Connecting Semantically Related Structures

D. Gueorguiev 10/22/21

Let us start with the V-particles V_k, V_1 and V_2 which are not composite and are ***related*** semantically. How to connect them?

```
   V_k              V_1              V_2
  /   \            /   \            /   \
V_1   V_2        V_2   V_k        V_1   V_k
```

What does it mean for two primitive V-particles V_1 and V_2 to be *related* semantically? Note that there is a difference between the terms *related semantically* and *semantically close*. The difference will become clear with the discussion here.

----needs clarification – should we do quantization in terms the new compound property color or energy level makes more sense

There should be sufficient attraction force $F^a(V_1, V_2)$ between them in at least one of the possible connectivity DAGs.

But what is this mysterious attraction force? How is it represented? Let us denote by G the smallest DAG which includes both V_1 and V_2.

~~The engagement of the V-particles in a parent-children ensemble is based on the property ***color***. The property ***color*** is a compound property of primitive V-particles. ***Color*** is made of a specific set of property keys forming a ***color basis***. Each primitive V-particle has a subset of property keys from the color basis. The parent-children ensemble like the ones depicted above are possible only when the colors of the participating particles are matching the expected color for the position (tree node) of each particle in the ensemble.~~

For each V-particle there are defined the following intrinsic quantities:

- Information content
- Valence

Information Content of a particle

//TO DO

Note on particle Valence

These are number of property subsets on each of which another particle may lock onto.

Let us take the example:

```
   am
 /    \
I      Dimitar
```

Let us denote with $\mathfrak{P}_1$ the property set of the verb which have gathered all properties dealing with subject matters – these are the properties which describe the plurality of the verb, the point of view,

and the kind subjects from semantical standpoint allowed to lock on this verb. Similarly, we denote with $\mathfrak{P}_2$ the properties of the verb which deal with object matter (the recipient of the verb action). In general for each V-particle we can have a finite number of property sets with a different A-particle latching onto each of them. The k sets $\mathfrak{P}_1$, $\mathfrak{P}_2$, …, $\mathfrak{P}_k$ from now on will be denoted as *Valence Sets*. Once one of the valence sets is occupied (i.e. locked onto by an A-particle) it may impose additional constraints on the properties related to the free *valence sets* and thus influence the choice of A-particle locking onto a free *valence set*. The maximum number of *valence sets* associated with a V-particle is an intrinsic property of the particle and it will be named *Valence*; for particle V its *Valence* will be denoted with $|V|$. Each V-particle "knows" when a connection (A-particle) has locked onto its properties and maintains an internal state recording the occupied number of *valence sets*. The relative position of the two particles in the semantic graph will depend on the value of *Valence* × *Information Content* for each of them. This property from now on will be known as *Particle Mass* and denoted by M:
$Particle\ Mass = Valence \times Information\ Content$ or in symbol notation:
$M_V\ = |V| \times IC$

Postulate: The parent particle has a larger mass, that is value of $Valence \times Information\ Content$, compared to its children.

The notion of semantic valence of a semantic structure

A newly formed semantic structure S_{new} has subsets of particles $\mathfrak{V}_1, \mathfrak{V}_2, \ldots, \mathfrak{V}_l$ each of which is *semantically linked* to a subset of particles from another (existing) semantic structure. To be specific, the particle subset $\mathfrak{V}_1$ of S_{new} is linked to the particle subset $\mathfrak{W}_1$ of the structure S_1, the subset $\mathfrak{V}_2$ of S_{new} is linked to the particle subset $\mathfrak{W}_2$ of the structure S_2, …, $\mathfrak{V}_l$ of S_{new} is linked to $\mathfrak{W}_l$ of S_l.
//TODO connect this with the paragraph *The notion of effective mass of semantic structure*

Note on Semantic Link between two structures

Semantic link represents specific relation between two semantic structures. For instance, *is-a* semantic link between two structures is established when each of the two structure denotes the same semantic concept. Of course, we could only know if a new concept denotes an old semantic concept when we make certain assumptions. Therefore, we assign a semantic significance vector $\mathbf{W}$ when we evaluate a semantic link between two structures. We associate a random variable with this semantic link. In order to find the set of most likely semantic links for given structure S we are going to build a spanning Bayesian network.

Is-relation:
$S_i \overset{p}{\leftrightarrow} S_j$: S_i *is-a* S_j with semantic significance vector W; the associated random variable will be denoted with $I_{i,j}$. The structures S_i and S_j have the same semantic meaning. Two semantic structures have the same semantic meaning when the semantic distance between them is small enough. Evaluating semantic distance involves evaluating their respective semantic signatures.

Has-a relation:
$S_i \rightarrow S_j$: S_i *has-a* S_j with semantic significance vector W; the associated random variable will be denoted with $I_{i,j}$. The structures S_i and S_j are related such that there exists substructure of S_i with close enough semantic distance to that of S_j.

The notion of mass of a semantic structure

When we talk about a mass of a semantic structure S we account for the following aspects which are relevant to the notion of mass:

- The aggregate mass of S, denoted with $M^*(S)$, is obtained by summing up the mass of all V-particles which belong to S. Obviously, the presence of V-particles with high information content and the presence of more verbs will increase the aggregate mass of the structure.
- The connectedness of the semantic structure to the enclosing context(s). The structure of the connections of S to the enclosing context(s) will determine which will be the parent semantic structure(s) of S. The inbound and outbound connections to S will determine how the effective mass of S, denoted with M^e, will change compared to the aggregate mass M^*.

//TODO

The notion of effective mass of a semantic structure

Let us have semantic structure S which is connected to a set of enclosing contexts C_1, C_2, …, C_K by similarity links (*is-a* relationships described earlier; for details see also Relations Between Semantic Structures). The directionality of the similarity links matters – the inbound similarity links, denoted by $SA^-(S)$, represent concepts which are defined outside of S and will lead to a decrease of the effective mass of S. On other side, the outbound links, denoted with $SA^+(S)$, represent concepts which are defined in S and are referred in other structure outside of S. So outbound links will lead to an increase of the effective mass of S. One can write the following relation for calculation of the effective mass M^e of the structure S:

$$M^e(S) = M^*(S) + \sum_{i=1}^{N^+} \mathcal{C}(S_i, C_k) \times M^e(S_i) \times SA^+(S_i) - \sum_{j=1}^{N^-} \mathcal{C}(S^j, C_k) \times M^e(S^j) \times SA^-(S^j)$$

Here $S_i \subset S$ are substructures of S which are referred to by outside structures within the context C_k.

//TODO

Recall that we postulated that the parent V-particle has the larger value of $Valence \times Information\ Content$ compared to its children.

Let us have the structures S_k, S_1 and S_2 which are close semantically. How to connect them?

S_k	S_1	S_2
/ \	/ \	/ \
S_1 S_2	S_2 S_k	S_1 S_k

---- S_1 ----------- > S_2 < -
| | AS_1 ^ | $AS_1, AS_2, \ldots, AS_k$ – directed association links connecting S_1 to S_2
| |___________| |
| AS_2 |
| . . . |
|____________________|
AS_k

The structure of an association link
Association link connects two V-particles on two different semantic structures.