

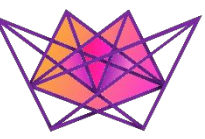
	<b>Evgeniy N. Pavlovskiy, Ph.D., head of the NSU Stream Data Analytics and Machine Learning lab.</b>			
<b>Quan</b>	1	0	1	1
<b>Tum</b>	0	1	0	0
<b>Se</b>	1	0	0	-1
<b>Man</b>	All	Things	Are	Number
<b>Tics</b>	1	0	Pythagoras	



# Bra-ket notation (Dirac, 1939)

- $V$  – vector (Hilbert) space,  $\mathbb{F}$  – field
- $|\psi\rangle$  – *pure state* (vector, or operator  $\mathbb{F} \rightarrow V$ ) – ket
- $\langle\psi|$  – *effect of state*  $|\psi\rangle$  (dual vector, dual operator  $V \rightarrow \mathbb{F}$  Hermitian conjugate) – bra
- *Inner product* of  $|\psi\rangle$  and  $|\varphi\rangle$  is  $\langle\psi|\varphi\rangle$





# Outer product

- *Outer product*  $|w\rangle\langle v|$  for  $w \in W, v \in V$  is a  $V \rightarrow W$  operator:

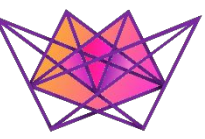
$$(|w\rangle\langle v|)(|v'\rangle) \equiv |w\rangle\langle v|v'\rangle = \langle v|v'\rangle|w\rangle$$

- Arbitrary  $A: V \rightarrow W$  can be written

in a basis  $\{|v_i\rangle\}_i$  for  $V$  and  $\{|w_j\rangle\}_j$  for  $W$ :

$$A = \sum_{ij} m_{ij} |w_j\rangle\langle v_i|,$$

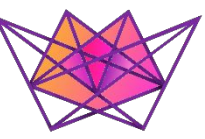
where  $m_{ij} = \langle w_j|A|v_i\rangle$



# Eigenvectors and eigenvalues

- $A = \sum_i \lambda_i |i\rangle\langle i|$ , if  $|i\rangle$  is an orthogonal basis in which  $A$  is diagonal.
- $|i\rangle$  – are *eigen vectors*
- $\lambda_i$  – are *eigen values*
- Easy to check:

$$A|i\rangle = \lambda_i|i\rangle$$



# Density operator (matrix)

- If  $|\varphi_i\rangle$  – are pure states,
- $\{p_i\}$  – are probabilities over them, then

$$\rho \equiv \sum_i p_i |\varphi_i\rangle\langle\varphi_i| \text{ is a } \textit{dense operator}$$

- *Positive operator.*

$$\langle v|A|v\rangle \geq 0 \text{ for all } v$$

- **Theorem:**  $\rho$  is a density operator iff it's a positive Hermitian operator with trace = 1.



# Trace inner product

- $A$  and  $B$  are density matrices same dimension

- $tr(A^T B)$

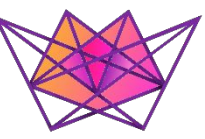
- $A = \sum_i p_i |i\rangle\langle i|$  and  $B = \sum_j q_j |j\rangle\langle j|$

- $tr(A^T B) = tr(\sum_i p_i |i\rangle\langle i| \sum_j q_j |j\rangle\langle j|) =$

$$= \sum_{ij} p_i q_j \langle i|j\rangle tr(|i\rangle\langle j|)$$

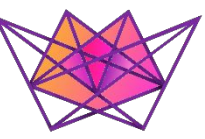
$$= \sum_{ij} p_i q_j \langle i|j\rangle \langle i|j\rangle$$

$$= \sum_{ij} p_i q_j \langle i|j\rangle^2$$



# Distributional Semantics

- ‡ “You shall know the word by the company it keeps” (Firth)
  - Obtain meaning high dimensional vector representations from large corpora automatically
- Compositionality
  - DS can not be applied for entire sentence (lack of frequency)
- Entailment
  - $w$  entails  $v$  if the meaning of a word  $w$  is included in the meaning of a word  $v$   
( $w$  is-a  $v$ ) – *subsumption relation*
    - non symmetric



# Distributional Inclusion Hypothesis

- If  $u$  is *semantically narrower* than  $v$ , then a significant number of salient distributional features of  $u$  are also included in the feature vector of  $v$ :
  - Hypothesis 1: If  $v \Rightarrow w$  then all the characteristic features of  $v$  is expected to appear in  $w$ .
  - Hypothesis 2: If all the characteristic features of  $v$  appear in  $w$ , then  $v \Rightarrow w$ .





# Category Theory

- A monoidal category  $\mathcal{C}$  is a category consisting of the following:
  - a functor  $\otimes: \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$  called the *tensor product*
  - an object  $I \in \mathcal{C}$  called the *unit object*
  - a natural isomorphism whose components  $(A \otimes B) \otimes C \xrightarrow{\alpha_{A,B,C}} A \otimes (B \otimes C)$  are called the *associators*
  - a natural isomorphism whose components  $I \otimes A \xrightarrow{\lambda_A} A$  are called the *left unitors*
  - a natural isomorphism whose components  $A \otimes I \xrightarrow{\rho_A} A$  are called the *right unitors*



# Category Theory

- The objects of the category are thought to be types of systems
- A morphism  $f: A \rightarrow B$  is a process that takes a system of type  $A$  to a system of type  $B$ .
- for  $f: A \rightarrow B$  and  $g: B \rightarrow C$ ,  $g \circ f$  is the composite morphism that takes a system of type  $A$  into a system of type  $C$  by applying the process  $g$  after  $f$ .
- Morphisms of type  $\psi: I \rightarrow A$  are called elements of  $A$ .



# Compact closed categories

- A monoidal category is *compact closed* if for each object  $A$ , there are also left and right dual objects  $A^r$  and  $A^l$ , and morphisms

$$\eta^l: I \rightarrow A \otimes A^l \quad \eta^r: I \rightarrow A \otimes A^r$$

$$\epsilon^l: A^l \otimes A \rightarrow I \quad \epsilon^r: A \otimes A^r \rightarrow I$$

- that satisfies

$$(1_A \otimes \epsilon^l) \circ (\eta^l \otimes 1_A) = 1_A$$

$$(\epsilon^r \otimes 1_A) \circ (1_A \otimes \eta^r) = 1_A$$

$$(\epsilon^l \otimes 1_{A^l}) \circ (1_{A^l} \otimes \eta^l) = 1_{A^l}$$

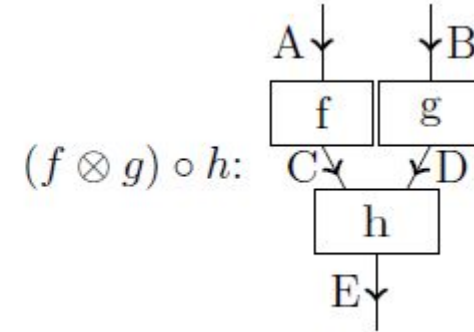
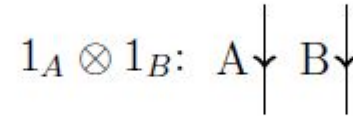
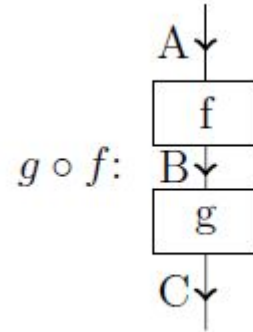
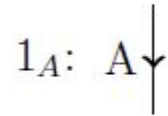
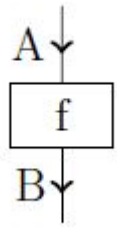
$$(1_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes 1_{A^r}) = 1_{A^r}$$

- The maps of compact categories are used to represent *correlations*, and in categorical quantum mechanics they model maximally entangled states.

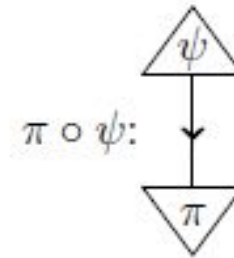
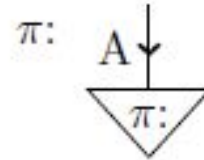
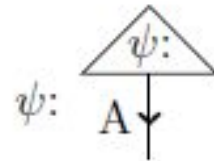


# Graphical calculus

$f: A \rightarrow B$



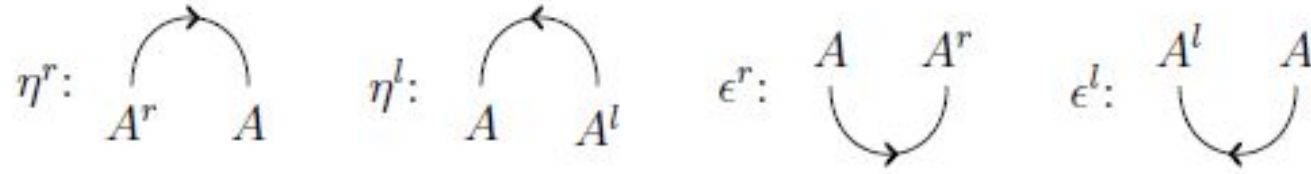
$\psi: I \rightarrow A$



$\pi: A \rightarrow I$



# Graphical calculus



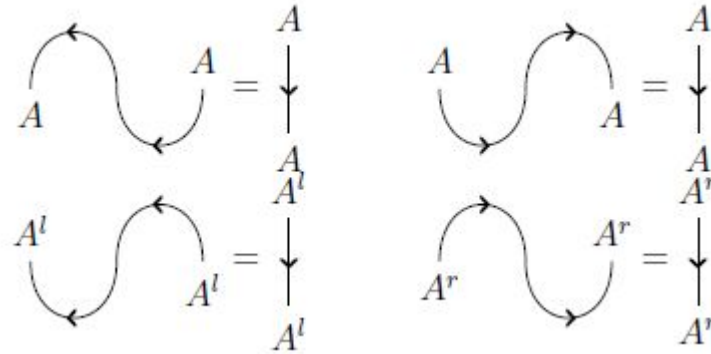
## Snake identities

$$(1_A \otimes \epsilon^l) \circ (\eta^l \otimes 1_A) = 1_A$$

$$(\epsilon^r \otimes 1_A) \circ (1_A \otimes \eta^r) = 1_A$$

$$(\epsilon^l \otimes 1_{A^l}) \circ (1_{A^l} \otimes \eta^l) = 1_{A^l}$$

$$(1_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes 1_{A^r}) = 1_{A^r}$$



## Swing rule





# Compositional Distributional Model

- Pregroup grammars (Lambek)
- A partially ordered monoid  $(P, \leq, \cdot, 1)$  consists of:
  - a set  $P$
  - a monoid multiplication operator  $\cdot : P \times P \rightarrow P$  satisfying the condition

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \text{ for all } a, b, c \in P$$

- and thy monoidal unit  $1 \in P$  where for all  $a \in P$ 
$$a \cdot 1 = a = 1 \cdot a$$
- a partial order  $\leq$  on  $P$



# Pregroup (Lambek, 2001)

- A *pregroup*  $(P, \leq, \cdot, 1, (-)^l, (-)^r)$  is a partially ordered monoid in which each element  $a$  has both a left adjoint  $a^l$  and a right adjoint  $a^r$  such that

$$a^l a \leq 1 \leq a a^l \text{ and } a a^r \leq 1 \leq a^r a$$

- Adjoints have properties:
  - Uniqueness: Adjoints are unique
  - Order reversal: If  $a \leq b$  then  $b^r \leq a^r$  and  $b^l \leq a^l$
  - The unit is self adjoint:  $1^l = 1 = 1^r$
  - Multiplication operation is self adjoint:  $(a \cdot b)^l = b^l \cdot a^l$  and  $(a \cdot b)^r = b^r \cdot a^r$
  - Opposite adjoints annihilate:  $(a^r)^l = a = (a^l)^r$
  - Same adjoints iterate:  $a^{ll} a^l \leq 1 \leq a^{rr} a^r, a^{lll} a^{ll} \leq 1 \leq a^{rrr} a^{rr}, \dots$

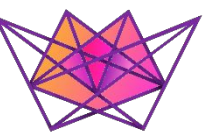


# Pregroup grammar

- $a \rightarrow b$  means  $a \leq b$  ( $a$  reduces to  $b$ )
- “John likes Mary”
- “John” and “Mary” assigned to type  $n$  (noun)
- “likes” is assigned to compound type  $(n^r s n^l)$
- “likes” takes a noun from the left and from the right, and returns a sentence

$$n(n^r s n^l)n \rightarrow 1s n^l n \rightarrow 1s1 \rightarrow s$$





# Basic types

- $n$  : noun
- $j$  : infinitive of the verb

$s$  : declarative statement  
 $g$  : glueing type



# Pregroups as compact closed categories

- $P$  is a concrete instance of a compact closed category

$$\eta^l = [1 \leq p \cdot p^l] \quad \epsilon^l = [p^l \cdot p \leq 1]$$

$$\eta^r = [1 \leq p^r \cdot p] \quad \epsilon^r = [p \cdot p^r \leq 1]$$

- Test snake identities:

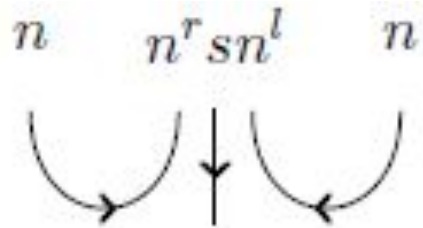
$$\begin{aligned} & (1^p \otimes \epsilon_p^l) \circ (\eta_p^l \otimes 1_p): \\ & p = 1p \leq pp^l p \leq p1 = p \end{aligned}$$

...



# Examples

- “John likes Mary”



- “John does not like Mary”





# FVect - finite dimensional vector space

- FVect – finite dimensional vector spaces over the base field  $\mathbb{R}$  together with linear maps, form a monoidal category
- FVect as a compact closed category.

Given a vector space  $V$  with basis  $\{\vec{e}_i\}_i$

$$\eta_V^l = \eta_V^r : \mathbb{R} \rightarrow V \otimes V$$
$$1 \mapsto \sum_i e_i \otimes e_i$$

$$\epsilon_V^l = \epsilon_V^r : V \otimes V \rightarrow \mathbb{R}$$
$$\sum_{ij} c_{ij} v_i \otimes w_j \mapsto \sum_{ij} c_{ij} \langle v_i | w_j \rangle$$



# $\mathbf{FVect} \times \mathbf{P}$ – categorical representation of meaning space

Objects in  $\mathbf{FVect}$  are of the form  $(V, p)$ , where  $V$  is the vector space representation of the meaning and  $p$  is the pregroup type. There exists a morphism  $(f, \leq) : (V, p) \rightarrow (W, q)$  if there exists a morphism  $f : V \rightarrow W$  in  $\mathbf{FVect}$  and  $p \leq q$  in  $\mathbf{P}$ .

The compact closed structure of  $\mathbf{FVect}$  and  $\mathbf{P}$  lifts componentwise to the product category  $\mathbf{FVect} \times \mathbf{P}$ :

$$\eta^l : (\mathbb{R}, 1) \rightarrow (V \otimes V, p \cdot p^l)$$

$$\eta^r : (\mathbb{R}, 1) \rightarrow (V \otimes V, p^r \cdot p)$$

$$\epsilon^l : (V \otimes V, p^l \cdot p) \rightarrow (\mathbb{R}, 1)$$

$$\epsilon^r : (V \otimes V, p \cdot p^r) \rightarrow (\mathbb{R}, 1)$$

**Definition 6.3.** An object  $(V, p)$  in the product category is called a **meaning space**, where  $V$  is the vector space in which the meanings  $\vec{v} \in V$  of strings of type  $p$  live.



# “From-the-meanings-of-words-to-the-meanings-of-the-sentence” map

- Let  $v_1 v_2 \dots v_n$  be a string of words, each  $v_i$  with a meaning space representation  $\vec{v}_i \in (V_i, p_i)$ . Let  $x \in P$  be a pregroup type such that  $[p_1 p_2 \dots p_n \leq x]$ . Then the meaning vector for the string is:

$$\overrightarrow{v_1 v_2 \dots v_n} \in (W, x) \equiv f(v_1 \otimes v_2 \otimes \dots \otimes v_n),$$

- where  $f$  is defined to be the application of the compact closure maps obtained from the reduction  $[p_1 p_2 \dots p_n \leq x]$  to the composite vector space  $V_1 \otimes V_2 \otimes \dots \otimes V_n$ .



# Example: “John likes Mary”

- It has the pregroup type  $nn^r sn^l n$
- vector representations  $\overrightarrow{John}, \overrightarrow{Mary} \in V$  and  $\overrightarrow{likes} \in V \otimes S \otimes V$
- The morphism in  $\mathbf{FVect} \times \mathbf{P}$  corresponding to the map is of type:

$$(V \otimes (V \otimes S \otimes V) \otimes V, nn^r sn^l n) \rightarrow (s, S)$$

- From the pregroup reduction  $[nn^r sn^l n \rightarrow s]$  we obtain the compact closure maps  $\epsilon^r 1 \epsilon^l$ . In  $\mathbf{FVect}$  this translates into:

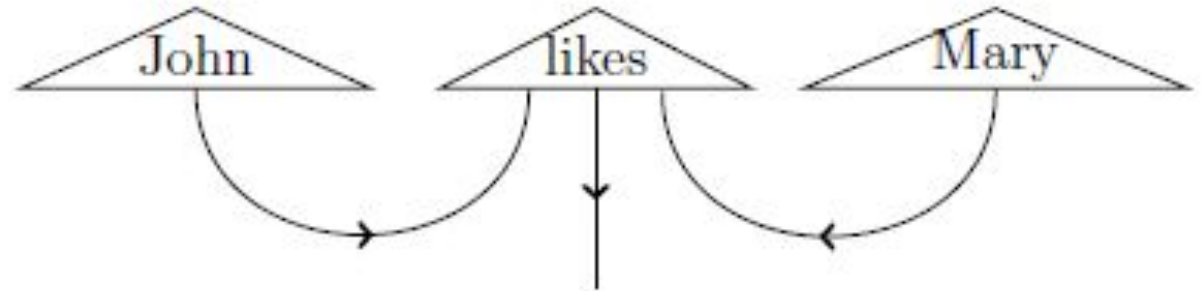
$$\epsilon_V \otimes 1_S \otimes \epsilon_V: V \otimes (V \otimes S \otimes V) \otimes V \rightarrow S$$



# Example: “John likes Mary”

•  $\overrightarrow{John} \otimes \overrightarrow{likes} \otimes \overrightarrow{Mary}$

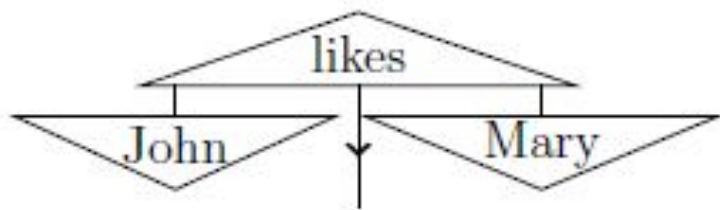
$$\overrightarrow{likes} = \sum_{ijk} c_{ijk} v_i \otimes s_j \otimes v_k$$



$$\overrightarrow{John\ likes\ Mary} = \epsilon_V \otimes 1_S \otimes \epsilon_V (\overrightarrow{John} \otimes \overrightarrow{likes} \otimes \overrightarrow{Mary})$$

$$= \sum_{ijk} \langle John | v_i \rangle s_j \langle v_k | Mary \rangle$$

$$(\langle \epsilon_V^r | \otimes 1_S \otimes \langle \epsilon_V^l |) \circ |\overrightarrow{John} \otimes \overrightarrow{likes} \otimes \overrightarrow{Mary}\rangle$$



$$(\langle \overrightarrow{John} | \otimes 1_S \otimes \langle \overrightarrow{Mary} |) \circ |\overrightarrow{likes}\rangle$$





# Example

$$\begin{aligned} \lceil eat \rceil &= (|\overrightarrow{sloths}\rangle|\overrightarrow{plants}\rangle + |\overrightarrow{lions}\rangle|\overrightarrow{meat}\rangle)(\langle\overrightarrow{sloths}| \langle\overrightarrow{plants}| + \langle\overrightarrow{lions}| \langle\overrightarrow{meat}|) \\ &= (|\overrightarrow{sloths}\rangle|\overrightarrow{plants}\rangle)(\langle\overrightarrow{sloths}| \langle\overrightarrow{plants}|) + \\ &\quad (|\overrightarrow{sloths}\rangle|\overrightarrow{plants}\rangle)(\langle\overrightarrow{lions}| \langle\overrightarrow{meat}|) + \\ &\quad (|\overrightarrow{lions}\rangle|\overrightarrow{meat}\rangle)(\langle\overrightarrow{sloths}| \langle\overrightarrow{plants}|) + \\ &\quad (|\overrightarrow{lions}\rangle|\overrightarrow{meat}\rangle)(\langle\overrightarrow{lions}| \langle\overrightarrow{meat}|) \\ &\sim (|\overrightarrow{sloths}\rangle \langle\overrightarrow{sloths}| \otimes |\overrightarrow{plants}\rangle \langle\overrightarrow{plants}|) + \\ &\quad (|\overrightarrow{sloths}\rangle \langle\overrightarrow{lions}| \otimes |\overrightarrow{plants}\rangle \langle\overrightarrow{meat}|) + \\ &\quad (|\overrightarrow{lions}\rangle \langle\overrightarrow{sloths}| \otimes |\overrightarrow{meat}\rangle \langle\overrightarrow{plants}|) + \\ &\quad (|\overrightarrow{lions}\rangle \langle\overrightarrow{lions}| \otimes |\overrightarrow{meat}\rangle \langle\overrightarrow{meat}|) \end{aligned}$$



# Example

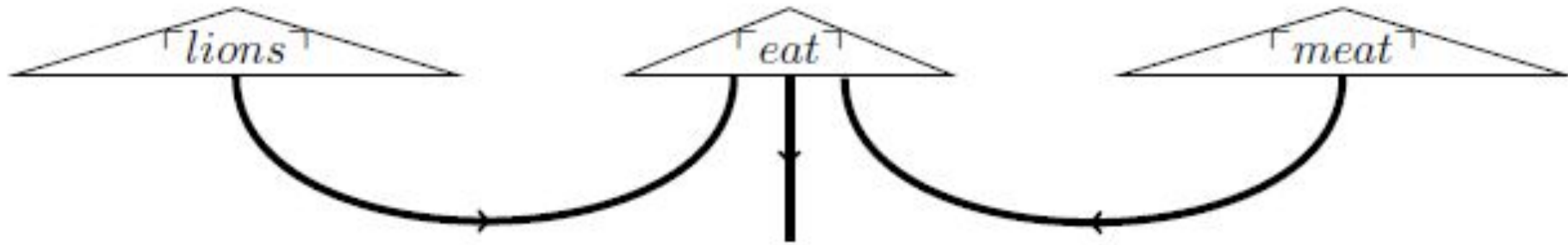
This is the density matrix representation of a pure composite state that relate “sloths” to “plants” and “lions” to “meat”. If we fix the bases  $\{\overrightarrow{lions}, \overrightarrow{sloths}\}$  for  $N_1$ , and  $\{\overrightarrow{meat}, \overrightarrow{plants}\}$  for  $N_2$ ,  $\lceil eat \rceil : N_1 \otimes N_1 \rightarrow N_2 \otimes N_2$  has the following matrix representation:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

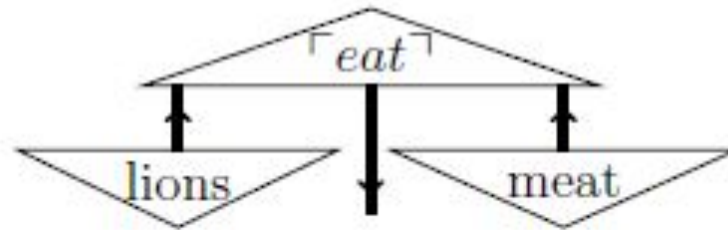
“Lions eat meat” . This is a transitive sentence, so as before, it gets assigned the pregroup type:  $nn^l sn^r n$ . The diagrammatic expression of the pregroup reduction is as follows:



# Example



This reduces to:





# Example

Explicit calculation gives:

$$\begin{aligned} & (\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)(\ulcorner lions \urcorner \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner) \\ &= \langle \overrightarrow{lions} | \overrightarrow{sloths} \rangle^2 \langle \overrightarrow{plants} | \overrightarrow{meat} \rangle^2 + \\ & \quad \langle \overrightarrow{lions} | \overrightarrow{sloths} \rangle \langle \overrightarrow{lions} | \overrightarrow{lions} \rangle \langle \overrightarrow{meat} | \overrightarrow{meat} \rangle \langle \overrightarrow{plants} | \overrightarrow{meat} \rangle + \\ & \quad \langle \overrightarrow{lions} | \overrightarrow{lions} \rangle \langle \overrightarrow{lions} | \overrightarrow{sloths} \rangle \langle \overrightarrow{meat} | \overrightarrow{meat} \rangle \langle \overrightarrow{plants} | \overrightarrow{meat} \rangle + \\ & \quad \langle \overrightarrow{lions} | \overrightarrow{lions} \rangle^2 \langle \overrightarrow{meat} | \overrightarrow{meat} \rangle^2 \\ &= 0 + 0 + 0 + 1 \\ &= 1 \end{aligned}$$



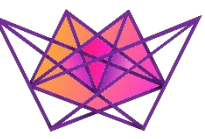
# Example

“Sloths eat meat” . This sentence has a very similar calculation to the one above, and has the result:

$$(\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)(\ulcorner sloths \urcorner \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner) = 0$$

“Mammals eat meat” . This sentence has the same pregroup types as the first sentence, and so has the same reduction map:

$$\begin{aligned} & (\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)(\ulcorner mammals \urcorner \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner) \\ &= (\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)\left(\left(\frac{1}{2}\ulcorner lions \urcorner + \frac{1}{2}\ulcorner sloths \urcorner\right) \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner\right) \\ &= \frac{1}{2}(\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)(\ulcorner lions \urcorner \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner) + \\ &\quad \frac{1}{2}(\epsilon_N^l \otimes 1_S \otimes \epsilon_N^r)(\ulcorner sloths \urcorner \otimes \ulcorner eat \urcorner \otimes \ulcorner meat \urcorner) \\ &= \frac{1}{2} \end{aligned}$$



# Readings

- Esma Balkir. Using Density Matrices in a Compositional Distributional Model of Meaning. // Master thesis. University of Oxford. 2014
- Joachim Lambek. Type grammars as pregroups. Grammars, 4(1):21{39, 2001.