

A Chatbot for Automatic Processing of Learner Concerns in an Online Learning Platform

Mamadou BAKOUAN¹

Laboratoire de Recherche en Informatique et
Télécommunication (LARIT)
Ecole Doctorale Polytechnique de l'INP-HB
Yamoussoukro, Côte d'Ivoire

Beman Hamidja KAMAGATE²

Laboratoire de Recherche en Informatique et
Télécommunication (LARIT)
Ecole Supérieure Africaine des TIC - ESATIC
Abidjan, Côte d'Ivoire

Tiemoman KONE³

Laboratoire d'Informatique, Signaux et Télécommunications
Institut de Recherche Mathématiques IRMA/UFHB
Abidjan, Côte d'Ivoire

Souleymane OUMTANAGA⁴, Michel BABRI⁵

Laboratoire de Recherche en Informatique et
Télécommunication (LARIT)
INP-HB
Yamoussoukro, Côte d'Ivoire

Abstract—In this article, we present a chatbot model that can automatically respond to learners' concerns on an online training platform. The proposed chatbot model is based on an adaptation of the similarity of Dice to understand the concerns of learners. The first phase of this approach allows selecting the pre-established concerns that the teacher has in a knowledge base which are closest to those posed by the learner. The second phase consists of selecting among these k most appropriate concerns based on a measure of similarity built on the concept of domain keywords. The experimentation of the prototype of this chatbot makes it possible to find the adequate answers. In the case, where the question refers to a question from the teacher, the learner is asked if the question identified is the one he was referring to. If he answers in the affirmative, the instructions associated with his request are sent to him. If not, the learner's concern is sent to the human tutor. The hybridization of this chatbot with the human agent comes to enrich the initial knowledge base of the chatbot. The results obtained with the concept based on the keywords of the domain are encouraging. The learner's comprehension rate is above 50% when applying the concept of domain keywords while the measure of Dice is below 50%.

Keywords—Metadata; ontologies; semantic similarity; natural language; semantic web; chatbot

I. INTRODUCTION

Chatbot are interactive virtual characters whose mission is to provide assistance to people in high-profile environments. Previous research has shown that this technology seems to have a positive influence on learning [1]. In addition, the presence of interactive virtual agents, also called Chatbot, taking on the role of guardian [2], seems to have positive effects on student engagement [3] and on the effectiveness of teaching [4]. In the education system in Côte d'Ivoire, the number of graduates is growing steadily, without a corresponding increase in the capacity of higher education institutions [5]. To face this situation, the government has opted for the integration of new technologies (ICT) in education through the interconnection of universities and public schools in Côte d'Ivoire [6]. This project should make it

possible to unclog university lecture halls by relying on distance learning and facilitate access to teaching resources. However, since 2015 the infrastructures of the e-Education project are not operational.

In this dynamic, the State uses e-learning through the creation of Université Virtuelle de Côte d'Ivoire (UVCI) [7]. One of UVCI's missions is to develop distance education in Côte d'Ivoire. This type of teaching is based on a set of platforms to facilitate access to learning resources for learners. In the pedagogical model of the UVCI, the human tutor plays the role of framer. It ensures the educational follow-up of the training. However, the response time of the physical tutor is low and the high number of students per physical tutor degrades the quality of the training. This sometimes gives rise to the feeling of abandonment in some students.

To remedy this, we offer a chatbot that helps to take care of students' concerns on a permanent basis. It is about lightening the task of teachers and tutors while contributing to the framing and effective management of student concerns. In the next section, we will describe the role of metadata and ontologies in how chatbot work. Then we will discuss the mechanism used by the chatbot to understand the sentences. Finally, we will see the experimentation of the prototype of the chatbot and the results.

II. LITERATURE REVIEW

Information systems have to evolve with certainty, their agility is a major requirement. Software architectures must therefore promote real flexibility and reusability to adapt to change. New software architectures have brought a real ability of an architecture to evolve in order to integrate some changes response to the complex need of integration of information systems. It is particularly in this context that the new generation of formal metadata system technologies and the semantic web, derived from the Service-Oriented Architectures paradigm, aims to respond in a relevant way to the question of interoperability related to the agility of chatbot systems.

A. Semantic Web Technology

The term semantic Web, ascribed to Tim Berners-Lee [8] in the W3C, first refers to the vision of the Web of tomorrow as a vast space of exchange of resources between human beings and machines allowing exploitation, qualitatively superior, large volumes of information and varied services. Virtual space, it should see, unlike the one we know today, the users discharged of a good part of their tasks of research, construction and combination of the results, thanks to the increased capacities of the machines to access the resources and to reason with them. The semantic web is structured in layers. These layers correspond to different categories of formalisms grouped into three levels. This is the naming / addressing level, the syntactic level and the semantic level. The semantic web respects an architecture (see Fig. 1). This figure represents the structure of semantic web components.

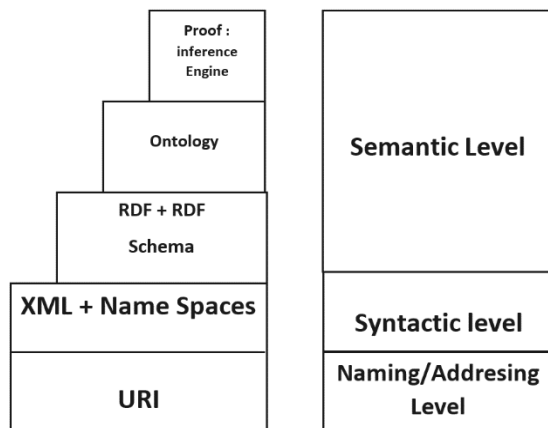


Fig. 1. Semantic web architecture [8].

Most of the languages standardized by W3C as part of the Semantic Web are XML dialects, such as RDF and RDFS. The RDFS provides basic elements for defining ontologies or vocabularies for structuring RDF resources. SPARQL is a query language for RDF. Like SQL for relational databases or Xpath and XQuery for XML documents, this language is used to retrieve information from RDF documents. The construction of ontologies and metadata requires consensus in order to avoid lexical ambiguities due to hyperonymies and polysemias. A metadata is literally a datum on a datum. That is, a structured set of information describing any resource [9]. The principle of metadata is about association a number of fields with resources for which values are assigned to each. These values can be given in a free format, as they can also conform to well-defined data formats. The operation consists of considering tags that are introduced in the files or in the appropriate programming languages. Tags have the effect of improving the efficiency of information searches compared to full-text searches. It is important to note that tagged digital resources carry with them their own metadata and those, when downloaded, copied, replicated, transmitted by email. This approach promotes interoperability for better exploitability of digital resources. Several standardization organizations have proposed and published metadata schemas that could be used by as many people as possible. We will analyze the different metadata schemas in the next section.

- LOM (Learning Object Metadata) [10]
- EAD (Encoded Archival Description) [11]
- Dublin Core [11]

The concept of metadata requires the definition of a kernel of standard and context-dependent information. This can make it difficult to exploit metadata in a learning model. To optimize this concept, the metadata schema used in learning models is enhanced by technology. Indeed, metadata is associated with domains of knowledge that can be conceptualized in ontologies.

Ontologies represent a source of very reliable and structured knowledge. For this reason and thanks to the initiatives of the Semantic Web, which brought the creation of thousands of domain ontologies, ontologies have been widely exploited in knowledge-based systems, and more specifically, for the calculation of semantic similarity. An ontology is formally defined as a pair (O, Lex) where O is an abstract ontology and Lex is a lexicon for O [12]. Let L be a logical language having formal semantics in which inference rules can be expressed. An abstract ontology is a structure $O = (C, \leq_C, R, \sigma, \leq_R, I, R)$ consists of:

Two disjoint sets C and R whose elements are respectively called Concepts and Relations;

A partial order \leq_C on C , called hierarchy of concepts or taxonomy;

A function $\sigma: C \times C$ called signature;

A partial order \leq_R on R , called hierarchy of relations where $r_1 \leq_R r_2$ implies $\sigma(r_1) \leq_C \sigma(r_2)$ with $r_1, r_2 \in R$.

A set I, R of inference rules expressed in the logical language L ;

The dom function: $R \rightarrow C$ with $\text{dom}(r) = \Pi_1(\sigma(r))$ returns the domain of r ;

The range function: $R \rightarrow C$ with $\text{rank}(r) = \Pi_2(\sigma(r))$ returns the scale of values of r ;

A lexicon for an abstract ontology $O = (C, \leq_C, R, \sigma, \leq_R, I, R)$ is a structure $Lex = (S_C, S_R, Re f_C, Re f_R)$ which consists on:

Two sets S_C and S_R whose elements are called signs, respectively for concepts and relations;

- Two relations $Re f_C \subseteq S_C \times C$ and $Re f_R \subseteq S_R \times R$, called assigning lexical references respectively for concepts and relationships;

From $Re f_C$ we define $\forall s \in S_C, Re f_C(s) = c \in C | (s, c) \in Re f_C$ and $Re f_C^{-1}(s) = s \in C | (s, c) \in Re f_C$

- From $Re f_R$ we define $\forall s \in S_R, Re f_R(s) = r \in R | (s, r) \in Re f_R$ and $Re f_R^{-1}(s) = s \in R | (s, r) \in Re f_R$

There are ontologies in different fields that support the design of learning systems including DogOnt ontology, SOUPA, CoBrA, CoDAMoS, etc. [13], [14].

- SOUPA (Standard Ontology for Ubiquitous and Pervasive Applications).
- DogOnt (Ontology Modeling for Intelligent Domestic Environments).
- CoDAMoS (Context-Driven Adaptation of Mobile Services).
- CoBrA (cobra-have overview).

In the literature several languages have been used for the description of ontologies. These languages include the eXtensible Markup Language [13], the Resource Description Framework (RDF) [15], the DAML + OIL (Darpa Modeling Language of Ontology + Ontology Inference Layer) [16] and OWL (Ontology Web Language) [14]. These languages offer different levels of expressiveness. Making yourself available to answer questions about distance learning activities related to a training module followed in a teaching platform are non-obvious tasks especially if the number of learners is important. Hence our idea, to integrate a chatbot whose role is collaboration and cooperation with the human tutor.

B. ChatBot

A Chatbot is a computer program capable of simulating a conversation with one or more users by voice or text exchange. Indeed, he plays the role of an assistant who aims to answer the questions put to him, while imitating human behavior [17]. The operating principle of a virtual guardian agent goes through three stages:

- The learner first sends questions that he would like to address to the agent.
- The agent receives the learner's question.
- He analyzes the question by consulting his knowledge base and finally provides an answer to the questions asked by the learner.

We could classify chatbot into two main categories:

- Virtual recommendation agents: This agent makes proposals to users in a virtual environment [18].
- Feedbacks chatbot: This agent makes feedbacks after performing an activity in a virtual environment [19].

Sassi researchers [18] propose a virtual recommendation agent that assists a user in his daily tasks, without any explicit request from the user. This agent aims to assist the user in his daily tasks thanks to his ability to perceive the state of the environment and to interact effectively according to the needs of the user.

Joanna's work [19] focused on the chatbot of Feedbacks. They provide a chatbot that can provide feedback to users after performing an activity in a virtual environment. Chatbot feedback and interpretation of user feedback is based on knowledge of the virtual environment. After analyzing the different works, we found that the proposed chatbot do not take into account the online learning environment. In the next section, we present some approaches for comparing texts. We will speak later of similarity between texts. The presented

approaches have been selected to best respond to the context. Thus, this document does not claim to give an exhaustive list of all the existing methods but tries to give an overview of the most used methods in the context of our study. In the next section, we will describe these different notions of similarity measure in sentences.

C. Similarity Measures Between Sentences

In automatic language processing, similarity measurement plays an important role and is one of the fundamental tasks. The automatic understanding of a sentence requires from the web agent different types of abilities: recognizing words and associating them with lexical information (morphological analysis); structure the sentence with a grammar (parsing), understand the sentence with semantic rules (semantic analysis) and take into account the context (pragmatic analysis). Huang [20] has shown that the performances of syntactic similarity based on the Jaccard index and the Dice index are very close and that they are significantly better than those of the Euclidean distance and the Levenshtein distance. The distance from Levenshtein is widely used in linguistics and bioinformatics as well as for the recognition of text blocks. Unfortunately, the computation time (complexity) is when applied to two sequences of approximately the same size. This is an obstacle in many practical applications.

In Christine's work [21], she proposes a method for measuring the semantic similarity between strings of characters. This method is based on the combination of Levenshtein's distance and Jaccard's index. This method has shortcomings when the strings correspond to names composed of several words. In addition, it requires a perfect match between each string in the two sets of strings. Thus, Hai-Hieu Vu and Jeanne Villaneau [22] proposed another method for measuring the semantic similarity between sentences that uses Wikipedia as the only linguistic resource. This method is based on a vector representation; it uses a random indexing to reduce the size of the manipulated spaces. Hai's method does not return a precise answer to the user. It returns to the user a Wikipedia article containing the elements of answer to his concern. The user is led to analyze this article in order to find an answer to his concern. Goutam Majumder and Partha Pakray [23] propose a method for calculating the semantic similarity between sentences based on the WordNet taxonomy. It allows to index, classify and put in relation the semantic and lexical contents of the English language. This method is not adapted to our context.

The similarity methods proposed in the research works are based on the TF-IDF method. TF-IDF (term frequency-inverse document frequency) is a weighting method used for finding information in the corpus. The TF-IDF method requires preprocessing of the corpus to determine the discriminating power of each word. While this pretreatment uses significant resources and lengthens the query processing time. The proposed chatbot model is an adaptation of the Dice measure based on the concept of domain keywords to understand the concerns of learners. The hybridization of the chatbot with the human agent comes to enricher the initial knowledge base of the chatbot.

III. MECHANISM USED BY THE CHATBOT TO UNDERSTAND SENTENCES

We propose a measure adaptation of Dice to calculate the similarity between sentences. This approach is based on the Dice index and the measure of similarity of the keywords of the domain. We will discuss the principle of the algorithm and the process of calculating the similarity between sentences.

A. Principle of the Algorithm

- The learner sends a question to the chatbot.
- The chatbot receives the learner's question.
- The chatbot analyzes the learner's question.
- Cleanup (Remove StopWord).
- Lemmatization (Convection of words in lemma).
- Selection of k questions (Comparison of words in common and select questions closest to the learner's question).
- Similarity based on domain words (Search among selected questions, one that is semantically close to the learner's question).
- Proposition of the question semantically close to the learner.
- The learner should confirm that the proposal corresponds to his / her concern or not.
- If the learner answers with "NO", his question is returned to a human agent.
- If the learner answers with "YES", the chatbot provides the answer to the learner's question.

B. Calculation of the Similarity between Sentences

The calculation of the similarity between sentences has been implemented by performing the following steps:

Phrase Labeling: This step deals with all of the sentences in the corpus (see Fig. 2) and converts each of their terms into lemmas. Lemmatization consists of finding the root of the bent verbs and bringing the plural and / or feminine words back to the singular masculine form (see Fig. 3).

Selection of k questions: A measure of similarity to select the k questions closest to the learners' preoccupation (see Fig. 4). This similarity approach is based on the measure of Dice. The measure of Dice calculates the similarity between two sentences Q_E and Q_S based on the number of terms common to Q_E and Q_S (see Fig. 4).

$$sim_{dice}(Q_E, Q_S) = \frac{2N_C}{S_E + S_S} \quad (1)$$

Q_E represents all the terms of the student's question.

S_E represents the number of terms after the lemmatization of the student's question.

Q_S represents all the terms of the teacher's question.

S_S is the number of terms after the lemmatization of the teacher question.

N_C is the number of terms common to Q_E et Q_S

Practical case of similarity of the Dice index between Q_E and Q_S :

Q_S "Example of question proposed by the teacher": Why I cannot read other students' posts in the forum?

Q_E "Example of student question": Unable to read messages from my fellow students in the forum.

Step 1: Cleaning the stopwords

Step 2: Converting the terms to lemma

Step 3: Analysis of terms common to Q_E and Q_S

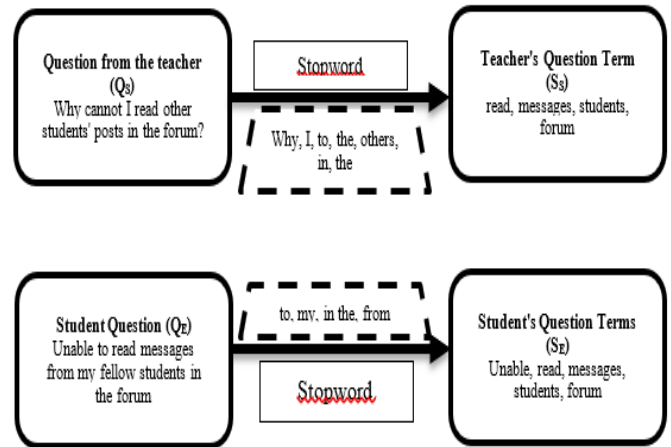


Fig. 2. Stopword cleaning process.

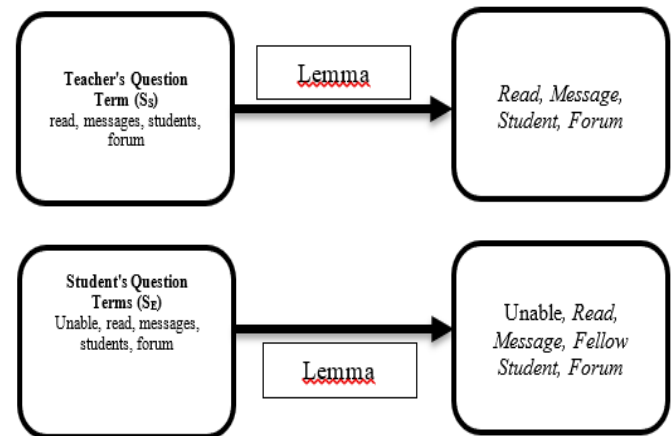


Fig. 3. Term lemmatization process.

The analysis of the terms common to Q_E and Q_S makes it possible to retain the k Q_S questions close to the Q_E questions. Then, a method of similarity based on the keywords of the domain allows to retain the Q_S question closest to the Q_E question (see Fig. 5).

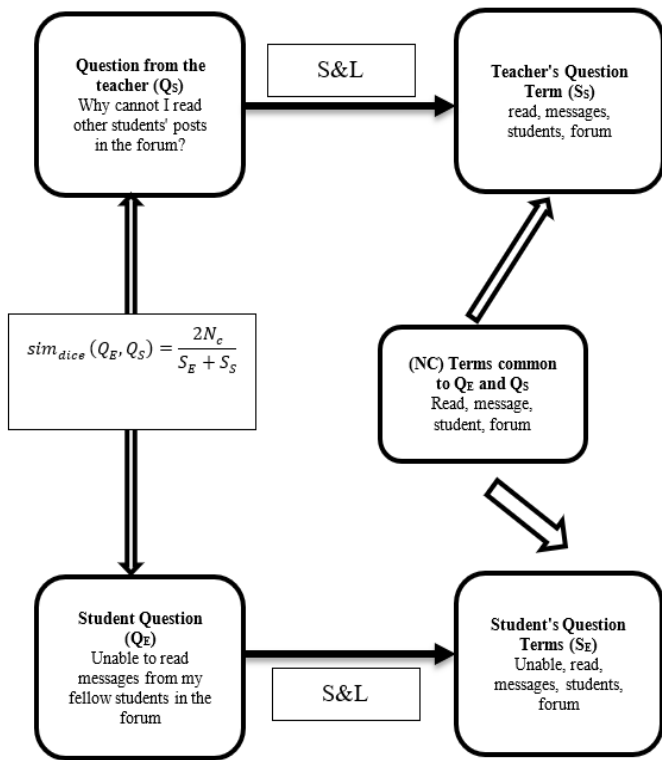


Fig. 4. Common terms analysis process.

Similarity measure built on the concept of domain keywords: this method consists in finding the Q_S question closest to the Q_E question by integrating the principle of the keywords of the domain.

MC represents all the keywords of the teacher's course.

Q_E represents all the terms of the student's question.

S_E represents the number of terms after the lemmatization of the student's question that belong to faith in Q_E and MC.

Q_S represents all the terms of the teacher's sth question with $\forall s \in \{1 \dots k\}$.

S_S represents the number of terms after the lemmatization of the teacher sth question that belong to the faith in Q_S and MC. T_S : represents all terms in common to S_S and S_E

$$S_E = \{t_i \in Q_E / t_i \in MC\} \quad (2)$$

$$S_E = Q_E \cap MC \quad (3)$$

$$S_S = \{t_i \in Q_S / t_i \in MC\} \quad (4)$$

$$S_S = Q_S \cap MC \quad (5)$$

$$T_S = |S_E \cap S_S| \quad (6)$$

$\forall s \in \{1 \dots k\} Q_E \cong Q_S$ if T_S is the maximum

Practical case of similarity between Q_E and Q_S :

Domain Keyword (s): Course, Email, Feedback, Duty, Grade, Medium, Test, Tool, Communication, Forum, Message, Discussion, Student, Publication

Question from the teacher:

Q1: Why cannot I read messages from other students in the forum?

Answer: In a Question & Answer forum, you must first contribute to the forum by submitting a contribution before having access to the messages of other students.

Q2: How can I keep up with the news of the forum (s) to which I subscribe?

Answer: In the "My classes" workspace, a message informs you of additions to the forum (s).

Q3: How to visualize all my publications in the forums?

Answer: In the tab "My page - My profile - Forum posts", you can view all contributions to forums, discussions launched, or answers given.

Question of the student:

Q1: Cannot read the messages of my fellow students in the forum

Q2: I cannot read messages from students in my group

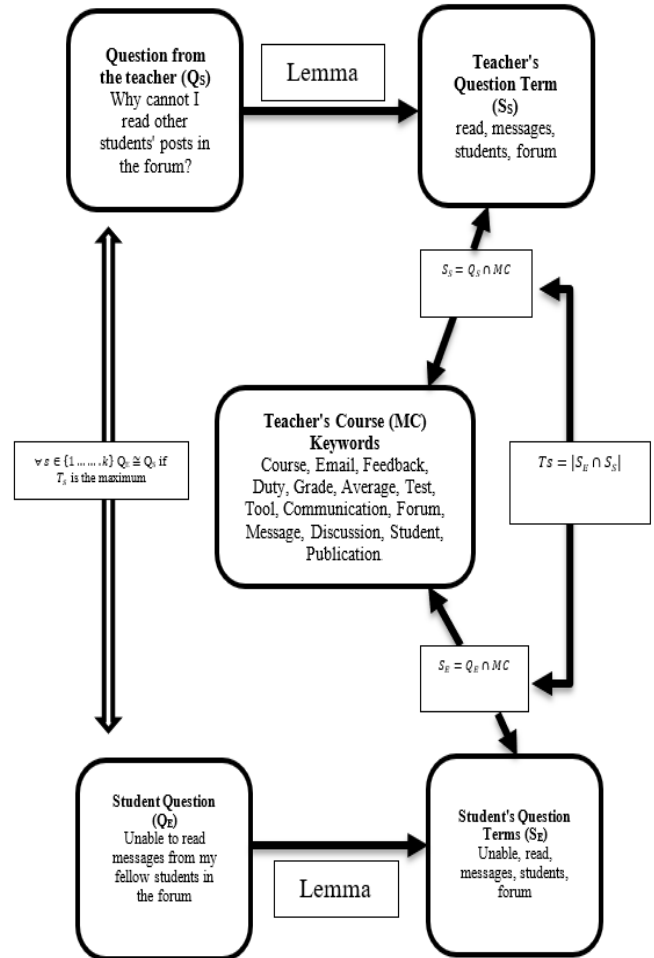


Fig. 5. Similarity measure based on domain keywords.

IV. GLOBAL OPERATION OF THE CHATBOT

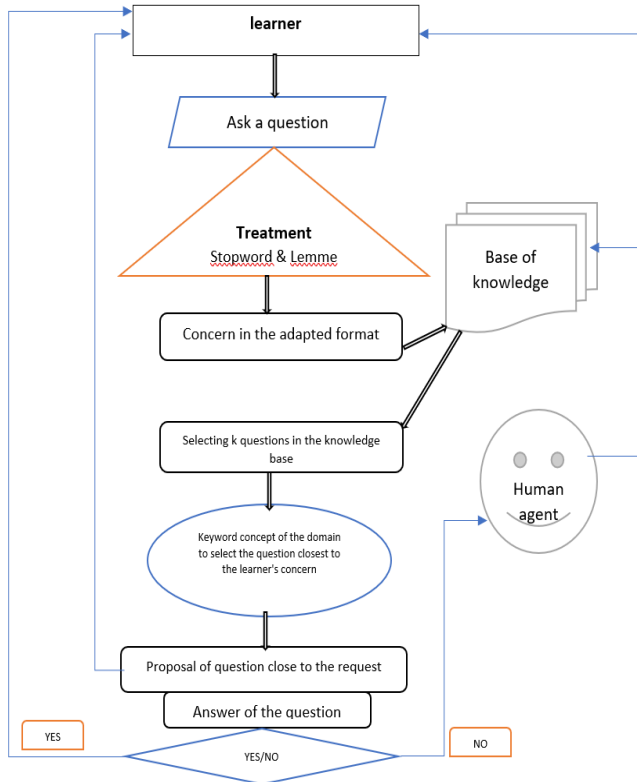


Fig. 6. Principle of function of the Chatbot.

Learner: The learner module allows to submit the concerns.

Ask a question: The module asks a question, it is the raw text of the learner's question.

Treatment: The treatment module takes into account the cleaning of the stopword and the lemmatization of the terms of the learner's preoccupation.

Lemmatization refers to the lexical analysis of the content of a text grouping the words of the same family. Each of the words of a content is thus reduced to an entity called lemma (canonical form).

Stopword: A stopword is a non-significant word in a text. It is opposed to a full word. The meaning of a word is evaluated from its distribution (in the statistical sense) in a collection of texts.

Concern in the adapted format: The learner's concern is converted into a format that allows the chatbot to understand it.

Knowledge base: The knowledge base brings together knowledge specific to the field of Université Virtuelle de Côte d'Ivoire, in a form usable by the chatbot. It contains rules that allow structuring of the data.

Selection of k questions in the knowledge base: This module allows you to browse the knowledge base in search of the teacher's questions that are close to the learner's preoccupation. We retain k questions that have high score and close to the learner's concern.

Keyword concept of the domain to select the question closest to the learner's preoccupation: Once the k closest questions are selected, we apply a concept based on the domain keyword principle. This approach selects the question of the teacher closest to the learner's concern.

Proposition of the question closest to the request: This module makes it possible to propose the question of the teacher to the learner. The learner is amenable to rewrite, if he answers by YES then the answer associated with the question of the teacher is returned to the learner. If the learner answers by NO then his concern is sent to a human agent for treatment.

Human Agent: When the chatbot does not have the answer to the learner's concern, the learner's concern is sent to the human agent who analyzes it and returns the appropriate answers. The responses of the human agent enrich the knowledge base.

Fig. 6 shows the overall operation of the chatbot and hybridization in the human agent to enrich the knowledge base.

V. EXPERIMENTATION

The experiment concerns the global operation of the prototype of the chatbot. The learner is connected to his workspace (Fig. 7) and he submits his concern to the chatbot. When he clicks the Enter key or the Submit button in the window, his concern is then converted into a language query (Fig. 8). Treatments are successively carried out as the suppression of stopwords then a lemmatization of the remaining terms.

As a result of these treatments, the query obtained is analyzed to obtain questions from the teacher close to the learner's question. Then, the treatment carried out makes it possible to find the question of the teacher closest to the question of the student. Once a question is selected, it is sent to the view to be returned to the learner (Fig. 9). Then the answer to this question will be analyzed and will return the instructions according to the following answer:

Yes: the appropriate instructions will be sent (Fig. 9)

No: the learner's concern is sent to a human tutor for analysis (Fig. 10)

The experiment is performed with the following hardware and tools: It is a Corei7 processor computer, 12GB RAM and 1TB hard drive, the object-oriented PHP programming language and a Database Management System MYSQL.



Fig. 7. Window allowing the learner to submit his concern to the chatbot.

Fig. 7 represents the window of dialogue with the learner. First a message of good is given then the student in the field seizes in the order to submit his concern to ChatBot. Finally, the learner clicks the submit button to validate his concern. As shown in Fig. 8, once the learner submits his preoccupation. This triggers the process of processing the quest. After treatment, the chatbot offers a response element to the learner. The learner has the opportunity to confirm the proposal of the chatbot. An answer item is returned to the learner based on the confirmation. If the learner answers by YES, he receives the answer adapted to his concern (Fig. 9) and if he answers by NO, the concern is sent to the human agent to have the adapted answer (Fig. 10).

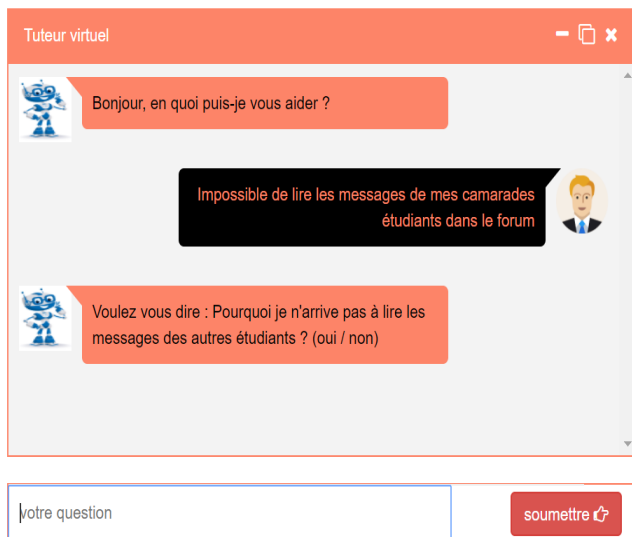


Fig. 8. Suggested question after the analysis of the learner's concern.

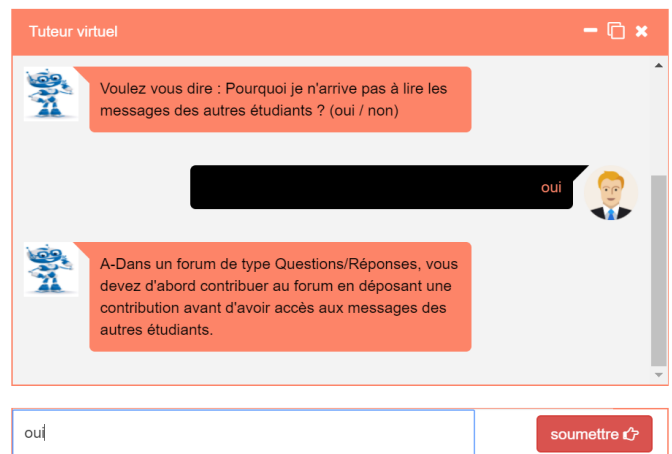


Fig. 9. The learner confirms the question proposal.

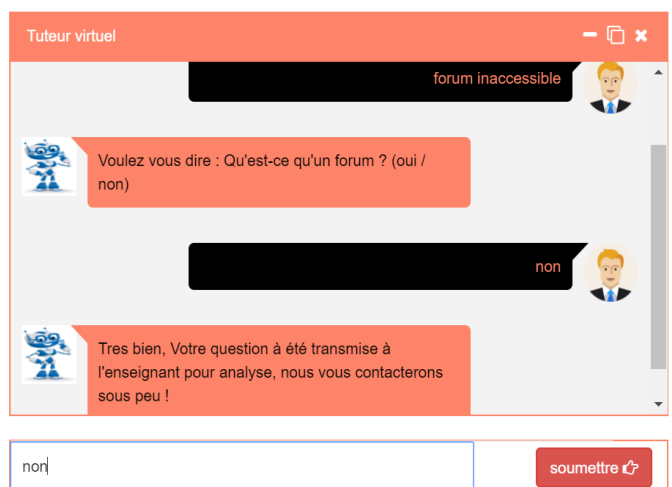


Fig. 10. The proposed question does not match the learner's concern.

The figures show the process used to respond to the learner. In the case where the question refers to a question from the teacher, the learner is asked if the question identified is the one he was referring to. If he answers yes, the instructions associated with his request are sent to him. If not, the learner's concern is sent to the human tutor.

Experimenting with the prototype of the chatbot makes it possible to find adequate answers to queries posted by the learner by applying our semantic similarity method.

A. Evaluating the Performance of the Chatbot Prototype

This assessment focuses on the learners' level of understanding of the learner's concerns. The tests are performed by applying the Dice Index method and the domain keyword-based concept (Dice Improvement) to the learner's concerns. The different tests are carried out with a series of concerns of the learners. It is a question of calculating the rate of comprehension of the concerns of the learners by the chatbot. The rate of comprehension of the questions based on the Dice index represents the ratio of the number of terms of the learner understood by the chatbot on the terms of the learner's question.

TE : The terms of the learner's question

TEC : Learner's terms understood

TC : Rate of understanding.

$$TC = \frac{TEC}{TE} \quad (7)$$

Calculation process of the understanding rate with the measure of Dice

TEC : 2 ; TE : 5 ; TC = 2/5

TC = 40%

The rate of understanding of the questions based on the concept of domain keywords (An improvement of the measure of Dice) represents the ratio of the number of terms of the teacher understood by the chatbot on the terms of the question of the teacher.

TS: The terms of the teacher's question that corresponds to TE

TSC: The terms of the teacher understood

TC: Rate of understanding.

$$TC = \frac{TSC}{TS} \quad (8)$$

Calculation process of the understanding rate with the concept of domain keywords

TSC: 3 ; TS : 7 ; TC = 3/7

TC = 43%

The table below represents the comprehension rate according to the number of questions asked by the learner (Table I).

NQ: The number of questions

MT: The method based on the index of Dice and concept of keywords of the domain

DICE: The measure of Dice

CMC: The concept based on the words of the domain

TC-DICE: Rate of understanding of the questions by applying the index of dice followed by the variation of the number of questions

TC-CMC: Rate of comprehension of the questions by applying the concept of words of the domain followed by the variation of the number of questions

TABLE I. RATE OF UNDERSTANDING OF THE QUESTIONS BASED ON THE DICE INDEX AND THE CONCEPT OF KEY WORDS IN THE FIELD

N Q	10		20		30		40		50	
	DIC E	CM C	DIC E	CM C	DIC E	CM C	DIC E	CM C	DIC E	CM C
T C	40 %	55 %	35 %	50 %	34 %	48 %	30 %	45 %	30 %	40 %

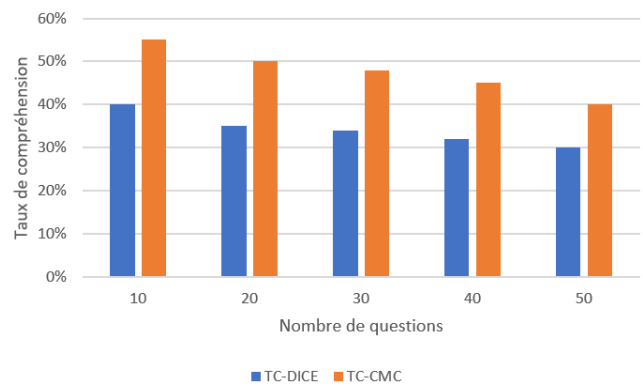


Fig. 11. Representation of the understanding rate of the questions based on the measure of Dice and the concept of keywords of the domain.

The graph above represents the rate of understanding of the questions based on the index of Dice and that based on concept of keywords of the domain (an improvement of the measure of Dice) (Fig. 11).

B. Results

Fig. 11 shows that the learner's comprehension rate is above 50% when applying the Dice Index method. In addition, the rate of understanding of the questions is weaker and weaker as the number of questions increases while the student's comprehension rate is above 50% when applying the concept of domain keywords. This concept is an improvement of the Dice Index. The results obtained with the concept based on the keywords of the domain are encouraging.

VI. CONCLUSION

In this paper, we presented work based on similarity measures to provide a chatbot with the ability to provide adequate responses in a learning interaction with learners. We have shown the steps to implement the prototype of the proposed chatbot that is an adaptation of the Dice Index. We also described the overall operation of the chatbot and the process used to address the learner's concern. The rest of the work consists in integrating the chatbot into the teaching of Université Virtuelle de Côte d'Ivoire and finish with the process of evaluation of learner's satisfaction.

REFERENCES

- [1] W. Johnson, P. Rizzo, W. Bosma, S. Kole, M. Ghijsen, and H. van Welbergen. Generating socially appropriate tutorial dialog. In ISCA Workshop on Affective Dialogue Systems, pages 254-264. Berlin, Heidelberg, 2004.
- [2] M. Limniou, D. Roberts, and N. Papadopoulos. Full immersive virtual environment CAVE [TM] in chemistry education. Computers & Education, 51 (2): 584-593, 2008.
- [3] A. Kokane, H. Singhal, S. Mukherjee, and G. Reddy. Effective elearning using 3D virtual tutors and webRTC based multimedia chat. In International Conference on Recent Trends in Information Technology (ICRTIT), pages 1-6, 2014.
- [4] J. Rowe, S. McQuiggan, B. Mott, and J. Lester. Motivation in narrative-centered learning environments. Proceedings of the workshop on narrative learning environments, AIED, pages 40-49, 2007.
- [5] DECO 2017, Directorate of Examinations and Concours (DECO) - MINISTRY OF NATIONAL EDUCATION OF CÔTE D'IVOIRE - The Territorial Assembly then created on that date, the twelve (12) first departments. The decree n ° 30004 / CAB1 of May 25, 1957 fixing the

- attributions of the Ministry of National Education, creates at the same time the DECO - December 2017
- [6] EGOUV 2015 Official Portal of the Government of Ivory Coast - 2nd DAYS OF RITER: UNIVERSITY IMPORTS NEW DIGITAL SOLUTIONS FOR HIGHER EDUCATION
- [7] UVCI 2015, The Virtual University of Côte d'Ivoire abbreviated UVCI is a public university whose "decree n ° 2015-775 of December 9, 2015 establishing, attributions, organization and functioning of an administrative public institution"
- [8] Tim 2001, The term semantic web was proposed by Tim Berners Lee in 2001 ("The Semantic Web," Scientific American Magazine, May 17, 2001) an evolution of the web that would allow the available data to be usable and interpretable by software agents.
- [9] Mohamed Tahar Ileh, Imad Saleh. Virtual tutoring approach in an e-learning platform based on an SMDF multi-agent Virtual Tutor. Tahar 2009 Formal Metas Data System SMDF, Semantic Web, Foccus, Model, Remote Tutoring, Moodle and e-learning. Volume 2, Number 1, Pages 3-18
- [10] HAYNES and David. Metadata for information management and retrieval. London: Facet, 2004, pages 186.
- [11] Mirna El-Hajj Barbar, Saleh Imad, Barbar Kablan and Youssef Moncef, "An Automata Model for the Educational Evaluation Process: Generating an Automata Markup Language (AML)", Proceedings of the H2PTM'03 Conference, 24-26 September 2003, Hermès publishing, 9 Pages.
- [12] Guizzardi, G., Almeida Falbo, R., and Pereira Filho, J. G. (2001). From domain ontologies to Object Oriented Frameworks. In G. Stumme, A. Maedche, & S. Staab (Eds.), Proceedings of the ONTO 2001 Workshop on Ontologies (pp. 1-14). (THIS Workshop Proceedings, Vol 48). Germany: CEUR.
- [13] Smith, M.K., Welty, C. and McGuinness, D.L. (2004). OWL: Ontology Web Language Guide. Technical Report, W3C: World Wide Web Consortium.
- [14] Marsh, J. (2001). XML base. Technical Report, W3C: World Wide Web Consortium.
- [15] Klyne, G. and Carroll, J. J. (2004). Resource Description Framework (RDF): Concepts and Abstract Syntax. Technical Report, W3C: World Wide Web Consortium.
- [16] Connolly, D., Harmelen, F. V., Horrocks, I., McGuinness, D.L., Patel-Schneider, P.F. and Stein, L.A. (2001). DAML + OIL: Reference Description. W3C Technical Report: World Wide Web Consortium.
- [17] Messous 2014, a virtual teaching agent - an interactive environment for human learning (EIAH) featuring an educational hypermedia and an animated conversational agent (ACA) operating in a web environment. Pages 16
- [18] Sassi 2014, Assistance and advice to users in the context of ambient intelligence: a study of uses in context: X-CAMPUS eXtensible Conversational Agent for Multichannel Proactive Ubiquitous Services. Hajer Sassi, Lille 1, 2014
- [19] Joanna Taoum, Bilal Nakhil, Elisabetta Bevacqua, Ronan Querrec. A design proposal for interactive virtual tutors in an informed environment. 11th Days of the French Association of Virtual Reality (AFRV 2016), Oct 2016, Brest, France.
- [20] Huang H, et al. (2008) The crystal structure and identification of NQM1 / YGR043C, a transaldolase from *Saccharomyces cerevisiae*. *Proteins* 73 (4): 1076-81
- [21] Christine Largeron, Bernard Kaddour, Maria Fernandez. SoftJaccard: a measure of similarity between sets of strings for the unification of named entities. *Extraction and Knowledge Management (EGC 2009)*, Jan 2009, Strasbourg, France. Cépaduès-Éditions, RNTI-E-15, pp.443-444, 2009.
- [22] Measure the similarity between sentences thanks to Wikipedia using a random indexing Hai-Hieu Vu, Jeanne Villaneau, Farida Said, Pierre-François Marteau 2015
- [23] Goutam Majumder, Partha Pakray, Alexander Gelbukh, David Pinto; Semantic Textual Similar Methods, Tools, and Applications: A Survey Goutan 2016