

Lambek Calculus and its modal extensions

Michael Moortgat, Utrecht University

WoLLIC 2023, Halifax

Plan

Ieri Categorical modalities, then and now.

- ▶ Soft Linear Logic ! and its subexponential, multimodal refinements
- ▶ Residuated families \diamond_i, \square_i

Oggi Dependency and function-argument structure.

- ▶ Dependency roles (subj, obj, ...) demarcating locality domains
- ▶ Rethinking constituency

Domani The neurosymbolic turn.

- ▶ Training data for type inference; constructive supertagging
- ▶ neural proof nets for parsing

A landscape of logics

Lambek calculi Identity $A \longrightarrow A$, composition $A \longrightarrow C$ if $A \longrightarrow B$ and $B \longrightarrow C$

Residuation: $B \longrightarrow A \backslash C$ iff $A \bullet B \longrightarrow C$ iff $A \longrightarrow C / B$

Options: \bullet associativity and/or commutativity; multiplicative unit

Substructural, sublinear a hierarchy of type logics reflecting different views on the **structure** of the assumptions Γ in sequent judgements $\Gamma \vdash A$.

LOGIC	Γ	ASS	COMM
LP	multiset	✓	✓
L	string	✓	-
NL	tree	-	-

► **(N)L**: syntactic types

NL types assigned to phrases (bracketed strings); **L**: types assigned to strings

► **LP** (aka unit-free MILL): semantic types aka unit-free MILL

The need for control

- ▶ languages exhibit phenomena that seem to require some form of
reordering, restructuring, copying
- ▶ global structural options are problematic
too little (undergeneration), too much (overgeneration)
- ▶ extended type language with modalities for structural control:
 - ▷ **licensing** structural reasoning that is lacking by default
 - ▷ **blocking** structural reasoning that would otherwise be available

Global associativity ☹️

Recall our relative clause example, derivable in **L** thanks to global associativity.

$$\begin{array}{c}
 \text{paper} \quad \text{that} \quad \text{Bob} \quad \text{rejected} \\
 \hline
 n \quad (n \backslash n) / (s / np) \quad np \quad (np \backslash s) / np \quad np \vdash np \\
 \hline
 \text{that} \cdot (Bob \cdot \text{rejected}) \vdash n \backslash n \quad \text{rejected} \cdot np \vdash np \backslash s \quad /E \\
 \hline
 \text{paper} \cdot (\text{that} \cdot (Bob \cdot \text{rejected})) \vdash n \quad \backslash E
 \end{array}$$

$\text{Bob} \cdot (\text{rejected} \cdot np) \vdash s$
 $(\text{Bob} \cdot \text{rejected}) \cdot np \vdash s$

$\text{that} \cdot (Bob \cdot \text{rejected}) \vdash s / np$
 $\text{Bob} \cdot \text{rejected} \vdash s / np$

$\text{that} \cdot (Bob \cdot \text{rejected}) \vdash n \backslash n$
 $\text{paper} \cdot (\text{that} \cdot (Bob \cdot \text{rejected})) \vdash n$

► **not enough** restricted to peripheral gaps, but

paper that Bob rejected __ immediately

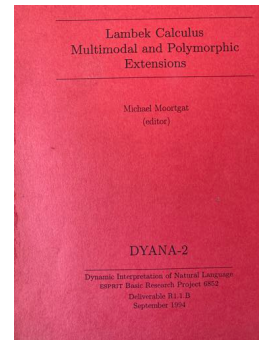
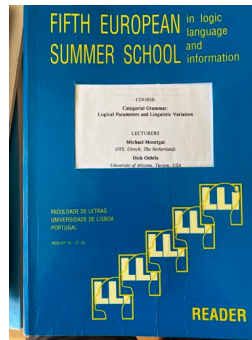
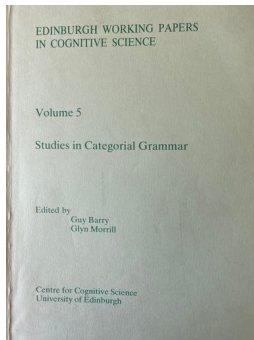
► **too much** insensitive to island constraints

paper that (Alice reviewed a thesis) and $_{(s \backslash s) / s}$ (B rejected __)

Vintage

The two views on modal extensions go back to the early 1990ies

- ▶ (Soft) Linear Logic ! and its subexponential, multimodal refinements
- ▶ Residuated families \diamond_i, \square_i



Morrill, Leslie, Hepple and Barry, 1990, Categorical Deductions and Structural Operations • MM & Oehrlé, 1993, ESSLLI Lisbon Lecture Notes • MM ed 1994, DYANA Report, Residuation in mixed Lambek systems, Controlling resource management

Modalities I: decomposing !

$$\begin{array}{ccc}
 \frac{! \Gamma \vdash A}{! \Gamma \vdash ! A} !R & \frac{\Gamma, A \vdash B}{\Gamma, ! A \vdash B} !L & \frac{\Gamma, A^n \vdash B}{\Gamma, ! A \vdash B} M \\
 \frac{\Gamma \vdash B}{\Gamma, ! A \vdash B} W & \frac{\Gamma, ! A, ! A \vdash B}{\Gamma, ! A \vdash B} C & \frac{\Gamma \vdash B}{! \Gamma \vdash ! B} SP
 \end{array}$$

Exponentials, multimodally Indexed $!_i$ for particular structural rules.

Cf Jacobs (1993,94) for syn/sem of $!_c, !_w$; fully generalized in Blaisdell et al 2022,23.

(Soft) linear logic ! Lafont 2004

terms: Baillot & Mogbil 2004

- ▶ Promotion ($!R$) is replaced by **soft** promotion (SP) (i.e. $!A \not\vdash !!A$); Dereliction ($!L$), Contraction, Weakening are replaced by Multiplexing (M)
- ▶ Cut elim/normalization: P
- ▶ Moot/Retoré 2019: SLL enough expressivity to specify **lexical** lambda terms
- ▶ SLL for **syntax**: ingenuity required for compatibility with non-comm, non-ass

Modalities II: residuated pairs

- ▶ The type language is extended with a pair of unary connectives \Diamond, \Box satisfying

$$\frac{\Diamond A \longrightarrow B}{A \longrightarrow \Box B}$$

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

compositions: $\Diamond \Box A \longrightarrow A$ (interior) $A \longrightarrow \Box \Diamond A$ (closure)

monotonicity: from $A \longrightarrow B$ infer $\Diamond A \longrightarrow \Diamond B, \Box A \longrightarrow \Box B$

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

$$A_{\Diamond}^r : (A \bullet B) \bullet \Diamond C \longrightarrow A \bullet (B \bullet \Diamond C)$$

$$C_{\Diamond}^r : (A \bullet B) \bullet \Diamond C \longrightarrow (A \bullet \Diamond C) \bullet B$$

Multimodal generalization families $\{\Diamond_i, \Box_i\}_{i \in I}$ for particular structural choices

\Diamond, \Box inverse duals



Relational semantics

Frames (W, R^2, R^3) . Valuation v sends types to subsets of W ,

$$\begin{aligned}v(A \bullet B) &= \{x \mid \exists yz. Rxyz \wedge y \in v(A) \wedge z \in v(B)\} \\v(C/B) &= \{y \mid \forall xz. (Rxyz \wedge z \in v(B)) \Rightarrow x \in v(C)\} \\v(A \setminus C) &= \{z \mid \forall xy. (Rxy \wedge y \in v(A)) \Rightarrow x \in v(C)\} \\v(\Diamond A) &= \{x \mid \exists y. (Rxy \wedge y \in v(A))\} \\v(\Box A) &= \{y \mid \forall x. (Rxy \Rightarrow x \in v(A))\}\end{aligned}$$

Soundness/completeness Kurtonina 1995 generalizing Došen 1992 for **(N)L(P)**

Extensions of **NL**_◊ with **weak Sahlqvist** postulates are complete w.r.t. the class of 2/3-ary frames satisfying the corresponding 1st order constraint effectively computable by the Sahlqvist-van Benthem algorithm.

Weak Sahlqvist postulates $A \longrightarrow B$ such that A is built out of single-use atoms and connectives \bullet, \Diamond ; B also is pure \bullet, \Diamond frm containing at least one occurrence of \bullet or \Diamond , with all atoms of B occurring in A .

Structural communication

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

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

► $\mathcal{L}_{/, \bullet, \backslash} \vdash A \longrightarrow B$ iff $\mathcal{L}'_{\diamond, \square, /, \bullet, \backslash} \vdash A^b \longrightarrow B^b$

inhibiting \cdot^b blocks applicability of structural option P

► $\mathcal{L}'_{/, \bullet, \backslash} \vdash A \longrightarrow B$ iff $\mathcal{L}_{\diamond, \square, /, \bullet, \backslash} + P_{\diamond} \vdash A^{\#} \longrightarrow B^{\#}$

licensing $\cdot^{\#}$ provides access to a controlled version of P

The $\cdot^{\#}$ direction of obtaining IL within MILL via ! exponential ($A \rightarrow B = !A \multimap B$).

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

Controlling Associativity

One schema serves for the licensing/inhibiting directions:

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

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

$$\not\models ((a \setminus b) \bullet (b \setminus c))^{\flat} \longrightarrow (a \setminus c)^{\flat}$$

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

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

N.D. Proofs and terms: syntactic calculi (N)L_{/, \}

Types, terms p atomic

$$A, B ::= p \mid A \backslash B \mid B / A \quad M, N ::= x \mid \lambda^r x. M \mid \lambda^l x. M \mid (M \ltimes N) \mid (N \rtimes M)$$

Wansing, 1990, Formulas-as-types for a Hierarchy of Sublogics of Int Prop Logic

Typing rules Axiom $x : A \vdash x : A$

var Γ, Δ all distinct

$$\frac{\Gamma \cdot x : A \vdash M : B}{\Gamma \vdash \lambda^r x. M : B / A} I/ \quad \frac{x : A \cdot \Gamma \vdash M : B}{\Gamma \vdash \lambda^l x. M : A \backslash B} I\backslash$$

$$\frac{\Gamma \vdash M : B / A \quad \Delta \vdash N : A}{\Gamma \cdot \Delta \vdash (M \ltimes N) : B} E/ \quad \frac{\Gamma \vdash N : A \quad \Delta \vdash M : A \backslash B}{\Gamma \cdot \Delta \vdash (N \rtimes M) : B} E\backslash$$

Compare: LP_→ **L** extended with product commutativity, a.k.a. MILL, Multiplicative Intuitionistic Linear Logic. In MILL, the slashes $/, \backslash$ collapse to linear implication \multimap .

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x. M : A \multimap B} (\multimap I) \quad \frac{\Gamma \vdash M : A \multimap B \quad \Delta \vdash N : A}{\Gamma, \Delta \vdash M N : B} (\multimap E)$$

Control operators: N.D. rules, terms

Structures Unary $\langle \rangle$ structural counterpart of \Diamond : $\Gamma, \Delta ::= A \mid \langle \Gamma \rangle \mid \Gamma \cdot \Delta$

$$\frac{\langle \Gamma \rangle \vdash A}{\Gamma \vdash \Box A} \Box I \qquad \frac{\Gamma \vdash \Box A}{\langle \Gamma \rangle \vdash A} \Box E$$

$$\frac{\Gamma \vdash A}{\langle \Gamma \rangle \vdash \Diamond A} \Diamond I \qquad \frac{\Delta \vdash \Diamond A \quad \Gamma[\langle A \rangle] \vdash B}{\Gamma[\Delta] \vdash B} \Diamond E \qquad \frac{\Gamma[\langle A \rangle] \vdash B}{\Gamma[\Diamond A] \vdash B} \Diamond E'$$

shorthand $(\Diamond E')$ if left premise of $(\Diamond E)$ is an axiom

Control operators: terms Terms: $M, N ::= x \mid \dots \mid \nabla M \mid \Delta M \mid \blacktriangledown M \mid \blacktriangle M$

$$\frac{\langle \Gamma \rangle \vdash M : A}{\Gamma \vdash \blacktriangle M : \Box A} \Box I \qquad \frac{\Gamma \vdash M : \Box A}{\langle \Gamma \rangle \vdash \blacktriangledown M : A} \Box E$$

$$\frac{\Gamma \vdash M : A}{\langle \Gamma \rangle \vdash \Delta M : \Diamond A} \Diamond I \qquad \frac{\Delta \vdash M : \Diamond A \quad \Gamma[\langle x : A \rangle] \vdash N : B}{\Gamma[\Delta] \vdash N[\nabla M/x] : B} \Diamond E$$

$\Diamond E$ officially: case ∇M of x in N

Controlled associativity/commutativity ☺

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

$$\begin{array}{c}
 \text{paper} \quad \frac{n}{\text{paper}} \quad \frac{\text{that} \quad (n \setminus n) / (s / \Diamond \Box np)}{\text{that} \cdot (\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \vdash n \setminus n} \quad \backslash E \\
 \hline
 \text{Bob} \quad \frac{np}{\text{Bob}} \quad \frac{\frac{\text{rejected} \quad \frac{\Box np \vdash \Box np}{\langle \Box np \rangle \vdash np} \quad \Box E}{(np \setminus s) / np} \quad / E \quad \frac{\text{immediately} \quad (np \setminus s) \setminus (np \setminus s)}{\text{immediately} \vdash np \setminus s} \quad \backslash E}{\text{rejected} \cdot \langle \Box np \rangle \vdash np \setminus s} \quad / E \\
 \hline
 \frac{\text{Bob} \cdot ((\text{rejected} \cdot \langle \Box np \rangle) \cdot \text{immediately}) \vdash s}{\text{Bob} \cdot ((\text{rejected} \cdot \text{immediately}) \cdot \langle \Box np \rangle) \vdash s} \quad C_{\Diamond}^r \\
 \hline
 \frac{\text{Bob} \cdot ((\text{rejected} \cdot \text{immediately}) \cdot \langle \Box np \rangle) \vdash s}{(\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \cdot \langle \Box np \rangle \vdash s} \quad A_{\Diamond}^r \\
 \hline
 \frac{(\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \cdot \langle \Box np \rangle \vdash s}{(\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \cdot \Diamond \Box np \vdash s} \quad \Diamond E' \\
 \hline
 \frac{(\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \cdot \Diamond \Box np \vdash s}{\text{Bob} \cdot (\text{rejected} \cdot \text{immediately}) \vdash s / \Diamond \Box np} \quad / I \\
 \hline
 \frac{\text{that} \cdot (\text{Bob} \cdot (\text{rejected} \cdot \text{immediately})) \vdash n \setminus n}{\text{paper} \cdot (\text{that} \cdot (\text{Bob} \cdot (\text{rejected} \cdot \text{immediately}))) \vdash n} \quad / E
 \end{array}$$

$$A_{\Diamond}^r : (A \bullet B) \bullet \Diamond C \longrightarrow A \bullet (B \bullet \Diamond C) \quad C_{\Diamond}^r : (A \bullet B) \bullet \Diamond C \longrightarrow (A \bullet \Diamond C) \bullet B$$

Proofs and terms

Adjusted lexical meaning recipe for the relative pronoun, $(n \setminus n) / (s / \diamond \square np)$

$$[\text{that}]^{lex} = \lambda v \lambda w \lambda z. ((w (\blacktriangledown \nabla z)) \wedge (v z))$$

► v of type $[s / \diamond \square np]^{lex} = \diamond \square e \rightarrow t$; w of type $[n]^{lex} = e \rightarrow t$

► z reusable $\diamond \square e$ variable distributed over the \wedge conjuncts

Proof term M , derivational $[M]^{der}$ and lexical $[M]^{lex}$ translations:

$$M = \text{paper} \rtimes (\text{that} \rtimes \lambda^r x. (\text{Bob} \rtimes ((\text{rejected} \rtimes (\blacktriangledown \nabla x))) \rtimes \text{immediately})) : n$$

$$[M]^{der} = ([\text{that}] \lambda x. (([\text{immediately}] ([\text{rejected}] (\blacktriangledown \nabla x))) [\text{Bob}])) [\text{paper}] : e \multimap t$$

$$[M]^{lex} = \lambda z. ((\text{PAPER} (\blacktriangledown \nabla z)) \wedge ((\text{IMMEDIATELY} (\text{REJECTED} (\blacktriangledown \nabla z))) \text{BOB})) : \diamond \square e \rightarrow t$$

From postulates to structural rules

Linearity general form of **linear** structural rules:

Moot 2002

$$\frac{\Gamma[\Xi[\Delta_1, \dots, \Delta_n]] \vdash A}{\Gamma[\Xi'[\Delta_{\pi_1}, \dots, \Delta_{\pi_n}]] \vdash A} R$$

► $\Xi[], \Xi'[]$ generalized contexts of arity n : $\mathcal{C} ::= [] \mid \langle \mathcal{C} \rangle \mid \mathcal{C} \cdot \mathcal{C}$ arity: # holes

► $\Xi[\Gamma_1, \dots, \Gamma_n]$ structure obtained by substitution of $\Gamma_1, \dots, \Gamma_n$ in $\Xi[]$ of arity n

Example controlled associativity/commutativity postulates in rule form

$$A_{\diamond}^r : (A \bullet B) \bullet \diamond C \longrightarrow A \bullet (B \bullet \diamond C) \quad C_{\diamond}^r : (A \bullet B) \bullet \diamond C \longrightarrow (A \bullet \diamond C) \bullet B$$

$$\frac{\Gamma[\Delta \cdot (\Delta' \cdot \langle \Delta'' \rangle)] \vdash A}{\Gamma[(\Delta \cdot \Delta') \cdot \langle \Delta'' \rangle] \vdash A} A_{\diamond}^r \quad \frac{\Gamma[(\Delta \cdot \langle \Delta'' \rangle) \cdot \Delta'] \vdash A}{\Gamma[(\Delta \cdot \Delta') \cdot \langle \Delta'' \rangle] \vdash A} C_{\diamond}^r$$

\rightsquigarrow replace formula vars by structure vars, \diamond, \bullet by their structural counterparts

Terms the linear structural rules leave the proof term unchanged

From postulates to structural rules (cont'd)

$$\frac{\Gamma[\Xi[\Delta_1, \dots, \Delta_n]] \vdash A}{\Gamma[\Xi'[\Delta_{\pi_1}, \dots, \Delta_{\pi_n}]] \vdash A} R$$

Linear, non-increasing R is non-increasing if $|\Xi'| \leq |\Xi|$

- ▶ number of unary $\langle \rangle$ in conclusion \leq in number of $\langle \rangle$ premise
- ▶ compare: $\Diamond(A \bullet B) \longrightarrow \Diamond A \bullet \Diamond B$ ✓; but not $\Diamond A \bullet \Diamond B \longrightarrow \Diamond(A \bullet B)$

Complexity, expressivity (Moot 2002) \mathbf{NL}_\Diamond + linear, non-increasing structural rules:

- ▶ decidable
- ▶ PSPACE complete
- ▶ recognizes the context-sensitive languages

Mildly CS fragments? Moot 2008, simulating TAGs \simeq 2-MCFG_{wn}

Controlling copying: lexicon or syntax?

Parasitic gaps felicitous only in the context of a primary gap, compare *c*, *d*

- | | | |
|----------|--|-------------|
| <i>a</i> | papers that Bob rejected $_$ (immediately) | gap |
| <i>b</i> | Bob left the room without closing the window | |
| <i>c</i> | *window that Bob left the room without closing $_$ | island |
| <i>d</i> | papers that reviewers rejected $_$ without reading $_$ (carefully) | pg: adjunct |
| <i>e</i> | security breach that a report about $_$ in the NYT made public $_$ | |

Reduction to lexical polymorphism

MM, Sadrzadeh, Wijnholds 2019

without^{*b,c*} :: $\square(X \setminus X)/Z, X = iv, Z = gp$ (gerund)
 without^{*d*} :: $\square((X/\diamond \square np) \setminus (X/\diamond \square np))/(Z/\diamond \square np)$

Semantically, with $[np \setminus s] = [gp] = N^* \otimes S$, $[\diamond \square np] = N$, without^{*d*} reduces to transitive verb coordination, i.e. $[\text{rejected}] \odot \neg[\text{reading}]$

$$(N \otimes S^* \otimes N) \otimes (N^* \otimes S \otimes N^*) \otimes (N \otimes S^* \otimes N)$$

Alternative: controlled contraction in syntax

Recall the postulates for regular gaps (no copying involved): controlled associativity A_{\diamond} , controlled commutativity C_{\diamond} allowing non-peripheral gaps.

$$A_{\diamond} : (A \bullet B) \bullet \diamond C \longrightarrow A \bullet (B \bullet \diamond C)$$

$$C_{\diamond} : (A \bullet B) \bullet \diamond C \longrightarrow (A \bullet \diamond C) \bullet B$$

We now add variants of A_{\diamond} , C_{\diamond} for the cases of extraction that involve copying:

$$A_{\diamond}^! : (\diamond A \bullet B) \bullet \diamond C \longrightarrow \diamond(A \bullet \diamond C) \bullet (B \bullet \diamond C)$$

$$C_{\diamond}^! : (A \bullet \diamond B) \bullet \diamond C \longrightarrow (A \bullet \diamond C) \bullet \diamond(B \bullet \diamond C)$$

- ▶ In addition to the principal gap, $A_{\diamond}^!$ and $C_{\diamond}^!$ drop a secondary gap in an island phrase (\diamond marked) that would be inaccessible without the principal gap.
- ▶ $A_{\diamond}^!$: pg precedes principal gap
- ▶ $C_{\diamond}^!$: pg follows principal gap

Illustration

$$\begin{array}{c}
\text{bob} \\
\hline
\text{np}
\end{array}
\frac{
\frac{
\frac{
\text{rejected} \quad \square np \vdash \square np \quad \square E
}{(np \backslash s) / np \quad \langle \square np \rangle \vdash np} / E
\quad
\frac{
\text{without} \quad \frac{\text{reading} \quad \square np \vdash \square np}{gp / np} \quad \langle \square np \rangle \vdash np \quad \square E
}{\square (s \backslash s) / gp \quad \text{reading} \cdot \langle \square np \rangle \vdash gp} / E
}{
\text{rejected} \cdot \langle \square np \rangle \vdash np \backslash s \quad \text{without} \cdot (\text{reading} \cdot \langle \square np \rangle) \vdash \square (s \backslash s) \quad \square E
} \backslash E
\frac{
\text{bob} \cdot (\text{rejected} \cdot \langle \square np \rangle) \vdash s \quad \langle \text{without} \cdot (\text{reading} \cdot \langle \square np \rangle) \rangle \vdash s \backslash s
}{
\text{bob} \cdot (\text{rejected} \cdot \langle \square np \rangle) \cdot \langle \text{without} \cdot (\text{reading} \cdot \langle \square np \rangle) \rangle \vdash s \quad A^\diamond, A^\diamond
}
\frac{
((\text{bob} \cdot \text{rejected}) \cdot \langle \square np \rangle) \cdot \langle (\text{without} \cdot \text{reading}) \cdot \langle \square np \rangle \rangle \vdash s \quad C_\diamond^!
}{
((\text{bob} \cdot \text{rejected}) \cdot \langle \text{without} \cdot \text{reading} \rangle) \cdot \langle \square np \rangle \vdash s \quad \diamond E, Ax
}
\frac{
\text{that} \quad ((\text{bob} \cdot \text{rejected}) \cdot \langle \text{without} \cdot \text{reading} \rangle) \cdot \diamond \square np \vdash s
}{
((\text{bob} \cdot \text{rejected}) \cdot \langle \text{without} \cdot \text{reading} \rangle) \vdash s / \diamond \square np \quad / I
}
\frac{
\text{paper} \quad (n \backslash n) / (s / \diamond \square np) \quad ((\text{bob} \cdot \text{rejected}) \cdot \langle \text{without} \cdot \text{reading} \rangle) \vdash s / \diamond \square np
}{
\text{that} \cdot \dots \vdash n \backslash n \quad / E
}
\frac{
\text{paper} \quad n \quad \text{that} \cdot \dots \vdash n \backslash n
}{
\text{paper} \cdot (\text{that} \cdot \dots) \vdash n \quad \backslash E
}$$

Blocking structural rules

Recall the **island violations** caused by (global or controlled!) associativity:

paper that (Alice reviewed a thesis) but_{(s\ s)/s} (Bob rejected __)

$$\begin{array}{c}
 \frac{\dots}{s} \quad \frac{\text{but}}{(s \setminus \Box s)/s} \quad \frac{\frac{B}{np} \quad \frac{\frac{\text{rejected}}{(np \setminus s)/np} \quad \frac{\Box np \vdash \Box np}{\langle \Box np \rangle \vdash np}}{\text{rejected} \cdot \langle \Box np \rangle \vdash np \setminus s} /E}{B \cdot (\text{rejected} \cdot \langle \Box np \rangle) \vdash s} \setminus E \\
 \frac{\dots}{s} \quad \frac{\text{but} \cdot (B \cdot (\text{rejected} \cdot \langle \Box np \rangle)) \vdash s \setminus \Box s}{\dots \cdot (\text{but} \cdot (B \cdot (\text{rejected} \cdot \langle \Box np \rangle))) \vdash \Box s} /E \\
 \frac{\dots \cdot (\text{but} \cdot (B \cdot (\text{rejected} \cdot \langle \Box np \rangle))) \vdash \Box s}{\langle \dots \cdot (\text{but} \cdot (B \cdot (\text{rejected} \cdot \langle \Box np \rangle))) \rangle \vdash s} \Box E \\
 \frac{\langle \dots \cdot (\text{but} \cdot (B \cdot \text{rejected})) \rangle \cdot \langle \Box np \rangle \vdash s}{\langle \dots \cdot (\text{but} \cdot (B \cdot \text{rejected})) \rangle \cdot \langle \Box np \rangle \vdash s} \text{???}
 \end{array}$$

◇ **as an obstacle** a modified type assignment imposes the desired island constraint:

► but :: (s \ \Box s)/s

Morill 1994

► \Box Elim seals off the conjunction as an island from which $\langle \Box np \rangle$ cannot escape

We will generalize this idea to demarcate **dependency domains** ...

Comparing RES and BANG

Correspondences Similarities more striking than differences, reading $!_i$ as $\Diamond_i \Box_i$
 Simulating $!_i$ properties as combinations of \Diamond, \Box logical and structural rules, e.g.

$$\frac{\Gamma \vdash B}{! \Gamma \vdash ! B} SP \qquad \frac{\frac{\Gamma \vdash B}{\langle \Box \rangle \Gamma \vdash B} \Box L \quad \frac{\langle \Box \Gamma \rangle \vdash B}{\Box \Gamma \vdash \Box B} \Box R}{\langle \Box \Gamma \rangle \vdash B} K$$

MM 1996

Differences some features of RES not shared by BANG

- ▶ licensing and blocking uses of modalities share same logical rules
- ▶ components \Diamond and \Box have individual uses, cf the dependency annotation

Resolution? Multitype approach, Palmigiano c.s., arguing that $!$ cannot be seen as primitive, but must be deconstructed in heterogeneous adjoint pair $\Diamond \blacksquare$

Dependency modalities

Heads vs dependents

Dependency roles articulate the linguistic material on the basis of two oppositions:

▶ head - **complement** relations

- ▷ verbal domain: subj, (in)direct object, ...
- ▷ nominal domain: prepositional object, ...

▶ **adjunct** - head relations

- ▷ verbal domain: (time, manner, ...) adverbial
- ▷ nominal domain: adjectival, numeral, determiner, ...

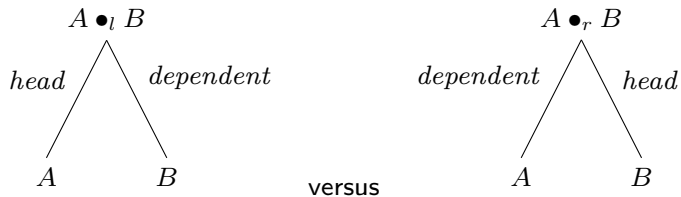
Compare: fa-structure: function vs argument

Orthogonality The fa and the dependency articulation are in general not aligned. This asks for a multidimensional type logic.

E.g. Determiner. Semantically, characteristic function of ($\llbracket N \rrbracket$, $\llbracket VP \rrbracket$) relation; morphologically, dependent on head noun.

DNL

Bimodal NL Moortgat & Morrill, 1991, Heads and phrases. Type calculus for dependency and constituent structure. Ms UU



$$\begin{aligned}
 A \longrightarrow C/_l B & \text{ iff } A \bullet_l B \longrightarrow C & \text{ iff } B \longrightarrow A \backslash_l C \\
 A \longrightarrow C/_r B & \text{ iff } A \bullet_r B \longrightarrow C & \text{ iff } B \longrightarrow A \backslash_r C
 \end{aligned}$$

- ▶ left vs right-headed •
- ▶ heads: $C/_l B$, $A \backslash_r C$; dependents: $C/_r B$, $A \backslash_l C$
- ▶ models: prosodic prominence, morphosyntactic government/rection, ...

Defining headed products

Left/right headed \bullet as composition of regular \bullet and modal marking of the dependent:

$$\text{left headed} := A \bullet \Diamond B \quad \text{right headed} := \Diamond A \bullet B$$

Residuation: translation of the slashes

recall: $\Diamond A \longrightarrow B$ iff $A \longrightarrow \Box B$

$$\frac{\frac{A \longrightarrow C/\Diamond B}{A \bullet \Diamond B \longrightarrow C}}{\Diamond B \longrightarrow A \setminus C} \quad \frac{\frac{A \longrightarrow \Box(C/B)}{\Diamond A \longrightarrow C/B}}{\Diamond A \bullet B \longrightarrow C}$$

$$\frac{\frac{\frac{A \bullet \Diamond B \longrightarrow C}{\Diamond B \longrightarrow A \setminus C}}{B \longrightarrow \Box(A \setminus C)}}{\frac{\frac{\Diamond A \bullet B \longrightarrow C}{B \longrightarrow \Diamond A \setminus C}}{B \longrightarrow \Diamond A \setminus C}}$$

Multimodal generalization families $\{\Diamond_d, \Box_d\}_{d \in \text{DepLabel}}$

- ▶ $\Diamond_d A \setminus C, C/\Diamond_d B$ head functor assigning dependency role d to its **complement**
- ▶ $\Box_d(A \setminus C), \Box_d(C/B)$ dependent functor projecting **adjunct** role d

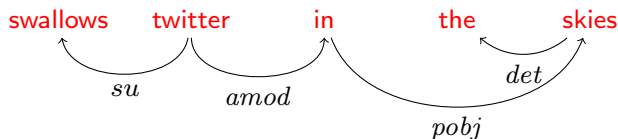
Dependency structure

Dependency-enhanced types:

$$\begin{array}{c}
 \frac{\frac{\text{swallows}}{np} \quad \frac{\text{twitter}}{\diamond_{su} np \backslash s}}{\langle \text{swallows} \rangle^{su} \vdash \diamond_{su} np \quad \diamond_{su} np \backslash s} \quad \diamond I \quad \backslash E \\
 \frac{\langle \text{swallows} \rangle^{su} \cdot \text{twitter} \vdash s}{(\langle \text{swallows} \rangle^{su} \cdot \text{twitter}) \cdot \langle \text{in} \cdot \langle \langle \text{the} \rangle^{det} \cdot \text{skies} \rangle^{pobj} \rangle^{amod} \vdash s} \quad \backslash E
 \end{array}$$

$$\begin{array}{c}
 \frac{\text{the}}{\square_{det}(np/n)} \quad \frac{\text{skies}}{n} \quad \square E \\
 \frac{\langle \text{the} \rangle^{det} \vdash np/n \quad \square E \quad \frac{\text{skies}}{n}}{\langle \text{the} \rangle^{det} \cdot \text{skies} \vdash np} \quad /E \\
 \frac{\text{in} \quad \square_{amod}(s \backslash s) / \diamond_{pobj} np \quad \langle \langle \text{the} \rangle^{det} \cdot \text{skies} \rangle^{pobj} \vdash \diamond_{pobj} np}{\text{in} \cdot \langle \langle \text{the} \rangle^{det} \cdot \text{skies} \rangle^{pobj} \vdash \square_{amod}(s \backslash s)} \quad \diamond I \\
 \frac{\text{in} \cdot \langle \langle \text{the} \rangle^{det} \cdot \text{skies} \rangle^{pobj} \vdash \square_{amod}(s \backslash s) \quad \square E}{\langle \text{in} \cdot \langle \langle \text{the} \rangle^{det} \cdot \text{skies} \rangle^{pobj} \rangle^{amod} \vdash s \backslash s} \quad \backslash E
 \end{array}$$

Induced dependency structure:



\leadsto within dependency domain, outgoing arcs from head to (head of) dependents

Extraction revisited

NL Relatives Dutch **left**-branch extraction via controlled associativity, commutativity

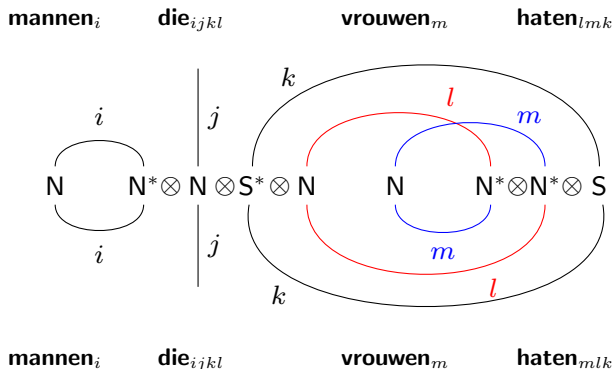
$$\Diamond_x A \bullet (B \bullet C) \longrightarrow (\Diamond_x A \bullet B) \bullet C \quad \Diamond_x A \bullet (B \bullet C) \longrightarrow B \bullet (\Diamond_x A \bullet C)$$

Relative pronoun: die :: $(n \setminus n) / (!_x n p \setminus s)$

$$!_x A \triangleq \Diamond_x \Box_x A$$

ambiguous between subj/obj relativization:

s subordinate clause, head-final



Extraction revisited (cont'd)

Dependency refinement **derivational** ambiguity is traded in for **lexical** ambiguity, to be resolved in the supertagging phase.

- ▶ NL is head-final: transitive verb type:

hate :: $\Diamond_{obj} np \backslash (\Diamond_{subj} np \backslash s)$

- ▶ two relative pronoun types: subject vs object relativization

die :: $\Box_{mod}(n \backslash n) / \Diamond_{body}(\Diamond_{subj} np \backslash s)$

die :: $\Box_{mod}(n \backslash n) / \Diamond_{body}(\Diamond_{obj} np \backslash s)$

Rethinking constituency

Associativity head + \diamond_d demarcated dependents constitutes dependency domain;
within these domains • associativity freely available.

Down the rabbit hole The above relpro types restrict access to immediate dependents of the rel clause body. $\text{die} :: \Box_{\text{mod}}(n \setminus n) / \Diamond_{\text{body}}(!x \Diamond_{\text{subj} \setminus \text{obj}} np \setminus s)$ reaches more deeply embedded hypotheses.

The *xleft* (derived) inference rule now has $\Gamma[\]$ traversing **unary**+binary structure:

$$\frac{\Gamma[A \cdot \Delta] \vdash B}{\Gamma[\Delta] \vdash !_x A \setminus B} \text{ xleft}$$

\rightsquigarrow requires extra postulate allowing \diamond_x to commute with dependency modalities \diamond_d for (all | some) $d \in \text{DepLabel}$:

$$\diamond_x A \bullet \diamond_d B \longrightarrow \diamond_d(\diamond_x A \bullet B)$$

$$\frac{\Gamma[\langle \langle \Delta \rangle^x \cdot \Delta' \rangle^d] \vdash A}{\Gamma[\langle \Delta \rangle^x \cdot \langle \Delta' \rangle^d] \vdash A} \text{ xleft'}$$

A neurosymbolic perspective

Challenges

Recall we write $L(G, B)$ for the strings of type B recognized by grammar G .

$w_1 \cdots w_n \in L(G, B)$ if the following hold:

- $(w_i, A_i) \in \text{Lex}$ for $1 \leq i \leq n$;
- $\Gamma_{[A_1, \dots, A_n]} \vdash B$, for Γ an antecedent structure with yield A_1, \dots, A_n

► type ambiguity: what is the right type for w_i given its context?

\leadsto supertagging

► structural ambiguity: what is the proper structure for Γ to derive B

\leadsto parsing

Training data: NL types in the wild

Type lexicon + derivations/ λ terms extracted from Lassy Small, gold standard treebank of written Dutch. 68782 samples.

- ▶ Lassy annotation: DAGs, nodes: categories, edges: dependency relations
- ▶ Re-entrancy: gaps, coordination, ('understood subjects' of non-finite verb forms)
- ▶ traditional dependency roles; can be mapped to UD Bouma & vNoord 2017

Lassy2Æthel extraction

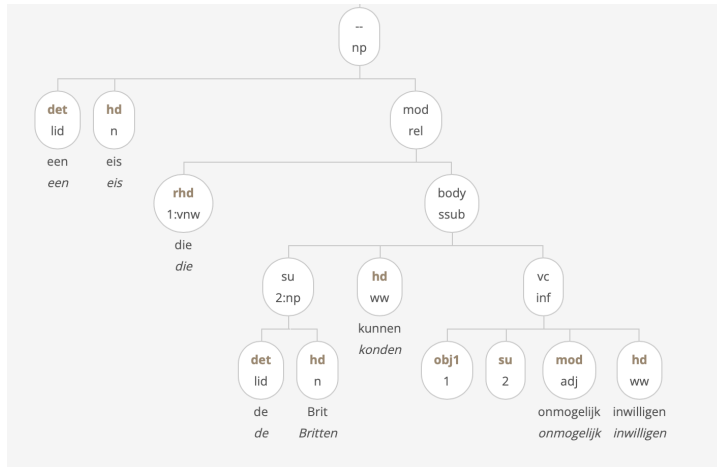
- ▶ non-directional syntax types: alignment with surface string left to neural parser
- ▶ modalities: dependency marking; structural control (extraction)
- ▶ finegrained result — compare CCG: categories 5292/1323, slashes(+ \diamond , \square) 29/2

Ref Kogkalidis, MM & Moot, 2020

Æthel: Automatically Extracted Typological Derivations for Dutch. [LREC](#).

<https://github.com/konstantinosKokos/aethel>

A demand that the British couldn't possibly grant: Lassy



- ▶ tree display format, avoiding crossing edges
 - ▶ re-entrancy relpro 'die' ~ obj1: gap hypothesis
 - ▶ re-entrancy su 'de Britten' ~ su: understood subject infinitive
- word order: position indices

A demand that the British couldn't possibly grant

Sample WR-P-E-I-0000015007.p.1.s.51.xml(27) ()

```
In [26]: sample = aethel[37628]
```

```
In [28]: sample.sentence
```

```
Out[28]: 'een eis die de Britten onmogelijk konden inwilligen .'
```

```
In [27]: list(sample.lexical_phrases)
```

```
Out[27]: [LexicalPhrase(string=een, type=det(N__NP), len=1),
LexicalPhrase(string=eis, type=N, len=1),
LexicalPhrase(string=die, type=(relcl(ox(x(oxobj1(VNW)))__SSUB))__mod(NP__NP), len=1),
LexicalPhrase(string=de, type=det(N__NP), len=1),
LexicalPhrase(string=Britten, type=N, len=1),
LexicalPhrase(string=onmogelijk, type=mod(INF__INF), len=1),
LexicalPhrase(string=konden, type=ovc(INF)__osu(NP)__SSUB, len=1),
LexicalPhrase(string=inwilligen, type=obj1(VNW)__INF, len=1),
LexicalPhrase(string=., type=PUNCT, len=1)]
```

```
In [22]: proof=sample.proof
```

```
In [23]: print(proof)
```

```
<c2, <c6, < <c5> mod, c7> vc, < <c3> det, c4> su> relcl> mod, <c0> det, c1 ⊢ ◻mod(c2 ◻relcl((λx0.case ◻x(x0) of x1
in (c6 ◻vc(◻mod(c5) (c7 ◻x(x1))) ◻su(◻det(c3) c4)))))) (◻det(c0) c1) : NP
```

$\nabla^{mod}(\text{die} \ \Delta^{relcl} (\lambda x_0. \text{case } \nabla^x x_0 \text{ of } x_1 \text{ in } (\text{konden} \ \Delta^{vc} (\nabla^{mod} \text{onmogelijk} (\text{inwilligen } \nabla^x x_1)) \ \Delta^{su} (\nabla^{det} \text{de Britten})))) (\nabla^{det} \text{een eis})$

Going neural

PhD project Konstantinos Kogkalidis

- ▶ Kogkalidis, 2023, Dependency as Modality, Parsing as Permutation.
Phd Thesis, Utrecht University. [url](#)
- ▶ Kogkalidis & MM, 2022, [arXiv](#)
Geometry-Aware Supertagging with Heterogeneous Dynamic Convolutions
- ▶ Kogkalidis, MM & Moot, 2020
Neural Proof Nets. CoNLL [url](#)

Code: <https://github.com/konstantinosKokos/spindle>

Integrating supertagging and neural parsing

Neural proof nets The parsing method uses LL proof nets. Proof net construction can be seen as a staged process:

- ▶ proof frame: forest of formula decomposition trees — supertagging ☺
- ▶ proof structure: p frame plus pairwise linking of in/out atoms
- ▶ proof net: p structure with successful traversal

MILL_—^{◇,□} lambda term as byproduct of traversal

Key neural methods

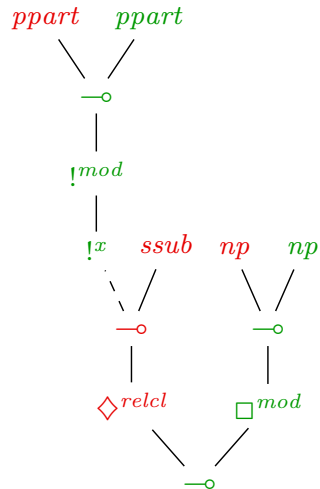
- ▶ supertagging: **parallel** tree decoding with dynamic graph convolutions
- ▶ axiom linking: Sinkhorn iterative method to approach double stochastic matrix
- ▶ verification: Lamarche traversal method

Lamarche 2008

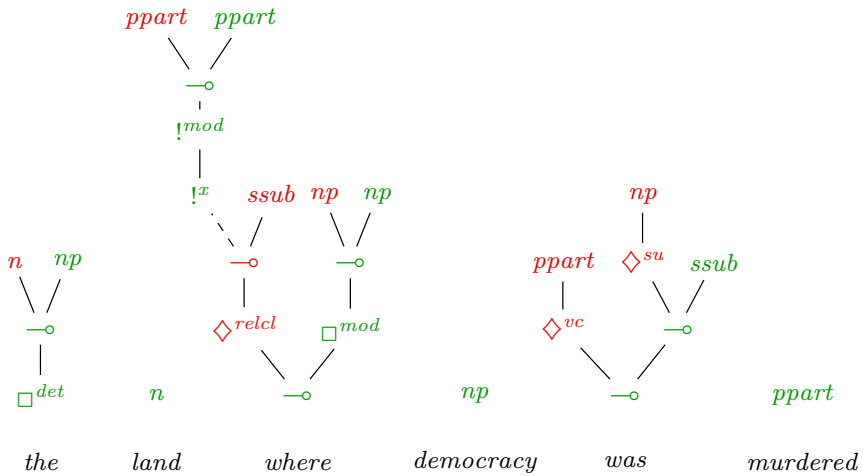
Supertag = polarized formula decomposition tree

Example the land **where** democracy was murdered

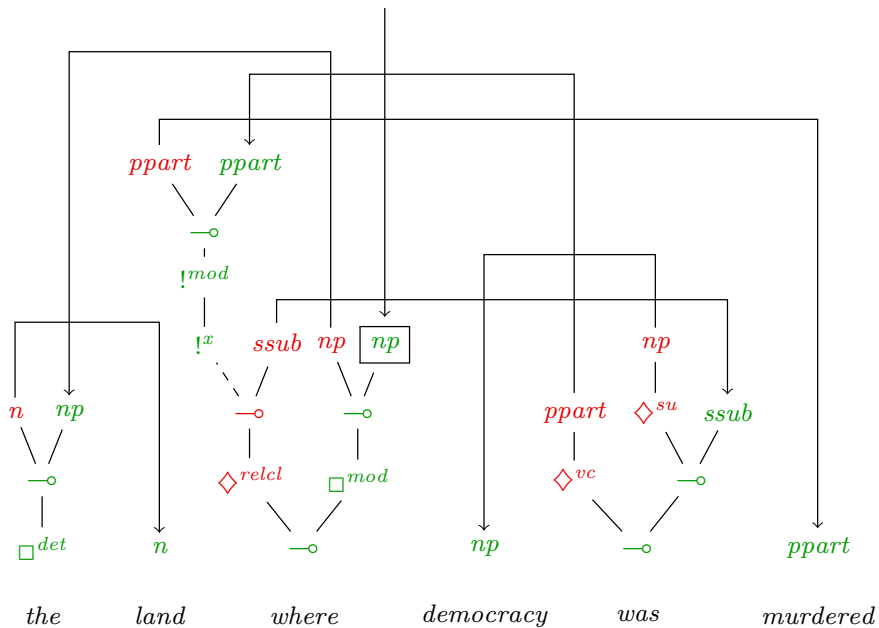
Polarities green: given, input; red: to prove, output



Proof frame

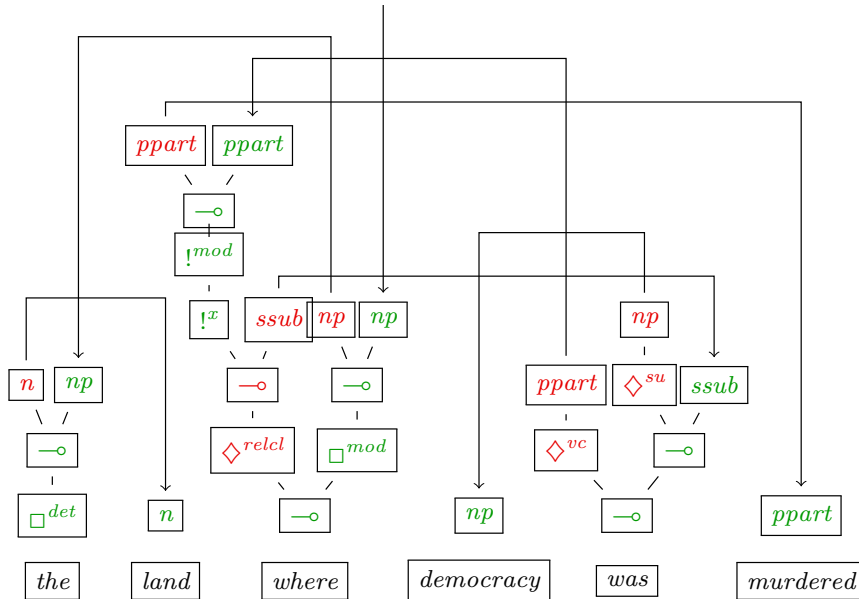


Proof frame \leadsto proof structure



Proof structure \rightsquigarrow proof net

the ESSLI2022 movie is [here](#)



$\nabla^{mod}(\text{where } \Delta^{relcl}(\lambda x_0. (\text{was } \Delta^{vc}(\nabla^{mod} \nabla^{mod} \nabla^x \nabla^x x_0 \text{ murdered}) \Delta^{su} \text{ democracy}))) (\nabla^{det} \text{the land})$

Conclusions

Some key themes of the talk:

- ▶ Logic: a modally enhanced multi-dimensional type logic
 - ▷ dependency structure \perp function-argument structure
 - ▷ linear \diamond, \square lambda terms as general-purpose recipes for meaning composition
 - ▷ where possible, confine non-linearity to lexical meaning recipes
- ▶ NLP: end-to-end compositionality:
 - ▷ obtain elementary word embeddings from data, and additionally
 - ▷ their types and their internal composition
 - ▷ neural parsing (grounded in/informed by) data-driven word representations

References

- [1] Guy Barry, Mark Hepple, Neil Leslie, and Glyn Morrill. Proof figures and structural operators for categorial grammar. In *Proceedings of EACL 1995*, pages 198–203, Berlin, 1991.
- [2] Eben Blaisdell. Decidable exponentials in nonassociative noncommutative linear logic. *CoRR*, abs/2306.12661, 2023.
- [3] Eben Blaisdell, Max I. Kanovich, Stepan L. Kuznetsov, Elaine Pimentel, and Andre Scedrov. Non-associative, non-commutative multi-modal linear logic. In Jasmin Blanchette, Laura Kovács, and Dirk Pattinson, editors, *Automated Reasoning - 11th International Joint Conference, IJCAR 2022, Haifa, Israel, August 8-10, 2022, Proceedings*, volume 13385 of *Lecture Notes in Computer Science*, pages 449–467. Springer, 2022.
- [4] Eben Blaisdell, Max I. Kanovich, Stepan L. Kuznetsov, Elaine Pimentel, and Andre Scedrov. Explorations in subexponential non-associative non-commutative linear logic. *CoRR*, abs/2307.03059, 2023.
- [5] Paula Czarnecka, Guy Emerson, and Ann Copestake. Words are vectors, dependencies are matrices: Learning word embeddings from dependency graphs. 2019.

- [6] Giuseppe Greco and Alessandra Palmigiano. Linear logic properly displayed. *ACM Trans. Comput. Logic*, 2022.
- [7] Bart Jacobs. Semantics of lambda-i and of other substructure lambda calculi. In Marc Bezem and Jan Friso Groote, editors, *Typed Lambda Calculi and Applications, International Conference on Typed Lambda Calculi and Applications, TLCA '93, Utrecht, The Netherlands, March 16-18, 1993, Proceedings*, volume 664 of *Lecture Notes in Computer Science*, pages 195–208. Springer, 1993.
- [8] Bart Jacobs. Semantics of weakening and contraction. *Ann. Pure Appl. Log.*, 69(1):73–106, 1994.
- [9] Konstantinos Kogkalidis and Michael Moortgat. Geometry-aware supertagging with heterogeneous dynamic convolutions. *CoRR*, abs/2203.12235, 2022.
- [10] Konstantinos Kogkalidis, Michael Moortgat, and Richard Moot. Æthel: Automatically extracted typelogical derivations for Dutch. In *Proceedings of The 12th Language Resources and Evaluation Conference, LREC 2020, Marseille*, pages 5257–5266. European Language Resources Association, 2020.
- [11] Konstantinos Kogkalidis, Michael Moortgat, and Richard Moot. Neural proof nets. In *CoNLL2020, Proceedings of the 24th Conference on Computational Natural Language Learning*, pages 26–40. Association for Computational Linguistics, 2020.

- [12] Natasha Kurtonina. *Frames and Labels. A Modal Analysis of Categorical Inference*. PhD thesis, OTS Utrecht, ILLC Amsterdam, 1995.
- [13] Natasha Kurtonina and Michael Moortgat. Structural control. In Patrick Blackburn and Maarten de Rijke, editors, *Specifying Syntactic Structures*, pages 75–113. CSLI, Stanford, 1997.
- [14] Yves Lafont. Soft linear logic and polynomial time. *Theor. Comput. Sci.*, 318(1-2):163–180, 2004.
- [15] François Lamarche. Proof Nets for Intuitionistic Linear Logic: Essential Nets. Research report, 2008.
- [16] Michael Moortgat. Multimodal linguistic inference. *J. Log. Lang. Inf.*, 5(3/4):349–385, 1996.
- [17] Michael Moortgat, Konstantinos Kogkalidis, and Gijs Wijnholds. Diamonds are forever. Theoretical and empirical support for a dependency-enhanced type logic. In R. Loukanova, P. Lumsdaine, and R. Muskens, editors, *Logic and Algorithms in Computational Linguistics 2021*, Studies in Computational Intelligence. Springer, (to appear).
- [18] Michael Moortgat and Glyn Morrill. Heads and phrases: Type calculus for dependency and constituent structure. Technical report, Research Institute for Language and Speech (OTS), Utrecht, 1991.

- [19] Michael Moortgat, Mehrnoosh Sadrzadeh, and Gijs Wijnholds. A Frobenius algebraic analysis for parasitic gaps. *FLAP*, 7(5):823–852, 2020.
- [20] Richard Moot. Lambek grammars, tree adjoining grammars and hyperedge replacement grammars. In Claire Gardent and Anoop Sarkar, editors, *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Frameworks, TAG+ 2010 , Tübingen, Germany, June 6-8, 2008*, pages 65–72. Association for Computational Linguistics, 2008.
- [21] Richard Moot. *Type-logical investigations: proof-theoretic, computational and linguistic aspects of modern type-logical grammars*. Habilitation à diriger des recherches, Université Montpellier, November 2021.
- [22] Richard Moot and Christian Retoré. Natural language semantics and computability. *J. Log. Lang. Inf.*, 28(2):287–307, 2019.
- [23] Glyn Morrill, N. Leslie, Mark Hepple, and Guy Barry. Categorical deductions and structural operations. In G. Barry and G Morrill, editors, *Studies in Categorical Grammar*, volume 5 of *Edinburgh Working Papers in Cognitive Science*, pages 1–21. Centre for Cognitive Science, 1990.
- [24] Mehrnoosh Sadrzadeh and Reinhard Muskens. Static and dynamic vector semantics for lambda calculus models of natural language. *J. Lang. Model.*, 6(2):319–351, 2018.