Computational Aspects of Connectionist Interaction Information Retrieval

Sándor Dominich, Zsolt Tuza[1]

Department of Computer Science
University of Veszprém
Egyetem u. 10.
8200 Veszprém, Hungary
{dominich}{tuza}@dcs.vein.hu

Abstract

Connectionism represents a soft computing technique that aims at enhancing retrieval effectiveness, and is, at the same time, very computation demanding. In IR, only recently has computational complexity of retrieval algorithms become a research issue, although its practical importance has long been recognized. The paper presents a methodical study of the computational complexity of a connectionist retrieval algorithm, the Associative Interaction retrieval method. After a short description of the method itself, the complexity of weights computation and "winner-takes-all"-based activation spreading (i.e., retrieval) are established. This is followed by an empirical estimate of the probability to have multiple maxima, and by an asymptotic estimate of the probability to have unique maximum.

1 Introduction

The application of soft computing techniques to Information Retrieval (IR) aims at enhancing retrieval performance by trying to capture aspects that could hardly be modelled by other means numerically; this latter is important because only this can yield implementable systems. One example for such a technique is based on fuzzy sets theory, which allows for expressing the inherent vagueness encountered in the relation between terms and documents, and fuzzy logic makes it possible to express retrieval conditions by means of formulas in the weighted Boolean model; an overview can be found in (Kraft, Bordogna, Pasi, et al., 1998).

Connectionism represents another approach. Basic entities (documents, terms) are represented as an interconnected network of nodes. Artificial Neural Networks (ANN) and Semantic Networks (SN) are two techniques used for this. For example, Wordnet (Miller, 1990) is an online dictionary based on SN. ANN learning allows to model relations between documents as well as documents and terms. It was used with the principal aim to increase the accuracy of document-term weights (Bartell, Cottrell and Belew, 1995; Belew, 1987, 1989; Bienner, Guivarch and Pinon, 1990; Cunningham et al., 1997; Fuhr and Buckley, 1991; Layaida et al., 1997; Kwok, 1990). ANNs were also applied to query modification aiming at enhancing retrieval performance (Crestani, 1993; Wong and Yao, 1990), and to retrieval from legal texts (Rose and Belew, 1991; Rose, 1994; Warner, 1993).

Connectionism allows for modelling adaptive clustering (i.e., a clustering in which the cluster structure is being developed in the presence of the query or user) has proved to be a viable approach to IR (Belew, 1989; Rose, 1994; Johnson et al., 1994, 1996). Retrieval is then viewed similar to that in fixed clustering: those documents are said to be retrieved which form the same cluster, i.e., 'nearest' to query (Cohen and Kjeldsen, 1987; Belew, 1989; Kwok, 1989, 1995; Doszkocs et al., 1990; Chen, 1994, 1995; Merkl, 1999; Wermter, 2000).

[1] Also affiliated with the Computer and Automation Institute (SZTAKI), Budapest

The IR literature and research only knows a few examples when computational complexity aspects, raised by the practical operation and scaling up analysis of the retrieval system, have been addressed: Google's PageRank (Hawelivala, 1999; Kim, 2002) although several connectionist retrieval systems have been developed thus far (Niki, 1994, 1997; SirWeb; SCALIR Rose, 1994) and the retrieval systems based on the I^2R (Interaction Information Retrieval) paradigm[2]. The I^2R paradigm was suggested in (Dominich, 1994) based on the concept of interaction according to the Copenhagen Interpretation in Quantum Mechanics (query: measuring apparatus, documents: observed system, retrieval: measurement). The idea of flexible, multiple and mutual interconnections from AI^2R also appear and are investigated by a number of researchers (e.g., Salton, Allan and Singhal, 1996; Salton, Singhal, Mitra and Buckley, 1997; Pearce and Nicholas, 1996; Carrick and Watters, 1997; Liu, 1997; Mock and Vemuri, 1997; Dominich, 1997, 2001). The retrieval effectiveness of I^2R has been analysed in detail in (Dominich, 2003) using both standard test collections and user experiments.

However, as yet there have not been any methodical evaluations as regards the computational complexity of a connectionist retrieval model and system. Thus, the aim of this paper is just this. Based on the I^2R, the paper shows how a detailed analysis of the computational complexity of such a method can be carried out.

2 Associative Interaction Information Retrieval

The documents are represented as a flexibly interconnected network of objects. The interconnections are adjusted each time a new object (e.g., a document) is fed into the network. The query is interconnected with the already interconnected objects. Thus, on the one hand, new connections develop (between the object-query and the other objects), and on the other hand, some of the existing connections can change — this represents an interaction between query and documents. Retrieval is defined as recalled memories: those documents are retrieved which belong to reverberative circles triggered by a spreading of activation started at the object-query. The reverberative circles correspond to clusters, which are not fixed as they develop in the presence of the query. This model will be referred to as Associative Interaction Information Retrieval (AI^2R).

Any object o_i, $i = 1, 2, ..., M$, is assigned a set of identifiers (e.g., keywords) t_{ik}, $k = 1, 2, ..., n_i$. There are weighted and directed links between any pair (o_i, o_j), $i \neq j$, of objects. The one is the ratio between the number f_{ijp} of occurrences of term t_{jp} in object o_i, and the length n_i of o_i, i.e. total number of terms in o_i:

$$w_{ijp} = \frac{f_{ijp}}{n_i}, \quad p = 1, ..., n_j \tag{1}$$

Because w_{ijp} is analogous to the probability with which object o_i 'offers' t_{jp} (or equivalently with which t_{jp} is extracted from o_i when being in o_j), the corresponding link may be viewed as being directed from object o_i towards object o_j.

The other weight, w_{ikj}, is the inverse document frequency. If f_{jik} denotes the number of occurrences of term t_{ik} in o_j, and df_{ik} is the number of documents in which t_{ik} occurs, then:

$$w_{jik} = f_{jik} \log \frac{2M}{df_{ik}} \tag{2}$$

Because w_{ikj} is a measure of how much content of object o_j is 'seen' (or 'mirrored' back) by term t_{ik}, the corresponding link may be viewed as being directed from o_i towards o_j. The other two connections — in the opposite direction — have the same meaning as above: w_{jik} corresponds to w_{ijp}, while w_{jpi} corresponds to w_{ikj} (Figure 1). Figure 2 shows a simple example.

---

[2] http://www.dcs.vein.hu/CIR

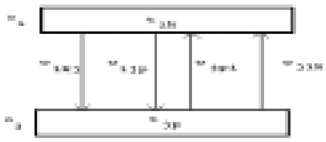

Figure 1. Associative Interaction Information Retrieval (AI^{2}R). Connections between an arbitrary object pair o_i and o_j (see text).

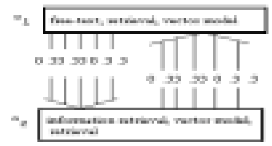

Figure 2. Associative Interaction Information Retrieval (AI^{2}R). Links with weights between object pair o_1 and o_2 (example). Links pointing in the same direction are grouped together, and shown by one big common arrow.

The process of answering a query (retrieval) is performed in two phases:

(i) Interaction. The query Q is incorporated first into the network. New weighted links appear between Q and the other objects, and some of the existing weights change (formula 2); this may be regarded as an expression of the network 'learning' the query. Figure 3 shows a simple example.

(ii) Retrieval. A spreading of activation takes place according to a winner-take-all strategy. The activation is initiated at the query, say o_j, and spreads over along the strongest connection thus passing on to another object, and so on. The total strength of the connection between any pair (o_i, o_j), $i \neq j$, of objects, and thus between the query and another object o_i is defined as follows:

$$\sum_{p=1}^{n_j} w_{jpi} + \sum_{k=1}^{n_i} w_{jik} \qquad (3)$$

The summations (formula 3) is made possible by the meaning associated to w_{jpi} and w_{jik} (formulas 1, 2); each represents a measure of the extent to which the query, represented by o_j, 'identifies' — the content of — o_i. After a finite number of steps the spreading of activation reaches an object already affected (in the worst case it passes through the entire network and eventually gets back to the query) thus giving rise to a loop called reverberative circle (as a model for short term memory). This is analogous to a local memory recalled by the query. The reverberative circle can be interpreted as an adaptive cluster associated to the query when this is present in the network. Those objects are said to be retrieved which belong to the same reverberative circle, and they are ranked in the order of maximal activation, i.e., in the order in which they are traversed. The same objects may not form the same cluster for a different query.

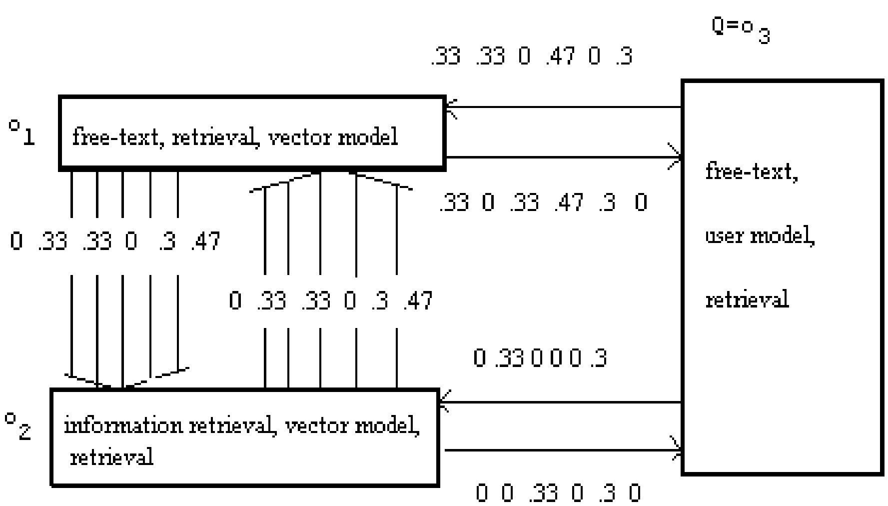

Figure 3. Associative Interaction Information Retrieval (AI^{2}R). All links having the same direction between Q and o_1, and Q and o_3 are shown as one single arrow to simplify the drawing. (a) Interaction. New connections between object-query Q(=o_3) and object-documents o_1 and o_2 are developed, and there is a changed link between o_1 and o_2 (0.47 instead of 0.3, see Figure 2 and formula 2). (b) Retrieval. The activation starts at Q, and spreads over to o_1 (total weight = .33+.33+.47+.3=1.43) from which to o_2, and then back to o_1. Q, o_1 and o_2 form a reverberative circle, and thus o_1 and o_2 will be retrieved.

3 Computational Complexity of AI^{2}R

As it could be seen in part 2 an algorithm which implements the method should compute a huge number of weights. Thus the question of tractability and hence complexity of such a computation arises, and it is answered in the following two theorems.

3.1 Complexity of Weights Computation

The complexity of weights calculation is given by:

THEOREM 1. The complexity of weights computation is polynomial.

Proof. As it can be seen from the formulas 1 and 2 there are $2 \times (n_i + n_j)$ number of weights between every pair (o_i, o_j), $i \neq j$, of which there are $\binom{M}{2}$, hence $2 \times (n_i + n_j) \times \binom{M}{2}$ has complexity $O(M^2N)$, where O denotes 'big-Oh', and $N = \max_{i,j}(n_i, n_j)$, i.e., the largest of object lengths. The computation of the sums of weights (formula 3) between a given object o_i and all the other objects o_j, of which there are $M - 1$, takes time $(n_i + n_j) \times (M - 1)$, and thus an upper bound for the computation of all sums in the network is $(n_i + n_j) \times (M - 1)^2 = O(NM^2)$ because i can vary, too, at most $M - 1$ times. Hence an overall upper bound for weights computation is $O(NM^2) + O(NM^2) = O(NM^2) = O(K^3)$, where $K = \max(N, M)$. ■

In other words the computation of weights is tractable.

3.2 Complexity of the Retrieval

The complexity of the retrieval process is given by the following results.

THEOREM 2. The links of maximum weights from all o_q can be found in $O(M^2)$ time.

Proof. For each o_q we open a stack $S(q)$ and a real variable $m(q)$. Initially $S(q)$ is empty and $m(q) = 0$. Scanning the weights w_{iq} for all the $M - 1$ values of i, we do nothing for $w_{iq} < m(q)$, but put o_i into $S(q)$ if $w_{iq} = m(q)$. Finally, if $w_{iq} > m(q)$, we first empty $S(q)$, then put o_i into $S(q)$ and redefine $m(q) := w_{iq}$. At the end of this procedure, the contents of $S(q)$ tells precisely the maximum-weight linkings at o_q. ■

THEOREM 3. One cycle can be retrieved in $O(M)$ steps.

Proof. Open a block $\underline{a} = a_1a_2\ldots a_M$, of M zeroes. Starting at o_q (that represents the query), rewrite $a_q = 1$. When at o_i, choose one $o_j \in S(i)$ (the top element). If $a_j = 1$, a cycle has been found. Otherwise we rewrite $a_j := 1$ and continue the search there. It is clear that a nonzero entry in $\underline{a}$ (and hence also a cycle) will be reached in at most M steps. ■

We define the graph G_{max} with vertex set $\{v_1,v_2,\ldots,v_M\}$ and arcs v_iv_j where $v_j \in S(i)$. By definition, a cycle is feasible with respect to the retrieval process if and only if it correspongs to a cycle in G_{max}. We denote by G'_{max} the subgraph of G_{max} containing those arcs which lie in at least one cycle.

THEOREM 4. The subgraph G'_{max}, and also a family of cycles covering all edges of G'_{max}, can be found in polynomial time.

Proof. Selecting any one arc v_iv_j, it can be tested in polynomial (in fact, linear) time whether there exists a directed path from v_j to v_i in G_{max} (as a necessary and sufficient condition for v_iv_j to lie in a cycle), and also such a path of minimum length can be found, applying Breadth-First Search. ■

3.3 Probability of Multiple Maxima

After weights summation there are $M - 1$ weighted links — $s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_M$ — from an object o_i to all the other objects. Because i varies from 1 to M there are at most $M \times (M - 1) = O(M^2)$ links to be evaluated in all (in a search). Depending on the multiplicity (i.e., unique, double, triple maximum, or higher) of the maximum of the sequence $s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_M$ the number of reverberative circles can increase. The number of retrieved objects depends on two factors: (a) the number of reverberative circles, and (b) the number of objects a reverberative circle contains. In order to render the influence of the first parameter simulations were carried out using a C program written for this purpose. *M* was taken 100,000, and different number of sequences of weights were generated at random. In each of these cases the maximum and its multiplicity was determined (Table 1).

Table 1. Simulation of the multiplicity of maxima. In 985 sequences out of 1000 sequences there was a unique maximum, in 14 cases there were double maxima, in 3 cases there was one triple maximum, and there were no maxima with multiplicity 4.

	Multiplicity of maxima Number of sequences			
Number of generated sequences	**1**	**2**	**3**	**4**
1 000	985	14	1	0
2 500	2469	30	1	0
10 000	9532	439	26	2

Drawing the empirical density function yields a curve represented by the thinner line in Figure 4. The value of the empirical density function on every interval Δx = (0, 1), (1, 2), (2, 3), (3, 4) is calculated separately using the usual ratio:

$$\frac{number\_of\_values}{length \times total\_number\_of\_values} \tag{4}$$

for each of the three cases (Table 1), and then the corresponding values are averaged. The empirical density function can be approximated by the function:

$$f(x) = u^2 e^{-u^{0.7} x} \tag{5}$$

After curve fitting (calculations carried out using standard Mathcad curve fitting) this becomes:

$$f(x) = 3.864 \cdot e^{-1.605x} \tag{6}$$

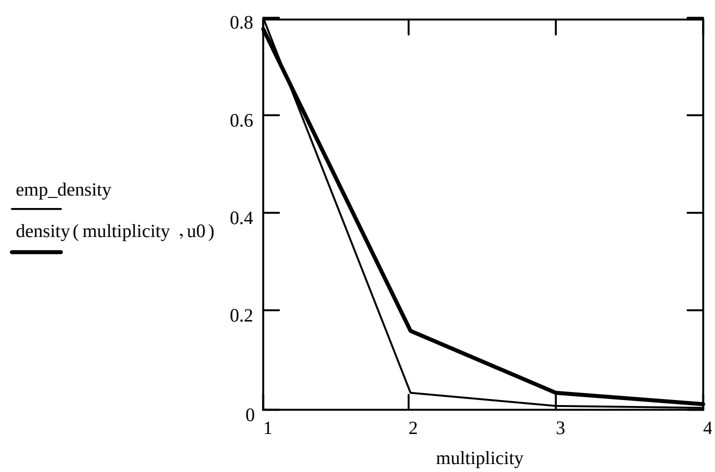

Figure 4. Empirical (thinner line) and estimated (thicker line) density functions for the multiplicity of maxima (see text).

which is an estimated density function, and thus the probability to have maximum with multiplicity 2 or 3 in a random sequence $s_1, ..., s_{i-1}, s_{i+1}, ..., s_M$ of weights can be estimated using the usual definition from probability:

$$\int_2^3 3.864 \; e^{-1.605 \; x} dx = 0.078 \tag{7}$$

The simulation results show that there always are a few multiple maxima, and their proportion is not high. The multiplicity increases with the number of sequences. (The probability of the multiplicity of maximum in a random sequence is an open, interesting and difficult mathematical problem.)

We conclude this section with an asymptotic estimate of the probability that, assuming uniform weight distribution, the maximum value is unique (that is, the retrieval procedure is continued in a unique direction).

Let S = $s_1 s_2 \ldots s_k$ be a randomly chosen sequence of length k where $s_i \in \{1, \ldots, n\}$ and Prob($s_i = j$) = $1/n$ for every i and j, independently for all i. (To simplify notation, we write k for $M - 1$; and n denotes the number of possible weight values w_{ij}.)

Suppose that the maximum occurs at a unique element of S, say $s_i = m$ and $s_j < m$ for all $j \neq i$ ($1 \leq j \leq k$). Each of the k positions in S is equally likely to occur as this particular i, and for every j we have Prob($s_j < m$) = $(m - 1)/n$. Thus,

$$\text{Prob}(S \text{ has unique maximum}) = P(n, k) = \frac{k}{n} \sum_{m=2}^{n} \left(\frac{m-1}{n} \right)^{k-1} \tag{8}$$

In order to obtain fairly tight asymptotic estimates, we write the right-hand side of (8) in the form

$$P(n, k) = \frac{k}{n} \sum_{q=1}^{n} \left(1 - \frac{q}{n} \right)^{k-1} \tag{9}$$

and apply the inequalities

$$e^{-\frac{uv}{1-u}} < (1-u)^v < e^{-uv} \tag{10}$$

As for an upper bound, we immediately obtain

$$P(n, k) < \frac{k}{n} \sum_{q \geq 1} e^{-\frac{q(k-1)}{n}} = \frac{k}{n} \frac{1}{e^{\frac{k-1}{n}} - 1} \tag{11}$$

Introducing the notation $x = k/n$, a convenient estimate seems to be

$$f(x) = \frac{x}{e^x - 1} \tag{12}$$

Lower bounds are somewhat more complicated. Since the leftmost and rightmost sides of (10) are not far from each other only if u is near zero, for the purpose we can split the sum in (9) into two parts, for q small and q large, respectively. To do this, a convenient threshold $b = \lfloor c\,(n/k) \log n \rfloor$ is taken, where c is a suitably chosen constant. Then, for $q > b$, all of the summands are smaller than n^{-c} and hence become negligible, depending on c that can be fixed with respect to the accuracy required. Assuming $q \le b$, and applying the lower bound in (10), we obtain

$$P(n,k) > \sum_{q=1}^{b} e^{-\frac{q(k-1)}{n\left(1-\frac{i}{n}\right)}} > \sum_{q>0}\left(e^{-\frac{k-1}{n\left(1-\frac{1}{b}\right)}}\right)^{q} - \sum_{q>b}\left(e^{-\frac{k-1}{n\left(1-\frac{1}{b}\right)}}\right)^{q} \tag{13}$$

Introducing the notation

$$y = e^{-\frac{k-1}{n\left(1-\frac{1}{b}\right)}}$$

for the expression occurring in the parentheses in (13), we obtain

$$P(n,k) > \frac{y}{1-y} - \frac{y^{b}+1}{1-y} \tag{14}$$

We note that the negative term in (14) is again $O(n^{-c})$, by the choice of b, i.e., it may become negligible, and then the formula simplifies to $y/(1-y)$.

If both k and n/k are large, the formula in (12) seems to be quite convenient for use. Having checked with $k = 30000$ and $n = 1000000$ (that is, six-digit accuracy in weights) as typical values, we obtain $f(0.03) \approx 0.985$ that happens to match completely with the first simulation result in Table 1. Note, however, that the other lines in the Table slightly deviate from this number.

4 Conclusions

The paper presented a methodical study of the computational complexity of a connectionist retrieval algorithm, the Associative Interaction retrieval method. After a short description of the method itself, the complexity of weights computation and "winner-takes-all"-based activation spreading (i.e., retrieval) were established. It wa shown that the computation of weights is tractable: it is polynomially upper bounded. The retrieval process is based on a "winner-takes-all" strategy, during which the maximum can be found in quadratic time, a cycle in linear time, and a family of cycles covering all vertices also in polynomial time. The practical importance of this result is that the retrieval method is tractable. This was followed by an empirical estimate of the probability to have multiple maxima. Simulations were used to generate weights, based on which the empirical density function was approximated using curve fitting, and thus the probability to have double or triple maxima was estimated to be relatively low (0.078). The practical meaning of this result consists in that running time is relatively stable and fluctuations are only rarely expected to happen; thus the user is not frustrated by having to wait for unpredictably varying running times. The paper concludes with an asymptotic estimate of the probability to have unique maximum, the theoretical results thus obtained match acceptably well the simulation results.

5 Acknowledgements

The authors would like to thank Ádám Nagy for simulations, and to acknowledge the support of grants OTKA T 037821 (National Foundation for Scientific Research, Hungary), NKFP OM-00359/2001 (Ministry of Education, National Research and Development Program, Hungary), and AKP 2001-140 (Hungarian Academy of Sciences).

References

Bartell, B., Cottrell, G.W. and Belew, R.K. (1995). Learning to retrieve information. *Proceedings of the Second Swedish Conference on Connectionism*, Skovde, Sweden, March.

Belew, R.K. (1987). A Connectionist Approach to Conceptual Information Retrieval. *Proceedings of the International Conference on Artificial Intelligence and Law* (pp. 116-126). Baltimore, ACM Press.

Belew, R.K. (1989). Adaptive information retrieval: Using a connectionist representation to retrieve and learn about documents. *Proceedings of the SIGIR 1989* (pp. 11-20). Cambridge, MA, ACM Press.

Bienner, F., Giuvarch, M. and Pinon, J.M. (1990). Browsing in hyperdocuments with the assistance of a neural network. *Proceedings of the European Conference on Hypertext* (pp. 288-297). Versailles, France.

Carrick, C. and Watters, C. (1997). Automatic Association of News Items. *Information Processing and Management*, 33(5), 615-632.

Chen, H. (1995). Machine Learning for information retrieval: Neural networks, symbolic learning and genetic algorithms. *Journal of the American Society for Information Science*, 46, 194-216.

Chen, H., Hsu, P., Orwig, R., Hoopes, L., and Nunamaker, J.F. (1994). Automatic concept classification of text from electronic meetings. *Communications of the ACM*, 37(10), 56-73.

Cohen, P., and Kjeldson, R. (1987). Information retrieval by constrained spreading activation in semantic networks. *Information Processing and Management*, 23, 255-268.

Crestani, F. (1993). Learning Strategies for an Adaptive Information Retrieval System Using Neural Networks. *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, CA, USA.

Cunningham S.J., Holmes G., Littin J., Beale R., and Witten I.H. (1997). Applying connectionist models to information retrieval. In Amari, S. and Kasobov, N. (Eds.) *Brain-Like Computing and Intelligent Information Systems* (pp 435-457). Springer-Verlag.

Dominich, S. (1994). Interaction Information Retrieval. *Journal of Documentation*, 50(3), 197-212.

Dominich, S. (1997). The Interaction-based Information Retrieval Paradigm. In Kent, A. (Ed.) *Encyclopedia of Library and Information Science*, Vol. 59, Suppl. 22. (pp. 218-236). Marcel Dekker, Inc., New York Basel Hong Kong.

Dominich, S. (2001). *Mathematical Foundations of Information Retrieval*. Kluwer Academic Publishers, Dordrecht, Boston, London.

Doszkocs, T., Reggia, J., and Lin, X. (1990). Connectionist models and information retrieval. *Annual Review of Information Science & Technology*, 25, 209-260.

Dulov, E.V., and Andrianova, N.A. (2002). Bounding the maximum value of the real-valued sequence. *Journal of Inequalities in Pure and Applied Mathematics*, vol. 3, no. 3, art. 44.

Fuhr, N. and Buckley, C. (1991). A probabilistic learning approach for document indexing. *ACM Transactions on Information Systems*, 9(3), 223-248.

Haveliwala, T.H. (1999). *Efficient Computation of PageRank*. Stanford University, http://dbpubs. stanford.edu:8090/pub/1998-31 (visited: 4 Nov 2002)

Johnson, A., Fotouhi, F., and Goel, N. (1994). Adaptive clustering of scientific data. *Proceedings of the 13th IEEE International Phoenix Conference on Computers and Communication* (pp. 241-247). Tempe, Arizona.

Johson, A., and Fotouhi, F. (1996). Adaptive clustering of hypermedia documents. *Information Systems*, 21, 549-473.

Kim, S.J., and Lee, S.H. (2002). An Improved Computation of the PageRank Algorithm. In: Crestani, F., Girolamo, M., and van Rijsbergen, C.J. (eds.) *Proceedings of the European Colloquium on Information Retrieval*. Springer LNCS 2291, pp: 73-85

Kraft, D.H., Bordogna, P. and Pasi, G. (1998). Fuzzy Set Techniques in Information Retrieval. *In*: Didier, D. and Prade, H. (Eds.) *Handbook of Fuzzy Sets and Possibility Theory*. *Approximate Reasoning and Fuzzy Infomation Systems*, (Chp. 8). Kluwer Academic Publishers, AA Dordrecht, The Netherlands.

Kwok, K.L. (1989). A Neural Network for the Probabilistic Information Retrieval. In Belkin, N.J. and van Rijsbergen, C.J. (Eds.) *Proceedings of the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 21-29). ACM Press, Cambridge, MA, USA.

Kwok, K.L. (1990). Application of Neural Networks to Information Retrieval. In Caudill, M. (Ed.) *Proceedings of the International Joint Conference on Neural Networks*, Vol. II (pp. 623-626). Hilldale, NJ, Lawrance Erlbaum Associates, Inc.

Kwok, K.L. (1995). A network approach to probabilistic information retrieval. *ACM Transactions on Information Systems*, 13(3), 243-253.

Layaida, R., Boughanem, M. and Caron, A. (1994). Constructing an Information Retrieval System with Neural Networks. *Lecture Notes in Computer Science*, 856, Springer.

Liu, G.Z. (1997). Semantic Vector Space Model: Implementation and Evaluation. *Journal of the American Society for Information Science*, 48(5), 395-417.

Merkl, D. (1999). Document classification with self-organizing maps. In Oja, E. and Kaski, S. (Eds.) *Kohonen Maps*. Elsevier, Amsterdam, The Netherlands.

Miller, G. (1990). Wordnet: An online lexical database. *International Journal of Lexicography*, 3(4), Special Issue.

Mock, K.J. and Vemuri, V.R. (1997). Information Filtering via Hill Climbing, Wordnet and Index Patterns. *Information Processing and Management*, 33(5), 633-644.

Niki, K. (1994). Connectionist Self-organized Information Retrieval System: SIR. *Proceedings of International Conference on Neural Information Processing* (pp. 1803-1808). Vol 3., http://www.etl.go.jp/etl/ninchi/niki/welcome.html

Niki, K. (1997). Sel-organizing Information Retrieval System on the Web: Sir-Web. In Kasabov, N. et al. (Eds.) *Progress in Connectionist-based Information Systems. Proceedings of the 1997 International Conference on Neural Information Processing and Intelligent Information Systems*, Vol. 2. (pp. 881-884). Springer, Singapore.

Pearce, C. and Nicolas, C. (1996). TELLTALE: Experiments in a Dynamic Hypertext Environment for Degraded and Multilingual Data. *Journal of the American Society for Information Science*, 47(4), 263-275.

Rose, D. E. (1994). *A symbolic and connectionist approach to legal information retrieval*. Hillsdale, NJ, Erlbaum.

Rose, D.E. and Belew, R.K. (1991). A connectionist and symbolic hybrid for improving legal research. *International Journal of Man-Machine Studies,* 35(1),1-33.

Salton, G., Allan, J. and Singhall, A. (1996). Automatic Text Decomposition and Structuring. *Information Processing and Management*, 32(2), 127-138.

Salton, G., Singhall, A., Mitra, M. and Buckley, C. (1997). Automatic Text Structuring and Summarization. *Information Processing and Management*, 33(2), 193-207.

SirWeb. http://www.etl.go.jp/etl/ninchi/niki/welcome-e.html

Wermter S. (2000). Neural Network Agents for Learning Semantic Text Classification. *Information Retrieval*, 3(2), 87-103.