Accounting for User's Knowledge and Search Goals in Information Retrieval Evaluation - Extended Abstract

Dima El-Zein, Célia da-Costa-Pereira

Université Côte d'Azur, Labortatoire I3S, CNRS, UMR 7271, France

Accounting for the user's cognitive aspects in the information retrieval field is still considered a challenge up until our days. Knowing that recent frameworks are trying to fill this gap, the bigger challenge remains to evaluate those frameworks and to measure the results' relevance in view of the user cognition. The majority of existing evaluation measures often consider isolated document-query environments. Traditional evaluation measures, for example, *precision* and *recall*, are not suitable to evaluate the quality of such IR algorithms. Goffman *et al.* recognised that the relevance of a document must be determined with respect to the documents appearing before it while Boyce *et al.* claimed that the change a document makes in the knowledge state must be reflected in the choice of document for the second position. The few measures that account for the user's cognitive aspects when evaluating the "relevance" of a result or ranking are limited to one search session, one query, or one search goal. The evaluation metric proposed by Clarke et al. for example systematically rewarded novelty and diversity; however, only one interaction session was considered. Also, the existing evaluation methods do not consider that the user can submit different queries. They neither track nor update the user's knowledge. Filling this gap is the main aim of our work.

We present our ongoing work on an evaluation framework for cognitive information retrieval systems that considers the user's previous knowledge and his/her search goal. The user's knowledge is a dynamic representation that changes after reading any result. Similarly, the progress towards the goal is re-assessed. The mentioned representations are not limited to one query or one search session but rather constructed throughout all the user-system interactions. This paper focuses on learning-related search tasks where the user's need (or goal) is to learn about some topic T. We adopted the user's learning model named *vocabulary learning*, which takes place at the lower level of Bloom's taxonomy. During a vocabulary learning task, in an information retrieval context, users interact with the search system to acquire keywords related to a topic T.

Search and Learning Goals: This model considers a user's information need as achieved when he/she learns about the keywords related to T. The learning goal is represented as a weighted set of keywords $VK^T = \{vk_1, \ldots, vk_{m^T}\}$, which we call *vocabulary keyword need*; m^T is the number of keywords representing the need. For example, the need to know about the topic "deep learning" might involve learning about *machine learning*, *programming*, and *calculus*. The keywords needed to learn a topic could be extracted implicitly from the query expressing the user's need, or explicitly from the search task text. A search task, for example, could be a school assignment, or a set of questions to answer. Once the keywords to be learned are defined, the number of occurrences the user has to read must also be decided. We adopt the hypothesis that to satisfy a *search need* or *learning goal*, any user should read a given number of occurrences of those keywords. We define the corresponding distribution vector *vocabulary occurrence need* $VO^T = \{vo_1, \ldots, vo_{m^T}\}$, ($vo_i \in \mathbb{N}$), noting how many instances of every keyword of VK a user has to read to achieve the learning goal T. To calculate the number of instances to be read, we suppose tn is the total number of keywords the user had to read, then its distribution among the keywords will be proportional to the weight w_i of each keyword in a corpus. We consider in this paper a personalised approach that accounts for the user's previous knowledge. Indeed, the user might already have some previous knowledge about those keywords, for example, having already read them in another document in the previous session.

User's Knowledge The user's knowledge is represented by the set of keywords he/she is knowledgeable about, stored in a knowledge base $KB = \{k_1, \ldots, k_p\}$. The amount of the user's knowledge about a word is measured by the cumulative count of the number of times the user has seen the word before; this measure is stored in $KO = \{ko_1, \ldots, ko_p\}$. We also assume that the user's knowledge of a keyword monotonically increases with each instance of it that the user reads. We made the hypothesis that users acquire their knowledge from the documents they read. When the user reads a document $d = \{dk_1, \ldots, dk_q\}$, the keywords dk_i contained in the document are taken into account in the user's knowledge - added to the user's knowledge base KB. The keywords occurrence is represented by $do = \{dko_1, \ldots, dko_q\}$. Several methods can be used to extract the keywords from the documents, for example, RAKE—Rapid Automatic Keyword Extraction Algorithm, TF-IDF, or KeyBERT. The user's previous knowledge will impact the number of occurrences he/she still has to read. Notice that, in order to solve the "cold start problem", we assume that the user's knowledge base is empty at the very first session, i.e. we suppose that the user did not read any document before the first search session ($KB = \emptyset$). However, unlike in the literature, we relax this hypothesis for the subsequent search sessions and suppose that the user can have some previous knowledge acquired in the previous sessions. In this case, we have that $KB \neq \emptyset$ i.e. $KB = \{k_1, \ldots, k_p\}$. The ith element of the user's *vocabulary occurrence need*, must consider the previous user's knowledge. The vocabulary need at the beginning of one session is calculated using the following equation, with $vk_i = k_j$.

$$vo_{ui} = \begin{cases} vo_i, & \text{if } ko_j = 0; \\ 0, & \text{if } ko_j > vo_i; \\ vo_i - ko_j, & \text{otherwise .} \end{cases} \tag{1}$$

The intuitive meaning of this formula is that if the user does not have any previous knowledge then the amount of information he/she has still to read corresponds to the original content of the *vocabulary occurrence need* i.e. $VO^T = VO_u^T$. On the contrary, if the user has already some previous knowledge that is relevant to the information need, the number of occurrences he/she has still to read should decrease with respect to VO^T.

The Evaluation Framework We present the features that our cognitive evaluation frameworks will take into account:

(I) *Measuring the gain brought by one document to a user.* Let $Gain_{VO}(d)$ denote the learning gain brought by a document d to a user u (user u is represented here by his need VO). It is considered to be the sum of gains brought by the keywords inside it $gain_{vo}(dk_i)$. Calculating the gain of every document will allow determining the most "relevant" document or the one to be returned first. In the *ideal* situation, the document should contain the exact number of occurrences needed, i.e. $\forall i, j$ such that $vk_i = dk_j$, we should have $vo_i = ko_j$. In this case, indeed, the gain would be equal to its maximum value. When the number of occurrences in the document is different (greater or lower) from the number of occurrences needed, which is very probable, a gain function must be applied. The ongoing work is studying the relevance of a document with respect to the user's need. The proportion of dk_i and vk_j will be taken into account. The results will allow us to come up with the gain formula accounting for the conditions above.

(II) *Measuring the Gain brought by a document at rank r.* Applying the gain function to each document allows us to determine the one to be ranked first. For the second and for every subsequent document, the gain measure should be reapplied with the user's updated knowledge state. The documents scoring higher will be ranked first. More precisely, before calculating the gain brought by a document, the VO_u must be updated according to the information proposed in the document before it (at rank r-1). For example, the knowledge state of a user after reading document a at rank $r = 1$, gets updated by adding all the keywords of document a and their occurrences. We can define the gain at rank r, $Gain[r]$, as the gain provided by the document at rank r: $Gain[r] = Gain_{VO}(d)$.

(2)

(III) *Calculating the cumulative and discounted Gains* We calculate the *cumulative gain vector* at rank r, $CG[r]$, as follows: $CG[r] = \sum_{j=1}^{r} Gain[j]$

In order to take into consideration the rank of the documents, before computing the cumulative gain vector, a discount may be applied at each rank to penalize documents lower in the ranking. We use the classical discount: $log_2(1 + r)$ to calculate the *discounted cumulative gain* at rank r. Finally, we normalize the calculated DCG by the ideal discounted cumulative gain vector. DCG': $n - DCG = \frac{DCG}{DCG'}$.

We discussed in this paper an ongoing work of an evaluation framework for retrieval algorithms that account for the user's cognition, especially the user's search goal and knowledge. The framework penalises the redundancy of information not only with respect to the previously proposed documents, but also to the user's knowledge, and helps in the direction of achieving a search goal. In our framework, the user's search need is represented by a defined number of occurrences of keywords the user needs to read. A document is considered relevant as long as it contains some needed keywords. After that point, we penalize for redundancy. We also consider the "changing characteristics" of the cognitive aspects. We can say that the to be proposed measure will be a first step toward evaluating the ranking and retrieval algorithms with respect to the user's cognition.

CIRCLE 2022 - Joint Conference of the Information Retrieval Communities in Europe, July 4-7, 2022, Samatan, France
*Corresponding author.
elzein@i3s.unice.fr (D. El-Zein); Celia.DA-COSTA-PEREIRA@univ-cotedazur.fr (C. da-Costa-Pereira)
https://www.i3s.unice.fr/~elzein/ (D. El-Zein); https://www.i3s.unice.fr/~cpereira/ (C. da-Costa-Pereira)
0000-0003-4156-1237 (D. El-Zein); 0000-0001-6278-7740 (C. da-Costa-Pereira)