January 1996

An Application of Plausible Reasoning to Information Retrieval
An Application of Plausible Reasoning to Information Retrieval

Farhad Oroumchian
Department Of Computer And Information Science
Syracuse University
Syracuse, USA, 13244
farhad@top.cis.syr.edu

Robert N. Oddy
School Of Information Studies
Syracuse University
Syracuse, USA, 13244 Syracuse, USA, 13244
rnoddy@mailbox.syr.edu

doi: 10.1145/243199.243271

1. Abstract

This work explores the use of plausible inferences as a means of retrieving relevant documents. Collins and Michalski’s theory of plausible reasoning has been modified to accommodate information retrieval. Methods are proposed to represent document contents by logical terms and statements, and queries by incomplete logical statements. Extensions to plausible inferences are discussed. Two versions of the extended plausible reasoning system were implemented, one using dominance weights (described in the paper) and the other using tf. Idf (Term Frequency Inverse Document Frequency) weights. Experiments were conducted using the titles and abstracts of the CACM collection and it was found that both versions of the extended plausible reasoning system are better than the vector space model and the system using dominance weights performed better than the system with tf. Idf weights.

2. Introduction

Applying computers to information retrieval converts the information problem into a representation problem where the information need is represented by a query and documents by only a few terms in conventional retrieval systems. The relevance judgment is modelled by a matching of representations of documents with queries. In this environment in which everything is only a representation of the reality, the usual way to predictively estimate the information seeker’s relevance judgment is a method to measure “how much” a document is “about” a query.

By this view, information retrieval research can be categorized according to the representations on which the authors focus. Natural Language Processing aims at improving representation of the content of documents. Thesaurus research examines query expansion to achieve an adequate representation of the information need of the user. Logic-based IR looks into ways to improve matching methods and to capture the notion of “aboutness”. Van Rijsbergen [12] develops a logic-based theory where retrieval is inference. Documents are represented partially in possible worlds and retrieved only if they imply a query. Another approach is to view the retrieval as a cognitive process, try to model the information seeker’s interest, and retrieve documents accordingly (e.g. Oddy [7]). These approaches are not mutually exclusive, there are many overlaps in their representations and it is possible to unify all these approaches under one framework.

Since people were doing Information Retrieval long before computers came along, it seems interesting to look at an example to see how reference librarians do their job and how these different approaches relate to that process. A reference librarian is not an authority or expert in the domain of subjects covered in the library’s collection, nor is (s)he familiar with all the jargon used in different subjects. (S)He depends more upon general and somewhat superficial familiarity than very specific and in depth knowledge of the subjects. A reference librarian also learns a great deal from library patrons. For example, (s)he learns successful search strategies in special domains and learns important concepts of the sublanguages used in subjects. Thus a reference librarian is able to reason with incomplete and uncertain and dynamically changing information.

In this study an attempt is made to simulate the reasoning aspect of a reference librarian when trying to reformulate a query and find other terms or references of interest to tie user. The theory of plausible reasoning developed by
Collins and Michalski [4] 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 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?”

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 questions.

3. Basics of the Theory of Plausible Reasoning
Collins and Michalski’s theory of plausible reasoning is empirically-based. 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 inference 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 system consist of basic expressions, operators, and certainty parameters. In the formal notation of the theory, the statement “coffee grows in the Lianos” might be written:

\[ \text{GROWS-IN (Lianos) = Coffee, } \gamma = 1.0 \]

This statement has the descriptor GROWS-IN applied to the argument Lianos and the referent Coffee. The certainty of the statement (Y) is 1.0 since it declares a fact about the Limos. The pair descriptor and argument is called a ~. 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 [4] are:

1. \( \gamma \) The degree of certainty or belief that an expression is true.
2. \( \phi \) Frequency of the referent in the domain of the descriptor (e.g. a large percentage of birds fly).
3. \( \tau \) Degree of typicality of a subset within a set (e.g. robin is a typical bird and ostrich is not a typical bird).
referents r1, r2, r3, {r2 ....}
  e.g., collie, brown and white,
  \{brown...\} (translation: brown plus other colors)
arguments a1, a2, F(a1)
  e.g., Fido, 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))
statements d1(a1) = \{ r1 \} : \gamma, \phi
  e.g., means-of-locomotion (bird) = \{ flying \} :
  certain, high frequency
  (translation: I am certain almost all birds fly)
dependencies between terms
  d1(a1) \langle-----\rangle d2(a1): \alpha, \beta, \gamma
  e.g., latitude (place) \langle-----\rangle average-temperature (place):
  moderate, moderate, certain (translation: I am certain that
  that latitude constrains average temperature with moderate
  reliability and that average temperature constrains the
  latitude with moderate reliability)
implications between statements
  d1(a1) = r1 \langle===> d2(a1)= r2 : \alpha, \beta, \gamma
  e.g., grain (place) = \{ rice,...\} \langle===> 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)

**FIGURE 1. Elements Of Expression In The Core Plausible Reasoning Theory.**

4. \( \ddot{\delta} \) Dominance of a subset in a set (e.g. chickens are not a
   large percentage of birds but are a large percentage of barn-
   yard fowl).

5. \( \sigma \) Degree of similarity of one set to another set.

6. \( \alpha \) 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.
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 [3] and [4] for an in-depth explanation or reference [6] for an implementation of the theory.

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

There have been a number of attempts to use models of uncertain reasoning for IR, with some similarities to plausible reasoning theory used here. For example, Croft et al have done experiments using a plausible inference approach to retrieval [5]. In their knowledge base, concepts are represented as frames. Relationships of synonym, is-a instance, part-of, cross-reference and their inverses have been used. They have used constrained spreading activation as their inference mechanism. Combining the term-based, nearest-neighbor and citation evidence has proved effective. (Other examples are GRANT [2], IOTA[11], RUBRIC[11]). What distinguishes the present work from earlier research is that it is based on a theory which is both empirically derived and very general, in the sense that it appears to subsume other models of IR reasoning.

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 this partial 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” citing an article by Bruce Croft can be represented as follows:

1. REF (Information Retrieval) = {doc#1}
2. REF (Non-classical Logic) = {doc#1}
3. REF (Non-classical Logic (Information Retrieval)) = {doc#1}
4. AUTH (doc#1) = {C. J. Van Rijsbergen}
5. CITE (doc#1) = {doc#2}

The first statement indicates the concept Information Retrieval is a reference for doc#1. The second statement states that the concept Non-classical Logic is a reference for doc#1. The third statement expresses that the term Non-classical Logic (Information retrieval) is a reference for doc#1. The fourth statement points out C. J. Van Rijsbergen as the author of doc#1. The last statement represents the citation of doc#2 (Croft’s article) by doc#1.
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:

\[ \text{REF (A-Subject)} = \{ \text{?} \} \]

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:

\[ \text{REF (Phrase)} = \{ \text{?} \} \]

For example:

\[ \text{REF (Tune Sharing System)} = \{ \text{?} \} \]

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:

\[ \text{REF( heuristic algorithm (graph coloring))} = \{ \text{?} \} \]

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 retrieved 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 2 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 references 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. 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 3. 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.

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 Acceptability can be computed globally to indicate the acceptance of a term or statement by the user population. On the other hand it can be computed individually for each user, then it will represent the point of view of that particular user over time or in a single session. These two being the extremes, there are many solutions in between. For example, acceptability could be recorded globally for the user population and for an individual user only if the user’s point of view differs from the population in a significant way. Another related question is, when to trust the

FIGURE 2. Finding References By Completing Incomplete Query Statement, Direct Approach.

user and when not to. Responses of a knowledgeable user can improve the quality of the knowledge base while those of a novice user may confuse the system. Similarly, “acceptable” could mean different things to different groups of user population. For example, in most computer literature, machine is a computer, however in the field of robotics a machine could be anything from a simple drill to a sophisticated robot. The issues in the computation of Acceptability such as whether to store it individually or globally, collect it only from expert or everyone, compute it for the whole collection or for the small subdomains are still open questions which need further investigation. In the rest of this paper, for the sake of simplicity, it is assumed that Acceptability represents the user population’s point of view. The methods of computation of certainty parameters are not discussed here since they are more experimental than theoretical.

The value of certainty (\(y\)) 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 1.1 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) [8]. 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 to find other relevant concepts and their associated references. As an example, let’s consider the query:

\[
\text{REF(algorithm (distributed)) = \{ ?\}}
\]

which expresses that there is an interest in references for distributed algorithms. Assuming 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 4 illustrates this transformation.

The example in Figure 4 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.
5. Experiment

Several experiments have been conducted using the CACM1 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, Recall, Normalized Precision and Normalized Recall [9]. Statistical significance is measured for normalized precision and recall values and precision values at each of 10 standard recall points, by using Wilcoxon Matched-Pairs Signed-Ranks test [10]. This test takes account of both the magnitude and the direction of the differences.

After processing the CACM collection a knowledge base was built which contains the documents, phrases and logical terms, X and BN relationships. A phrase is defined to be a combination of nouns and adjectives which are located next to each other. A term consists of two phrases related by a clue in the text, e.g., the relationship between “User Interface” and “Windows Operating System” signified by preposition “of” in sentence fragment “User Interface of Windows Operating System”. An X relationship is the relationship between parts of a phrase and the phrase itself, e.g., the relationship between “32-bit” and “32-bit Operating System”. Broader-Narrower relationship is a relationship between two phrases where one ends with the other. For example, the noun phrase Relational Database has a narrower interpretation than the phrase Database, similarly the phrase language has a broader meaning than the phrase data sublanguage which has in turn a broader meaning than the phrase data sublanguage.
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 called Dominance implemented as:

\[ D = \frac{F_{\text{child}} \cdot C_{\text{child}}}{\sum_{\text{child}} F_{\text{child}} \cdot C_{\text{child}}} \]

where \( F_{\text{child}} \) is frequency of a child node among other children of the parent node, \( C_{\text{child}} \) is the confidence on a particular child and the sum is over all children of the parent node.

The other is tf.idf weight [37] which is an abbreviation for “term frequency times inverse document frequency”.

1. The CACM collection consists of 3204 surrogates from Communication of ACM and 64 queries.

The vector space model implemented in these experiments also uses this measure. It is computed as:

\[ w_{ij} = \frac{N}{df_j} \log \left( \frac{N}{df_j} \right) \]

where \( w_{ij} \) is the weight of phrase \( j \) in document \( i \), \( f_{ij} \) is the frequency of phrase \( i \) in document \( j \), \( df_i \) stands for the number of documents in the collection in which phrase \( j \) has occurred, and \( N \) is the total number of documents in the collection.

Two plausible retrieval systems were built, one based on dominance the other based on tf.idf. 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: [8]

- Finding referents and completing incomplete query statements, direct approach by using only terms.
- Finding referents and completing incomplete query statements, indirect approach.
- SPEC-based Referent Transform.
- SPEC-based Argument Transform.
- 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:
  \[
  \text{IF} \quad a \quad \text{SPEC} \quad A \quad \text{AND} \\
  \quad b \quad \text{SPEC} \quad B \quad \text{THEN} \\
  \quad a(b) \quad \text{<--------> A(B)}
  \]
Two retrieval systems have been implemented based on the vector space model, one of them represents documents and queries as vectors of words in a multi-dimensional space. The other one uses vectors of phrases. These phrases have been obtained from the plausible reasoning model. The words have been produced by breaking up the phrases. Therefore, both models are representing the documents by the same set of phrases and words.

Since the vector space model assumes no relationship between index terms, in contrast to the plausible reasoning model which relies on a semantic network of term-term and Term document relationships, it seemed that the comparison of these systems may not reveal as much information as one would hope in such an experiment, therefore it was suggested that the knowledge of relationships should be made available to the vector model too. It was hoped that by doing this the effects of the knowledge of the relationships would be neutralized and the difference in their reasoning would be revealed.

The following procedure explains how every query given to the vector model is extended in order to neutralize the effects of the semantic network of the plausible retrieval:

1. Phrases of each query were looked up in the knowledge base. For each query phrase \( q \) we identify all phrases \( p \) such that:
   1.1. There is a chain of \( X \) arcs from \( q \) to \( p \); or
   1.2. There is a chain of BN arcs from \( q \) to \( p \); or
   1.3. There is a chain as in 1.1, 1.2 from \( q \) to a phrase \( p' \) and also \( p \) and \( p' \) form a logical term.

2. These phrases were added to the appropriate queries in the vector space model.

3. These extended queries were submitted to both implementations of the vector space model.

5.1 Results

Several experiments are conducted to answer the research questions: ‘is plausible reasoning more effective than vector retrieval?” and sub questions such as “Does the dominance parameter perform as well as \( \text{tf. Idf weights} \)”. 

![Figure 5. A comparison of plausible reasoning systems with vector space model.](image)
Figure 5 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 precision are computed for each query at standard recall points by interpolation and then averaged over the queries for each standard point.

Table 1 and Table 2 list for average precision values for each standard recall level and the Z and p values of the significant test for the plausible retrieval systems in comparison to vector space implementations. A star next to a p value indicates a significant difference.

As it is presented in the last rows of Table 3 and Table 4, both implementations of plausible reasoning extended with vector retrieval’s output performed significantly better than vector model in terms of normalized precision and recall.

To investigate whether the difference between two implementations of plausible reasoning is statistically significant, normalized precision and recall values are calculated for both systems without appending the output of the vector space model to their output. Table 5 includes normalized precision end recall values for both implementations and the difference between them.
TABLE 2. Average Precision, Z And P Values At Standard Recall Points For Vector Space Model (using Single Words) And Both Implementations Of Plausible Reasoning.

<table>
<thead>
<tr>
<th>Recall Level</th>
<th>Vector Model (Single Words)</th>
<th>Extended Plausible Reasoning, with dominance weight</th>
<th>Extended Plausible Reasoning, with tf.idf weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Precision</td>
<td>Average Precision</td>
<td>Average Precision</td>
</tr>
<tr>
<td>0.1</td>
<td>0.51979</td>
<td>0.63482</td>
<td>0.60777</td>
</tr>
<tr>
<td>0.2</td>
<td>0.40858</td>
<td>0.50734</td>
<td>0.44642</td>
</tr>
<tr>
<td>0.3</td>
<td>0.30662</td>
<td>0.37492</td>
<td>0.35967</td>
</tr>
<tr>
<td>0.4</td>
<td>0.24519</td>
<td>0.26745</td>
<td>0.26080</td>
</tr>
<tr>
<td>0.5</td>
<td>0.21374</td>
<td>0.22256</td>
<td>0.22214</td>
</tr>
<tr>
<td>0.6</td>
<td>0.14800</td>
<td>0.15449</td>
<td>0.15588</td>
</tr>
<tr>
<td>0.7</td>
<td>0.07411</td>
<td>0.07472</td>
<td>0.07472</td>
</tr>
<tr>
<td>0.8</td>
<td>0.04512</td>
<td>0.04903</td>
<td>0.04903</td>
</tr>
<tr>
<td>0.9</td>
<td>0.01301</td>
<td>0.01508</td>
<td>0.01508</td>
</tr>
<tr>
<td>1.0</td>
<td>0.00161</td>
<td>0.00201</td>
<td>0.00201</td>
</tr>
</tbody>
</table>

TABLE 3. Results of Significant Tests 1-4. The Vector Space Model (Using Phrases) Vs. Both Implementations Of Plausible Reasoning.

<table>
<thead>
<tr>
<th>Test</th>
<th>Extended Plausible Reasoning with dominance</th>
<th>Extended Plausible Reasoning with tf.idf</th>
</tr>
</thead>
<tbody>
<tr>
<td>test1(N.Precision)</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>test2(N.recall)</td>
<td>30</td>
<td>28</td>
</tr>
</tbody>
</table>

No. of queries favor plausible reasoning: 30
No. of queries favor vector system: 29
No. of queries tied: 30
Z value: 4.22
P: < 0.00003

No. of queries favor plausible reasoning: 9
No. of queries favor vector system: 10
No. of queries tied: 9
Z value: 3.86
P: > 0.000005

No. of queries favor plausible reasoning: 9
No. of queries favor vector system: 9
No. of queries tied: 9
Z value: 4.04
P: < 0.00003

No. of queries favor plausible reasoning: 9
No. of queries favor vector system: 11
No. of queries tied: 9
Z value: 3.78
P: > 0.00007

No. of queries favor plausible reasoning: 9
No. of queries favor vector system: 11
No. of queries tied: 9
Z value: 3.78
P: < 0.00011
Since no references are retrieved for 8 queries, these two tests are conducted on only 40 queries. The results of the tests presented in Table 5 show that the system with the dominance weights has outperformed the system with tf. idf weights.
6. Conclusion
This research has investigated the application of Collins and Michalski’s empirically-based theory of plausible reasoning. It has been acknowledged that retrieval is logic but relevance is not. It has been shown that retrieval can be done through plausible inferences similar to those used by people. A document representation method has been proposed which uses logical terms, statements and phrases to represent documents and it has been demonstrated that this method provides a better representation for documents than using single words.

In these experiments simple clue-based methods were used in extracting relationships that were needed for knowledge base and reasoning. Small samples taken from the relationships revealed that only 90% of certain relationships are accurate. Also many relationships were not recognized in this implementation. Nevertheless, plausible reasoning was able to cope with inaccuracies and outperform the vector space model. Better NLP techniques that more reliably and completely summarize document content with more logical terms and statements may contribute to better performance of plausible retrieval systems through enriching the documents’ representation.

As is suggested in tables 1-4 both plausible reasoning implementations extended by the vector space models are significantly better than their counterpart vector space models alone. Table 5 demonstrates that dominance weights have performed better than tf. idf weights in these experiments. So the dominance weight, as used here, can compete with the well-known tf. idf weight which has been shown [9] to have good performance. Plausible reasoning ranks documents differently than the vector space model, often documents with lower ranks in vector model have been ranked higher by plausible reasoning. Although the CACM collection used in the experiments includes the relevance judgments, it lacks the ranking of them (i.e. knowledge of whether one document is more relevant than another). So it is not possible to qualitatively analyze different methods in regard to the ranking they provide and the ranking the users want to see. The characteristics of plausible reasoning make it a viable choice for many applications in IR. For example, there are companies which furnish 10 to 20 news briefs daily to their users, these systems usually build a profile for each user and retrieve a short list of documents relevant to this profile every day. Plausible reasoning has the potential to learn a user’s interests, build a profile and retrieve references based on relevance to this profile. Its high precision and small outputs makes it suitable for this kind of environment.


