

## Signatures of Semantic Structures

D. Gueorguiev 2/6/2022

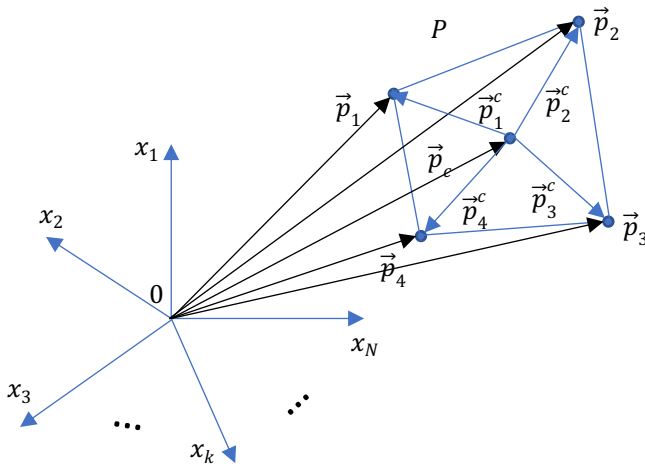
The position, the shape of a Semantic Structure and the volume it takes up in Semantic Space is ever evolving. The Signature of a Semantic Structure represents static view of the position in Semantic Space of that Semantic Structure at specific moment of time.

### The Signature of Semantic Properties

Each property can be represented as a collection of  $N = N(P)$  points ( $N \geq 1$ ) in Semantic Space plus the position of its centroid. Thus, each property can be represented by the following  $(N + 1) \times L$  matrix:

$$\mathbf{P} = [\mathbf{p}_c \ \mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_N]$$

Here the column vector which corresponds to the centroid  $\vec{p}_c$  will be denoted with subscript  $c$  :  $\mathbf{p}_c$ . Each column vector  $\mathbf{p}_i$ ,  $i = 1, N$  represents a point in the  $L$  dimensional Semantic Space relative to its centroid  $\mathbf{p}_c$ . The property  $P$  can be visualized as an  $L$ -dimensional  $K$ -polytope. With  $\vec{p}_i$ ,  $i = 1..4$  are denoted the vertices of the polytope. With  $\vec{p}_c$  we denote the center of mass (centroid) of the polytope which we will discuss in detail later.



We have stated that each vertex in the property polytope represents a semantic aspect of the property. The metric  $|\vec{p}_i^c| = |\vec{p}_i - \vec{p}_c|$  represents the type (or kind) of the semantic aspect  $\vec{p}_i$ . The value of the semantic aspect is represented by the position of  $\vec{p}_i$  with respect to the semantic axes  $x_1, \dots, x_N$ . In other words the  $K - 1$ -tuple  $(\theta_1, \theta_2, \dots, \theta_{K-1})$  where each  $\theta_j$ ,  $j = 1..K - 1$  denotes the angle between  $\vec{p}_i$  and the semantic axis  $x_j$  uniquely identifies a specific value for the semantic aspect  $\vec{p}_i$ . For details refer to [Modeling Attractive and Repulsive Forces in Semantic Properties](#).

## The Energy Signature of a Semantic Property

With every property there is an associated energy vector which defines the current energy signature of the property  $P$ :

$$\vec{E}_P = [E_1, E_2, \dots, E_{N(P)}]$$

The energy signature of the property describes how the energy is distributed among the property aspects.

## The Signature of Property Association Link

For the structure shown below

$$\begin{array}{c} \mathbf{w}_{1,2} \\ P_1 - A_{12} - P_2 \end{array}$$

Let the dimension of the signature matrix for property  $P_1$  is  $K_1 \times N$  and the dimension of the signature matrix for property  $P_2$  is  $K_2 \times N$ . With  $\mathbf{w}$  it is denoted the semantic significance vector pertaining to the association but it is not relevant for the current discussion it will be omitted in the further results.

The signature of the property association  $A_{12}$  is

$$\mathbf{A}_{12} = [\mathbf{p}_1^1 \mathbf{p}_2^1 \dots \mathbf{p}_k^1 | \mathbf{p}_1^2 \mathbf{p}_2^2 \dots \mathbf{p}_l^2]$$

where  $k < K_1$  and  $l < K_2$ . Here the points in semantic space  $\mathbf{p}_1^1, \mathbf{p}_2^1, \dots, \mathbf{p}_k^1$  are a subset of the semantic aspects of property  $P_1$  and the points  $\mathbf{p}_1^2, \mathbf{p}_2^2, \dots, \mathbf{p}_l^2$  are a subset of the semantic aspects of property  $P_2$ . In a sense the property association acts as a filter which exposes only certain aspects of each property to generate binding force. In the expression for  $\mathbf{A}_{12}$  above the symbol  $|$  denotes a column separator separating the points pertaining to the property  $P_1$  and those pertaining to the property  $P_2$ . The last expression can be written concisely in block matrix notation as

$$\mathbf{A}_{12} = [\mathbf{P}^1 | \mathbf{P}^2]$$

## The Signature of Primitive Semantic Particles

Below are shown two examples of primitive semantic particles with their property trees:



## The Signature of Compound Semantic Particles

Let us have the compound particle  $V_{comp}$  represented by its elementary particle sequence and semantic tree  $stree(V_{comp})$ :

$$stree(V_{comp}) =$$

```

      /---V1---\
     /   |   \
    V2  V5  V6-
   / \   / | \
  V3 V4 V7 V8 V9

```

The property tree for each  $V$ -particle  $V_k, k = 1..9$  are given with the algebraic notation discussed in [Semantic Tree Operations](#).

$ptree(V_k) = \sum_{\mathbf{k} \in \mathfrak{T}(V_k), i \in \mathbb{P}(V_k)} (\mathbf{k}, P_i)$ . Here  $\mathbf{k}$  denotes the path  $(k_{l_1}, k_{l_2}, \dots, k_{l_h})$  constructed by branching consecutively along the  $k_{l_1}$ -th branch from the top level, then the  $k_{l_2}$ -th branch from the lower level and finally  $k_{l_h}$ -th branch from the  $h$ -th level. The set  $\mathfrak{T}(V_k)$  denotes the set of all paths from the root to a leaf in the property tree of  $V_k$ . The set  $\mathbb{P}(V_k)$  denotes the indices of the vertices in the property tree of  $V_k$ .

$V_1$  $  \begin{array}{c}  P_1 \ P_2 \ P_3 \\  \backslash \   \ / \ \dots \\  P_k \ \dots \ o \ \dots P_i \\  \dots \ / \   \ \backslash \ \dots \\  P_{j+1} \ P_j \ P_{j-1}  \end{array}  $	$V_2 \ \backslash$  $  \begin{array}{c}  P_1 \ \dots \ o \ \dots P_2 \\  \backslash \   \ \backslash \\  P_3 \ P_4 \ P_5 \ \dots \\  \backslash \ / \ \backslash \ \backslash \\  P_6 \ P_7 \ P_8 \ P_9 \\    \\  P_{10}  \end{array}  $	$\dots$  $V_9$  $  P_1 \ \dots \ o \ \dots P_2  $
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------

Expressing the property tree of  $V_1$  with the algebraic notation discussed earlier:

$$ptree(V_1) = \sum_{k_j \in \mathfrak{T}(V_1), i \in \mathbb{P}(V_1)} (k_j, P_j)$$

Similarly,  $ptree(V_2)$  is given with

$$\begin{aligned}
 ptree(V_2) = & (1, P_0) + (k_1, P_1) + (k_2, P_2) + (k_3, P_3) + (k_4, P_4) + (k_5, P_5) + (k_3 k_1, P_6) + (k_3 k_2, P_7) \\
 & + (k_5 k_1, P_8) + (k_5 k_2, P_9) + (k_3 k_1 k_1, P_{10})
 \end{aligned}$$

Here  $P_0$  is  $text(V_2)$ .

Now if we expand the property trees for each  $V$ -particle in the semantic tree for the composite particle  $V_{comp}$  we will have a larger augmented property tree. This augmented property tree represents the semantic structure of  $V_{comp}$  and can be recorded in a matrix form which is the semantic signature of  $V_{comp}$ . The semantic signature matrix of  $V_{comp}$  will have the following structure in block matrix notation:

$$ssig(V_{comp}) = [\mathbf{P}_0 \mid \mathbf{A}_{0,1} \mid \mathbf{P}_1 \mid \mathbf{P}_0 \mid \mathbf{A}_{0,2} \mid \mathbf{P}_2 \mid \mathbf{P}_0 \mid \mathbf{A}_{0,3} \mid \mathbf{P}_3 \mid \dots \mid \mathbf{P}_p \mid \mathbf{A}_{p,q} \mid \mathbf{P}_q]$$

Here the matrices  $\mathbf{P}_i$  represent the signatures of the properties  $P_i$ . The size of  $\mathbf{P}_i$  is  $K_i \times N$ .

Here the symbol  $\mid$  denotes a separator column.

The last matrix can be rewritten in block matrix notation:

$$ssig(V_{comp}) = [\mathbf{B}_1 | \mathbf{B}_2 | \mathbf{B}_3 | \dots | \mathbf{B}_q]$$

$$\mathbf{B}_1 = [\mathbf{P}_0 | \mathbf{A}_{0,1} | \mathbf{P}_1], \mathbf{B}_2 = [\mathbf{P}_0 | \mathbf{A}_{0,2} | \mathbf{P}_2], \mathbf{B}_3 = [\mathbf{P}_0 | \mathbf{A}_{0,3} | \mathbf{P}_3], \dots, \mathbf{B}_q = [\mathbf{P}_p | \mathbf{A}_{p,q} | \mathbf{P}_q]$$

Here the block matrix  $\mathbf{B}_1$  fully describes the property  $P_1$  including how it is connected to the property tree  $ptree(V_1)$ . Similarly,  $\mathbf{B}_2$  and  $\mathbf{B}_3$  fully describes the properties  $P_2$  and  $P_3$  and their connectivity to  $ptree(V_1)$ . Finally,  $\mathbf{B}_q$  fully describes the property  $P_q$  and its connectivity to  $ptree(V_9)$ . From now on we will denote the block matrices  $\mathbf{B}_i$  as *semantic elements* of  $V_{comp}$ .

*Statement:* Every semantic particle, primitive or composite, can be represented as a sequence of *semantic elements*.

*Definition:* Semantic distance between two semantic elements  $B_1$  and  $B_2$

The semantic element  $B_1$  represents two properties -  $P_i$  and  $P_j$  connected through association link  $A_{i,j}$ . The properties  $P_i$  and  $P_j$  are represented by their property signature matrices  $\mathbf{P}_i$  and  $\mathbf{P}_j$ . The association link  $A_{i,j}$  is represented with its association matrix  $\mathbf{A}_{i,j}$  and semantic significance vector  $\mathbf{w}_{i,j}$ . (Note: Sometimes for clarity all vectors in a block matrix representing semantic element will be denoted with the vector symbol  $\vec{\phantom{x}}$  when clear distinction needs to be made). For details refer to the document [Note On Binding of Association Property to Semantic Properties](#). Similarly, the semantic element  $B_2$  represents the properties  $P_k$  and  $P_l$  connected through association link  $A_{k,l}$ . As before the properties  $P_k$  and  $P_l$  are represented by their property signatures  $\mathbf{P}_k$  and  $\mathbf{P}_l$ . The association link  $A_{k,l}$  is represented with its association matrix  $\mathbf{A}_{k,l}$  and semantic significance vector  $\mathbf{w}_{k,l}$ .

Let  $\mathbf{B}_1$  denotes the matrix of the first semantic element  $B_1$  given with  $\mathbf{B}_1 = [\mathbf{P}_i \ \mathbf{A}_{i,j} \ \mathbf{P}_j]$

Let  $\mathbf{B}_2$  denotes the matrix of the second semantic element  $B_2$  given with  $\mathbf{B}_2 = [\mathbf{P}_k \ \mathbf{A}_{k,l} \ \mathbf{P}_l]$

Then the semantic distance between the two is given with:

$$sdist(B_1, B_2) = sdist(P_i, P_k) + sdist(A_{i,j}, A_{k,l}) + sdist(P_j, P_l)$$

where

$$sdist(P_i, P_k) = |\mathbf{p}_i - \mathbf{p}_k|, \quad sdist(P_j, P_l) = |\mathbf{p}_j - \mathbf{p}_l|$$

$$sdist(A_{i,j}, A_{k,l}) = |\mathbf{w}_{i,j} - \mathbf{w}_{k,l}| \times sdist(\mathbf{a}_{i,j}, \mathbf{a}_{k,l})$$

*Definition:* The semantic distance of two semantic matrices  $\mathbf{a}$  and  $\mathbf{b}$  which have the same number of columns is given with:

$$sdist(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^n |\vec{\mathbf{a}}_i - \vec{\mathbf{b}}_i| \text{ where } \mathbf{a} = [\vec{\mathbf{a}}_1 \ \vec{\mathbf{a}}_2 \ \dots \ \vec{\mathbf{a}}_n] \text{ and } \mathbf{b} = [\vec{\mathbf{b}}_1 \ \vec{\mathbf{b}}_2 \ \dots \ \vec{\mathbf{b}}_n].$$

In the block matrix for  $ssig(V_{comp})$   $\mathbf{p}_0$  denotes the signature column vector of the property  $P_0$ ,  $\mathbf{a}_{0,1}$  denotes the association matrix of the arc between property  $P_0$  and property  $P_1$ ,  $\mathbf{a}_{p,q}$  denotes the association matrix of the arc between property  $P_p$  and  $P_q$ . Let us denote the number of rows of  $ssig(V_{comp})$  by  $N$  and the number of columns by  $M$ .

The semantic signature matrix  $ssig(V_{comp})$  can be decomposed as a sum of two intrinsic structural matrices – property signature matrix  $psig(V_{comp})$  and connectivity signature matrix  $csig(V_{comp})$ :

$$ssig(V_{comp}) = psig(V_{comp}) + csig(V_{comp})$$

$$psig(V_{comp}) = [\mathbf{p}_0 \ 0 \ \mathbf{p}_1 \ \mathbf{p}_0 \ 0 \ \mathbf{p}_2 \ \mathbf{p}_0 \ 0 \ \mathbf{p}_3 \ \dots \ \mathbf{p}_p \ 0 \ \mathbf{p}_q]$$

$$csig(V_{comp}) = [0 \ \mathbf{a}_{0,1} \ 0 \ 0 \ \mathbf{a}_{0,2} \ 0 \ 0 \ \mathbf{a}_{0,3} \ 0 \ \dots \ 0 \ \mathbf{a}_{p,q} \ 0]$$

Let us denote by  $psig(P_1, V_{comp})$  the augmented semantic property signature of property  $P_1$  with respect to  $V_{comp}$ . It is given with:

$$psig(P_0, V_{comp}) = [\mathbf{p}_0 \ 0 \ 0 \ \mathbf{p}_0 \ 0 \ 0 \ \mathbf{p}_0 \ 0 \ 0 \ \dots \ 0 \ 0 \ 0]$$

Similarly,

$$psig(P_1, V_{comp}) = [0 \ 0 \ \mathbf{p}_1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \dots \ 0 \ 0 \ 0]$$

$$psig(P_q, V_{comp}) = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \dots \ 0 \ 0 \ \mathbf{p}_q]$$

Then obviously:

$$psig(V_{comp}) = \sum_{k \in \mathbb{S}(V_{comp})} \sum_{i \in \mathbb{P}(V_k)} psig(P_i, V_{comp})$$

Here  $\mathbb{S}(V_{comp})$  denotes the set of the indices of all semantic particles which the composite  $V_{comp}$  is composed from.

Another way to partition the signature matrix into block matrices is:

$$ssig(V_{comp}) = [\mathbf{V}_1 \ \mathbf{A}_{1,2} \ \mathbf{V}_2 \ \mathbf{A}_{1,3} \ \mathbf{V}_3 \ \dots \ \mathbf{A}_{6,8} \ \mathbf{V}_8 \ \mathbf{A}_{6,9} \ \mathbf{V}_9]$$

The block matrix  $\mathbf{V}_1$  represents the property tree of the particle  $V_1$  and it is given by:

$$\mathbf{V}_1 = [\mathbf{p}_0 \ \mathbf{a}_{0,1} \ \mathbf{p}_1 \ \mathbf{p}_0 \ \mathbf{a}_{0,2} \ \mathbf{p}_2 \ \mathbf{p}_0 \ \mathbf{a}_{0,3} \ \mathbf{p}_3 \ \dots \ \mathbf{p}_0 \ \mathbf{a}_{0,k} \ \mathbf{p}_k]$$

The block matrix  $\mathbf{A}_{1,2}$  describes the connection between the particles  $V_1$  and  $V_2$  connecting the root property  $\mathbf{p}_0$  of  $V_1$  and the root property  $\mathbf{p}_{k+1}$  of  $V_2$ . It is given with:

$$\mathbf{A}_{1,2} = [\mathbf{p}_0 \ \mathbf{a}_{0,k+1} \ \mathbf{p}_{k+1}] \text{ //TODO: expand it – the matrix structure is more complicated!}$$

### Properties of the signature matrix

Here are some interesting properties of  $ssig(V_{comp})$ :

The number of rows  $N$  in  $ssig(V_{comp})$  is  $3 \times$  the number of arcs in the augmented property tree of  $V_{comp}$ .

The rank of

TO DO: finish the property section

### Asymptotic closeness of semantic structures

Let us have two semantic structures S1 and S2.

$$ssig(S_1) = \begin{bmatrix} \mathbf{V}_{k_1} & \mathbf{A}_{k_1, k_2} & \mathbf{V}_{k_2} & \mathbf{A}_{k_2, k_3} & \mathbf{V}_{k_3} & \dots & \mathbf{A}_{k_p, k_q} & \mathbf{V}_{k_q} \end{bmatrix}$$

$$ssig(S_2) = \begin{bmatrix} \mathbf{V}_{l_1} & \mathbf{A}_{l_1, l_2} & \mathbf{V}_{l_2} & \mathbf{A}_{l_2, l_3} & \mathbf{V}_{l_3} & \dots & \mathbf{A}_{l_r, l_s} & \mathbf{V}_{l_s} \end{bmatrix}$$

Uniform asymptotic closeness

$K$ -level uniform asymptotic closeness