Assessing Similarity Between Two Ontologies: 
The Use of the Integrity Coefficient

Aly Ngoné Ngom
LANI, Gaston Berger University, Saint-Louis, Sénégal
SET, Sup de Co of Dakar, Sénégal
Papa Ousseynou Mbaye
SET, Sup de Co of Dakar, Sénégal
Ibrahima Gaye
Alioune Diop University of Bambey, Sénégal

Abstract
The aim of this paper is to propose a new coefficient of integrity $I_{new}$ for improving $N_{Plus}$ measure which is an improvement of the $T_{Ngom}$ measure. In $N_{Plus}$ measure, the coefficient of integrity used ($I$) decreases and tends to 0 fastly when we just add some concepts for extending set of resemblance of ontologies. To fix this problem, we introduce $R$, the coefficient of representativeness of concepts added in the ontology for its extension. $I_{new}$ decreases slowly compared to $I$ and depends to the cardinality of the ontology extended and the number of concepts added to it too.

2012 ACM Subject Classification Theory of computation → Operational semantics; Information systems → Data structures

Keywords and phrases Semantic similarity measures, Ontologies similarities, Tversky ’s measures, Concepts similarities

Digital Object Identifier 10.4230/OASIcs.SLATE.2022.1

Acknowledgements We want to thank Sup De Co of Dakar which supports all fund of the registration.

1 Introduction

Ontologies allow to formalize knowledge related to the description of the world by making them accessible and shareable across the Web. They introduce the semantic layer into the architecture of the on based-systems [1]. When several ontologies are used for an application, it is necessary that these ontologies present some similarity. The assessing of similarity between ontologies may be very interesting. Indeed, it can make easy the choice of ontologies in the case of elaboration of a system, which uses them. In addition, it can help to evaluate the ontology evolution by comparing its different versions. In [6], we have proposed an approach for assessing similarity between two ontologies $O_1$ and $O_2$. This approach has given good results but doesn’t take into account some properties (relations, axioms) for extending the formed sets (set of resemblance and set of differences) and improving the proposed measure. In [5], a new measure $N_{Plus}$ is proposed in the goal to improve the $T_{Ngom}$ measure proposed in [6]. $N_{Plus}$ measure improves $T_{Ngom}$ in the fact of it takes into account “is-a” relation to enlarge the formed set of resemblance by extending ontologies. In the formula of the $N_{Plus}$ measure, a coefficient of integrity $I$ is defined. The integrity coefficient of an ontology $O_j$ named $I_j$ is a value which is related to the number of concepts of $O_j$ ($n_j$) that we have to add to another ontology $O_i$ in goal to extend it ($i,j \in \mathbb{N}$). The larger is $n_j$, the smaller is $I_j$. The major problem of $I$ is that it decreases fastly and tends fastly to 0 when we just add some concepts. To fix this problem, we propose in this paper a new coefficient of integrity $I_{new}$, which will decrease slowly in relation to the added concepts for extending ontologies and the size (cardinality) of ontologies.
Similarity Between Two Ontologies

This paper continues by presenting our research context. Section 2 presents the related work. Then, we present the new coefficient of integrity $I_{\text{new}}$ in section 3 before making its experimentation in section 4. We end with conclusion and perspectives of this work.

## 2 Related Works

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

Maedche and Staab [3] propose a method for comparing two ontologies. This method is based on two levels:
- the **Lexical level** which consists of investigation on how terms are used to convey meanings;
- the **Conceptual level** which is the investigation of what conceptual relations exist between terms.

The Lexical comparison allows to find concepts by assessing syntactic similarity between concepts. It is based on Levenshtein [2] edit distance (ed) formula which allows to measure the minimum number of change required to transform one string into another, by using a dynamic programming algorithm. The Conceptual Comparison Level allows to compare the semantic of structures 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 to assess 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 the method only considers the presence of the concept in the hierarchy.

In [6], we propose an approach which assesses similarity between two ontologies. The approach is based on set theory, edges based [4] and feature-based similarity based [7]. It is composed of three steps. The first step consists to determine the different sets : set of concepts shared by the two ontologies and sets of concepts specific to each ontology. In the second step, we have assessed the average similarity values between concepts for each sets thanks to semantic similarity measures. Finally, in the last step, we have assessed similarity between ontologies by redefining Tversky’s measure relying to the two first steps. Figure 1 represents ontologies $O_1$ and $O_2$.

In figure 1, 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 $O_1$ ;
- $(O_1 \cap O_2) = \{B, D, F, G\}$ : set of concepts present in $O_1$ and $O_2$.

The three steps can be summarized in below:
- **The Step 1** consists to determine the sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O_1 \cap O_2)$.
- Once the sets are determined, we assess the 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.
To assess similarity between two ontologies, we have defined a measure which readjust the Tversky measure. This measure takes into account shared features and differences of ontologies. Applying the Tversky measure, the similarity between $O_1$ and $O_2$ is given by the formula 1.

$$T_{\text{vr}}(O_1, O_2) = \frac{f(O_1 \cap O_2)}{f(O_1) + f(O_2) + \beta f(O_1 \setminus O_2)}$$

Instead of the function $f$, we use Wu and Palmer [8] semantic similarity measure. for every determined set, we have computed the average of the similarity values between concepts. Using Wu and Palmer similarity measure, the similarity between two concepts $c_1$ and $c_2$ is given by the formula 2.

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

The concept $c_3$ 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 1 becomes formula 3.

$$T_{\text{Ngom}}(O_1, O_2) = \frac{\theta \cdot \overline{\text{Sim}}(O_1, O_2) + \omega \cdot \overline{\text{Sim}}(O_2)}{\theta \cdot \text{card}(O_1) + \omega \cdot \text{card}(O_2) + \alpha \cdot \overline{\text{Sim}}(O_1 \setminus O_2) + \beta \cdot \overline{\text{Sim}}(O_2 \setminus O_1)}$$

with:

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

and where:

- $\overline{\text{Sim}}(O_1)$ (respectively $\overline{\text{Sim}}(O_2)$) is the average value of similarity between concepts $(x_i, x_j)$ in ontology $O_1$ (respectively $(x_i, x_j)$ in ontology $O_2$), $i, j \in \mathbb{N}$ and $i \neq j$.
- $\overline{\text{Sim}}(O_1 \setminus O_2)$ (respectively $\overline{\text{Sim}}(O_2 \setminus O_1)$) 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 $O_1$), $i, j \in \mathbb{N}$ and $i \neq j$.
- the coefficients $\theta$, $\omega$, $\alpha$ and $\beta$ allow to take into account the similarity values in relation to the number of concepts of the sets and number of concepts of ontologies.

**Figure 1** Representation of ontologies $O_1$ and $O_2$ with Tversky’s feature model.
1:4  Similarity Between Two Ontologies

The measure presented by formula 3 respects this 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 $O_1 = O_2$.

The method we have proposed in [6] gives satisfactory results. Indeed, it allows to assess similarity between two ontologies while taking into account the semantic links that exist between the concepts in ontologies. However, it doesn’t take into account properties of concepts for extending the formed sets (set of resemblance and set of difference).

For taking into account properties of concepts for extending the formed sets (set of resemblance and set of difference), we have improved the $T_{Ngom}$ measure [6] by proposing in [5] an approach which assesses similarity between ontologies by extending set of resemblance thanks to the is-a relation of ontologies. The improved approach can be summarized in five steps:

- **Step 1** consists to determine the sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O_1 \cap O_2)$.
- Once the sets are determined, we assess the average of the semantic similarity values between concepts of each set in **step 2**.
- In **step 3**, we extend ontologies $O_1$ and $O_2$ by using the set $(O_1 \cap O_2)$. In this step, for each concept $c$ of $(O_1 \cap O_2)$, we search there sons $x_i \ (i \in \mathbb{N})$ in $O_1$ (respectively $O_2$) and we add them as sons of $c$ in $O_2$ (respectively $O_1$) 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 $(O'_1 \cap O'_2)$ shared by the two ontologies.
- In **step 4**, we determine $(O'_1 \cap O'_2)$ 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 the results of the step 2 and 4 in our measure which is a redefinition of the $T_{Ngom}$ measure [6].

In summary, for assessing similarity between ontologies, we use sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O'_1 \cap O'_2)$; 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 $(O'_1 \cap O'_2)$. Figure 2 represents the different we use for assessing similarity between ontologies $O_1$ and $O_2$.

![Figure 2](image-url)  
**Figure 2** Representation of extensions 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 $O_1$;
- $(O'_1 \cap O'_2) = \{B, C, D, E, F, G\}$ : set of concepts present in $O'_1$ and $O'_2$. 

In Figure 2, we represent three parts:

- $(O_1 \setminus O_2)$: concepts present in $O_1$ and not in $O_2$;
- $(O_2 \setminus O_1)$: concepts present in $O_2$ and not in $O_1$;
- $(O'_1 \cap O'_2)$: concepts present in both $O'_1$ and $O'_2$. 

In summary, for assessing similarity between ontologies, we use sets $(O_1 \setminus O_2)$, $(O_2 \setminus O_1)$ and $(O'_1 \cap O'_2)$; 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 $(O'_1 \cap O'_2)$.
The measure is given by the formula 4:

\[
N_{\text{Plus}}(O_1, O_2) = \frac{\theta \sigma_{O_1'} \cap O_2 + \omega \sigma_{O_2'} \cap O_1}{\theta \sigma_{O_1'} \cap O_2 + \omega \sigma_{O_2'} \cap O_1 + \alpha \vartheta_{O_1 \setminus O_2} + \beta \vartheta_{O_2 \setminus O_1}}
\]

with:

- \( \theta = \frac{\text{cardinality}(O_1 \cap O_2)}{\text{cardinality}(O_1) + n_1 + n_2} \)
- \( \omega = \frac{\text{cardinality}(O_2 \cap O_1)}{\text{cardinality}(O_2) + n_1 + n_2} \)
- \( \alpha = \frac{\text{cardinality}(O_1 \setminus O_2)}{\text{cardinality}(O_1)} \)
- \( \beta = \frac{\text{cardinality}(O_2 \setminus O_1)}{\text{cardinality}(O_2)} \)
- \( I_1 = \frac{1}{1 + n_2} \)
- \( I_2 = \frac{1}{1 + n_1} \)
- \( \sigma_{O_1'} \) (respectively \( \sigma_{O_2'} \)) is the average value of similarity between concepts \((x_i, x_j)\) in ontology \(O_1'\) (respectively \((x_i, x_j)\) in ontology \(O_2'\)). \( i, j \in \mathbb{N} \) and \( i \neq j \).
- \( \vartheta_{O_1 \setminus O_2} \) (respectively \( \vartheta_{O_2 \setminus O_1} \)) 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 \(O_1\)). \( i, j \in \mathbb{N} \) and \( i \neq j \).
- \( \text{cardinality}(O) \) is the number of elements (concepts) of the set (ontology) \(O\);
- \( I_i \) : Integrity coefficient of Ontology \(O_i\) \((i \in \mathbb{N})\);
- \( n_i \) : number of concepts of \(O_i\) added for extending \(O_j\) \((i, j \in \mathbb{N})\);
- As in \([6]\), \( \theta, \omega, \alpha \) and \( \beta \) are parameters which allow to take into account the similarity values in relation to the number of concepts of the sets 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 \(O_i\) in goal to extend it \((i, j \in \mathbb{N})\). The larger is \(n_j\), the smaller is \(I_i\). We have the expression 5:

\[
\begin{align*}
\lim_{n \to \infty} I &= \lim_{n \to \infty} \frac{1}{1 + n} = 0; \\
\lim_{n \to 0} I &= \lim_{n \to 0} \frac{1}{1 + n} = 1;
\end{align*}
\]

with \((n \in \mathbb{N})\).

We note that measure presented by formula 4 like formula 3 respects this properties:

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

The \( N_{\text{Plus}} \) measure presented in [5] improves \( T_{N_{\text{gomo}}} \) measure of [6] because it takes into account is-a relation of ontologies to extend set of resemblance of ontologies by adding concepts. The limits of this measure is in the integrity coefficient of ontologies \(I\). \( I \) decreases fastly if we add just a few of concepts. Figure 3 shows the evolution of \( I_i \) when we add concepts for extending set of resemblance.

In Figure 3, we remark that \( I \) decreases very fastly. When we add concepts in an ontology for extending the resemblance set, we note that variation of \( I = \frac{1}{1 + n} \):

- \( I = 1 \) if \( n = 0 \);
- \( I = 0,5 \) if \( n = 1 \);
- \( I = 0,33 \) if \( n = 2 \);
- ...
- \( I = 0,1 \) if \( n = 9 \).
The integrity coefficient is equal to 0.5 when we just add one concept. It means that with one concept added, we lose 50% of the integrity of the ontology. We need to find a good coefficient formula for improving $N_{Plu+}$.

3 New definition of integrity coefficient of ontology ($I_{new}$)

In section 2, we presented $N_{Plu+}$ measure of [5] which improves $T_{Ngom}$ measure of [6] by extending set of resemblance of ontologies. In definition of $N_{Plu+}$ formula, we introduced $I$ which represents the integrity coefficient. The integrity coefficient ($I$) is a value introduced to express how an ontology loses its integrity when concepts are added in the goal to extend set of resemblance shared with another ontology. This value is between 0 and 1. Initially, when there is no concept added to the ontology, the value of $I$ is 1. This value decreases as we add concept to the ontology. Figure 3 shows that $I$ decreases fastly to 0 when we add just a lot of concepts. To fix this problem, we propose to redefine $I$. We introduce $R$ for expressing coefficient of representativeness of concepts added in the ontology for its extension. $R$ is expressed by the formula 6:

$$R = \frac{n}{\text{cardinality}(O) + n}$$

with:
- $n$ is the number of concepts added to the ontology $O$ for its extension ($n \in \mathbb{N}$);
- $\text{cardinality}(O)$ represents the number of concepts of the ontology $O$;

We redefine $I$ with $I_{new}$ in terms of $R$ according to the formula 6 and we obtain the formula 7:

$$I_{new} = 1 - R$$

$$I_{new} = 1 - \frac{n}{\text{cardinality}(O) + n}$$

$$I_{new} = \frac{\text{cardinality}(O)}{\text{cardinality}(O) + n}$$

The new coefficient of integrity $I_{new}$ represented by formula 7 is between 0 and 1 as shown in below:

$$\lim_{n \to \infty} I_{new} = \lim_{n \to \infty} \frac{\text{cardinality}(O)}{\text{cardinality}(O) + n} = 0;$$

$$\lim_{n \to 0} I_{new} = \lim_{n \to 0} \frac{\text{cardinality}(O)}{\text{cardinality}(O) + n} = 1;$$

with ($n \in \mathbb{N}$).
For studying the evolution of \( I_{\text{new}} \), we use 4 functions \( I_1, I_2, I_3 \) and \( I_4 \), and we draw there evolutions curves. Functions are defined in below:

\[
I_1 = \frac{1}{1 + \frac{n}{100}} \quad \text{The same formula of I in [5]};
\]

\[
I_2 = \frac{100}{100 + n} \quad \text{cardinality}(O) = 100
\]

\[
I_3 = \frac{500}{500 + n} \quad \text{cardinality}(O) = 500
\]

\[
I_4 = \frac{1000}{1000 + n} \quad \text{cardinality}(O) = 1000
\]

Figure 4 represents the evolutions of different curves of functions \( I_1, I_2, I_3 \) and \( I_4 \).

In Figure 4, we remark that \( I_1 \) represented by \( I_1 \) decreases fastly than \( I_2, I_3 \) and \( I_4 \). We note that, the new coefficient of integrity \( I_{\text{new}} \) improves \( I \) and its evolution relies to the cardinality of ontology extended and the number of concepts added to it.

4 Simulations

In this section, we experiment our methodology. We compare \( T_{\text{Ngom}} \) and \( N_{\text{Plus}} \) measures using coefficient of integrity \( I \) proposed in [5] and the new coefficient of integrity \( I_{\text{new}} \) proposed in this paper. We assess similarity of ontologies by using ontologies extracted from WordNet\(^1\). Figure 5 to 12 represent ontologies that we use to simulate the measures.

Ontologies are explained 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 \( O_9 \).

\(^1\) http://wordnet.princeton.edu
Table 1 gives results of comparisons between ontologies using $T_{Ngom}$, $N_{Plus}$ and $N_{PlusI_{new}}$. $N_{PlusI_{new}}$ represents $N_{Plus}$ with the new coefficient of integrity $I_{new}$. Note that similarities between ontologies $O_3$ and $O_4$, and between $O_3$ and $O_5$ are assessed in [5].

In table 1 we note that $N_{Plus}$ improves $T_{Ngom}$ by considering “is-a” relations to extends ontologies. In the same way, $N_{PlusI_{new}}$ improves $N_{Plus}$ thanks to the improvement of the coefficient of integrity. Similarity of ontologies is better with $N_{PlusI_{new}}$. Values of similarities between ontologies are improved. The correlation coefficient between these three measures presented in table 2 is very good and very close to 1 ($\text{correlation}(N_{Plus}, N_{PlusI_{new}}) =$
<table>
<thead>
<tr>
<th></th>
<th>$T_{Ngom}$</th>
<th>Hierarchies added to ontologies for extensions</th>
<th>$N_{Plus}$</th>
<th>$N_{Plus}I_{new}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_3$, $O_4$</td>
<td>0.95</td>
<td>in $O_4$: $h_1$(instrumentality, gun), $h_2$(vehicle, boat), $h_3$([cutlery, eating_utensil], table_knife)</td>
<td>none</td>
<td>0.98</td>
</tr>
<tr>
<td>$O_3$, $O_5$</td>
<td>0.57</td>
<td>in $O_5$: $h_1$(conveyance, mail), $h_2$(conveyance, public_transport), $h_3$(conveyance, hosebox), $h_4$(wheeled_vehicle, rolling_stock)</td>
<td>in $O_6$: $h_1$(wheeled_vehicle, motor)</td>
<td>0.74</td>
</tr>
<tr>
<td>$O_3$, $O_6$</td>
<td>0.76</td>
<td>in $O_6$: $h_1$(instrumentality, gun), $h_2$(vehicle, boat)</td>
<td>none</td>
<td>0.88</td>
</tr>
<tr>
<td>$O_3$, $O_7$</td>
<td>0.6</td>
<td>in $O_7$: $h_1$([cutlery, eating_utensil], table_knife)</td>
<td>none</td>
<td>0.83</td>
</tr>
<tr>
<td>$O_3$, $O_8$</td>
<td>0.49</td>
<td>in $O_8$: $h_1$([cutlery, eating_utensil], table_knife), $h_2$([cutlery, eating_utensil], plate)</td>
<td>none</td>
<td>0.74</td>
</tr>
<tr>
<td>$O_3$, $O_9$</td>
<td>0.47</td>
<td>in $O_9$: $h_1$([cutlery, eating_utensil], table_knife), $h_2$([cutlery, eating_utensil], late), $h_3$([cutlery, eating_utensil], bowl)</td>
<td>none</td>
<td>0.74</td>
</tr>
<tr>
<td>$O_3$, $O_{10}$</td>
<td>0.43</td>
<td>in $O_{10}$: $h_1$([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)</td>
<td>none</td>
<td>0.71</td>
</tr>
<tr>
<td>$O_7$, $O_8$</td>
<td>0.85</td>
<td>in $O_8$: $h_1$([cutlery, eating_utensil], plate), $h_2$([cutlery, eating_utensil], bowl)</td>
<td>none</td>
<td>0.93</td>
</tr>
<tr>
<td>$O_7$, $O_{10}$</td>
<td>0.76</td>
<td>in $O_{10}$: $h_1$([cutlery, eating_utensil], plate), $h_2$([cutlery, eating_utensil], bowl), $h_3$([cutlery, eating_utensil], spoon), $h_4$([cutlery, eating_utensil], glass)</td>
<td>none</td>
<td>0.89</td>
</tr>
</tbody>
</table>
Similarity Between Two Ontologies

Figure 11 Ontology $O_9$.

Figure 12 Ontology $O_{10}$.

Table 2 Coefficient correlation between measures $T_{Ngom}$, $N_{Plus}$ and $N_{Plus}I_{new}$.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{Plus} - N_{Plus}I_{new}$</td>
<td>0.99</td>
</tr>
<tr>
<td>$N_{Plus} - T_{Ngom}$</td>
<td>0.98</td>
</tr>
<tr>
<td>$T_{Ngom} - N_{Plus}I_{new}$</td>
<td>0.99</td>
</tr>
</tbody>
</table>

0.99, correlation($N_{Plus}, T_{Ngom}$) = 0.98 and correlation($T_{Ngom}, N_{Plus}I_{new}$) = 0.99. With $N_{Plus}I_{new}$ measure, the correlation is better with $N_{Plus}$ and $T_{Ngom}$ measures. We can say that $N_{Plus}I_{new}$ is central with respect to the two measures $T_{Ngom}$ and $N_{Plus}$.

5 Conclusion

In this paper, we have proposed a new coefficient of integrity $I_{new}$ for the improvement of $N_{Plus}$ measure proposed in [5]. In [5], the coefficient of integrity $I$ introduced in $N_{Plus}$ measure decreases and tends to 0 fastly. When we add just some concepts, the integrity becomes very poor. To fix this problem, we introduced $R$, the coefficient of representativeness of concepts added in the ontology for its extension. $I_{new}$ decreases slowly compared to $I$ and depends to the cardinality of the ontology extended and the number of concepts added to it too. Figure 4 shows how $I_{new}$ decreases slowly relative to $I$. For simulations, we used wordnet and compared $N_{Plus}$, $T_{Ngom}$ and $N_{Plus}I_{new}$. $N_{Plus}I_{new}$ is the measure which uses $I_{new}$ and $N_{Plus}$ uses $I$. The results of simulations show that $N_{Plus}I_{new}$ improves $N_{Plus}$ thanks to $I_{new}$. The assessing of the coefficient of correlation between the three measures gives as results, good correlation between measures. According to the coefficient of correlation, we remark that $N_{Plus}I_{new}$ is central with respect to the two measures $T_{Ngom}$ and $N_{Plus}$. In our future works, we will propose a methodology for extracting the smaller sub ontology of an ontology which represents a large domain. This sub ontology will be used for assessing similarity between ontologies because with a large ontology, the set of concepts not shared by ontologies will be very large and the similarity will be very poor.

References


