Pregroups and distributional semantics

Antonin Delpeuch

March 25, 2013

Meaning as a distribution?

The distributional hypothesis

You shall know a word by the company it keeps. (Firth)

Words that occur in the same contexts tend to have similar meanings.

Requirements

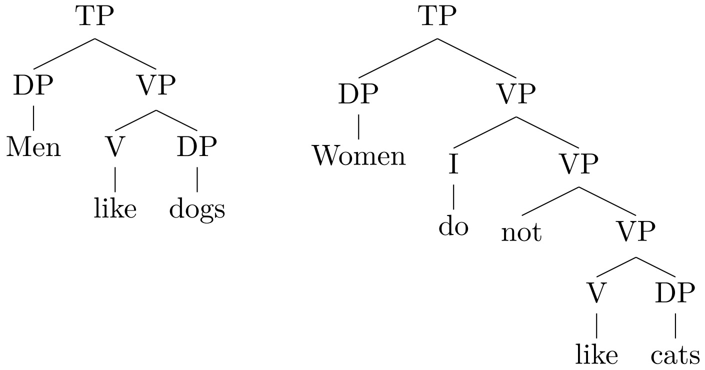
Nice features for a semantic :

- compositional
- quantitative
- enables comparison of meanings

Requirements

Nice features for a semantic :

- compositional
- quantitative
- enables comparison of meanings



Lambek pregroups

- **Lambek calculus** (1958) : $a \cdot b$, a/b, $b\backslash a$.
  Three binary operations : $\cdot$, $/$ and $\backslash$.

- **Lambek calculus** (1958) : $a \cdot b$, a/b, $b\backslash a$.
  Three binary operations : $\cdot$, $/$ and $\backslash$.
- **Lambek pregroup** (1999) : $a \cdot b$, $a \cdot b^l$, $b^r \cdot a$.
  One binary operation, $\cdot$, and two unary operations, giving the left and the right adjoints.

Lambek pregroups

- **Lambek calculus** (1958) : $a \cdot b$, a/b, $b\backslash a$.
  Three binary operations : $\cdot$, $/$ and $\backslash$.
- **Lambek pregroup** (1999) : $a \cdot b$, $a \cdot b^l$, $b^r \cdot a$.
  One binary operation, $\cdot$, and two unary operations, giving the left and the right adjoints.

In both cases, the reduction rules are formalized through a partial order.

Axioms of a pregroup

Definition

A pregroup is a monoid with a partial order $\leq$ and two unary operations ^{l} and ^{r} such that :

- if $a \leq b$ then $ca \leq cb$ and $ac \leq bc$ for all c
- $x^l x \leq 1 \leq x x^l$ and $x x^r \leq 1 \leq x^r x$

	Calculus	**Pregroup**
(a)	$z \to z$	reflexivity of $\leq$
(b)	$(xy)z \to x(yz)$	associativity of $\cdot$
(c)	if $xy \to z$, then $x \to z/y$	$x = x1 \leq xyy^l = zy^l$
(d)	if $x \to z/y$, then $xy \to z$	$xy = zy^l y \leq z1 = z$
(e)	if $x \to y$ and $y \to z$ then $x \to z$	transitivity of $\leq$

Simplified grammar of English : n for nouns, s for sentences, j for infinitive verbs.

men	like	cute	dogs
n	$n^r s n^l$	$n n^l$	n

Examples

Simplified grammar of English : n for nouns, s for sentences, j for infinitive verbs.

men	like	cute	dogs
n	$n^r s n^l$	$n n^l$	n

Examples

Simplified grammar of English : n for nouns, s for sentences, j for infinitive verbs.

men	like	cute	dogs
n	$n^r s n^l$	$n n^l$	n

men	do	not	kill	cute	dogs
n	$n^r s j^l n$	$n^r j j^l n$	$n^r j n^l$	$n n^l$	n

Simplified grammar of English : n for nouns, s for sentences, j for infinitive verbs.

men	like	cute	dogs
n	$n^r s n^l$	$n n^l$	n

men	do	not	kill	cute	dogs
n	$n^r s j^l n$	$n^r j j^l n$	$n^r j n^l$	$n n^l$	n

Simplified grammar of English : n for nouns, s for sentences, j for infinitive verbs.

men	like	cute	dogs
n	$n^r s n^l$	$n n^l$	n

men	do	not	kill	cute	dogs
n	$n^r s j^l n$	$n^r j j^l n$	$n^r j n^l$	$n n^l$	n

Definition

A **monoidal category** is a category with a monoid law $\otimes$ on the objects, such that for each pair of arrows ($f : A \to B$, $g : C \to D$) there exists $f \otimes g : A \otimes C \to B \otimes D$ and such that

$$(f_1 \otimes g_1) \circ (f_2 \otimes g_2) = (f_1 \circ f_2) \otimes (g_1 \circ g_2)$$

In a set theoretic interpretation : $f \otimes g$ is just the parallel evaluation of f and g.

Definition

A **compact closed category** is a monoidal category such that for each object A, there exists :

- A^l and A^r
- $\eta^l : I \to A \otimes A^l$
- $\eta^r : I \to A^r \otimes A$
- $\epsilon^l : A^l \otimes A \to I$
- $\epsilon^r : A \otimes A^r \to I$

and the *yanking equalities* are satisfied.

The yanking equalities

The following equations are called *yanking equalities* :

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

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

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

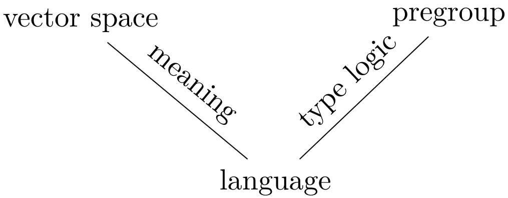
$$(1_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes 1_{A^r}) = 1_{A^r}$$

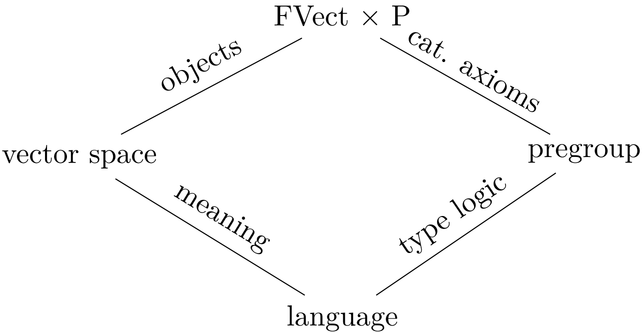
Vector spaces

The category of vector spaces is a compact closed category :

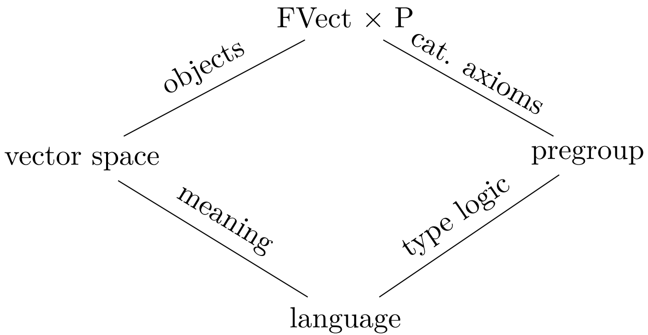
- for the monoid law $\otimes$ (Kronecker product) ;
- with $A^l = A^r = A^*$ (dual) ;
- with $\epsilon^l, \epsilon^r, \eta^l, \eta^r$ correctly defined.

Semantics




FVect × P
objects
cat. axioms
vector space
pregroup
meaning
type logic
language

Semantics



men	like	dogs
$(\overrightarrow{men} \in W, n)$	$(\overrightarrow{like} \in W \otimes S \otimes W, n^r s n^l)$	$(\overrightarrow{dogs} \in W, n)$

Morphism associated to a reduction

If $a \leq b$ holds, then it can be deduced from the axioms :

$$x^l x \leq 1 \leq x x^l$$

$$x x^r \leq 1 \leq x^r x$$

Similarly, in vector spaces category described before, all the morphisms can be expressed using ϵ^l, ϵ^r, η^l and η^r (and 1_A, of course).

When $a \leq b$, we denote by $[a \leq b]$ the morphism of vector spaces associated to this reduction.

Meaning of a sentence

The meaning of the sentence $(v_1 \in W_1, p_1) \otimes \ldots \otimes (v_n \in W_n, p_n)$ is

$$[p_1 \cdot \ldots \cdot p_n \leq s](W_1 \otimes \ldots \otimes W_n)$$

.

$S = \mathbb{R}, \; W = \mathbb{R}^4$
Basis vectors for W : $\overrightarrow{female}$, $\overrightarrow{africa}$, $\overrightarrow{dog}$ and $\overrightarrow{idea}$.

$$\overrightarrow{women} \qquad \overrightarrow{dogs} \qquad \overrightarrow{like}$$

$$\begin{pmatrix} 0.8 & 0.05 & 0.05 & 0.1 \end{pmatrix} \quad \begin{pmatrix} 0.1 & 0 & 0.9 & 0 \end{pmatrix} \quad \begin{pmatrix} 0.1 & 0.2 & 0.6 & 0.1 \\ 0.25 & 0.5 & 0 & 0.25 \\ 0.4 & 0 & 0.6 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{pmatrix}$$

An example (cont.)

Type reduction : $n\ n^r s n^l\ n \leq s$.
Morphism : $[nn^r sn^l n \leq s] = f : W \otimes W \otimes S \otimes W \otimes W \to S$.

$$f = \epsilon^r_W \otimes 1_S \otimes \epsilon^l_W$$

Meaning :

$$(\epsilon^r_W \otimes 1_S \otimes \epsilon^l_W)(\overrightarrow{women} \otimes \overrightarrow{like} \otimes \overrightarrow{dogs})$$

$$= (\epsilon^r_W \otimes 1_S \otimes \epsilon^l_W)(\sum_{i,j} c_{ij} \overrightarrow{women} \otimes \overrightarrow{u}_i \otimes \overrightarrow{v}_j \otimes \overrightarrow{dogs})$$

$$= \sum_{i,j} c_{ij} \langle \overrightarrow{women} | \overrightarrow{u}_i \rangle \langle \overrightarrow{u}_j | \overrightarrow{dogs} \rangle$$

$$= 0.47025$$

Bob Coecke, Edward Grefenstette, and Mehrnoosh Sadrzadeh.
Lambek vs. lambek: Functorial vector space semantics and string diagrams for lambek calculus.
Annals of Pure and Applied Logic, 2013.

J. Lambek.
Pregroup grammars and chomsky's earliest examples.
Journal of Logic, Language, and Information, 17(2):pp. 141–160, 2008.

Benjamin C. Pierce.
Basic Category Theory for Computer Scientists (Foundations of Computing).
The MIT Press, 1 edition, August 1991.