# A Utility-Theoretic Analysis of Expected Search Length

Peter Bollmann
Technische Universität Berlin
Fachbereich Informatik, FR 5-11
Franklinstrasse 28/29
D-1000 Berlin 10

Vijay V. Raghavan
The Center for Advanced Computer Studies
University of Southwestern Louisiana
Lafayette, LA 70504-4330

## Abstract

In this paper the expected search length, which is a measure of retrieval system performance, is investigated from the viewpoint of axiomatic utility theory. Necessary and sufficient criteria for the expected search length to be an ordinal scale and sufficient criteria that it is a ratio scale are given.

## 1) Introduction

Evaluation measures play an important role in information retrieval. Recall, precision, fallout, E-measure, expected search length, recall-precision graph, etc. are examples of such measures. Although some measures, e.g. recall-precision graph, are based on more than one number, it is both convenient and attractive to adopt single-number measures. In particular, it is easier to analyze and characterize their properties and, hence, ensure that a measure most appropriate for the situation at hand is chosen.

Single-number evaluation measures can be defined as a mapping from retrieval results obtainable from an IR system, into the real numbers. Since it is intended that higher values of the measure be accorded to better retrieval results (or, the opposite for measures such as fallout), it is clear that every measure induces a preference relation on the set of retrieval results. Just as there is a preference relation connected with any measure, there is also a preference relation that represents the quality assessments made by an evaluator (or, a typical user). Thus, in choosing a measure we must ensure that a good match exists between these two preference relations. This process requires methods both for describing properties of (evaluator) preference relations and for verifying whether a given set of properties are consistent with the preference relation induced by some measure.

The early investigations aimed at connecting preference relations to measures of retrieval evaluation, based on measurement theory, were due to [CHERNIAVSKY70, RIJSBERGEN74]. More recent efforts in this direction have led to the use of elementary viewpoints to describe a given evaluation measure [BOLLMANN81] and to the identification of general properties that characterize a large family of evaluation measures [BOLLMANN84].

Another interesting approach to the investigation of evaluation measures involves the identification in detail of specific properties of evaluator preference relation and then determining whether a given utility function is consistent with these properties. Methods of axiomatic utility theory may be used in this context [FISHBURN79, KRANTZ71]. Several utility-theoretic investigations of evaluation measures have been reported [KRAFT73, KRAFT78a, KRAFT78b]. Those approaches started by explicitly giving the utility function which defines the value of retrieving one relevant document and the cost of retrieving one nonrelevant document. In this paper we take a utility-theoretic approach to give criteria for the use of the expected search length (*esl*), introduced by Cooper [COOPER68], as an evaluation measure. In contrast to the earlier utility-theoretic investigations, we start with a preference relation given by the user on a set of probability measures and derive from this the existence and the form of the utility function. As an additional result we get information specifying what kind of scale (ordinal, ratio) is obtained by the use of the expected search length for performance evaluation.

In the next section, basic concepts and notations are introduced. In the third section, necessary and sufficient conditions for the *esl* to be an ordinal scale, are obtained. We then establish, in section 4, sufficient criteria that *esl* may be treated as a ratio scale. The final section presents a summary of results and their implications.

## 2) Basic Concepts and Notations

The expected search length, or *esl*, is an evaluation measure for information retrieval results. It is defined on the weak ordering of relevant and nonrelevant documents produced as the retrieval result by the system. For such a weak order, *esl* is the expected number of nonrelevant documents the user retrieves before he/she retrieves $i$ relevant ones. Because *esl* has a different value for every $i$, we shall sometimes write $esl_i$ in case $i$ is of importance. In order to represent the weak ordering of the documents we shall use the following notation. We shall denote a relevant document by a + sign and a nonrelevant document by a − sign. Then, the retrieval result (which is a weak ordering) will be given as a distribution of + signs and − signs into ranks. We assume that there are no empty ranks.

*Example 2.1*
Let us consider the weak order given by

$$\Delta = \left( \frac{+}{-} \mid \frac{+}{--} \mid \frac{++}{---} \right).$$

Here the ranks are divided by vertical lines. The document collection consists of 10 documents, four of which are relevant and 6 are not. The first rank contains one relevant and one nonrelevant document, the second rank one relevant and two nonrelevant documents and the third rank two relevant and three nonrelevant documents. The user starts inspecting the retrieval result with the first rank. Documents within the same rank are obtained at random with respect to relevance. Hence, if the user inspects the first document it is relevant with probability ½ and it is nonrelevant with probability ½. If the user stops inspecting after having retrieved one relevant document then $esl_1\ (\Delta) = 1/2$. □

In case there are less than $i$ relevant documents in the collection we assume that the user obtains all nonrelevant documents. It is easily seen from example 2.2 that every retrieval result is connected with a probability measure $P$ on the number of retrieved nonrelevant documents.

*Example 2.2:* Let us consider the weak order Δ of example 2.1. Assuming again that the user stops after two relevant documents, we have

P(user obtains no nonrel. doc.) $= 0$

P( " " one " " ) $= 1/3$

P( " " two " " ) $= 1/3$

P( " " three " " ) $= 1/3$

P( " " $k \geqslant 4$ " " ) $= 0$ .

Based on the above probability measure, we can clearly see that $esl_2\ (\Delta) = 2$. □

## 3) Expected Search Length as Ordinal Scale

Here we want to give necessary and sufficient criteria that $esl$ may be used as an ordinal scale. To this end we introduce the notion of a simple probability measure. This and other concepts that are used in this chapter can be found in [FISHBURN79].

**Definition 3.1**: A simple probability measure $P$ on the set $X$ of elementary events is a probability measure with the additional property that there exists a finite set $A \subseteq X$ with $P\ (A) = 1$.

In the context of expected search length every retrieval result can be considered as a simple probability measure. To this end we define $X = \{x_0, x_1, x_2, \ldots\}$, where $x_j$ is the event that user retrieves $j$ nonrelevant documents and $P\ (x_j)$ is the corresponding probability. Because every retrieval result contains only a finite number of nonrelevant documents, we obtain a simple probability measure.

*Example 3.1* : Let us assume that a user wants exactly one relevant document. Let

$$\left( - \;\middle|\; \begin{array}{c} ++ \\ --- \end{array} \;\middle|\; \begin{array}{c} -- \\ -- \end{array} \right)$$

be a retrieval result. The elementary events $x_j$, $j = 0, 1, 2, \ldots$, are defined as follows:

$x_j$ — the event that the user gets exactly $j$ nonrelevant documents before he obtains one relevant document.

With these elementary events we obtain

$P\ (x_0) = 0$

$P\ (x_1) = 2/5$

$P\ (x_2) = 3/10$

$P\ (x_3) = 1/5$

$P\ (x_4) = 1/10$

$P\ (x_j) = 0, \qquad j \geqslant 5.$

We see that $P\ (A) = 1$ for $A = \{x_1, x_2, x_3, x_4\}$. Hence $P$ is a simple probability measure. □

If we assume that the user stops searching after having retrieved $i$ relevant documents and the elementary events are the numbers of retrieved nonrelevant documents, then every retrieval result can be mapped uniquely to a simple probability measure. The converse does not hold. There are simple probability measures for which no retrieval result exists.

*Example 3.2:* Let us assume the user wants exactly one nonrelevant document, and let us further assume $P\ (x_1) = P\ (x_3) = 1/2$ and $P\ (x_j) = 0$ for the rest. Then in the first rank, which contains a relevant document, there must be nonrelevant documents. There are only three possible cases which are that the relevant document is together with one, two or three nonrelevant documents in one rank because $P\ (x_j) = 0$ for $j \geqslant 4$. We see that for none of these cases the assumed simple probability distribution can be obtained. □

In the following we extend our universe of objects to all simple probability measures on $X = \{x_0, x_1, x_2, \ldots\}$. This can be considered as all possible lotteries on numbers of nonrelevant documents.

First we want to define the expectation of a real function $f$ on $X$.

**Definition 3.2**: Let $P$ be a probability measure on $X$ and $f$: $X \rightarrow \mathbf{R}$ a real function. Then

$$E(f, P) = \sum_{x \epsilon X} f(x).P(x)$$

is the expectation of $f$ with respect to $P$.

*Example 3.3:* Let $P$ be the simple probability measure obtained from a retrieval result and $f\ (x_j) = -j$, the negation of the number of retrieved nonrelevant documents. Then $E\ (f, P)$ is the negation of the expected number of nonrelevant documents, and hence $E\ (f, P)$ is inversely related to the search length. □

Next we want to consider preference relations on simple probability measures. Let $\mathbf{P}$ be the set of all simple probability measures on $X$ and let $P$, $Q$ and $R$ be simple probability measures in $\mathbf{P}$. Then,

$$P \cdot\!> Q$$

means that the user strictly prefers $P$ over $Q$, and

$$P \sim Q$$

if and only if not $P \cdot\!> Q$ and not $Q \cdot\!> P$. This means that $P$ and $Q$ are equally good for the user. We write

$$P \cdot\!\geqslant Q$$

if and only if $P \cdot\!> Q$ or $P \sim Q$. For the structure $(\mathbf{P}, \cdot\!\geqslant)$, we assume that following axioms to hold:

i) $P \cdot\!\geqslant Q$ or $Q \cdot\!\geqslant P$ for all $P, Q \epsilon \mathbf{P}$ (completeness)

ii) $(P \cdot\!\geqslant Q$ and $Q \cdot\!\geqslant R) \Rightarrow P \cdot\!\geqslant R$ for all $P, Q, R \epsilon \mathbf{P}$ (transitivity)

Hence $(\mathbf{P}, \cdot\!\geqslant)$ is a weak order.

**Definition 3.3:** $(\mathrm{P}, \cdot>)$ is a strict weak order if and only if the following axioms hold for all $P, Q, R \in \mathrm{P}$:

i) $P \cdot> Q \Rightarrow$ not $Q \cdot> P$ (asymmetry)

ii) ((not $P \cdot> Q$) and (not $Q \cdot> R$)) $\Rightarrow$ (not $P \cdot> R$) (negative transitivity).

Lemma 3.1: $(\mathrm{P}, \cdot\geqslant)$ is a weak order if and only if $(\mathrm{P}, \cdot>)$ is a strict weak order.

Proof: The proof can be found in [FISHBURN79].

Next we want to define the convex combination of two simple probability measures.

**Definition 3.3**: Let $P, Q \in \mathrm{P}$ be two simple probability measures and $0 < \alpha < 1$ a real number. Then the convex combination $\alpha P + (1 - \alpha) Q$ is a function which assigns to every $B \subseteq X$ the number $\alpha P (B) + (1 - \alpha) Q (B)$.

The convex combination of two simple probability measures is again a simple probability measure. For $x_j \in X$, $\alpha P (x_j) + (1 - \alpha) Q (x_j)$ is the probability that the user gets $x_j$ nonrelevant documents if he encounters $P$ with probability $\alpha$ and $Q$ with probability $1 - \alpha$.

Now we ask for necessary and sufficient criteria that $E (f, P)$, which in our case is related to the expected search length, may be used as an ordinal scale. Hence we ask for criteria such that

$$P \cdot\geqslant Q \Leftrightarrow E (f, P) \geqslant E (f, Q) \text{ , for } all \ P, Q \in \mathrm{P},$$

holds or equivalently that

$$P \cdot> Q \Leftrightarrow E (f, P) > E (f, Q) \text{ for } all \ P, Q \in \mathrm{P}$$

holds. The answer is partially given in the following theorem which can be found in [FISHBURN79].

Theorem 3.1: Let $\mathrm{P}$ be the set of all simple probability measures on X and $\cdot>$ be a binary relation on $\mathrm{P}$. There exists a real function $f$: $\mathrm{X} \rightarrow \mathbf{R}$ which is uniquely defined up to positive linear transformations with

$$P \cdot> Q \Leftrightarrow E (f, P) > E (f, Q)$$

for all $P, Q \in \mathrm{P}$ if and only if the following conditions hold for all $P, Q, R \in \mathrm{P}$

i) $(\mathrm{P}, \cdot>)$ is a strict weak order.

ii) $Q \cdot> P$ and $0 < \alpha < 1 \Rightarrow$
$\alpha Q + (1 - \alpha) R \cdot> \alpha P + (1-\alpha) R$

iii) If $R \cdot> Q \cdot> P$ then there exists $0 < \alpha, \beta < 1$ such that
$\beta P + (1-\beta) R \cdot> Q \cdot> \alpha P + (1- \alpha) R$.

Proof: The proof can be found in (FISHBURN79).

Condition ii) is some monotonicity axiom which says that convex combination with the same simple probability measure (especially retrieval result) does not change preference. Condition iii) says that a bad retrieval result does not matter as long as its probability $\beta$ is small enough. Additionally it says that it matters if its probability $\alpha$ is large enough. In other words, combining a small amount of $P$ (a poor result) with $R$ can still leave it better than $Q$, whereas combining a sufficiently large portion $P$ will certainly pull down $R$ (a good result) vis-a-vis $Q$. We feel that this may be the most debatable condition.

*Example 3.4:* Let $P$, $Q$ and $R$ be situations where 100, 3 and 2 nonrelevant documents respectively are retrieved for sure. Then $R \cdot\!> Q \cdot\!> P$ holds. Let $\alpha$ be 1/100. Then the expected search length says that it is better to retrieve 100 nonrelevant documents with probability 1/100 and two nonrelevant documents with probability 99/100 than to get 3 nonrelevant documents for sure. □

By theorem 3.1 the existence of the function $f$, which is useful as an evaluation measure, implies that the evaluator preference relation, $\cdot\!>$, satisfies the criteria (i), (ii) and (iii). We, in addition impose the restriction that $f\ (x_j) = -j$ be the specific form of $f$. It is shown below that if $\cdot\!>$ satisfies one additional criterion then the function $f(x_j) = -j$ will yield the desired behavior. More specifically, $f(x_k)$ is shown to be of the form $\gamma$ k + $\delta$, where $\gamma, \delta \in \mathbf{R}$ and $\gamma < 0$. Now if the condition in theorem 3.1 that $f$ is unique up to positive linear transformation is applied, the stated result follows. Incidentally this latter condition ensures that $f$ is an interval scale. To find conditions for $f$, we define a sequence of simple probability measures $P_k$, $k = 0, 1, 2, \ldots$ with

$$P_k\ (x_j) = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases} .$$

Hence $P_k$ corresponds to the retrieval result where we retrieve $k$ nonrelevant documents for sure. We get

$$E(f, P_k) = f(x_k) .$$

We now assume

$$\tfrac{1}{2}\ P_{k-1} + \tfrac{1}{2}\ P_{k+1} \sim P_k,$$

which means that the user is indifferent between retrieving $k - 1$ and $k + 1$ nonrelevant documents each with probability ½ versus $k$ nonrelevant documents for sure. This implies

$$E(f, \tfrac{1}{2} P_{k-1} + \tfrac{1}{2} P_{k+1}) = E\ (f, P_k)$$

and hence

$$\tfrac{1}{2} f\,(x_{k-1}) + \tfrac{1}{2} f\,(x_{k+1}) = f\,(x_k).$$

This difference equation in $k$ has as a general solution

$$f(x_k) = \gamma k + \delta \qquad \gamma,\ \delta \in \mathbf{R}$$

Because the user wants to retrieve as few nonrelevant documents as possible, we obtain as a consequence

$$f\ (x_k) = \gamma k + \delta. \qquad \gamma,\ \delta \in \mathbf{R}\ ,\ \gamma < 0.$$

Since, by theorem 3.1, $f(x_k)$ is uniquely defined up to positive linear transformations, we can choose

$$f(x_k) = -k.$$

Furthermore, if for all $P, Q \in \mathrm{P}$

$$P \mathrel{\cdot\!\geqslant} Q <\!=\!> \sum_{k=0}^{\infty} -k\ P\ (x_k) \geqslant \sum_{k=0}^{\infty} -k\ Q\ (x_k)\ ,$$

we obtain

$$\tfrac{1}{2} P_{k-1} + \tfrac{1}{2} P_{k+1} \sim P_k.$$

Hence the above given additional condition is both necessary and sufficient that *esl* reflects the preference relation $\mathrel{\cdot\!\geqslant}$ on P.

## 4) Expected Search Length as Ratio Scale

Here we want to give sufficient criteria that *esl* may be used as a ratio scale. We want to remember that statements like "Retrieval result $\Delta_1$ is 20% better than retrieval result $\Delta_2$" are not meaningful for ordinal or interval scales but that they are meaningful for ratio scales. In order to get criteria for a ratio scale we want to consider the following situation. Assume that a user who is interested in getting $i$ relevant documents obtains $j$ nonrelevant documents for retrieval result $P$, and $k$ nonrelevant documents with respect to another retrieval result $Q$. Then for inspecting $P$ <u>and</u> $Q$ he/she retrieves $j+k$ nonrelevant ones. Here we make the assumption that the sets of nonrelevant documents in $P$ and $Q$ are disjoint. In the case that this does not hold we can assume that two different users inspect $P$ and $Q$ and we ask for the total loss. If $P$ and $Q$ are two probability distributions then the convolution $P * Q$ is the probability distribution for random variable defined as the sum of the random variables associated with $P$ and $Q$. The convolution of two simple probability distributions is again a simple probability distribution. The convolution of two retrieval results is however not always a retrieval result.

*Example 4.1:* Let $P$ and $Q$ be defined as

$$\begin{array}{ll} P(x_0) = 1/2 & Q(x_0) = 1/3 \\ P(x_1) = 1/4 & Q(x_1) = 1/3 \\ P(x_2) = 1/4 & Q(x_2) = 1/3 \\ P(x_j) = 0, j \geqslant 3 & Q(x_j) = 0, j \geqslant 3. \end{array}$$

Then we obtain

$$P * Q\,(x_0) = P(x_0)\cdot Q(x_0) = 1/6$$

$$P * Q(x_1) = P\,(x_0).\ Q\,(X_1) + P\,(x_1).\ Q\,(x_0) = 1/6 + 1/12 = 1/4$$

The general formula is

$$P * Q(x_k) = \sum_{r+t=k} P\,(x_r)\,Q\,(x_t). \qquad \square$$

Next we want to define the concept of a risk structure [Krantz71].

**Definition 4.1:** Let P be a nonempty set of probability distributions that is closed under convolution * and let $\cdot\!\geqslant$ be a binary relation on P. (P, $\cdot\!\geqslant$, *) is a risk structure if and only if the following axioms hold for all $P, Q, R, R' \in$ P:

i) $\cdot\!\geqslant$ is a weak order

ii) $P \cdot\!\geqslant Q <=> P * R \cdot\!\geqslant Q * R$ (Monotonicity)

iii) If $P \cdot\!> Q$ then there exists a positive integer $n$ such that for all $R, R' \in$ P

$$\underbrace{P*P*P*\ldots*P}_{n}*R \ \cdot\!\geqslant \underbrace{Q*Q*Q*\ldots*Q}_{n}*R'$$

holds. (Archimedian axiom). □

The first axiom has already been introduced in the context of ordinal scales. The second axiom is an axiom of monotonicity. The Archimedian axiom can be interpreted as a thought experiment. Assume we have two systems $\Sigma$ and $\Sigma'$. With $\Sigma$ we obtain retrieval result $R$ and with $\Sigma'$ we obtain $R'$. Furthermore assume that we have an arbitrary sequence of queries and, for every query, suppose $\Sigma$ yields $P$ and $\Sigma'$ yields $Q$. Then considering the evaluation of the overall result by Archimedian axiom the performance of $\Sigma$ is not worse than that of $\Sigma'$.

*Lemma 4.1*: If for (P, $\cdot\!\geqslant$, *) the conditions of theorem 3.1 hold, $E(-k, P)$ is an order preserving mapping, and * is the convolution, then (P, $\cdot\!\geqslant$, *) is a risk structure.

*Proof:* From theorem 3.1 we know that there exists $f: X \rightarrow \mathbf{R}$ with $f(x_k) = -k$ such that

$$P \cdot\!\geqslant Q <=> E\ (f, P) \geqslant E\ (f, Q)$$

For the convolution

$$E\ (f, P\ *\ Q) = E\ (f, P) + E\ (f, Q)$$

holds because $f$ is linear in $k$.

Firstly, we want to prove monotonicity.

$$\begin{aligned} P \cdot\!\geqslant Q <=> & \ E\ (f, P) \geqslant E\ (f, Q) \\ & E\ (f, P) + E\ (f, R) \geqslant E\ (f, Q) + E\ (f, R) \\ & E\ (f, P\ *\ R) \geqslant E\ (f, Q\ *\ R) \\ & P\ *\ R \cdot\!\geqslant Q\ *\ R \end{aligned}$$

For the proof of the Archimedian axiom, we have to assume $E\ (f, P) > E\ (f, Q)$. This implies that for some integer $n \geqslant 1$

$$n\ E\ (f, P) + E\ (f, R) \geqslant n\ E\ (f, Q) + E\ (f, R')$$

holds, no matter how large $E\ (f, R)$ and $E\ (f, R')$ are. This yields

$$E\ (f, \underbrace{P\ *\dots*\ P}_{n} *R) \geqslant E\ (f, \underbrace{Q\ *\dots*\ Q}_{n}*R')$$

and hence

$$\underbrace{P\ *\dots*\ P}_{n} *\ R \cdot\!\geqslant \underbrace{Q\ *\dots*\ Q}_{n} *R' \qquad \square$$

Now we want to apply the following theorem [KRANTZ71].

<u>Theorem 4.1</u> If (P, $\cdot\!\geqslant$, *) is a risk structure then there exists a real function $\mu$: P $\rightarrow \mathbf{R}$ such that for all $P, Q \in$ P

i) $P \cdot\!\geqslant Q <=> \mu\ (P) \geqslant \mu\ (Q)$

ii) $\mu\ (P*Q) = \mu(P) + \mu\ (Q)$

holds. $\mu$ is uniquely defined up to multiplication with a positive constant. □

The application of this theorem is analogous to how theorem 3.1 was applied. Since we have shown $E(f, P * Q) = E(f, P) + (E(f, Q)$ holds if (P, $\gtrdot$, *) is a risk structure, where $f$ is as defined, the function $\mu$ of theorem 4.1 can be given the form $E(-k, P)$. For this function to be an appropriate measure, the criteria mentioned for *esl* to be an ordinal scale as well as those in definition 4.1 must hold. Furthermore the condition in theorem 4.1, that $\mu$ is unique up to multiplication by a positive constant, leads to the conclusion that, under the specified criteria, $\mu$ may be used as a ratio scale.

## 5) Summary

In order to obtain criteria for the use of expected search length, we have considered this evaluation measure in the more general context of probability distributions on numbers of nonrelevant document. For use as an ordinal scale, the necessary and sufficient criteria are the conditions of theorem 3.1 and the condition that retrieving $k-1$ and $k+1$ nonrelevant documents respectively with probability ½ each is equivalent to retrieving $k$ nonrelevant documents for sure. For use as a ratio scale, we have to include the additional necessary condition that in the case of inspecting two retrieval results the retrieved nonrelevant documents sum up.

These criteria are important for the proper use of the measurement values. If we disagree with any of the criteria for a measure to be an ordinal scale, we may not use expected search length at all. If on the other hand we agree, then the expected search length may be used as ordinal scale. In this case we have to bear in mind that certain statistics such as the arithmetic mean are not meaningful for ordinal scales and that we have to restrict ourselves to rank order statistics. In the that case we additionally agree to the criteria for a ratio scale, most of the mathematical operations such as the arithmetic mean, the geometric mean, percentages, etc., can be made on the measurement values. Hence, we are convinced that our contribution is valuable in deciding whether the choice of the *esl* as an evaluation measure is appropriate.

## REFERENCES

[COOPER68] Cooper, W.S., "Expected Search Length: A Single Measure of Retrieval Effectiveness Based on the Weak Ordering Action of Retrieval Systems," *American Documentation*, Vol. 19, 1968, pp. 30-41.

[KRAFT73] Kraft, D.H., "A Decision Theory View of the Information Retrieval Situation: An Operations Research Approach," *Journal of the ASIS*, 1973, pp. 368-376.

[KRAFT78a] Kraft, D.H., and Bookstein, A., "Evaluation of Information Retrieval Systems: A Decision Theory Approach," *Journal of the ASIS*, 1978, pp. 31-40.

[KRAFT78b] Kraft, D.H., "A Threshhold Rule Applied to the Retrieval Decision Model," *Journal of the ASIS,* 1978, pp. 77-80.

[FISHBURN79] Fishburn, P.C., *Utility Theory for Decision Making,* New York, 1979.

[KRANTZ71] Krantz, D.H., Luce, R.D., Suppes, P. and Tversky, A., *Foundations of Measurement,* Vol. I, New York and London, 1971.

[CHERNIAVSKY70] Cherniavsky, V.S., and Lakhuty, D.G., "Problem of Evaluating Retrieval Systems I," *Naucho-Techniceskaya Informazia*, Ser. 2, 24-30 (in Russian): English translation, *Automatic Documentation and Mathematical Linguistics*, 4, 9-26, 1970.

[BOLLMANN81] Bollmann, P., and Cherniavsky, V.S., "Measurement-Theoretical Investigation of the MZ-Metric," in R.R. Oddy et al. (Ed.), *Information Retrieval Research*, Butterworths, 1981.

[BOLLMANN84] Bollmann, P., "Two Axioms for Evaluation Measures in Information Retrieval," *Proceedings of the Third Joint BCS and ACM Symposium,* in *Research and Development in Information Retrieval*, Cambridge (UK), 1984.