

Dominant Meaning Method for Intelligent Topic-Based Information Agent towards More Flexible MOOCs

Mohammed Abdel Razek^{1,2}

¹Deanship of E-Learning and Distance Education, King Abdul-Aziz University, Jeddah, KSA

²Math & Computer Science Department, Faculty of Science, Azhar University, Cairo, Egypt

Email: maabdulrazek1@kau.edu.sa, abdelram@azhar.edu.eg

Received 24 August 2014; revised 18 September 2014; accepted 6 October 2014

Academic Editor: Dr. Steve S. H. Ling, University of Technology Sydney, Australia

Copyright © 2014 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The use of agent technology in a dynamic environment is rapidly growing as one of the powerful technologies and the need to provide the benefits of the Intelligent Information Agent technique to massive open online courses, is very important from various aspects including the rapid growing of MOOCs environments, and the focusing more on static information than on updated information. One of the main problems in such environment is updating the information to the needs of the student who interacts at each moment. Using such technology can ensure more flexible information, lower waste time and hence higher earnings in learning. This paper presents Intelligent Topic-Based Information Agent to offer an updated knowledge including various types of resource for students. Using dominant meaning method, the agent searches the Internet, controls the metadata coming from the Internet, filters and shows them into a categorized content lists. There are two experiments conducted on the Intelligent Topic-Based Information Agent: one measures the improvement in the retrieval effectiveness and the other measures the impact of the agent on the learning. The experiment results indicate that our methodology to expand the query yields a considerable improvement in the retrieval effectiveness in all categories of Google Web Search API. On the other hand, there is a positive impact on the performance of learning session.

Keywords

Massive Open Online Courses, MOOCs, Search Engine, Hypermedia Systems, Web-Based Services, Query Expansion, Probabilistic Model, Information Retrieval, Search Engine

1. Introduction

Agent technology has been suggested by experts to be a promising approach to help in overcoming the challenges of the modern e-learning environment [1]. It is now a very important issue due to the increasing in the use of new technology and distributed artificial intelligence technology for new learning environment. Business Dictionary [2] defines an agent technology as “computer programs that use artificial intelligence technology to ‘learn’, and automate certain procedures and processes”. It uses re-determined inventory levels on the Internet to automate ordering process. It emerges rapidly as one of the leading technologies for overcoming the uncertainty in a dynamic environment to help in improving distributed systems [3]. The domain of information on the web is suitable for an agent-based approach because information online is regularly distributed in dynamic varying environments.

Nowadays, information agent technology has been a long-term development and we still need it to deal with the huge information increasing on the web. One of these new application domains of e-learning is Massive Open Online Courses (MOOCs). Mass [4] found that MOOCs need more learning resources to help students in acquiring more knowledge. The web includes huge and different types of information. A small amount of information is suitable for learners. This is because web search engines are not able to classify such information related to learner’s needs. Intelligent Topic-Based Information Agent (ITBIA) offers an updated knowledge including various types of resource (alternative explanations, examples, exercises, images, applets, etc.) for students. The algorithm formulates Hierarchy of Dominant Meanings Tree using a textual course material (e.g., video transcripts) and queries from topics extracted from discussion postings between students in the course. Then it searches Google using these queries, and inserts the resulting content into categorized lists. During the discussion, the ITBIA employs a current topic to recognize the intended meaning of a query. ITBIA offers a top ranked website to students to look at them and enrich the discussion. Various approaches have been introduced in the field of information retrieval. Even though these approaches are successful, from time to time they are not able to offer precise information to the user [5]. This paper presents a mechanism based on a dominant meaning approach for searching the web.

Two experiments were done to measure a precision and to examine the performance of the ITBIA to retrieve relevant documents.

In the rest of the paper, Section 2 details the related work for the topic of MOOCs and search engine. Section 3 introduces Intelligent Topic-Based Information Agent and its frameworks. In Section 4, the methodology used by tutor agent is presented. Finally, Section 5 introduces the results of the experiments and its discussions.

2. Related Work

Because Intelligent Topic-Based Information Agent is first introduced in MOOCs, this section surveys previous work in three folds: agent technology, search engine and MOOCs.

Agent technology and their applications in information retrieval and the web information have been studied for more decades. A number of applications which apply agent technologies have already been reported in the literature such as [1] [3] and [6]. This paper studies the way of presenting more up-to-date web information related to the topic taught on MOOCs to be used as learning recourse. Usual search engines store the textual content of the web pages in a database and they revisit a prearranged record of files including keywords suggested during a query. Agent seeker [7] is a multi-agent platform that helps users to retrieve and index, web documents based on meta-search engine. The ITBIA is also devoted to practically execute a query coming from the rest of the platform, in particular from the query agent [7]. Our type of agent follows the four type of activity such as aggregation, remixing, repurposing, and sharing [8] to describe information retrieved from the web. Lin *et al.* [9] proposes a query expansion technique by mining extra query terms. They depend on the fuzzy rules to compute the weights of the additional query terms. They use these additional query terms along with the original query terms for refining the performance of search systems. Ciaramita *et al.* [10] uses a hierarchical learning architecture to define lexical semantic classification for formulate a semantic query by mining training sentences from a dictionary ontology—Word Net. Akrivas *et al.* [11] based on an inclusion relation, suggest a query expansion method which is built as a fuzzy set of semantic entities. They use the user’s profile to personalize the query. In contrast, Razek *et al.* [12] uses dominant meaning to determine the meaning of text retrieved from Google. Based on dominant meaning, he used dominant meaning to individualize a query and search result. Google uses query expansion to enable the search to routinely give additional terms to a user’s search query. When query

expansion is enabled, the appliance can expand two types of terms: words that share the same word stem as the word given by the user. For example, if a user searches for “engineer”, the appliance could add “engineers” to the query. Terms of one or more space-separated words those are synonymous or closely related to the words given by the user. For example, if a user searches for “office building”, the query could expand to include “office tower” [13]. In contrast, our methodology of expansion is built on the dominant meaning of the query. Accordingly, if the query is about “engineer”, the proposed search engine combines “computer engineer” to the query.

MOOCs is an open access web-based courses built on the typical quality of a behaviorist, cognitivist, and connectivist philosophy. MOOCs are built upon conventional online learning techniques but, scaled to a large number of students and backed up by a strong technology medium. Daniel [14] divides MOOCs into two categories: one is connectivism-based and the second is extensions of conventional online courses. Kop *et al.* [8] emphasizes that connectivist MOOCs have four types of activity, such as aggregation, remixing, repurposing, and sharing. Based on the same idea, our proposed agent has a typical action. MOOCs’ students feel to be alone in the process of learning because MOOCs focus on teaching more than learning. With MOOCs, the interaction between students and teacher is lost, overcome this interaction with peer forums, where student can contribute with questions and receive answers from their classmates.

Based on similar methods for social networking strategies, MOOCs use benefits of subject matter experts to engage students with others throughout the world with some organizing subgroups specific to their learning goals [15]. The first trying to offer a free MOOC course was in 2011 by a group of universities such as the University of Pennsylvania, Princeton University, Stanford University, and the University of Michigan [16]. In 2012, the MIT and Harvard University came mutually to offer a free course with the same manner. In 2013, the American Council on Education (ACE) recommended that institutions grant credit for five online courses on Coursera [3]. These collections of world-class universities have caused the start of increasingly greater interest among higher education leaders in such massive open online courses. However, video lectures are the fundamental of MOOCs, there are other technologies used to create a rich learning environment by combining PowerPoint slides, notes, or animated illustrations on digital whiteboards [17].

MOOCs try to maximize the usage of the asynchronous discussion forums to reimburse for the lack of synchronous interaction [17]. Other than video lectures, one of the characteristics of MOOCs is reducing instructor contact with individual students; students frequently depend on self-organized study and collaborative learning [16]. Gillet [18] uses a Connectivist MOOCs platform to create Personal Learning Environments (PLE). PLE platform is designed for fully self-regulated learning activities. The social media features are built-in PLE to boost opportunistic interaction and informal exchanges between students. PLE satisfies Kop activities and presents more media related to the topic taught. MOOCs provide influence technologies to create a rich learning environment by adding in some components, such as, professor speaking directly to the camera while accompanied by PowerPoint slides, notes, or animated illustrations on digital whiteboards [17]. Massis [4] mentions that MOOCs may give students the chance to offer the learning for all people anytime and anywhere. He also discovers that the development of MOOCs is built as an integral component and librarians have a chance to extend their influence. He finds that MOOCs is in need for more learning resources to help students in acquisition more knowledge.

3. Intelligent Topic-Based Information Agent Overview

The goal this section is to present the roles of the proposed intelligent agents to create advanced AI programs that can function usefully in web environment. Agent technologies are applied successfully on environments where it is dangerous or where it is difficult to place human being. It is clear that the choice of information agents to help in retrieving learning recourses from the internet is because the internet is virtual environment which do not exist in our physical live. The human needs huge amount of time to just read or analysis, remark a part of the information coming from the internet. Such environment needs an intelligent agent to retrieve suitable information to meet every user’s needs.

The Intelligent Topic-Based Information Agent (ITBIA) is devoted to control the metadata coming from the internet of the topic domain, and organize the results of a topic’s query. This mechanism is a conceptual model that defines the structure and behavior of the ITBIA agent. **Figure 1** illustrates ITBIA architecture and the flow of how it works. ITBIA platform works like a meta-search engine. Intelligent Topic-Based Information Agent follows the following stages:

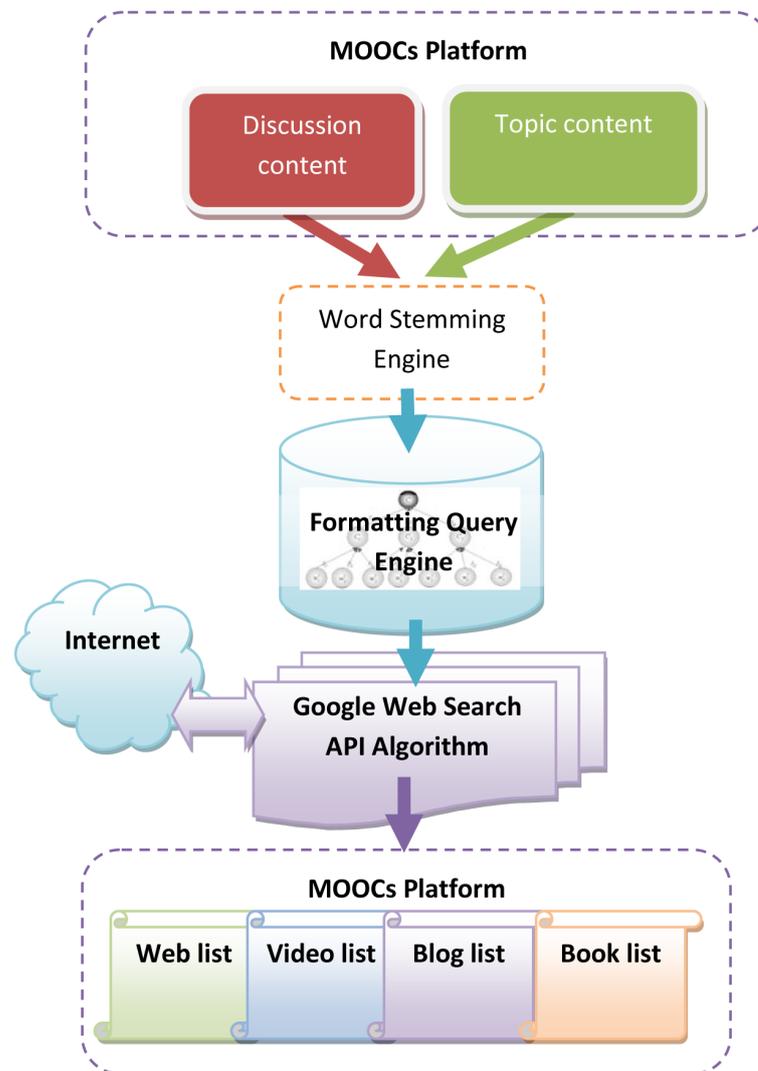


Figure 1. Intelligent topic-based information agent.

- **Extracting Query**—this is to extract some words from the discussion between learners, stem, filter, and eliminate stop words. All MOOCs courses include pre-recorded video lectures, and almost all courses have textual course material (e.g. video transcripts) included blank-audio described the video lectures. In each course, students will also have access to a discussion forum, where he/she can ask questions of course staff and discuss the course topics with his/her peers. The ITBIA has two tasks: building hierarchy of dominant meanings, and constructing a query. The methodology of building hierarchy of dominant meanings, and constructing a query are described on the next section.
- **Formatting Query**—this is to reformulating the query based on the dominant meaning technique. The concept of the query is constructed from a discussion between MOOCs' students and the application domain of the topic taught.
- **Dispatch Query**—this is to send the query to. It uses Google Web Search API, rather than using a set of online engines. ITBIA is configured search across Google for different types of search such as local search, web search, video search, blog search, news search, image search, and book search [13].
- **Display Results**—this is to employ and re-rank the results into a uniform format for displaying to learners and deleting duplicates. According to the versatility of the content on the web, it makes sense for us to use Google Web Search API to retrieve a diversity of media related to topic taught in MOOCs. After formatting the query, it is sent to Google by the Google Web Search API Algorithm (GWSAPI). GWSAPI is embedded

as a JavaScript in MOOCs webpage.

The user interface of ITBIA is built as a dynamic search box in a simulation of Coursera webpage downloaded in our Apache server, as pictured in **Figure 2**. The MOOCs platform contains the user interface of ITBIA. The user interface of ITBIA is built as a dynamic search box in a simulation to Coursera webpage downloaded in our Apache server. After formatting the query, it is sent to Google by the Google Web Search API Algorithm (GWSAPI). GWSAPI is embedded as a JavaScript in a tested MOOCs webpage in our server. The course in of Blended Learning: Personalizing Education for Students created by [19]. The ITBIA demonstrates search results in MOOCs webpage classified by its format and its importance.

3.1. Intelligent Topic-Based Information Agent Methodology

The approach described in this section employs the approach used to create dominant meaning query and is similar to the other works cited before, in particular [12] and [20], however, it is inspired by a specific application domain coming from textual course material (e.g. video transcripts) in MOOCs. The framework of MOOCs Tutors Agent consists of the following operations: drag some words from the discussion between learners, get the corresponding dominant meanings of the current concept, build a query related to the set of words reformulated from the dominant meaning tree, and re-rank the results.

3.1.1. Dominant Meaning Method for Constructing Hierarchy Tree

In this subsection, we present methodology of a query construction to retrieve the most likely document on line to match the topic been taught.

The proposed approach is applied on “Blended Learning: Personalizing Education for Students” MOOCs course through Coursera. The course is offered to anyone with a strong interest in blended learning, participants from the education technology sector, or non-profits sectors [21]. It deals with how to explore the topics of blended learning and how to develop blended learning models. It describes some various types of blended learning models and the best practices from real schools using these models.

The course includes pre-recorded video lectures along with video transcripts for the audio-back. In the course, participants have access to a discussion forum, where they can ask questions and discuss topics with other contributors.

To build the dominant meaning tree, we identify the topic of each class, and use video transcripts as a set of documents for each topic in this domain [20]. Stop words are those that occur commonly but are too general—such as “*the, is, on, or, of, how, why, etc*”. Those words were developed as a list and removed from the collection. Consider that the course contains m topics $\{T_k\}_{k=1}^m$. Each topic is represented by a given set of concepts represented by video transcript $T_k = \{C_{j,k} \mid j = 1, \dots, r\}$. Each video transcript is represented by a set of words

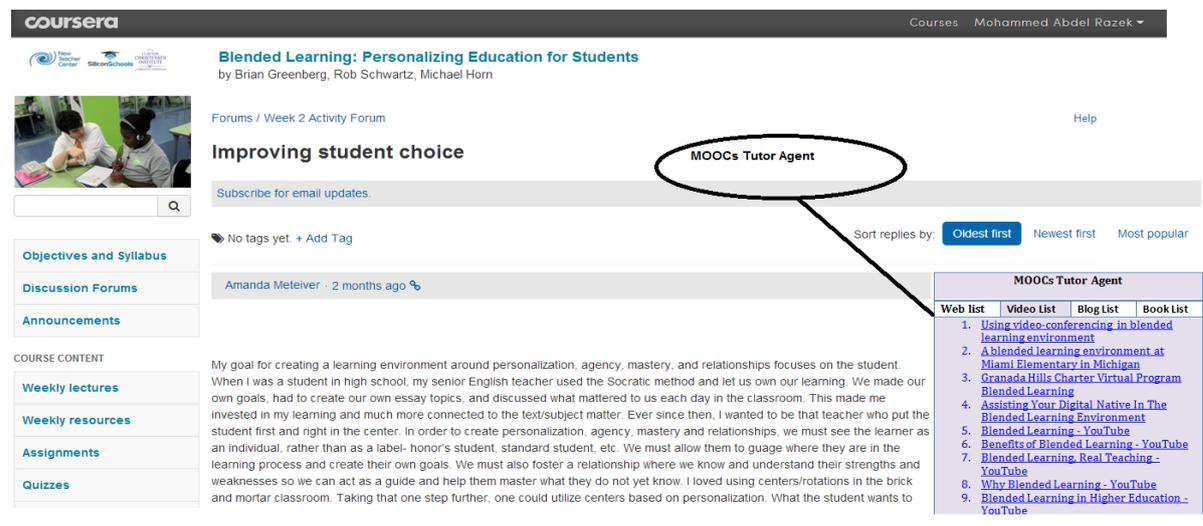


Figure 2. Coursera platform with embedding MOOCs tutors agent.

$C_{j,k} = \{w_{i,j,k} | i=1, \dots, n_{j,k}\}$, where $w_{i,j,k}$ represents the frequency of word w_j occurs in document C_k and topic T_k . To formulate the query, we need to choose top five words that match the top values of the Dominant Meaning Probability (DMP). Where the dominant meaning probability is defined as, and as shown in **Figure 3**:

$$P_{i,j,k}(w_i | C_j, T_k) = 1 - \frac{1}{\ln(2)} \left| \ln \left(\frac{w_{i,j,k}}{n_{j,k}} \right) \right| \quad (1)$$

where $i=1, \dots, n_{j,k}$, $j=1, \dots, r$, $k=1, \dots, m$, and $0 \leq P_{i,j,k}(w_j | C_k, T_i) \leq 1$.

Accordingly, a function called $\Psi_{j,k}$ Total Value of the Concept is calculated as follows:

$$\Psi_{j,k}(C_j, T_k) = \frac{1}{n_{j,k}} \left[\sum_{i=1}^{n_{j,k}} P_{i,j,k}(w_i | C_j, T_k) \right] \quad (2)$$

where $j=1, \dots, r$, $k=1, \dots, m$.

3.1.2. How ITBIA Formatting Query

The ITBIA gets some words through discussion between learners; the dilemma is to represent the query and discover which topic properly represents these words. To improve the query, ITBIA do not intentionally give attention to common words called stop words. The stop words are the words that they are not strongly affected on a search, and therefore they can significantly be got rid. To formulate the query, the ITBIA traverses the Dominant Meaning Tree in **Figure 3** using the formatting query algorithm. The algorithm starts with the word represents the topic T_k node as starting point. It choose the concept C_k which has a biggest value of the function $\Psi_{j,k}$, and then choose a word which has a biggest value of a $P_{i,k,j}$ till it find the matching word, and then formulate the query from its siblings.

Formatting Query Algorithm (Word w , Topic T_k , $\Psi_{j,k}$, $P_{i,k,j}$)

1. Sort $\Psi_{j,k}$ Descending according to j , Return $\Psi - List$;
2. Sort $P_{i,k,j}$ Descending according to i , Return $P - List$;
3. While $\Psi - List \neq []$ do begin
 1. Remove the first concept from $\Psi - List$, call it X.
 2. While $P - List$ in concept X $\neq []$ do begin
 1. Remove the first word from $P - List$ in concept X, call it Wx.
 2. If Word w matches Wx, announce success.
 3. Return words in the list $P - List$.

3.1.3. Re-Ranking Web Objects Results

Our proposed probability definition should improve retrieval effectiveness by making imposing some constraints on submitted queries and retrieved documents. Following the above Formatting Query Algorithm, suppose that the algorithm return the concept C_i , therefore the query is represented by the vector $V(w_{1,i,k}, w_{2,i,k}, \dots, w_{M,i,k})$. In general, the results of the query is represented by web objects $\{D_s\}_{s=1}^{s=R}$. These web objects are divided by four categories based on Google API search such as the images list, web list, video list, and blog list. To filter the stream of the research results, we evaluate the importance of each document

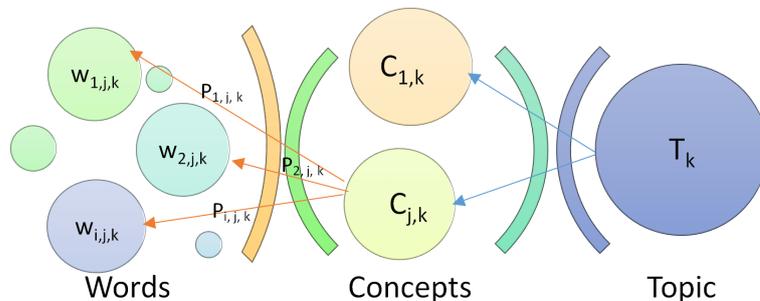


Figure 3. Dominant meaning tree.

against the query. The distance meaning function $Dis(D_s|C_l)$ is to compute the relevance of document D_s to the concept C_l , as follows:

$$Dis(D_s|C_l) = \frac{1}{M} \left[\sum_{j=1}^{j=M} \frac{F(w_{j,l,k}|D_s)}{F_s} \right] \quad (3)$$

where, $F_s = \text{Max}_{j=1, \dots, M} \{F(w_{j,l,k}|D_s)\}$, $\forall s = 1, \dots, R$; where, the function $F(w|D)$ represents the frequency of the word w which appears in document.

4. Experimental Evaluation of ITBIA

There are two experiments conducted on the Intelligent Topic-Based Information Agent: one to measure the improvement in the retrieval effectiveness and the other measures the impact of ITBIA on the learning using it through MOOCs and whether retrieved materials are relevant to students. The paper used the distance meaning function defined in the previous section to determine which document is relevant and which is not relevant from the document retrieved using Google Web Search API. Therefore, it used a precision to examine the performance of the ITBIA to retrieve relevant documents. The ‘‘Precision’’ measure can be defined as follows [22]:

$$\text{Precision} = \frac{\text{The number of retrieved relevant documents}}{\text{The number of retrieved documents}}$$

The definition is considered that the precision is the fraction of retrieved documents that are relevant to the find.

4.1. Data Set

The data set is collected using Google Web Search API Google. The dataset collected is contain 697 web objects. We take one month from Feb. 2014 to March 2014 to collect and classify the data. **Table 1** presents their main features and the number of documents and it is categorized into 4 classes and 5 query concepts.

4.1.1. Effectiveness of the Methodology Used in ITBIA

This experiment is conducted to show the effectiveness of the methodology used in searching in ITBIA. The experiment is to compare our different query expansion methods, referring to previous section for the relative performances. We used the materials of a course named ‘‘Blended Learning: Personalizing Education for Students’’ in Coursera [21] to build the dominant meaning tree. We used video transcripts for each six domain in the course as a training set for building the concepts and its related words in the tree.

Based on the definition of the Distance Meaning Function, we consider the document D is relevant if $0.5 \leq Dis(D|C) \leq 1$ and is irrelevant if $0 \leq Dis(D|C) < 0.5$.

The results are filtered and classified based on Google Web Search API, such as web list, video list, blog list, and book list. Later, the distance meaning values are calculated for the four lists for both query without expansion, and query dominant meaning expansion using Equation (2). **Table 2** shows performance improvement when the query is expanded by using dominant meanings.

Table 1. Number of document in data set.

	Concepts	Web objects				Total
		Web list	Video list	Blog list	Book list	
Q1	Blended learning models	49	37	25	24	135
Q2	Role of the student	44	38	32	33	147
Q3	Role of the teacher	38	35	31	29	133
Q4	Change at the school level	47	37	33	31	148
Q5	Hardware, software and space	41	35	31	27	134
	Total	219	182	152	144	697

Table 2. Number of document in data set based on the dominant meanings.

Concepts	Web objects				
	Query without expansion		Query dominant meaning expansion		
	Irrelevant	Relevant	Irrelevant	Relevant	
Q1	Blended learning models	55	80	21	114
Q2	Role of the student	80	67	40	107
Q3	Role of the teacher	77	56	34	99
Q4	Change at the school level	83	65	39	109
Q5	Hardware, software and space	89	45	37	97
	Total	384	313	224	473

Figure 4 summarizes the results of using our dominant meaning query expansion methods, measuring their relative performance in terms of the number of relevant documents retrieved. It seems that the dominant meaning methods which take into consideration the meaning of the query are better than the traditional expansion for the query.

Based on these results, the query without expansion did not provide any noticeable performance improvement for any of the five topics. However, the best result for the measure was at the topic “blended learning models” in **Figure 4**. We believe this is because the Blended learning subject is popular topic on the web. In the same result, the topic “hardware, software and space” has the lowest average value among the topics. In general, our expansion using dominant meaning shows high performance, and it improves the results with around 38% more than the result without expansion.

Table 3 and **Figure 5** show the performance of the dominant meaning in terms of the precision of results. They display the precision of the results of the original queries and with the dominant meaning query expansion. In general, the performance of search results with the dominant meaning query expansion has precision higher than the search results without query expansion.

Apparently, the improving in the query of the blended learning concept gains (0.84) considered as a highest value precision. Obviously, the well-known terms online contains more domain knowledge than the small ones. Accordingly, it seems that the performance increases with the importance of the topic online. It is possible to return this growing to the huge number of web sites that contain books related to the subject of the blended learning. Moreover, it can be seen easily that the improvement in the five concepts are 25%, 27%, 32%, 29%, and 38% “blended learning models”, “role of the student”, “role of the teacher”, “change at the school level”, “hardware, software and space”, respectively. In addition, the average improvement for all topics and lists is around (30%) which is considerable results based on the high quality of the Google search engine and its fast-paced work environment.

4.1.2. The Impact of ITBIA on the Learning Using It through MOOCs

This section investigates ITBIA effectiveness factors influencing learner satisfaction on the learning session through MOOCs. Effectiveness investigation measures how well ITBIA achieves its objective. It is expected that effectiveness investigation is correlated with learner satisfaction. Seventy distance learning students were recruited for this test; all participants were between the ages of 16 and 21 years old with a median age of 19. There were 70 students included 42 girls and 28 boys. A simulation web site was created to mimic coursera course “Blended Learning: Personalizing Education for Students” [21]. The website was hosting on our Apache server, and the students were invited to navigate the course and answer on a questionnaire. An online survey is conducted, and 10 questions are returned, as shown in **Table 4**. The survey included 5-point LIKERT type scale, which has been anchored to measure all of the scales with the statements: “Strongly Disagree” = 5, “Disagree” = 4, “Neutral” = 3, “Agree” = 2 and “Strongly Agree” = 1.

As shown in **Figure 6**, the results indicate that there were significant in the response of the learner that ITBIA gives more courses materials which were relevant and useful. Learners were asked if they thought that ITBIA was valuable for learning session and helped students to have valuable learning experiences during the course. A

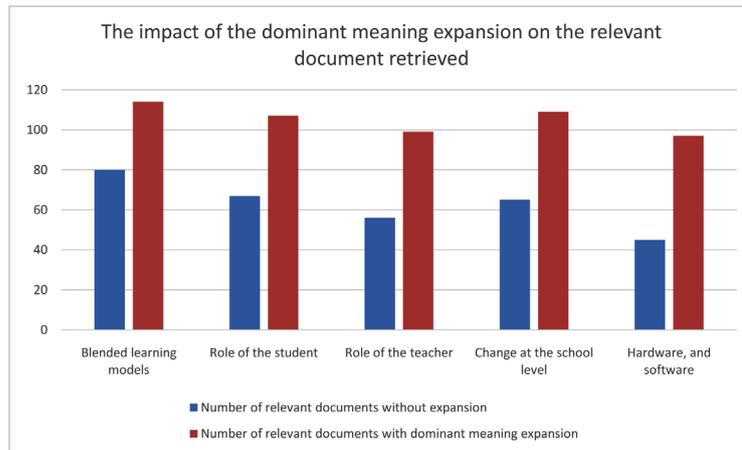


Figure 4. The impact of the dominant meaning expansion on the relevant document retrieved.

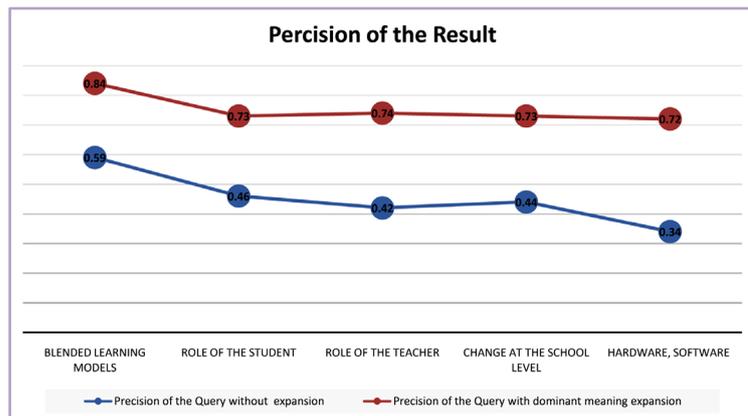


Figure 5. Percision of the results.

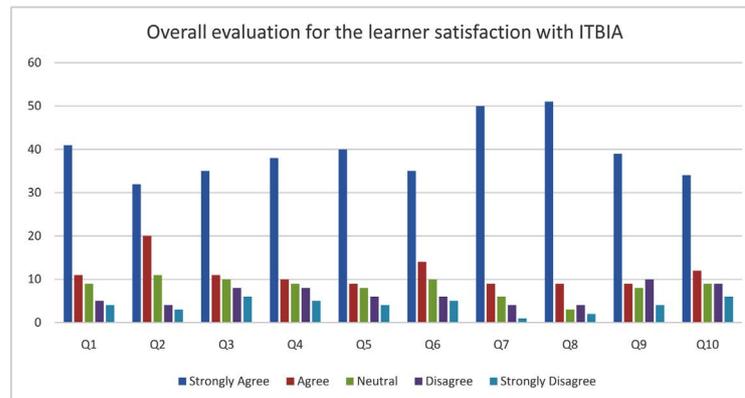


Figure 6. Overall evaluation for the learner satisfaction with ITBIA.

significant majority (84%) included who agreed and strongly agreed confirmed that it is valuable, while seven percent included who strongly disagreed and disagreed responded no.

When learners were asked if the ITBIA's web search results helped them in the learning session. There was 87% included who agreed and strongly agreed reported that it helped them, while 9% included who disagreed and strongly disagreed of learners indicated that the ITBIA could not help them. During the survey, learners

Table 3. Precision of the results.

	Query concept	Precision of the query without expansion	Precision of the query with dominant meaning expansion	Improvement in the precision
Q1	Blended learning models	0.59	0.84	0.25
Q2	Role of the student	0.46	0.73	0.27
Q3	Role of the teacher	0.42	0.74	0.32
Q4	Change at the school level	0.44	0.73	0.29
Q5	Hardware, software and space	0.34	0.72	0.38

Table 4. Criteria of evaluating learner satisfaction for ITBIA.

Question number	Criteria
Q1	I think ITBIA gives more courses materials which were relevant and useful
Q2	I think ITBIA helps me to have valuable learning experiences from my courses
Q3	I think ITBIA's web search results helped me in the learning session
Q4	The procedure of search in ITBIA was clearly presented
Q5	It is easy for me to complete the learning session by using ITBIA
Q6	The ITBIA encouraged me to become actively involved in the Coursera discussions
Q7	I think ITBIA is needed to massive open online courses
Q8	I think ITBIA's features are reliable
Q9	I think ITBIA's features are accessible
Q10	I can accomplish ITBIA tasks quickly

were asked if they thought that if it was easy for him to complete the learning session by using ITBIA. Results revealed 73% confirmed that it has flexible navigation. Where there was 15% responded no. Moreover, the respond on “the ITBIA encouraged me to become actively involved in the Coursera discussions” has an average of 70% on a scale of 1 (strongly agree) along with the scale of “agree” while there was 16% who do not agree along with strongly disagree.

Finally, to test the entire survey, we use Cronbach's alpha. The test is to be conducted on all subscales combined. The reliability coefficient for the instructor subscale is for this test is 0.91 which it indicates a good internal consistency.

5. Conclusion

In this paper, we described Intelligent Topic-Based Information Agent that autonomously creates a query, searches the web, retrieves information, resorts the results, and recommends contents to users of the massive open online courses. The research results clearly establish the potential of using agent technology to improve the performance of retrieving well-suited information from the web according to topic which has been taught. Using information agent technology in this paper provides users with an effective way of finding out required results from Google search engine. To overcome these challenges, the framework of Intelligent Topic-Based Information Agent was presented and the algorithm of the query expansion technique was presented and implemented. Two experiments were conducted on the Intelligent Topic-Based Information Agent: one measures the improvement in the retrieval effectiveness and the other measures the impact of ITBIA on the learning using it through MOOCs. The first experiment gave a test query based on the topic taught in the proposed course in Coursera; our result indicated gains in average of roughly 30% with regard to the precision test. The second experiment showed the significance of using ITBIA on the learning session of MOOCs. In general, the result showed that the proposed agent performs efficiently.

References

- [1] Chin, D.B., Dohmen, I.M. and Schwartz, D.L. (2013) Young Children Can Learn Scientific Reasoning with Teachable Agents. *IEEE Transactions on Learning Technologies*, **6**, 248-257. <http://dx.doi.org/10.1109/TLT.2013.24>
- [2] Business Dictionary (2014) Definition of Agent Technology. <http://www.businessdictionary.com/definition/agents-technology.html>
- [3] Chen, B. and Cheng, H.H. (2013) A Review of the Applications of Agent Technology in Traffic and Transportation Systems. *IEEE Transactions on Learning Technologies*, **11**, 485-497.
- [4] Massis, B.E. (2013) What's New in Libraries MOOCs and the Library. *New Library World*, **114**, 267-270. <http://dx.doi.org/10.1108/03074801311326894>
- [5] Joshi, K., Verma, A., Kandpal, A, Garg, S., Chauhan, R. and Goudar, R. (2013) Ontology Based Fuzzy Classification of Web Documents for Semantic Information Retrieval. *IEEE-2013 Sixth International Conference on Contemporary Computing (IC3)*, 8-10 August 2013, 1-5. <http://dx.doi.org/10.1109/IC3.2013.6612160>
- [6] Gregoire, P.L., Desjardins, C., Laumonier, J. and Chaib-draa, B. (2007) Urban Traffic Control Based on Learning Agents. *IEEE Intelligent Transportation Systems Conference*, Washington DC, 30 September-3 October 2007, 916-921.
- [7] Vecchiola, C., Grosso, A., Passadore, A. and Boccalatte, A. (2009) Agent Service: A Framework for Distributed Multi-Agent System Development. *International Journal of Computers and Applications*, **31**, 202-2968.
- [8] Kop, R., Fournier, H. and Mak, S.F.J. (2011) A Pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant Support on Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, **12**.
- [9] Lin, H. and Wang, L. (2008) Query Expansion for Document Retrieval by Mining Additional Query Terms. *Information and Management Sciences*, **19**, 17-30.
- [10] Ciaramita, M., Hofmann, T. and Johnson, M. (2003) Hierarchical Semantic Classification: Word Sense Disambiguation with World Knowledge. <http://cs.brown.edu/~th/papers/CiaHofJonIJCAI2003.pdf>
- [11] Akrivas, G., Wallace, M., Andreou, G., Stamou, G. and Kollias, S. (2002) Context-Sensitive Semantic Query Expansion. 2002 *IEEE International Conference on Artificial Intelligence Systems (ICAIS'02)*, Divnomorskoe, September 2002, 109-114.
- [12] Razek, M., Frasson, C. and Kaltenbach, M. (2006) Dominant Meanings towards Individualized Web Search for Learning Environments. In: Magoulas, G.D. and Chen, S.Y., Eds., *Advances in Web-Based Education: Personalized Learning Environments*, IDEA Group Publishing, 46-69.
- [13] Google (2014) Web Search API Custom Ranking. <https://developers.google.com/custom-search/docs/ranking>
- [14] Daniel, J. (2012) Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education*, **3**, 4-24.
- [15] McAuley, A., Stewart, B., Siemens, G. and Cormier, D. (2010) The MOOC Model for Digital Practice. SSHRC Application, Knowledge Synthesis for the Digital Economy.
- [16] Voss, B.D. (2013) Massive Open Online Courses (MOOCs): A Primer for University and College Board Members. Association of Governing Boards (AGB) of Universities and Colleges. http://agb.org/sites/agb.org/files/report_2013_MOOCs.pdf
- [17] Kerry, W. (2013) Academic Libraries in the Age of MOOCs. *Reference Services Review*, **41**, 576-587.
- [18] Gillet, D. (2013) Personal Learning Environments as Enablers for Connectivist MOOCs. *Proceedings of the 12th International Conference on Information Technology Based Higher Education and Training*, Antalya, 10-12 October 2013, 1-5. <http://dx.doi.org/10.1109/ITHET.2013.6671026>
- [19] Greenberg, B., Schwartz, R. and Horn, M. (2013) The Course in of Blended Learning: Personalizing Education for Students. Massive Open Online Course. <https://www.coursera.org/course/blendedlearning>
- [20] Razek, M. (2013) Towards More Efficient Image Web Search. *Intelligent Information Management*, **5**, 196-203. <http://dx.doi.org/10.4236/iim.2013.56022>
- [21] Severance, C. (2012) Teaching the World: Daphne Koller and Coursera. *IEEE Computer Society*, **45**, 8-9.
- [22] Meng, W., Yu, C. and Liu, K. (2002) Building Efficient and Effective Meta-Search Engines. *ACM Computing Surveys*, **34**, 48-89. <http://dx.doi.org/10.1145/505282.505284>

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or [Online Submission Portal](#).

