World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

A study of methodologies for assessing semantic similarity between two ontologies

Aly Ngoné NGOM [1, 5, *], Mamadou DIOP [2], Ibrahima GAYE [3], Fatoumata WILANE [3] and Guidedi KALADZAVI [4]

[1] *Renewable Energy & Energy Efficiency Laboratory (labEREE), International Institute of Water and Environmental Engineering (2iE), Science Street, Ouagadougou 01 BP 594, BURKINA FASO.*
[2] *Eco-Materials and Sustainable Habitats Laboratory (LEMHaD), International Institute of Water and Environmental Engineering (2iE), Science Street, Ouagadougou 01 BP 594, BURKINA FASO.*
[3] *ICT for Development Research Team (TIC4Dev), Alioune DIOP University of Bambey, BP: 30 Bambey, SENEGAL*
[4] *Computer Science Research Laboratory, University of Maroua, CAMEROON.*
[5] *Numerical analysis and Computer Science Laboratory (LANI), Gaston Berger University, BP 234 Saint - Louis, SENEGAL*

World Journal of Advanced Research and Reviews, 2025, 25(03), 721-731

Publication history: Received on 28 January 2025; revised on 04 March 2025; accepted on 06 March 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.25.3.0693

Abstract

This paper aims at presenting a methodology to assess similarity between two ontologies. Many works on similarities between ontologies have been done. That is not the case between ontologies. This approach based on set theory, edges and feature similarity consists of two parts. Firstly, we determine set of concepts shared by two ontologies and Their differences before assessing similarity between the ontologies. Secondly the methodological part will be an extension of the first one. In this part, the method of assessing similarity values between two ontologies is extended by considering the "is-a" relation between concepts two ontologies.

Keywords: Ontologies similarity; Concepts similarity; Set theory; Semantic similarity

1. Introduction

This paper is based on work originally presented in 4th Annual Conf. on Computational Science & Computational Intelligence (CSCI'17) [1]. We use as the basis of our work [2]. Ontologies allow formalizing knowledge related to the description of the world for access and sharing purposes across the Web. They introduced the semantic layer into the architecture of on based-systems [3]. With the advent of the Semantic Web, research led to the modelling of various ontologies, sometimes for the same domain. However, all these ontologies are sometimes heterogeneous (different terms for the same concept, different relations for the same association, different languages, etc.) and this faces the integration problem. Indeed, we distinguish several tasks that imply collaborative use of various ontologies. When several ontologies are used for an application, it becomes necessary they present some similarity. The assessing of similarity between ontologies may be very interesting. Indeed, it can facilitate the choice of ontologies in the case of elaborating a system, which uses them. In addition, it can help to evaluate the ontology evolution by comparing its different versions. This paper aims at presenting a method for assessing similarity between two ontologies. The approach is based on the set theory, edges-based semantic similarity [4] and feature-based similarity [5]. The last part of the analysis is schemed as follows. Section 2, entitled Conceptual Framework, presents the definitions of some core elements. Section 3 reviews some existing methods devoted to the evaluation of similarity between ontologies. Then, the Section 4 depicts our Methodology. Section 5 is devoted to some experiments to evaluate and validate this methodology. In section 6, we analyse results we get from the experimentation. The paper ends with a conclusion and the following work in Section 7.

* Corresponding author: Aly Ngoné NGOM

2. Conceptual framework

We present in this section the definition of some core concepts, which can help for a better understanding of the paper.

2.1. Ontology

The foundational definition of ontology as "an explicit specification of a conceptualization" has been proposed by Gruber [6,7]. The exact meaning depends on the understanding of the terms "specification" and "conceptualization". According to Genesereth and Nelson [8], conceptualization is a "set of objects, concepts, and other entities that are presumed to exist in some areas of interest and the relationships that hold them". In the Gruber's definition does not show that if specification depends on the logical view of ontologies. That is why Guarino and Giaretta introduced the logical theory instead of mere specification. Afterward, Borst [9] enriches the previous definition by considering consensual fact related to knowledge modeling discipline characteristics, such as sharing and reuse. For him, "Ontologies are defined as a formal specification of a shared conceptualization". Finally, Studer et al. [10] merge the existing definitions. For them, "An ontology is a formal, explicit specification of a shared conceptualization". They put emphasis on the necessity of formal, explicit and shared paradigms. Even if, it's the merging of the existing definitions, it seems consensual. It is more cited in recent years, demonstrating its compliance with the expectations of the knowledge-based systems designers [11]. In addition, the explicitness, formality and share-ability of knowledge features in an ontology are carried out by five elements [12]: concepts, relations, functions, axioms, and instances. Refering to [13], we define an ontology O as:

- $O = \{C, H, R\}$
- C is a set of concepts.
- $H \subset C \times C$ is the taxonomy of the ontology. $h = (c_1, c_2) \in H$ means that c_2 is subsumed by c_1.
- $R \subset C \times C$ represents set of relations between two concepts. In this paper, we only consider the "is-a" relation.

2.2. Concept

A concept constitutes a think about something, semantically evaluable and communicable [14]. It can be abstract or concrete, elementary (electron) and composite (atom), real or fiction. In short, a concept is a notion that represents synonymous terms or terms representing the same thing in different languages. A concept could be the description of a task, a fact, a function, an action, a strategy, a process, etc. For example, in an ontology of a library, a "book" can be considered as a concept, which refers "livre" in French, "book" in English, to the term "Buch" in Deutsch, to the term "libro" in Spanish, to the term "Derewel" in Mafa (Cameroonian local language), etc. Thereby, it enables to the ontology on-based intelligent agents to reason and to inter-comprehend (semantic interoperability) on knowledge as would humans do.

2.3. Semantic Similarity

Semantic similarity measures are functions widely used in several informatics fields among which natural language processing (NLP), Bioinformatics, Information Retrieval... They allow to determine similarity between terms or concepts which have no syntaxical similarity. From an ontology point of view, [15, 16] consider that two concepts are similar if they are "geographically" close to in a conceptual hierarchy. Thus, there is semantic similarity between two concepts (for example, movie dog and comic dog) if:

- from an intentional point of view, the two concepts share a large descriptive and functional properties proportion;
- from an expressional point of view, both concepts share a large proportion of the terms they denote (for example, Dog, Toutou, Crab, etc.);
- from an extensional point of view, the two concepts share a large proportion of their instances (eg Snowy, Rantanplan, Idefix, etc.).

In [17], semantic similarity measures are classified into three groups:

- edge-based semantic similarity measures which derive on the calculation of the distance between concepts including the number of arcs which separate them;
- semantic similarity measures based on the amount of concepts shared information through the use of information theory, hence the notion of Information Content;
- hybrid measures based on the combination of the two above mentioned families or on the use of various methods.

3. Related Work

Several works are dedicated to the evaluation of similarity between two concepts in an ontology. However, there are not many works dealing with evaluation of similarity between ontologies. The following are some works about similarity between two ontologies.

Maedche and Staab [15] propose a method for comparing two ontologies. This method is based on two levels:

- the Lexical level which consists of investigating on how terms are used to convey meanings;
- the Conceptual level which consists of investigating on the conceptual relations between given terms.

The Lexical comparison allows finding concepts by assessing syntaxical similarity between concepts. It is based on Levenshtein [16] edit distance (ed) formula which allows measuring the minimum number of changes required to get one string into another, by using a dynamic programming algorithm. The Conceptual Comparison Level allows comparing the structures semantic of two ontologies. Authors use Upwards Cotopy (UC) to compare the Concept Match (CM). Then, they use the CM to determine the Relation Overlap (RO). Finally they assess the average of RO. This approach allows assessing similarity between two ontologies by using the Lexical and Conceptual Comparison Level. However, if we reverse the position of some concepts in the hierarchy, we can get the same results because this method only considers the presence of the concept in the hierarchy.

In [18], authors implement an online ontology comparison tool which can give a numeric measurement of the difference between two ontologies. The given tool is based on senses refinement (SR) algorithm. It uses concepts and senses retrieved from WordNet [19].

The algorithm that implements SR considers the subsumption relation "is-a" (hyponymy) and constructs a set of concepts for each ontology (the source ontology and the target ontology). Each set contains concepts of ontology and synsets of concepts. A synset is a set of synonymous concepts. Since a concept can have several meanings in WordNet (polysemy), then the algorithm chooses concepts of the synset that is related to the same semantics as the studied concept. Once sets of concepts have been formed for each ontology, the ontologies are compared, by assessing their difference. The difference value is obtained by applying the Tversky measure [5].

The [18] method allows comparing two ontologies on the basis of their difference. This method uses set theory as our proposition in this paper. But the only result we get refers to the value of difference between the two ontologies. Contrary to our method which evaluates the similarity of the ontologies by taking into account their differences.

4. Methodology

4.1. Principle

The approach we propose is based on set theory, edges based semantic similarity [4] and feature-based similarity [5]. We consider ontology as a set of concepts linked together by semantic relations. The main aim of this paper is to compare two ontologies. For this, we compare sets of elements of ontologies by using feature based similarity rules. Feature-based similarity was introduced by Tversky [5]. In his work, Tversky assesses similarity between objects by taking into account their common points and their differences. Figure 1 represents Tversky's feature model.

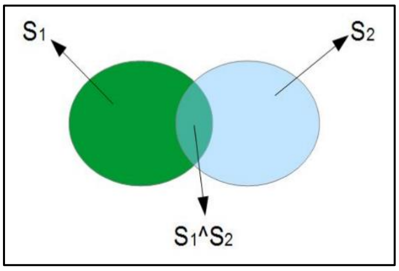

Figure 1 Example of Tversky's feature model

In this figure, we have:

- S_1 and S_2 are sets of elements;

- $(S_1 \setminus S_2)$ (respectively $(S_2 \setminus S_1)$) represents set of elements present in S_1 and not in S_2 (respectively present in S_2 and not in S_1) ;
- $(S_1 \wedge S_2)$ is the intersection between S_1 and S_2 ; i.e the common elements of sets S_1 and S_2.

The Tversky measure is given by the formula 1.

$$Tvr\ (S_1 , S_2) = \frac{f\ (S_1 \cap S_2)}{f\ (S_1 \cap S_2) + \alpha\ f\ (S_1 \setminus S_2) + \beta\ f(S_2 \setminus S_1)} \text{............(1)}$$

In the formula 1, we have:

- f represents a function that reflects the salience of a set of features;
- α and β are parameters which allow expressing the non-resemblance factors between S_1 and S_2.

In our case, we have to assess similarity of two ontologies (O_1 and O_2). By analogy with the Tversky's feature model, figure 2 gives representation of ontologies (O_1 and O_2.

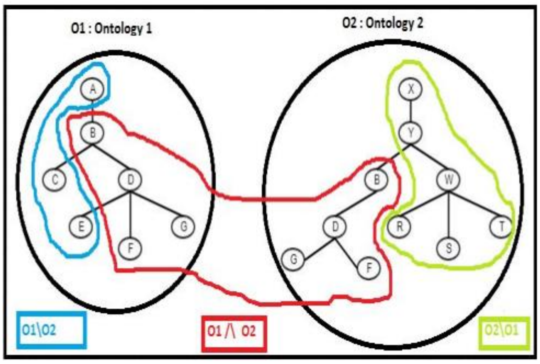

Figure 2 Representation of ontologies O_1 and O_2 with Tversky's feature model.

In figure 2, we distinguish three parts:

- $(O_1 \setminus O_2)$ = {A,C,E} : set of concepts present in O_1 and not in O_2;
- $(O_2 \setminus O_1)$ = {R,S,T,W,X,Y } : set of concepts present in O_2 and not in O_1;
- $(O_1 \wedge O_2)$ = {B,D,F,G} : set of concepts present in O_1 and O_2.

This approach can be summarized in 3 steps:

- The **Step 1** consists to determine the sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O_1 \wedge O_2)$.
- Once the sets are determined, we assess average of the semantic similarity values between concepts of each set in the **step 2**.
- Finally, in the **step 3**, we assess similarity between ontologies by using the results of the step 2 in our measure which is a redefinition of the Tversky measure.

In the following sub-section, we will deal with functions that we will use, in the steps 2 and 3, to assess similarity between the ontologies.

4.2. Measures

To assess similarity between two ontologies, we define a measure which readjusts the Tversky measure. We rely on the Tversky measure because it is a reference in the feature-based similarity case. In addition, Tversky measure inspired many works like [20] and [21]. The measure we propose takes into account shared features and differences of ontologies. Referring to figure 2, we have the following sets : $(O_1 \wedge O_2)$ = {B,D,F,G}, $(O_1 \setminus O_2)$ = {A,C,E} and $(O_2 \setminus O_1)$ = {R,S,T,W,X,Y }. Applying the Tversky measure, the similarity between O1 and O2 is given by the formula 2.

$$Tvr\ (O_1 , O_2) = \frac{f\ (O_1\ \cap\ O_2)}{f\ (O_1\ \cap\ O_2) +\ \alpha\, .\, f\ (O_1 \setminus O_2) + \beta .\, \mathrm{f}(O_2 \setminus O_1)} \qquad (2)$$

Instead of the function f, we will use one of the edge based semantic similarity measures that we studied in [4]. for every determined set, we compute the average of the similarity values between concepts. In [4], we studied edge-based semantic similarity measures. In [22] and [23], we used the measure of Zargayouna and Salotti [24] which extends the measure of Wu and Palmer [25]. The measure of Zargayouna and Salotti presents good correlation with human judgment defined by Miller and Charles [26] but the problem is this measure doesn't take into account the similarity of concepts, which are not in different hierarchy. In this paper, we use the measure of Wu and Palmer because it presents good correlation with Miller and Charles human judgment. Using Wu and Palmer similarity measure, the similarity between two concepts c_1 and c_2 is given by the formula 3.

$$Sim\ (c_1\ , c_2) = \frac{2\ \times\ depth(c_3)}{depth(c_1) +\ depth(c_2)} \tag{3}$$

The concept c3 represents the Least Common Subsumer (LCS) of concepts c_1 and c_2.

By replacing the terms of the Tversky measure with the average of the similarity values between concepts of the determined sets, formula 2 becomes formula 4.

$$T_{Ngom}\ (O_1\ , O_2) = \frac{\theta . \overline{x_{O_1}}\ +\ \omega . \overline{x_{O_2}}}{\theta . \overline{x_{O_1}}\ +\ \omega . \overline{x_{O_2}} +\ \alpha . \overline{y_{(O_1 \setminus O_2)}} + \beta\ \overline{z_{(O_1 \setminus O_2)}}} \tag{4}$$

With:

- $\theta\ = \frac{cardinality\ \ (O_1 \cap O_2)}{cardinality\ (O_1)}$
- $\omega\ = \frac{cardinality\ \ (O_1 \cap O_2)}{cardinality\ (O_2)}$
- $\alpha\ = \frac{cardinality\ (O_1 \setminus O_2)}{cardinality\ (O_1)}$
- $and\ \beta\ = \frac{cardinality\ (O_2 \setminus O_1)}{cardinality\ (O_2)}$
- cardinality(O) is the number of elements (concepts) of the set (ontology) O;

and where:

- $\overline{x_{O_1}}$ (respectively $\overline{x_{O_2}}$) is the average value of similarity between concepts (x_i,x_j) in ontology O_1 (respectively (x_i,x_j) in ontology O_2). i, j ∈ N and i ≠ j.
- $\overline{y_{(O_1 \setminus O_2)}}$ (respectively $\overline{z_{(O_1 \setminus O_2)}}$) is the average value of similarity between concepts (y_i,y_j) (respectively (z_i,z_j) present in ontology O_1 but not in O_2 (respectively present in ontology O_2 but not in O_1). i, j ∈ N and i ≠ j.
- the coefficients θ, ω, α and β allow to take into account the similarity values in relation to the number of concepts of the sets' concepts and number of concepts of ontologies.

The measure presented by formula 4 respects these properties:

- the measure is symmetric: $T_{Ngom}\ (O_1\ , O_2) =\ T_{Ngom}\ (O_2\ , O_1)$;
- the measure is bounded between 0 and 1;
- if $T_{Ngom}\ (O_1\ , O_2) = 1$ then $O_1\ =\ O_2$.

4.3. Extension with is-a relation

4.3.1. Principle

In [2], we proposed an improvement of the assessment of the similarity between ontologies. To improve the methodology, we use the "is-a" relation to extend set of concepts which is shared by ontologies. For extending the set of concepts shared by ontologies, we do a mapping between concepts of ontologies. If there is syntaxical correspondence between concepts, we proceed to the extension of the set. We can summarize the methodology in five steps:

- **Step 1** and **2** of the initial methodology will be kept.
- In **step 3**, we extend ontologies O_1 and O_2 by using the set (O_1 ∧O_2). In this step, for each concept c of the set (O_1 ∧O_2), we search their sons x_i (i ∈ N) in O_1 (respectively O_2) and we add them as sons of c in O_2 (respetively O_1) if they don't exist in this ontology. At the end of this step, we obtain two ontologies: O'_1 (respectively O'_2)

which extends O_1 (respectively O_2) with concepts of O_2 (respectively O_1). Thus, extension of ontologies allows us to determine the set of concepts $(O'_1 \wedge O'_2)$ shared by two ontologies.

- In **step 4**, we determine $(O'_1 \wedge O'_2)$ which is the set of shared concepts by ontologies O'_1 and O'_2.
- Finally, in the **step 5**, we assess similarity between ontologies by using results of the step 2 and 4 in our measure which is a redefinition of the T_{Ngom} measure [1].

In summary, for assessing similarity between ontologies, we use sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O'_1 \wedge O'_2)$; i.e. we consider the difference between O_1 and O_2 by using sets $(O_1 \setminus O_2)$ and $(O_2 \setminus O_1)$, and the resemblance between the two ontologies by using set $(O'_1 \wedge O'_2)$. Figure 3 represents the difference we use for assessing similarity between ontologies O_1 and O_2.

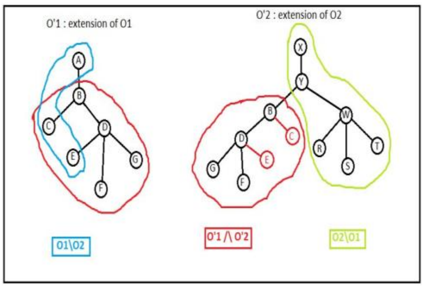

Figure 3 Representation of extensions of ontologies O'_1 and O'_2 with Tversky's feature model

In figure 3, we distinguish three parts:

- $(O_1 \setminus O_2)$ = {A,C,E} : set of concepts present in O_1 and not in O_2;
- $(O_2 \setminus O_1)$ = {R,S,T,W,X,Y } : set of concepts present in O_2 and not in O_1;
- $(O'_1 \wedge O'_2)$ = {B,C,D,E,F,G} : set of concepts present in O'_1 and O'_2.

4.3.2. Measures

The measure we present in this paper is an improvement of our measure T_{Ngom} [4] which redefines Tversky's [22] similarity measure. For assessing similarity between two ontologies O_1 and O_2, our measure takes into account the difference between the two ontologies by assessing the average similarity values of sets $(O_1 \setminus O_2)$ and $(O_2 \setminus O_1)$, and their common concepts by extending the ontologies (O_1 to O'_1 and O_2 to O'_2) and assessing the average similarity values of the set of extended ontologies common concepts of $(O'_1 \wedge O'_2)$. We use Wu and Palmer measure [25] for computing semantic similarity between concepts of sets in ontologies. The measure is given by the formula 5:

$$N_{Plus}(O_1, O_2) = \frac{(\theta.\overline{x_{O'_1}} + I_2) + (\omega.\overline{x_{O'_2}} + I_1)}{(\theta.\overline{x_{O'_1}} + I_2) + (\omega.\overline{x_{O'_2}} + I_1) + \alpha.\overline{y_{(O_1 \setminus O_2)}} + \beta\,\overline{z_{(O_1 \setminus O_2)}}} \quad (5)$$

with:

- $\theta = \frac{cardinality\ (O'_1 \cap O'_2)}{cardinality\ (O_1) + n_1 + n_2}$
- $\omega = \frac{cardinality\ (O'_1 \cap O'_2)}{cardinality\ (O_2) + n_1 + n_2}$
- $\alpha = \frac{cardinality\ (O_1 \setminus O_2)}{cardinality\ (O_1)}$
- $\beta = \frac{cardinality\ (O_2 \setminus O_1)}{cardinality\ (O_2)}$
- $I_1 = \frac{1}{1 + n_2}$
- $I_2 = \frac{1}{1 + n_1}$

- $\overline{x_{O'_1}}$ (respectively $\overline{x_{O'_2}}$) is the average value of similarity between concepts (x_i,x_j) in ontology O'_1 (respectively (x_i,x_j) in ontology O'_2). i, j ∈ N and i ≠ j.
- $\overline{y_{(O_1 \setminus O_2)}}$ (respectively $\overline{z_{(O_1 \setminus O_2)}}$) is the average value of similarity between concepts (y_i,y_j) (respectively (z_i,z_j) present in ontology O_1 but not in O_2 (respectively present in ontology O_2 but not in O_1). i, j ∈ N and i ≠ j.
- cardinality(O) is the number of elements (concepts) of the set (ontology) O;
- I_i : Integrity coefficient of Ontology O_i (i ∈ N);
- n_i : number of concepts of O_i added for extending O_j (I,j ∈ N);
- As in [4], the parameters θ, ω, α and β allow to take into account the similarity values in relation to the number of concepts of the sets' concepts and number of concepts of ontologies.

The integrity coefficient of ontology (I_i) is a value which is related to the number of concepts of ontology O_j (n_j) that we have to add to O_i for extending it (i,j ∈ N). The larger is n_j, the smaller is Ii. We have the expression 6:

$$\begin{cases} \lim_{n\to\infty} I = \lim_{n\to\infty} \frac{1}{1+n} = 0 \,; \\ \lim_{n\to 0} I = \lim_{n\to 0} \frac{1}{1+n} = 1 \,; \end{cases} \text{with } (n \in N). \quad (6)$$

We note that measure presented by formula 5 like formula 4 respects this properties:

- The measure is symmetric: $N_{Plus}(O_1, O_2) = N_{Plus}(O_2, O_1)$;
- The measure is bounded between 0 and 1;
- if $N_{Plus}(O_1, O_2) = 1$ then $O_1 = O_2$.

5. Experimentations

In this section, we experiment our methodology. We compare T_{Ngom} and N_{Plus} measures by assessing similarities between ontologies extracted from Wordnet. Table 1 lists 8 ontologies used to compare measures.

Table 1 List of ontologies

Ontologies (O)	Concepts (C)	Hierarchies (H)
O_3	instrumentality, [conveyance, transport], vehicle, wheeled_vehicle, motor, [bike, bicycle], [car, auto], truck, article, ware, [cutlery, eating_utensil], fork	h_1 =(instrumentality, [conveyance, transport]), h_2 =([conveyance, transport], vehicle), h_3 =(vehicle, wheeled_vehicle), h_4 =(wheeled_vehicle, motor), h_5 =(wheeled_vehicle, [bike, bicycle]), h_6 =(motor, [car, auto]), h_7 =(motor, truck), h_8 =(article, ware), h_9 =(ware, [cutlery, eating_utensil]), h_{10} =([cutlery, eating_utensil], fork)
O_4	instrumentality, [conveyance, transport], vehicle, wheeled_vehicle, motor, [bike, bicycle], [car, auto], truck, article, ware, [cutlery, eating_utensil], fork, gun, boat, table_knife	h_1 =(instrumentality, [conveyance, transport]), h_2 =([conveyance, transport], vehicle), h_3 =(vehicle, wheeled_vehicle), h_4 =(wheeled_vehicle, motor), h_5 =(wheeled_vehicle, [bike, bicycle]), h_6 =(motor, [car, auto]), h_7 =(motor, truck), h_8 =(article, ware), h_9 =(ware, [cutlery, eating_utensil]), h_{10} =([cutlery, eating_utensil], fork), h_{11} =(instrumentality, gun), h_{12} =(vehicle, boat), h_{13} =([cutlery, eating_utensil], table_knife)
O_5	instrumentality, conveyance, mail, public_transport, hosebox, vehicle, wheeled_vehicle, bus, train, bicycle, car, rolling_stock	h_1 =(instrumentality, coveyance), h_2 =(conveyance, mail), h_3 =(conveyance, public_transport), h_4 =(conveyance, hosebox), h_5 =(conveyance, vehicle), h_6 =(public_transport, bus), h_7 =(public_transport, train), h_8 =(vehicle, wheeled_vehicle), h_9 =(wheeled_vehicle, car), h_{10} =(wheeled_vehicle, bicycle), h_{11} =(wheeled_vehicle,

		rolling_stock)
O_6	instrumentality, [conveyance, transport], vehicle, wheeled_vehicle, motor, [bike, bicycle], [car, auto], truck, gun, boat	h_1 =(instrumentality, [conveyance, transport]), h_2 =([conveyance, transport], vehicle), h_3 =(vehicle, wheeled_vehicle), h_4 =(wheeled_vehicle, motor), h_5 =(wheeled_vehicle, [bike, bicycle]), h_6 =(motor, [car, auto]), h_7 =(motor, truck), h_8=(instrumentality, gun), h_9 =(vehicle, boat)
O_7	article, ware, [cutlery, eating_utensil], fork, table_knife	h_1 =(article, ware), h_2=(ware, [cutlery, eating_utensil]), h_3 =([cutlery, eating_utensil], fork), h_4 =([cutlery, eating_utensil], table_knife)
O_8	article, ware, [cutlery, eating_utensil], fork, table_knife, plate	h_1 =(article, ware), h_2 =(ware, [cutlery, eating_utensil]), h_3 =([cutlery, eating_utensil], fork), h_4 =([cutlery, eating_utensil], table_knife), h_5 =([cutlery, eating_utensil], plate)
O_9	article, ware, [cutlery, eating_utensil], fork, table_knife, plate, bowl	h_1 =(article, ware), h_2 =(ware, [cutlery, eating_utensil]), h_3 =([cutlery, eating_utensil], fork), h_4 =([cutlery, eating_utensil], table_knife), h_5 =([cutlery, eating_utensil], plate), h_6 =([cutlery, eating_utensil], bowl)
O_{10}	article, ware, [cutlery, eating_utensil], fork, table_knife, plate, bowl, spoon, glass	h_1 =(article, ware), h_2 =(ware, [cutlery, eating_utensil]), h_3 =([cutlery, eating_utensil], fork), h_4 =([cutlery, eating_utensil], table_knife), h_5 =([cutlery, eating_utensil], plate), h_6 =([cutlery, eating_utensil], bowl), h_7 =([cutlery, eating_utensil], spoon), h_8 =([cutlery, eating_utensil], glass)

We explain table 1 as following:

- Ontologies O_3 and O_5 are fragments of Wordnet;
- ontology O_4 is obtained by adding 3 concepts to O_3 (***gun***, ***boat*** and ***table_knife***);
- O_6 is a sub-ontology of O_4 with concepts: ***instrumentality***, [***conveyance***, ***transport***], ***vehicle***, ***wheeled_vehicle***, ***motor***, [***bike***, ***bicycle***], [***car***, ***auto***], ***truck***, ***gun***, ***boat***;
- O_7 is a sub-ontology of O_4 with concepts: ***article***, ***ware***, [***cutlery***, ***eating_utensil***], ***fork***, ***table_knife***;
- ontology O_8 is obtained by adding concept ***plate*** to O_7;
- ontology O_9 is obtained by adding concepts ***bowl*** to O_8;
- finally, ontology O_{10} is obtained by adding concepts ***spoon*** and ***glass*** to O_9.

Table 2 gives results of comparisons between ontologies using T_{NGOM} and N_{Plus}. Note that similarities between ontologies O_3 and O_4, and between O_3 and O_5 are assessed in [2].

Table 2 Results of comparisons of ontologies with measures $\boldsymbol{T_{NGOM}}$ and $\boldsymbol{N_{Plus}}$.

	T_{NGOM}	**Hierarchies added to ontologies for extensions**		N_{Plus}
(O_3, O_4)	0.95	**in** O_3 : h_1 (instrumentality, gun), h_2 (vehicle, boat), h_3 ([cutlery, eating_utensil], table_knife)	**in** O_4 : none	0.98
(O_3, O_5)	0.57	**in** O_3: h_1 (conveyance, mail), h_2 (conveyance, public_transport), h_3 (conveyance, hosebox), h_4 (wheeled_vehicle, rolling_stock)	**in** O_5: h_1 (wheeled_vehicle, motor)	0.74
(O_3, O_6)	0.76	**in** O_3 : h_1 (instrumentality, gun), h_2 (vehicle, boat)	**in** O_6 : none	0.88
(O_3, O_7)	0.6	**in** O_3: h_1 [cutlery, eating_utensil], table_knife)	**in** O_7 : none	0.83

(O_3, O_8)	0.49	**in** O_3: h_1 ([cutlery, eating_utensil], table_knife), h_2 (([cutlery, eating_utensil], plate)	**in** O_8 : none	0.74
(O_3, O_9)	0.47	**in** O_3: h_1 ([cutlery, eating_utensil], table_knife), h_2 ([cutlery, eating_utensil], late), h_3 ([cutlery, eating_utensil], bowl)	**in** O_9 : none	0.74
(O_3, O_{10})	0.43	**in** O_3: h_1 ([cutlery, eating_utensil], table_knife), h_2 ([cutlery, eating_utensil], late), h_3 ([cutlery, eating_utensil], bowl), h_4 ([cutlery, eating_utensil], spoon), h_5 ([cutlery, eating_utensil], glass)	**in** O_{10} : none	0.71
(O_7, O_9)	0.85	**in** O_7: h_1 (([cutlery, eating_utensil], plate), h_2 (([cutlery, eating_utensil], bowl)	**in** O_9 : none	0.93
(O_7, O_{10})	0.76	**in** O_7 : h_1 ([cutlery, eating_utensil], plate), h_2 ([cutlery, eating_utensil], bowl), h_3 ([cutlery, eating_utensil], spoon), h_4 ([cutlery, eating_utensil], glass)	**in** O_{10} : none	0.89

6. Analysis

In this section, we analyze results obtained from table 2 of section 5.

The assessment of the similarity between O_3 and O_4 gives good results for T_{NGOM} and N_{Plus}.measures (T_{NGOM} (O_3, O_4) = 0.95 and N_{Plus}. (O_3, O_4) = 0.98). The difference is not so important. The difference between ontologies is that O_4 contains three more concepts than O3 (***gun***, ***boat*** and ***table_knife***).

The similarity value between O_3 and O_5 is the average with T_{NGOM} measure (T_{NGOM} (O_3, O_5) = 0.57). After extension of ontologies, the similarity value increase with N_{Plus}measure (N_{Plus} (O_3, O_5) = 0.74).

The similarity value between O_3 and O_6 is good with T_{NGOM}measure (T_{NGOM} (O_3, O_6) = 0.76). When we add concepts to O_6, the similarity value decreases and tends toward low values (T_{NGOM} (O_3, O_7) = 0.6, T_{NGOM} (O_3, O_8) = 0.49, T_{NGOM} (O_3, O_9) = 0.47 and T_{NGOM} (O_3, O_{10}) = 0.43). When we consider "is-a" relation to extend ontologies, similarity values decrease but stay good with N_{Plus}measure (N_{Plus} (O_3, O_6) = 0.88, N_{Plus} (O_3, O_7) = 0.83, N_{Plus} (O_3, O_8) = 0.74, N_{Plus} (O_3, O_9) = 0.74 and N_{Plus} (O_3, O_{10}) = 0.71).

The similarity values between O_7 and O_9, and between O_7 and O_{10} are good with T_{NGOM} (T_{NGOM} (O_7, O_9) = 0.85 and T_{NGOM} (O_7, O_{10}) = 0.76). The values increase when we consider "is-a" relation and N_{Plus} measure (N_{Plus} (O_7, O_9) = 0.93 and N_{Plus} (O_7, O_{10}) = 0.89).

7. Conclusion

In this paper, we proposed a method for assessing similarity between two ontologies. The adopted approach is based on set theory, edges-based semantic similarity [4] and feature-based similarity [5]. This method has two parts. The first one can be summarized in 3 steps. In the step 1, we determined the sets of concepts which characterize the concepts shared by the two ontologies and the sets of concepts that are different from them. In the step 2, we assessed the average of the semantic similarity values between concepts of each set we determined in step 1. We have used Wu and Palmer [25] semantic similarity, an edge-based semantic similarity measure to compute similarity between concepts of the sets in an ontology, before assessing the average value of similarity for each set. Finally, in step 3, we adjusted the Tversky measure with N_{Plus} to evaluate the similarity between two ontologies. The second part of this method is an extension of the first one. In this part, we have extended the method of assessing similarity values between two ontologies by considering the "is-a" relation. This part can be summarized in 5 steps. As the first part of the method, steps 1 and 2 successively determine sets of concepts shared by ontologies and compute average of the semantic similarity values between concepts of each set. Differences between the two parts of the method become in the steps 3 which consists to extends ontologies O_1 and O_2 by using concepts subsumed by concepts of set ($O_1 \wedge O_2$). After this step, we have obtained O'_1 and O'_2 which are respectively extensions of O_1 and O_2. In the step 4, we have determined ($O'_1 \wedge O'_2$) the set of concepts share by O'_1 and O'_2. In the last step (step 5), we assessed similarity value of ontologies by using N_{Plus}measure. We experimented the method by assessing similarity values between ontologies extracted from Worldnet. We find that similarity values increase with N_{Plus}. In perspectives, we will propose an approach to assess similarity between an ontology and a speech in text format to check if the text and the ontology refer to the same theme.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare no conflicts of interest related to this work.

References

[1] A. N. Ngom, F.K. Sangare and M. Lo, "Proposition of a method for assessing Similarity Value Between Two Ontologies.", 4th Annual Conf. on Computational Science & Computational Intelligence (CSCI'17)|Dec 14-16, 2017 | Las Vegas, Nevada, USA, pages 174-179, 2017.

[2] A. N. Ngom, G. Kaladzavi, F.K. Sangare and M. Lo, "Assessing Similarity Value Between Two Ontologies.",In Proceedings of the 10th International Joint Cnference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3k 2018) - Volume 2:KEOD, pages 343-350, 2018.

[3] Khadim Dramé. "Contribution à la construction d'ontologies et à la recherche d'information : application au domaine médical.", Université De Bordeaux, 2014.

[4] A. N. Ngom, "Etude des mesures de similarité sémantique basée sur les arcs", CORIA, Paris, France, pp: 535 - 544, 2015.

[5] A. Tversky, "Features of similarity", Psychological Review, 84(4), pp: 327-352, 1977.

[6] Thomas, R. Gruber, "Atranslation approach to portable ontology specifications.", Knowledge Acquisition journal,Vol. 5, No.2, pp:199-220, 1993.

[7] Thomas, R. Gruber, "Toward principles for the design of ontologie sused for knowledge sharing.", International Journal of Human Computer Studies,Vol. 43, No.5, pp:907-928, 2005.

[8] Michael R., Genesereth & Nils. J. Nilsson, "Logical foundations of artificial intelligence.", Morgan Kaufmann Publishers Inc, 95 First Street, Los Altos, CA 94022, 1987.

[9] Gaoussou Camara "Conception d'un système de veille épidémiologique à base d'ontologies: Application à la schistosomiase au Senegal.", PhD thesis, Université Pierre Marie-Curie & Université Gaston Berger, Saint-Louis, 2013.

[10] Rudi, Studer, V. Richard, Benjamins, & Dieter, Fensel, "Knowledge engineering: Principles and methods.", IEEE Transactions on Data and Knowledge Engineering, 25 edition, 1998.

[11] Ana, Gomez-Perez & V.Richard, Benjamins, "Overview of knowledge sharing and reus components: Ontologies and problem- solving methods.", In Proceedings of the IJCAI99 workshop on Ontologies and problem-Solving Methods (KRR5), Stockholm (Suede), pp:111-115, 1999.

[12] Naouel, Ouafek, "Les ontologies spatiales : la generation automatique d'ontologies spatiales à partir du modèle conceptuel spatio-temporel mads.", Master's Thesis, University of Mentouri, Algeria.

[13] A. Maedche and S. Staab, "Ontology Learning for the Semantic Web", IEEE Intell. Syst. 16, 72 - 79. doi:10.1109/5254.920602, 2001.

[14] Feng Y., Bagheri E., Ensan F. & Jovanovic J., "The state of the art in semantic relatedness : a framework for comparison.", Knowledge Engineering Review, doi:10.1017/S0269888917000029, 2017.

[15] A. Maedche and S. Staab, "Measuring Similarity between Ontologies.", In: Gómez-Pérez A., Benjamins V.R. (eds) Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web. EKAW 2002. Lecture Notes in Computer Science, vol 2473. Springer, Berlin, Heidelberg,pp 251-263, 2002.

[16] I. V. Levenshtein. Binary Codes capable of correcting deletions, insertions, and reversals., Cybernetics and Control Theory, 10(8):707Ű710, 1966.

[17] Aly Ngoné NGOM “Une approche d’évolution d’ontoogie basée sur les mesures de similarité sémantique.", PhD thesis, Université Gaston Berger, Saint-Louis, 2018.

[18] James Z. Wang and Farha Ali. “An Efficient Ontology Comparison Tool for Semantic Web Applications.", Proceeding WI ’05 Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence Pages 372-378, 2005.

[19] Christine Fellbaum (ed.) “WordNet: An Electronic Lexical Database.", The MIT Press, May 1998.

[20] Giuseppe PIRRO, “A sémantic similarity metric combining features and intrinsic information content.", Language and Cognitive Processes, (6)p., pp: 1 - 28, 1991.

[21] G. PIRRO and J. EUZENAT, “A feature and information theoretic framework for semantic similarity and relatedness.", Data and Knowledge Engineering 68(11), pp: 1289 - 1308, 2009.

[22] A. N. Ngom, Y. Traore, P.F. Diallo, F.K. Sangare and M. Lo, “A method to update an ontology : simulation", Int’l Conf, Information and Knowledge Engineering, IKE’16, Las Vega, Nevada, USA, pp: 92 - 96, 2016.

[23] A. N. Ngom, P. F. Diallo, F. Kamara-Sangaré, M. LO, “A method to validate the insertion of a new concept in an ontology.", SITIS 2016 : The 12th International Conference on Signal Image Technology and Internet Systems, Naples, Italy, pp: 275 - 281, 2016.

[24] H. Zargayouna and S. Salotti, “Indexation sémantique de documents XML", PHD Thesis, Paris XI University, France, pp: 112 - 115, 2005.

[25] Z. Wu and M. Palmer, “Verbs semantics and lexical selection.", U.A.f.C.L. Stroudsburg, PA (ed.), In Proceedings of the 32nd annual meeting on ACL, volume 2 de ACL ’94.,pp: 133 - 138, 1994.

[26] G.A. Miller and W. Charles, “Contextual Correlates of Semantic Similarity", Language and Cognitive Processes, (6)p., pp: 1 - 28, 1991.