Categories in syntax and semantics

Michael Moortgat

CCSS Science Jam, March 1, 2018

Setting the stage

NL semantics: two views

Distributional model Words that occur in similar contexts have similar meaning.

- words as vectors in a high-dimensional meaning space; bases: context words
- strong quantitative model; applications in technology, cogsci
- bad at capturing grammatical dependencies: **man bites dog** $\simeq$ **dog bites man**

Compositional model The meaning of a complex expression is a function of the meaning of its parts and the rules that put them together.

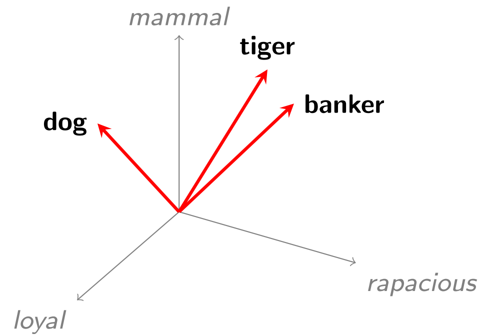
- formal semantics, interpretation as structure-preserving map Source $\xrightarrow{h}$ Target
- strong on derivational semantics: proofs-as-programs
- weak/agnostic on lexical semantics: "the meaning of life is: LIFE"

“You know a word by the company it keeps” — let’s see

```
 no party animosity , I read it with .... satisfaction , as the result of good
sted . In its general principles and .... outlines it was conformable to such
 a citizen who , by a long course of .... actions , regulated by prudence , ju
ll not be without effect . With this .... example before me , with the sense a
bation implied by your suffrage is a .... consolation to me for the past , and
manufactures , education , and other .... objects within each State . In time
 of which the highest interests of a .... and free people are intimately conne
cannot enter on these duties without .... anxiety for the result . From a just
 attention is naturally drawn to the .... causes which have contributed in a p
fair and honorable treaty , and with .... advantage to the original States ; t
th and efficiency as a member of the .... community of nations . Here too expe
conflicts the United States received .... injury from several of the parties .
 Lakes and beyond the sources of the .... rivers which communicate through our
ly interested in preserving it . The .... agricultural interest of the nation
the North engaged in navigation find .... encouragement in being made the favo
d ruin . Let us , then , look to the .... cause , and endeavor to preserve it
```

Rapacious bankers, loyal dogs

Example **dog, tiger, banker** set out in a 3D space spanned by context features *mammal, loyal, rapacious* real life: hundreds, thousands



- geometric view of similarity: orientation, cosine $\angle$
- ☺ vector arithmetic: **king** $-$ **man** $+$ **woman** $\simeq$ **queen**
- ☹ composition: $+$, $\odot$ commutative; same space for nouns, verbs, . . . ?

A marriage of opposites

Compositional ⓄⓄ distributional 2010– . . . Coecke, Sadrzadeh, Clark (Oxbridge); Baroni, Bernardi (Trento) . . .

From types to tensors Key ideas:

- categorial syntax: the category of a word reflects its combinatorial possibilities;
- structural correspondence: types / interpreting semantic spaces

Source: syntax	Target: FVect
basic types np, n, s, . . .	basic spaces N, S, . . .
complex types $A/B, B\backslash A$, . . .	tensor spaces $\mathsf{A}_1 \otimes \cdots \otimes \mathsf{A}_n$
n-ary function types	tensors of rank $n+1$
function application	tensor contraction

slashes=directional implications:

instead of $A \multimap B$, syntax has A/B "A over B" vs $B\backslash A$ "B under A"

The slide ~~you were~~ linguists are afraid of

From linear maps . . . a $(q \times p)$ matrix A encodes a linear map: an instruction that transforms each vector $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p$ in a vector $\mathbf{y} = (y_1, \ldots, y_q) \in \mathbb{R}^q$. We write

$$\mathbf{y} = A\mathbf{x}$$

for the (inner) product of matrix A with vector $\mathbf{x}$: $y_j = \sum_{k=1}^{p} a_{jk} x_k$

. . . to NL meanings tensor contraction generalizes matrix multiplication to spaces of higher rank, encoding multilinear maps:

- rank 2: adjective, type n/n, space N $\otimes$ N: $\textbf{loyal}_{ij}\ \textbf{dog}_j$
- rank 3: transitive verb, type $(np\backslash s)/np$, space N $\otimes$ S $\otimes$ N: $\textbf{man}_i\ \textbf{bites}_{ijk}\ \textbf{dog}_k$
- rank 4: adverb, type $(np\backslash s)\backslash(np\backslash s)$, space N $\otimes$ S $\otimes$ N $\otimes$ S: $\textbf{barks}_{ij}\ \textbf{loudly}_{ijkl}$

repeated indices are summed over, Einstein sommation notation

An NWO OC project

State-of-the-art Existing proposals are problematic in two respects:

Syntax Lambek calculus, pregroup grammars: lacking expressivity;
CCG: spurious ambiguity, incompleteness

FVect degenerate model (Compact Closed Category), collapses $\otimes$, $\oplus$
types $A/(B\backslash C)$, $(A/C) \otimes B$ indistinguishable semantically: $\mathsf{A} \otimes \mathsf{C}^* \otimes \mathsf{B}$

Project objectives Moving beyond the state-of-the-art: overall objective is

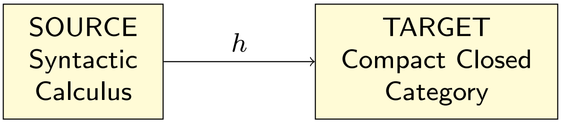
to improve the accuracy of a compositional vector-based interpretation
by employing a more sophisticated grammar logic

To concretize this, there is also a practical goal,

to create a collection of computational tools and datasets
for the compositional distributional study of Dutch.

The plan

Compositionality as structure preserving map/homomorphism relating Syn/Sem:



- Source: categorial grammar, modelling phrases and their composition/concatenation
- Target: meaning composition, FVect/linear maps (abstractly: CCC)

We distinguish two phases in the interpretation process:

- Derivational semantics:

  relates types/proofs of syntactic source to their counterparts in FVect

- Lexical semantics:

  provides concrete vector-based interpretations for the constants (words)

Source: Syntactic Calculus

Joachim ('Jim') Lambek

- ° Leipzig, 5 December 1922, † Montreal, 23 June 2014
- mathematician, McGill University, Montreal
- known for his work on category theory:

  Introduction to Higher Order Categorical Logic (with Phil Scott)
- inventor of 'parsing-as-deduction' method in computational linguistics

Lambek's program for syntax

'Parsing-as-deduction':

- categories, parts of speech $\rightsquigarrow$ logical formulas
- the judgement whether a phrase is well-formed takes the form of a deduction in the grammatical type logic.

THE MATHEMATICS OF SENTENCE STRUCTURE*

JOACHIM LAMBEK, McGill University

The definitions [of the parts of speech] *are very far from having attained the degree of exactitude found in Euclidean geometry.*

—Otto Jespersen, 1924.

1. Introduction. The aim of this paper is to obtain an effective rule (or algorithm) for distinguishing sentences from nonsentences, which works not only for the formal languages of interest to the mathematical logician, but also for natural languages such as English, or at least for fragments of such languages. An attempt to formulate such an algorithm is implicit in the work of Ajdukiewicz.† His method, later elaborated by Bar-Hillel [2], depends on a kind of arithmetization of the so-called *parts of speech*, here called *syntactic types*.‡

Categories / types

Atoms a small set of primitive types for phrases that we judge 'complete'

- s sentence
- np names
- n common nouns

Types The full set of types is given by the grammar below.
We use X as ranging over atomic types; A, B stand for arbitrary types.

$$A, B \quad ::= \quad X \mid A \otimes B \mid A/B \mid B\backslash A$$

- $A \otimes B$: 'A and then B'; sequencing, non-commutative
- A/B: 'A over B'; consumes B on its right to produce A
- $B\backslash A$: 'B under A'; consumes B on its left to produce A

A/B, $B\backslash A$: fractions, numerator A, denominator B

Syntactic Calculus, .58, .61

Lambek's type logic comes in two versions: id/composition of derivability arrow and

- **NL** [L61]: types are assigned to phrases; pure logic of residuation;
- **L** [L58]: types assigned to strings; associative $\otimes$, with unit I

RES	$A \longrightarrow C/B$ iff $A \otimes B \longrightarrow C$ iff $B \longrightarrow A\backslash C$
UNIT	$I \otimes A \longleftrightarrow A \longleftrightarrow A \otimes I$
ASS	$A \otimes (B \otimes C) \longleftrightarrow (A \otimes B) \otimes C$

example:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{np/n \longrightarrow np/n}}{(np/n)\otimes n \longrightarrow np}\,\triangleright^{-1} \qquad \dfrac{\dfrac{\dfrac{}{np\backslash s \longrightarrow np\backslash s}}{np\otimes(np\backslash s)\longrightarrow s}\,\triangleleft^{-1}}{np \longrightarrow s/(np\backslash s)}\,\triangleright}{(np/n)\otimes n \longrightarrow s/(np\backslash s)}\,\circ}{(\underbrace{(np/n)}_{\text{the}} \otimes \underbrace{n}_{\text{temperature}}) \otimes \underbrace{(np\backslash s)}_{\text{rises}} \longrightarrow s}\,\triangleright^{-1}$$

Overgeneration

Syntactic Calculus: beautiful maths — empirical reality: more recalcitrant ...

Unit Vanishing Modifiers, because $I : A/A \longrightarrow$ for arbitrary A,

$$\begin{array}{ll} (OK) & \text{very}_{(n/n)/(n/n)} \otimes \text{smart}_{n/n} \otimes \text{student}_n \to n \\ (Bad) & \text{very}_{(n/n)/(n/n)} \otimes \text{student}_n \to n \end{array}$$

Associativity for arbitrary A, B, one has

$$A/A \longrightarrow B/(A\backslash B))/A \qquad B/B \longrightarrow B/(A\backslash B))/A$$

providing a join type $X := (s/(np\backslash s))/np$ for Unlikely Conjuncts:

$$(\text{Alice\_thinks}_{s/s} \otimes \text{or}_{(X\backslash X)/X} \otimes \text{a\_friend\_of}_{np/np}) \otimes \text{Bob}_{np} \otimes \text{left}_{np\backslash s} \to s$$

Pregroup grammars generally used for compositional distributional modelling

- PG: bi-Lambek + Excluded Middle + collapse $\otimes = \oplus$, $O = I$
- hence: inherit the UNIT, ASS problems from **L**

$\mathbf{NL}_\diamond$: logic and control

The lesson from Linear Logic keep non-associative, non-commutative base logic **NL**; introduce possibilities for reordering, restructuring in a controlled form.

Modalities Unary pair of residuated operations:

$$\mathsf{RES}_\diamond \qquad \Diamond A \longrightarrow B \quad \text{iff} \quad A \longrightarrow \Box B$$

compositions: $\Diamond\Box A \rightarrow A \rightarrow \Box\Diamond A$; $\Diamond\cdot$, $\Box\cdot$ monotonic.

Licensing reordering, restructuring e.g. postulates allowing leftward extraction:

$$\Diamond A \otimes (B \otimes C) \longrightarrow (\Diamond A \otimes B) \otimes C \qquad \Diamond A \otimes (B \otimes C) \longrightarrow B \otimes (\Diamond A \otimes C)$$

Structural control Kurtonina & MM '97: embedding theorems **NL** $\rightleftarrows$ **L**

Running example: relative clauses NL/E

Consider Stieg Larsson's bestseller "Män som hatar kvinnor"

	The girl with the dragon tattoo	(non-compositional)
Dutch:	mannen$_n$ die$_{??}$ vrouwen$_{np}$ haten$_{np\backslash(np\backslash s)}$	OV order!
English:	men$_n$ who$_{??}$ hate$_{(np\backslash s)/np}$ women$_{np}$	subject rel
	men$_n$ who(m)$_{??}$ women$_{np}$ hate$_{(np\backslash s)/np}$	object rel

NL challenges:

- single type assignment for RelPro, accounting for subj/obj ambiguity

  die :: $(n\backslash n)/(\Diamond\Box np\backslash s)$ versus

  who :: $(n\backslash n)/(np\backslash s)$, whom :: $(n\backslash n)/(s/\Diamond\Box np)$
- capture intersective interpretation of Noun+Rel in distributional model
- capture the disambiguating effect of particular lexical choices

  "doctor who cured the patient" vs "whom the patient consulted" (in NL)

Dutch derivational ambiguity, prooftheoretically

$$\dfrac{\vdots}{\dfrac{\dfrac{}{n\backslash n \longrightarrow n\backslash n} \qquad \dfrac{\vdots}{np \otimes (np\backslash(np\backslash s)) \longrightarrow \Diamond\Box np\backslash s}\ \text{vrouwen haten}}{}}$$

$$\dfrac{\vdots}{n \otimes (((n\backslash n)/(\Diamond\Box np\backslash s)) \otimes (np \otimes (np\backslash(np\backslash s)))) \longrightarrow n}$$

after a common start the derivations diverge (left: subject rel, righ: object rel):

$$\dfrac{\dfrac{\dfrac{}{np \longrightarrow np} \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{np \longrightarrow np}}{\Box np \longrightarrow \Box np}\,\Box}{\Diamond\Box np \longrightarrow np}\,\triangledown^{-1} \qquad \dfrac{}{s \longrightarrow s}}{np\backslash s \longrightarrow \Diamond\Box np\backslash s}\,\backslash}{np\backslash(np\backslash s) \longrightarrow np\backslash(\Diamond\Box np\backslash s)}\,\backslash}{np \otimes (np\backslash(np\backslash s)) \longrightarrow \Diamond\Box np\backslash s}\,\triangleleft^{-1}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{np \longrightarrow np}}{\Box np \longrightarrow \Box np}\,\Box}{\Diamond\Box np \longrightarrow np}\,\triangledown^{-1} \qquad \dfrac{}{np\backslash s \longrightarrow np\backslash s}}{np\backslash(np\backslash s) \longrightarrow \Diamond\Box np\backslash(np\backslash s)}\,\backslash}{\Diamond\Box np \otimes (np\backslash(np\backslash s)) \longrightarrow np\backslash s}\,\triangleleft^{-1}}{np \otimes (\Diamond\Box np \otimes (np\backslash(np\backslash s))) \longrightarrow s}\,\triangleleft^{-1}}{\Diamond\Box np \otimes (np \otimes (np\backslash(np\backslash s))) \longrightarrow s}\,\widehat{\sigma}^{l}_{\diamond}}{np \otimes (np\backslash(np\backslash s)) \longrightarrow \Diamond\Box np\backslash s}\,\triangleleft$$

die :: $(n\backslash n)/(\Diamond\Box np\backslash s)$

Human language processing $\neq$ proof theory

The derivation above is not meant as a realistic model of human language processing.

- incrementality
  - NLP proceeds in time, producing intermediary results along the way

    garden path effects: 'the horse raced past the barn fell'
  - logical derivations can only start when the complete goal phrase is given
- top-down versus bottom-up
  - logic: decomposes a complex structure up to axioms; goal driven; td

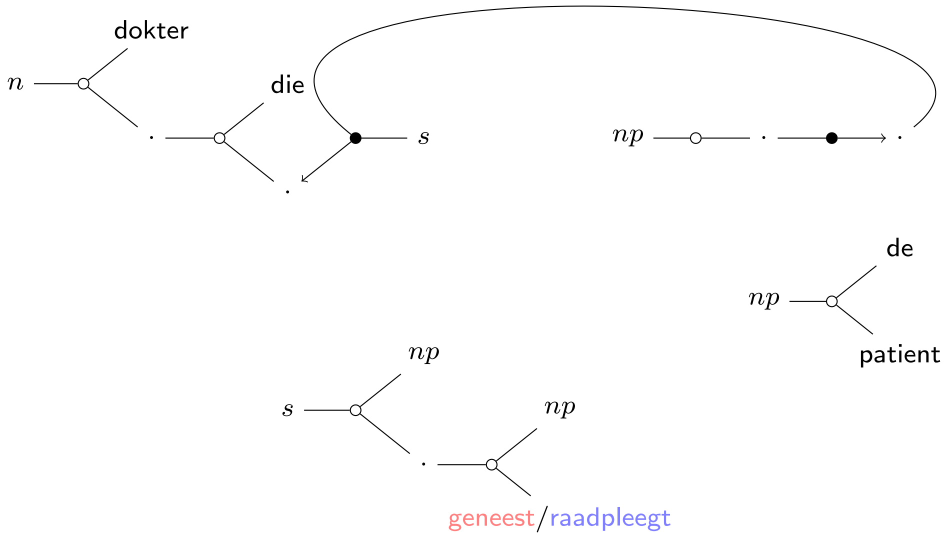
    residuation shifting: artefact, different 'views' of same structure
  - NLP: puts together a structure out of lexical parts; data driven; bu

Proof nets a graphical calculus that doesn't suffer from the above defects.

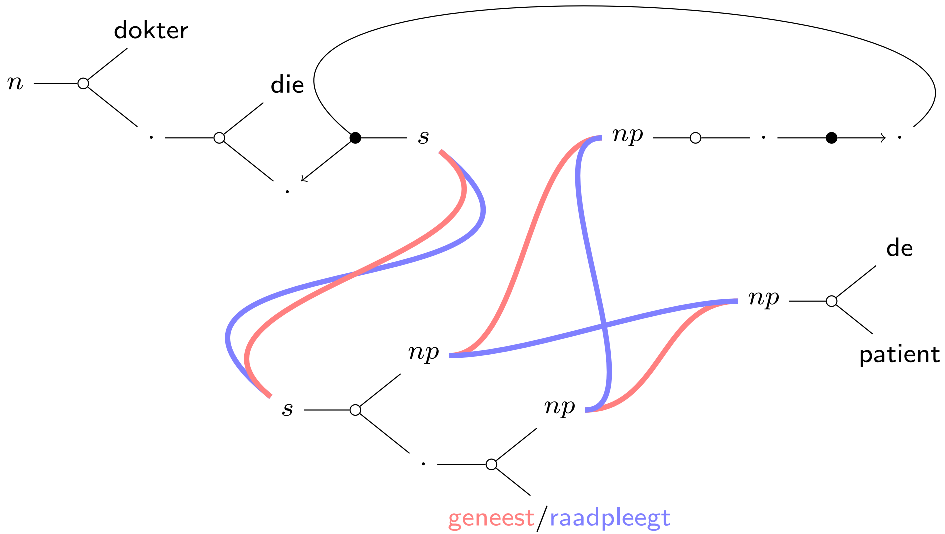
R. Moot, Proof nets for display logic; M. Moortgat & R. Moot, Proof nets for Lambek-Grishin calculus. arXiv

Proofs-as-pictures

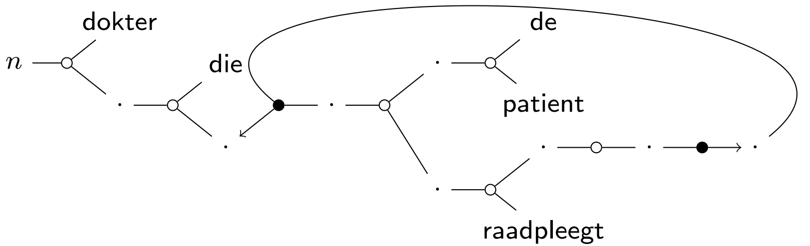
“geneest”: cures, “raadpleegt”: consults



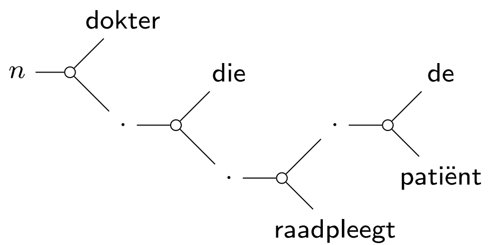
Proofs-as-pictures



From proof structure to proof net



The candidate p.s. qualifies as a net after structural rewriting, ○● logical contractions



contractions $\sim$ monotonicity; structural rewriting: $\Diamond A \otimes (B \otimes C) \rightarrow B \otimes (\Diamond A \otimes C)$

Target: FVect

Compact closed category, FVect

Compact closed category (CCC) is monoidal, i.e. it has an associative $\otimes$ with unit I; and for every object there is a left and a right adjoint satisfying

$$A^l \otimes A \xrightarrow{\epsilon^l} I \xrightarrow{\eta^l} A \otimes A^l \qquad A \otimes A^r \xrightarrow{\epsilon^r} I \xrightarrow{\eta^r} A^r \otimes A$$

In a *symmetric* CCC, the tensor moreover is commutative, and we can write A^* for the collapsed left and right adjoints.

FVect, linear maps concrete instance of sCCC

- unit I: the field $\mathbb{R}$; bases: fixed, so $V^* \cong V$ one can ignore $\cdot^*$
- ϵ map: inner products; η map: identity tensor (with $\lambda = 1$) or multiples

$$\epsilon_V : V \otimes V \to \mathbb{R} \quad \text{given by} \quad \sum_{ij} v_{ij}(\vec{e}_i \otimes \vec{e}_j) \mapsto \sum_i v_{ii}$$

$$\eta_V : \mathbb{R} \to V \otimes V \quad \text{given by} \quad \lambda \mapsto \sum_i \lambda(\vec{e}_i \otimes \vec{e}_i)$$

From syntax to semantics

$\mathbf{NL}_\diamond \xrightarrow{\lceil\cdot\rceil}$ sCCC We show how types/proofs of the syntactic source $\mathbf{NL}_\diamond$ are mapped to the corresponding types/proofs of the interpreting sCCC.

Types $\lceil\cdot\rceil$ assigns a vector space to the type atoms of $\mathbf{NL}_\diamond$: $\lceil np \rceil = \lceil n \rceil =$ N, $\lceil s \rceil =$ S; for complex types we set

- the syntactic control operators are semantically vacuous:

$$\lceil \Diamond A \rceil = \lceil \Box A \rceil = \lceil A \rceil$$

- syntactic concatenation and its residuals $\leadsto$ tensor product in sCCC/FVect:

$$\lceil A \otimes B \rceil = \lceil A \rceil \otimes \lceil B \rceil$$

$$\lceil A/B \rceil = \lceil A \rceil \otimes \lceil B \rceil^* \quad \lceil A \backslash B \rceil = \lceil A \rceil^* \otimes \lceil B \rceil$$

Interpretation: proofs

Syntactic derivations $f : A \longrightarrow B$ in $\mathsf{NL}_\diamond$ are interpreted as linear maps.

Identity, composition $\lceil 1_A \rceil = 1_{\lceil A \rceil}, \quad \lceil g \circ f \rceil = \lceil g \rceil \circ \lceil f \rceil$

Residuation

$$\left\lceil \frac{f : A \otimes B \longrightarrow C}{\triangleright f : A \longrightarrow C/B} \right\rceil$$

$$\lceil \triangleright f \rceil = \lceil A \rceil \xrightarrow{1_{\lceil A \rceil} \otimes \eta_{\lceil B \rceil}} \lceil A \rceil \otimes \lceil B \rceil \otimes \lceil B \rceil^* \xrightarrow{\lceil f \rceil \otimes 1_{\lceil B \rceil^*}} \lceil C \rceil \otimes \lceil B \rceil^*$$

$$\left\lceil \frac{g : A \longrightarrow C/B}{\triangleright^{-1} g : A \otimes B \longrightarrow C} \right\rceil$$

$$\lceil \triangleright^{-1} g \rceil = \lceil A \rceil \otimes \lceil B \rceil \xrightarrow{\lceil g \rceil \otimes 1_{\lceil B \rceil}} \lceil C \rceil \otimes \lceil B \rceil^* \otimes \lceil B \rceil \xrightarrow{1_{\lceil C \rceil} \otimes \epsilon_{\lceil B \rceil}} \lceil C \rceil$$

similarly for $\triangleleft, \triangleleft^{-1}$

Interpreting proofs

Monotonicity The case of parallel composition is immediate: $\lceil f \otimes g \rceil = \lceil f \rceil \otimes \lceil g \rceil$.

For / we have

$$\left\lceil \frac{f : A \longrightarrow B \quad g : C \longrightarrow D}{f/g : A/D \longrightarrow B/C} \right\rceil$$

where $\lceil f/g \rceil =$

$$\begin{array}{c}
\lceil A \rceil \otimes \lceil D \rceil^* \\
\Big\downarrow \lceil f \rceil \otimes \eta_{\lceil C \rceil} \otimes 1_{\lceil D \rceil^*} \\
\lceil B \rceil \otimes \lceil C \rceil^* \otimes \lceil C \rceil \otimes \lceil D \rceil^* \\
\Big\downarrow 1_{\lceil B \rceil \otimes \lceil C \rceil^*} \otimes \lceil g \rceil \otimes 1_{\lceil D \rceil^*} \\
\lceil B \rceil \otimes \lceil C \rceil^* \otimes \lceil D \rceil \otimes \lceil D \rceil^* \\
\Big\downarrow 1_{\lceil B \rceil \otimes \lceil C \rceil^*} \otimes \epsilon_{\lceil D \rceil} \\
\lceil B \rceil \otimes \lceil C \rceil^*
\end{array}$$

similarly for $g \backslash f$

Interpretation: proof nets

The interpretation $\lceil \cdot \rceil$ follows the inference steps of the deductive system for syntax.

The proof net method allows for a more direct interpretation:

- residuation steps are factored out: different 'views' on same structure
- proof net is entirely determined by
  - axiom linkings
  - $\circ\bullet$ contractions, possibly mediated by structural rewriting

We implement this direct-style interpretation by means of a generalized Kronecker delta function for the axioms linkings of a net.

Axiom links, Kronecker delta

Axioms $\lceil 1_A \rceil = 1_{\lceil A \rceil}$, in FVect: the identity map on vector space $\lceil A \rceil$.

This map is represented by square matrix $\mathbb{1}$ of $dim(\lceil A \rceil)$ with entries given by the Kronecker delta function:

$$\mathbb{1}_{i,j} = \delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Proofs Derivations with n axioms are interpreted by means of an identity tensor

$$T = (A_1 \otimes \cdots \otimes A_n) \otimes (A_1 \otimes \cdots \otimes A_n)$$

with entries given by a generalized Kronecker delta $\delta^{\bar\imath}_{\bar\jmath}$.

$\bar\imath, \bar\jmath$: sequences of indices $\bar\imath_1, \ldots \bar\imath_n$ (inputs), $\bar\jmath_1, \ldots \bar\jmath_n$ (outputs) with the $\bar\imath_k, \bar\jmath_k$ ranging over $dim(A_k)$.

$$\delta^{\bar\imath}_{\bar\jmath} = \begin{cases} 1 & \text{if } \bar\imath_k = \bar\jmath_k \text{ for } 1 \leq k \leq n \\ 0 & \text{otherwise} \end{cases}$$

- residuation rules: change dom/range signature, leave δ unchanged;
- monotonicity rules merge premise delta's: conclusion $\delta^{\bar\imath,\bar\imath'}_{\bar\jmath,\bar\jmath'}$ from premises $\delta^{\bar\imath}_{\bar\jmath}$, $\delta^{\bar\imath'}_{\bar\jmath'}$

Two axioms, three proofs

$$\dfrac{\dfrac{\overline{np \longrightarrow np} \quad \overline{s \longrightarrow s}}{\dfrac{np\backslash s \longrightarrow np\backslash s}{\dfrac{np \otimes (np\backslash s) \longrightarrow s}{np \longrightarrow s/(np\backslash s)} \triangleright} \triangleleft^{-1}} (\backslash)}{}$$

From the axioms, we obtain a matrix $M \in \mathsf{N} \otimes \mathsf{S}$, a vector $V \in \mathsf{S}$, a rank-3 tensor $C \in \mathsf{S} \otimes \mathsf{N} \otimes \mathsf{S}$, depending on how the resources are distributed over input/output.

$$\text{dream}^{np\backslash s} \longrightarrow np\backslash s \quad \xrightarrow{\lceil\cdot\rceil} \quad \textbf{dream}^{\mathsf{N}\otimes\mathsf{S}}_{i,j} \xmapsto{\delta^{k,j}_{i,l}} M_{k,l} \in \mathsf{N} \otimes \mathsf{S}$$

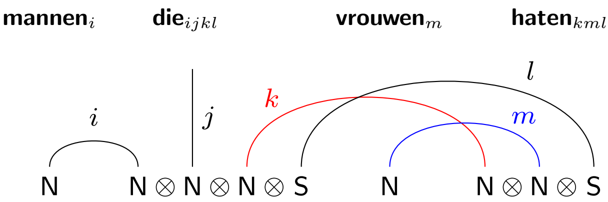
$$\text{poets}^{np} \otimes \text{dream}^{np\backslash s} \longrightarrow s \quad \xrightarrow{\lceil\cdot\rceil} \quad \textbf{poets}^{\mathsf{N}}_{i} \otimes \textbf{dream}^{\mathsf{N}\otimes\mathsf{S}}_{j,k} \xmapsto{\delta^{i,k}_{j,l}} V_l \in \mathsf{S}$$

$$\text{poets}^{np} \longrightarrow s/(np\backslash s) \quad \xrightarrow{\lceil\cdot\rceil} \quad \textbf{poets}^{\mathsf{N}}_{i} \xmapsto{\delta^{i,l}_{k,j}} C_{j,k,l} \in \mathsf{S} \otimes \mathsf{N} \otimes \mathsf{S}$$

Larsson again

Repeated indices: tensor contraction.

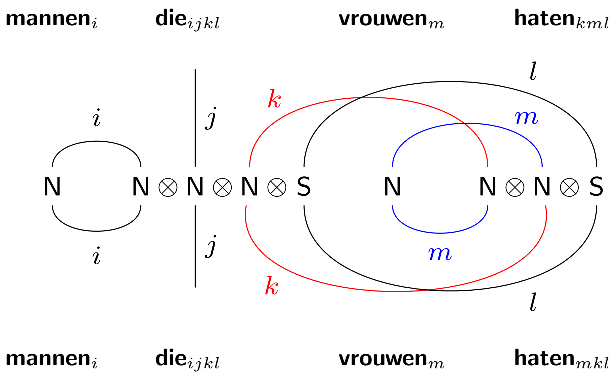
Top: object relative (women hating men)



Larsson again

Repeated indices: tensor contraction.

Top: object relative (women hating men); bottom: subject rel (men hating women).



Word meanings

The derivational semantics is parametric w.r.t. the meaning of the lexical constants.

- non-logical constants $\sim$ open class items

  high-dimensional word embeddings harvested from corpus data

- logical constants $\sim$ closed class items (function words)

  predefined maps, independent of distribution

Relative pronoun meaning

Wanted: rank-4 tensor **die** $\in \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{S}$.

In FVect, intersective effect via $\odot$: preserves context features with non-0 values.

- elementwise multiplication expressible in terms of tensor contraction:

  $v \odot w = v^\top \times C \times w$, for $C \in \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{N}$ with ones on diagonal, 0's elsewhere

- relative clause body: vector w via rank reduction $\mathsf{N} \otimes \mathsf{S} \mapsto \mathsf{N}$

  $w = M \times \mathbf{1}$, summing rows, with $M \in \mathsf{N} \otimes \mathsf{S}$ and $\mathbf{1}$ the all-ones vector in S

Does this make sense? Let the data speak . . .

□

References

[1] Marco Baroni, Raffaela Bernardi, and Roberto Zamparelli. Frege in Space: a Program for Compositional Distributional Semantics. *Linguistic Issues in Language Technology*, 9:241–346, 2014.

[2] 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*, 164(11):1079–1100, 2013.

[3] Michael Moortgat and Gijs Wijnholds. Lexical and Derivational Meaning in Vector-Based Models of Relativisation. CoRR abs/1711.11513, 2017.

[4] Richard Moot and Christian Retoré. Proof Nets for the Multimodal Lambek Calculus: From Theory to a Wide-Coverage Categorial Parser. Chapter 7, The Logic of Categorial Grammar. LNCS 6850, 2012.

[5] Mehrnoosh Sadrzadeh, Stephen Clark, and Bob Coecke. The Frobenius Anatomy of Word Meanings I: Subject and Object Relative Pronouns. *Journal of Logic and Computation*, pages 1293–1317, 2013.