Daena: International Journal of Good Conscience. 9(3)76-85. Diciembre 2014. ISSN 1870-557X

Expertise Assessment: A Quantitative Approach Using Natural Semantic Networks

Evaluación Del Conocimiento: Un Enfoque Cuantitativo Utilizando Redes Semánticas Naturales

Garza, Sara E. & Torres Guerrero, Francisco

Abstract. Evaluating the knowledge that an individual or group has over a specific domain is an important—yet challenging—task. *Natural semantic networks* have been used to capture the long-term knowledge of a group of subjects with respect to a particular topic, and allow to assess a group's level of expertise by comparing its network against a network collected from experts. While approaches for making this comparison usually rely on a *qualitative appreciation*, our approach is *quantitative*, since it provides a *degree of similarity* between pairs of networks by means of graph theory and information retrieval. We show the feasibility of this approach by comparing a set of networks from different topics; for each pair, one of the networks belongs to a group of students (unknown expertise) and the other belongs to a group of teachers.

Keywords. Bipartite graphs, expertise, feature vectors, natural semantic networks.

Resumen. Una tarea relevante, aunque también retadora, es poder evaluar el conocimiento que un individuo o grupo posee sobre un dominio específico. Las *redes semánticas naturales* han sido creadas para capturar el conocimiento de largo plazo de un grupo de sujetos con respecto a un tema en particular; estas redes permiten conocer el nivel de experiencia (conocimiento) de un grupo al comparar su red contra un red obtenida de expertos. Mientras que esta comparación normalmente se hace de manera cualitativa, nuestro enfoque es *cuantitativo*, puesto que calcula un *grado de similitud* entre pares de redes por medio de la teoría de grafos y la recuperación de información. Mostramos la factibilidad de este enfoque al comparar redes de diferentes temas. Para cada par de redes, una de ellas pertenece a un grupo de estudiantes (experiencia que nos interesa evaluar) y la otra pertenece a un grupo de maestros.

Palabras clave. Experiencia, grafos bipartitos, redes semánticas naturales, vectores de características.

Introduction

Evaluating knowledge and skills is a challenging task. Organizations, for example, aim to hire the best elements for a particular job with a limited number of recruiters, time, and resources. A similar problem arises in schools, where there is a need to accurately assess what students have learnt while also assuring that this is a long-term knowledge (and not only what they have studied the night before the exam). Automation and innovation are thus required in the evaluation process.

Natural semantic networks (which we will also refer to as NSN's) study long-term memory by gathering the socio-cognitive perspective of a group on a given topic. For example, an NSN collected from undergraduate students for the topic "Artificial Intelligence" would reveal what the students know about this subject, and an NSN collected from a programmer on "Object Oriented Programming" would reveal if the programmer is able to construct inter-connected pieces of knowledge from this domain.

Comparison of natural semantic networks allows estimating similarities between groups. Indeed, measuring the distance between the network of a group whose level of expertise is unknown against the network of a group of experts permits to discover the former group's level of expertise. E.g., we could evaluate how far the students are from the teacher, how far an apprentice is from a master or—in general—how far an individual or group is from a desirable level of expertise.

Methods for comparing natural semantic networks conventionally involve a qualitative appreciation or the use of statistical methods to find variability. However, up to now, there are basically no metrics that yield the *degree of similarity* between a pair of networks. We tackle this area of opportunity.

We propose an approach for estimating a group's level of expertise within a particular topic by calculating the similarity between its natural semantic network and the one collected from experts; the latter network is created as a *master architecture*. To calculate similarity, we model each network as a *bipartite graph* and extract a *feature vector* from this graph. Our approach considers both *content* and *structure* from the networks, and it heavily relies on graph theory and information retrieval.

Background

The current section provides basic foundations on natural semantic networks and bipartite graphs.

Natural semantic networks

Natural semantic networks (NSN's), introduced by Figueroa et al. (1981), reflect the knowledge of a population with respect to a topic or domain by gathering responses from a sample group; the former reveals their long-term memory on the topic. To generate a natural semantic network, a set of *participants* (20-40, usually) is given a set of *target concepts* (6-10). For every target concept, each participant must provide a set of individual words that come to mind when that target concept is presented (there is a time limit of 60-90 seconds to write down these words); the words in this set as known as *definers*. Once the definers are given, the participant is expected to also provide a score (using a scale 1-10) for each one of these, where the score represents how important or how close the definer is with respect to the target concept.

When all participants have completed the aforementioned tasks, it is possible to calculate the total score for each definer. This score is known as its *m-value*, and—for a given target concept—only the *k* definers with the highest *m*-value are kept (normally, *k*=10). This top *k* becomes the target concept's *SAM group*, where "SAM" stands for *Semantic Analysis of M-value* (López and Theios, 1992).

Let us also note that a definer could belong to more than one target concept. It is thus possible to have multiple *m*-values and multiple SAM groups for a single definer. The former gives rise to another important metric within the NSN domain: the *f-value*. Such value is simply the number of times that a definer can be found in the network.

A final consideration regards the NSN *connectedness*; in that sense, every SAM group must have at least one definer with an *f*-value higher than 1, i.e. there should not be isolated groups within the network. Every group not complying with this rule is eliminated, and the NSN has to be constantly refined until this criterion is met.

Table 1 shows a fragment of a natural semantic network. As we can see, ECOLOGY and ENVIRONMENT are the target concepts (there are two SAM groups, consequently), and each of these is associated with four definers; the definers *Nature* and *Animals* are present in both SAM groups and this is reflected by their *f*-values. Furthermore, the definers are placed in decreasing order according to their *m*-value.

Table 1. Fragment of an NSN on ecology.

ECOLOGY		
F	Definer	M
1	Recycle	50
2	**Nature**	30
2	**Animals**	20
1	Plants	10

ENVIRONMENT		
F	Definer	M
2	**Nature**	50
2	**Animals**	30
1	Water	20
1	Plants	10

Bipartite graphs

Networks are mathematically represented with *graphs*. A graph consists of a set of *vertices* (entities) and a set of *edges* (connections). When each edge is associated with a weight (numerical label), the graph is *weighted*.

A *bipartite graph* is a graph whose vertex set is composed of two disjoint vertex subsets, and whose edge set only contains edges that join vertices from different subsets. Even though this conception may seem sophisticated, bipartite graphs are actually very common. One of the most representative examples is given by the actor-movie network, in which the first vertex subset is given by actors and the second one is given by movies; an edge exists between an actor and a movie when the former has participated in the latter. Other examples of bipartite graphs include researcher-publication, train-station, and protein-reaction networks (Newman, 2010).

For simplicity, it is common to work with a bipartite graph's *projections*. A projection is a unipartite graph (i.e., a graph with a vertex set from a single type) where one of the vertex subsets is joined *through* the other. To illustrate this concept, consider the previously mentioned actor-movie network. From this bipartite graph, it is possible to extract the actor projection and the movie projection. The actor projection is a graph whose vertices represent actors and whose edges join pairs of actors that have appeared in the same movies (i.e. actors joined through movies); the movie projection, on the other hand, is a graph whose vertices represent movies and whose edges join pairs of movies where the same actors have appeared (i.e. movies joined through actors).

Document similarity with the vector space model

The *vector space model* is, so far, one of the central models for information retrieval (Baeza-Yates and Ribeiro-Neto, 1999). This model views a document as a *bag of words* (a representations where order is not important) and extracts a *weighted feature vector* from this bag, where each vector's length is equal to the vocabulary (unique words) of the whole document collection. The weight for a given word in a particular document indicates how important the word is in that document.

A common metric for calculating similarity between document vectors is the cosine similarity, which (as the name suggests) computes the cosine of the angle between the vectors. This quantity is computed by means of the *dot product* of the vectors and a normalization operation:

$$\cos(a, b) = \frac{a \cdot b}{|a| \times |b|} = \frac{\sum_M w_{i,a} \times w_{i,b}}{\sqrt{\sum_M w_{i,a}^2} \times \sqrt{\sum_M w_{i,b}^2}}, \tag{1}$$

where a and b are the document vectors, M is the length of the vocabulary, and $w_{i,x}$ is the weight for word i in document x (either a or b, in this case). Let us note that a similarity of 1 indicates that the documents are identical, and a similarity of 0 indicates that the documents have no words in common.

Related Work

In the area of NSN comparison, Sanchez et al. (2013) contrast the NSN's of two distinct groups by means of an index that calculates the ratio of common edges with respect to the total amount possible (similar to the Jaccard index). Our work is also inherently related to *graph matching* (Bunke, 2000), which can be exact or inexact. While the first addresses problems related to graph *isomorphisms* (detecting if two graphs are equal), the second attempts to provide the number of operations needed to turn one graph into another (graph edit distance) or a degree of resemblance between graphs (graph similarity). The works by Dehmmer and Emmert (2007) and Qureshi et al. (2007) both extract feature vectors for calculating graph similarity; while the former utilizes vertex degree (i.e. the number of connected edges), the latter uses statistical and symbolic features for object recognition. Meanwhile, the approach by Champin and Solnon (2003) first obtains different mappings for the pair of graphs and then computes similarity with a psychologically-sustained metric. With regard to semantic data similarity, Bergmann and Gil (2011) focus on semantic workflow retrieval by building graphs with different types of vertices and edges.

Methodology

To perform a quantitative pairwise comparison between NSN's, we first consider the creation of a *master architecture* on the topic, which is a network constructed with more time and less restrictions. Using the target concepts from this architecture, the group's network (whose level of expertise is so far unknown) is collected. Both networks (master and unknown) are then modeled as concept-definer bipartite graphs and for each graph the definer projection is extracted. This projection is, in turn, modeled as a vector whose features are vertices and edges; the weight of each feature is calculated from the *f* and *m*-values. The weight vectors are finally compared via cosine similarity. Let us explain each step with more detail.

Our pairwise comparison is between an *expert* network and an *unknown-expertise* network. For constructing the latter, we use the process described in the Background. For constructing the former (expert network), instead of a conventional NSN, we create a *master architecture*. A master architecture, in essence, is a natural semantic network: it has concepts, definers, *f*-values, and *m*-values. However, its construction process is more flexible: there is no time limit for placing the definers, the target concepts can be modified, and several iterations can be carried out for creating a connected network (on the contrary of a conventional NSN, which is usually a *one-shot*). The master architecture, logically, needs to be developed by a set of experts on the topic (at least one).

Once the two networks to be compared are available, we model each one of them as a *weighted bipartite graph*. In this case, one of the vertex subsets is given by *target concepts* and the other vertex subset is given by *definers*; an edge joins a concept with a definer when both belong to the same SAM group. The edge weight corresponds to the *m*-value of the definer in the concept's SAM group. Figure 1 illustrates the bipartite graph for the network presented in Table 1.

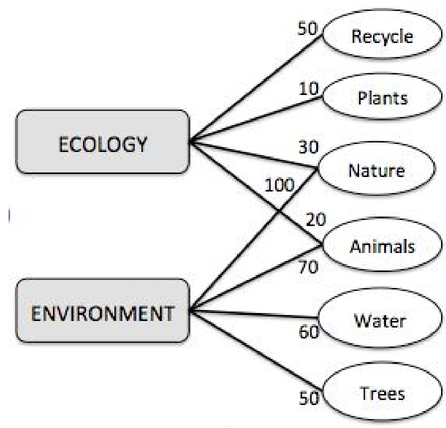


Figure 1. Bipartite representation for the ecology network (see Table 1).

For networks on the same topic, target concepts do not vary. The core of the network (and source for differences) lies, therefore, in the set of definers. It is for this reason that we select to work with the *definer projection* from the concept-definer bipartite graph; as the reader could infer, the vertices of this projection represent definers and the edges represent definers sharing one or several concepts (i.e., being part of the same SAM group). While extracting the definer projection is merely trivial, finding an equivalent for the bipartite weights is not.

To calculate edge weights for the definer projection, we assume that definers are closer or more similar to each other if the difference in their *m*-values is small. As a result, we first compute the relative difference between definers; for a definer d_a and a definer d_b in a SAM group of a concept *c*, this difference would be calculated as

$$\delta_r(d_a, d_b) = \left|\frac{m_c^a - m_c^b}{m_c^{\max} - m_c^{\min}}\right|, \qquad (2)$$

where m_c^a is the *m*-value for d_a in *c*, m_c^b is the *m*-value for d_b in *c*, and $m_c^{\max}$ and $m_c^{\min}$ are, respectively, the maximum and minimum *m*-values of *c*'s group. Since the difference between definers is actually a distance, we obtain the relative similarity by taking the complement of $\delta_r(d_a, d_b)$:

$$\mathrm{sim}(d_a, d_b) = 1 - \delta_r(d_a, d_b). \qquad (3)$$

Because one same pair of definers can appear in several groups, we calculate the overall similarity between d_a and d_b as the average of their relative similarities in the set $C_{a,b} \subseteq C$ of SAM groups that contains both of them. An edge weight is, therefore, calculated with

$$w_{a,b} = \frac{\sum_{c \in C_{a,b}} \mathrm{sim}(d_a, d_b)}{|C_{a,b}|}. \tag{4}$$

Since a weight of 0 typically indicates the absence of an edge, we set $\mathrm{sim}(d_a, d_b)$ as half of the second lowest similarity in c's group when the numerator in Eq. 2 is $m_c^{\max} - m_c^{\min}$. To illustrate these calculations, see Figure 2 (continuation from the ecology example).

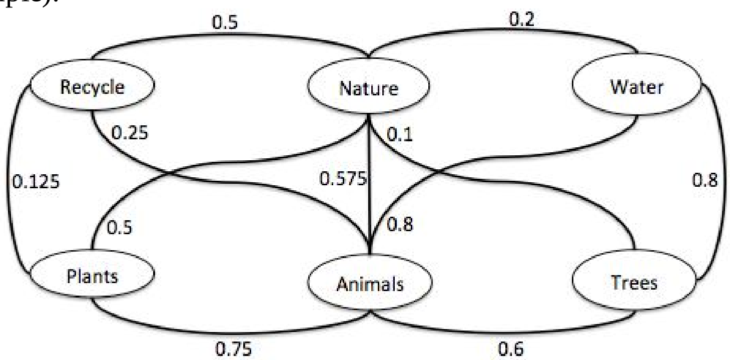


Figure 2. Definer projection for the ecology network (see Table 1 and Figure 1).

To calculate pairwise similarity in NSN's, it is easier to compact the definer graph into a *weighted feature vector*. In this case, the two types of features in the vector are given by vertices and edges. We define the weight of an edge feature *{a,b}* as the edge weight $w_{a,b}$ in the definer graph. For each vertex feature d_a, we define the weight as the *relative f-value* of d_a (which could actually be seen as the centrality of d_a). Let us denote this quantity as

$$\phi_a = \frac{f_a}{f_{\max}}, \tag{5}$$

where f_a is the *f*-value of d_a and $f_{\max}$ is the highest *f*-value found in the network. As an example, consider the "Recycle" definer from the ecology network portrayed in Table 1 (see also Figure 2 for the definer projection). The weight for this vertex feature would be $\frac{1}{2} = 0.5$. For the edge feature "Recycle-Animals", the weight would be $1 - \left[\frac{50-20}{50-10}\right] = 0.25$.

Figure 3 shows the weight vector corresponding to the ecology network.

animals	nature	plants	...	animals_ nature	animals_ plants	animals_ recycle	...
1	1	0.5	...	0.575	0.75	0.25	...

vertices

edges

Figure 3. Weight vector for the ecology network.

Results

To test our approach, we have collected NSN's from two different topics and built a visual similarity matrix (Torres and Garza, 2014). Since networks from the same topic should be more similar to each other and less similar to networks from other topics (regardless of how close or far the unknown-expertise group is from the experts), the similarity matrix is expected to show a clear *block-diagonal*; the former implies that the proposed metric works correctly—at least in a rough sense. Furthermore, we visualized the networks as *tag clouds* to provide a qualitative analysis over their similarities and differences.

For validation purposes, we collected networks in the academic context; we were initially interested on contrasting teacher-versus-student knowledge (other contexts are left for future work). The two selected subjects were *Object Oriented Programming* and *Adaptive Systems Programming* (which is also known as *Artificial Intelligence* or *Intelligent Systems*); both subjects belong to the undergraduate level. Teachers with a PhD on the subject created the master architectures, and the unknown-expertise networks were gathered from students of the same university (UANL) who were taking the courses. Results are shown in Figure 4.

From the visual similarity matrix (presented in Figure 4), there is an outstanding main diagonal and an outstanding block-diagonal. The former is natural, since the similarity between a network and itself is always 1.0—hence the black color. The interesting result lies in the rest of the matrix. As we can see from the illustration, the similarity between networks on the same subject (grayish cells) is higher than the one between networks on different subjects (white cells). In fact, the resulting similarity between the teacher and student networks for *Object Oriented Programming* was 0.2, and the resulting similarity for the *Adaptive Systems Programming* networks was 0.1; in contrast, the similarity between networks of different subjects was, on average, 0.0007—almost a hundred times smaller. However, according to our metric, student knowledge is still far from expert knowledge on these subjects. Let us make a deeper analysis by visualizing the networks.

To visualize the collected networks, we have used *tag clouds*, which highlight frequent terms. In our case, the terms of the cloud are the definers of each network, and their frequency is given by the *f*-value. Our visualizations were constructed using the Wordle tool (http://wordle.net)—a tool that creates aesthetic tag clouds. The visualizations are shown in Figures 5 and 6.

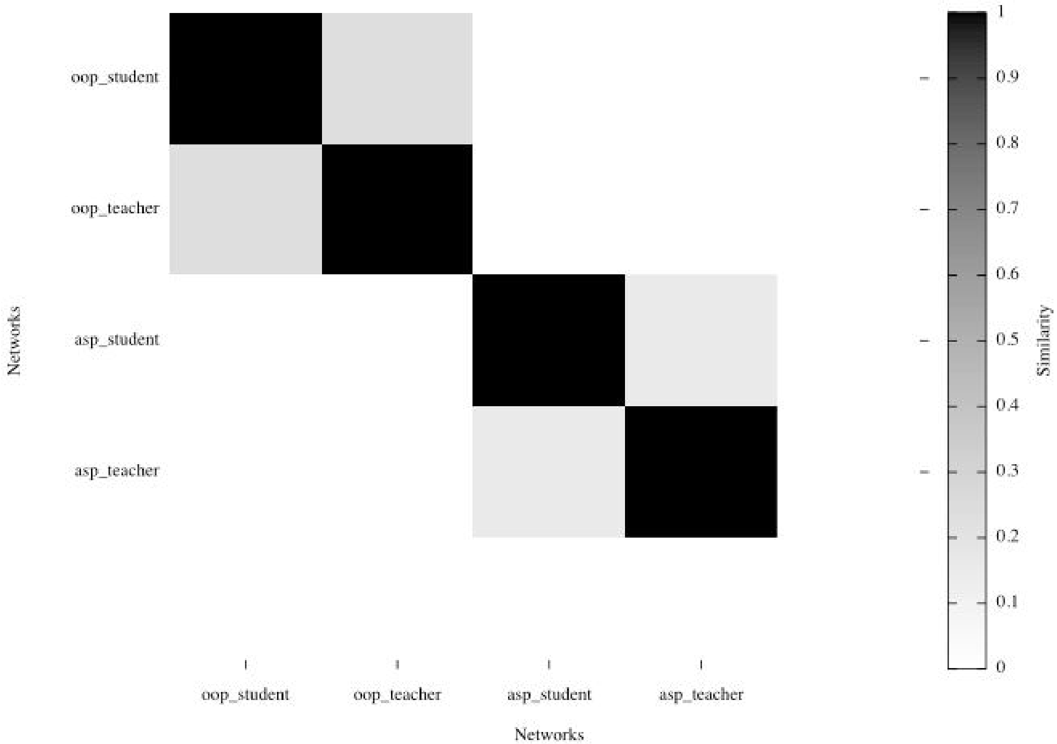
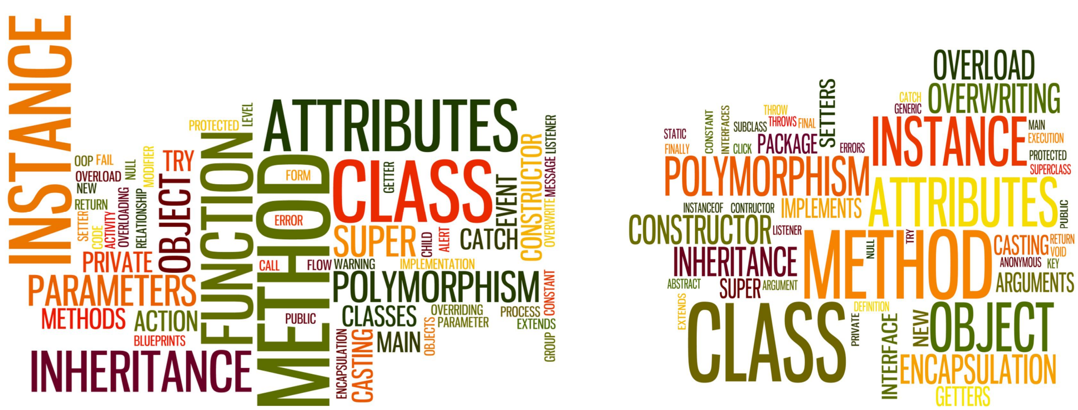


Figure 4. Visual similarity matrix for natural semantic network comparison. The oop_student and oop_teacher are the NSN's for *Object Oriented Programming* (unknown-expertise and master architecture, respectively), and the asp_student and asp_teacher are the NSN's for *Adaptive Systems Programming*.



(a) Student network (b) Master architecture

Figure 5. NSN's for "Object Oriented Programming" (tag clouds).

(a) Student network (b) Master architecture

Figure 6. NSN's for "Adaptive Systems Programming" (tag clouds).

For the *Object Oriented Programming* networks, similarities can be observed more easily, since terms such as "class", "object", "method", "instance", and "polymorphism" are present in both networks and have alike frequencies; however, it is interesting to note that students view OOP more in terms of functions (probably from previous classes on structured programming), and this definer is missing in the master architecture. Also, students are more acquainted with the "parameters" definer, while the master architecture instead contains the "arguments" definer; while these are interchangeable, a teacher might as well explain that the latter term is a synonym of the former, making the overall concept thus clearer for new groups.

For the *Adaptive Systems Programming* networks, the differences are more visible. For instance, it is easy to note that the importance given to terms in common is uneven (e.g. "network" and "system"). In that sense, it was more frequent to find common terms with a low importance (e.g. "butterfly" for chaos theory). We can also note that students tended to concentrate their vocabulary into fewer terms, while the teacher (master architecture) had a wider vocabulary—an effect that is natural, considering that novices usually start with only an overview of the topic in question. A final observation concerns the most frequent terms of both networks; while the students view *Adaptive Systems* as "intelligent algorithms", the teacher views these as "random systems". Appreciating such differences permits to have an implicit feedback on the course and, therefore, make the necessary adjustments.

Discussion

With a similarity metric such as the one presented, it is possible to provide a coarse comparison and establish a general view of how close a pair of networks is. In that sense, our quantitative approach provides an overall insight of the comparison. When this overall insight can be quite useful at an initial stage, a framework with multiple levels of granularity is highly desirable. Also, a more specialized visualization framework is also desirable for aiding a more agile comparison.

Recommendations

We have presented an approach to assess the expertise of a group or individual. The approach consists of collecting the knowledge of this group or individual by means of a *natural semantic network* and then contrasting it against the natural semantic network collected from experts. We have specifically proposed to perform the comparison in a quantitative way, via a *similarity metric*. The pairwise similarity is calculated by modeling the networks as bipartite graphs and then extracting weighted feature vectors from these graphs. We have tested this approach over a set of networks gathered in the academic context (student knowledge vs. teacher knowledge).

There are several lines for future work. One of these concerns the design of a *soft similarity metric*, i.e. a metric able to tolerate messy writing and some degree of ambiguity. Another line concerns the construction of a comparison framework, which we believe could be founded on fuzzy graphs. A third line includes the use of meta-information (e.g. response times and scoring patterns) to carry out the comparison. Finally, the present research was performed in an academic context; it would be interesting to test the approach in an organizational context, i.e. compare junior versus senior workers.

References

Baeza-Yates, R. and Ribeiro-Neto, B. (1999). Modern information retrieval. ACM Press, NY, USA.

Bergmann, R. and Gil, Y. (2011). Retrieval of semantic workflows with knowledge intensive similarity measures. In Case-Based Reasoning Research and Development, pages 17–31. Springer.

Bunke, H. (2000). Graph matching: Theoretical foundations, algorithms, and applications. In Proc. Vision Interface, volume 2000, pages 82–88.

Champin, P. A. and Solnon, C. (2003). Measuring the similarity of labeled graphs. In Case-Based Reasoning Research and Development, pages 80–95. Springer. 86.

Dehmer, M. and Emmert-Streib, F. (2007). Comparing large graphs efficiently by margins of feature vectors. Applied mathematics and computation, 188(2): 1699–1710.

Figueroa, J., G., G., and Solís, V. (1981). Una aproximación al problema del significado: las redes semánticas. Revista latinoamericana de psicología, 13(3):447–458.

Lopez, E. O. and Theios, J. (1992). Semantic analyzer of schemata organization (SASO). Behavior Research Methods, Instruments, & Computers, 24(2): 277–285.

Newman, M. (2010). Networks: An introduction. Oxford University Press, New York, NY, USA.

Qureshi, R. J., Ramel, J., Cardot, H., and Mukherji, P. (2007). Combination of symbolic and statistical features for symbols recognition. ICSCN 2007, pp. 477–482.

Sánchez Miranda, M. P., de la Garza González, A., and Ramírez, E. O. L. (2013). Simulaciones computacionales sobre cuestiones ambientales en dos grupos de contraste. Liberabit, 19(2):223–233.

Sara E. Garza & Francisco Torres Guerrero

Facultad de Ingeniería Mecánica y Eléctrica
Universidad Autónoma de Nuevo León (UANL)
San Nicolás de los Garza, Nuevo León
Mexico
saraelena@gmail.com franciscot@gmail.com