Semantic Spaces: Measuring the Distance between Different Subspaces

G. Zuccon[1], L. A. Azzopardi[1], and C. J. van Rijsbergen[1]

{guido, leif, keith}@dcs.gla.ac.uk
Dept. of Computing Science, University of Glasgow,
Glasgow, United Kingdom

Abstract. Semantic Space models, which provide a numerical representation of words' meaning extracted from corpus of documents, have been formalized in terms of Hermitian operators over real valued Hilbert spaces by Bruza et al. [1]. The collapse of a word into a particular meaning has been investigated applying the notion of quantum collapse of superpositional states [2]. While the semantic association between words in a Semantic Space can be computed by means of the Minkowski distance [3] or the cosine of the angle between the vector representation of each pair of words, a new procedure is needed in order to establish relations between two or more Semantic Spaces. We address the question: how can the distance between different[1] Semantic Spaces be computed? By representing each Semantic Space as a subspace of a more general Hilbert space, the relationship between Semantic Spaces can be computed by means of the subspace distance. Such distance needs to take into account the difference in the dimensions between subspaces. The availability of a distance for comparing different Semantic Subspaces would enable to achieve a deeper understanding about the geometry of Semantic Spaces which would possibly translate into better effectiveness in Information Retrieval tasks.

1 Introduction

Semantic Space techniques map words in a high dimensional vector space [4]. The map is usually built by computing lexical co–occurrences between words appearing in the same context where each vector is assigned to a word and represents the co–occurrences between the word and others. In this work, we consider a particular instance of a Semantic Space, the Hyperspace Analogue to Language (HAL). The HAL space is created through the co–occurrence statistics within a corpus of documents. This space has been used as a representation model of semantic memory [5] and has been shown to be compatible with human reasoning in cognitive science [6]. Within the area of Information Retrieval, HAL has been used to perform information inference for query expansion [7].

[1] We refer to different Spaces, not different instances of the same space, i. e. the same space rescaled.

In Semantic Spaces (like HAL) words (or concepts [8]) are represented by points in a high dimensional vector space: their position in the space is related to their meaning and inter–relationships. The former can be inferred by examining the components of the high dimensional vector associated with a word, while the latter can be exploited by a similarity measurement between word vectors. For example, in [3] the authors propose adopting the **Minkowski distance**, defined with respect to two vectors $\mathbf{u_i}$ and $\mathbf{v_i}$ of the Semantic Space as:

$$d_M = \sqrt[r]{\sum(|\mathbf{u_i} - \mathbf{v_i}|)^r} \quad (1)$$

Comparing the word vectors is one way to derive meaning from the Semantic Space. An alternative is to compare subspaces of documents, or sets of documents. While Semantic Spaces provide a representation of knowledge generated from a sample of text, a problem arises when we consider two or more Semantic Spaces that have been generated from independent samples of text. Specifically, how do we compare one Semantic Space with another? Once again, the simplest solution is to consider the distance between the representation of the same word vector in the two Semantic Spaces, using for example the Minkowski distance. However, this naïve treatment may be inappropriate, because different words used in the same sense will not be compared. For example, *cat* and *kitten* are semantically related in the context of the concept feline, and thus we would expect them to share the same vector representation. However, when computing the distance in a naïve way we do not take into account such relationships. Then, if in document d_1 we refer to the concept of feline with the common word *cat*, while in d_2 we refer to the close concept but using the term *kitten*, we might not capture the semantic relationship between the two documents. To avoid such problem, we propose to compute the distance between Semantic Spaces not relying on word–representation similarity, but on the more general subspace distance. The subspace correspondent to a document or to a set of documents conveys the meaning expressed by the text traces; comparing subspaces then would provide a distance based on the meaning/topic area associated *to the set* as opposed to the word level.

The paper continues as follow. In Section 2 we illustrate a formalization of Semantic Spaces in terms of Quantum Theory (QT) as it has been introduced in [2]. Moreover, we briefly present how to derive a numeric representation of a Semantic Space from a corpus of documents. In Section 3, several measures to compute the distance between subspaces are illustrated, guiding the reader to the definition of a metric which allows comparisons between Semantic Subspaces. Section 4 illustrates and discusses the preliminary experiments using subspace distance. The paper concludes providing a discussion of the distance between Semantic Subspaces, stating the objects of future investigations (Section 5).

2 Semantic Spaces: a Hilbert space representation

In the following, the formalization of Semantic Space in terms of Hilbert spaces [2] is presented. Consider a n–dimensional (real valued) Hilbert space H, in which the inner product is represented by the Euclidean scalar product. In the following we limit our focus at real valued Hilbert spaces, discarding the analysis of complex valued spaces. Such a limitation is driven by the fact that the spaces are built from statistical data from texts, which uses only real values. Nevertheless, it is clear that complex numbers plays an important role in the description of states of a QT systems [9]. Each dimension of the Hilbert space H corresponds to a word in the vocabulary of a corpus of documents. The global Semantic Space, i.e. the Semantic Space derived considering the whole corpus, is denoted by $\hat{\mathcal{S}}$. The Semantic Space derived from document d of the considered corpus is represented by S_d. Similarly the Semantic Space associated to a word w belonging to the vocabulary V of the corpus is denoted by S_w. It is clear that $\hat{\mathcal{S}}$ is a subspace of the Hilbert space H since its vectors are instances of the vectors in H; in particular, $\hat{\mathcal{S}}$ is a n—dimensional subspace. Similarly S_d is a m–dimensional subspace of H. Note that the subspace relationship $S_d \subseteq \hat{\mathcal{S}}$ always holds.

We briefly illustrate the procedure to form the high dimensional matrix which corresponds to the HAL representation of the corpus of documents[2]. A window of text is passed over each document in the collection in order to capture co–occurrences of words. The length of the window is set to l: a typical value of l is 10; different values capture different levels of relationship between words. Words that co–occur into a window do so with a strength inversely proportional to the distance between the two co–occurring words. A thorough study which investigates the most effective function for encoding the inverse proportional weighting can be found in [10, Chapter 8.5]. By sliding the window over the whole collection and recording the co–occurrence values, a co–occurrence matrix A can be created. Since in our approach, as well as in [1, 2, 7], we are not interested in the order of the co–occurrences, in contrast with the work of Gärdenfors [8], therefore we can compute a symmetric matrix by means of $\hat{S} = A + A^T$, and then normalise the columns.

A symmetric matrix obtained by the illustrated procedure is associated to each subspace and is denoted with the same symbol assigned to the subspace: it is clear from the use if it refers to the subspace itself or to its symmetric HAL matrix. Note that subspace S_d can be defined as the range, or the complement of the range, of matrix S_d. The symmetric matrices $\hat{\mathcal{S}}$ and each S_d, S_w are Hermitian linear operators. The following relations between the previous linear operators hold:

$$\hat{\mathcal{S}} = \sum_{d \in C} S_d, \tag{2}$$

$$\hat{\mathcal{S}} = \sum_{w \in V} S_w \tag{3}$$

where C is a corpus of documents. In the rest of this paper the focus will be on subspaces referring to document or set of documents.

[2] The interested reader should refer to [3] for a complete investigation of the procedure.

3 A distance measure between (HAL) spaces

We aim to define a distance measure between Semantic Spaces, in order to be able to geometrically compare Semantic Spaces generated by different sources of evidence, i.e. compare subspaces formed with different subsets of documents.

Consider the general case of comparing the subspaces S_a and S_b derived by different sets of documents (a more particular case is when the set D associated to S_d contains only one document). We can associate to each subspace a $n \times n$ projector operator P. Then the **inner product** between two subspaces of H is the trace inner product for projection matrices:

$$\langle S_a, S_b \rangle = tr(P_a^* P_b) = tr(P_a P_b) \tag{4}$$

The appropriate candidate as distance between Semantic Subspaces has to satisfy several characteristics. Firstly, it would be desirable that the measure turns to be a metric. The inner product between two subspaces is not a metric: the inner product of P_a with itself is maximal rather then minimal. Nonetheless, it represents a measure of the *similarity* between the two subspaces: it is matter of fact that the measure proposed at the end of this Section employs the inner product between projectors of subspaces. An additional constraint to the measure has to be added. When comparing Semantic Subspaces, obtained for example from two documents, it is not guaranteed that they have the same number of dimensions, on the contrary it is frequently the case that the basis for such subspaces differ remarkably. Thus, a right candidate to measure the distance between two Semantic Subspaces should be able to capture differences in the dimensions of the basis of the Semantic Subspaces. The angle between the vectors of the subspaces is a key factor not only for the inner product between projectors, but for a whole family of measures based on the **principal (or minimal) angles**.

Definition 1. *For nonzero subspaces S_a and $S_b \subseteq \mathcal{S}$, the principal angle between S_a and S_b is defined as the number $0 \leq \theta \leq \frac{\pi}{2}$ that satisfies*

$$\cos\theta = \max_{\mathbf{a}\in S_a, \mathbf{b}\in S_b, \|\mathbf{a}\|=\|\mathbf{b}\|=1} \mathbf{a}^T\mathbf{b} \tag{5}$$

The principal angle θ is 0 if and only if $S_a \cap S_b \neq \mathbf{0}$, while $\theta = \frac{\pi}{2}$ if and only if $S_a \perp S_b$. It is worthwhile to reformulate definition 1 in terms of projectors; this leads to the following theorem (where the proof is shown in [11])

Theorem 1. *If P_a and P_b are the orthogonal projectors onto S_a and S_b respectively, then*

$$\cos\theta = \|P_a P_b\| = \|P_b P_a\| \tag{6}$$

These principle angles are related to the eigenvalues of $P_a P_b$: in fact, the first m (where m is the minimum between the subspace dimensions of S_a and S_b) eigenvalues of $P_a P_b$ are $\cos^2\theta_1, \ldots, \cos^2\theta_m$. We are however interested in comparing subspaces which have different dimensions, i.e. they do not have the same basis dimension. Unfortunately, the behaviour of a measure based on the principle angles is quite controversial if the subspaces have a different dimension. In fact, the principal angles are defined just for the minimum between the subspace dimensions: thus the measure does not take into consideration all the dimensions of both subspaces. For example, consider two subspaces: S_a of dimension p, S_b of dimension r such that $r \geq p$. Subspace S_b is built such that its first p basis vectors are the same of S_a, while the other $r - p$ basis vectors are arbitrarily constructed. Consider the **geodesic distance** [12] as measure based on principal angles; the measure is defined by:

Definition 2. *Let S_a and S_b be two subspaces and $\theta_1, \ldots, \theta_m$ be the m principal angles between S_a and S_b (where m is the dimension of the smallest subspace). The geodesic distance between S_a and S_b is*

$$d_g(S_a, S_b) = \sqrt{\theta_1^2 + \ldots + \theta_m^2} \tag{7}$$

If S_a and S_b are constructed as illustrated before, then the geodesic distance between S_a and itself will be 0 since each θ_i is zero implying that $S_a \equiv S_a$. However, when measuring the distance between S_a and S_b based on principal angles, we find that all the p angles that are computed are equal to the null angle 0, since S_b shares p basis vectors with S_a. Thus, the measure does not take into account the $r - p$ basis vectors of S_b that are not shared with S_a.

A distance measure based on the principal angles between subspaces, such as the geodesic distance, is then significant if and only if the subspaces have the same dimensions: this is unlikely to be the case when comparing different Semantic Spaces. We refer to such problem as the Zero Distance Problem: the bigger the difference in the number of dimensions of the two subspaces, the greater the extent of the problem since the number of discarded dimensions in the computation of the distance grows. Measures based on the principal angles, such as the geodesic distance, are generally affected by the Zero Distance Problem: the solution to the problem passes through the **chordal distance** [13], a monotonic function of the inner product.

Definition 3. *The chordal Grassmannian distance between two subspaces S_a and S_b is given by means of the associated projectors P_a and P_b by*

$$d_c(P_a, P_b) = \sqrt{m - tr(P_a P_b)} \tag{8}$$

As for the previous measures, also in the case of the chordal distance a difference in the dimensionality of the subspaces S_a and S_b is only partially taken into account: in fact the product $P_a P_b$ depends on the degree of association (or similarity) between the two subspaces, comparing each dimension, but the number m in equation 8 refers to the dimension of the subspaces S_a and S_b, thus needing to be of the same dimension. Anyway, the chordal distance opens the path to the definition of a distance that is not restricted by the dimensionality of the subspaces to be compared. The first step is to introduce the Hausdorff distance which measures the distance between two compact subsets of the space. In our case, we consider the L_2–Hausdorff distance between a vector u_i and a subspace V which is expressed by $d_H(u_i, V) = min||u_i - v||$, where $v \in V$ and $||.||$ is the Euclidean norm. We now have to consider the subspace distance between subspaces of the same dimension proposed in [14]:

Definition 4. *The subspace distance $d_s(S_a, S_b)$ for two p–dimensional subspaces S_a and S_b is defined as*

$$d_s(S_a, S_b) = \sqrt{\sum_{i=1}^{p} d_H^2(\mathbf{u_i}, S_b)} \tag{9}$$

where

1. $\mathbf{u_1}, \ldots, \mathbf{u_i}, \ldots, \mathbf{u_p}$ *is an orthonormal basis for S_a, and*
2. $d_H(\mathbf{u_i}, S_b)$ *is the Hausdorff distance from the end point of the basis vector $\mathbf{u_i}$ to subspace S_b.*

Such distance has been extended to the case where the subspaces have different dimensions. Let $\mathbf{v_1}, \ldots, \mathbf{v_i}, \ldots, \mathbf{v_r}$ be an orthonormal basis for S_b, then

Definition 5. *The subspace distance $d_s(S_a, S_b)$ between the p–dimensional subspaces S_a and the r–dimensional subspace S_b is defined as*

$$d_s(S_a, S_b) = \sqrt{\max(p, r) - \sum_{i=1}^{p}\sum_{j=1}^{r} (\mathbf{u_i}^T \mathbf{v_j})^2} \tag{10}$$

The introduced distance has several properties. In primis, it is invariant to the choice of the orthonormal basis for the subspaces S_a and S_b. Furthermore, it is symmetric and not negative, in particular $d_s(S_a, S_b) = 0$ if and only if $S_a \equiv S_b$. The upper bound for the subspace distance is given by $d_s(S_a, S_b) \leq \sqrt{\max(p, r)}$ and corresponds to the orthogonality condition $S_a \perp S_b$. Finally, as proved in [15], the subspace distance satisfies the triangle inequality, and thus it is a proper metric defined on subspaces.

The next step is to express the subspace distance in terms of projector operators, and thus finding a relationship with the chordal distance. As demonstrated in [16], equation 10 can be re-written as:

$$d_s(S_a, S_b) = \sqrt{\frac{1}{2} tr \left[(\Lambda_p - \Lambda_r)^2 + (S_a S_a^T - S_b S_b^T)^2 \right]} \tag{11}$$

where $\Lambda_i = diag(1, \ldots, 1, 0, \ldots, 0)$ is a diagonal matrix with i 1's and $n - i$ 0's elements and S_a, S_b are the symmetric HAL matrices associated to the corresponding subspaces. With some algebraic calculation and since the matrix products $S_a S_a^T$ and $S_b S_b^T$ are the projectors P_a and P_b respectively, the subspace distance can be stated as:

$$d_s(S_a, S_b) = \sqrt{\max(p, r) - tr(P_a P_b)} \tag{12}$$

The proposed subspace distance might be employed to compute the distance between Semantic Subspaces, aiming to obtain a more precise measurement of separation than using a naïve distance based on comparison between single word representations, e.g. the Minkowski distance.

Comparing equation 12 and 8, both formulating the chordal distance between two subspaces, it appears clear the strong relationship between the two distances, differing in taking into account the maximum of the subspace dimensions.

Each rank d projector represent a basis of a Hilbert subspace and can be regarded as a d–(hyper)plane: this provides an embedding of the Grassmannian of d–plane into a flat vector space. Thus, the rank d projector will sit on a sphere in this flat space, more precisely it will be point on the surface of a sphere, and its Euclidian distance provides us with a chordal distance between projectors. The chordal distance has been successfully used to study the packing problems for n–planes, where the aim is to find a set of hyper-planes such that the minimum distance between each pair of planes in the set is as large as possible [17]. Since the chordal distance provides a natural measure of the distance between bases of the same rank in a Hilbert spaces, it has been used to detect Mutually Unbiased Bases (MUB) [18], i.e. bases which spans planes totally orthogonal between them. This condition is reached when the chordal distance between the two bases is maximum.

Previous research in QT focused on the derivation of a suitable measure to judge the distance between quantum states of different preparations. Such a measure can be used to characterize the degree of distinguishability between states (and related preparations). In fact, because of the statistical error introduced when measuring frequencies of possible outcomes for a finite ensemble of identically prepared systems, it is generally difficult to distinguish between preparations that slightly differ [19]. The measure thus is used to judge the degree of separation between states. This is the underlying idea of the *statistical distance* between quantum preparations presented in [19], and is determined entirely by statistical fluctuations. However, it turns out that the statistical distance provides an identical result to the measure of the angle between rays in a Hilbert space associated with the pure quantum states of the preparations. Computing the distance between Semantic Space representations of a word (in terms of HAL subspaces) is similar to measuring the angle between the representative rays spanned by the word in its Hilbert space representations. Another distance that is related to the evaluation of the distinguishability of two quantum states is the so called Bures distance [20]. It measures the distance of the associated density operators ρ_1 and ρ_2 by the formula $d_B(\rho_1, \rho_2) = \sqrt{2}[1 - tr((\rho_1^{1/2}\rho_2\rho_1^{1/2})^{1/2})]^{1/2}$. The Bures distance has been interpreted as a generalization of transition probabilities to mixed states [21].

4 Pilot experiment

We conducted a pilot investigation in order to examine how well subspace distance performs. In particular, we experimentally demonstrate that related documents are at a closer subspace distance between each other than not related ones. As baseline for the comparison we employed the Minkowski distance with $r = 2$ (Euclidian distance) between Semantic Space representations of words. In the following we describe the details of the experiment, discussing how Semantic Subspaces have been generated and in which terms we compare the subspace distance against the baseline.

We employ a standard IR collection, namely WSJ 87–92, as source of documents used to generate the Semantic Subspaces. This collection has more than 170 thousand newspaper articles, containing over 226 thousand unique terms. For the purpose of our pilot study we consider two subsets of document, R and N. Set R contains only documents that have been judged relevant (by human assessors) to a query q, while set N containing those documents judged as irrelevant to the same query, each query belonging to one of the TREC 1 topics. All the documents have been processed through stop-word removal and stemming.

Two methods for generating the Semantic Spaces have been employed, both inspired by the HAL paradigm. In both methods a window of text is passed over the text. The window size is 11, 5 words on the left and 5 on the right of the target word. We adopted an inverse proportional function to score the strength of co-occurrences with the target word, i.e. closer the co-occurent term is to the target word and higher is the score attributed to the pair. The only difference between the two methods is represented by the text over which the window is passed. The first method, which sticks to the definition of the generating procedure for HAL, passes the window over all the text contained in a document. On the contrary, the second method, which is partially inspired by [2] and [22], passes the window over traces of text extracted from the document. Such traces are extracted considering windows centered on target words. For each TREC 1 topic, target words are extracted from the description of the topic itself. As well as the documents, also the target words are pre-processed by matching against a stop-word list and by stemming. From now on, we refer to this Semantic Subspace generation method as *HAL traces*.

In tables 1 and 2, we report the preliminary results obtained by our study using topic 51 of the WSJ 87–92 TREC collection. The values presented in table 1 contains the average distance values obtained employing *HAL traces*, while table 2 refers to the average distances calculated using the classic HAL representation. In both tables, the Euclidean distance has been calculated as the square root of the squared difference between selected word representations associated with two HAL spaces. The values obtained where then averaged among the documents contained in the set and reported into the tables. Instead, for what concern the subspace distance, the reported values refer to the average over the correspondent set of documents of the following formula:

$$sim_s(S_a, S_b) = 1 - \frac{\sqrt{\max(p, r) - \sum_{i=1}^{p} \sum_{j=1}^{r} (\mathbf{u_i}^T \mathbf{v_j})^2}}{\sqrt{\max(p, r)}} \quad (13)$$

which expresses the similarity (driven by the subspace distance) between subspaces S_a and S_b. A value of this similarity close to 0 means that the two subspaces are almost orthogonal, with $sim_s(S_a, S_b) = 0$ representing the case $S_a \perp S_b$, while a value close to 1 represents high degree of similarity. Thus, the two distances are not directly comparable. However, it is possible to understand the behavior of the two measure in discriminating between relevant and not relevant documents.

Table 1. Average distance between sets of relevant documents (R) and not relevant documents (N) obtained by the subspace distance (*Subspace* in the table) and the Euclidean distance (*Euclidean*) computed over subspaces generated by the *HAL traces* paradigm.

	R		N	
	Subspace	Euclidean	Subspace	Euclidean
R	0.0376 ± 0.0116	9.3910 ± 4.6994	0.0182 ± 0.0072	6.7059 ± 5.5936
N	0.0182 ± 0.0072	6.7059 ± 5.5936	0.0386 ± 0.0093	3.3816 ± 2.0667

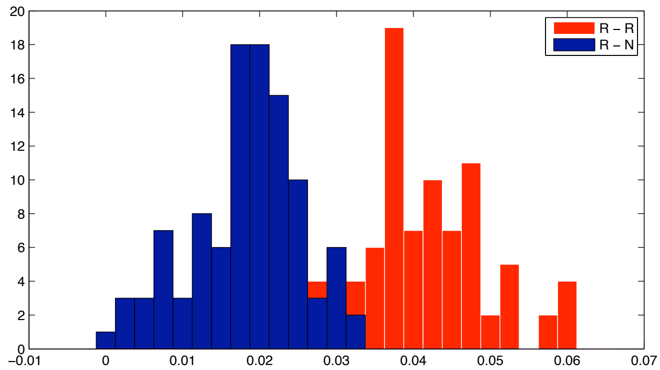
Table 2. Average distance between sets of relevant documents (R) and not relevant documents (N) obtained by the subspace distance (*Subspace* in the table) and the Euclidean distance (*Euclidean*) computed over subspaces generated by the traditional *HAL* paradigm.

	R		N	
	Subspace	Euclidean	Subspace	Euclidean
R	0.1504 ± 0.0142	19.8121 ± 4.9445	0.0124 ± 0.0068	5.7710 ± 5.1289
N	0.0124 ± 0.0068	5.7710 ± 5.1289	0.1181 ± 0.0173	3.5376 ± 2.4407

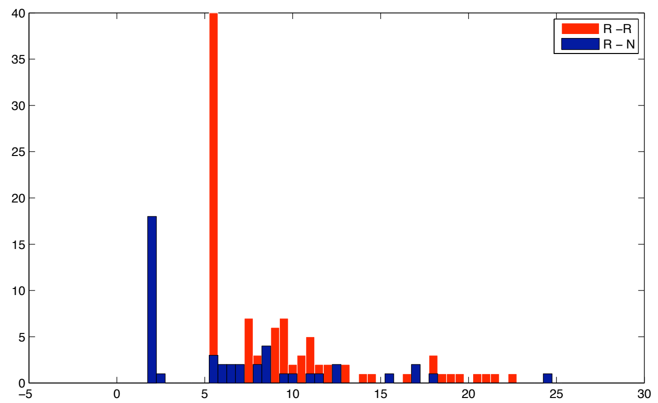
4.1 Discussion of the preliminary results

The results of the preliminary experiments reported in this paper refer to topic 51 of TREC 1. The results show that the subspace distance is able to discriminate between subspaces associated with relevant documents and the ones generated from non relevant documents. In fact, in accordance with the values reported for the subspace distance, the degree of Semantic similarity between the non relevant set of documents (labeled N) and the relevant set (labeled R) is lower (0.0182 for *HAL traces*, 0.0124 for *classic HAL*) than the similarity among occurrences of relevant documents (*HAL traces*: 0.0376, *classic HAL*: 0.1504) or not relevant documents (*HAL traces*: 0.0386, *classic HAL*: 0.1181). The same result is not achieved by the Euclidean distance. For Semantic Subspaces generated by *HAL traces* and by the traditional approach, the Euclidian distance between subspaces belonging to R is higher than the accumulated average distance between R subspaces and N ones.

Fig. 1. Frequencies distribution of pairwise subspace distances (a) and Euclidean distances (b) between subspaces belonging to the set of relevant documents (R) and the non relevant (N) for topic 51. The subspace generation paradigm adopted is *HAL traces*.



(a) Distribution of frequencies for the subspace distance.



(b) Distribution of frequencies for the Euclidean distance.

From the tables is possible to evince that the subspace distance tends to flatten the distance among subspaces to the range $[0.9, 1.0]$, while the Euclidean distance is able to provide a greater range of values, making easy to detect significant differences between subspaces.

Fig. 1 illustrates the frequencies distribution of pairwise distance values (obtained by the subspace distance (a) the Euclidean distance (b)) between Semantic Subspaces generated using the paradigm *HAL traces*, although rather similar figures are obtained when considering subspaces generated by the standard HAL derivation. The figures can be interpreted as follow. Subspaces associated to relevant documents (R) are on average at a closer subspace distance to each other than to non relevant documents (N) (see Fig. 1 (a)). using the Euclidean distance the separation between R and N is not as distinct. This suggests that the subspace distance will be more effective in discriminating relevant documents from non relevant.

5 Conclusion and Future Work

In this work a distance based on the chordal distance has been introduced in order to compare Semantic Subspaces constructed from subsets of a document corpus. Our approach allows to compare directly two sets of documents though their subspace distance, whereas [3] only deals with comparing a word and its meaning. Geometrically, this corresponds in considering the projection of a subspace into another, rather than the intersection between two subspaces.

Future work will be directed towards applying the proposed measure in a number of retrieval applications in order to determine its effectiveness.

6 Acknowledgement

This work was supported in part by the EPSRC Renaissance project (grant number EP/F014384/1) and by the EPSRC grants EP/E002145/1 and EP/F014708/1. The authors are thankful to prof. D. Song for the useful discussions and the guidance provided to make the first steps into the investigation of Semantic Subspaces, and to Alvaro Huertas-Rosero for the valuable suggestions. Moreover, the authors are grateful to the IRF[3] (Information Retrieval Facility) for the resources and support provided for carrying out the preliminary experiments.

References

1. Bruza, P.D., Cole, R.J.: Quantum Logic of Semantic Space: An Exploratory Investigation of Context Effects in Practical Reasoning. In: We Will Show Them: Essay in Honour of Dov Gabbay. Volume 1. College Pubblications (2005) 339–361

[3] http://www.ir-facility.org/

2. Bruza, P.D., Woods, J.: Quantum Collapse in Semantic Space: Interpreting Natural Language Argumentation. In: Proceedings of the 2nd QI Symposium. (2008) 141–147
3. Lund, K., Burgess, C.: Producing High-dimensional Semantic Spaces from Lexical Co-occurrence. Behavior Research Methods **28**(2) (1996) 203–208
4. Osgood, C., Suci, G., Tannenbaum, P., Date, P.: The Measurement of Meaning. University of Illinois Press (US) (1957)
5. Burgess, C., Livesay, K., Lund, K.: Explorations in Context Space: Words, Sentences, Discourse. Discourse Processes **25**(2 & 3) (1998) 211–257
6. Landauer, T.K., Foltz, P.W., Laham, D.: An Introduction to Latent Semantic Analysis. Discourse Processes **25**(2 & 3) (1998) 259–284
7. Song, D., Bruza, P.D.: Discovering Information Flow Using High Dimensional Conceptual Space. In: Proceedings of the 24th ACM SIGIR. (2001) 327–333
8. Gärdenfors, P.: Conceptual Spaces: The Geometry of Thought. MIT Press, (US) (2000)
9. Bowman, G.E.: Essential Quantum Mechanics. Oxford University Press (UK) (2008)
10. Sahlgren, M.: The Word-Space Model. PhD thesis, Stockholm University (2006)
11. Ipsen, I.C.F., Meyer, C.D.: The Angle Between Complementary Subspaces. American Mathematical Monthly **102**(10) (1995) 904–914
12. Wong, Y.C.: Differential Geometry of Grassmann Manifolds. In: Proceedings of the National Academy of Science. Volume 57. (1967) 589–594
13. Bengtsson, I., Bruzda, W., Ericsson, A., Larsson, J.A., Tadej, W., Zyczkowski, K.: Mubs and Hadamards of Order Six. ArXiv Quantum Physics e-prints (2006)
14. Wang, L., Wang, X., Feng, J.: Subspace Distance Analysis with Application to Adaptive Bayesian Algorithm for Face Recognition. Pat. Rec. **39**(3) (2006) 456–464
15. Sun, X., Wang, L., Feng, J.: Further Results on the Subspace Distance. Pat. Rec. **40**(1) (2007) 328–329
16. Sun, X., Cheng, Q.: On Subspace Distance. LNCS **4142** (2006) 81–89
17. Conway, J.H., Hardin, R.H., Sloane, N.J.A.: Packing Lines, Planes, etc.: Packings in Grassmannian Spaces. Experimental Mathematics **5**(2) (1996) 139–159
18. Bengtsson, I., Bruzda, W., Ericsson, A., Larsson, J.A., Tadej, W., Życzkowski, K.: Mutually Unbiased Bases and Hadamard Matrices of Order Six. Journal of Mathematical Physics **48**(5) (2007)
19. Wootters, W.K.: Statistical Distance and Hilbert Space. Phys. Rev. D **23**(2) (1981) 357–362
20. Braunstein, S.L., Caves, C.M.: Statistical Distance and the Geometry of Quantum States. Phys. Rev. Lett. **72**(22) (1994) 3439–3443
21. Uhlmann, A.: The "Transition Probability" in the State Space of a *-algebra. Reports on Mathematical Physics **9** (1976) 273–279
22. Huertas-Rosero, A.F., Azzopardi, L.A., van Rijsbergen, C.J.: Characterising through Erasing: A theoretical framework for representing documents inspired by quantum theory. In: Proceedings of the 2nd QI Symposium. (2008) 160–163