

Syntactic Control for Compositional Vector Space Models

Michael Moortgat¹ Gijs Wijnholds²

Utrecht Institute of Linguistics
M.J.Moortgat@uu.nl

ILLC Amsterdam
gijswijnholds@gmail.com

TACL 2015
June 21, 2015

Standard Theory

Interpretation

Problems

Source Extensions

Extending the Interpretation

Conclusion

Lambek's program

- ▶ grammar: universal type logic + language-specific lexicon
- ▶ type logic: 'deductive system as category'
- ▶ base system: biclosed tensor category (non-ass, non-comm!)
- ▶ add-ons: postulates attributing extra properties to \otimes :
 - ▶ unit
 - ▶ associativity
 - ▶ commutativity

The Lambek Hierarchy

$$\frac{}{1_A : A \rightarrow A} \text{ Id} \quad \frac{f : A \rightarrow B \quad g : B \rightarrow C}{g \circ f : A \rightarrow C} \text{ T}$$

► NL:

$$\frac{f : A \otimes B \rightarrow C}{\triangleright f : A \rightarrow C/B} \text{ Res} \quad \frac{f : A \otimes B \rightarrow C}{\triangleleft f : B \rightarrow A \setminus C} \text{ Res}$$

The Lambek Hierarchy

$$\frac{}{1_A : A \rightarrow A} \text{Id} \quad \frac{f : A \rightarrow B \quad g : B \rightarrow C}{g \circ f : A \rightarrow C} \text{T}$$

► **NL:**

$$\frac{f : A \otimes B \rightarrow C}{\triangleright f : A \rightarrow C/B} \text{Res} \quad \frac{f : A \otimes B \rightarrow C}{\triangleleft f : B \rightarrow A \setminus C} \text{Res}$$

► **L: NL +**

$$a_{A,B,C} : (A \otimes B) \otimes C \longleftrightarrow A \otimes (B \otimes C) : a_{A,B,C}^{-1}$$

The Lambek Hierarchy

$$\frac{}{1_A : A \rightarrow A} \text{Id} \quad \frac{f : A \rightarrow B \quad g : B \rightarrow C}{g \circ f : A \rightarrow C} \text{T}$$

► **NL:**

$$\frac{f : A \otimes B \rightarrow C}{\triangleright f : A \rightarrow C/B} \text{Res} \quad \frac{f : A \otimes B \rightarrow C}{\triangleleft f : B \rightarrow A \setminus C} \text{Res}$$

► **L: NL +**

$$a_{A,B,C} : (A \otimes B) \otimes C \longleftrightarrow A \otimes (B \otimes C) : a_{A,B,C}^{-1}$$

► **LP: L +**

$$c_{A,B} : A \otimes B \rightarrow B \otimes A$$

The Lambek Hierarchy

$$\frac{}{1_A : A \rightarrow A} \text{Id} \quad \frac{f : A \rightarrow B \quad g : B \rightarrow C}{g \circ f : A \rightarrow C} \text{T}$$

► **NL:**

$$\frac{f : A \otimes B \rightarrow C}{\triangleright f : A \rightarrow C/B} \text{Res} \quad \frac{f : A \otimes B \rightarrow C}{\triangleleft f : B \rightarrow A \setminus C} \text{Res}$$

► **L:** **NL** +

$$a_{A,B,C} : (A \otimes B) \otimes C \longleftrightarrow A \otimes (B \otimes C) : a_{A,B,C}^{-1}$$

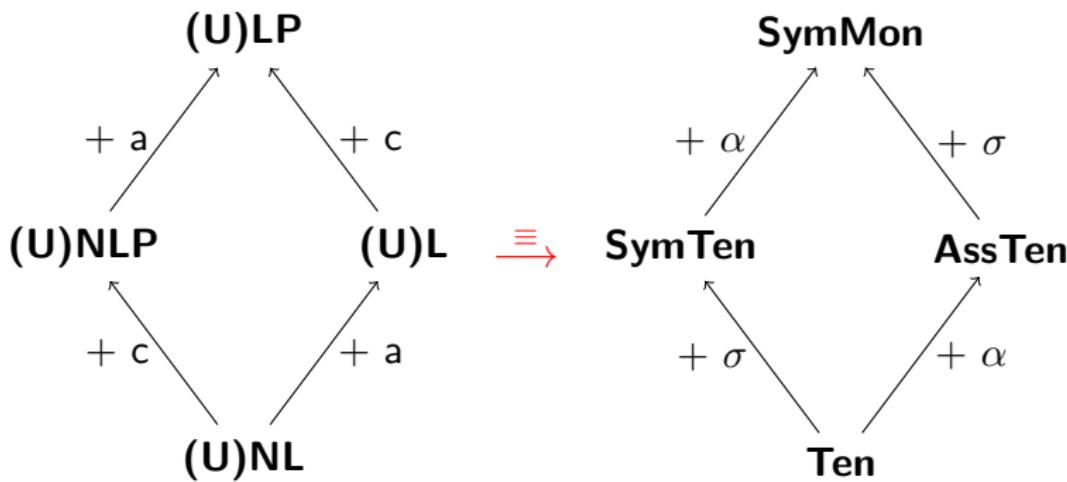
► **LP:** **L** +

$$c_{A,B} : A \otimes B \rightarrow B \otimes A$$

► **U(N)L(P):** add unit I

$$l_A : I \otimes A \longleftrightarrow A : l_A^{-1} \quad r_A : A \otimes I \longleftrightarrow A : r_A^{-1}$$

The Lambek Hierarchy: logics, categories



- ▶ **Ten** : bi-closed tensor category
- ▶ **Mon** : bi-closed monoidal category
- ▶ **Sym, Ass** respectively symmetry and associativity.

Lambek Grammars

- ▶ Obtained by adding a lexicon to the logic:

john : np

mary : np

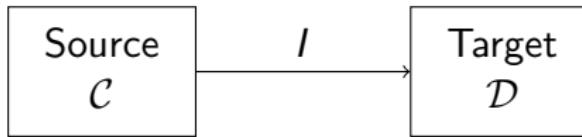
loves : $(np \setminus s)/np$

- ▶ a sequence of words $w_1 \cdots w_n$ is a well-formed expression of type B just in case the lexicon assigns types A_i to w_i and there is a proof $A^\otimes \rightarrow B$, where A^\otimes is some product of the A_i

$$\frac{\frac{\frac{1_{np} : np \rightarrow np \quad Id \quad \frac{1_{np} : np \rightarrow np \quad Id \quad 1_s : s \rightarrow s \quad Id}{1_{np} \setminus 1_s : np \setminus s \rightarrow np \setminus s} M\backslash}{(1_{np} \setminus 1_s) / 1_{np} : (np \setminus s) / np \rightarrow (np \setminus s) / np} M/}{\triangleright^{-1}((1_{np} \setminus 1_s) / 1_{np}) : (np \setminus s) / np \otimes np \rightarrow np \setminus s} r}{\triangleleft^{-1} \triangleright^{-1} ((1_{np} \setminus 1_s) / 1_{np}) : np \otimes (((np \setminus s) / np) \otimes np) \rightarrow s} r$$

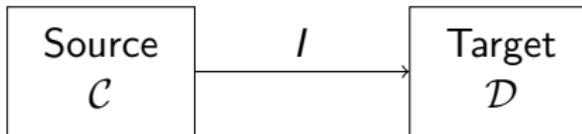
Compositional interpretation

- ▶ Montague: homomorphism relating source/target algebras
- ▶ Lambek: functorial transition from source to target category



Compositional interpretation

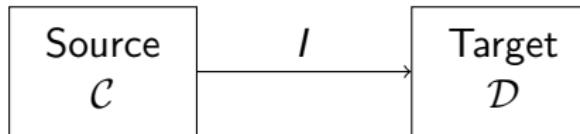
- ▶ Montague: homomorphism relating source/target algebras
- ▶ Lambek: functorial transition from source to target category



Target in standard Montague Grammar: set-theoretic semantics, built from universe E , truth values $\{0, 1\}$.

Compositional interpretation

- ▶ Montague: homomorphism relating source/target algebras
- ▶ Lambek: functorial transition from source to target category

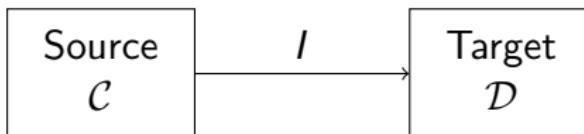


Target in standard Montague Grammar: set-theoretic semantics, built from universe E , truth values $\{0, 1\}$.

- ▶ E.g. $I(\text{man}) = \{e \in E \mid e \text{ is a man}\}$

Compositional interpretation

- ▶ Montague: homomorphism relating source/target algebras
- ▶ Lambek: functorial transition from source to target category



Target in standard Montague Grammar: set-theoretic semantics, built from universe E , truth values $\{0, 1\}$.

- ▶ E.g. $I(\text{man}) = \{e \in E \mid e \text{ is a man}\}$
- ▶ We now have $I(\triangleleft^{-1} \triangleright^{-1} ((1_{np} \setminus 1_s) / 1_{np})) = \lambda a \otimes F \otimes b.F \ b \ a$
and $I(\text{john loves mary}) = \mathbf{j} \otimes \mathbf{love} \otimes \mathbf{m}$, so we get:

$$\llbracket \text{john loves mary} \rrbracket = \lambda a \otimes F \otimes b.F \ b \ a (\mathbf{j} \otimes \mathbf{love} \otimes \mathbf{m}) = \mathbf{love}(\mathbf{m})(\mathbf{j})$$

which will be 1 precisely when **j** and **m** are in the love relation.

Vector Space Models

- ▶ In the Montagovian world, semantics is limited to binary truth $\{\}$.
- ▶ In general, the target structure can be any bi-closed tensor/monoidal category.
- ▶ Adopting the “meaning as use” view, we would like to use finite-dimensional vector spaces as a target structure.
- ▶ The category **FVect** is in fact a bi-closed monoidal category.
- ▶ The meaning of words is induced from a corpus by means of co-occurrence counts.
- ▶ One can now reason about meaning *similarity*.

Cf. [Coecke et al 2010, 2013]

Illustrating Vector Space Models

We use two basic spaces N, S spanned by entities and $\vec{1}$, resp:

- We enhance the running example with vectors:

$$\begin{array}{lll} \text{john} & : & np \\ \text{mary} & : & np \\ \text{loves} & : & (np \setminus s)/np \end{array} \quad : \quad \begin{array}{l} \vec{n_{13}} \in N \\ \vec{n_7} \in N \\ .8(\vec{n_{13}} \otimes \vec{1} \otimes \vec{n_7}) + .2(\vec{n_7} \otimes \vec{1} \otimes \vec{n_{13}}) \\ + \sum_{ij} w_{ij}(\vec{n_i} \otimes \vec{1} \otimes \vec{n_j}) \in N \otimes S \otimes N \end{array}$$

- Application is now computed using inner products:

$$\begin{aligned} & \vec{n_{13}} \otimes .8(\vec{n_{13}} \otimes \vec{1} \otimes \vec{n_7}) + .2(\vec{n_7} \otimes \vec{1} \otimes \vec{n_{13}}) + \sum_{ij} w_{ij}(\vec{n_i} \otimes \vec{1} \otimes \vec{n_j}) \otimes \vec{n_7} \\ \mapsto & .8\langle \vec{n_{13}} | \vec{n_{13}} \rangle \langle \vec{n_7} | \vec{n_7} \rangle \vec{1} + .2\langle \vec{n_{13}} | \vec{n_7} \rangle \langle \vec{n_{13}} | \vec{n_7} \rangle \vec{1} + \sum_{ij} w_{ij} \langle \vec{n_{13}} | \vec{n_i} \rangle \langle \vec{n_j} | \vec{n_7} \rangle \vec{1} \\ = & .8\vec{1} + 0\vec{1} + 0\vec{1} \\ = & .8\vec{1} \end{aligned}$$

Illustrating Vector Space Models

We use two basic spaces N, S spanned by entities and $\vec{1}$, resp:

- We enhance the running example with vectors:

$$\begin{array}{lll} \text{john} & : & np \\ \text{mary} & : & np \\ \text{loves} & : & (np \setminus s)/np \end{array} \quad : \quad \begin{array}{l} \vec{n_{13}} \in N \\ \vec{n_7} \in N \\ .8(\vec{n_{13}} \otimes \vec{1} \otimes \vec{n_7}) + .2(\vec{n_7} \otimes \vec{1} \otimes \vec{n_{13}}) \\ + \sum_{ij} w_{ij}(\vec{n_i} \otimes \vec{1} \otimes \vec{n_j}) \in N \otimes S \otimes N \end{array}$$

- Application is now computed using inner products:

$$\begin{aligned} & \vec{n_{13}} \otimes .8(\vec{n_{13}} \otimes \vec{1} \otimes \vec{n_7}) + .2(\vec{n_7} \otimes \vec{1} \otimes \vec{n_{13}}) + \sum_{ij} w_{ij}(\vec{n_i} \otimes \vec{1} \otimes \vec{n_j}) \otimes \vec{n_7} \\ \mapsto & .8\langle \vec{n_{13}} | \vec{n_{13}} \rangle \langle \vec{n_7} | \vec{n_7} \rangle \vec{1} + .2\langle \vec{n_{13}} | \vec{n_7} \rangle \langle \vec{n_{13}} | \vec{n_7} \rangle \vec{1} + \sum_{ij} w_{ij} \langle \vec{n_{13}} | \vec{n_i} \rangle \langle \vec{n_j} | \vec{n_7} \rangle \vec{1} \\ = & .8\vec{1} + 0\vec{1} + 0\vec{1} \\ = & .8\vec{1} \end{aligned}$$

- Question: Who is the lucky one?

Problems with the standard Lambek systems

Overgeneration because of global structural postulates

- ▶ global Associativity: grammaticality is preserved under arbitrary restructuring ↴
- ▶ global Commutativity: grammaticality is preserved under arbitrary reordering ↴

Undergeneration for logics respecting word order

- ▶ recognizing capacity **(N)L**: strictly context-free
- ▶ attested patterns beyond CF: w^2 , $a^n b^n c^n$, ...

Extended type logics

Three strategies for dealing with over/undergeneration issues:

- ▶ unary modalities for structural control: \Diamond, \Box
- ▶ arrow reversal symmetry: \otimes, \oplus
- ▶ primitive (pre/post) negations instead of units: $0., .0, 1., .1$

Unary modalities for structural control

- ▶ We add a pair of unary connectives \Diamond, \Box satisfying

$$\frac{\Diamond A \rightarrow B}{A \rightarrow \Box B} \text{ rp}$$

- ▶ Logic: \Diamond, \Box form a residuated pair. One easily shows

$$\Diamond \Box A \rightarrow A \rightarrow \Box \Diamond A \text{ (compositions)}$$

from $A \rightarrow B$ infer $\Diamond A \rightarrow \Diamond B, \Box A \rightarrow \Box B$ (mon)

- ▶ Structure: *global* rules $\rightsquigarrow \Diamond$ controlled *restricted* versions, e.g.

$$\begin{aligned} A^\diamond : \quad & (A \otimes B) \otimes \Diamond C \rightarrow A \otimes (B \otimes \Diamond C) \\ C^\diamond : \quad & (A \otimes B) \otimes \Diamond C \rightarrow (A \otimes \Diamond C) \otimes B \end{aligned}$$

Illustration: non-peripheral extraction

$\Diamond \Box np$: 'moveable' np ; key-and-lock: $\Diamond \Box np \rightarrow np$, once in place.

		found	
		$(np \setminus s) / np \quad \Diamond \Box np \rightarrow np$	
		<hr/>	
		found $\otimes \Diamond \Box np \rightarrow np \setminus s$	
		<hr/>	
Alice	np	$(\text{found} \otimes \Diamond \Box np) \otimes \text{there} \rightarrow np \setminus s$	there
		<hr/>	
		Alice $\otimes ((\text{found} \otimes \Diamond \Box np) \otimes \text{there}) \rightarrow s$	C^\diamond
		<hr/>	
		Alice $\otimes ((\text{found} \otimes \text{there}) \otimes \Diamond \Box np) \rightarrow s$	A^\diamond
		<hr/>	
		(Alice $\otimes (\text{found} \otimes \text{there})) \otimes \Diamond \Box np \rightarrow s$	
		<hr/>	
what	$np / (s / \Diamond \Box np)$	Alice $\otimes (\text{found} \otimes \text{there}) \rightarrow s / \Diamond \Box np$	
		<hr/>	
		what $\otimes (\text{Alice} \otimes (\text{found} \otimes \text{there})) \rightarrow np$	

Compare: 'Alice found *something interesting* there'

A general theory of structural control

Let $\mathcal{L}' = \mathcal{L} + P$ for some structural postulate P (Ass, Comm).

Kurtonina & Moortgat 1997: two types of modal translation to relate $\mathcal{L}, \mathcal{L}'$:

- ▶ $\mathcal{L}_{/, \otimes, \backslash} \vdash A \rightarrow B$ iff $\mathcal{L}'_{\Diamond, \Box, /, \otimes, \backslash} \vdash A^\flat \rightarrow B^\flat$
inhibiting: \cdot^\flat blocks applicability of structural option P
- ▶ $\mathcal{L}'_{/, \otimes, \backslash} \vdash A \rightarrow B$ iff $\mathcal{L}_{\Diamond, \Box, /, \otimes, \backslash} + P_\Diamond \vdash A^\sharp \rightarrow B^\sharp$
licensing: \cdot^\sharp provides access to a *controlled* version of P

We illustrate with **NL** vs **L**.

Controlling \otimes Associativity

One schema serves for the licensing/inhibiting directions:

$$\begin{aligned} p^\natural &= p \\ (A \otimes B)^\natural &= \Diamond(A^\natural \otimes B^\natural) \\ (A/B)^\natural &= \Box A^\natural / B^\natural \\ (B \setminus A)^\natural &= B^\natural \setminus \Box A^\natural \end{aligned}$$

Controlling \otimes Associativity

One schema serves for the licensing/inhibiting directions:

$$\begin{array}{rcl} p^\natural & = & p \\ (A \otimes B)^\natural & = & \Diamond(A^\natural \otimes B^\natural) \\ (A/B)^\natural & = & \Box A^\natural / B^\natural \\ (B \setminus A)^\natural & = & B^\natural \setminus \Box A^\natural \end{array}$$

- ▶ expressing **NL** in **L**: \Diamond blocks applicability of Ass, e.g.

$$\not\vdash ((a \setminus b) \otimes (b \setminus c))^\flat \rightarrow (a \setminus c)^\flat$$

Controlling \otimes Associativity

One schema serves for the licensing/inhibiting directions:

$$\begin{array}{rcl} p^\natural & = & p \\ (A \otimes B)^\natural & = & \Diamond(A^\natural \otimes B^\natural) \\ (A/B)^\natural & = & \Box A^\natural / B^\natural \\ (B \setminus A)^\natural & = & B^\natural \setminus \Box A^\natural \end{array}$$

- ▶ expressing **NL** in **L**: \Diamond blocks applicability of Ass, e.g.

$$\not\vdash ((a \setminus b) \otimes (b \setminus c))^\flat \rightarrow (a \setminus c)^\flat$$

- ▶ expressing **L** in **NL**: \Diamond provides access to controlled Ass

$$\Diamond(\Diamond(A \otimes B) \otimes C) \longleftrightarrow \Diamond(A \otimes \Diamond(B \otimes C)) \quad (\text{Ass}^\diamond) = (\text{Ass})^\sharp$$

Arrow reversal symmetry: fusion vs fission

- ▶ cotensor \oplus , left/right difference \otimes, \oslash satisfying

$$\frac{C \rightarrow B \oplus A}{B \oslash C \rightarrow A} \ dr \qquad \frac{C \rightarrow B \oplus A}{C \oslash A \rightarrow B} \ dr$$

arrow reversal: compare $A \otimes (A \setminus B) \rightarrow B$ vs $B \rightarrow (B \oslash A) \oplus A$

Arrow reversal symmetry: fusion vs fission

- ▶ cotensor \oplus , left/right difference \otimes, \oslash satisfying

$$\frac{C \rightarrow B \oplus A}{B \oslash C \rightarrow A} \ dr \qquad \frac{C \rightarrow B \oplus A}{C \oslash A \rightarrow B} \ dr$$

arrow reversal: compare $A \otimes (A \setminus B) \rightarrow B$ vs $B \rightarrow (B \oslash A) \oplus A$

- ▶ add-ons: linear distributivities, respecting order, structure.
G1, G3: mixed Ass; G2, G4: mixed Comm.

$$\frac{A \otimes B \rightarrow C \oplus D}{C \oslash A \rightarrow D/B} \ G1$$

$$\frac{A \otimes B \rightarrow C \oplus D}{C \oslash B \rightarrow A \setminus D} \ G2$$

$$\frac{A \otimes B \rightarrow C \oplus D}{B \oslash D \rightarrow A \setminus C} \ G3$$

$$\frac{A \otimes B \rightarrow C \oplus D}{A \oslash D \rightarrow C/B} \ G4$$

(the bidirectional case requires modal control)

Illustration: Japanese questions

Question word: infix. E.g. 'What did John buy?'

'John-ga **nani-o** katta ka' — lit. 'John **what** bought Q'

Illustration: Japanese questions

Question word: infix. E.g. 'What did John buy?'

'John-ga nani-o katta ka' — lit. 'John what bought Q'

$$\begin{array}{c}
 \vdots \qquad \vdots \\
 \hline
 n \otimes (a \otimes (a \backslash (n \setminus s))) \rightarrow s \quad q \rightarrow (q \oslash wh) \oplus wh \\
 \hline
 s \backslash q \rightarrow (n \otimes (a \otimes (a \backslash (n \setminus s))))) \backslash ((q \oslash wh) \oplus wh) \\
 \hline
 (n \otimes (a \otimes (a \backslash (n \setminus s))))) \otimes (s \backslash q) \rightarrow (q \oslash wh) \oplus wh \quad G1 \\
 \hline
 (q \oslash wh) \otimes (n \otimes (a \otimes (a \backslash (n \setminus s)))) \rightarrow wh / (s \backslash q) \\
 \hline
 n \otimes (a \otimes (a \backslash (n \setminus s))) \rightarrow (q \oslash wh) \oplus (wh / (s \backslash q)) \quad G2 \\
 \hline
 (q \oslash wh) \otimes (a \otimes (a \backslash (n \setminus s))) \rightarrow n \backslash (wh / (s \backslash q)) \\
 \hline
 a \otimes (a \backslash (n \setminus s)) \rightarrow (q \oslash wh) \oplus (n \backslash (wh / (s \backslash q))) \quad G1 \\
 \hline
 (q \oslash wh) \otimes a \rightarrow (n \backslash (wh / (s \backslash q))) / (a \backslash (n \setminus s)) \\
 \hline
 (n \otimes (((q \oslash wh) \otimes a) \otimes (a \backslash (n \setminus s)))) \otimes (s \backslash q) \rightarrow wh
 \end{array}$$

q: yes/no question; wh: constituent question; a: direct object

Negations as Primitives

(Co)tensor units lead to overgeneration (because of $I \rightarrow A/A$ etc)

- ▶ We add primitive (pre/post) negations ${}^0\cdot, {}^0., {}^1\cdot, {}^1.$ satisfying

$$\frac{A \rightarrow B^0}{B \rightarrow {}^0 A} \text{ gc}$$

$$\frac{{}^1 A \rightarrow B}{B^1 \rightarrow A} \text{ dgc}$$

Negations as Primitives

(Co)tensor units lead to overgeneration (because of $I \rightarrow A/A$ etc)

- ▶ We add primitive (pre/post) negations ${}^0\cdot, {}^0., {}^1\cdot, {}^1.$ satisfying

$$\frac{A \rightarrow B^0}{B \rightarrow {}^0A} \text{ gc} \qquad \frac{{}^1A \rightarrow B}{B^1 \rightarrow A} \text{ dgc}$$

- ▶ Logic: (dual) Galois connected pairs. One easily shows

$$A \rightarrow {}^0(A^0), \quad A \rightarrow ({}^0A)^0 \text{ and } {}^1(A^1) \rightarrow A, \quad ({}^1A)^1 \rightarrow A$$

antitonicity: from $A \rightarrow B$ infer ${}^0B \rightarrow {}^0A$, etc

Negations as Primitives

(Co)tensor units lead to overgeneration (because of $I \rightarrow A/A$ etc)

- ▶ We add primitive (pre/post) negations ${}^0\cdot, {}^0., {}^1\cdot, {}^1.$ satisfying

$$\frac{A \rightarrow B^0}{B \rightarrow {}^0 A} \text{ gc} \qquad \frac{{}^1 A \rightarrow B}{B^1 \rightarrow A} \text{ dgc}$$

- ▶ Logic: (dual) Galois connected pairs. One easily shows
 $A \rightarrow {}^0(A^0)$, $A \rightarrow ({}^0 A)^0$ and ${}^1(A^1) \rightarrow A$, $({}^1 A)^1 \rightarrow A$
 antitonicity: from $A \rightarrow B$ infer ${}^0 B \rightarrow {}^0 A$, etc
- ▶ Structure: collapse to 2-negation or 1-negation system

$$A^{lr} \leftrightarrow A \leftrightarrow A^{rl} \qquad A \leftrightarrow A^{\perp\perp}$$

via negation variants of the Grishin distributivities

Compositional interpretation for the extended logics

How can we adapt the functorial transition to vector-based models of meaning to the extended type logics?

- ▶ control modalities \Diamond, \Box
- ▶ cotensor, left/right difference \oplus, \otimes, \oslash
- ▶ primitive (pre/post) negations

Interpreting $NL_{\Diamond,\Box}$

- ▶ We add a basic space C , and set

$$\begin{aligned} I(\Diamond A) &= C \otimes I(A) \\ I(\Box A) &= C^* \otimes I(A) \end{aligned}$$

- ▶ Assuming that C expresses some *context*, the abstract meaning recipe of “what Alice found” becomes

what ($\lambda c \otimes F$. **found** ($F c$) **alice**)

- ▶ So, we interpret “what” as a vector that supplies a context and a context interpretation by which to evaluate the probability by which alice found some object.
- ▶ Context: “Once upon a time there were two keys, and Alice found the first key. It is not hard to determine **what Alice found.**”

Illustration

- ▶ C spanned by vectors \vec{obj} , \vec{pers} , and assume:

$$\vec{c} = 0.66\vec{obj} + 0.33\vec{pers}$$

$$\vec{F} = \vec{obj} \otimes \vec{key}_1 + \vec{obj} \otimes \vec{key}_2 + 0(\vec{obj} \otimes \vec{alice}) + \dots + \vec{pers} \otimes \vec{alice}$$

$$\vec{found} = \vec{alice} \otimes \vec{1} \otimes \vec{key}_1 + 0(\vec{alice} \otimes \vec{1} \otimes \vec{key}_2) + 0(\vec{key}_1 \otimes \vec{1} \otimes \vec{alice}) + \dots$$

- ▶ Interpreting the context against the context interpretation vector yields

$$\vec{F} \cdot \vec{c} = 0.66\vec{key}_1 + 0.66\vec{key}_2 + 0.33\vec{alice}$$

- ▶ We interpret “what” as the vector that evaluates its argument against \vec{c} and \vec{F} :

$$\begin{aligned}
 & \text{what } (\lambda c \otimes F. \text{ found } (F \cdot c) \text{ alice}) \\
 \mapsto & \langle \vec{alice} | \vec{alice} \rangle \langle \vec{key}_1 | 0.66\vec{key}_1 + 0.66\vec{key}_2 + 0.33\vec{alice} \rangle \vec{1} + \\
 & 0 \langle \vec{alice} | \vec{alice} \rangle \langle \vec{key}_2 | 0.66\vec{key}_1 + 0.66\vec{key}_2 + 0.33\vec{alice} \rangle \vec{1} + \dots \\
 & = 0.66 \vec{1}
 \end{aligned}$$

Interpreting Fission

- ▶ We can interpret the $(\oplus, \otimes, \oslash)$ family using the standard tensor product on vector spaces:

$$\begin{aligned} I(B \oplus A) &= I(B) \otimes I(A) \\ I(A \setminus B) &= I(A)^* \otimes I(B) \\ I(A \otimes B) &= I(A)^* \otimes I(B) \\ &\vdots \end{aligned}$$

- ▶ Linear distributivities are interpreted using the symmetry and associativity transformations of **FVect**:

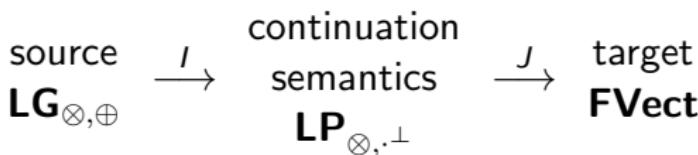
$$\begin{aligned} g_1 : A \otimes (B \otimes C) &\rightarrow (A \otimes B) \otimes C &\rightsquigarrow \alpha_{A^*, B, C}^{-1} \\ g_2 : A \otimes (C \otimes B) &\rightarrow C \otimes (A \otimes B) &\rightsquigarrow \alpha_{C, A^*, B} \circ (\sigma_{A^*, C} \otimes id_B) \circ \alpha_{A^*, C, B}^{-1} \\ &\vdots \end{aligned}$$

Cf. [Wijnholds 2014]

Avoiding collapse

Under the above interpretation, the source distinction between the \otimes and \oplus families collapses at the target level.

With a two-step interpretation one can avoid this collapse:



- ▶ intermediate level: continuation semantics for \mathbf{LG} , with linear \otimes and implication w.r.t. response type \perp ($A^\perp := A \multimap \perp$)
- ▶ source distinction \otimes, \oplus is retained in the interpretation

$$I(A \otimes B) = I(A) \otimes I(B) \quad vs \quad I(A \oplus B) = I(A)^\perp \otimes I(B)^\perp$$

Cf [Moortgat 2009]

Interpreting Negations

- Given a dagger structure induced by \cdot^\dagger , we can interpret negations:

$$\begin{aligned} I(A^0) &= I(A)^\dagger \\ I(A^1) &= I(A)^\dagger \\ &\vdots \end{aligned}$$

- Again, in the transition to the target interpretation, the source distinctions between pre/post negations are lost
- Linguistic application: lexical ambiguity
- Either **FVect** or **FHilb** allow Selinger's **CPM** construction, which can be used to measure ambiguity.
- Viz. 'Queen rules' vs 'Queen rules England'.

Cf. [Selinger 2007, Piedeleu et al. 2015]

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

- ▶ the resulting logics have enhanced syntactic control, avoiding the over/undergeneration issues of Lambek's original systems

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

- ▶ the resulting logics have enhanced syntactic control, avoiding the over/undergeneration issues of Lambek's original systems
- ▶ they still allow a functorial transition to vector-based models of meaning

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

- ▶ the resulting logics have enhanced syntactic control, avoiding the over/undergeneration issues of Lambek's original systems
- ▶ they still allow a functorial transition to vector-based models of meaning
- ▶ in the mapping to the target, many source distinctions are lost

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

- ▶ the resulting logics have enhanced syntactic control, avoiding the over/undergeneration issues of Lambek's original systems
- ▶ they still allow a functorial transition to vector-based models of meaning
- ▶ in the mapping to the target, many source distinctions are lost
- ▶ can one find empirical support for *enriched* quantitative models, retaining more of the source structure?

Summarizing, future work

We have discussed an extended vocabulary of logical constants:

positive formulas: $A \otimes B, A \oslash B, B \oslash A; \Diamond A, A^1, {}^1A;$

negative formulas: $A \oplus B, A \backslash B, B / A; \Box A, {}^0A, A^0$

- ▶ the resulting logics have enhanced syntactic control, avoiding the over/undergeneration issues of Lambek's original systems
 - ▶ they still allow a functorial transition to vector-based models of meaning
 - ▶ in the mapping to the target, many source distinctions are lost
 - ▶ can one find empirical support for *enriched* quantitative models, retaining more of the source structure?
- ... future work

Questions?

-  A. Bastenholf.
Categorial Symmetry.
PhD thesis, Utrecht University, 2013.
-  R. Bernardi and M. Moortgat.
Continuation semantics for the Lambek-Grishin calculus.
Inf. Comput., 208(5):397–416, 2010.
-  B. Coecke, M. Sadrzadeh, and S. Clark.
Mathematical foundations for a compositional distributional model of meaning.
CoRR, abs/1003.4394, 2010.
-  B. Jacobs.
Semantics of weakening and contraction.
Annals of Pure and Applied Logic, 69(1):73 – 106, 1994.
-  L. Kallmeyer.
Parsing Beyond Context-Free Grammars.
Springer, 2010.
-  N. Kurtonina and M. Moortgat.
Structural control.
In P. Blackburn and M. de Rijke, editors, *Specifying Syntactic Structures*, pages 75–113. CSLI, Stanford, 1997.
-  J. Lambek.
On the calculus of syntactic types.
In R. Jacobson, editor, *Structure of language and its mathematical aspects*, pages 166–178. American Mathematical Society, 1961.
-  M. Moortgat.
Symmetric categorial grammar.
J. Philosophical Logic, 38(6):681–710, 2009.
-  R. Piedeleu, D. Kartsaklis, B. Coecke, and M. Sadrzadeh.
Open system categorial quantum semantics in natural language processing.
CoRR, abs/1502.00831, 2015.
-  G. Wijnholds.
Categorical foundations for extended compositional distributional models of meaning.
MSc Thesis, ILLC, University of Amsterdam, 2014.