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 $(\mathbf{r}_a, \mathbf{r}_b)$:

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

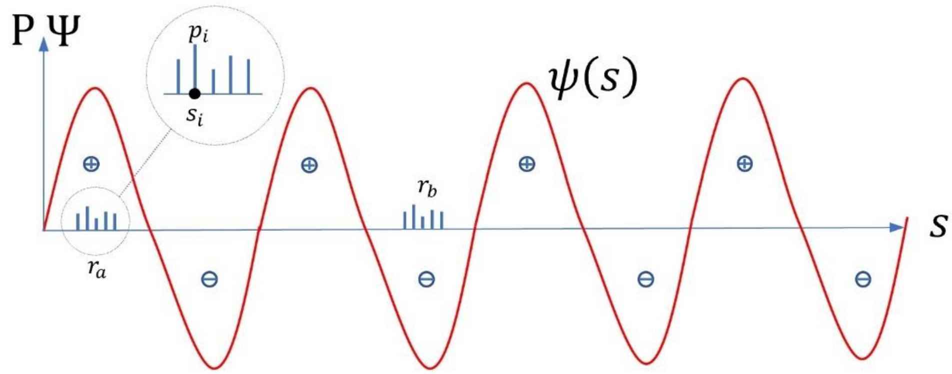
The relevant region pairs $(\mathbf{r}_a, \mathbf{r}_b)$ are defined as follows. Let us sort the pairs of regions from $\mathbf{p}_1$ and $\mathbf{p}_2$ by the absolute value of the binding force.

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

The question now is how 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 be able to interact with the surrounding *semantic energy* which will span the entire horizontal axis s and propagate as a *spectrum of energy waves*. One of those *semantic energy waves* is depicted in red in the Figure below. Each region $\mathbf{r}_a$ will be sensitive to specific frequency ω_a of the surrounding semantic energy function Ψ. The frequency sensitivity of each region is encoded in the region values p_i, $i = 1..\dim(\mathbf{r}_a)$ (or in a short notation $i \in a$). The semantic energy function $\Psi(s)$ will, in general, span all dimensions of the semantic space i.e. the values of the integer axis s. Obviously, each frequency component of the energy function ψ_a will be periodic function along the semantic dimensions axis s. As we said the Semantic Energy function $\Psi(s)$ can be decomposed in spectral components; thus the amplitude A, the frequency ω and phase φ are somehow encoded in the semantic energy wave ψ_a to which the region $\mathbf{r}_a$ is sensitive. The region $\mathbf{r}_a$ need to encode only the frequency ω_a to each it is sensitive and this will be sufficient to determine the amplitude A_a and the phase φ_a of the selected energy wave ψ_a. Hence, we can write:

$\psi_a = \psi_a(s; A_a, \omega_a, \varphi_a)$ and $\mathbf{r}_a = \mathbf{r}_a(s; \omega_a)$

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 $p_j, j = 1..\dim(\mathbf{r}_b)$ which are *sensitive to the energy wave* ψ_b.

So each semantic region $\mathbf{r}$ in a property is sensitive to a semantic energy wave with specific frequency encoded in the region values $p_j, j = 1..\dim(\mathbf{r})$. The semantic energy stored in the region depends on *the position* and *the width* of the semantic region along the semantic axis s.

Definition: Semantic energy of a property region
The Semantic energy $f(\mathbf{r})$ stored in the region $\mathbf{r}$ will be the sum of all values of the energy wave to which the region is sensitive $\psi(s_j)$ along all semantic dimensions $s_j, j = 1..\dim(\mathbf{r})$ which the region spans:

$$f(\mathbf{r}) = \sum_{j=1}^{\dim(\mathbf{r})} \psi(s_{\mathbf{r}} + s_j)$$

Here with $s_{\mathbf{r}}$ we denote the semantic dimension from which the region $\mathbf{r}$ starts.
Using an abbreviated notation instead of explicitly using the starting semantic dimension for a region we write:

$$f(\mathbf{r}) = \sum_{j \in \mathcal{S}(\mathbf{r})} \psi(s_j)$$

Here the set $\mathcal{S}(\mathbf{r})$ contains all semantic dimensions which the region $\mathbf{r}$ covers.

Definition: Semantic energy of a property $\mathbf{p}$
The semantic energy $f(\mathbf{p})$ stored in property $\mathbf{p}$ is given with the sum of the semantic energies stored in all relevant regions (the regions which energies are not in the first ℓ-quantile):

$$f(\mathbf{p}) = \sum_{\mathbf{r} \in \mathcal{R}(\mathbf{p})} f(\mathbf{r})$$

Here with $\mathcal{R}(\mathbf{p})$ we have denoted the set of all regions included in property $\mathbf{p}$.

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 <u>*are the same*</u> i.e. $\omega_a = \omega_b = \omega$.

Let us denote with f_a the following sum $f_a = \sum_{i \in a} \psi_a(s_i)$; Here $\psi_a(s_i)$ is the value of the energy wave at the i-th dimension of region $\mathbf{r}_a$
Let us denote with f_b the following sum $f_b = \sum_{j \in b} \psi_b(s_j)$; Here $\psi_b(s_j)$ is the value of the energy wave at the j-th dimension of region $\mathbf{r}_b$

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. $f_a f_b < 0$. In other words, either $f_a > 0 \wedge f_b < 0$ or $f_a < 0 \wedge f_b > 0$.

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 $f_a f_b > 0$ 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 of a property

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 $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_l$, $l \leq k$. 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 $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_m$, $m \leq k$.
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 $\aleph_i$:
$\mathbf{r}_i = \aleph_i \aleph_i \dots \aleph_i$ ($\aleph_i$ repeats α times). The atomic block $\aleph$ contains only few data points over a subset of semantic dimensions and encodes the frequency of the energy wave.

The amount of the stored energy wave in the region is proportional to the size (length) of the region. In other words, the larger is the repetition count α the larger will be the amount of the stored energy. Based on this we can write the energy wave $\psi_{\mathbf{r}}$ of which region $\mathbf{r}$ is sensitive to as:

$\psi_{\mathbf{r}}(s) = A\sin(\omega(\aleph)s + \varphi)$ 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 wave for region $\mathbf{r}$ and semantic mass of the region.

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

Relation between Binding/Repulsive Force and Semantic Energy of a property

Definition: **Mirror unit region** of a given property region $\mathbf{r}$: we denote it with $\mathbb{I}(\mathbf{r})$ and when it is obvious from the context to which region we are referring we will just use the symbol $\mathbb{I}$. Mirror unit region $\mathbb{I}(\mathbf{r})$ is a region which is sensitive to the same frequency ω as the region $\mathbf{r}$ it is mirroring. The semantic energy of the mirror unit region is with opposite sign to that of $\mathbf{r}$ and the absolute value of its semantic energy is 1:

$$\left|f(\,\mathbb{I}(\mathbf{r}))\right| = 1,\ sign\left(f(\,\mathbb{I}(\mathbf{r}))\right) = -sign(f(\,\mathbf{r}))$$

Statement: The semantic energy $f(\mathbf{r}) = \sum_{j\in\mathcal{S}(\mathbf{r})}\psi(s_j)$ stored in region $\mathbf{r}$ is equal to the binding force $f^+(\mathbf{r},\mathbb{I}(\mathbf{r}))$ between $\mathbf{r}$ and $\mathbb{I}(\mathbf{r})$ i.e.:

$$f(\mathbf{r}) = f^+(\mathbf{r},\mathbb{I})$$

Definition: **Mirror unit property** of a given property P: we denote it with $\mathbb{I}(\mathbf{p})$ and when it is obvious from the context to which region we are referring we will just use the symbol $\mathbb{I}$. Mirror unit property $\mathbb{I}(\mathbf{p})$ is such property which has the same number of regions as $\mathbf{p}$ and those regions of the mirror property are sensitive to the same set of frequencies as the regions of $\mathbf{p}$. Written succinctly:

$$\forall\,\mathbf{r}\in\mathbf{p}\ \exists\,\mathbb{I}(\mathbf{r})\in\mathbb{I}(\mathbf{p}) \wedge \forall\,\mathbb{I}(\mathbf{r})\in\ \mathbb{I}(\mathbf{p})\ \exists\,\mathbf{r}\in\mathbf{p}$$

Statement: The semantic energy $f(\mathbf{p}) = \sum_{\mathbf{r}\in\mathcal{R}(\mathbf{p})} f(\mathbf{r})$ stored in property $\mathbf{p}$ is equal to the binding force $f^+\big(\mathbf{p}, \mathbb{I}(\mathbf{p})\big)$ between $\mathbf{p}$ and $\mathbb{I}(\mathbf{p})$ i.e.:

$$f(\mathbf{p}) = f^+(\mathbf{p}, \mathbb{I})$$

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 semantic mass for the new ensemble is determined as we already know the semantic masses of all properties in the new ensemble. For details about determining of the semantic mass of an ensemble of properties follow the discussion in On The Semantic Position Of Semantic Structures.

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 Case:
there are two P-particles with the same semantic mass which happens to be the largest mass in the particle tree. Then the particle with the lower semantic energy will be closer to the semantic center than the particle with the higher semantic energy. (not sure about the energy condition). An ensemble in which the heaviest two particles are having the same masses and same energies is ill-formed and one of the two properties has to have a region discarded so it will end up with lower semantic mass.

Step 4. Let us have two particles P-particles P_j and P_k such that $M_{P_j} \geq M_{P_k}$. Let us denote with P_i the particle with the closest but larger semantic mass than that of P_j and P_k. Thus $M_{P_i} > M_{P_j} \geq M_{P_k}$. Then each one of the following configurations are possible:

a)	*b)*	*c)*	*d)*	*e)*	*f)*	*g)*
P_i	P_i	P_i	P_i	P_i	P_i	P_i
/ \	/ \	|	|	|	|	x x
P_j P_k	P_k P_j	P_j	P_k	P_j	P_k	P_j P_k
		|	|	x	x	
		P_k	P_j	P_k	P_j	

Case a) will occur when there is non-zero binding force $f_{i,j}^{+} = f^{+}(\mathbf{p}_i, \mathbf{p}_j) > 0$ between P_i and P_j and also between P_i and P_k - $f_{i,k}^{+} = f^{+}(\mathbf{p}_i, \mathbf{p}_k) > 0$. In this case either $M_{P_j} > M_{P_k}$or $M_{P_j} = M_{P_k}$and $f_{i,j}^{+} > f_{i,k}^{+}$.

Case b) will occur when there is non-zero binding force $f^{+}(\mathbf{p}_i, \mathbf{p}_j) > 0$ between P_i and P_j and also between P_i and P_k - $f^{+}(\mathbf{p}_i, \mathbf{p}_k) > 0$. In this case $M_{P_j} = M_{P_k}$and $f_{i,k}^{+} > f_{i,j}^{+}$.

Case c) will occur when $f_{i,j}^{+} > 0$, $f_{i,k}^{+} = 0$, and $f_{j,k}^{+} > 0$ when either $M_{P_j} > M_{P_k}$or $M_{P_j} = M_{P_k}$.

Case d) will occur when $f_{i,k}^{+} > 0$, $f_{i,j}^{+} = 0$, and $f_{j,k}^{+} > 0$ when $M_{P_j} = M_{P_k}$.

Case e) will occur when $f_{i,j}^{+} > 0$ and $f_{i,k}^{+} = 0$

Case f) will occur when $f_{i,j}^{+} = 0$ and $f_{i,k}^{+} > 0$

Case g) will occur when $f_{i,j}^{+} = 0$ and $f_{i,k}^{+} = 0$

//TODO: finish this