
Exp#6	$w_i * c_i$	$w_i * (1-c_i)$	< >0**	$m1 \otimes m2 \otimes m3 \otimes ...m5$

different query terms. For complete details of local approach experiments refer to [7].
In Exp#1 there is no preference among inferences; therefore the confidences are not weighted. The uncommitted belief is one minus the total belief assigned. The differences between this experiment and pessimistic approach in [6] are only the difference between global and local approach but the combination rule is the same. In Exp#2 to Exp#6, the confidence coming from better inferences (inferences correctly identifying relevant documents) are preferred over the confidence coming from other inference. In Exp#2, the weights of inferences are taken into consideration to decrease the uncommitted belief. The mass of null set is not equal to zero but at the end, the results are normalized. In Exp#3, the weights of inferences are taken into consideration to increase the uncommitted belief. The mass of null set is equal to zero**.** In Exp#4, everything is the same as Exp#2**,** but the uncommitted belief is more than Exp#2. In Exp#5, the uncommitted belief is not taken into account. For the second level of combination, an addition operator is used to combine the weighted confidences. Exp#6 is similar to Exp#2 but the results of mass functions are not normalized. In this experiment, we have considered the open world assumption [9] to model the capability of the reference librarian in this collection. The open world assumption reflects the idea that Ω might not contain the actual world [10]. So, Instead of normalization, we only do not consider $m_i(\phi)$ in computing normalization factor and divide the mass with a value less than the value that used for normalization in other experiments.

7 Results

In the experiments conducted it was demonstrated that the answer to the research questions: "Is plausible reasoning more effective than vector retrieval? ", and sub-questions such as "Does the dominance parameter perform as well as tf.idf weigths" is yes [5][6]. Figure 8 illustrates a comparison of plausible reasoning systems based on dominance and tf.idf weights with vector space systems using words and phrases. For this graphical representation, simple precisions are computed for each query at standard recall points by interpolation and then averaged over the queries for each standard point.
Other experiment focused on answering the research questions: "What is the best combination approach and method for PLIR system?". There are no differences in the performance of the two global approaches. The problem both have is that they assign the same confidence value to groups of documents. Therefore their output consists of buckets of documents. In table 1, the experiment Exp#0 is the global approach. For comparing the performances of the experiments with each other and to see how different methods break the ties, a modified version of precision is used. The definition of this precision is as below:

$$RP_i = r_i * p_i / r$$

Where:
r_i : number of retrieved documents at ith rank
p_i : precision at ith rank
r : total number of retrieved documents

This measure prefers experiments that have more relevant documents in higher ranks and ranks have lesser documents assigned to them. Figure 9 shows a comparison of overall performance on all queries.
Figure 10 depicts the query-by-query results of comparing the local approaches to the global approach (Exp#0). In all experiments with local approaches the ranking has improved over the global approaches. Each one of experiments produced more ranks with lesser number of documents in each rank. In these experiments Exp#6 has produced better results. Since the test data was small and a good number of documents were only retrieved by a single inference, therefore we do not feel comfortable to generalize the results. However, we believe this result may be reproducible in larger experiments. Therefore we are repeating the same tests on *data* collection *of TREC-9 Filtering Track*.

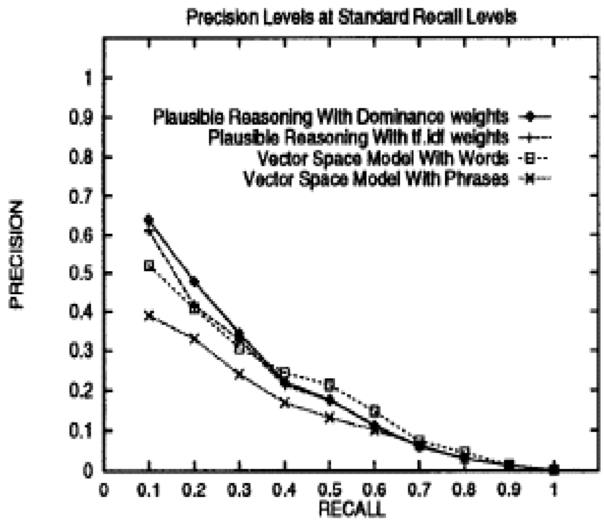

Fig.8 A comparison of plausible reasoning systems with vector space model.

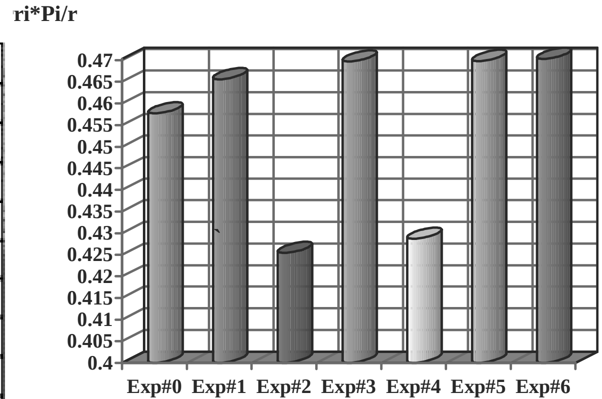

Fig 9 A Comparison of overall perfomance on all queries for the experiments with both global and local combination approach

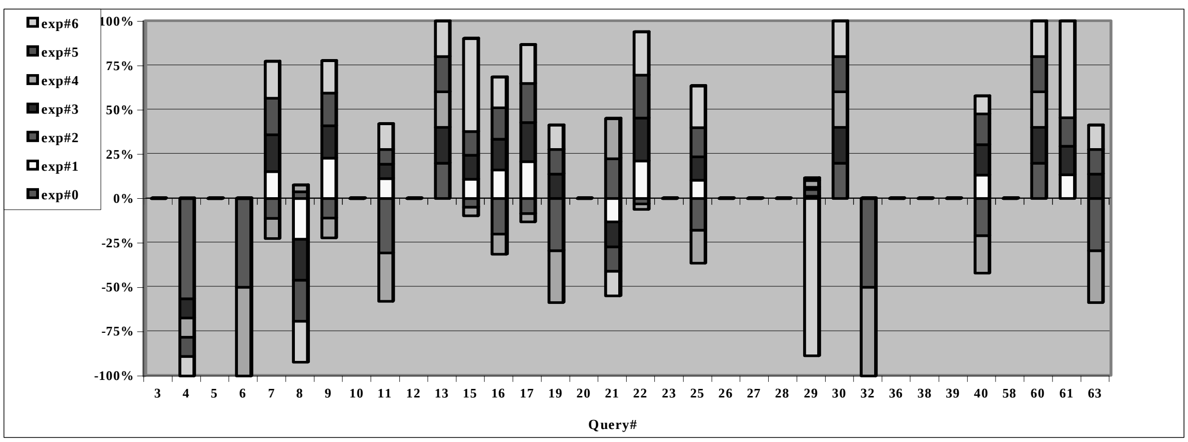

Fig.10 **The difference of comparing measure with Exp#0**

8 Conclusion

A series of experiments were conducted to improve the quality of ranking of the PLIR system. Since PLIR retrieves documents through plausible inferences, these inferences could be considered as sources of evidence of relevance of a document to a query or query term. The application of Dempster-Shafer theory of evidence was investigated for combining the evidences gathered through plausible inferences. Several experiments were conducted with different assumptions and settings.

In general, the local approaches for combining confidence values produce more qualitative ranking than global approaches. It seems that the application of DS theory in combining the evidences has contributed positively in this regards. It seems that the local approaches push the non-relevant documents to lower ranks. Therefore with a good method of computing a query-based threshold, one could eliminate many of non-relevant documents.

In these experiments we only touched the misleading inferences problem. Misleading inferences generate misleading evidences that can be taken into account when all evidences of relevance are combined. We also want to experiment with the misleading inferences by using user's relevance feedback and considering their weight in the evidence combination formula.

Another interesting idea that we just started to play with is: the open world assumption. In future we plan to use both ideas in combining evidences and computing the confidence of relevance. We also consider the DS combination function to combine the evidences that affect the relevance feedback to update the query instead of utilizing the Rocchio[8] formula.

References:

[1] A. Collins, and R. Michalski, The logic of Plausible Reasoning A core theory, Cognitive Science, vol. 13, 1989, pp. 1-49

[2] A. Collins, M. H. Burstejn, Modeling A Theory Of Human Plausible Reasoning, Artificial Intelligence III, 1988.

[3] A. Dempster, "Upper and lower probabilities induced by a multivalued mapping," Ann. Math. Stat., vol. 38, no. 2, pp. 325–339, 1967

[4] K. Dontas, M. Zemankova, APPLAUSE: An Implementation Of The Collins-Michalski Theory Of Plausible Reasoning, Proceedings of the 3rd International Symposium on Methodologies for Intelligent Systems, Torino, Italy.

[5] F. Oroumchian, Information Retrieval By Plausible Inferences: An Application Of The Theory Of Plausible Reasoning Of Collins And Michalski, Ph.D. Dissertation, School Of Computer And Information Science, Syracuse University, 1995.

[6] F. Oroumchian, and R.N. Oddy, An Application Of Plausible Reasoning To Information Retrieval, proc. of the 19th ACM SIGIR Conference on Research and Development in Information Retrieval, Zurich, Switzerland, pp. 244-252, 1996

[7] Farhad Oroumchian, Babak.N Araabi, Elham Ashoori, Using Plausible Inferences and Dempster shafer Theory Of Evidence For Filtering, 4th International Conference on Recent Advances in Soft Computing (RASC2002), Nottingham, United Kingdom, Dec 12-13. 2002

[8] J. Rocchio, "Relevance feedback in Information Retrieval,proc of the SMART Retrieval System -Experiments in Automatic Document process, pp 313-322, 1971.

[9] Ph. Smets, The nature of the unnormalized beliefs model, In D. Dubois, M. P. Wellman, B. D'Ambrosio, and Ph. Smets, editors, Uncertainty in Artificial Intelligence 92, Morgan Kaufman, pp. 292–297, San Mateo, CA, 1992.

[10] Ph. Smets, Belief Functions and The Transferable Belief Model, page at the web site of the Imprecise : 31/10/2002

[11] C. J. Van Rijsbergen, Toward An Information Logic, In N. J. Belkin &C. J. Van Rijsbergen (Eds.), Proceedings Of The 12th Annual International SIGIR Conference, 1989, Cambridge, MA.