The Signature of Semantic Structures

D. Gueorguiev 10/2/2021

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.

$ptree(V_k) \;=\; \sum_{\boldsymbol{k}\in\mathfrak{T}(V_k), i\in\mathbb{p}(V_k)} (\boldsymbol{k}, P_i)$. Here $\boldsymbol{k}$ 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.

```
        V1                         V2  `              ...        V9

   P1 P2  P3                 P1--- o --- P2              P1--- o --- P2
     \ | /  ..                    / | \
 Pk --- o --- Pi               P3  P4  P5_
  .. / | \  ..                / \       |  \
  Pj+1 Pj  Pj-1              P6  P7     P8   P9
                             |
                             P10
```

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

$$ptree(V_2) = (1, P_0) + (k_1, P_1) + (k_2, P_2) + (k_3, P_3) + (k_4, P_4) + (k_5, P_5) + (k_3k_1, P_6) + (k_3k_2, P_7) + (k_5k_1, P_8) + (k_5k_2, P_9) + (k_3k_1k_1, P_{10})$$

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:

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

The last matrix can be rewritten in block matrix notation:

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

$\mathbf{B}_1 = [\mathbf{p}_0\ \ \mathbf{a}_{0,1}\ \ \mathbf{p}_1]$, $\mathbf{B}_2 = [\mathbf{p}_0\ \ \mathbf{a}_{0,2}\ \ \mathbf{p}_2]$, $\mathbf{B}_3 = [\mathbf{p}_0\ \ \mathbf{a}_{0,2}\ \ \mathbf{p}_3]$, . . . , $\mathbf{B}_q = [\mathbf{p}_p\ \ \mathbf{a}_{p,q}\ \ \mathbf{p}_q]$

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 signatures $\mathbf{p}_i$ 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 $\overrightarrow{}$ when clear distinction needs to be made.) 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

$$sdist(P_i, P_k) = |\mathbf{p}_i - \mathbf{p}_k|,\ \ sdist(P_j, P_l) = |\mathbf{p}_j - \mathbf{p}_l|$$
$$sdist(A_{i,j}, A_{k,l}) = |\mathbf{w}_{i,j} - \mathbf{w}_{k,l}| \times sdist(\mathbf{a}_{i,j}, \mathbf{a}_{k,l})$$

Definition: The semantic distance of two semantic matrices $\mathbf{a}$ and $\mathbf{b}$ which have the same number of columns is given with:
$sdist(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} |\vec{\mathbf{a}}_i - \vec{\mathbf{b}}_i|$ where $\mathbf{a} = [\vec{\mathbf{a}}_1\ \vec{\mathbf{a}}_2 \dots \vec{\mathbf{a}}_n]$ 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:

$$\mathbf{A}_{1,2} = [\mathbf{p}_0 \;\; \mathbf{a}_{0,k+1} \; \mathbf{p}_{k+1}]$$

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 × 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.

$$ssig(S_1) = \left[\mathbf{V}_{k_1}\ \mathbf{A}_{k_1,k_2}\ \mathbf{V}_{k_2}\ \mathbf{A}_{k_1,k_3}\ \mathbf{V}_{k_3}\ \dots\ \mathbf{A}_{k_p,k_q}\ \mathbf{V}_{k_q}\right]$$

$$ssig(S_2) = \left[\mathbf{V}_{l_1}\ \mathbf{A}_{l_1,l_2}\ \mathbf{V}_{l_2}\ \mathbf{A}_{l_1,l_3}\ \mathbf{V}_{l_3}\ \dots\ \mathbf{A}_{l_r,l_s}\ \mathbf{V}_{l_s}\right]$$

Uniform asymptotic closeness

K-level uniform asymptotic closeness