Revising Conceptual Similarity by Neural Networks

Arianna Pavone and Alessio Plebe
Department of Cognitive Science, University of Messina, via Concezione n.6/8, 98122, Messina, Italy
{apavone, aplebe}@unime.it

Keywords: Conceptual Similarity Network, Semantic Pointer Architecture, Neural Engineering Network, Similarity Model.

Abstract: Similarity is an excellent example of a domain-general source of information. Even when we do not have specific knowledge of a domain, we can use similarity as a default method to reason about it. Similarity also plays a significant role in psychological accounts of problem solving, memory, prediction, and categorisation. However, despite the strong presence of similarity judgments in our reasoning, a general conceptual model of similarity has yet to be agreed upon. In this paper, we propose an alternative, unifying solution in this challenge in concept research based on the recent Eliasmith's theory of biological cognition. Specifically we introduce the Semantic Pointer Model of Similarity (SPMS) which describes concepts in terms of processes involving a recently postulated class of mental representations called *semantic pointers*. We discuss how such model is in accordance with the main guidelines of most traditional models known in literature, on the one hand, and gives a solution to most of the criticisms against these models, on the other. We also present some preliminary experimental evaluation in order to support our theory and verify whether similarities derived by human judgments can be compatible with the SPMS.

1 TRADITIONAL MODELS OF SIMILARITY

A model of similarity $\mathcal{M}$ should describe how the elements of a universal set of entities $\mathcal{E}$ are represented and organized in our cognitive system. Based on this representation, given two elements $\text{A}, \text{B} \in \mathcal{E}$, the model provides a way to compute similarity between A and B.[1] Formally the model defines a *similarity function*, $\delta : \mathcal{E} \times \mathcal{E} \rightarrow [0..M]$, which associates to each ordered pair, $(\text{A}, \text{B}) \in \mathcal{E}$, a similarity value $\delta(\text{A}, \text{B})$, where $[0..M] \subset \mathbb{R}$ is the range in which the degree of similarity varies. The value $\delta(\text{A}, \text{B}) = 0$ implies no similarity, while $\delta(\text{A}, \text{B}) = M$ is for the maximum similarity.

In some cases a *dissimilarity function* $\bar{\delta}(\text{A}, \text{B})$ is defined, which is always inversely proportional to the corresponding similarity value. When it is defined on the same range of variability we have $\bar{\delta}(\text{A}, \text{B}) = M - \delta(\text{A}, \text{B})$.

Despite the strong presence of similarity judgments in our reasoning, an accurate model of similarity has yet to be agreed upon. However a number of theoretical accounts of similarity have been proposed in the last decades. See (Holyoak et al., 2012) and (Hahn, 2003) for a detailed survey on similarity models.

One of the most influential models for similarity is the *geometric model* (GMS) (Carroll and Wish, 1974; Torgerson, 1958; Torgerson, 1965), also known as the *mental distance model*, where entities are represented as points in a n-dimensional space and their similarity is inversely proportional to the distance between the corresponding points.

Thus in a GMS which assumes a mental space of dimension n, each entity $\text{A} \in \mathcal{E}$ is represented by a point with n coordinates, $\text{A} = \langle a_1, a_2, .., a_n \rangle$, and the mental distance function $\bar{\delta}$ is a metric in the mathematical sense of the term. This implies that the dissimilarity $\bar{\delta}(\text{A}, \text{B})$, between two entities $\text{A}, \text{B} \in \mathcal{E}$, is associated with the distance of the corresponding points:

$$\bar{\delta}(\text{A}, \text{B}) = \left[\sum_{k=1}^{n} |a_k - b_k|^r \right]^{\frac{1}{r}} \quad (1)$$

where r is a parameter that allows different spatial metrics to be used. The most common spatial metric is the *Euclidean metric* (with $r = 2$) where the dis-

[1] In general, the fact that A is introduced before B can be a relevant factor, since the similarity between A and B may be different from the similarity between B and A if $\mathcal{M}$ admits non-symmetric judgments

tance between two points is the length of the straight line connecting them.

Another common spatial metric is the *city-block metric* (with $r = 1$) where the distance between two points is the sum of their distances on each dimension. It is appropriate for psychologically separated dimensions such as color and shape, brightness and size or any given set of separated measurable perceptions.

The most influential alternative to a GMS is the *feature based model* (FMS), also known as *feature contrast model* or simply *contrast model*, which has been introduced by Tversky in (Tversky, 1977). The underlying idea is that subjective assessments of similarity not always satisfy mathematical rules required by a GMS (Carroll and Wish, 1974).

In a FMS any entity is represented as a collection of features and the similarity between two entities $\mathrm{A}, \mathrm{B} \in \mathcal{E}$ is expressed as a linear combination of the measure of the common and distinctive features. Specifically if we denote by $[\mathrm{A} \cap \mathrm{B}]$ the number of features in common and by $[\mathrm{A} - \mathrm{B}]$ the number of features that are in A and not in B, then the similarity $\delta(\mathrm{A}, \mathrm{B})$ is computed by:

$$\delta(\mathrm{A}, \mathrm{B}) = \alpha[\mathrm{A} \cap \mathrm{B}] - \beta[\mathrm{A} - \mathrm{B}] - \gamma[\mathrm{A} - \mathrm{B}] \quad (2)$$

where α, β, and γ are weights for the common and distinctive components.

One advantage of the feature based model is that it can account for violations in any of the metric distance axioms. For instance it is able to implement asymmetric similarity between entities since β may be different from γ.

The premise that entities can be described in terms of constituent features has been a powerful idea in cognitive psychology and has influenced much work.[2]

In the *alignment based model* (AMS), also known as *structural models*, comparison of two entities is computed not just by matching their features, but by determining how such features correspond to, or align with, one another. One of the most interesting aspects of AMS is that they make purely relational similarity possible (Falkenhainer et al., 1989). Infact matching features have a greater influence on similarity when they belong to parts that are placed in correspondence. However parts tend to be placed in correspondence if they have many features in common (Markman and Gentner, 1993b).

Another interesting aspect of AMS is that it concerns alignable and non-alignable differences (Markman and Gentner, 1993a). Non-alignable differences between two entities are attributes of one entity that have no corresponding attribute in the other entity. Consistent with the role of structural alignment in similarity comparisons, alignable differences influence similarity more than non-alignable differences (Markman and Gentner, 1996) and are more likely to be encoded in memory (Gentner and Markman, 1997).

Extending the alignment based account of similarity, a *transformational model* (TMS) (Wiener-Ehrlich et al., 1980) assumes that a cognitive system uses an elementary set of transformations when computing the similarity between two entities. In this context, if we assume that each entity is described by a sequence of features, the similarity is assumed to decrease monotonically as the number of transformations required to make one sequence of features identical to the other increases.

Thus while the correspondences among features of two entities are explicitly stated in an alignment based model, such correspondences are implicit in a transformational based model. For this reason a potentially fertile research direction is to combine the alignment based account for representing the internal structure of the features of entities with the constraints that transformational accounts provide for establishing a psychologically plausible measure of similarity (Mitchell, 1993).

2 AN ACCOUNT OF SIMILARITY BASED ON THE SPA

Eliasmith's theory of biological cognition (Eliasmith, 2013) introduces *semantic pointers* as neural representations that carry partial semantic content and are composable into structures necessary to support complex cognition. Such representations result from the compression and recursive binding of perceptual, lexical, and motor representations, effectively integrating traditional connectionist and symbolic approaches. In this section we briefly introduce the reader to the semantic pointer theory and describe how they are used to represent concepts in Eliasmith's theory of biological cognition. Then we give our account of similarity based on semantic pointers.

[2]Neural network representations, for instance, are often based on features, where entities are broken down into binary vectors of features where ones identify the presence of features and zeros their absence. In such context the similarity distance value is typically computed using the fairly simple function, i.e. the *Hamming distance*, formally given by $\delta(\mathrm{A}, \mathrm{B}) = [\mathrm{A} - \mathrm{B}] + [\mathrm{B} - \mathrm{A}]$.

2.1 Representing Concepts as Semantic Pointers

In its most basic form a *semantic pointer* can be thought of as a compressed representation of a specific concept, acting as a summary of it. On a cognitive level it can be represented by the activity of a large population of neurons induced by a perceptual stimulus, while mathematically it can be represented as a vector in an n-dimensional space, where the value n is closely related to the size of the population of neurons that give rise to this compressed representation.

Typically, such representations are generated from perceptual inputs. A classic example is what happens with the vision of an object, within one's visual field, although similar representations can be generated through hearing, touch or other perceptual stimulus. In brief, visual perception initially gives rise to the activity of a very large population of neurons on the first level of the visual cortex and will then be encoded by a high-dimensional semantic pointer. Through some transformations the neural populations in the underlying layers of the visual cortex make it possible to obtain increasingly compact representations of the input, providing, to all effects, semantic summaries of the original input. It is therefore possible that at the end of the perception process, through such compression links, a very compact representation of the perceptual input is produced.

This transformation of the input representation is biologically consistent both with the decrease in the number of neurons observed in the deeper hierarchical layers of the visual cortex, and with the development of hierarchical statistical models of neural inspiration for the reduction of dimensionality (Hinton and Salakhutdinov, 2006).

Such representations are referred to as "pointers" since, as the equivalent in computer science, they can be used to target representations at lower levels in the compression network (Hinton and Salakhutdinov, 2006), while are associated with the term "semantic" because they hold semantic information about the states they represent by virtue of being non-arbitrarily related to these states through the compression process.

One of the main characteristics of semantic pointers lies in the possibility of being linked together in more structured representations containing lexical, perceptual, motor or other kind of information. And it is important to keep in mind that any given semantic pointer can be manipulated independently of the network used to generate it, and that the structured representations resulting from such a binding are themselves semantic pointers.

In Eliasmith's theory of biological cognition the binding of two semantic pointers is performed using a process called circular convolution (Eliasmith, 2004; Eliasmith, 2013) and is indicated by the symbol $\circledast$. It is not our interest to go into technical details of this operation in this work and for this reason we can limit ourselves to saying that the circular convolution can be thought of as a function that merges, by binding them together, two input vectors into a single output vector of the same dimension. The result of this binding operation is a single representation that captures the relationships between the input pointers involved in the binding.

For instance, assuming *taste*, *size*, *color*, *salty*, *tiny* and *white* are all semantic pointers representing the corresponding concepts, a representation of the concept "salt" may be given by $salt = taste \circledast salty + size \circledast tiny + color \circledast white$.

The fact that the dimension of the resulting pointer is the same as the dimension of those involved in the binding implies that part of the information contained in the starting pointers has been lost. The pointer obtained from the binding can therefore be seen as a compressed representation of the relationship that binds two or more concepts together.

The binding process can be repeated an indefinite number of times and, more interestingly, it can be reversed (by the unbinding operation) to obtain one of the semantic pointers that have given rise, through the binding, to the compressed representation of a more complex concept. Of course, since part of the information has been lost during the binding process, such reconstruction can only be done in an approximate way: more technically the result of applying transformations to structured semantic pointers is usually noisy.

Returning to the previous example, the concept "salt" is linked to the concept "white" being one of its constituent, binded to the concept of "color". More formally $salt \circledast color^{-1} \approx white$. For the reasons just mentioned, the result is noisy and approximate, however it can be "cleaned up" (Eliasmith, 2013) to the nearest allowable representation by means of a clean-up memory. Mapping a noisy or partial vector representation to an allowable representation plays a central role in several other neurally inspired cognitive models and there have been several suggestions as to how such mappings can be done (Eliasmith, 2013), including Hopfield networks, multilayer perceptrons, or any other prototype-based classifier.

2.2 An Account of Similarity based on the SPA Framework

In this section we introduce the *Semantic Pointer Model of Similarity* (SPMS), a similarity model that associates the degree of similarity of concepts to the distance between the corresponding semantic pointers in a system of mental spaces. This definition brings the new model of similarity very close to the traditional GMS, however it has also some points in common with other classic similarity models.

As before, let $\mathcal{E}$ be a universal set of entities. The SMPS is a similarity model constituted by a system, $\mathcal{S}$, of contextual mental spaces of different size, populated by a set, Σ, of allowable concept representations. In this context we can think of "allowable representations" as those semantic pointers that are associated with concepts in $\mathcal{E}$, but more generally also with sentences, sub-phrases, or other structured representations that are pertinent to the current context. Ultimately, what makes a representation allowable is its inclusion in a clean-up memory.

Using an extended notation we can define a clean-up memory over Σ as a mapping, $\Sigma(v) = x$, which associates any n-dimensional semantic pointer v with its closest allowable representation $x \in \Sigma$. Therefore $\Sigma(x) = x$, for each $x \in \Sigma$.

The set Σ of semantic pointers can be partitioned in two subsets: the set of primitive representations, denoted by Σ^0, and the set of structured representations, denoted by $\Sigma^{\circledast}$. The set Σ^0 includes all semantic pointer which represent elementary concepts, such as *size*, *color*, *big*, *black*, etc. The set $\Sigma^{\circledast}$ includes all semantic pointers which represent structured concepts obtained by binding other pointers, such as *sugar* = *size*⊛*tiny* + *taste*⊛*sweet* + *color*⊛*white*. Formally:

$$\Sigma^{\circledast} = \big\{ v \in \Sigma \mid \quad v = x \circledast y + w, \quad \text{for some } x, y, w \in \Sigma \cup \{\varepsilon\} \big\} \tag{3}$$

For each geometric space $S \in \mathcal{S}$, we indicate by $|S|$ its dimension. Similarly, for each $v \in \Sigma$, we indicate by $|v|$ the size of the semantic pointer v. Each pointer $v \in \Sigma$ lies in a geometric space $S \in \mathcal{S}$. In such a case we say, using an extended notation, that $v \in S$ or that $\sigma(v) = S$. We have $|S| = |v|$ for any $v \in S$.

The coexistence of several geometric spaces of different dimensions allows representations at different levels of detail: larger spaces would host semantic pointers capable of representing complex concepts and at a greater level of detail; mental spaces of smaller dimensions would instead host smaller semantic pointers which refer to a summary of the corresponding concepts. spaces are represented with larger radii. Thus, as in an FMS, any semantic pointer may be a well structured representation obtained by the binding of its constitutive elements, which actually represent their constitutive features. Therefore, the larger the number of such constitutive features, the larger the size of the resulting semantic pointer should be in order to allow the discrimination of information related to the constituent elements. In fact, too much compressed semantic pointers would not be able to hold enough information.

Fig.1 shows an example of a partial cross-section of SPMS with 6 spaces and 13 semantic pointers. It also presents 7 Examples which highlight some of the main features of the SPMS model.

In the SPMS different comparison spaces operate on different contexts of judgment, grouping the semantic pointers on the basis of their semantic category or on the basis of the level of detail through which the concepts are represented, and therefore on the basis of the size of the semantic pointers.

What is particularly important for the definition of the SPMS and which significantly differentiates it from the models presented up to now is the possibility that several representations of the same concept may be present in the model. Specifically in the SPMS the same concept may appear in two or more mental spaces: it can be present in several mental spaces depending on the context in which this concept is evaluated; it can be present in mental spaces of different size depending on the level of detail with which this concept is evaluated. This implies that, unlike what happens in the GMS where the judgement is limited to a single space of comparison, the similarity between two concepts also depends on the context in which this judgment is made and on the degree of detail with which the concepts are evaluated. It may in fact happen that two objects have different measures of similarity between them if they appear (and are compared) in different spaces. For each context, such as appearance, color, taste, etc. the space of comparison may change. For example, *salt* and *sugar* could be judged very similar in terms of shape but different in terms of taste (see Example 3).

It may also happen that two objects have different measures of similarity between them if they appear (and are compared) in mental spaces with different size, representing concepts at different levels of detail. For example, *jaguars* and *leopards* can be considered similar animals to a superficial analysis, but very different in terms of relating behavior, lifestyle and habitat (see Example 4).

Formally we define a mapping $\rho : \Sigma \to \mathcal{E}$, which associates to any $v \in \Sigma$ the concept that it represents, while we indicate by $\vec{V}(\text{A})$ the set of all allowable representations in Σ of a concept $\text{A} \in \mathcal{E}$. Formally, for each $\text{A} \in \mathcal{E}$, we have:

$$\vec{V}(\text{A}) = \{v \in \Sigma \mid \rho(v) = \text{A}\} \quad (4)$$

However, for a concept A, only one representation can be involved in a neural process. Which of the representations in $\vec{V}(\text{A})$ is used may depend on several factors, including the context or the degree of prominence of the concept. Formally, given a concept represented by an input stimulus $\text{A} \in \mathcal{E}$, we indicate by $\sigma(\text{A})$ the semantic vector in $\vec{V}(\text{A})$ which is induced by the concept A.

Semantic pointers may be connected by hierarchical and structural edges. In a SPMS *hierarchical edges* are those links that make up the compression network, connecting semantic pointers of different size representing the same concept with different levels of detail (see Example 1). Such edges can be traversed in both directions allowing to move along the compression network. Formally the set of hierarchical edges L_h is a set of ordered couples of elements in Σ. An edge in L_h is of the form (v_i, v_j), where $|v_i| > |v_j|$ and $\rho(v_i) = \rho(v_j)$. For each $v \in \Sigma$, we have

$$|\{(v,u) \in L_h\}| \leq 1 \text{ and } |\{(u,v) \in L_h\}| \leq 1.$$

In a SPMS *structural edges* connect two semantic pointers, appearing in the same or in different mental spaces, with the semantic pointer obtained by their binding. When the results of the binding is a semantic pointer lying in a different mental space, such edges represent a connection between the spaces of the component pointers to that of the resulting pointer (see Example 2). The set of structural edges L_s is a set of ordered couples of elements in Σ. An edge in L_s is of the form (u, v), where $v = u \circledast x$ for some $x \in \Sigma$. Therefore we have $|v| = |u| = |x|$. Formally:

$$L_s = \big\{(u,v) \in (\Sigma \times \Sigma) \mid \exists\, x \in \Sigma \text{ such that } u \circledast x = v\big\}. \quad (5)$$

Therefore we have $|v| = |u| = |x|$.

When two semantic pointers lie in the same geometric space their similarity can be computed by means of the distance between the corresponding points, as in the case of the GMS. Thus, as happen in a FMS, when the compared semantic pointers are obtained by binding sets of constituent elements with a substantial intersection we may aspect they turn out to be very close to each other within the same geometric space and therefore perceived within the model as very similar. On the other hand, when the semantic pointers are obtained by the binding of very different vectors, it is plausible to think that they are located far within the same geometric space or even in two different spaces.

Computing the distance, or similarity, between representations lying in different mental spaces may require more work than the simple computation of the distance between two points. We argue that in such cases the similarity is related with the shortest path $\tau(x,y)$ which connects the first concept x to the closest representation y of the second concept, in terms of number of hierarchical or structural edges.

More formally, assuming that A and B are two concepts in $\mathcal{E}$, represented by two perceptual stimulus, and $\sigma(\text{A}) = x$ is the semantic pointer induced by the concept A, with $x \in \vec{V}(\text{A})$. The dissimilarity value $\bar{\delta}(\text{A},\text{B})$ is computed as follows:

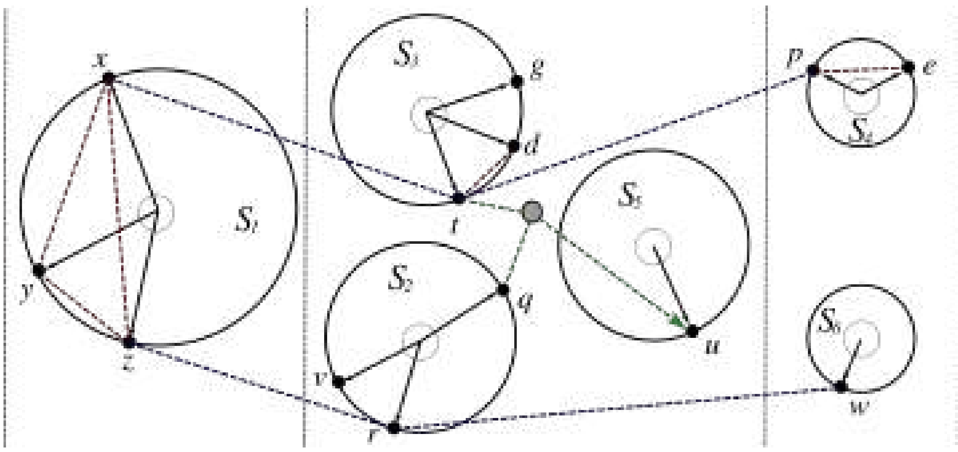
$$\bar{\delta}(\text{A},\text{B}) = \begin{cases} \left[\displaystyle\sum_{k=1}^{n} |x_k - y_k|^2\right]^{\frac{1}{2}} & \text{if } \exists\, y \in \Big(\sigma(x) \cap \vec{V}(\text{B})\Big) \\ \displaystyle\min_{y \in \vec{V}(\text{B})} \Big(\tau\big(x,y\big)\Big) & \text{otherwise} \end{cases} \quad (6)$$

Thus, as occurs in the traditional GMS, the smaller the distance between two semantic pointers within the same mental space, the more similar are the two concepts represented by the semantic pointers. Consequently the farther the distance the less similar they are. The key difference between the SPMS and the traditional GMS lies in the fact that the latter assumes the existence of a single mental space while the SPMS provides for the presence of a system of comparison spaces.

Equation (6) also highlights a common feature between the SPMS and TMS in which similarity is assumed to decrease monotonously as the number of transformations required to make one sequence of features identical to another increases. Even in the new model, in fact, one could imagine how the similarity can decrease in a monotonous way as the number of transformations required to pass from the context to the other increases.

More interestingly, according to equation (6), in the SPMS it seems to matter which of the two concepts is introduced first in the judgment, namely the question "how similar is A to B?" is semantically different from the question "how similar is B to A?". Such aspect is strongly connected with the symmetrical property of the similarity relationship, often criticized in the GMS.

We will cover this and other aspects in more detail in the next section.



Example 1. $x,t,p \in \vec{V}(\text{China})$*, with* $|x| > |t| > |p|$. *The pointer x represents the concept at a high level of detail while p represent the concept at a low level of detail. The edges* (x,t) *and* (t,p) *are hierarchical edges of the compression network.*

Example 2. $t \in \vec{V}(\text{color})$ *and* $q \in \vec{V}(\text{red})$. *The concept* red-color *is represented by a semantic pointer, u, obtained by the binding of t and q, i.e.* $u = t \circledast q$. *We have* $|t| = |q| = |u|$. *The edges* (t,u) *and* (q,u) *are structural edges.*

Example 3. $r,t \in \vec{V}(\text{sugar})$ *while* $q,d \in \vec{V}(\text{salt})$. *Their representations when comparing* salt *and* sugar *in terms of shape* (S_3) *are closer than when comparing the same concepts in terms of taste* (S_2)*.*

Example 4. $x,t \in \vec{V}(\text{leopard})$ *while* $y,d \in \vec{V}(\text{jaguar})$. *Their representations when comparing* leopard *and* jaguar *in short* (S_3) *are closer than when comparing the same concepts in details* (S_1)*.*

Example 5. $x,t,p \in \vec{V}(\text{China})$ *while* $e \in \vec{V}(\text{North Korea})$. *The similarity between* China *and* North Korea *corresponds to* $sp(x,e)$*, while the similarity between* North Korea *and* China *corresponds to the distance between e and p. Therefore* $\delta(\text{China}, \text{North Korea}) < \delta(\text{North Korea}, \text{China})$.

Example 6. $x \in \vec{V}(\text{triangle})$, $y \in \vec{V}(\text{square})$ *and* $z \in \vec{V}(\text{rectangle})$. *All such pointers lie in the same mental space. Therefore the triangle inequality holds, i.e.* $\bar{\delta}(\text{triangle}, \text{rectangle}) > \bar{\delta}(\text{triangle}, \text{square}) + \bar{\delta}(\text{square}, \text{rectangle})$.

Example 7. $r,t \in \vec{V}(\text{donut})$, $v \in \vec{V}(\text{life-ring})$ *and* $d \in \vec{V}(\text{cookie})$. *The concepts* donut *and* life-ring *are very close in* S_2 *since they share the same shape, while* donut *and* cookie *are very close in* S_3 *since they are both pastries. However r is very distant from d indicating that* life-ring *is judged not similar to* cookie.

Figure 1: A partial cross-section of a system of metric spaces in an SPMS. The figure depicts 6 mental spaces and 13 pointers. For graphic reasons metric spaces are represented by circles where larger metric spaces are represented with larger radii. Different mental spaces are connected by hierarchical edges (blue dashed lines) and by structural edges (green dashed arrows). Pointers are identified by black arrows while internal distances in a metric space are depicted by red lines.

3 THEORETICAL ADVANTAGES OF SMPS

Over the years a number of criticisms against traditional models for similarity has been raised, especially against the GMS, and making an exhaustive list of all the objections would be too difficult and is a goal that goes beyond the scope of this work. In this section we briefly analyze some of the main objections raised against the traditional models in a more detailed way and show how the SPMS may naturally overcome known drawbacks of such standard models.

Specifically our analysis mainly focuses on the objections to the three axioms of the geometric model, the reflexive relation, the symmetric relation and the triangular inequality, on the objection relating to the limit to the number of nearest neighbors that can be assigned to a single concept and to the objection related to the lack of specific structure in feature-based models.

Distance Minimality. One of the first objections raised against the geometric model of similarity is that relating to the minimality of distance which imposes that $\delta(\mathrm{A}, \mathrm{A}) = 0$ for any $\mathrm{A} \in \mathcal{E}$. At the basis of this objection, the hypothesis has been advanced that some concepts are, at a perceptual level, more similar to other concepts rather than to themselves.

Just to cite a famous example, Podgorny and Garner hypothesized in their study (Podgorny and Garner, 1979) that the letter “S” is more similar to itself than the letter “W” is to itself or, even more surprising, that the letter “C” is more similar to the letter “O” than “W” is to itself. This hypothesis is made on the basis of an experimental study in which the reaction time has been used as a measure of similarity of two perceptual inputs: in this context, longer reaction times indicate a lower degree of similarity while shorter reaction times indicate a degree of higher similarity.

In the SPMS the higher reaction time in response to a perceptual input can be justified by the fact that detailed and more complex perceptive inputs take longer to be processed by the visual cortex. The letter “W” is undoubtedly more graphically complex than the letter “O” and it is plausible that, given the richness in the details, the decomposition of the starting semantic pointer into its constituent elements is an operation that takes a long time to perform. On the other hand much simpler perceptual inputs can reach a higher level of compression and can therefore be processed faster.

In other words, there is a significant amount of time it takes a semantic pointer to go through a transformation in the brain. Therefore, depending on the size and nature of the input, compressing or decompressing some semantic pointers can take longer than others.[3]

Symmetric Property. The symmetric property in a GMS has also often been criticized. This property implies that, for any $\text{A}, \text{B} \in \mathcal{E}$, the measure of similarity of A towards B should be the same if computed for B towards A. Obviously, in a geometric model this property always holds since the distance from x to y is the same as that from y to x. However some studies highlight that, from a perceptual point of view, this property is not always valid. A famous example in this direction is that presented in (Laakso and Cottrell, 2005), according to which it is assumed that North Korea is perceived much more like China than China can be perceived similar to North Korea. An other example is that presented in (Polk et al., 2002) who found that when the frequency of colors is experimentally manipulated, rare colors are judged to be more similar to common colors than common colors are to rare colors.

Of course, such criticisms are based on the idea that the judgments of similarity are all formulated in the same mental space of comparison. By adopting the SPMS and the idea that the spaces of comparison may change with the context or with the size of the semantic pointers, these problems can be overcome.

[3] This also suggests that the reaction time (Podgorny and Garner, 1979) is not the best measure of similarity for computing the similarity between two concepts, at least not in all cases.

Discussing the same example presented in (Laakso and Cottrell, 2005), it can easily be assumed that China is better known than North Korea. In other words, people generally know much more details about China (size, history, language, currency, culture, etc.) than about North Korea. China is therefore a more relevant concept than North Korea and this implies that the semantic pointer representing China can be much richer than the semantic pointer representing the North Korea.

Referring to Example 5, we can assume that the concept of China has representations starting from the highest levels of the compression networks while the concept of North Korea has only a more abstract representation which resides in lower levels. Since in the SPMS the similarity value depends on which of the two concepts is introduced first, if the more relevant concept is introduced first we start our computation from an higher dimensional mental space and we need to move to the lower levels of the network. Unlike if the less relevant concept is introduced first, our computation may start (and end) on the lower levels.

Triangle Inequality. Regarding the triangle inequality in (Tversky and Gati, 1982) they found some violations when it is combined with an assumption of segmental additivity. Specifically it turns out that in the standard GMS, given three concepts A, B and C, we would aspect that $\delta(\text{A},\text{B}) + \delta(\text{B},\text{C}) \geq \delta(\text{A},\text{C})$ (see Example 6).

However, it is easy to find violations of these assumptions in the common perception of similarity between concepts. Consider for instance the three concepts *life-ring*, *donut* and *cookie*. We can assume that *life-ring* and *donut* are judged similar due to their common shape, while *donut* and *cookie* are judged similar since they are both pastries. However, it is difficult to assume that the distance between *life-ring* and *cookie* is small enough to justify the triangular inequality (see Example 7). Once again the problem from which this objection arises is that it is assumed that the comparison is made within the same geometric space. If, on the other hand, one accepts that different comparisons can take place in different mental spaces, there is an easy justification for this inconsistency.

Number of Closest Neighbors. in (Tversky and Hutchinson, 1986) the authors suggested that a GMS also imposes an upper limit on the number of points that can share the same closest neighbor. A much more restrictive limit is implied in the hypothesis that the data points represent a sample of a continuous distribution in a multidimensional Euclidean space. For instance, under the constraint that there must be a minimum angle of ten degrees between two pointers, a 2D space contains at most 36 pointers while a 3D space contains at most 413 pointers.

By analyzing 100 datasets in (Tversky and Hutchinson, 1986) the authors showed that many conceptual input sets exceed such geometric-statistical limit.

For the sake of completeness we mention that Carol Krumhansl proposed in (Krumhansl, 1978) a solution allowing a variable nummber of closest neighbors to improve the geometric similarity models in order to solve this specific problem. On the basis of this proposal, the dissimilarity between two concepts is modeled in terms of both the distance between elements in a mental space and the spatial density in the proximity of the compared elements. In this context, spatial density is by introducing *spatial density* as the number of elements positioned in proximity to the element. By including spatial density violations of the principles of minimality, symmetry and triangular inequality can also potentially be explained, as well as part of the influence of context on similarity. However, the empirical validity of the spatial density hypothesis has been heavily criticized (Tversky and Gati, 1982; Corter, 1988)

The SPMS model naturally justifies such criticism. If we hypothesize the presence of different representations of the same concept $\mathrm{A} \in \mathcal{E}$, in several mental spaces, the number of neighbors increases proportionally with the cardinality of $\vec{V}(\mathrm{A})$. Furthermore, a pointer x in a mental space of dimension n has, by definition, a greater number of closest neighbors than those of a pointer y lying in a mental space of dimension m, with $m < n$.

Unstructured Representations. One of the main criticisms leveled against the FMS is that it uses unstructured representations of concepts, while it is simply given by a set of unrelated features. To solve this problem, over the years some solutions have been proposed based mainly on two basic ideas, namely that of organizing the characteristics on the basis of a propositional structure and that of organizing them on the basis of a hierarchical structure.

In the propositional approach (Palmer, 1975) the characteristics that are part of a concept are related to each other by statements drawn mainly from the visual domain, such as *above*, *near*, *right*, *inside*, etc. In a very similar way in the hierarchical approach, characteristics represent entities that are incorporated into each other, that is, related to each other by statements such as *part-of* or *a-type-of*.

The SPMS model natively uses a propositional approach. Indeed, it has been shown in various studies how structured representations of concepts, such as those we find formulated in natural language used every day, are essential for the explanation of a wide variety of cognitive behaviors. The possibility of being able to link two different vector representations at the base of the SPMS is of fundamental importance in the definition of our similarity model since, if we are able to link pointers together, then we are able to define a role for each pointer that is a component of a complex structure, tagging pointers of content with pointers having a structural role.

4 EXPERIMENTAL EVALUATION

In this section we present and discuss some preliminary experimental evaluation in order to support the theory underlying the SPMS proposed in this paper.

The SimLex-999 (Hill et al., 2015) and the SimVerb-3500 (Gerz et al., 2016) datasets have been used as benchmarks to verify whether semantic similarities derived by human judgments can be compatible with the SPMS. SimLex-999 provides human ratings for the similarity of 999 words pairs while SimVerb-3500 provides human ratings for the similarity of 3,500 verb pairs. In both datasets the judgments are given in the range form 0 to 10. For our experimental verification we hypothesized the existence of several contexts, identified by the clusters of terms semantically close to each other. These clusters have been computed by selecting all terms in the dataset whose similarity judgment is higher than a bound b. In our experiments this bound has been set to 6. We will refer to such clusters as contexts.

Formally, if T is the set of all terms in the dataset, the model $\mathcal{S}$ consists of all those mental spaces S such that $|\{u \in S : \delta(u,v) \leq b\}| \geq 2$ and $|\{u \in S : \delta(u,v) > b\}| = 0$, for each $v \in S$.

Highly significant in our evaluation is the possibility of considering different meanings of a term as different representations, thus allowing the same term to be included in more than one context. For example in our simulation the term “participate” in the SimLex-999 dataset turns out to appear in two separate contexts, based on its two main meanings, and specifically {“cooperate”, “participate”} and {“add”, “attach”, “join”, “participate”}. Similarly the verb “object” in the SimVerb-3500 turns out to appear in two contexts {“differ”, “object”, “argue”, “disagree”} and {“deny”, “reject”, “object”, “decline”, “refuse”}.

We compute the average and maximum errors (e_{avg} and e_{max}) as the divergence from the average distance between two different mental spaces. Formally, assume E_t is the set of edges connecting two separate mental spaces and let $E_t^{(i,j)} = |E_t \cap (S_i \times S_j)|$ be the set of external edges connecting the spaces S_i and S_j.

Table 1: Experimental results obtained by clustering the terms in the SimLex-999 and SimVerb-3500 datasets to verify whether semantic similarities derived by human judgments can be compatible with the proposed similarity model.

	Simlex-999	SimVerb-3500
Overall terms / relations	1028 / 999	827 / 3500
Selected terms / relations	521 / 324	687 / 1025
Number of contexts	207	344
Minimum/Maximum context size	2 / 9	2 / 32
Terms in one set	516	346
Terms in more than one set	5	339
Average/Maximum error e_{avg}/e_{max}	0.25 / 0.72	0.42 / 0.91

Then we compute the error e_{ij} between S_i and S_j as

$$e_{ij} = \frac{\sum_{(u,v)\in E_t^{(i,j)}} \left(\bar{\delta}(u,v) - \delta_{avg}^{ij}\right)}{|E_t^{(i,j)}|},$$

where

$$\delta_{avg}^{ij} = \frac{\sum_{(u,v)\in E_t^{(i,j)}} \bar{\delta}(u,v)}{|E_t^{(i,j)}|}$$

Therefore the errors e_{avg} and e_{max} are computed as

$$e_{avg} = \frac{\sum_{S_i,S_j\in\mathcal{S}} e_{ij}}{|\{e_{ij}\ :\ e_{ij} > 0\}|},$$
$$e_{max} = \max\left\{e_{ij}\ :\ S_i,S_j \in \mathcal{S}\right\}$$

Table 1, presenting the results obtained in our simulation, shows how the error obtained by adapting the similarity judgments in the datasets to our model is quite negligible, and supports the idea that the SPMS can justify many of the similarity judgements obtained empirically.

5 CONCLUSIONS

We introduced the Semantic Pointers Model of Similarity (SPMS), a unifying solution in one of the most relevant challenges of concept research. Our model is based on the recent Eliasmith's theory of biological cognition, where concepts are represented as semantic pointers, and assumes the coexistence of several contextual mental spaces of different size, populated by a set of allowable concept representations. We proposed a mathematical formulation of the model and of the way the similarity distance is computed, highlighting the features that the new model has in common with traditional models and how many of the criticisms raised over the years towards the latter find a natural settlement. Our preliminary experimental investigations show how the SPMS is able to adequately model the human judgments of similarity present in two of the most recent datasets available in the literature. In our future work we intend to create a more sophisticated computational model based on the SPMS.

REFERENCES

Carroll, J. D. and Wish, M. (1974). Models and methods for three-way multidimensional scaling. *Contemporary developments in mathematical psychology*, 2:57–105.

Corter, J. E. (1988). Testing the density hypothesis: Reply to krumhansl. *Journal of Experimental Psychology: General*, 117:105–106.

Eliasmith, C. (2004). Learning context sensitive logical inference in a neurobiolobical simulation. *AAAI Fall Symposium - Technical Report*.

Eliasmith, C. (2013). *How to Build a Brain: A Neural Architecture for Biological Cognition*. Oxford University Press.

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41(1):1–63.

Gentner, D. and Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52:45–56.

Gerz, D., Vulić, I., Hill, F., Reichart, R., and Korhonen, A. (2016). SimVerb-3500: A large-scale evaluation set of verb similarity. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2173–2182, Austin, Texas. Association for Computational Linguistics.

Hahn, U. (2003). *Similarity*. London: Macmillan.

Hill, F., Reichart, R., and Korhonen, A. (2015). SimLex-999: Evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics*, 41(4):665–695.

Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507.

Holyoak, K. J., Morrison, R. G., Goldstone, R. L., and Son, J. Y. (2012). Similarity.

Krumhansl, C. (1978). Concerning the applicability of geometric models to similarity data : The interrelationship between similarity and spatial density.

Laakso, A. and Cottrell, G. (2005). *Churchland on connectionism*. Cambridge University Press.

Markman, A. and Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language*, 32(4):517–535.

Markman, A. and Gentner, D. (1993b). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25(4):431–467.

Markman, A. and Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & cognition*, 24(2):235—249.

Mitchell, M. (1993). *Analogy-Making as Perception: A Computer Model*. MIT Press, Cambridge, MA, USA.

Palmer, S. (1975). *Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction*, pages 279–307.

Podgorny, P. and Garner, W. R. (1979). Reaction time as a measure of interintraobject visual similarity: Letters of the alphabet. *Perception & Psychophysics*, 26:37–52.

Polk, T. A., Behensky, C., Gonzalez, R., and Smith, E. E. (2002). Rating the similarity of simple perceptual stimuli: Asymmetries induced by manipulating exposure frequency. *Cognition*, 82:B75–B88.

Torgerson, W. (1958). Theory and methods of scaling.

Torgerson, W. S. (1965). Multidimensionsal scaling of similarity. *Psychometrika*, 30:379–393.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84:327–352.

Tversky, A. and Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological review*, 89:123–54.

Tversky, A. and Hutchinson, J. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review*, 93:3–22.

Wiener-Ehrlich, W. K., Bart, W. M., and Millward, R. (1980). An analysis of generative representation systems. *Journal of Mathematical Psychology*, 21(3):219–246.