Intelligent Information Retrieval Using an Inductive Learning and a Neural Network Model

Kim, Seonghee*

Contents

1. Introduction

The major purpose of an information retrieval (IR) system is to respond to a request for information about a particular topic by retrieving a set of related documents. Currently, most operating IR systems are based on Boolean logic. The Boolean retrieval process produces a subset of document files based on a match or no-match selection between query terms and index terms of the document. In other words, IR systems based on a Boolean retrieval model assume that the query and documents are exactly represented in terms of a set of keywords or in natural languages (Belkin et al. 1982). Consequently, the system produces a set of related documents that exactly match the query while rejecting all other partial or non-matching documents. This typical IR system based on Boolean logic reveals inherent problems associated with retrieval effectiveness.

While there are no absolute remedies to all the problems associated with information retrieval systems, a few works in this field have concentrated on

* Lecturer, Dept. of Lib. & Info. Science, ChoongAng University.

improving system effectiveness by introducing new models of document retrieval and new search strategies based on these models (Salton and McGill 1983; Van Rijsbergen 1979). Specifically, the use of partial matching and ranking techniques for information retrieval has been explored successfully (Coyle 1985; Doszkocs et al. 1990; Larson 1992). These models include the probabilistic model (Bookstein and Swanson 1974), extended Boolean models (Salton, Fox, and Wu, 1983; Salton and McGill, 1983) and vector space representations (Salton, Wong, and Yang, 1975; Wong, Ziarko, and Wong 1985). Each of these models is based on keyword retrieval that operates at a symbolic, text-matching level, yet ignores any semantic and contextual information in the retrieval process (Watters 1989).

In addition to the models mentioned above, many researchers have applied neural networks[1] to the area of information retrieval since 1980. Current research conducted by Mozer (1984), Belew(1986), and Wilkinson and Hingston (1992) demonstrate that neural networks are applicable to information retrieval. While these studies based on neural networks can solve the problems of incomplete query representation, the inconsistent indexing problem still remains unsolved.

In this paper I describe and demonstrate the possibility of integrating an inductive learning technique with a neural network for intelligent information retrieval in order to solve both inconsistent indexing and incomplete query problems. Since an inductive learning technique has the ability to identify a document's most significant indexing terms with some relationship to their semantic significance, it provides a potential solution to the problem of inconsistent indexing. Moreover, the importance of semantic and contextual information may be reflected in the selection of index terms. Therefore, the combination of neural networks and inductive learning techniques promises to be particularly useful in areas that demand flexible inferencing and reasoning when incomplete query and inconsistent indexing problems are present. The model is based on recent developments with

1) A neural network: an information processing system consisting of a number of very simple and highly interconnected processors called neurode.

the advent of parallel distributed processing (PDP) and its attempts to model human language. The goal of such PDP systems is not necessarily just to store static knowledge, but to generate knowledge inferences by the interaction of many units' activation levels. I will discuss the previous works relevant to out study in the next section. In section 3, I will briefly introduce the methodology I employed. A description of the proposed intelligent IR system will be given in section 4. Conclusions and future research directions are provided in sections 5.

2. Previous works

As indicated above, this paper is intended to describe and demonstrate the use of neural network and inductive learning techniques in information retrieval. It is therefore important to have an understanding of the current state of knowledge of neural networks for information retrieval. A survey of the literature, however, reveals that information on the application of neural networks to problems of information retrieval is limited, and that such applications fall short of representing a true connectionists paradigm. Such a paradigm juxtaposes IR models of language, learning, competence, and performance with the human model. Mozer (1984) described an early implementation of a document retrieval system using neural networks. The network consisted of document and term (query) nodes (these were weighted connections between document and term nodes). Mozer's experiment suggested the possibility of using the internal structure of a document collection as an additional source of information for retrieval.

Belew (1987) has done some pioneering work in the use of neural networks for information retrieval, and has established an "interactive" network. In this system, nodes include both term and author. Links are weighted in correspondence to a certain word frequency measure derived from automatic analysis of the text. The researcher also introduced a modified correlation learning rule to improve its performance: one which models the behavior of its user population.

While these studies show the potential of neural networks for information retrieval, some researchers have gone beyond stand alone neural network applications and have integrated the neural network concept with existing information retrieval architectures, such as the probabilistic and vector space models. Kwok (1989; 1990) shows how a probabilistic retrieval formulation, using single terms as document components, can be implemented in a neural network environment. A modified Hebbian learning algorithm is used in the system design whereby a 3-layer (query, index term, and document) network is implemented in which queries are connected to index terms which are then connected to documents. These connections are bi-directional going back and forth between connected units. Queries and documents are regarded as neurons of the same category, and stand either as input or output units. Intra-layer connections are disallowed. Kowk's experimental model shows substantial improvement of IR effectiveness.

More recently, Wilkinson and Hingston (1991; 1992) use the Cosine measure in a neural network model for document retrieval. They also implemented a 3-layer network representation. The resulting document retrieval system's performance was tested using a standard document collection developed by Salton (1968). The researchers show that all major improvements to the performance of the network are based on strategies that have been shown to be of value in already existing vector space models.

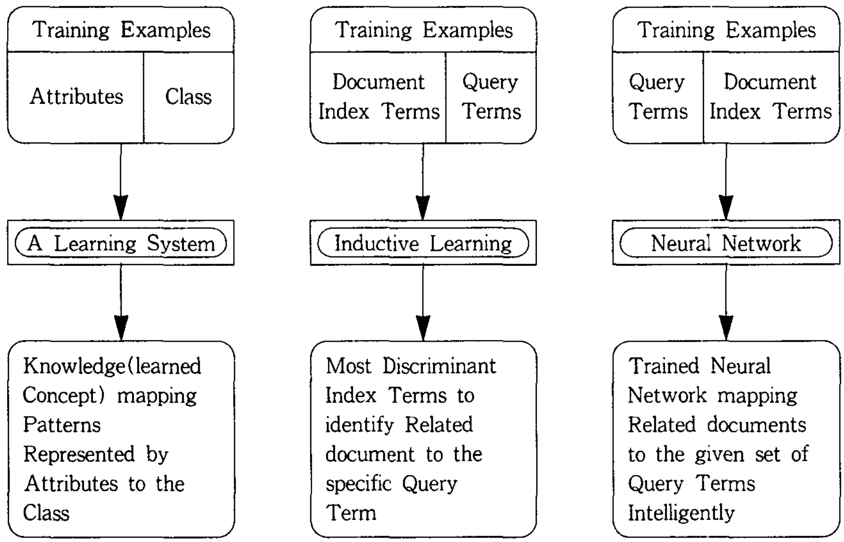
In none of the studies identified above did the researchers fully demonstrate that neural networks improve retrieval effectiveness. They did not also attempt to control a number of factors that affect retrieval effectiveness when neural networks are applied to information retrieval. For example, according to Cherkassky and Vassilas (1989), a neural network based on a backpropagation learning algorithm is strongly affected by network topology (the number of nodes and layers) and the choice of learning parameters.

While these studies do imply that the use of appropriate neural networks can enhance the performance of traditional "exact-match" techniques, they failed to fully show the potential of neural networks to map and retrieve relevant documents even when an incomplete query input is given. Finally, these studies do not tackle the problem of inconsistent indexing. As indicated earlier, one reason for low retrieval effectiveness is related to inconsistent or incomplete indexing. It is well known that, in a controlled vocabulary search, semantic and syntactic ambiguities are caused by the inconsistency of human indexers. Also, natural language retrieval presents additional problems of term imprecision resulting from linguistic ambiguities which result from a lack of synonyms and homograph control. It is of little surprise that, given these conditions, experiments involving retrieval system performance, have indicated negative results (Katzer et al. 1982; Pao and Worthen 1989). It is the purpose of this study to show that a combined method of neural networks and inductive learning techniques is desirable for effective retrieval. In other words, a combination of the neural networks and inductive learning methods may have the potential for improving recall, as well as precision measures of retrieval, and may yield valuable insights into the design of an "intelligent information retrieval system".

3. Suggested methodology

The proposed information retrieval system in this study is based on a hybrid model consisting of a neural network and an inductive learning algorithm or system (for an illustration refer to Figure 1) A learning system is a computer program that makes decisions based on the accumulated experience contained in successfully solved cases. Among the many types of learning, one can examine learning from examples. Widely used learning algorithms, in this category, include inductive learning and neural networks. For this type of learning, bundles of training examples (observations of successfully solved cases) are given to the learning system. Each training example is represented by a set of attribute values describing the observation and its decision class. Then, the system extracts the knowledge (learned concept for the decision). To provide a theoretical background on the suggested methodology, I briefly describe inductive learning and neural network models in this section.

〈Figure 1〉 A Learning System



3.1 Inductive Learning

Inductive learning is a process of acquiring knowledge by drawing inductive inferences from training examples (Michalski, 1983). Such a process involves operations of generalizing, specializing, transforming, correcting and refining knowledge representations. The input to an inductive learning algorithm consists of three parts: 1) a set of training examples, 2) generalization rules and other transformation rules, and 3) the criteria for a successful inference (Park et al.1990). Each training example has two components: a database consisting of a set of attributes, each with an assigned value; and the classification decision made by a domain expert according to the given data case. The output generated by this inductive learning algorithm is a set of decision rules consisting of inductive concept definition for each of the classes. The basis of the induction task is a set of positive and negative training examples. In the case of data collected for this study, the positive examples are documents related to the given query term, and the negative examples are all irrelevant documents.

Learning programs falling into this category include AQ-Star(Michalski 1983), PLS (Rendell 1986), and ID3 (Quinlan 1986). A detail description of the learning programs can be found in Shaw et al. (1990). The inductive learning program used in this study is Quinlan's (1988; 1992) C4.5, a descendant of ID3.

The induction method is based on the process of dividing a group of training examples by the value of a selected attribute, in the hope that the examples in a subgroup would belong to the same class. This program generates a classifier in the form of a decision tree. The *decision tree* structure includes:

- a leaf, indicating a class or
- a decision node, specifying a test to be performed on a single attribute value, with one branch and subtree for each possible outcome of the set.

C4.5 uses a method of selecting the most discriminant attribute and its threshold value to form the node of a tree using the *information gain ratio* criterion (Quinlan 1988) based on the concept of entrophy (Shannon 1948; 1951).

Inductive learning in this study is proposed to improve the retrieval effectiveness of the information retrieval system. Rather than relying on the given indexing system, the proposed system relies on index terms chosen to discriminate the most related documents as to the given query. In representing documents, the system ignores the dictated potentially inconsistent indexing scheme. Instead, it utilizes an inductive learning algorithm that isolates the most discriminant indexing terms which are related to given query terms. The selected indexing terms, are then ranked by their potential importance in discrimination so that the degree of relatedness can be controlled by adjusting the term inclusion boundary. By doing so, the proposed system can maintain its retrieval effectiveness even in the presence of the inconsistent and/or incomplete indexing. Semantic and syntactic ambiguity can also be mitigated by using these ranked discriminant index terms because ambiguity of terms or inconsistency of indexing are absorbed and imbedded in the rank which clarifies and quantifies the relatedness.

3.2 Neural network model

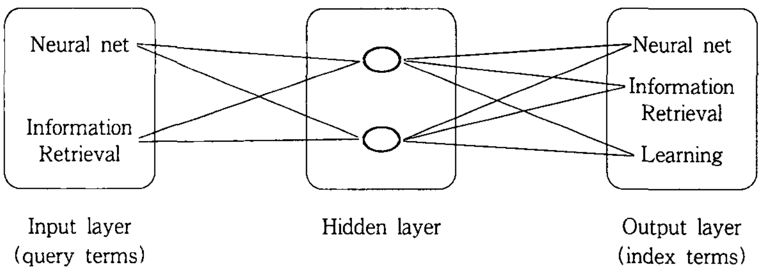
A neural network is an information processing system consisting of a number of very simple and highly interconnected processors called neurodes. The neurodes are connected with each other by many weighted links, over which signals can pass. There are several layers of neurodes in the network. Generally, the connection between neurodes occurs between layers although some networks allow the connection within one layer. The control parameters in building the network include : 1) the network topology, 2) the learning or training algorithm, and 3) learning and momentum ratios

A special class of neural networks called <u>multi-layer feed-forward neural nets</u>, has proved its utility and power in the development of complex classification systems (Wilson et al. 1992, Pao 1989), and is used in this study. A feed-forward network with appropriately linked weights can be used to model the causal relationship between a set of query terms and a set of related documents. The correlation is then fed into the inductive learning component of the retrieval system. The architecture of this system, as the name implies, consists of multiple layers of neurodes as shown in Figure 2.

These layers are : (1) an input layer that introduces information from the environment to the network, (2) an output layer that holds the responses of the network to a given pattern, and (3) middle or hidden layers that are any layers between the input and output layers. Each unit within the middle and the output layers can have a threshold, usually referred to as *biases*, associated with it.

Neural networks with hidden layers have the ability to develop internal representations. The middle layer neurodes are often characterized as features detectors that combine raw observations into higher order features, thus permitting the network to make reasonable generalizations (Salchenberger et al. 1992). Since there is no rigorous way of deriving a right number of hidden layers and neurodes, in later phases of our research we will conduct an experimental study to determine the suitable network topology.

〈Figure 2〉 An example of a Neural Network



The outputs of nodes in one layer are transmitted to nodes in another layer through links that amplify or attenuate such outputs through weight factors. Except for the input layer nodes, the net input to each node is the sum of the weighted outputs of the nodes in the prior layer. For example, the net input to a node in layer j is

$$net_j = -i\ w_{ij}\ O_i$$

The output of node j is

$$O_j = f\ (net_j)$$

where f is the activation function. Activation functions map a neurode's input to its output. It is generally a threshold level of a neurode's activity at which the neurode will output a signal. Usually, a neural network model starts with a random set of weights and a training algorithm is used to adjust these weights. In this study, the backpropagation learning algorithm (Rumelhart, Hinton, and Williams,

1985) is used to perform the training requirements. It has been widely used in the development of many neural network applications today.

From the perspective of information retrieval(IR), new IR techiques based on neural networks can deal with poorly defined user queries. Unlike current IR systems which are based on a match or non-match decision, the proposed system will ferret out related relevant documents using the query terms given, regardless the degree of their completeness. Search results can be exhaustive or specific depending on the requirement of the search problem.

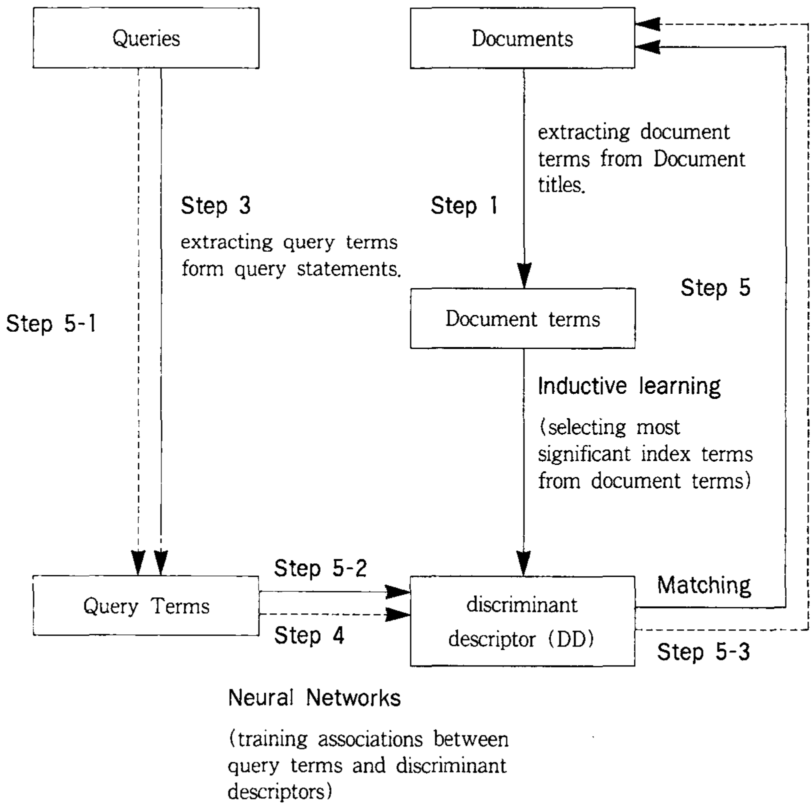
4. Design of the proposed IR system

The proposed information retrieval system is based on a hybrid model consisting of an inductive learning and neural network system. The logical flowchart for our proposed information retrieval system is shown in Figure 3. To develop the proposed system, the queries and target document database are selected initially (see appendix I). At first, we select all words, other than so called noise words or position-holders, such as prepositions and articles from the document database to form a set of document index terms. Likewise, we select a set of query terms. Next, for each query term, we identify a relevant document subset. Using the inductive learning algorithm described above, we extract the most discriminating indexing terms that distinguish the relevant documents out of the given document database. The inductive learning technique is used as a preprocessor to create a structured set of discriminant indexing terms which when grouped by their relative importance, enable a stepwise increase or decrease of document representation.

After the set of discriminant indexing terms are identified, a process of training the neural network is initiated. The input layer of the network consists of all query terms, whereby the output layer consists of all discriminant indexing terms. When trained, the network can respond effectively to a given query in enumerating relevant documents.

〈Figure 3〉 Proposed IR system

The neural network inductive learning model



4.1. Illustration of the proposed system using case examples

In this section, we explain step by step how our proposed system works. Suppose one of the queries for training reads "I want to know what neural networks are and what kinds of neural networks exist". From this query, the term "neural networks" is identified. Correspondingly, the relevant set of documents are selected from the given document database. The relevant document titles are as follows:

1. Understanding neural networks
2. Back-propagation neural network.
3. Artificial neural networks.
4. Unsupervised learning
5. Learning internal presentations by error propagation.

From these set of documents, the following document index terms are extracted:

1. understanding
2. neural networks
3. back-propagation
4. artificial
5. unsupervised
6. learning
7. internal representations
8. error-propagation

Next, the inductive learning algorithm is applied to extract discriminant index terms (most significant) among those described above. The resulting induction is shown in Figure 2. Through this process the terms "neural networks" and "learning" are identified as the most discriminating. Implied is that the proposed system can produce more relevant documents even if the document terms are different from the query terms, (contrasted with current IR systems which cannot find those documents whose titles contain the term "learning", but do not contain the term "neural networks" (i.e. document #4 and #5). The following is an example of an "IfThenElse" statement found in the decision tree:

IF index term includes "neural networks"

THEN it is a relevant document for the query containing the term "neural network"

ELSE IF index term includes "learning"

THEN it is a relevant document for the query containing the term "neural network"

ELSE it is a irrelevant document for the query containing the term "neural network"

Thus, a simple neural network is built to map the queries and discriminant terms. The training starts by feeding a pair of query and corresponding related documents, as represented by their respective discriminant index terms (refer to Figure 4). After the network is trained, the proposed system can provide the most effective way of matching subsets of documents to a given query regardless of its completeness.

4.2 The Implementation

I have implemented the first version of the intelligent information retrieval system on a Apple Macintosh Quadra and a Unix-based DEC station 5000/200 workstation. The initial test has been limited to the systems inductive learning component and involves using only 4 queries and 16 document titles containing 27 document terms, each of which is represented by a attribute in inductive learning component of the system. First, inductive learning method, C4.5 was applied to select the most significant index terms of the 16 document titles. As a result, 25 discriminant descriptors were extracted. Then, a fully connected neural network of 4 input nodes, one hidden layer with 3 nodes, and 3 output node were used to train the query terms and discriminant descriptors. (We selected only 3 discriminant descriptors for the pilot study; In a full-fledged mode, all 25 discriminant descrptors will be used). A learning rate of 0.7 and momentum of 0.9, were chosen to control the learning process. The stopping criteria was set such that the maximum error for each pattern in the training set did not exceed 0.05, or for the maximum total number of iterations not exceed 1,000.

5. Discussion

I have shown that an inductive learning method and a neural network can be used for information retrieval in a flexible fashion that allows for a variety of inputs to influence the final output. The important properties of the proposed intelligent information retrieval system are two-fold:

First, the proposed system compensates for inaccuracy or incompleteness in the query. Traditional IR systems requires an almost complete set of keywords that relate to the user's query. In this system, on the other hand, users are able to start retrieving relevant documents by using only the query terms that occur to them initially, even if those queries are not complete. During the course of retrieval, an active query term (i.e. "neural networks" in the previous example) can activate the other relevant query terms (i.e. "learning"). The retrieved set of documents will be ones that match either the query terms that were initially active, or the ones which were internally activated (induced query terms). Therefore, the system has a capability of flexible interpretation of the query. The query is used as a quide in the retrieval process, steering the systems attention to fruitful directions, not to a rigid creterion that must be matched exactly. This flexibility allows the system to help compensate for inaccurate and incomplete queries.

Second, the system can compensate for inconsistency in the indexing of document collections. The proposed system builds its own mapping structure between document terms. That is, the inductive learning technique has the ability to distinguish the most significant indexing terms with various degrees of their semantic significance. Thus, the meaning attached to a word by the user need not be equivalent to the meaning stored in the documents. Therefore, the intelligent information retrieval system can help to overcome inconsistency and incompleteness in the indexing of the collection.

Since the proposed system is an initial prototype, there are number of phases yet to follow before complete system performance can be tested. The first phase is to study and agree on definitions of complete and incomplete query representations. To measure retrieval effectiveness in this context, I will conduct tests using stepwise controls of the degree of incompleteness. The second phase is related to the optimal design of the neural network topology. In this study, the system is based on a three layer network. According to Cherkassky and Vassilas (1980), a neural network based on back-propagation learning is strongly affected by the number of layers and the choice of learning parameters. The implication is that better approaches need to be developed to determine the optimal number of layers and learning parameters. Since the variability of the number of layers and learning parameters is still an unsolved issue, some exploratory experiments need to be performed.

Bibliography

Bates, M. 1986. "Subject access in online catalogs: A design model," Journal of the American Society for Information Science, 37: 357-376.

Belew, R.K. 1986. Adaptive Information Retrieval: Machine Learning In Associative Networks. Ph.D. Dissertation, University of Michigan, Ann Arbor.

Belew, R.K. 1987. "A connectionist approach to conceptual information retrieval," Proc. First International Conference on Artificial Intelligence and Law, ACM: 116-125.

Belkin, N.J., Addy, R.N., and Brooks, H.M. 1982. "Ask for Information Retrieval: Part 1. Background and Theory," Journal of Documentation, 38: 145-164.

Blazek, R. and Bilal, D. 1988. "Problems with OPAC: A Case Study of an Academic Library," RQ, 28: 169-178.

Bookstein, A. and Swanson, D.R. 1974. "Probabilistic Models for Automatic Indexing," Journal of the American Society for Information Science, 25: 312-318.

Borgman, C.L. 1986, "Why are online catalogs are hard to user?: Lessons learned from information retrieval studies," Journal of the American Society for Information Science, 37: 384-400.

Cherkassky, V. and Vassilas, N. 1989. "Performance of Back Propagation Networks for Associative Database Retrieval," Proc. International Joint Conference on Neural Networks, Volume 1.

Coyle, K. 1985. "Record matching: A discussion," Information Technology and Libraries, 4: 57-59.

Doszkocs, T.E., Regiia, J., and Lin, X. 1990. "Connectionist Models and Information Retrieval," Annual Review of Information Science and Technology, 25.

Katzer, J., McGill, M.J., Tessier, J.A., Frakes, W., and Dasgupta, P. 1982. "A Study of the Overlap Among Documents Representations," Information Technology: Research and Development, 4: 261-274.

Kuhltau, C.C. 1991. "Inside the search process: Information seeking from the user's perspective," Journal of the American Society for Information Science, 42: 361-371.

Kwok, K.L. 1989. "A Neural Network for Probabilistic Information Retrieval," 12th International Conference on Research and Development in Information Retrieval, Cambridge, Massachusetts, 21-30.

Kwok, K.L. 1990. "Application of Neural Network to Information Retrieval," International Joint Conference on Neural Networks: Volume II: 623-626.

Larson, R.R. 1992. "Evaluation of advanced retrieval techniques in an experimental online catalog," Journal of The American Society for Information Science, 43: 34-53.

Rumelhart, D.E. and McClelland, J.L. 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. I: Foundations, Bradford: Cambridge, MA.

Michalski, R.S. 1983. "A Theory and Methodology of Inductive Learning," in Michalski, R.S., Carbonell, J., and Mitchell, T. (eds.), Machine Learning 1: 47-80.

Mori, H., Cheng Long Chung, Yousuke Kinoe, and Yoshio Hayashi 1990. "An Adaptive Document Retrieval System Using Neural Network," International Journal of Human-Computer Interaction, 2(3): 267-280.

Mozer, M. 1984. "Inductive Information Retrieval Using Parallel Distributed Computation," Technical Report, La Jolla, CA: ICS, UCSD.

Pao, Y.H. 1989. Adaptive Pattern Recognition and Neural Networks. Addison-Wesley: Reading, MA.

Pao, Miranda Lee and and Worthen, Dennis B. 1989. "Retrieval effectiveness by semantic and citation searching," Journal of the American Society for Information Science, 40: 226-235.

Park, S.C., Piramuthu, S., Raman, N., and Shaw, M.J. 1990. "Integrating Inductive Learning and Simulation in Rule-Based Scheduling," in Lecture Notes in Artificial Intelligence #462: Expert Systems in Engineering, Subseries or

Lecture Notes in Computer Science, Gottlob, G., and Neill, W. (eds.), Springer-Verlag: 152-167.

Quinlan, J.R. 1988. "Decision Trees and Multi-Valued Attributes," Machine Learning, 11: 305-318.

Quinlan, J.R. 1986. "Induction of Decision Trees", Machine Learning, 1: 82-106.

Quinlan, J.R. 1992. C4.5: Programs for Machine Learning, San Mateo, CA: Morgan Kaufmann.

Radecki, T. 1988. "The Trends in Research on Information Retrieval - The Potential For Improvements in Conventional Boolean Retrieval Systems," Information Processing & Management, 24(3): 219-227.

Rendell, L. 1986. "A General Framework for Induction and a Study of Selective Induction," Machine Learning, 1(2): 177-226.

Rumelhart, D.E.; Hinton, G.E., and Williams, R.J. 1985. "Learning Internal Representations by Error Propagation," Technical Report, La Jolla, CA: ICS, UCSD.

Salchenberger, L.M., Cinar, E.M., and Lash, N.A. 1992. "Neural Networks: A New Tool For Predicting Thrift Failures," Decision Science, 23: 899-916.

Salton, G. 1968. Automatic Information Organization and Retrieval, New York: McGraw Hill.

Salton, G. and McGill, M.J. 1983. Introduction to Modern Information Retrieval, New York: McGraw-Hill, P. 448.

Salton, G., Fox, E.A., and Wu, H. 1983. "Extended Boolean Information Retrieval," Communications of the ACM, 26(11): 1022-1036.

Salton, G., Wong, A., and Yang, C.S. 1975. "A vector space model for automatic indexing," Communications of the ACM, 18: 613-620.

Shaw, M.J., Gentry, J.A., and Piramuthu, S. 1990. "Inductive learning Methods for Knowledge-Based Decision Support: A Comparative Analysis," Computer Science in Economics and Management, 3: 147-165.

Van Rijsbergen, C.J. 1979. Information Retrieval, Second edition, Butterworths.

Watters, C.R. 1989. "Logic framework for information retrieval," Journal of The

American Society for Information Science, 40: 311-324.

Wilkinson, R. and Hingston, P. 1992. "Incorporating The Vector Space Model in a Neural Network used for Document Retrieval," Library HI Tech, 10: 69-75.

Wilkinson, R. and Hingston P. 1991. "Using the Cosine Measure in a Neural Network for Document Retrieval," 14th Annual International ACMSIGIR Conference on Research and Development in Information Retrieval: 202-210.

Wilson, R. and Sharda, R. 1992. "Neural Networks," OR/MS Today: 36-42.

Wong, S.K., Ziarko, W., and Wong, P.C.N. 1985. "Generalized vector space model in information retrieval," Proceedings of the Seventh International Conference on Information Storage and Retrieval: 18-25

초 록

귀납학습과 신경망조직을 이용한 지능형 정보검색*

김 성 희**

불리언 논리에 기초한 현재 정보검색 시스템은 두 가지 본질적인 문제점 - 1) 부정확하거나 불완전한 질의 표현과 2) 일관성 없는 색인 - 이 있다. 많은 연구자들이 신경망조직(neural network) 이 정보검색에 있어서 불완전한 질의표현 문제를 해결할 수 있다고 주장해 온 반면, 일관성없는 문제는 아직 해결하지 못한 채 남아있다. 본고에서는 이러한 두 가지 문제점을 해결하기 위해 신경망 조직과 귀납학습이 소개되고 있다. 또한 이 논문에서는 신경망 조직이 어떻게 귀납학습과 통합해서 효율적인 정보검색시스템에 응용될 수 있는지를 보여주고 있다.

* 이 논문은 한국문헌정보학회 1995년 춘계학술발표회에서 발표된 것임.

** 중앙대학교 강사.