A Novel Global Measure Approach based on Ontology Spectrum to Evaluate Ontology Enrichment

Karim Kamoun
Faculty of Sciences of Tunis, University of Tunis El Manar, Tunisia

Sadok Ben Yahia
Faculty of Sciences of Tunis, University of Tunis El Manar, Tunisia

ABSTRACT
In the context of ontology evolution in real world applications, particularly in the field of semantic web, ontologies are called to change in their structure as well as their semantic. It is necessary to evaluate the quality based on stability to make analysis to get appropriate enrichment manner for ontology evolution.
In this paper, we introduce a new approach with that aims making three contributions. First, we present a new aspect of ontology quality based on its stability. Second, we present a new notion called *ontology spectrum* which can be used for analyzing ontology stability. Third, we provide an experimental method to evaluate this new aspect of quality within two processes: individual measure based on semantic similarity measures and global measure based on ontology spectrum.

Keywords
Ontology evaluation, semantic similarity measure, ontology enrichment, ontology quality, ontology stability.

1. INTRODUCTION
The Web is evolving toward the Semantic Web, in which the semantics of Web content is defined, making the Web meaningful, understandable, and machine-processable [1]. Ontologies stand as a key component of the Semantic Web since they are the backbone of knowledge representation; thus, they can be incorporated into computer-based applications and systems to facilitate data annotation [2], decision support [3,4], information retrieval, and natural-language processing [5] and serve as an integral part of the Semantic Web.
Ontologies are a very advanced tool of knowledge representation, for organizing the knowledge of a wide area of expertise. However, the sources of information, usually documentaries, constantly changing both the used vocabulary and the meaning of the elements containing it. Ontologies need, therefore, to be kept up-to-date so that the dependent systems remain operational. Following an update or more particular, an enrichment, ontology should be evaluated in order to examine its quality. it is important to determine whether that ontology is actually improving over time (becoming more complete and correct), or whether changes to the ontology had a negative effect, making the ontology less cohesive and correct and decreasing its quality.
A wide variety of evaluation methods have been proposed such as described in [6, 7, 8, 9]. These methods differ in that they have different interpretations of quality. In this paper, we look for evaluation approaches based on the use of ontology in real world context. In this context, related works are considered to be individual measures [8,10,12] that evaluate individual concepts or query results of the ontology.
In this work, we introduce a novel global evaluation measure approach for ontology evolution particularly for enrichment. This global measure will consider the ontology stability with metric independent to the number of concepts, in order to make enrichment evaluation between ontologies with different cardinalities.

The remainder of this paper is organized as follows. Section 2 positions this paper within the related work and motivates our proposed approach. In section 3, we describe the new aspect of quality based on ontology stability within individual similarities measure. This is followed, in section 4, by a description of new approach, called *global stability measure*, for analyzing ontology stability based on ontology spectrum. In section 5, we present and discuss the results of two experiments in order to validate our approach. Section 6 briefly recalls our contributions and sketches avenues for future work.

2. RELATED WORK
In this section, we scrutinize the related work that in snugness to our work. This state of the art is focused on two parts: the ontology evaluation and semantic similarity measures.

2.1 Ontology evaluation
Ontology's evaluation is an area of research that has emerged fairly recently, several approaches exists in the literature putting the focus on that topic[6,7,8,9] These approaches mainly differ in the way and the criteria chosen to evaluate the evolved ontology. We distinguish various evaluation approaches into five categories which are global standard approach, application based approach, data driven approach, human judgment approach and structural approach.

1- The first approach, called the *gold-standard* approach [10], attempts to assess the quality of ontologies using "gold standard" ontology. In this approach, the gold standard ontology is regarded as a well constructed one. It could be another existing ontology, or it could be taken statistically from a corpus of documents or prepared by a domain expert. The concepts of a constructed ontology are evaluated by comparing them vs those of the gold standard ontology. Typically, the gold standard approach is used to evaluate an ontology generated by a learning process.

2- The second one is an application-based approach [10] in which the quality of the ontology is evaluated based on its actual use in a real-world application [11]. The output of the application or its performance on the given task might be better or worse depending on the ontology of use. Ontologies may therefore be evaluated simply by plugging them into an application and evaluating the results of such application. Orme et al.[8], examine the quality, completeness, and stability of ontology data as far as ontology evolves. They propose a metrics suite, based on standard software quality concepts, to assess the complexity and cohesion of ontology data on one hand, and determine ontology completeness and Stability for evolved ontology on the other hand.

3- The third approach is data-driven since it evaluates the quality of ontology by measuring the fit between the ontology and the corpus of a problem domain to which it refers. In this approach ontology is evaluated only at the lexical level [12].

4- The fourth approach relies on human judgment. In the latter, the evaluation is carried out by domain experts whom try to evaluate how well the ontology meets a set of predefined criteria, standards, and requirements. Our approach belongs the second category.

5- The fifth approach used metrics operating following criteria [14]: complexity, cohesion, modularity, abstraction and taxonomy. In the same vision, the approach OntoQA (Metric-Based Ontology Quality Analysis) by Tartir et al [15], is worth of mention. The latter proposes a battery of metrics divided into two related categories: schema metrics and instance metrics. The first category evaluates ontology design and its potential for rich knowledge representation. The second category evaluates the placement of instance data within the ontology and the effective usage of the ontology to represent the knowledge modeled in the ontology.

In our work we stand within the second category described by Brank et al [10,12], the application-based approach category, since we consider that ontology was found to be used by an application, so any changes in ontological data directly affects ontology based systems. Therefore, it would be more reliable to evaluate the ontology with respect to its actual use and test its performance against expected results by users. Ontology-based applications, as do other approaches, generally use semantic similarity measures that explore the ontology and fulfill the needs of the application.

2.2 Semantic similarity measures

Generally ontology has a structure of concepts which the relation of subsumption (subClassOf) is the primary relationship. This structure defines the semantics of these concepts. The measures that exploit this structure are called semantic measures of concepts. Semantic measures can thus evaluate a link between two concepts of the same ontology by exploiting their relationship.
According to [16], various forms of semantic similarity measures can be classified in to three types: measures that focus in the characteristic of ontology's entities, semantic relationship measures and informational content measures.

For the first, based on characteristic of ontology's entities, the similarity measure between two concepts is defined as a process of matching characteristics. The given similarity is based on both common and different characteristics of those two concepts [17, 18, 19, 20].

For the second, based on semantic relationship, Rada et al. [21] proposed a metric to measure conceptual distance between A and B in hierarchical "is-a' 'semantic nets. The distance between A and B is equal to the minimum number of edges separating A and B. Wu and Palmer[22] depending on *mscs*(*Ci;Cj*) which refers to the most specific subsume (the lowest common ancestor in the tree) of both concepts *Ci* and *Cj*. Other works looking at improving measurement accuracy by considering other semantic links in addition to subsumption [23,24].

The third type, based on informational content, distinguishes between two categories of measures. Those using textual corpus and others using ontology structure:

- For the first category, Resnik [25], Jiang and Conrath [26] associates a probability p with concepts in a "is-a" hierarchy to denote the likelihood of encountering an instance of a concept c in a textual corpus.
- For the second category, Seco et al [27], Blanchard et al [28] present, respectively, new method for computing the information content of concept by considering only the taxonomic structure of the ontology.

In the context of second category with information content based on ontology structure, Blanchard et al [28] proposes four hypothesis of instance distributions. The first hypothesis, P_p, focuses on a uniform distribution among the concepts with the same profoundness; the information content of a concept depends basically only on its profoundness. The second hypothesis, P_s, implies an uniform distribution among the set of sons of each concept, the informational content of a concept depends on the number of sibling of the subsuming concepts. The third hypothesis, P_g, proposes an uniform distribution among the set of leaves of the ontology. The more leaves a concept has, the less important its informational content would be. The fourth and the last hypothesis is P_h, it focuses on the principle that the concepts of the same height should carry the same information content. The same authors Blanchar et al propose a new measure PSS "the Proportion of Shared Specificity" [28] which takes into account the density of links in the graph between two concepts. This measure is based on hypothesis P_s described above and takes the form of the Dice measurement.
We can characterize all those described measures as individual measures of similarities. This can be explained by the fact that it can individually measure the similarity of concepts. The quality evaluation approach we are going to propose can be applied for various similarity measures in particular the PSS measure.
As we have described previously, we focused in this paper on the evaluation quality of ontology based on its actual use in real word applications. Ontology based application can use semantic similarity measure to exploit ontology structure. In the next section, we present the aspect of quality based on stability which can be compute using semantic similarity measure.

3. APPROACH OF STABILITY BASED ON INDIVIDUAL MEASURES

The most useful approach of ontology quality evaluation is the one based on the use of the ontology in real world application. The user of ontology based system is interested in the response to their request queries.

Thus, in this work, we are interested in this category of quality evaluation of ontology

In the context of ontology enrichment, we present evaluation quality of ontology based on the stability of response for simplified request queries. This response is based on similarity measures to get suitable concept results from those queries. For complex queries, it can be considered as combination of simplified ones. We call this evaluation approach: a based individual measure evaluation. If this measure still unchanged for ontology enrichment, the ontology will be considered to be stable. Indeed, according to query responses, a stable ontology has not significant changes throw enrichment.

We look for the stability of the results regarding ontology evolution or more precisely ontology enrichment. Hence, we propose a new approach which can be presented as an objective to choose the best way to enrich the ontology.

Let us consider two ontologies O1 and O2 where O2 is an evolution of O1 after enrichment of this one. N stands for the cardinality of O1. We compute the similarity of the common set of concepts between O1 and O2 in order to find the semantic and structural stability of concepts following evolution. This similarity is computed using the average of the similarities between the concepts of different ontologies.

$$Sim\_common(O_1,O_2)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\left|sim(c_i^{O_1},c_j^{O_1})-sim(c_i^{O_2},c_j^{O_2})\right|}{n^2} \quad (1)$$

where n is the cardinality or the number of concepts contained in O_1 and O_2 is the enrichment result of O_1 $(O_1 \subset O_2)$ $C_i^{O_1}$ represents the concept C_i in ontology O_1 and *Sim_common* is the semantic similarity measure between two concepts.

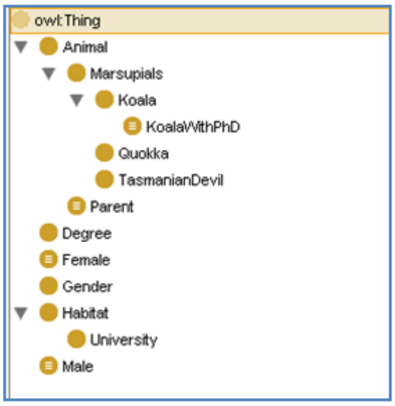

Fig1:Initial ontology Koala(vi).owl.

The value of *Sim_common* belongs to [0,1]. Whenever this value tends to 0, this means that the concepts of the ontology maintain almost the same semantic values after evolution.

We take an illustrative example of a simple ontology named koala.owl [29] defined by Knublauch in the reference site of Protege-OWL. The ontology Koala.owl includes 20 concepts except the concept of the virtual root (owl: Thing). It describes the concepts related to humans and marsupials (subclasses of mammals). We have removed randomly from this ontology seven concepts in order to obtain an initial ontology koala (vi) that includes 13 concepts which we enrich with 7 concepts that we have removed to finally reach our pristine ontology koala.owl. The final ontology is the enrichment result with the sub trees of concepts Person and Forest (c.f. Figure 1and 2).

In our context, to evaluate the quality of the ontology after enrichment, we must rely on a set of queries to evaluate their results regarding the initial ontology and its enrichment. The queries are mainly based on research concepts using semantic similarity measures. To be generic over queries submitted by the user, we make an exhaustive and simplified comparison of requests by similarity between concepts. Indeed, any query is a combination of a simple query search of a single concept. Applying the semantic similarity measure based on the information content PSS (Proportion of shared specificity) [28], described above, to both ontologies koala (vi) and koala.owl we obtain two tables containing the values of similarity measures between pairs of concepts belonging to the same ontology.

Tables 1 and 2 represent part of the similarity measures of common pairs of ontology's concepts before enrichment koala(vi).owl and after enrichment koala.owl.

Interestingly enough, there are pairs of concepts whose similarity values has changed after enrichment. For example, the pair of concepts (Parent, Marsupials), the similarity value in Table 1 is equal to 0.721, this value becomes 0.62 in the second table, which makes a difference of 0.101.

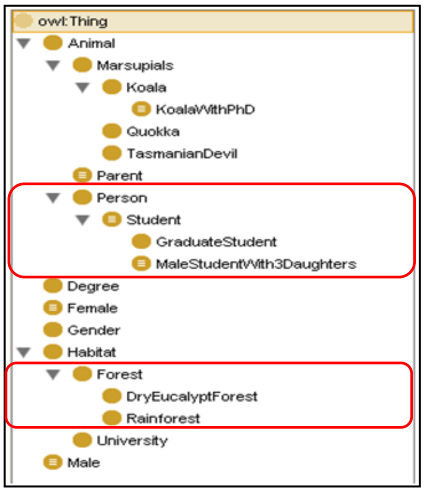

Fig2: Enrichment ontology Koala.owl.

Table 1.Similarity measure of concepts pairs of ontology Koala(vi).owl.

	Koala	**Marsupials**	**University**	**Parent**	**Animal**	**………**	**Habitat**
Koala	1	0,819	0	0,591	0,667	…….	0
Marsupials	**0,819**	1	0	0,721	0,838	…...	0
University	0	0	1	0	0	……	1
Parent	0,591	**0,721**	0	1	0,883	……	0
Animal	0,667	0,838	0	0,883	1	…….	0
…						…….	
Habitat	0	0	1	0	0	…….	1

Table 2. Similarity measure of concepts pairs of ontology Koala.owl after enrichment.

	Koala	**Marsupials**	**University**	**Parent**	**Animal**	**………**	**Habitat**
Koala	1	0,840	0	0,521	0,620	…….	0
Marsupials	**0,840**	1	0	0,620	0,765	…...	0
University	0	0	1	0	0	……	0,838
Parent	0,521	**0,620**	0	1	0,765	……	0
Animal	0,620	0,765	0	0,765	1	…….	0
. .						…….	
Habitat	0	0	0,838	0	0	…….	1

Whereas the pair of concepts (Koala, Marsupials) the difference of both measurements before and after enrichment is equal to 0.021 (= 0.84 - 0.819). These variations can be explained by the fact that the measure used PSS is based on the information content using the P_s hypothesis which depends on the number of brothers of the subsuming concepts and the enrichment affected this structure.

In order to check the similarity measures difference of common pairs of concepts between two ontologies, we calculated the average similarity measure previously shown in formula 1 by *Sim_common*, the result is: *Sim_common*(Koala(vi),Koala) = 0,013. The variation of *Sim_common* value depends essentially on the semantic similarity measure used. There is some measures which are more sensitive to some types of changes in the ontology structure during enrichment. In our example, the semantic similarity measure used did not really affect the value of *Sim_common*, since the enrichment did not touch significantly the structural relationships on which depends the similarity measure PSS.

Therefore, we evaluated the semantic stability for common concepts between ontology Koala (vi) and its enrichment Koala. In this individual measure based approach, the new concepts added (like the concept Person or Forest) in the ontology Koala, can be evaluated only after a new evolution. This constraint is the major limitation to use individual measures of similarities (*Sim_common*) to evaluate the ontology semantic stability. Looking for an alternative manner to evaluate semantic stability even for new added concepts, would be a good challenge.

In the case where these new concepts maintain almost the same semantic and structural aspects of the ontology, the ontology will be considered as stable. In the opposite case, ontology changes the appearance of its structure and therefore the enrichment causes a loss of ontology stability. For this requirement, we define a new global similarity measure of stability instead of the individual one based on the common concept set.

Figure 3 shows an example of ontology O_1, consisting of four concepts C_1, C_2, C_3 and C_4, which has been an enrichment of 3 other concepts C_5, C_6 and C_7 and obtain a new ontology O_2. The stability evaluation using individual similarity measure consider only the set of common concepts C_1,C_2,$C_3$and C_4, while the use of a global similarity measure should deal, in addition, with new concepts added C_5 , C_6 and C_7.

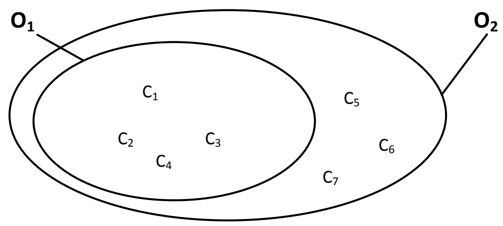

Fig 3: Limit of individual similarities measures

4. GLOBAL STABILITY MEASURE

In this section, we present a novel measure called global similarity measure. It's based on the notion of semantic similarity measures frequency of ontology.

We consider individual similarity measure for each both concepts of ontology. Their values can describe the ontology stability. To be independent to the number of concepts, we define a global measure based on the frequency of individual similarity measure. Those frequencies are computed by intervals. Thus, in each interval of similarity, we compute the number of concepts similarity measure values. This new measure is interpreted by the notion of frequency measures that defines the spectrum of ontology.

The frequency notion consists of counting the number of similarity measures values belonging sample interval, with size *Δs*, included in [0..1]. We take the example of Koala ontology [29]. Using the semantic similarity measure PSS [28] and considering intervals of size Δs = 0.05, we get the table of values (c.f. Table 3) and the histogram shown in Figure 4.

Table 3. Similarity values frequency.

Interval	0	]0, 0,05]	]0,05 , 0,1]	]0,1 , 0,15]	]0,15 , 0,2]	]0,2 , 0,25]	]0,25 , 0,3]	]0,3 , 0,35]	]0,35 , 0,4]	]0,4 , 0,45]	]0,45 , 0,5]	]0,5 , 0,55]	]0,55 , 0,6]	]0,6 , 0,65]	]0,65 , 0,7]	]0,7 , 0,75]	]0,75 , 0,8]	]0,8 , 0,85]	]0,85 , 0,9]	]0,9 , 0,95]	]0,95 , 1]
frequency	146	0	0	0	0	0	0	0	0	0	8	12	4	11	2	8	5	7	6	0	22

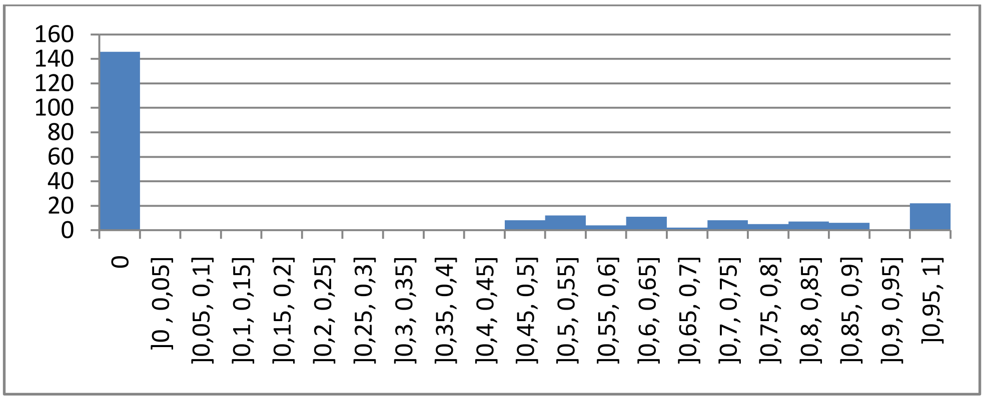

Fig 4: Spectrum of ontology Koala.owl.

For the similarity interval] 0.5, 0.55], there are 12 measures of concepts pairs belonging this interval. We also note a significant number (146) of measures set to 0, this reflects that many concepts have no semantic relationship with the similarity measure used.

Each ontology can be characterized as a spectrum of similarity measure (see figure 4). In order to measure stability, we compare ontology spectrum features before and after enrichment. If we keep the same pace of spectrum, then we can estimate that the ontology remains stable, because similarities proportions between concepts are maintained.

In order to have a simplified measure of similarity measures frequency spectrum, we rely on the formalism described in the next subsection.

4.1 Stability measure formalism

Frequency spectrum of ontology $Spect_O(S)$ compute the number of similarity values between all concepts $sim_O(C_i, C_j)$ in interval of size Δs. For a given value of similarity ***s***, the following expression is used to calculate the number of concepts with a similarity measure including the interval $\left]\left\lfloor \frac{s}{\Delta s} \right\rfloor * \Delta s, \left(\left\lfloor \frac{s}{\Delta s} \right\rfloor + 1\right) * \Delta s\right]$.

$$\forall s \in [0,1], Spect_O(s) = \sum_{i=1}^{n}\sum_{j=1}^{n} \delta\left(sim_O(C_i,C_j) - \left\lfloor \frac{s}{\Delta s} \right\rfloor * \Delta s\right) * \delta\left(\left(\left\lfloor \frac{s}{\Delta s} \right\rfloor + 1\right) * \Delta s - sim_O(C_i,C_j)\right) \quad (2)$$

$$With \begin{cases} \delta(x) = 1 \, si \; x > 0 \\ \delta(x) = 0 \, \mathrm{si} \; \mathrm{x} \leq 0 \end{cases}$$

Where Sim_O is the similarity function between concepts, it can be computed by using the similarity measure PSS.

For example, for the similarity value s = 0,12 and Δs = 0,05, we have the lower bound of the interval:

$$\left\lfloor \frac{s}{\Delta s} \right\rfloor * \Delta s = \left\lfloor \frac{0{,}12}{0{,}05} \right\rfloor * 0{,}05 = 2 * 0{,}05 = 0{,}1$$

and the upper bound:

$$\left(\left\lfloor \frac{s}{\Delta s} \right\rfloor + 1\right) * \Delta s = \left(\left\lfloor \frac{0{,}12}{0{,}05} \right\rfloor + 1\right) * 0{,}05 = 3 * 0{,}05 = 0{,}15$$

Thus, the value s = 0,12 belongs to the interval] 0.1, 0.15].

The function δ having the value 1, allows to consider the similarities in the specified interval. In fact, we have

$$\delta\left(sim_O(C_i,C_j) - \left\lfloor \frac{s}{\Delta s} \right\rfloor * \Delta s\right) = 1 \text{ if } sim_O(C_i,C_j) > \left\lfloor \frac{s}{\Delta s} \right\rfloor * \Delta s$$

Otherwise, the value of δ will be equal to 0.

The spectrum computes similarity values frequencies that depend on the number of concepts including ontology. In order to make this function independent of the ontology size, we give its normalized expression:

$$\forall s \in [0,1], SpectN_O(s) = \frac{Spect_O(s)}{n^2} \quad (3)$$

where n is the number of concepts for a given ontology O and s is a semantic similarity measure.

To evaluate ontology stability during enrichment process, we determine the difference between spectrums before and after enrichment respectively defined by O_1 and O_2. The following expression computes the average differences between normalized spectrums of two ontologies O_1 and O_2.

$$SimGlob(O_1,O_2)=\sum_{i=1}^{\left\lfloor \frac{1}{\Delta s}\right\rfloor}\frac{\left|SpectN_{O_1}(s_i)-SpectN_{O_2}(s_i)\right|}{\left\lfloor \frac{1}{\Delta s}\right\rfloor} \quad (4)$$

With $s_i = i*\Delta s$ et $s_i \in [0,1]$.

We note that if SimGlob converges to 0, we can deduce that we have two similar spectrums representative of two ontologies, and therefore the ontology O1 maintaines the same variation of semantic similarity measures even after enrichment to the ontology O2. This can be resulted in a stability of ontology regarding enrichment.

5. SIMULATION AND VALIDATION

Our new global evaluation approach is evaluated with various simulations to compare with the classical individual one as in [10, 12].

We realized automatic generator of virtual ontology reduced in taxonomies based on structural properties. The generator creates random taxonomic structures of concepts, built based on a subsumption hierarchy (tree structure), and export the results into ontologies described with OWL language.

The generator takes as input a number of structural criteria that must be followed during the construction of random ontologies. The criteria chosen are :

- The number of nodes or concept of the ontology;
- The number of leaves of the tree representing the ontology;
- The depth of the tree;
- The Average sons per node ;
- The ratio R= $\dfrac{Max(number \text{ sons by concept (not including leaves)})}{Min(number \text{ sons by concept (not including leaves)})}$
- The variance E: it's a difference between the depth of the tree built and the minimum depth of a leaf.

The generator will be used to randomly generate an initial ontology and also enrichment by adding random concepts.

In all simulations, we rely on two types of enrichment: poor and important. For poor enrichment, the size of the initial ontology is 50 concepts, and then five enrichments are performed each time with an addition of a single concept. Concerning important enrichment, we start with ontology including 15 concepts and then we enriched it five times with adding successively an important number of concepts: 30, 50, 90, 120 and 150.
We first compare our proposed approach on the ontology stability, based on global similarity with the individual approach similar to several research such as described in [10]. Next, we study the impact of enrichment volume on the stability of global similarity value.

5.1 Comparison between global similarity and individual similarity

In the following, we will look for the correlation interpreting the difference between global and individual similarity in relation to the importance of concepts numbers enriched with. To increase the number of simulation samples for a more significant study of correlation, we start with three initial ontologies randomly constructed, for each one we made five enrichments which give a total of 15 samples.

In the case of poor enrichment, the correlation value of Peterson is 0,94 close to 1 (figure 5). This indicates that the global and individual similarity measure have the same semantic interpretation. However, global similarity presents an additional advantage over the individual one in the ability to evaluate in addition the quality of new enriched concepts. Individual similarity evaluates the quality in the point of view of stability by computing similarity between two ontologies with same size. That is why, it is applied to the common set of concepts between ontology and its enrichment. Whereas, the global similarity is independent of the concepts number in ontology, and can be applied in general case.

For important enrichment, there will be a more significant difference between ontology and its enrichment. Because global similarity processes all concepts and individual similarity deals only with the common set of concepts between ontologies, the correlation value has declined significantly and is equal to 0.756 (figure 6) compared to 0.941 in the previous figure (figure 5) for poor enrichment.

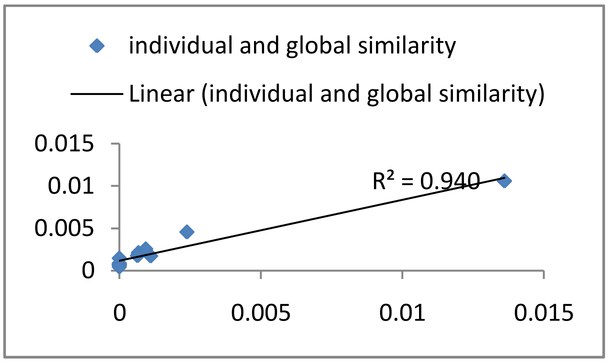

Fig 5: Correlation between individual and global similarity measure for poor enrichment.

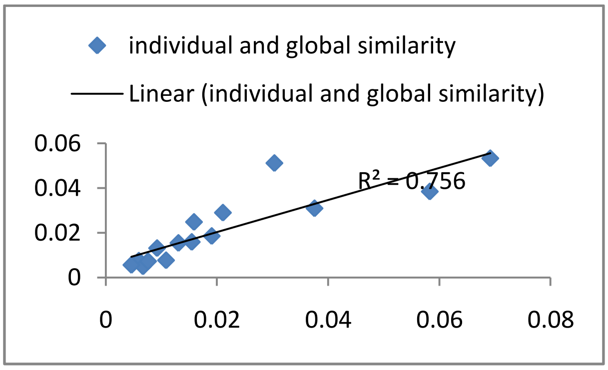

Fig 6: Correlation between individual and global similarity measure for important enrichment.

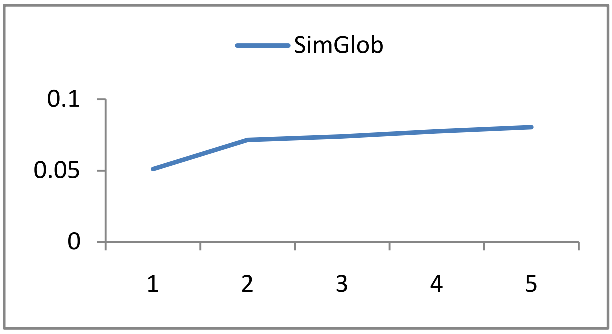

Fig 7: Global similarity measure between initial ontology and after a number of enrichment.

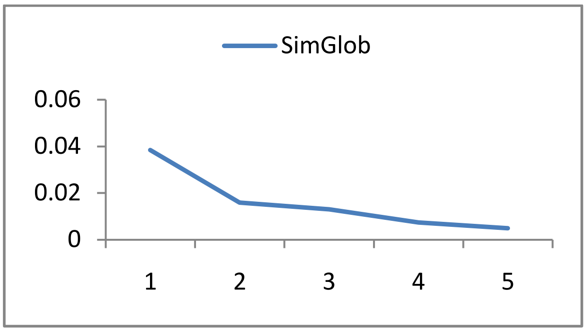

Fig 8: Global similarity measure between two successive ontologies before and after enrichment.

Since the proportion of all common concepts becomes less important than the enriched concepts, the individual similarity measure becomes inefficient. This reflects the weakness of the correlation with global similarity measure.

So, whenever the number of enriched concepts tends to 0, the individual measure would be more accurate for evaluating ontology stability. In this case, we showed that our global similarity approach has very strong correlation with the individual measure. In general case, when the enrichment is important, the individual measure is limited to interpreting the quality based on stability, the correlation becomes lower. Thus, our global similarity approach is highlighted.

5.2 Impact of enrichment volume on the stability of global similarity measure

In this simulation, we take an initial ontology containing 20 concepts and five enrichments are done each time with the addition of five concepts. We focus on the variation of global similarity measure in relation with enrichment. In Figure 7, we determine the global similarity between initial ontology and ontology after k enrichments; k varies from 1 to 5. The similarity is gradually increased, but with a smaller gradient, we note that the ontology become more stable relative to the increase of its size. Figure 8 presents the overall similarity between two successive ontologies regarding enrichment. This difference becomes smaller, confirming the same interpretation as the previous figure 7 that the ontology is gradually stabilized after successive enrichment.

6. CONCLUSION

To evaluate ontology stability with regard to enrichment, we have proposed two general approaches: the individual measure as classical one and the global measure. The global one describes ontology with frequency spectrum of concepts similarity measure. This new measure approach reflects simple request query looking for concepts of ontology. It is considered in the context of using ontology in real world application. Our spectrum evaluation approach gives an automatic stability evaluation of ontology with regard to enrichment. We performed the efficiency of global measure approach by simulations making random ontology construction and enrichment.

Our approach has similar results to classical one in the case of poor enrichment. However, it overcomes the limitation of classical approach to evaluate a higher enrichment.

Moreover, we have here restricted ourselves to a hierarchical structure deduced from the “is-a” link and the use of only one semantic similarity measure PSS. Although this structure is known to be the most structuring of a real-life ontology. As future work, we will attempt to generalize our approach to a graph structure to simultaneously take other links into account. In addition we will try to use other similarity measure and combine them in order to exploit all aspects of ontology structure.

7. REFERENCES

[1] T. Berners-Lee, J. Hendler and O. Lassila. The Semantic Web, *Scientific American* 284:34–43, 2001.

[2] MA. Musen. Ontologies in biomedicine. *AMIA 2008 tutorial T26*. Washington, DC, November 9; 2008.

[3] AC. Yu. Methods in biomedical ontology. *Journal Biomed Inform*;39:252–66, 2006

[4] MA. Musen Scalable software architectures for decision support. Methods Inf Med; 38:229–38, 1999.

[5] DL Rubin, SE Lewis, CJ Mungall, S. Misra, Westerfield M Ashburner, et al. National Center for Biomedical Ontology: advancing biomedicine through structured organization of scientific knowledge. *OMICS 2006*; 10:185–98, 2006.

[6] A. Burton-Jones, VC. Storey, V. Sugumaran, P Ahluwalia. A semiotic metrics suite for assessing the quality of ontologies. *In proceeding of Data and Knowledge Engineering 2005;* 55(1): 84–102, 2005.

[7] A. Gangemi, C. Catenacci, M. Ciaramita, J. Lehmann, A theoretical framework for ontology evaluation and validation. In *Proceedings of the Semantic Web Applications and Perspectives (SWAP),* 2nd Italian Semantic Web Workshop, Trento, Italy, 2005.

[8] A.M. Orme, H. Yao, and L.H. Etzkorn, Indicating ontology data quality, stability, and completeness throughout ontology evolution, *Journal of Software Maintenance*, 49-75, 2007

[9] G. Beydoun, A.A. Lopez-Lorca, F.G. Sánchez, and R. Martínez-Béjar, "How do we measure and improve the quality of a hierarchical ontology?", *Journal of Systems and Software*, 2363-2373, 2011.

[10] J Brank., M Grobelnik., D Mladenic., A Survey of Ontology Evaluation Techniques, *in Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)*, Ljubljana, Slovenia, 2005.

[11] J. Yu, J. Thom, A. Tam, Requirements-oriented methodology for evaluating ontologies, *Information Systems* 34:766–791, 2009

[12] J. Brank, D. Mladenic, M. Grobelnik, Gold standard based ontology evaluation using instance assignment, *Proceedings of the 4th International Workshop on Evaluation of Ontologies for the Web (EON)* at the 15th International World Wide Web Conference, Edinburgh, UK, 2006.

[13] A. Baneyx, J. Charlet, Evaluation, évolution et maintenance d'une ontologie en medecine : état des lieux et expérimentation. *Revue I3* ; *SI 2006 special issue on Ontological ressources*, 2006

[14] R. Djedidi and M.A. Aufaure. Patrons de gestion des changements owl. In Fabien L. Gandon, editor. *In proceedings of'knowledge engineering (IC)*, 145–156. PUG, 2009.

[15] S Tartir, IB Arpinar, M Moore, AP Sheth, B Aleman-Meza. OntoQA: Metric-based ontology quality analysis, *In Proceedings of IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources*, 2005.

[16] E. Blanchard, M. Harzallah, P. Kuntz and H. Briand. Sur l'évaluation de la quantité d'information d'un concept dans une taxonomie et la proposition de nouvelles mesures. *Special issue "knowledge modeling" journal of new information technologies (RNIT).* Cepadues (12), 127-145, 2008.

[17] Tversky, A. Features of similarity. *Psychological Review 84*(4), 327–352.

[18] P Jaccard, , Distribution of the alpine flora in the dranse's basin and some neighbouring regions (in french). *Bulletin de la Soc. Vaudoise Sci. Nat.* (37), 241–272, 1901

[19] L. R. Dice, Measures of the amount of ecologic association between species. *Ecology 26*(3), 297–302, 1945.

[20] A.Ochiaï. Zoogeographic studies of the soleoid fishes found in japan and its neighbouring regions. *Bulletin of the Japanese Society for Scientific Fisheries 22*, 526–530. 1957.

[21] R. Rada, H. Mili, E. Bicknell, and M. Blettner. Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, Jan/Feb 1989.

[22] Z. Wu and M. Palmer. Verb semantics and lexical selection. *In proceedings. of the 32nd annual meeting of the associations for Comp. Linguistics*, 133–138. 1994

[23] P. H. Ganesan, Garcia-Molina and J. Widom. Exploiting hierarchical domain structure to compute similarity. *ACM Trans. on Information Systems 21*(1): 64–93, 2003.

[24] A. G. Maguitman, F. Menczer, H. Roinestad, and A. Vespignani. Algorithmic detection of semantic similarity. *In proceedings of the 14th int. conf. on world wide web*, 107–116. ACM Press, 2005

[25] P. Resnik, Semantic similarity in a taxonomy : An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, 11: 95–130. 1999

[26] J. J. Jiang, and D. W. Conrath. Semantic similarity based on corpus statistics and lexical taxonomy. *Proceeding of int. conf. on Research in Computational Linguistics*, 19–33, 1997.

[27] N. Seco, T. Veale, and J. Hayes. An intrinsic information content metric for semantic similarity in wordnet. *In proceedings of the 16th European conf. on artificial intelligence*, 1089–1090, 2004.

[28] E. Blanchard, M. Harzallah and P. Kuntz. A generic framework for comparing semantic similarities on a subsumption hierarchy. *In proceedings of 18th European Conference on Artificial Intelligence (ECAI)*,20-24, 2008.

[29] Ontology library of Protégé OWL: http://protege.stanford.edu/plugins/owl/owl-library/koala.owl.