Measurement 118 (2018) 73–82

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

A mathematical evaluation for measuring correctness of domain ontologies using concept maps

Rizwan Iqbal[a,b,*], Masrah Azrifah Azmi Murad[b], Layth Sliman[c], Clay Palmeira da Silva[d]

[a] Department of Computer Engineering, Bahria University, Karachi Campus, Karachi, Pakistan
[b] Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[c] EFREI, Engineering School of Information and Digital Technologies, 30 Avenue de la République, 94800 Villejuif, France
[d] Université François Rabelais Tours 30 Avenue du Monge, bâtiment L, Tours, France

A R T I C L E I N F O

Keywords:
Ontology engineering
Concept mapping
Ontology evaluation
Closeness index
Similarity index

A B S T R A C T

There is a need for further research in the area of ontology evaluation specifically dealing with ontology development exploiting concept maps. The existing literature on ontology evaluation primarily emphasis on ontology formalisation as well as on performing logical inferences, which is usually not directly relevant for concept maps as they are commonly exploited as communication instruments for learning purposes. Commonly used techniques for evaluating concept maps for knowledge assessment may be adopted for a kind of criteria-based evaluation of a domain concept map with respect to a particular aspect. However, this makes its validity limited to a particular aspect or criteria. This paper presents a mathematical ontology evaluation technique to measure the correctness of domain ontologies engineered using concept maps. It is based on the notion of merging two different mathematical measures, namely closeness index and similarity index to come up with a combined index that takes different criteria or aspects into account while performing ontology evaluation. Therefore, the proposed technique makes the evaluation process more reliable and robust. Two case studies were conducted employing the proposed technique for evaluating two different domain ontologies that were engineered using concept maps. Calculations and results from the case studies showed that depending on the correctness of individual ontology, different values of combined Index was calculated manifesting the measure of correctness of each individual ontology in a quantifiable form. Moreover, the results depict that the technique provides in-depth evaluation, it is easy to adopt, requires no special skills, and is conveniently replicable.

1. Introduction

The term ontology was adopted from philosophy and recently it is commonly used in different fields, including computer science. One of the earliest definition of ontology in computer science was presented by Gruber [1], he defined ontology as ”an explicit specification of a conceptualization”. It is reported to be the most accepted and frequently quoted definition in the ontology community [2]. In the last two decades, ontologies have gained significant attention and they are being extensively used in many different fields like, knowledge management, artificial intelligence, information retrieval, natural language processing, e-Commerce, information integration, e-learning, database design, geographical information systems and etc [3–6].

Ontologies are considered as essential parts of intelligent information systems, where they are utilized by knowledge engineers to come up with problem solving approaches and reasoning services. Ontology is a declarative piece of knowledge which can be reused as well as shared. Moreover, they are an effective way to share and disseminate knowledge. They play an important role to envision and materialize the vision of the Semantic Web [7]. Furthermore, these structures of knowledge are machine understandable and machine readable in nature.

Ontology engineering is the discipline that particularly explores the methodologies, approaches and tools employed for developing and maintaining ontologies [8]. Due to constantly growing popularity and applicability of ontologies over the years, several methodologies have been proposed to date for facilitating the ontology development process. In the pursuit of bringing improvement in the ontology development process, the practitioners of the field have supported the notion of merging different methodologies and techniques together. This opens an avenue for new and different design ideas, making the ontology development process more effective [3,9,10].

Initially, concept maps were used in the field of education and learning but soon they got popular in other areas because of their

* Corresponding author at: Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
E-mail address: mail.rizwaniqbal@yahoo.com (R. Iqbal).

https://doi.org/10.1016/j.measurement.2018.01.009
Received 30 October 2015; Received in revised form 21 October 2017; Accepted 8 January 2018
Available online 09 January 2018
0263-2241/

flexible and intuitive nature. The effectiveness and usefulness of concept maps in various areas have been validated through several experiments [11–13]. Concept maps can be exploited in different ways for enhancing learning. They enable learners to give a structure to their thinking [14]. In addition, they facilitate learners in memorizing and deeply understanding information [13,15,16]. A concept map can be defined as a simple graphical representation to show meaningful relationships between two or more concepts linked together [17]. The relationships can be effectively represented by using linking words which form a semantic unit. For instance, if there are two concepts, ”car” and ”driver”, they will be connected to each other using the linking word ”drives”. The concepts are usually represented as boxes or circles and the relationships are represented using lines or arcs.

In the recent years, the use of graphical knowledge representation formalisms carry high potential in the context of ontology engineering and knowledge management domains [18]. Bearing the benefits of graphical knowledge representation formalisms in mind, concept maps have been explored and used for facilitating different aspects of ontology development.

2. Ontology engineering with concept maps

As mentioned earlier, concept maps are able to represent meaningful relationships between concepts linked by words to from a semantic unit. The concepts are included in circles or boxes while relations between concepts are represented by links connecting the boxes. The links are labeled, describing the relation between two concepts. Propositions result from the phrases composed by the concepts and the link label. Apart from having flexible, intuitive and adaptive nature, concept maps have also been reported to have structural resemblence with hierarchal structure of ontology [18,19]. For this reason, they can be effectively exploited as a knowledge acquisition tool, and as an intermediate representation for developing and visualizing ontologies. This is the major factor for their widespread acceptance and popularity in the field of ontology engineering.

The work reported by [20,21] allow experts to update knowledge in a knowledge base by using graphical representation. These systems demonstrated encouraging results. However, prior training by users is required before using the system and the target group is limited to experts. The work presented in [22] uses concept maps for developing ontologies in OWL-DL. This work uses concept maps as a means of graphically representing the ontology. However, the knowledge engineer has to handle the syntax errors at his end.

The systems reported in [23,24] provide users an environment for collaborative construction of ontologies. These systems make use of heuristics for transforming URI’s and use English like terms instead of using OWL operators. Although they empower users to develop ontologies collaboratively, no details regarding mechanism for argumentation and conflict resolution is mentioned.

The work presented in [25,17] reported that concept maps can be used for developing ontologies in a collaborative environment by experts, spread at distinct remote locations. It facilitates sharing of knowledge and bridges the communication gap between experts and knowledge engineers. However, issues regarding manual extraction of terms and ontology evaluation needs to be addressed.

Similarly, the tool proposed in [26] is a step ahead in the same direction. The tool automatically converts the concept maps to OWL ontologies. The experts create the concept maps and then a knowledge engineer uses the tool to perform conversion of maps to ontologies. Finally, the resulting ontologies from the tool are inputted in another tool for performing ontology alignment. However, this system needs a knowledge engineer to perform the conversion of concept maps to ontologies.

Another system presented in [27] employs the notion of controlled natural language (Rabbit) for guiding experts for ontology development. This system facilitates users for developing ontologies in different ways. For instance, it prompts errors when using controlled language and allows tracking classes that have been mentioned. However, similar to systems presented in [20,21], this system also requires users to undergo training.

In order to overcome the need of a knowledge engineer for performing conversion of concept maps to ontologies another method was proposed in [28]. This method uses concept maps as a means of defining domain knowledge by experts, followed by an application that analyzes the concept maps using a set of questions. These questions are based on hierarchy of the map and keywords used by the expert. The answers to these questions lower the map ambiguity and allow the expert to further improve the map. Although the method reduces the need for a knowledge engineer, issues like collaborative construction and ontology evaluation still remains unaddressed.

Existing literature shows that most current approaches employing concept maps are more likely to be used for facilitating a particular aspect of ontology development rather than taking into account the whole ontology development process. For instance, ontology evaluation is considered as one of the essential development oriented activities in the whole ontology development process [8]. Its main goal is to assess that whether the implemented ontology complies with the requirements, specification or criteria for which it was anticipated to satisfy. Ontology evaluation leads to identifying vulnerabilities in the implemented ontology resulting in modification or refinement in the implemented ontology.

Current researches in the field of ontology discuss different evaluation techniques or approaches and defines different quality criteria including coverage, consistency, completeness, computational efficiency [29–32]. Depending on the ontology type and adopted methodology for ontology development different evaluation methods will be suitable [29,33,34]. Therefore, there is an absence of one widely established ontology evaluation technique or approach. The existing literature on ontology evaluation primarily emphasis on ontology formalisation as well as on performing logical inferences, which is usually not directly relevant for concept maps that are commonly exploited as communication instruments for learning purposes [35]. There is a need for further research in the area of ontology evaluation specifically dealing with ontology development exploiting concept maps [35]. Some commonly used techniques for evaluating concept maps for knowledge assessment have been based on scoring methods [36–39]. Such techniques may be adopted for a kind of criteria-based evaluation of a domain concept map with respect to a certain aspect. However, most of such scoring techniques focus more on structural aspects and remain insufficient for giving full evidence of validity of a concept map representing a knowledge domain [35]. Moreover, some more recent techniques presents propositions of a concept map as a correct-incorrect discrimination task which may also employ weights based on the correctness of propositions [40–42]. Such techniques focus more on the content which is in line with the notion of semantic evaluation as they detect to what degree a created ontology reflects the knowledge of a domain.

However, there is a need to come up with a more robust technique for ontology evaluation as different techniques focus more on a particular aspect of the concept map. Therefore, in order to address the aforementioned gap, a mathematical ontology evaluation technique for measuring correctness of domain ontologies engineered using concept maps is proposed that takes different criteria or aspects into account while performing ontology evaluation.

Apart from concept maps and ontologies, in many other application domains like social network studies, and chemicals the data comprising of strings (representing binary variables) are used to show either the presence or absence of certain attributes. In these domains, similarity coefficients and scoring methods are largely exploited for computing similarities between different types of data [43]. Moreover, in the recent years, rank aggregation methods have been employed for aggregating results from gene expression microarray studies in the

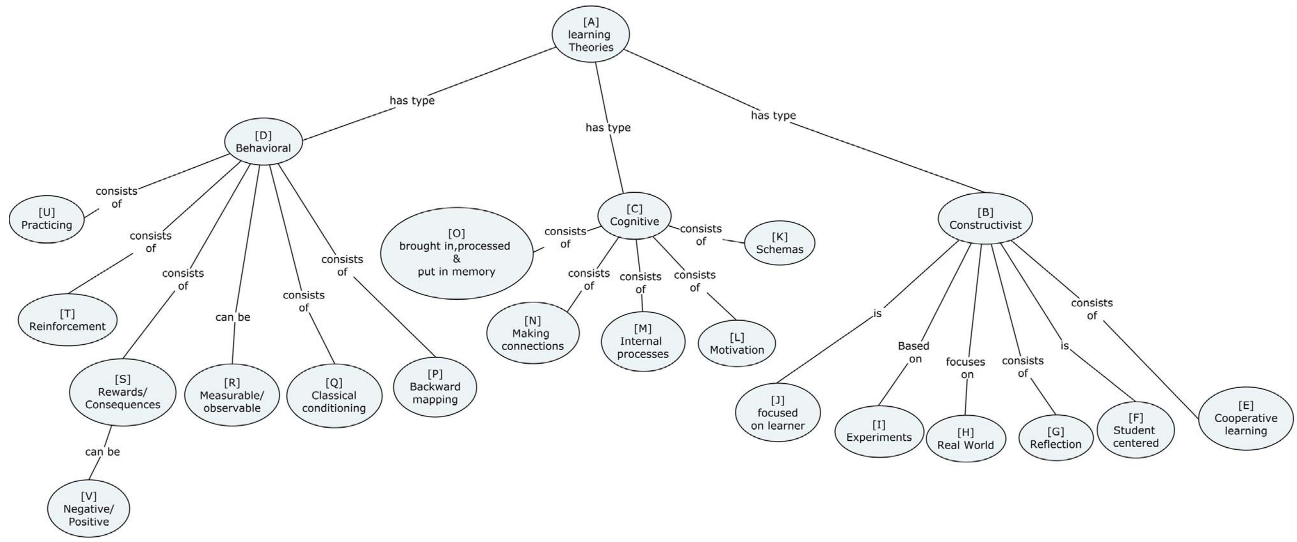

Fig. 1. Expert map in node form.

biological domain [44].

Rank aggregation methods can be broadly divided into three categories: distributional based [45], heuristic [46], and stochastic optimization algorithms [47]. Thurstone's model [48] is an example of distributional based algorithms [49] and over the years many extensions have also been proposed to it. For the category of heuristic algorithms, two types of heuristic algorithms, namely Borda's methods [50] and Markov chain based methods are well known [46]. In the case of stochastic optimization algorithms adapting cross entropy Monte Carlo (CEMC) approach have demonstrated efficient results [51].

3. Proposed technique

A mathematical evaluation technique for measuring correctness of domain ontologies engineering using concept maps is proposed herein. Correctness of a domain ontology can be described as the level of resemblance or similarity between the conceptualization (defined by domain experts) for a particular domain and the ontology developed for that specific domain. The technique combines two different measures, namely closeness Index [39] and similarity index [42]. Both the indexes take different aspects into consideration while calculating the values. The closeness index takes into consideration the closeness or structural resemblance in between the concept maps, while the similarity index is a weighted approach based on the correctness of propositions of the concept maps. The combining of two different measures makes the evaluation process more reliable and robust. The technique is mathematical as well as graphical in nature and does not require any particular skills or expertise for performing evaluation. It can be effectively performed under the supervision of domain experts. Evaluating ontologies using the proposed technique results in more reliable and authentic domain ontologies. The evaluated ontologies can be distributed and reused with confidence.

In the proposed technique, the expert map refers to a concept map for a particular domain which is created under the supervision of domain experts and knowledge engineers. It is created before the domain ontology is being implemented. The final map refers to a concept map which represents the implemented ontology in a map form. It is created after the ontology is implemented using an ontology editor. The transformation from implemented ontology to a map form can be accomplished by the aid of plugins (e.g. Ontograf plugin in Protégé ontology editor). These plugins generate diverse combination of graphics manifesting the concepts, instances, relationships, and attributes of the implemented ontology. The graphical output represent the implemented ontology in a form which is identical to concept maps. Therefore, this enables the final map to be easily compared with the expert map.

The first step is to calculate the closeness index. For calculating the closeness index, the expert map should be considered as $G_e = (V_e, E_e)$ where V_e and E_e are the set of nodes and relationship links in the map, respectively. Similarly, the final map which represents the implemented ontology (derived from graphics manifesting the concepts, instances, relationships, and attributes of the implemented ontology) is represented as $G_o = (V_o, E_o)$. To compare the maps, at first each of them is searched for nodes that are connected to each node n from $V = V_e \cup V_o$. The sets of such nodes are represented as $N_n^{(E)}$ and $N_n^{(O)}$. After the sets of adjacent nodes for a given node are determined, the intersection of the two sets ($I_n = N_n^{(E)} \cap N_n^{(O)}$) and their union ($U_n = N_n^{(E)} \cup N_n^{(O)}$) is determined.

Now that we have the I_n and U_n, the closeness index for any node n is defined as $C_n = \frac{|I_n|}{|U_n|}$. After the closeness indexes for all nodes in the two concept maps are calculated, the closeness index of the two concept maps can be defined as:

$$C(G_e, G_O) = \frac{1}{|V|} \sum_{i \in V} C_i, \text{ where } V = V_e \cup V_o. \ \ 0 \le C \le 1 \tag{1}$$

The next step is to calculate the similarity index. For calculating the similarity index, let $G_e = (V_e, E_e)$ be an expert map. If $(v_i, v_j) \in V_e$ and $e_{ij} \in E_e$, then (v_i, e_{ij}, v_j) represents a proposition in G_e if the relation link e_{ij} connects two concept nodes v_i, v_j. Any proposition (v_i, e_{ij}, v_j) of the expert map can be compared with the propositions in any other concept map . Based on this idea, the propositions in the expert map can be compared with the propositions in the final map. From the resulting comparison it can be decided that if a proposition (v_i, e_{ij}^*, v_j) in the final map is correct, partially correct, or incorrect. The following procedures show that how the comparison can be performed.

1. If there is a proposition (v_i, e_{ij}^*, v_j) in the final map, then
   (a) If $e_{ij}^* = e_{ij}$, (v_i, e_{ij}, v_j) is correct
   (b) If $e_{ij}^* = \phi$, (v_i, e_{ij}, v_j) is partially correct
   (c) If $e_{ij}^* \neq e_{ij}$, (v_i, e_{ij}, v_j), shows a misconception about (v_i, e_{ij}, v_j)
2. If there is a proposition (v_j, e_{ij}^*, v_i) in the final map, then
   (a) If $e_{ij}^* = e_{ij}$ or $e_{ij}^* \neq \phi$, (v_i, e_{ij}, v_j) is partially correct
   (b) If $e_{ij}^* \neq e_{ij}$, (v_i, e_{ij}, v_j), shows a misconception about (v_i, e_{ij}, v_j)
3. If the proposition (v_i, e_{ij}^*, v_j) or (v_j, e_{ij}^*, v_i) does not exist at all in the compared map then (v_i, e_{ij}, v_j) is completely neglected or absent

In order to quantify similarity between final map and expert map, propositions can be scored according to the correctness of the

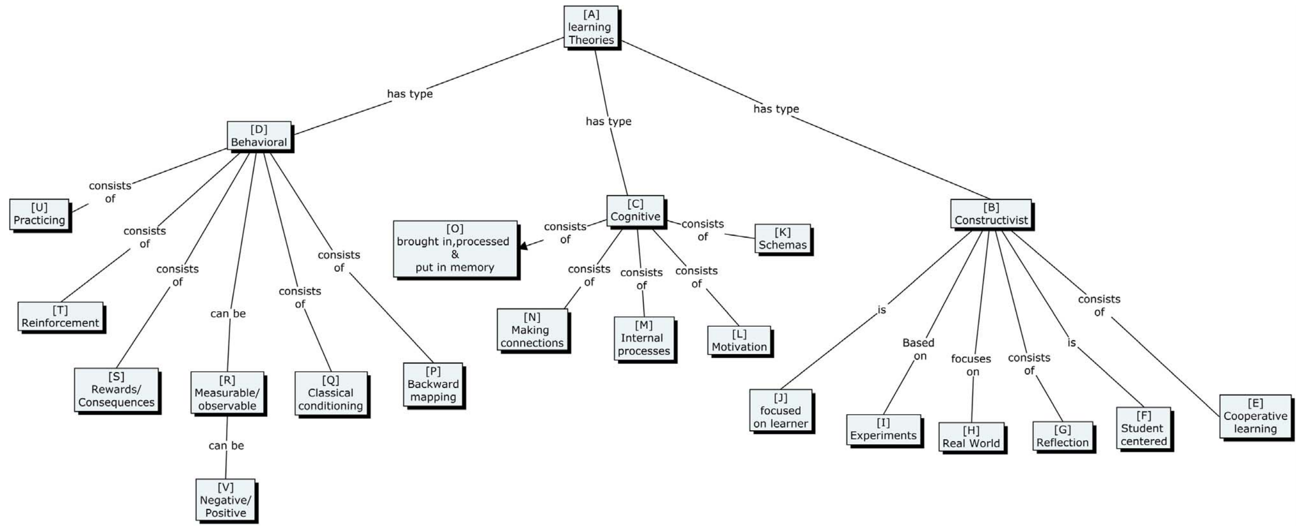

Fig. 2. Final map in node form.

propositions. If proposition $v_{pi}^{*} = (v_i, e_{ij}, v_j)$ existing in the final map is correct, v_{pi}^{*} is scored by the weight of the corresponding proposition defined in the expert map. If proposition v_{pi}^{*} in the final map is partially correct, v_{pi}^{*} it is scored by half the weight of the expert map proposition. If v_{pi}^{*} does not belong to either of the types mentioned above, this proposition receives a zero score.

By applying the aforementioned principles, we can define $score(v_{pi}^{*})$ as the score assigned to any proposition v_{pi}^{*} in the final map. The formula for calculating the $score(v_{pi}^{*})$ is one of the following three conditions. Assume that v_{pi}^{*} is a proposition in the expert map and $W(v_{pi}^{*})$ is its weight.

1. If v_{pi}^{*} is a correct proposition, $score(v_{pi}^{*}) = W(v_{pi}^{*})$
2. v_{pi}^{*} is a partially correct proposition, $score(v_{pi}^{*}) = \frac{1}{2}W(v_{pi}^{*})$
3. If v_{pi}^{*} is neither a correct proposition nor a partially correct proposition, $score(v_{pi}^{*}) = 0$

After calculating all the scores for the final map, similarity index can be calculated using the following equation.

Table 1
Calculations for closeness index.

n	$N_n^{(E)}$	$N_n^{(O)}$	I_n	U_n	C_n
A	{B, C, D}	{B, C, D}	{B, C, D}	{B, C, D}	1
B	{A, E, F, G, H, I, J}	{A, E, F, G, H, I, J}	{A, E, F, G, H, I, J}	{A, E, F, G, H, I, J}	1
C	{A, B, D}	{A, B, D}	{A, B, D}	{A, B, D}	1
D	{A, B, C}	{A, B, C}	{A, B, C}	{A, B, C}	1
E	{B}	{B}	{B}	{B}	1
F	{B}	{B}	{B}	{B}	1
G	{B}	{B}	{B}	{B}	1
H	{B}	{B}	{B}	{B}	1
I	{B}	{B}	{B}	{B}	1
J	{B}	{B}	{B}	{B}	1
K	{C}	{C}	{C}	{C}	1
L	{C}	{C}	{C}	{C}	1
M	{C}	{C}	{C}	{C}	1
N	{C}	{C}	{C}	{C}	1
O	{C}	{C}	{C}	{C}	1
P	{D}	{D}	{D}	{D}	1
Q	{D}	{D}	{D}	{D}	1
R	{D}	{D,V}	{D}	{D,V}	0.5
S	{D,V}	{D}	{D}	{D,V}	0.5
T	{D}	{D}	{D}	{D}	1
U	{D}	{D}	{D}	{D}	1
V	{S}	{S}	{S}	{S}	1

$$S = \frac{\sum_{foralli,j} score(v_{pi}^{*})}{\sum_{foralli,j} W(v_{pi}^{*})}, \quad 0 \le S \le 1 \tag{2}$$

The final step is to calculate the value of combined index. After the values for both, the closeness index and similarity index are calculated, the combined index can be calculated using the following equation. It is arithmetical average of both the indexes.

$$CI = \frac{(C+S)}{2}, \quad 0 \le CI \le 1 \tag{3}$$

In the above equation, C and S represent the values of closeness index and similarity index, respectively. The combined index takes two different mathematical measures (focusing on different aspects) into consideration rather than considering only one single measure. Therefore, this strengthens the depth and breadth of evaluation performed on the ontology.

4. Case study 1: Evaluation of Instructional Technology ontology

The objective of this section is to present the implementation and results of the proposed technique for ontology evaluation. With this aim, it was applied to evaluate an ontology engineered using concept maps for the domain of Instructional Technology. The development of expert map for the Instructional Technology domain and later evaluation of the Instructional Technology ontology was done under the supervision of domain experts from the Faculty of Education, Universiti Putra Malaysia, Malaysia.

The Instructional Technology ontology was implemented using Protégé ontology editor [52]. It is an open-source ontology editor consisting of a wide range of useful plugins. The final map was derived using the OntoGraf plugin [53] available in Protégé ontology editor. OntoGraf plugin generates diverse combination of graphics manifesting the concepts, instances, relationships, and attributes of the ontology. It displays the graphics in a form which is identical to mind maps or concept maps.

4.1. Calculation of closeness index

Let's consider the expert map for the Instructional Technology ontology to be $G_e = (V_e, E_e)$, where V_e and E_e are the set of nodes and relationship links in the map, respectively. Fig. 1 represents the expert map $G_e = (V_e, E_e)$ in the form of nodes and relationships. Similarly, the final map representing the implemented Instructional Technology ontology is considered as $G_o = (V_o, E_o)$. Fig. 2 represents the final map in

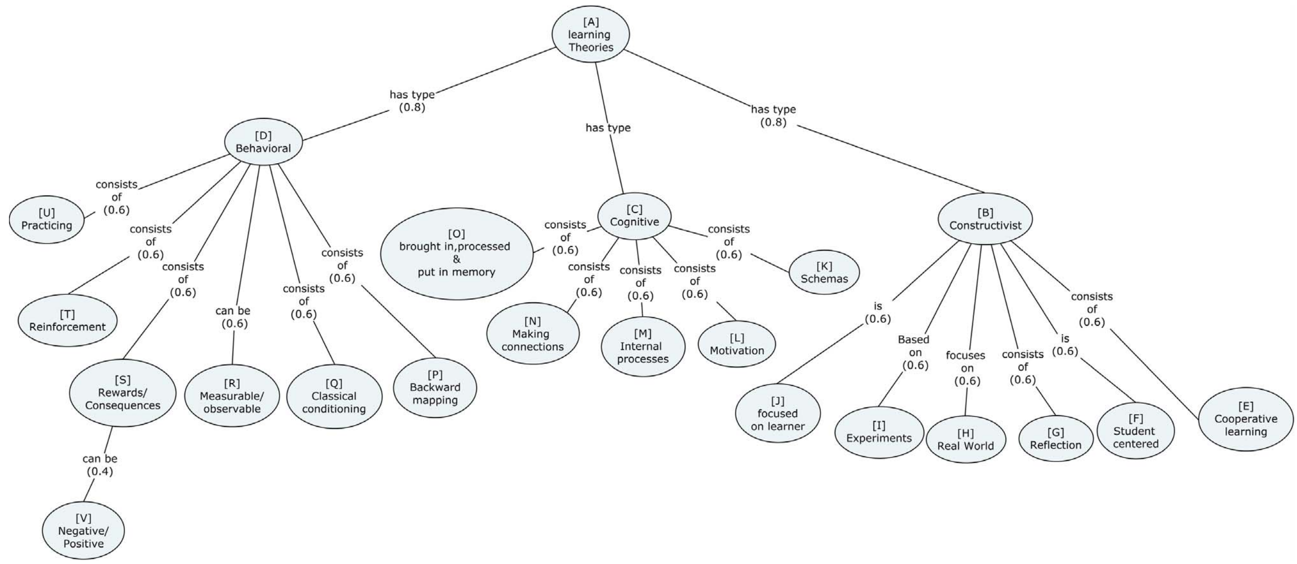

Fig. 3. Expert map with weights.

the form of nodes and relationships.

To compare the maps, at first each of them is searched for nodes that are connected to each node n from $V = V_e \cup V_o$. The sets of such nodes are represented as $N_n^{(E)}$ and $N_n^{(O)}$. After the sets of adjacent nodes for a given node are determined, the intersection of the two sets $(I_n = N_n^{(E)} \cap N_n^{(O)})$ and their union $(U_n = N_n^{(E)} \cup N_n^{(O)})$ is determined. After the values of I_n and U_n have been calculated, the closeness index for any node n is defined as $C_n = \frac{|I_n|}{|U_n|}$. Once the closeness indexes for all nodes in the expert map and final map are calculated, the closeness index for the two maps can be calculated. For the case of Instructional Technology ontology, Table 1 shows the calculations for $N_n^{(E)}$,$N_n^{(O)}$,U_n,C_n, and I_n. By substituting these values in Eq. (1), the closeness index between the expert map and final map is found to be $C(G_e,G_O) = 0.95$. The value of closeness index ranges between 0 and 1. A higher value of closeness index indicates a higher likeness between the maps. Therefore, the value shows that both the maps are not identical to each other with respect to closeness measure.

4.2. Calculation of similarity index

For calculating the similarity index, let $G_e = (V_e,E_e)$ represent the expert map for the case of Instructional Technology ontology. If $(v_i,v_j) \in V_e$ and $e_{ij} \in E_e$, then (v_i,e_{ij},v_j) represents a proposition in G_e and e_{ij} connects two concept nodes v_i,v_j. Any proposition (v_i,e_{ij},v_j) in the expert map can be compared with the propositions in the final map. Therefore, it can be determined that if a proposition (v_i,e_{ij}^*,v_j) in the final map is correct, partially correct, or incorrect. The following procedures were carried out to perform comparison between the expert map and final map for the case study.

1. If there is a proposition (v_i,e_{ij}^*,v_j) in the final map, then
   (a) If $e_{ij}^* = e_{ij}$,(v_i,e_{ij},v_j) is correct
   (b) If $e_{ij}^* = \phi$,(v_i,e_{ij},v_j) is partially correct
   (c) If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
2. If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
   (a) If $e_{ij}^* = e_i j$ or $e_{ij}^* \neq \phi$,(v_i,e_{ij},v_j) is partially correct
   (b) If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
3. If the proposition (v_i,e_{ij}^*,v_j) or (v_j,e_{ij}^*,v_i) does not exist at all in the final map then (v_i,e_{ij},v_j) is completely neglected or absent.

By comparing the final map with expert map, propositions in the final map are scored according to the correctness of its propositions. If proposition $v_{pi}^* = (v_i,e_{ij},v_j)$ existing in the final map is correct, v_{pi}^* is scored by the weight of the corresponding proposition defined in the expert map. If proposition v_{pi}^* in the final map is partially correct, v_{pi}^* is scored by half the weight of the expert map proposition. If v_{pi}^* does not belong to either of the types mentioned above, this proposition receives a zero score.

We define $score(v_{pi}^*)$ as the score assigned to any proposition v_{pi}^* existing in the final map. The formula for calculating the $score(v_{pi}^*)$ is one of the following three conditions assuming that v_{pi}^* is a proposition in the expert map and $W(v_{pi}^*)$ is its weight.

1. If v_{pi}^* is a correct proposition, $score(v_{pi}^*) = W(v_{pi}^*)$
2. If v_{pi}^* is a partially correct proposition, $score(v_{pi}^*) = \frac{1}{2}W(v_{pi}^*)$
3. If v_{pi}^* is neither a correct proposition nor a partially correct proposition, $score(v_{pi}^*) = 0$

Fig. 3 represents the expert map with weights which was developed under the supervision of domain experts. These weights were assigned

Table 2
Total score for propositions in the final map.

Final map propositions (v_{pi}^*)	Proposition Correctness	$score(v_p i^*)$
[node A] has type [node B], has type [node C], and has type [node D]	Correct	0.8 + 0.8 + 0.8
[node B] consists of [node E], is [node F], consists of [node G], focuses on [node H], based on [node I], and is [node J]	Correct	0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6
[node C] consists of [node K], consists of [node L], consists of [node M], consists of [node N], and consists of [node O]	Correct	0.6 + 0.6 + 0.6 + 0.6 + 0.6
[node D] consists of [node P], consists of [node Q], can be [node R], consists of [node S], consists of [node T],and consists of [node U]	Correct	0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6
[node R] can be [node V]	Misconception	0

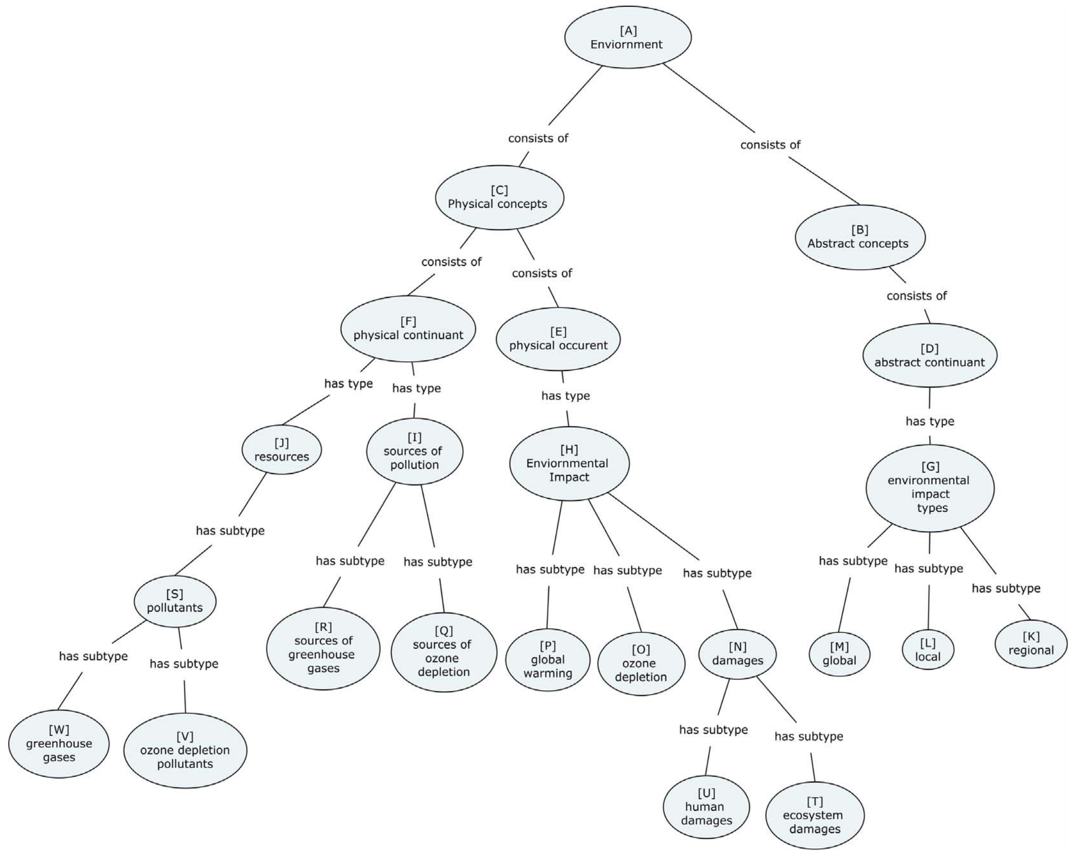

Fig. 4. Expert map in node form.

by the domain experts. The total weight of all the propositions in the expert map is 13. The weights in the final map were calculated under the supervision of domain experts, employing the aforementioned principles. Table 2 shows the total score calculated for the propositions in the final map. The accumulated total of the weights in the final map is found to be 12.6 (0.8 + 0.8 + 0.8 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6 + 0.6).

The value of similarity index ranges between 0 and 1. A higher value of similarity index means a higher resemblance between the maps. By substituting the above calculated values in Eq. (2), the value of similarity index between the two maps is calculated to be $S = \frac{12.6}{13} = 0.96$. Therefore, the obtained value indicates that both the expert map and final map are not completely alike with respect to similarity measure.

4.3. Calculation of combined index

After calculating values for both, the closeness index and similarity index, next step is to calculate the value of combined index. The value of combined index ranges between 0 and 1. A higher value of combined index indicates a higher resemblance between the maps. By substituting the values of closeness index and similarity index in Eq. (3), the value of combined index for both the maps is calculated to be $CI = \frac{(0.95 + 0.96)}{2} = 0.96$. This combined index is arithmetical average of both the indexes. The obtained value shows that the expert map and final map are not identical to each other. The ontology has been robustly evaluated based on a combination of two different mathematical measures that has strengthened the depth and breadth of evaluation performed on the ontology. It also manifests that the evaluated ontology should be subjected to minor modifications before it can be distributed and reused with confidence.

As mentioned earlier, evaluation was performed under the supervision of domain experts. It is worth mentioning that the experts did not report any problem or confusion during any step of the evaluation process. Moreover, the experts expressed that due to the graphical and adaptive nature of concept maps only a simple explanation was enough to carry out the whole evaluation process.

5. Case study 2: Evaluation of Sustainabilty concepts ontology

The objective of this section is to present the implementation and results of the proposed technique for evaluating an ontology engineered using concept maps for the domain of Sustainability concepts. The development of expert map for the Sustainability domain and later evaluation of the Sustainability concepts ontology was done under the supervision of domain experts from Faculty of Engineering, Universiti Putra Malaysia, Malaysia.

The ontology was implemented using Protégé ontology editor [52] and the final map was derived using the OntoGraf plugin [53] available in Protégé ontology editor.

5.1. Calculation of closeness index

Let's consider the expert map for the Sustainability concepts ontology to be $G_e = (V_e,E_e)$, where V_e and E_e are the set of nodes and relationship links in the map, respectively. Fig. 4 represents the expert map $G_e = (V_e,E_e)$ in the form of nodes and relationships. Similarly, the final map representing the implemented ontology is considered as $G_o = (V_o,E_o)$. Fig. 5 represents the final map in the form of nodes and relationships.

To compare the maps, at first each of them is searched for nodes that are connected to each node n from $V = V_e \cup V_o$. The sets of such nodes are represented as $N_n^{(E)}$ and $N_n^{(O)}$. After the sets of adjacent nodes for a given node are determined, the intersection of two sets

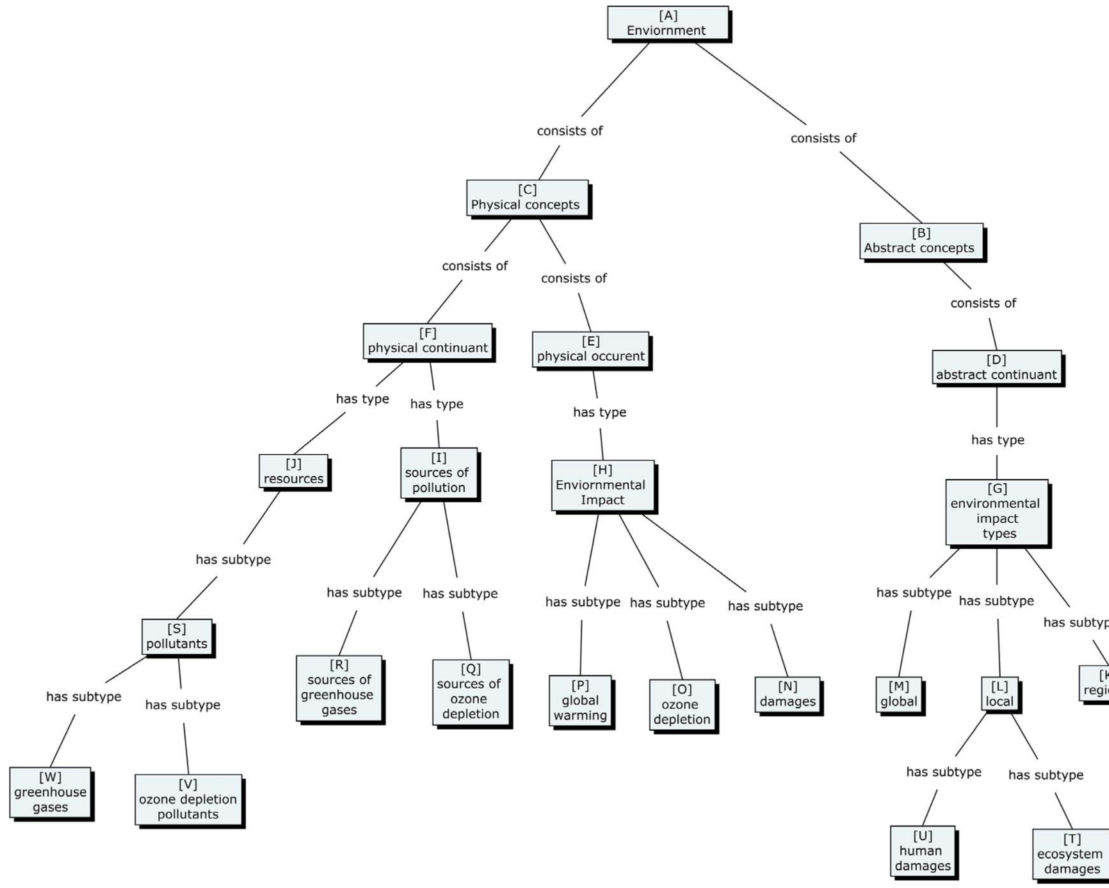

Fig. 5. Final map in node form.

($I_n = N_n^{(E)} \cap N_n^{(O)}$) and their union ($U_n = N_n^{(E)} \cup N_n^{(O)}$) is determined. After the values of I_n and U_n have been calculated, the closeness index for any node n is defined as $C_n = \frac{|I_n|}{|U_n|}$. Once the closeness indexes for all nodes in the expert map and final map are calculated, the closeness index for the two maps can be calculated. For the case of Sustainability concepts ontology, Table 3 shows the calculations for $N_n^{(E)}$,$N_n^{(O)}$,U_n,C_n, and I_n. By substituting these values in Eq. (1), the closeness index between the expert map and final map is found to be $C(G_e,G_O) = 0.84$. Therefore, this value shows that both the maps are not completely identical to each other with respect to closeness measure.

5.2. Calculation of similarity index

For calculating the similarity index, let $G_e = (V_e,E_e)$ represent the expert map for the domain of sustainability. If $(v_i,v_j) \in V_e$ and $e_{ij} \in E_e$, then (v_i,e_{ij},v_j) represents a proposition in G_e and e_{ij} connects two concept nodes v_i,v_j. Any proposition (v_i,e_{ij},v_j) in the expert map can be compared with the propositions in the final map. Therefore, it can be determined that if a proposition (v_i,e_{ij}^*,v_j) in the final map is correct, partially correct, or incorrect. The following procedures were carried out to perform comparison between the expert map and final map for the case study.

1. If there is a proposition (v_i,e_{ij}^*,v_j) in the final map, then
   (a) If $e_{ij}^* = e_{ij}$,(v_i,e_{ij},v_j) is correct
   (b) If $e_{ij}^* = \phi$,(v_i,e_{ij},v_j) is partially correct
   (c) If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
2. If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
   (a) If $e_{ij}^* = e_ij$ or $e_{ij}^* \neq \phi$,(v_i,e_{ij},v_j) is partially correct
   (b) If $e_{ij}^* \neq e_{ij}$,(v_i,e_{ij},v_j), shows a misconception about (v_i,e_{ij},v_j)
3. If the proposition (v_i,e_{ij}^*,v_j) or (v_j,e_{ij}^*,v_i) does not exist at all in the final map then (v_i,e_{ij},v_j) is completely neglected or absent.

By comparing the final map with expert map, propositions in the final map are scored according to the correctness of its propositions. If proposition $v_{pi}^* = (v_i,e_{ij},v_j)$ existing in the final map is correct, v_{pi}^* is scored by the weight of the corresponding proposition defined in the expert map. If proposition v_{pi}^* in the final map is partially correct, v_{pi}^* is scored by half the weight of the expert map proposition. If v_{pi}^* does not belong to either of the types mentioned above, this proposition receives a zero score.

We define $score(v_{pi}^*)$ as the score assigned to any proposition v_{pi}^* existing in the final map. The formula for calculating the $score(v_{pi}^*)$ is one of the following three conditions assuming that v_{pi}^* is a proposition in the expert map and $W(v_{pi}^*)$ is its weight.

1. If v_{pi}^* is a correct proposition, $score(v_{pi}^*) = W(v_{pi}^*)$
2. If v_{pi}^* is a partially correct proposition, $score(v_{pi}^*) = \frac{1}{2}W(v_{pi}^*)$
3. If v_{pi}^* is neither a correct proposition nor a partially correct proposition, $score(v_{pi}^*) = 0$

Fig. 6 represents the expert map with weights which was developed under the supervision of domain experts. These weights were assigned by the domain experts. The total weight of all the propositions in the expert map is 10.2. The weights in the final map were calculated under the supervision of domain experts, employing the aforementioned principles. Table 4 shows the total score calculated for the propositions in the final map. The accumulated total of the weights in the final map is found to be 9.4 (0.8 + 0.8 + 0.6 + 0.6 + 0.6 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0.4 + 0 + 0 + 0.4 + 0.4).

The value of similarity index ranges between 0 and 1. A higher value of similarity index means a higher resemblance between the maps. By substituting the above calculated values in Eq. (2), the value

Table 3
Calculations for closeness index.

n	$N_n^{(E)}$	$N_n^{(O)}$	I_n	U_n	C_n
A	{B, C}	{B, C}	{B, C}	{B, C}	1
B	{A, D}	{A, D}	{A, D}	{A, D}	1
C	{A, E, F}	{A, E, F}	{A, E, F}	{A, E, F}	1
D	{B, G}	{B, G}	{B,G}	{B, G}	1
E	{C, H}	{C, H}	{C, H}	{C, H}	1
F	{C, I, J}	{C, I, J}	{C, I, J}	{C, I, J}	1
G	{D, K, L, M}	{D, K, L, M}	{D, K, L, M}	{D, K, L, M}	1
H	{E, N, O, P}	{E, N, O, P}	{E, N, O, P}	{E, N, O, P}	1
I	{F, Q, R}	{F, Q, R}	{F, Q, R}	{F, Q, R}	1
J	{F, S}	{F, S}	{F, S}	{F, S}	1
K	{G}	{G}	{G}	{G}	1
L	{G}	{G, T, U}	{G}	{G, T, U}	0.333
M	{G}	{G}	{G}	{G}	1
N	{H, T, U}	{H}	{H}	{H, T, U}	0.333
O	{H}	{H}	{H}	{H}	1
P	{H}	{H}	{H}	{H}	1
Q	{I}	{I}	{I}	{I}	1
R	{I}	{I}	{I}	{I}	1
S	{J, V, W}	{J, V, W}	{J, V, W}	{J, V, W}	1
T	{N}	{L}	{Ø}	{N, L}	0
U	{N}	{L}	{Ø}	{N, L}	0
V	{S}	{S}	{S}	{S}	1
W	{S}	{S}	{S}	{S}	1

of similarity index between the two maps is calculated to be $S = \frac{9.4}{10.2} = 0.92$. Therefore, the obtained value indicates that both the expert map and final map are not identical with respect to similarity measure.

5.3. Calculation of combined index

After calculating values for both, the closeness index and similarity index, next step is to calculate the value of combined index. As mentioned earlier, the value of combined index ranges between 0 and 1. By substituting the values of closeness index and similarity index in Eq. (3), the value of combined index for both the maps is calculated to be $CI = \frac{(0.84 + 0.92)}{2} = 0.88$. This combined index is arithmetical average of both the indexes. The obtained value shows that the expert map and final map are not completely identical to each other. It is evident that the proposed technique has strengthen the evaluation process resulting in a quantitative form depicting that vulnerability exists in the implemented sustainability concepts ontology. Therefore, this ontology needs to be modified and corrected before it can be distributed and reused. It is worth mentioning that evaluation was performed under the supervision of domain experts and they did not report any problem or confusion during any step of the evaluation process.

6. Results and discussion

The evaluation results from the two case studies show that both the ontologies are not completely correct. It means that in both cases the expert map and final map are not identical to each other. For the first case study which was for the domain of Instructional Technology, the values of closeness index and similarity index were found to be 0.95 and 0.96, respectively. After calculating these two values, the value of combined index for the Instructional Technology ontology was found to be 0.96. Similarly, for the second case study which was for the domain of Sustainability concepts, the values of closeness index and similarity index were found to be 0.84 and 0.92, respectively. After calculating these two values, the value of combined index for the Sustainability

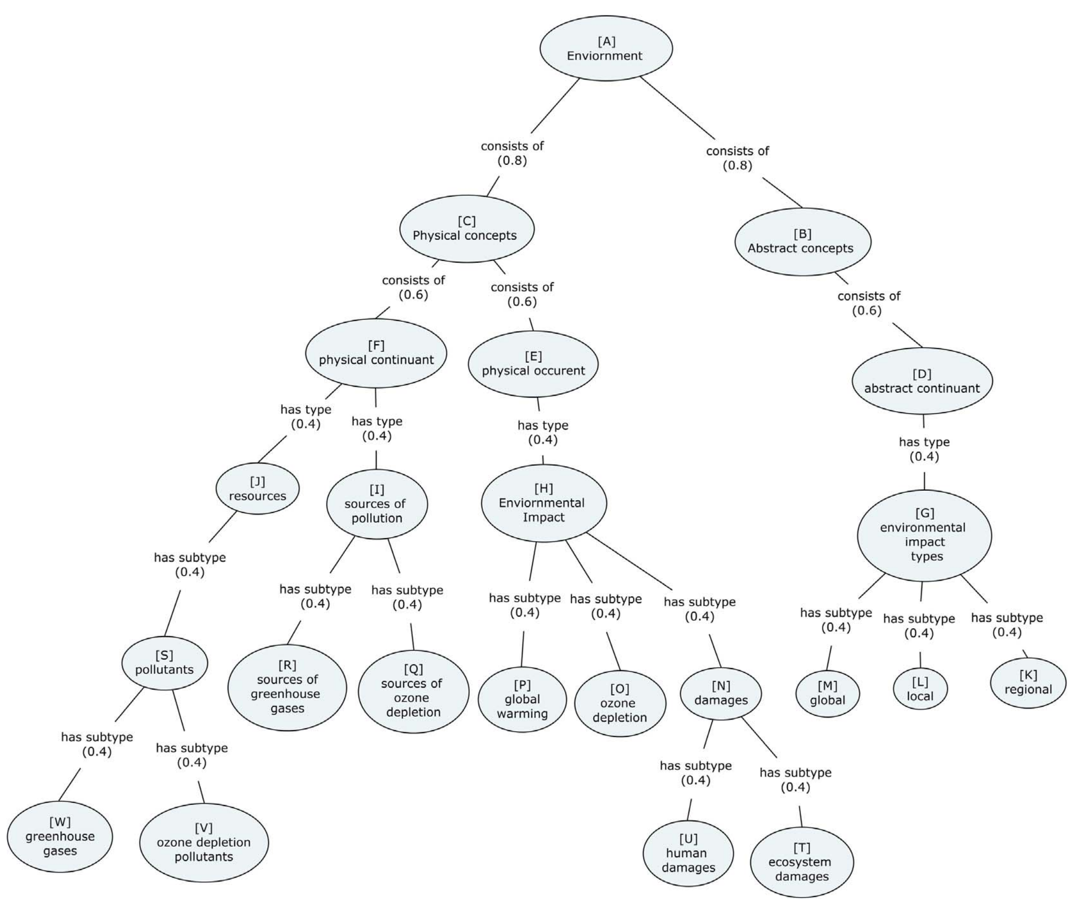

Fig. 6. Expert map with weights.

Table 4
Total score for propositions in the final map.

Final map propositions (v_{pi}^{*})	Proposition Correctness	$score(v_p i^{*})$
[node A] consists of [node B] and [node C]	Correct	0.8 + 0.8
[node B] consists of [node D]	Correct	0.6
[node C] consists of [node E] and [node F]	Correct	0.6 + 0.6
[node D] has type [node G]	Correct	0.4
[node E] has type [node H]	Correct	0.4
[node F] has type [node I] and [node J]	Correct	0.4 + 0.4
[node G] has subtype [node K], has subtype [node L], and has subtype [node M]	Correct	0.4 + 0.4 + 0.4
[node H] has subtype [node N], has subtype [node O], and has subtype [node P]	Correct	0.4 + 0.4 + 0.4
[node I] has subtype [node Q], and has subtype [node R]	Correct	0.4 + 0.4
[node J] has subtype [node S]	Correct	0.4
[node L] has subtype [node T], and has subtype [node U]	Misconception	0 + 0
[node S] has subtype [node V], and has subtype [node W]	Correct	0.4 + 0.4

concepts ontology was found to 0.88. By comparing the values of combined index for both the ontologies, it is evident that Instructional Technology ontology is more correct compared to the Sustainability concepts ontology. Therefore, the Instructional Technology ontology should be subjected to minor modifications while the Sustainability concepts ontology requires more modification to make it free from errors.

For instance, if the closeness index technique [39] would have been applied alone then the concept maps would have been compared based on a node matching scheme and focus only on the structure, at first each of them is searched for nodes that are connected to each node n. After the sets of adjacent nodes for a given node are determined, the intersection of the two sets and their union is determined and the semantical aspect would have been compromised.

On the other hand, if only the similarity index technique [42] would have been applied it can be decided from the resulting comparison of the concept maps that if a proposition is correct, partially correct, or incorrect focusing on the semantical aspect. This measure opts for a weighted approach in which resemblance is quantified based on the correctness of the propositions. However, employing this technique alone would have been resulted in neglecting the structural aspect. The rationale behind merging two different measures is due to the fact that both measures take different criteria's or aspects into account while performing the ontology evaluation and their combination results in more robust evaluation.

7. Conclusion

This paper have presented two case studies employing the proposed evaluation technique for two different domain ontologies namely Instructional Technology ontology, and Sustainability concepts ontology. The proposed technique empowered to mathematically measure the correctness of both domain ontologies in a quantifiable form combining two different measures. Results from the first case study showed that the value of combined index for the evaluated Instructional Technology ontology was calculated to be 0.96. On the other hand, results from the second case study showed that the value of combined index for the evaluated Sustainability concepts ontology is 0.88. This combined index is arithmetical average of both the indexes. These values manifest that Instructional Technology ontology is found to be more accurate compared to the Sustainability concepts ontology. However, minor changes have to be incorporated in the former ontology and a more detailed rework should be done in the latter ontology before both these ontologies can be distributed and reused. Furthermore, the technique can be confidently employed in the future to evaluate any domain ontology engineered using concept maps without the hassle of particular technical expertise.

Acknowledgement

This research was supported by Ministry of Science, Technology and Innovation (MOSTI), Malaysia. Grant No: 01-01-04-SF1629.

References

[1] T.R. Gruber, A translation approach to portable ontology specifications, Knowl. Acquis. 5 (2) (1993) 199–220.
[2] O. Corcho, M. Fernández-López, A. Gómez-Pérez, Methodologies, tools and languages for building ontologies. where is their meeting point? Data Knowl. Eng. 46 (1) (2003) 41–64.
[3] R. Iqbal, M.A. Azmi Murad, A. Mustapha, N. Mohd Sharef, An analysis of ontology engineering methodologies: a literature review, Res. J. Appl. Sci. Eng. Technol. 6 (16) (2013) 2993–3000.
[4] I.A. Al-Baltah, A. Ghani, A. Azim, W.A. Rahman, W. Nurhayati, R. Atan, A comparative study on ontology development methodologies towards building semantic conflicts detection ontology for heterogeneous web services, Res. J. Appl. Sci. Eng. Technol. 7 (13) (2014) 2674–2679.
[5] D. Fensel, F. Van Harmelen, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider, Oil: An ontology infrastructure for the semantic web, IEEE Intell. Syst. (2) (2001) 38–45.
[6] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: principles and methods, Data Knowl. Eng. 25 (1) (1998) 161–197.
[7] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, Sci. Am. 284 (5) (2001) 28–37.
[8] A. Gómez-Pérez, M. Fernández-López, O. Corcho, Ontological Engineering: With Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web, Springer-Verlag, 2004.
[9] G. Brusa, M. Laura Caliusco, O. Chiotti, Towards ontological engineering: a process for building a domain ontology from scratch in public administration, Expert Syst. 25 (5) (2008) 484–503.
[10] R. Iqbal, M.A. Azmi Murad, A. Mustapha, N. Mohd Sharef, An ontology engineering approach with a focus on human centered design, 13th International Conference on Intelligent Systems Design and Applications (ISDA), 2013, IEEE, 2013, pp. 86–91.
[11] M. Ruiz-Primo, On the use of concept maps as an assessment tool in science: What we have learned so far, Revista Electrónica de Investigación Educativa 2 (1) (2000) 1–24.
[12] D.L. Darmofal, D.H. Soderholm, D.R. Brodeur, Using concept maps and concept questions to enhance conceptual understanding, in: Frontiers in Education, 2002. FIE 2002. 32nd Annual, 2002, pp. T3A–1.
[13] A. Cicognani, et al., Concept mapping as a collaborative tool for enhanced online learning, Educ. Technol. Soc. 3 (3) (2000) 150–158.
[14] J. Biggs, C. Tang, Teaching for Quality Learning at University, McGraw-Hill-International, 2011.
[15] L. Cardellini, Conceiving of concept maps to foster meaningful learning: an interview with joseph D. Novak, J. Chem. Educ. 81 (9) (2004) 1303.
[16] D.H. Jonassen, Assessing cognitive structure: verifying a method using pattern notes, J. Res. Dev. Educ. 20 (3) (1987) 1–14.
[17] J.D. Novak, A.J. Cañas, The theory underlying concept maps and how to construct them, Tech. rep., Florida Institute for Human and Machine Cognition, 2006.
[18] A. Soares, C. Sousa, Using concept maps for ontology development: a case in the work organization domain, Innovation in Manufacturing Networks, Springer, 2008, pp. 177–186.
[19] K.-K. Chu, C.-I. Lee, R.-S. Tsai, Ontology technology to assist learners navigation in the concept map learning system, Expert Syst. Appl. 38 (9) (2011) 11293–11299.
[20] P. Clark, J. Thompson, K. Barker, B. Porter, Knowledge entry as the graphical assembly of components, Proceedings of the 1st International Conference on Knowledge Capture, ACM, 2001, pp. 22–29.
[21] V. Chaudhri, K. Murray, J. Pacheco, P. Clark, B. Porter, P. Hayes, Graph-based acquisition of expressive knowledge, Engineering Knowledge in the Age of the Semantic Web, Springer, 2004, pp. 231–247.
[22] H. Gómez-gauchía, B. Díaz-agudo, P. González-Calero, Two-layered approach to knowledge representation using conceptual maps and description logics, in:

Proceedings of the First Int. Conference on Concept Mapping, vol. 2, 2004.
[23] P. Hayes, R. Saavedra, T. Reichherzer, A Collaborative Development Environment for Ontologies (CODE), in: Semantic Integration Workshop (SI-2003), Citeseer, 2003, p. 139.
[24] P. Hayes, T.C. Eskridge, M. Mehrotra, D. Bobrovnikoff, T. Reichherzer, R. Saavedra, Coe: Tools for collaborative ontology development and reuse, in: Knowledge Capture Conference (K-CAP), 2005.
[25] A.G. Castro, P. Rocca-Serra, R. Stevens, C. Taylor, K. Nashar, M.a. Ragan, S.-A. Sansone, The use of concept maps during knowledge elicitation in ontology development processes–the nutrigenomics use case, BMC Bioinformat. 7 (2006) 267.
[26] V. Brilhante, G. Macedo, S. Macedo, Heuristic transformation of well-constructed conceptual maps into owl preliminary domain ontologies, in: Workshop on Ontologies and their Applications, WONTO, 2006.
[27] V. Dimitrova, R. Denaux, G. Hart, C. Dolbear, I. Holt, A.G. Cohn, Involving domain experts in authoring OWL ontologies, Lecture Notes in Computer Science, vol. 5318, Springer, 2008, pp. 1–16.
[28] R.R. Starr, J.M. De Oliveira, Concept maps as the first step in an ontology construction method, Inform. Syst. 38 (5) (2013) 771–783.
[29] A. Gómez-Pérez, Evaluation of ontologies, Int. J. Intell. Syst. 16 (3) (2001) 391–409.
[30] L. Obrst, et al., The evaluation of ontologies. toward improved semantic interoperability, in: C.J.O. Baker, K.-H. Cheung (eds.), Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences, 2007.
[31] D. Vrandecic, Ontology evaluation (doctoral dissertation), karslruhe institute of technology, 2010, from <www.aifb.kit.edu/images/b/b5/ontologyevaluation.pdf> (retrieved june 11, 2012).
[32] A. Zouaq, R. Nkambou, Evaluating the generation of domain ontologies in the knowledge puzzle project, IEEE Trans. Knowl. Data Eng. 21 (11) (2009) 1559–1572.
[33] J. Brank, M. Grobelnik, D. Mladenić, A survey of ontology evaluation techniques, in: Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005), 2005, pp. 166–170.
[34] V. Pammer, P. Scheir, S. Lindstaedt, Ontology coverage check: support for evaluation in ontology engineering, in: Proceedings of 2nd Workshop on Formal Ontologies Meet Industry (FOMI), 2006, pp. 123–134.
[35] C.M. Steiner, D. Albert, Validating domain ontologies: a methodology exemplified for concept maps, Cogent Educ. 4 (1) (2017) 1263006.
[36] J.D. Novak, D.B. Gowin, Learning How to Learn, Cambridge University Press, 1984.
[37] M.A. Ruiz-Primo, R.J. Shavelson, Problems and issues in the use of concept maps in science assessment, J. Res. Sci. Teach. 33 (6) (1996) 569–600.
[38] A. Valerio, D.B. Leake, A.J. Cañas, Automatic classification of concept maps based on a topological taxonomy and its application to studying features of human-built maps, Proceedings of the 3rd International Conference on Concept Mapping, 2008, pp. 122–129.
[39] T.E. Goldsmith, P.J. Johnson, W.H. Acton, Assessing structural knowledge, J. Educ. Psychol. 83 (1) (1991) 88.
[40] C.M. Steiner, Beziehung zwischen Concept Maps und Kompetenzstrukturen [Interrelation between concept maps and competence structures] (Unpublished diploma thesis). University of Graz, Graz., na, 2004.
[41] C.M. Steiner, D. Albert, J. Heller, Concept mapping as a means to build e-learning, Adv. Prin. Effect. e-Learn. (2007) 59–111.
[42] K.-E. Chang, Y.-T. Sung, R.-B. Chang, S.-C. Lin, A new assessment for computer-based concept mapping, J. Educ. Technol. Soc. 8 (3) (2005) 138–148.
[43] R. Todeschini, V. Consonni, H. Xiang, J. Holliday, M. Buscema, P. Willett, Similarity coefficients for binary chemoinformatics data: overview and extended comparison using simulated and real data sets, J. Chem. Inform. Model. 52 (11) (2012) 2884–2901.
[44] S. Lin, Rank aggregation methods, Wiley Interdiscipl. Rev.: Comput. Stat. 2 (5) (2010) 555–570.
[45] L.L. Thurstone, A law of comparative judgment, Psychol. Rev. 34 (4) (1927) 273.
[46] C. Dwork, R. Kumar, M. Naor, D. Sivakumar, Rank aggregation methods for the web, Proceedings of the 10th International Conference on World Wide Web, ACM, 2001, pp. 613–622.
[47] S. Lin, J. Ding, Integration of ranked lists via cross entropy monte carlo with applications to mrna and microrna studies, Biometrics 65 (1) (2009) 9–18.
[48] L.L. Thurstone, Rank order as a psycho-physical method, J. Exp. Psychol. 14 (3) (1931) 187.
[49] L.L. Thurstone, L.V. Jones, The rational origin for measuring subjective values, J. Am. Stat. Assoc. 52 (280) (1957) 458–471.
[50] J.C. de Borda, Memoire sur les elections au scrutin, Histoire de l'Academie Royale des Sciences, Paris, 1781.
[51] L.-Y. Deng, The cross-entropy method: a unified approach to combinatorial optimization, monte-carlo simulation, and machine learning, 2006.
[52] Protégé IDE, 2015. <http://protege.stanford.edu>.
[53] OntoGraf Plugin, 2015. <http://protegewiki.stanford.edu/wiki/OntoGraf>.