Combining Semantic Query Disambiguation and Expansion to Improve Intelligent Information Retrieval

Bilel Elayeb[1,3(✉)], Ibrahim Bounhas[2], Oussama Ben Khiroun[1], and Narjès Bellamine Ben Saoud[1,4]

[1] RIADI Research Laboratory, ENSI Manouba University, 2010 Manouba, Tunisia
Bilel.Elayeb@riadi.rnu.tn,
oussama.ben.khiroun@gmail.com,
Narjes.Bellamine@ensi.rnu.tn
[2] LISI Laboratory of Computer Science for Industrial Systems, ISD Manouba University, 2010 Manouba, Tunisia
Bounhas.Ibrahim@yahoo.fr
[3] Emirates College of Technology, P.O. Box: 41009, Abu Dhabi, UAE
[4] Higher Institute of Informatics (ISI), Tunis El Manar University, 1002 Tunis, Tunisia

Abstract. We show in this paper how Semantic Query Disambiguation (SQD) combined with Semantic Query Expansion (SQE) can improve the effectiveness of intelligent information retrieval. Firstly, we propose and assess a possibilistic-based approach mixing SQD and SQE. This approach is based on corpus analysis using co-occurrence graphs modeled by possibilistic networks. Indeed, our model for relevance judgment uses possibility theory to take advantage of a double measure (possibility and necessity). Secondly, we propose and evaluate a probabilistic circuit-based approach combining SQD and SQE in an intelligent information retrieval context. In this approach, both SQD and SQE tasks are based on a graph data model, in which circuits between its nodes (words) represent the probabilistic scores for their semantic proximities. In order to compare the performance of these two approaches, we perform our experiments using the standard ROMANSEVAL test collection for the SQD task and the CLEF-2003 benchmark for the SQE process in French monolingual information retrieval evaluation. The results show the impact of SQD on SQE based on the recall/precision standard metrics for both the possibilistic and the probabilistic circuit-based approaches. Besides, the results of the possibilistic approach out-perform the probabilistic ones, since it takes into account of imprecision cases.

Keywords: Semantic Query Disambiguation · Semantic query expansion · Word sense disambiguation · Information retrieval · Possibility theory · Probability theory · Semantic graph · Semantic proximity

1 Introduction

Information Retrieval Systems (IRS) stay suffer from many challenges especially related to users' queries. In fact, IRS users mainly express their needs within short queries which can also contain ambiguous terms. Consequently, IRS results can

include several irrelevant documents (noise) due to the limit context provided by such queries. This noise decrease search efficiency and open the doors of two problems to be solved: Semantic Query Expansion (SQE) and Semantic Query Disambiguation (SQD) in order to improve search results.

The Semantic Query Disambiguation process [1, 2] is based on Word Sense Disambiguation (WSD) task. Indeed, Word sense disambiguation consists of selecting the suitable sense of a word given its context [3]. In fact, WSD stays as the main problem in natural language processing (NLP) and has a great influence in several related applications such as mono- and cross-language information retrieval, information extraction, machine translation (MT), content analysis, word processing, lexicography and the semantic Web applications.

Recently, WSD field has been mainly improved thanks to SensEval and SemEval competitions. For example, some works confirmed that the efficiency of MT systems has been considerably enhanced, thanks to the incorporation of a WSD task; supporting the translation process [4, 5]. However, in the information retrieval field the WSD task shows also its importance in two ways: (i) query terms can have closely related senses with other words not exist in the query. Consequently, retrieval recall can be enhanced if we take into account of these semantic links between words; and (ii) queries and documents terms can have multiple senses which decrease the retrieval precision [6]. Selecting the correct sense for both queries and documents terms may significantly enhance retrieval precision by decreasing noise in search results.

In general, WSD systems support IR systems (IRS) by identifying the suitable senses of queries and documents terms during search process. On the one side, querying step is improved by identifying the correct sense of each query term given its context. On the other side, correct senses of documents terms should be also identified in order to suitably index them given their context. Both queries and documents terms disambiguation tasks should be done before starting retrieval process. Nevertheless, this conclusion was not approved in some early research works such as [7, 8] in which search effectiveness cannot be improved despite the incorporation of a WSD system in their IRS. On the contrary, other IRS such as in [9–14] justified their efficiency enhancement thanks to the integration of WSD systems.

The Semantic Query expansion is the process of reformulating the set of the original user's query terms adding to them some other terms from their context [15, 16]. This technique aims to enhance search effectiveness in information retrieval task. In case of Web search engine, query expansion includes assessing a user's original query terms and expanding the retrieval query to match further documents. In fact, Query expansion implicates many other methods such as: (i) Re-weighting the original query terms; (ii) Stemming every term in the query in order to identify all the different morphological forms of terms; (iii) Identifying spelling errors and automatically retrieving for the corrected form or proposing it in the results; and (iv) Searching synonyms of original query terms in order to enrich the query context.

However, query expansion task can reformulate the original query by adding some ambiguous terms. This problem cannot be solved without a query disambiguation task. This relationship and dependency between these two tasks prove the need to mix them together for the purpose of improving IR efficiency. [15] and [17, 44] proposed respectively SQE and WSD approaches based on possibilistic networks. However they

did not apply their WSD algorithm on query disambiguation. They also used dictionaries as lexical resources.

This paper is a fully revised version of the conference paper [18], in which we briefly presented a combined approach for SQD and SQE tasks using possibilistic networks and applied on an extracted co-occurrence graph. We also tested possibilistic networks for enhancing IR results, by studying many combinations of scenarios of SQD, SQE and relevance feedback. In this paper, we mainly address the following new issues: (i) we explain the theoretical contribution of possibility theory compared to probability theory; (ii) we propose and assess a second probabilistic circuit-based approach mixing SQD and SQE to improve efficiency of intelligent information retrieval. In this approach, both SQD and SQE tasks are based on a dictionary modeled by a graph, in which circuits between its nodes (words) represent the probabilistic scores for their semantic proximities; (iii) we compare the performance of these two approaches by performing our experiments using the standard ROMANSEVAL test collection for the SQD task and the CLEF-2003 test collection for the SQE process in French Monolingual IR evaluation; and (iv) we propose more perspectives for future investigations.

The paper is organized as follows. We review in Sect. 2 previous works using SQD and SQE in intelligent IR. In Sect. 3, we present the co-occurrence graph model used as a resource for both SQE and SQD tasks. Section 4 details the possibilistic and the circuit-based approaches for combining SQE and SQD. Experimental results, comparative study between these two approaches and their discussion are provided in Sect. 5. Finally, Sect. 6 concludes this paper by evaluating our work and proposing some directions for future research.

2   Related Works

In this literature review, we firstly study the most important approaches of WSD improving information retrieval efficiency in Sect. 2.1. Secondly, query expansion techniques and their impact on the performance of IRS are presented in Sect. 2.2. Finally, approaches combining SQD and SQE to improve IR are discussed in Sect. 2.3.

2.1   Semantic Query Disambiguation in IR

Word sense disambiguation (WSD) is a generally known task in natural language processing (NLP) problems and IR [19]. According to the survey presented in [3, 44], WSD seriously depends on knowledge resources which are classified into two groups: structured resources (such as thesauri, electronic dictionaries, etc.) and unstructured resources (such as corpora documents).

Query disambiguation task stays a serious challenge in information retrieval process. That's why several previous works have studied the advantages and the disadvantages of integrating a SQD task in IRS. For example, the authors in [7] matched queries' terms meanings with documents' terms senses in order to take advantage of WSD in IR. However, their results are not very convincing because of the limit sense

provided by query terms, which present some disambiguation. In order to confirm the impact of WSD on IR, Sanderson in [20, 21] took advantage of a set of pseudo-words to identify query terms meanings. Nevertheless, he confirmed the important need of high accuracy WSD systems able to improve IR effectiveness.

Schütze and Pedersen in [9] didn't use predefined sense inventories, but they exploited the sense inventory directly from the text retrieval collection. Indeed, and based on the correspondences of their contexts, every word and its occurrences were clustered into senses. Authors proved via their experiments that retrieval effectiveness has been enhanced thanks to the support of WSD task. Besides, IR performance was also increased as a result of using the combination of sense-based ranking and word-based ranking. Nevertheless, the sense inventory is mainly dependent on the used collection. Consequently, it is not easy to enlarge the text collection without re-playing preprocessing task. Further, the clustering process of each word is a hard task and a time consuming step.

On the other hand, the corpus SemCor was manually sense annotated in order to discuss the impact of a wrong WSD on IR [10]. Indeed, authors represented documents and queries with correct meanings as well as synonym sets (synsets) to achieve important enhancements in IR. In fact, thanks to the use of this synset representation, experimental results proved that IR effectiveness still enhanced even they used a WSD with an error rate between 40 % and 50 %. Afterward, the authors in [22] confirmed the discriminative effect of part-of-speech (POS) information in IR tasks.

Besides, senses predefined in hand-crafted sense inventories are also used to disambiguate both queries and documents terms. In fact, identify the correct senses for documents' terms improve indexing task which cannot alone enhance the whole IR performance without a query disambiguation process. For example, and in order to disambiguate the polysemous nouns given their context, Voorhees in [8] took advantage of the hyponymy "IS-A" relation existing in WordNet [23]. All experimental results showed that the stem-based retrieval outperformed the sense-based retrieval. However, these results cannot be improved using a wrong WSD system.

Both documents and queries terms are disambiguated in [11] using a fine-grained sense inventory with an accuracy of 62.1 %. Their experiments using the TREC collections accomplished important enhancements and outperformed a standard term based vector space model. But, the general poor performances of their system and their baseline approach make not easy to objectively evaluate the exact impact of WSD in IR efficiency.

Alternatively, Kim et al. in [12] proposed a coarse-grained sense tagging technique using WordNet to tag words with 25 root senses of nouns. They exploited the stem-based index method and assign a weight to document's term according to its sense matching result with the query. Experimental results, performed using the TREC collections, showed that their coarse-grained sense tagging technique achieved significant improvement since it was flexible and consistent. Moreover, they concluded that drawbacks caused by inaccurate WSD performance can be overcome by the incorporation of senses into the classical stem-based index.

Recently, Zhong and Ng in [14] approved the relevance of WSD task to enhance IR efficiency. Authors presented and tested a technique for senses annotations applied to short queries. In fact, they integrated WSD into the language modeling method to

information retrieval [24]. Moreover, they took advantage of sense synonym relationships to more increase the IR effectiveness. Experimental tests using TREC collections proved that supervised WSD performed better results than the two other WSD baselines and considerably enhanced IR performance.

The state-of-the-art IR systems using WSD confirmed that the word sense errors can simply cancel its encouraging effect. Consequently, it is relevant to decrease the destructive effect of wrong disambiguation. One of the possible solutions consists in the incorporation of senses into traditional term index such as stem-based index. Besides, the investigation of semantic relationships between senses considerably improves IR performance. These semantic relations have showed to be useful for query expansion task in IR.

2.2 Semantic Query Expansion in IR

Semantic Query Expansion (SQE) is one of the most popular technique has been used in IR systems to enhance their effectiveness by satisfying their users' needs. Carpineto and Romano in [16] classified SQE into two principal methods: automatic query expansion (ASQE), and interactive query expansion (ISQE), which depend on user assistance. In both cases, SQE can be accomplished by several methods such as utilization of external linguistic resources (thesauri, dictionaries, ontologies, etc.), corpus analysis and relevance feedback techniques [25]. Indeed, Manning et al. in [25] classified SQE methods based on relevance feedback into three principal classes: (i) In the first class called "user relevance feedback", the returned results take into account the user's judgment; (ii) In the second class called "indirect relevance feedback" (or implicit relevance feedback), we took advantage of indirect sources of evidence such as number of hits on web page's links; and (iii) In the third class called "pseudo relevance feedback" (or blind relevance feedback), the IRS exploited the top k most relevant retrieved documents in order to expand the original query. Therefore, a set of candidate terms from these documents is added using often variants of Rocchio algorithm [26]. Even though relevance feedback may decrease noise in IR results, all these methods do not provide a solution to precisely find the suitable sense of the query terms, therefore requiring other techniques for query disambiguation.

Many SQE approaches existing in the literature have used external linguistic resources such as WordNet on English IRS [16, 27, 28]. However, these approaches are based on poor, uncertain and unclear data, while possibility theory is naturally suitable for this type of application; because it permits to express ignorance, imprecision and uncertainty [29]. In fact, it provides two kinds of relevance: (i) plausible relevance quantified by the possibility, trends to remove non-semantically similar terms (irrelevant ones); and (ii) necessity relevance increases our belief in terms not removed by possibility measure. Based on these advantages provided by the possibility theory, Ben Khiroun et al. in [30] proposed and evaluated a possibilistic approach for semantic query expansion. They later extend their approach in [15] by proposing and assessing a new possibilistic IRS which takes advantage, combine and compare the possibilistic and the probabilistic circuit-based approaches for semantic query expansion [31, 32]. Indeed, authors took advantage of the dependencies relationships between the query

terms and the articles of a dictionary to model their possibilistic network. Consequently, they investigated possibility and necessity measures to compute the corresponding possibilistic semantic similarity between terms. In fact, the SQE technique consists of injecting into the original query the most possibly and necessarily articles selected from the dictionary. Besides, SQE process was enhanced by incorporating a reweighting model which provides to the original and new query terms some relative importance. The possibilistic and the probabilistic circuit-based approaches for SQE were firstly compared in terms of their impact to IR performance. Secondly, authors mixed these two approaches by assessing two different aggregation methods. They also improved IR efficiency by integration a reweighting query terms technique in the possibilistic matching model existing in [32] to increase the performance of the expansion task. Experimental results using the standard "LeMonde94" test collection and the French dictionary "Le Grand Robert" showed partial enhancement of the results of some test queries. These enhancements, not seen at the global level of analysis, approved that the performance of any semantic query expansion technique depends on the nature of the test queries in the test collection. Moreover, query expansion task can induce noise in the search results because of the injection of polysemous words. To reduce this problem, it is suitable to incorporate a semantic disambiguation mechanism solving the problem of word sense disambiguation before and/or after the expansion task.

2.3   Combining SQD and SQE in IR

Several approaches in the literature studied the impact of SQD with SQE in IR performance using knowledge sources from thesauri. Indeed, some thesauri-based methods accomplished enhancements in IR efficiency by expanding the disambiguated query terms with synonyms and some other information from WordNet [13, 27, 33]. Besides, document expansion also benefited from the investigation of knowledge sources from WordNet which consequently prove enhancements in IRS performance [34, 35].

On the other hand, Pinto and Pérez-sanjulián in [36] exploited WordNet as external linguistic resource for both WSD and SQE. They approved the necessity of incorporating a WSD task in SQE process in order to increase IR performance. Experimental results are achieved using short and long queries from the TREC-8 text collection. These results confirmed that SQE applied on both short and long queries is not sufficient to increase IR efficiency. However, identifying the suitable sense of each ambiguous query term using a set of extracted synonyms from WordNet can mainly contribute to improve IR performance. Consequently, retrieval effectiveness was significantly improved for short queries than long ones.

Moreover, Paskalis and Khodra proposed, tested and evaluated in [2] several scenarios on IR process by using WSD, SQE, stemming and a relevance feedback technique. For WSD task, they investigated an extended implementation of Lesk algorithm [19] in order to identify the correct meaning of each query and document terms. For SQE task, they firstly exploited a co-occurrence based thesaurus built automatically from the documents collection. Secondly, they took advantage of a

pseudo relevance feedback technique using a set of top relevant documents in order to extract some representative terms from them. These terms are finally injected in the original query to improve expansion process.

Recently, authors in [17, 44] and [15] proposed and evaluated respectively a possibilistic approach for WSD and a possibilistic approach for semantic query expansion (SQE). Both of them exploited a possibilistic network in order to compute possibilistic scores between French words using the French dictionary "Le Grand Robert" as an external linguistic resource. Indeed, in the possibilistic WSD approach, authors benefited from the double relevance measure (possibility and necessity) between words and their contexts. Experimental results are done using the standard ROMANSEVAL test collection. Experiments proved a promote enhancements in terms of disambiguation rates of French words. This disambiguation performed better on nouns as they are most frequent among the existing words in the context.

In [18], authors studied the impact of Word Sense Disambiguation (WSD) on Query Expansion (SQE) for monolingual intelligent information retrieval. The proposed approaches for WSD and SQE are based on corpus analysis using co-occurrence graphs modeled by possibilistic networks. Indeed, the model for relevance judgment uses possibility theory to take advantages of a double measure (possibility and necessity). Experiments are performed using the standard ROMANSEVAL test collection for the WSD task and the CLEF-2003 benchmark for the SQE process in French monolingual Information Retrieval (IR) evaluation. The results showed the positive impact of WSD on SQE based on the recall/precision standard metrics.

3   Model Architecture and Knowledge Representation

In order to have a generic data representation that can be used for SQE, SQD and relevance feedback, we opted for a graph model that uses co-occurrences between term nodes. These relations are extracted from corpora to model contextual and similarity links. Thus, these relations are useful to compute the similarity between the terms of the queries (in the case of expansion) or between terms and senses (in the case of disambiguation).

To perform co-occurrence graph construction, we consider that two nodes are related if they exist in the same sentence. The edges are bi-oriented and weighted by the normalized co-occurrence frequency of the related terms. On the other hand, ambiguous words are related with their appropriate senses in the dictionary as considered in the following:

- T: the set of terms in the corpus
- S: the set of senses in the dictionary
- A node t_i is related to a node t_j if t_i and t_j co-occur in the same sentence; where $\{t_i, t_j \in T\}$.
- A node t_i is related to a node s_j if t_i is an ambiguous term and s_j represents a sense of t_i; where $\{t_i \in T\}$ and $\{s_j \in S\}$.

The process in Fig. 1 presents the different resources used in the SQD task, SQE and pseudo relevance feedback.
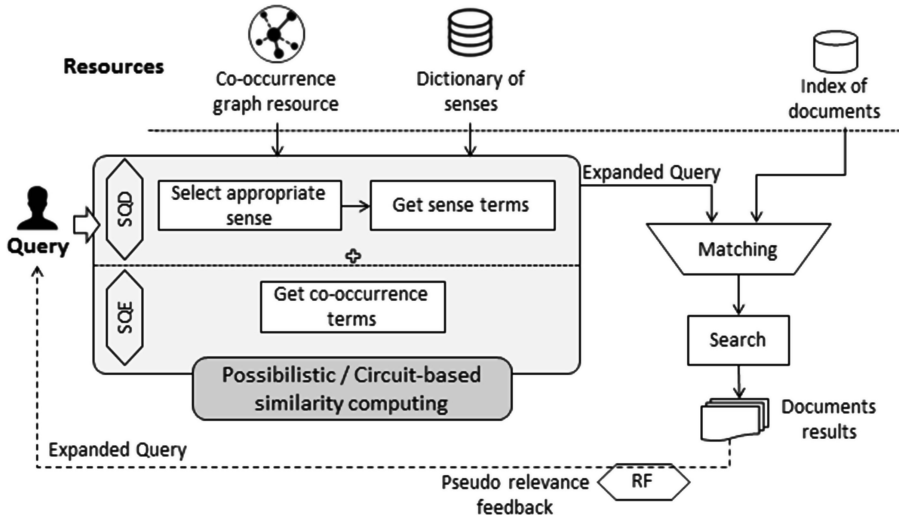
Fig. 1. Sematic query expansion using disambiguation process.

The QE module is executed to generate an expanded query starting from the initial query. In the case of ambiguous terms, the disambiguation module is used before applying QE. Thus, the best sense node having the greater possibilistic or probabilistic score is selected and the terms existing in its definition are used for expanding the original query. For both QE and QD processes, the co-occurrence graph is used to achieve possibilistic and probabilistic circuit-based calculus. Afterwards, the expanded query is matched with documents to achieve results as the classical IR process.

A pseudo relevance feedback is applied at the end of the process by extracting the most significant terms from the top first returned documents. The whole process may be iterated.

4   Possibilistic and Probabilistic Approaches for Combined SQD and SQE

We present in this section two approaches for combined SQD and SQE and we introduce an illustrative example.

4.1   A Possibilistic Approach for SQD and SQE

We based our approach on the possibilistic theory introduced by Zadeh [37] and developed by several authors [38, 39] in order to compute terms similarity in both SQE and SQD tasks. We adapted the possibilistic model architecture of Elayeb in [15] to be applied on co-occurrence graphs. Thus, we define the Degree of Possibilistic Relevance (*DPR*) for each co-occurrence graph' node n_j given a query $Q = (t_1, t_2, \ldots, t_T)$ by:

$$DPR(nj) = \Pi(n_j|Q) + N(n_j|Q) \tag{1}$$

Where $\Pi(n_j|Q)$ and $N(n_j|Q)$ represent respectively the possibility and necessity measures. The possibility measure allows to reject the non-relevant nodes identify the relevant nodes (those who are not close to the context of the query and may not be used to expand or disambiguate it). However, the necessity reinforces the relevance of the most important nodes. The two measures are computed as follows:

$$\Pi(n_j|Q) = \Pi(t_1|n_j) * \ldots * \Pi(t_T|n_j) = nft_{1j} * \ldots * nft_{Tj} \tag{2}$$

$$N(n_j|Q) = 1 - (1 - \phi n_{1j}) * \ldots * (1 - \phi n_{Tj}) \tag{3}$$

Where nft_{ij} represents the normalized frequency of query terms in the co-occurrence graph:

$$nft_{ij} = \frac{tf_{ij}}{\max_k(tf_{kj})} \tag{4}$$

In the formula (4), tf_{ij} is the weight of the edge relating the nodes t_i and n_j (i.e. the number of times the two nodes co-occur).

And:

$$\phi n_{ij} = Log_{10}\left(\frac{nCN}{nN_i}\right) * nft_{ij} \tag{5}$$

Where:
nCN = total number of nodes in the co-occurrence graph related to the query terms;
nN_i = number of nodes related to the term t_i.

Using the log function (such as in TF-IDF) allows to compute the discriminative power of the query terms. Thus, we select the graph nodes which are closest to the most discriminative items of the contextual information represented in the query.

4.2 A Probabilistic Approach Using Circuit-Based Measure for SQD and SQE

Elayeb studied in [31, 32] the query expansion problem and its impact on a possibilistic information retrieval system. His method is based on counting circuits in a graph generated from a dictionary. Indeed, in the graph of dictionary words maintain relationships that sometimes make circuits. For a given term t_i of an initial query Q^{old}, using the graph of the dictionary we compute the score of semantic proximity of term t_i with any other term t_j according to following formula [31, 32]:

$$Sem\_Pr\,ox(t_i, t_j) = \frac{Number\_of\_Circuits(t_i, t_j)}{Maximum\_Number\_of\_Circuits\_in\_the\_Graph} \tag{6}$$

Where: *Number_of_Circuits(t_i, t_j)* represents the number of circuits starting from the node t_i and passing through the node t_j in the graph of dictionary (i.e. $t_i \rightarrow \ldots \rightarrow t_j \rightarrow \ldots \rightarrow t_i$).

For the SQD task, we consider a sense S_i corresponding to an ambiguous word in the query Q. The semantic proximity of S_i to Q is generalized from the Eq. (6) as follow:

$$\text{Sem\_Pr}\,ox(S_i, Q) = \sum_{s_{ij} \in S_i} \sum_{t_k \in Q} \text{Sem\_Pr}\,ox(s_{ij}, t_k) \qquad (7)$$

The maximum length of circuit is one of important parameter in this distance. In fact, more the circuit is long more there is chance to mix various groups of meanings. However, taking into account only too short circuits would cause to cluster terms related to the same hyperonym into different groups. More details about the regrouping principle can be found in [31, 32], where author specify that the maximum length of circuit that we can take into account is about 4 edges.

4.3   Illustrative Example

Let us consider the following query admitting that it contains an ambiguous word:

    Les règles d'orthographe et de ponctuation pour la langue
allemande ont été considérablement simplifiées

Which may be translated as follows:

    The rules of spelling and punctuation for the German language
has been considerably simplified

The query is tokenized and lemmatized ignoring stop words (like pronouns, articles, etc.) as follow:

    règle (**rule**), orthographe (**spelling**), ponctuation (**punctu-
ation**), langue (**language**), allemand (**German**), cosidérable
(**considerable**), simple (**simple**)

The output query contains the ambiguous word "simple" (simple). So the WSD is executed and the sense having the best possibilistic score from ROMANSEVAL dictionary is selected (in this example we consider the sense "AII1"):

    AI2 Qui n'est formé que […]
    AI3 Qui suffit à soi seul […]
    AII1 Qui est facile à comprendre […]

Translated as:

    AI2 Which is formed only by[…]
    AI3 Sufficient to itself alone […]
    AII1 That is easy to understand […]

So the corresponding terms in the definition "AII1" are injected in query using the possibilistic approach (Fig. 2).
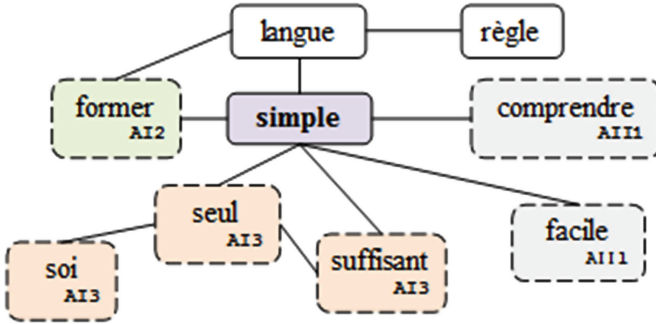
Fig. 2. A sample of the co-occurrence graph.

On the other hand, we consider this sample overview of the graph to compute the semantic proximity by using the circuit-based approach.

When enumerating the number of circuits for the three senses "AI2", "AI3" and "AII1", the sense "AI3" containing the words "seul" (alone), "soi" (itself) and "suffisant" (sufficient) has the highest semantic proximity for circuit-based computation.

Thus, this sense is the best one collating to the query context. So, the terms of the sense "AI3" are selected by the circuit-based approach for SQD task and are added to the query before expansion.

5   Experimental Results

In this section, we evaluate and compare the contribution of the possibilistic and the circuit-based approaches on both SQD and SQE tasks.

5.1   Experimental Settings

We used two test collections to experiment our proposed approach to study the impact of query disambiguation on the expansion process in French language: CLEF2003 and ROMANSEVAL.

On the one hand, the CLEF2003 test collection provides necessary tools for the evaluation of information retrieval systems on large corpora including a set of documents, a set of queries and the list of relevant documents for each query. Each query is represented in the XML format by a title containing its terms, a description and a detailed narrative text. The CLEF2003 collection for French language is composed of Le Monde 94, ATS 94, and ATS95 sub-collections forming 57 test queries and more than 300 MB of data [40].

On the other hand, the ROMANSEVAL test collection is useful for evaluating WSD approaches: it provides the necessary resources for WSD including a set of documents and a list of test sentences containing ambiguous words. A set of 60 ambiguous words distributed on three grammatical categories (20 nouns, 20 adjectives, 20 verbs) were annotated by 6 members in accordance with the senses. Each word occurrence may have one or several labels of sense or none [41].

In all our experiments, we focused only on queries from CLEF2003 test collection which contains ambiguous terms included in ROMANSEVAL test collection.

We used the Terrier experimental platform for IR to evaluate our system [42]. Different common IR measures where used like Recall/Precision, R-precision and Mean Average Precision (MAP) (for more details about state of the art IR measures see [25]). The Okapi (BM25) matching model and the Snowball stemmer (integrated in Terrier) are used for all experimentations.

In order to perform pseudo relevance feedback based on the document collection, we used the Bo1 (Bose-Einstein 1) pseudo relevance feedback method implemented in the Terrier information retrieval platform [42]. The default settings are specified as follows: the number of terms to expand a query is set to 10 and the number of top-ranked documents from which these terms are extracted is limited to 3 documents.

5.2   Evaluating SQD and SQE

This section summarises and discusses the overall performance of the various performed tests. Table 1 reports the main runs and evaluation scores for each one. For both possibilistic and circuit-based approaches, we performed two scenarios: 1- by applying the query expansion alone ("Poss_QE" and "Circuit_QE"); 2- by disambiguating the query before expansion ("Poss_QD&E" and "Circuit_QD&E"). The baseline scenario refers to the initial query without expansion or disambiguation.

Table 1. Overview of the results of the possibilistic and the probabilistic approaches.

Method		MAP	R-precision
Possibilistic	Poss_QE	0,5083	0,4742
	Poss_QD&E	**0,5124**	**0,4760**
Probabilistic	Circuit_QE	0,4920	0,4633
	Circuit_QD&E	0,5071	0,4642
Baseline		0,5487	0,5174

The last two columns present the Mean Average Precision (it is the mean of the average precision scores for each query) and the exact precision (R-Precision is the precision at rank R; where R is the total number of relevant documents) values [25].

The application of query expansion presents a performance decrease for all tests by adding new terms. However, possibilistic expansion method shows slightly better results than the circuit-based expansion method. The application of query disambiguation contributes as well for improving the retrieval results when comparing the 4 tests.

As a preliminary interpretation, this overall negative performance of query expansion (with and without query disambiguation), compared to the baseline test, could be explained by the generation of noise in search results (so lower precision values).

Oviglie et al. noted in [43] that the number of expansion terms for optimal precision varies widely across systems and topic (query) sets. Applying query expansion on long queries (that contain more than 10 words) may produce noisy and non-interpretable results as studied by Pinto and Pérez in [36]. So, we limited the number of expansion

terms to the quarter of the query' length in order to reduce the noise phenomenon according to the experimental results in [18].

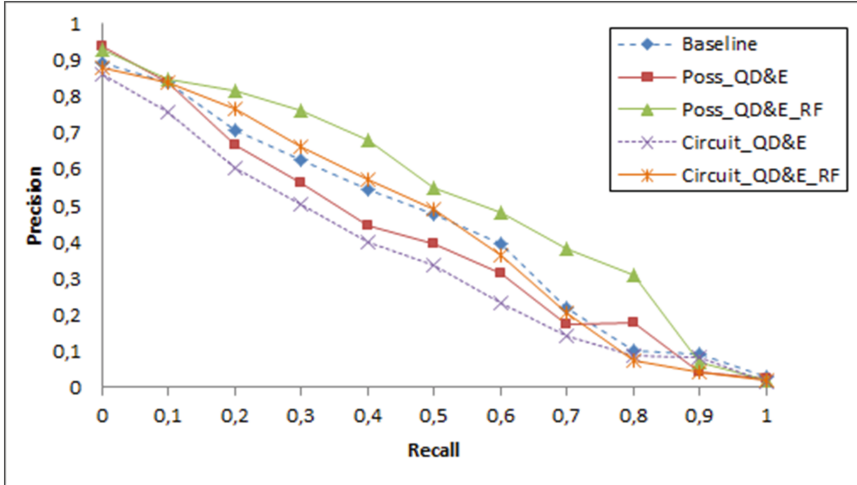We conducted a more detailed analysis by examining the Recall/Precision curve in Fig. 3.

Fig. 3. Recall-Precision curve comparing different tests.

So, focusing on the test scenario "Poss_QD&E", in which we applied both SQD and SQE, we can confirm that the query expansion combined with disambiguation is better than the baseline at high recall levels (i.e. initially better at retrieving the relevant documents).

In these detailed tests, we applied also the pseudo relevance feedback after disambiguating and expanding the queries of the test set. The application of relevance feedback with SQD and SQE improves the information retrieval performance for both possibilistic ("Poss_QD&E_RF") and probabilistic circuit-based ("Circuit_ QD&E_RF") approaches. Nevertheless, the possibilistic approach outperforms the circuit-based one. Indeed, the former method refines the search of new terms (respectively senses) for semantic expansion (respectively query disambiguation) by taking into account a double measurement of semantic proximity between the co-occurrence graph nodes.

6   Discussion and Future Works

This work presents and compares possibilistic and probabilistic approaches based on a co-occurrence graph resource. Thus, we compared the impact of word sense disambiguation in IR performance when applying query expansion and relevance feedback. The graph used in the different approaches was prepared from the collection of documents in ROMANSEVAL test collection.

Afterwards, this resource is used to choose the best candidates for disambiguation and expansion tasks. The computed score for semantic similarity depends on the used approach. The results show that the possibilistic one is finer than the probabilistic circuit-based one. This is explained by the fact that possibility and necessity measures increase the relevance of correct senses/terms and penalize the scores of the remaining ones.

Furthermore, we showed the important contribution of pseudo relevance feedback in the presented experiments of this paper. The same positive role of relevance feedback was observed in the works of Paskalis and Khodra [2]. Indeed, we join the fact that it should be better to focus on this technique to improve IR performance in parallel with word sense disambiguation methods.

In order to have a wider comparative study, we aim to compare in future works the impact of changing the knowledge source used for SQD and SQE tasks such as dictionary in place of co-occurrence graphs. However, this may present coverage problem especially for modern terms and proper nouns. As a second perspective, we aim to expand the proposed models from monolingual context to cross-lingual one by using other adapted corpora such as the SemEval corpus.

References

1. Krovetz, R.: Homonymy and polysemy in information retrieval. In: Proceedings of the 8th Conference on European Chapter of the Association for Computational Linguistics, pp. 72–79. Association for Computational Linguistics, Stroudsburg, PA, USA (1997)
2. Paskalis, F.B.D., Khodra, M.L.: Word sense disambiguation in information retrieval using query expansion. In: International Conference on Electrical Engineering and Informatics (ICEEI), pp. 1–6 (2011)
3. Navigli, R.: Word sense disambiguation: a survey. ACM Comput. Surv. (CSUR) **41**, 1–69 (2009)
4. Chan, Y.S., Ng, H.T.: Word sense disambiguation improves statistical machine translation. In: 45th Annual Meeting of the Association for Computational Linguistics (ACL-2007), pp. 33–40 (2007)
5. Carpuat, M., Wu, D.: Improving statistical machine translation using word sense disambiguation. In: The 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007), pp. 61–72 (2007)
6. Chifu, A.-G., Ionescu, R.-T.: Word sense disambiguation to improve precision for ambiguous queries. Cent. Eur. J. Comput. Sci. **2**, 398–411 (2012)
7. Krovetz, R., Croft, W.B.: Lexical ambiguity and information retrieval. ACM Trans. Inf. Syst. **10**, 115–141 (1992)
8. Voorhees, E.M.: Using WordNet to disambiguate word senses for text retrieval. In: Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 171–180. ACM, New York, NY, USA (1993)

9. Schütze, H., Pedersen, J.O.: Information retrieval based on word senses (1995)
10. Gonzalo, J., Verdejo, F., Chugur, I., Cigarrin, J.: Indexing with WordNet synsets can improve text retrieval. In: Proceedings of the COLING-ACL Workshop on Usage of WordNet in Natural Language Processing Systems, pp. 38–44 (1998)
11. Stokoe, C., Oakes, M.P., Tait, J.: Word sense disambiguation in information retrieval revisited. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 159–166. ACM, New York, NY, USA (2003)
12. Kim, S., Seo, H., Rim, H.: Information retrieval using word senses: root sense tagging approach. In: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 258–265 (2004)
13. Liu, S., Yu, C., Meng, W.: Word sense disambiguation in queries. In: Proceedings of the 14th ACM International Conference on Information and Knowledge Management, pp. 525–532. ACM, New York, NY, USA (2005)
14. Zhong, Z., Ng, H.T.: Word sense disambiguation improves information retrieval. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers – vol. 1, pp. 273–282. Association for Computational Linguistics, Stroudsburg, PA, USA (2012)
15. Elayeb, B., Bounhas, I., Ben Khiroun, O., Evrard, F., Bellamine-BenSaoud, N.: Towards a possibilistic information retrieval system using semantic query expansion. Int. J. Intell. Inf. Technol. 7, 1–25 (2011)
16. Carpineto, C., Romano, G.: A survey of automatic query expansion in information retrieval. ACM Comput. Surv. (CSUR) 44, 1–50 (2012)
17. Ben Khiroun, O., Elayeb, B., Bounhas, I., Evrard, F., Bellamine-BenSaoud, N.: A possibilistic approach for automatic word sense disambiguation. In: Proceedings of the 24th Conference on Computational Linguistics and Speech Processing (ROCLING), pp. 261–275, Taiwan (2012)
18. Ben Khiroun, O., Elayeb, B., Bounhas, I., Evrard, F., Bellamine-BenSaoud, N.: Improving query expansion by automatic query disambiguation in intelligent information retrieval. In: The 6th International Conference on Agents and Artificial Intelligence (ICAART 2014), pp. 153–160. Angers, Loire Valley, France (2014)
19. Banerjee, S., Pedersen, T.: An adapted lesk algorithm for word sense disambiguation using WordNet. In: Gelbukh, A. (ed.) CICLing 2002. LNCS, vol. 2276, pp. 136–145. Springer, Heidelberg (2002)
20. Sanderson, M.: Word sense disambiguation and information retrieval. In: Croft, B.W., van Rijsbergen, C.J. (eds.) Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '94), pp. 142–151. Springer, New York (1994)
21. Sanderson, M.: Retrieving with good sense. Inf. Retr. 2, 49–69 (2000)
22. Gonzalo, J., Peñas, A., Verdejo, F.: Lexical ambiguity and information retrieval revisited. In: Proceedings of the Joint SIGDAT Conference on Empirical Methods in NLP and Very Large Corpora (EMNLP/VLC), pp. 195–202 (1999)
23. Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.J.: Introduction to WordNet: an on-line lexical database. Int. J. Lexicogr. 3, 235–244 (1990)
24. Ponte, J.M., Croft, W.B.: A language modeling approach to information retrieval. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 275–281. ACM, New York, NY, USA (1998)
25. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press, New York (2008)

26. Rocchio, J.: Relevance Feedback in Information Retrieval. The SMART Retrieval System, pp. 313–323. Prentice-Hall, Englewood Cliffs (1971)
27. Voorhees, E.M.: Query expansion using lexical-semantic relations. In: Croft, B.W., van Rijsbergen, C.J. (eds.) Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '94), pp. 61–69. Springer, New York (1994)
28. Smeaton, A.F.: Using NLP or NLP resources for information retrieval tasks. In: Strzalkowski, T. (ed.) Natural Language Information Retrieval, pp. 99–111. Kluwer Academic Publishers, Dordrecht (1997)
29. Boughanem, M., Brini, A., Dubois, D.: Possibilistic networks for information retrieval. Int. J. Approx. Reason. **50**, 957–968 (2009)
30. Ben Khiroun, O., Elayeb, B., Bounhas, I., Evrard, F., Bellamine Ben Saoud, N.: A possibilistic approach for semantic query expansion. In: The 4th International Conference on Internet Technologies and Applications (ITA 2011), Wrexham Wales (UK), pp. 308–316 (2011)
31. Elayeb, B.: SARIPOD: Système multi-Agent de Recherche Intelligente POssibiliste de Documents Web. Ph.D. thesis, INP Toulouse France (2009)
32. Elayeb, B., Evrard, F., Zaghdoud, M., Ben Ahmed, M.: Towards an intelligent possibilistic web information retrieval using multiagent system. Interact. Technol. Smart Educ. (ITSE), Spec. Issue N. Learn. Support Syst. **6**, 40–59 (2009)
33. Fang, H.: A re-examination of query expansion using lexical resources. In: Proceedings of ACL-08: HLT, pp. 139–147 (2008)
34. Cao, G., Nie, J.-Y., Bai, J.: Integrating word relationships into language models. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 298–305. ACM, New York, NY, USA (2005)
35. Agirre, E., Arregi, X., Otegi, A.: Document expansion based on WordNet for robust IR. In: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pp. 9–17. Association for Computational Linguistics, Stroudsburg, PA, USA (2010)
36. Pinto, F.J., Pérez-sanjulián, C.F.: Automatic query expansion and word sense disambiguation with long and short queries using WordNet under vector model. Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos. **2**, 17–23 (2008)
37. Zadeh, L.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. **1**, 3–28 (1978)
38. Dubois, D., Prade, H.: Possibility theory and its application: where do we stand. Mathw. Soft Comput. **18**, 18–31 (2011)
39. Dubois, D., Prade, H.: Possibility theory. In: Meyers, R.A. (ed.) Computational Complexity, pp. 2240–2252. Springer, New York (2012)
40. Braschler, M., Peters, C.: CLEF 2003 methodology and metrics. In: Peters, C., Gonzalo, J., Braschler, M., Kluck, M. (eds.) CLEF 2003. LNCS, vol. 3237, pp. 7–20. Springer, Heidelberg (2004)
41. Segond, F.: Framework and results for French. Comput. Humanit. **34**, 49–60 (2000)
42. Ounis, I., Lioma, C., Macdonald, C., Plachouras, V.: Research directions in terrier: a search engine for advanced retrieval on the web. CEPIS Upgrad. J. **8**, 49–56 (2007)
43. Ogilvie, P., Voorhees, E., Callan, J.: On the number of terms used in automatic query expansion. Inf. Retr. **12**, 666–679 (2009)
44. Elayeb, B., Bounhas, I., Ben Khiroun, O., Evrard, F., Bellamine-BenSaoud, N.: A comparative study between possibilistic and probabilistic approaches for monolingual word sense disambiguation. Knowl. Inf. Syst. (2014). doi:10.1007/s10115-014-0753-z