# FLEXICON: AN EVALUATION OF A STATISTICAL RANKING MODEL ADAPTED TO INTELLIGENT LEGAL TEXT MANAGEMENT

DAPHNE GELBART J. C. SMITH

UNIVERSITY OF BRITISH COLUMBIA
FACULTY OF LAW ARTIFICIAL INTELLIGENCE RESEARCH PROJECT
VANCOUVER, CANADA V6T 1Y1

## ABSTRACT

The FLEXICON system was designed to provide legal professionals with an effective and easy-to-use legal text management tool. This paper discusses the structured knowledge representation model designed for the FLEXICON system serving both as an internal knowledge representation scheme, in conjunction with statistical ranking, and as an external representation used to summarize legal text for rapid evaluation of the search results. The model is evaluated and compared to alternative information retrieval models. Experimental test data is presented to demonstrate the model's retrieval effectiveness in comparison to boolean search.

## 1. INTRODUCTION

The FLEXICON system [Gelbart and Smith, 1990; Gelbart and Smith, 1991a, Gelbart and Smith 1991b, Gelbart and Smith, 1992] was designed to provide legal professionals with a tool which supports all aspects of legal text management for large databases. Our design goal was to present users with a system that offers a simple user-interface, maximal assistance in query formulation, effective retrieval and the means to evaluate the relevance of retrieved documents. The system must operate efficiently on large databases with minimal maintenance and lend itself to implementation for personal computers and CD-ROM optical disks. In addition, the system should address the special needs of legal professionals and operate on any legal text.



The ideal system might receive queries in natural language specifying the issues of interest to the user and intelligently identify documents addressing the issues specified by the query. Furthermore, in examining multi-issue documents, the system could selectively identify portions of the text that exactly match the user's intent. Such a system is, at present, not feasible, except when limited to very narrow domains and when supported by extensive investment of time and cost in system development and maintenance.

Legal cases, as produced by the courts, generally lack defined structural elements such as titles, abstracts and subsections. Computers do not possess the true intelligence required to express the meaning of large volumes of unstructured text in a form which can be maneuvered automatically to produce accurate relevancy determination. What they can do, however, is to automatically analyze and process the text in order to unfold its structure and locate units of text that convey its meaning, thus creating a structured representation of the text. This representation can be directly maneuvered by a text management system to locate best matching documents that satisfy a user's request.

This approach permits legal text to be automatically classified to identify the subdomains of law or issues to which it pertains, standard header information such as the style of cause, trial date and jurisdiction of legal cases, can be automatically extracted. It can be subdivided into paragraphs containing significant statements and those paragraphs can be further subdivided into classes such as factual information, legal discussion and special issues and finally, content-conveying clauses such as concept phrases, factual phrases and citations can be identified and used to represent the case and their interrelation can be recorded in a conceptual network . Once the structure is unfolded, its units can be used as effective index terms and as building blocks to define meaningful queries and the associations between terms can be used to refine the queries and enable the identification of relevant text beyond a simple keyword retrieval.

This approach was selected in designing the FLEXICON system. FLEXICON was described in detail in [Gelbart and Smith, 1991a]. It is based on a structured adaptation of the Vector Space model [Salton and McGill, 1983] to the legal domain using intelligence to automatically determine the structure underlying legal text and allowing users to form structured queries which convey the meaning of the user's needs. FLEXICON is being developed in the C and C++ programming languages under MicroSoft Windows and was designed to run on IBM personal computers.

The remaining of this paper compares FLEXICON to existing knowledge representation methods. Experimental data is then presented to demonstrate the retrieval effectiveness of FLEXICON and compare it to other models.

## 2. KNOWLEDGE REPRESENTATION MODELS FOR TEXT MANAGEMENT

Three main models have been developed for large-scale information retrieval systems [Turtle and Croft, 1992].

### 2.1 THE BOOLEAN MODEL

The boolean model is the standard for current large-scale operational information retrieval systems. Boolean retrieval is based on "exact match" between the search query and the documents in the database, which is expressed as words or phrases combined by boolean operators. The model requires no special structuring of text, allowing users to search unprocessed document titles, abstracts or their full text. Most boolean systems, however, apply some text processing procedures designed to enhance the retrieval efficiency, including noise word stripping, an inverted index and word location information. In contrast, boolean queries have a distinct structure which allows the specification of a search in terms of keywords occurring in the text and the boolean operators AND, OR and NOT. Most Boolean systems also allow users to include adjacency operators which are useful in specifying multiple-word search terms.

The main problem associated with the boolean model is the rigidity of its operators, resulting in the partition of the database into relevant documents and non relevant documents with no distinction of their level of relevancy to the query specification, corresponding to low recall or low precision of retrieval [Blair and Maron 1985; Herman 1989; Turtle and Croft, 1991] This problem is further complicated by the user interfaces associated with boolean queries, which require a level of understanding of boolean logic for the formulation of effectual queries which legal and paralegal users often do not posses.

In an attempt to overcome the main problem of boolean search: the rigidity of its operators, Salton, Fox and Wu [1983] have presented the Extended Boolean model. In this model each boolean operator has a degree of fuzziness or weight indicating how strictly the interpretation of the operator should be applied. The formulation of extended boolean queries, however, is even more difficult for users than that of traditional boolean systems.

In an attempt to overcome users' difficulty with the formulation of effective boolean queries, various works were reported on the translation of queries written in natural language to boolean queries. Das-Gupta [1987] developed an algorithm based on syntactic and semantic information to locate the presence of implied AND and OR operators in natural language text. Smith [1990] has presented a complex algorithm for a full translation of natural language queries to boolean queries. Once translated to boolean query, however, the "hit or miss" strategy prevails.

### 2.2 THE VECTOR SPACE MODEL

The Vector Space model developed by Salton [Salton and McGill, 1983], solves the main problems presented by boolean search. It is based on the "best match" concept using statistical information to select a ranked list relevant documents. The model represents both documents and queries as vectors in a multidimensional space, the dimensions of which are words used to represent the texts. Both document and query terms can be associated with weights that are computed on the basis of the statistical distribution of the terms in the database, to account for their importance. The relevancy of retrieved documents to the user's request is determined by comparing the document and query vectors using a correlation similarity measure such as the Cosine formula.

The model is highly effective and compares well with more complex representations based on extensive natural language processing [Dillon and Gray, 1982; Fagan 1987, Smeaton, 1986] and knowledge based representations [Gey and Chan,1989]. It satisfies many of the design requirements for the FLEXICON system by lending itself to automatic indexing, providing effective retrieval and a simple query-formulation user interface and by producing a ranked list of relevant documents. Moreover, the model can be efficiently implementated for personal computers.

Nevertheless, the model, as represented by the SMART system [Salton and McGill, 1983] does have some drawbacks for legal text management, primarily the representation of documents and queries as single-word terms which results in a significant loss of meaning from the contents of legal documents and queries. Like many other disciplines, the legal domain is associated with a professional jargon. Legal phrases, which represent a statement of the applicable law, cannot be adequately represented by their components which are often high-frequency low-content words. Case and statute citations, which can serve as excellent search terms, having neither synonyms nor homographs and serving as short coded expressions standing for previously decided issues [Tapper [1984], would be eliminated as noise words using the pure Vector Space model.

### 2.3 PROBABILISTIC MODELS

Probabilistic information retrieval models are based on the Probability Ranking principle [Robertson, 1977] which ranks legal documents according to their probability of relevance to the query given every available source of information. The model estimates the probability of relevance of a text to the query, on the basis of the statistical distribution of terms in relevant and non-relevant text, given an uncertainty associated with the representation of both the source text and the information need, as well as the relevance relationship between them.

Turtle and Croft have recently introduced the Inference Net model [Turtle, 1990; Turtle and Croft, 1991] which is based on Bayesian inference networks [Pearl, 1988], representing documents, queries and the user's information need as directed acyclic dependency graphs. They report improved retrieval effectiveness compared to traditional boolean search, the Vector Space model as represented by the SMART system and previous probabilistic models, primarily due to the use of multiple information sources in determining the relevancy of a document to the user's information needs. In future work, we will compare the retrieval effectiveness of the Inference Net model to that of FLEXICON whose knowledge representation is based on a simpler and easier to implement model.

## 3. THE FLEXICON KNOWLEDGE REPRESENTATION MODEL

As mentioned above, the knowledge representation scheme selected for the FLEXICON system is based on an adaptation of the Vector Space model to the legal domain, using intelligence to automatically determine the structure underlying legal text. Structural representations were also presented in other works [Bing, 1987], [Hafner, 1982]. However FLEXICON constructs structured representations automatically from text. Cases and other legal documents can be represented by document profiles that preserve the meaning of legal text and contain all the information necessary and sufficient to match documents with a user's query.

### 3.1 THE QUADRANT STRUCTURE

Four parameters are used to represent document profiles: Legal concepts, factual terms, case citations and statute citations. The fact terms tell the factual story on which a case is based. Legal concepts constitute a statement of the applicable law and the resolution to the issues when the law has been applied to the facts. Case citations stand for related issues previously decided on point or by analogy. Statute citations are references to applicable legislation. The four types of profile terms are weighted and sorted by factors reflecting their distribution in the processed document and the data collection. Section 4 below describes the text processing procedures employed to automatically extract the four types of terms from text.

### 3.2 THE GLOBAL INDEX

A global index consisting of all the terms that occur in the data collection is generated, for each quadrant, as the union of individual document representations. The added structure produced by including multi-word terms in the global index improves the retrieval effectiveness significantly by mapping high-frequency low-content terms which serve as poor discriminators into useful indexing terms.

### 3.3 QUERY STRUCTURE

User defined queries use the same representation as legal documents and are formed by selecting index terms from the quadrant dictionaries. Queries are compared to document profiles and best-matching documents are returned to the user, ranked by decreasing order of similarity to the query. In addition to queries entered directly by users, natural language queries can be automatically processed by FLEXICON and converted to the four lists of weighted terms, using the same technique employed to analyze and process documents. While we expect natural language queries to be inferior to those entered directly by users, they can be used by amateur users or as a query formulation tool. An additional tool is "retrieval by example" i.e. retrieving cases similar to sample database cases entered by the user. FLEXICON will automatically formulate a query consisting of the supersets of terms occurring in the document profiles of those cases. Queries generated by those tools can be edited and customized to the user's needs.

### 3.4 A Conceptual Network

An important feature of the FLEXICON document and query representation is the use of a conceptual network of related terms and documents.

#### 3.4.1 The Related Terms Thesaurus

FLEXICON automatically creates associations between quadrant terms by linking terms that tend to statistically co-occur in many documents. These associations, referred to as the Related Terms Thesaurus form a network structure associating each important legal concept, case, and statutory provision with related concepts, cases and statutes. Thus the FLEXICON search queries are related through their profiles to a matrix of interconnected profiles and can be refined by selectively adding terms from the thesaurus.

#### 3.4.2 Citation Cross-Reference

A single case name cited in a legal document represents the information content of that entire case. The section of a statute cited in a legal document is a link to the information contained in the entire statute. The FLEXICON user can follow Hypertext links from citations to the information contents which they represent. Furthermore, the FLEXICON dictionary of citations provides citation cross-reference information by providing links between a given cases and cited or citing cases, which can be followed by the user.

#### 3.4.3 Relevance Feedback

Additional interrelation between query terms and document queries can be accomplished by the use of Relevance feedback. Terms found in profiles of documents retrieved by an initial query can be used to refine that query and allow the search to be repeated until the user is satisfied.

#### 3.4.4 Case Classification

In addition to correlating related terms, FLEXICON also recognizes associations between legal cases. A legal system is made up of many doctrinal sub-systems such as private law and public law, which in turn are made up of further sub-systems such as property, contract, tort, constitutional law, administrative law, or criminal law. FLEXICON automatically classifies legal cases into domains of law by associating key legal concepts, leading cases, statutes and significant fact terms with relevant domains and measuring the extent that the four type of terms appearing in the text of the case match the terms in the domain lists. Case classification will be used to filter out non-relevant material and to localize and limit a search to specified domains.

The presentation of legal information as a matrix of interrelated conceptual doctrinal systems and sub-systems, full text of cases and complete statutes results in a highly intelligent and effective information management system.

### 3.5 Documentation

In addition to improving legal text indexing and query formulation, the FLEXICON knowledge representation also provides an important documentation tool, by serving as the basis for the automatic generation of electronic case "headnotes" which we call *flexnotes*. The flexnotes are designed as a means to rapidly decide the relevance of retrieved cases. A flexnote consists of case header information, a classification of the subject of law, the listing, in four quadrants, of the most significant concepts, facts, case citations and statute citations in decreasing order of a system-computed term's weight and key paragraphs automatically extracted by the system. The flexnote is followed by the text of the case with numbered paragraphs which the user can browse, following Hypertext links. A sample flexnote is presented in [Gelbart and Smith 1991a].

Important paragraph extraction in FLEXICON was developed in analogy to case headnotes produced by legal publishers which are often based on selected extracts representing important issues discussed in the case. FLEXICON examines, for each paragraph, significant legal phrases, fact terms and citation patterns, as defined by document profiles. The program also examines paragraph position, continuity and length as well as special phrases used by judges in significant paragraphs.

## 4. Legal Text Processing in FLEXICON

The structured FLEXICON case profiles are created using three major text processing functions. The recognition of case and statute citations in legal text is accomplished by template-matching functions and simple rules. Concept terms are extracted from the text by matching sections of the text with terms contained in a phrase dictionary and fact terms are the remaining words and phrases following noise word removal.

### 4.1 RECOGNIZING LEGAL CONCEPTS

Legal concepts are automatically extracted from text by matching sections of the text with terms contained in the Legal Phrase Dictionary, a domain lexicon of phrases used by legal professionals. Each term in the dictionary consists of a stem, which is matched to the text, and a concept phrase, to which the stem is linked. Most concept phrases in the dictionary are associated with more than one stem, thus allowing users to retrieve documents containing terms that are synonymous or semantically similar to concept phrases selected as search terms by the user. The dictionary also distinguishes between entries which require an exact match versus entries which allow the matched information to appear in the text in any order, to be suffixed and to be separated by noise words. The Concept dictionary is currently manually constructed by the FLAIR team. The FLEXICON technology can be used to assist in the dictionary's construction and maintenance by processing electronic versions of leading cases, scholarly articles and case comments to extract legal concepts, facts and citations. Because fact terms are defined as neither legal phrases, nor citations or noise words, we can expect that any missed concept phrases will show up as fact terms. Those will then added to the dictionary.

### 4.2 AUTOMATIC RECOGNITION OF FACT PHRASES

A typical case contains a factual story, a description of the set of legal issues which the story gives rise to, a statement of the applicable law, and a resolution of the issues when the law has been applied to the facts. After the removal of legal phrases, recognized by the Legal Phrase Dictionary, and of case and statute citations that are recognized by template matching functions, the remaining text represents the facts of the case. Single-word fact terms can easily be generated by removing noise words as determined by a noise word list. However, an improved indexing and query representation can be achieved by incorporating multi-word fact terms in both document and query representations. Unlike the recognition of legal concept phrases, which is dictionary-based, the recognition of fact phrases must be based on automatic sentence construct analysis.

[Dillon and Gray [1983] and Fagan [1987] have demonstrated the feasibility of automated phrase recognition in text and the superiority of statistically-based phrase recognition over the syntactically-based approach but did not demonstrate a significant affect on the retrieval effectiveness. Smeaton [1991] as well as Croft, Turtle and Lewis [1991], extended the phrase-based representations to both query and documents and, using structured queries as search terms, obtained substantial retrieval effectiveness gain.

Our fact phrase recognition method combines term distribution and proximity information with a lexicon of noise words. We define categories of noise words and use them as a "glue" connecting fact terms into phrases. "Joiners" (e.g.: "by", "of") join two fact terms, "modifiers" (e.g. "extended", "civil") qualify or constrain terms, and "pure noise" is eliminated. The program also uses a set of rules to identify classes of noise terms such as names, numbers, etc., which are retained in the context of phrases and citations. For example, "ways and means committee" is an effective term while "ways" and "means" are noise words and "committee" is a relatively low-content term. Likewise, noise words are retained as part of citations. While this simple approach to phrase recognition can also result in meaningless or useless terms or over specific terms, their occurrence can be minimized by the use of Corpus Filtering, i.e. the elimination of terms whose collection frequency falls below a given threshold and by applying constraints on the length of automatically determined terms.

Phrase terms are currently weighted by simply counting their occurrence in a document and their components are not included as separate terms unless they occur independently in the document. In the future we will also retain phrase components whose collection frequency is below a predefined threshold thus allowing users to search also for single word terms that serve as good discriminators.

### 4.3 AUTOMATED EXTRACTION OF SPECIFIC INFORMATION

In addition to the dictionary-based extraction of legal phrases, the automatic identification of fact phrases and the template-based recognition of citations, FLEXICON also automatically recognizes, and extracts from legal text, header information such as the styles of cause of legal cases, dates, jurisdictions and the judges that heard the cases. As well, we have done preliminary work on automatic interpretation, and extraction from text, of specific useful information such as the case outcome, the amount of damages awarded and noting up information [Deedman, Gelbart and Coleman, 1991].

## 5. TESTING THE FLEXICON RETRIEVAL EFFECTIVENESS

We have transformed FLEXICON from a prototype operating on a database of 50 cases to a system designed to operate on large databases. While the effectiveness of retrieval of the Vector Space model was documented by Salton [Salton and McGill, 1983] and the feasibility of using the model to search very large databases, including a legal database of 40,000 cases, was demonstrated by Herman and Candela [1989], we have conducted controlled experiments to test the effectiveness of the FLEXICON search and compare it to boolean search. Future testing will be conducted to compare FLEXICON to additional search methodologies such as the SMART system and probabilistic retrieval models.

Retrieval effectiveness is usually determined via precision/recall measurements. Recall indicates the robustness of the search, measuring the fraction of relevant documents retrieved out of all the relevant documents in the data collection. Precision indicates the accuracy of the search, measuring the fraction of relevant documents retrieved out of the entire set of retrieved documents. Plotting recall versus precision, we can examine the search precision at various recall levels and follow the overall pattern of the FLEXICON search. By recording recall/precision measurements at discrete output levels (e.g., for the first 20, 40, 60, 80 and 100 documents retrieved), we can compare the search results to additional search methodologies at reasonable output levels.

### 5.1 THE EXPERIMENT

A moderately sized data collection consisting of 1000 cases in the domain of pure economic loss plus general British Columbia cases, was selected as a test collection. The cases were automatically analyzed and flexed. A team of lawyers and law students have developed eight queries in natural language and ranked the relevancy of cases in the test collection for each query using guidelines designed to reduce the subjectivity of this determination (e.g.: most relevant cases being on-point). The students were then asked to produce the best FLEXICON and boolean queries corresponding to the natural language queries. Those preparing boolean queries were instructed to try various combinations and select the best query. The students formulating FLEXICON queries were encouraged to use relevance feedback to refine their queries. We ran the queries using the FLEXICON system and an LQ boolean search (by Liam Quin) which offers a standard boolean search, allows the use of multi-word search terms (as a substitute to proximity operators) and offers a simple document ranking on the basis of the occurrence of query terms in retrieved documents. Our test procedures automatically produce recall/precision tables and graphs by comparing the search results, at incremental output levels, to the manually-produced relevancy ratings.

### 5.2 RESULTS

Figure 1 demonstrates a sample natural language story (1.a) which serves as the basis for the best FLEXICON (1.b) and boolean (1.c) queries developed by the testing team. No effort was spared in producing the boolean queries. The FLEXICON query shows the four quadrants, including a case citation. While case citations can be included in boolean queries, they are seldom used because of the need for exact match. In FLEXICON, the user may enter a familiar citation, by consulting the term dictionary, the related-terms thesaurus (when completed) or by applying relevance feedback. See [Gelbart and Smith 1991a] for a complete description of the FLEXICON search scenario.

Figure 2 demonstrates average recall/precision results for eight queries comparing the FLEXICON search (table 2.a) to boolean search (table 2.b). The results are plotted in graph 2.c. Relevancy ratings were determined in two ways. In the first, only very relevant cases ("on point") were considered. In the second, both very relevant and fairly relevant cases were considered ("on point" or "by analogy"). As figure 2.c shows, the FLEXICON search generally manifests significantly higher precision at various levels of recall, using both methods. Furthermore, the FLEXICON queries were easy to formulate while the compilation of the corresponding boolean search queries required substantial effort and repetition. As well, the evaluation of the relevancy of retrieved cases was largely facilitated by the rapid examination of flexnotes and the use of Hypertext links to browse full text.

### 5.3 DISCUSSION

While an attempt to improve recall generally reduces precision and vice versa, the FLEXICON search provides good recall without compromising much on precision. Furthermore, because the FLEXICON search provides a list of documents, ranked according to their relevancy, the user can view relevant documents up to the desired level of recall and precision. For example, in viewing only the first set of 20 documents, the user is expected to achieve good search precision and acceptable recall. Viewing 60 documents, however, will provide better recall but reduce the precision. A user may select the first scenario when searching for the most relevant documents while using the second scenario when he or she must retrieve every relevant case. When using the second scenario, the user can quickly scan through a large number of cases by scanning the flexnotes and using the HYPERTEXT feature to jump to specific portions of the text.

## 6. CONCLUSION

The FLEXICON knowledge representation model is well suited for the representation of legal text in conjunction with information management. The model's superiority over boolean search is demonstrated in terms of the knowledge structuring, the user interface the retrieval effectiveness and the ranking of relevant documents. The model, as adapted for the legal domain, improves upon the original Vector Space model, as demonstrated by the SMART system, in incorporating intelligent structuring of both documents and queries. Probabilistic models, especially the Inference Network model which forms the basis for the WIN system recently introduced by the West Publishing Company, also appear to provide good representations of text. Both models intelligently incorporate structure in the document and query representations, offer elegant and easy-to-use interfaces, allow the incorporation of multiple information sources such as thesauri and produce a ranked list of relevant cases. However, FLEXICON may have the advantage of a simpler and more intuitive model. Furthermore, the FLEXICON user interface, while allowing users to enter queries in natural language, offers a higher degree of interaction with the user by providing additional query formulation and refinement tools via a related-terms thesaurus and relevance feedback. Future work will include comparison of the FLEXICON retrieval effectiveness with that of the Inference Network model used by the WIN system, using a procedure similar to that described in section 5 above.

ACKNOWLEDGMENTS

This research was supported by the Law Foundation of British Columbia and the Social Sciences and Humanities Research Council of Canada.

## REFERENCES

Bing, Jon, "Designing Text Retrieval Systems for Conceptual Searching", in Proc. 1st International Conference on Artificial Intelligence and Law, Boston, 1987.

Blair, D. C. and M.E. Maron, "An evaluation of retrieval effectiveness for a full-text document retrieval system", Communications of the ACM, Vol 28 Number 3, March 1985.

Croft, W.B., Turtle H.R., and Lewis D.D., "The use of Phrases and Structured Queries in Information Retrieval", In Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval, (1991).

Das-Gupta, P., "Boolean Interpretation of Conjunctions for Document Retrieval", Journal of the American Society for Information Science, Vol 38, 1987.

Deedman C., Gelbart D. and Coleman M., "SLATE: Specialized Legal Automated Term Extraction", Technical Report, University of British Columbia Faculty of Law.

Dillon M. and Gray S., "FASIT: A Fully Automatic Syntactically Based Indexing System", Journal of the American Society of Information Science 34, (1983).

Hafner, Carole, "Conceptual Organization of Case Law Knowledge Bases", in Proc. 1st International Conference on Artificial Intelligence and Law, Boston, 1987.

Fagan J., "Experiments in Automatic Phrase Indexing for Document Retrieval: A Comparison of Syntactic and Non-Syntactic Methods. Ph.D. Thesis, Technical Report 87-868, Cornell University, Computer Science Department, 1987.

Gelbart, Daphne and J.C. Smith, "Towards a Comprehensive Legal Information Retrieval System", in Proceedings of International Conference on Database and Expert Systems Applications, Vienna, 1990.

Gelbart, Daphne. and J.C. Smith "Beyond Boolean Search: FLEXICON, a Legal Text-Based Intelligent System", Proceedings of the Third International Conference on Artificial Intelligence and Law, Oxford, ACM Press, 1991.

Gelbart Daphne and J.C. Smith, "Current Issues in Text Retrieval: FLEXICON, a Legal Text-Based Intelligent System", in Proc AAAI-91, Natural Language Text Retrieval Workshop, Anaheim, July 1991.

Gey, Fredric and Chan Wingkei, "Comparing Vector Space Retrieval with the RUBRIC Expert System", SIGIR FORUM, Volume 23, No 1, 1989.

Herman, Donna and Gerald Candela, "A Very Fast Prototype Retrieval System Using Statistical Ranking", SIGIR FORUM, Vol 23, No 4, 1989.

Pearl, J., "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.", Morgan Kaufman, 1988.

Riesbeck C. and Schank R., "Comprehension by Computer: Expectation-Based Analysis of Sentences in Context". Computer Science Research Report 78, Yale University, New Haven.

Robertson, S.E., "The Probability Ranking Principle in IR", Journal of Documentation. 33:294-304. 1977.

Salton, G. and M.J. McGill. Introduction to Modern Information Retrieval, McGraw-Hill, 1983.

Smeaton A., "Information Retrieval Research: How it Might Affect the Practicing Lawyer", in Law, Decision-making and MicroComputers", (Nagel S., eds.), 1991.

Smith, M.E., "Aspects of the P-Norm Model of Information Retrieval: Syntactic Query Generation, Efficiency and Theoretical Properties. Ph.D. Thesis, Computer Science Department, Cornell University.

Tapper, Collin, "An experiment with citation vectors", Data Processing and The Law, 1984.

Turtle H.R. and Croft, W.B. "Evaluation of an Inference Network-Based Retrieval Model", ACM Trans. Inf. Syst. 3 (1991), 187-222.

Turtle H.R. and Croft W.B. "A comparison of Text Retrieval Models", The Computer Journal, Vol 35, No 3, 1992.

## Figure 1 FLEXICON and Boolean Queries

(1.a) Natural Language Query

Ms. Johnson, heir to the Johnson & Johnson Baby Powder fortune, asked Mr. Caldwell to draft her will. Mr. Caldwell is a solicitor with Blackbeard, Caldwell & Co. and had been Ms. Johnson's solicitor for the past ten years. Ms. Johnson told Mr. Caldwell that she wanted her entire estate to be divided equally between her two children, David and Chrissy. Ms. Johnson specifically told Mr. Caldwell that she didn't want her husband, Mr. Jones, to inherit anything from her estate. Ms. Johnson and her husband had chronic marital problems; Mr. Caldwell had been asked to draft a separation agreement for Ms. Johnson on four prior occasions. Ms. Johnson had received most of her fortune from her family, and she wanted to pass this on to her children. Mr. Caldwell assured Ms. Johnson that her estate would go to her children. Mr. Caldwell prepared the will and took it over to Ms. Johnson to sign. Ms. Johnson read and signed the will. Dave and Chrissy were present; they were home visiting their mother. Mr. Caldwell asked them both to witness the will. Ms. Johnson died one year later from a massive heart attack. On application the court of probate found that the will was not valid, and that Ms. Johnson died intestate. Under the Estate Administration Act the estate was apportioned between Mr. Jones, David and Chrissy. David and Chrissy are suing Mr. Caldwell in negligence. They allege that he was negligent in carrying out his duties as a solicitor, and that his negligence caused them to suffer economic loss. David and Chrissy would have received a much larger portion of their mother's estate if Mr. Caldwell had adequately performed his duties, and not allowed the beneficiaries to witness the will.

(1.b) Best FLEXICON Query

| LEGAL CONCEPTS | CASE CITATIONS |
|---|---|
| <H> attestation<br><H> beneficial interest<br><H> solicitor's negligence | <H> Ross v Caunters |
| **STATUTE CITATIONS** | **FACTS** |
| | <H> solicitor<br><H> client<br><H> intestator |

(1.c) Best Boolean Queries

[solicitor or lawyer] and [will or inheritance or estate or testamentary gift or bequest] and [beneficiary or third party beneficiaries or heir] and [defective or defective will or negligence or void or professional negligence or solicitor negligence] and [economic loss or suffer loss]

[solicitor or lawyer or solicitor-client relationship] and [will or defective will or bequest or estate or inheritance or testamentary gift]

Figure 2 Recall/Precision measurements

(2.a) FLEXICON Average Recall & Precision

| Very Relevant Average Recall | Very Relevant Average Precision | Relevant Average Recall | Relevant Average Precision |
|---|---|---|---|
| 10 | 90 | 10 | 97 |
| 20 | 87 | 20 | 91 |
| 30 | 78 | 30 | 76 |
| 40 | 71 | 40 | 71 |
| 50 | 70 | 50 | 63 |
| 60 | 67 | 60 | 53 |
| 70 | 64 | 70 | 44 |
| 80 | 45 | 80 | 34 |
| 90 | 38 | 90 | 28 |
| 100 | 5 | 100 | 2 |

(2.b) Boolean Average Recall & Prercision

| Very Relevant Average Recall | Very Relevant Average Precision | Relevant Average Recall | Relevant Average Precision |
|---|---|---|---|
| 10 | 92 | 10 | 83 |
| 20 | 53 | 20 | 56 |
| 30 | 51 | 30 | 42 |
| 40 | 40 | 40 | 28 |
| 50 | 25 | 50 | 29 |
| 60 | 20 | 60 | 16 |
| 70 | 21 | 70 | 9 |
| 80 | 7 | 80 | 0 |
| 90 | 3 | 90 | 0 |
| 100 | 0 | 100 | 0 |

(2.c) Average Recall and Precision Results