Note on modeling binding and repulsion force in semantic properties

D. Gueorguiev 1/17/2022

We already have stated that the internal structure of a semantic property can be represented by a set of semantic regions occupying a subset of semantic dimensions. Each region denotes a specific semantic aspect of the property. Thus, the total binding / repulsion force is equal to the sum of the of the binding forces between all relevant region pairs minus the sum of the repulsion forces between all relevant region pairs $(\mathbf{r}_a, \mathbf{r}_b)$:

$$f(\mathbf{p}_1, \mathbf{p}_2) = \sum_{a,b} f^+(\mathbf{r}_a, \mathbf{r}_b) + \sum_{c,d} f^-(\mathbf{r}_c, \mathbf{r}_d)$$

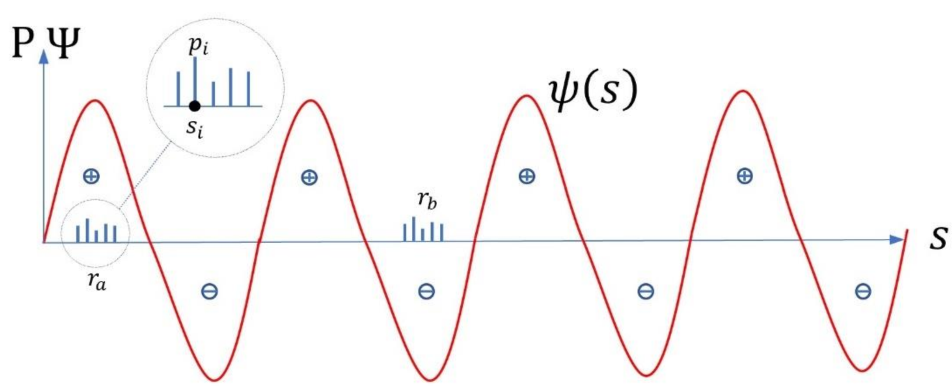
The relevant region pairs $(\mathbf{r}_a, \mathbf{r}_b)$ are defined as follows. Let us sort the pairs of regions from $\mathbf{p}_1$ and $\mathbf{p}_2$ by the absolute value of the binding force.

Definition: *relevant region pair* $(\mathbf{r}_a, \mathbf{r}_b)$ is such pair which has absolute binding force value ***not exclusively in*** the first ℓ -quantile for some $\ell > 0$. In other words, all region pairs which are ***only in*** the first ℓ-quantile are *irrelevant*.

The question now is how do we want to model the binding / repulsion force between a pair of regions. Here we are proposing a possible way to calculate the binding and repulsion forces and will discuss why it is useful to be done this way.
A pair of regions $(\mathbf{r}_a, \mathbf{r}_b)$ from the properties $\mathbf{p}_1$ and $\mathbf{p}_2$ are depicted on the discrete horizontal axis s on the Figure below. The horizontal axis s is discrete in nature and represents the entire set of semantic dimensions for every point in Semantic Space. Let us imagine that region $\mathbf{r}_a$, composed of a set of semantic values p_i, $i = 1..\dim(\mathbf{r}_a)$, will somehow *generate a semantic energy* ψ_a which will span the entire horizontal axis s and propagate as wave. This wave is depicted in red in the Figure below. Each region will encode the parameters of the impulse $\psi_a(s)$ in its values p_i. $\psi_a(s)$ will, in general, span all dimensions of the semantic space i.e. the integer coordinate s. Obviously, ψ_a will be periodic function along the semantic dimensions axis s. As we said the amplitude A, the frequency ω and the phase φ of the impulse are somehow encoded in a portion of each region values. Hence, we can write:

$\psi_a = \psi_a(s; A, \omega, \varphi)$ and $\mathbf{r}_a = \mathbf{r}_a(A, \omega, \varphi)$

Now let us introduce the second region $\mathbf{r}_b$ coming from the other property $\mathbf{p}_2$.



The region $\mathbf{r}_b$ is composed of a set of semantic values p_j, $i = 1..\dim(\mathbf{r}_b)$ which *generate an energy wave* ψ_b.

We postulate that the two regions will interact with each other through binding or repulsive force only if the frequencies of the corresponding energy waves ***are the same*** i.e. $\omega_a = \omega_b = \omega$.

We postulate that region $\mathbf{r}_a$ will attract region $\mathbf{r}_b$ *iff*:

1. The frequencies of the corresponding to each region energy wave are the same i.e. $\omega_a = \omega_b = \omega$
2. $f_a = \sum_{i \in a} \psi_a(s_i) > 0$; $\psi_a(s_i)$ is the value of the energy wave at the i-th dimension of region $\mathbf{r}_a$
3. $f_b = \sum_{j \in b} \psi_b(s_j) < 0$; $\psi_b(s_j)$ is the value of the energy wave at the j-th dimension of region $\mathbf{r}_b$

Then the attraction force between the two regions $\mathbf{r}_a$ and $\mathbf{r}_b$ will be given by the product of the absolute values:

$$f^+(\mathbf{r}_a, \mathbf{r}_b) = |f_a||f_b|$$

If $sign(f_a) = sign(f_b)$ then we have a repulsive force instead of attracting one:

$$f^-(\mathbf{r}_a, \mathbf{r}_b) = -|f_a||f_b|$$

If we have more than one region with the same frequency ω in one of the properties we sum them up and then multiply with the sum of the regions of the other property:

$$f(\mathbf{p}_1, \mathbf{p}_2; \omega) = \sum_{a,b} f^+\big(\mathbf{r}_a(\omega), \mathbf{r}_b(\omega)\big) + \sum_{c,d} f^-\big(\mathbf{r}_c(\omega), \mathbf{r}_d(\omega)\big)$$

Finally, the total binding/repulsive force between the two properties is given as the sum of all binding/repulsive forces on all frequencies:

$$f(\mathbf{p}_1, \mathbf{p}_2) = \sum_{\omega} f(\mathbf{p}_1, \mathbf{p}_2; \omega)$$

Relation between Semantic Mass and Semantic Energy

Let us consider a property P which has some number of non-zero regions:
$\mathbf{p} = [\mathbf{r}_1, \mathbf{0}, \mathbf{r}_2, \mathbf{0}, \dots, \mathbf{0}, \mathbf{r}_k]^T$ where $\mathbf{r}_i$, $i = 1..k$ are the proper regions in the property signature $\mathbf{p}$.
Let us group the regions by two different criteria – size and frequency. The sets in which the grouping of the regions is done by region size will be denoted with $\mathcal{S}$. The sets in which the grouping is done by the frequency of semantic energy will be denoted with $\mathcal{F}$.
For the grouping by region size we have the following sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_l$, $l \le k$. Here the regions which have the largest size are in set $\mathcal{S}_1$**,** the second largest regions are in $\mathcal{S}_2$, …, and the smallest regions are in $\mathcal{S}_l$**.**

For the grouping by energy wave frequency we have the following sets $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_m$, $m \le k$**.**
Every region which generates a wave with the largest frequency ω_1 will be in set $\mathcal{F}_1$, every region which generates wave with second largest frequency ω_2 will be in set $\mathcal{F}_2$,…, and the regions with the lowest frequency ω_m will be in set $\mathcal{F}_m$.

Each semantic region $\mathbf{r}_i$ is constructed by a repetition of a single atomic block $\aleph_i$:

$\mathbf{r}_i = \aleph_i\aleph_i \ldots \aleph_i$ ($\aleph_i$ repeats α times). The atomic block $\aleph$ contains only few data points over a subset of semantic dimensions and encodes the frequency and phase of the energy wave.
The amplitude of the generated energy wave is proportional to the size (length) of the region. In other words, the larger is the repetition count α the larger will be the amplitude A of the energy wave.
Based on this we can write the energy function $\psi_{\mathbf{r}}$ of region $\mathbf{r}$ as:

$\psi_{\mathbf{r}}(s) = A(\alpha)\sin\big(\omega(\aleph)s + \varphi(\aleph)\big)$ where $\mathbf{r} = \aleph\aleph\ldots\aleph$ ($\aleph$ repeats α times)

and for the semantic mass $M_{\mathbf{r}}$ of the region $\mathbf{r}$:

$M_{\mathbf{r}} = C(\aleph)M(\alpha)$ where $C(\aleph)$ is some coefficient which possibly depends on the semantic values contained in $\aleph$ and $M(\alpha)$ is monotonously increasing function of the repetition count α.
These two expressions for $\psi_{\mathbf{r}}$ and $M_{\mathbf{r}}$ give us the relationship between semantic energy and semantic mass.

Let us return on our previous example with two P-particles each of which has a single region. Particle P_1 has a region $\mathbf{r}_1$ and Particle P_2 has a region $\mathbf{r}_2$. If the two regions are not built from atomic blocks which encode the same frequency the binding/repelling force between the two regions will be 0. Let us assume that the two regions are built from the same atomic block $\aleph$ so they share frequency ω and phase φ. Let us denote by α_1 the repetition count of $\aleph$ in region $\mathbf{r}_1$ and by α_2 the repetition count of $\aleph$ in region $\mathbf{r}_2$. Let us denote by $s_{i_1}, s_{i_2}, \ldots, s_{i_m}$ the semantic dimensions which region $\mathbf{r}_1$ is spanning. Similarly, by $s_{j_1}, s_{j_2}, \ldots, s_{j_n}$ we denote the semantic dimensions which region $\mathbf{r}_2$ is spanning. Clearly either m is multiple of n or n is multiple of m.

//TODO: finish this

Constructing the property tree: constraints and inequalities based on binding force

Let is consider the property tree of a V-particle:

V_1

P_0
P_1--- o --- P_2 <----------- P_{11}
/ | \
P_3 P_4 P_5_
/ \ | \
P_6 P_7 P_8 P_9
|
P_{10}

Let us imagine we want to add new P-particle to the property tree. The following steps toward forming a new ensemble take place:

Step 1. All P-particles which are about to participate in the new ensemble become disassociated / disentangled
Step 2. The common center of mass for the new ensemble is determined.

Step 3. The particle with the largest mass will be closest to the common center of gravity and will be root to the property tree

Special Cases:

a) If there are P-particles with the same mass which is the largest mass in the particle tree then the ensemble is ill-formed and one of the two properties has to have a region discarded so it will end up with lower semantic mass.

b) Let there be two P-particles P_j and P_k with the same mass which is smaller than the mass of the root P-particle P_0. Let the particle with the closest but larger semantic mass than that of P_j and P_k is P_i. Thus we have $M_{P_j} = M_{P_k} < M_{P_i}$. either will be immediate descendants of the particle with the closest larger semantic mass

//TODO: finish this