Lambek-Grishin Calculus: Focusing, Display and Full Polarization

Giuseppe Greco¹ Michael Moortgat² Valentin D. Richard³ Apostolos Tzimoulis¹

¹ Vrije Universiteit, the Netherlands
² Institutes of Linguistics OTS, Universiteit Utrecht, the Netherlands
³ ENS Paris-Saclay, France

TACL 2022, Coimbra 23 June





école _____ normale _____ supérieure _____ paris _ saclay ____

Identity of proofs

Identifying or telling apart proofs has far-reaching consequences.

- Philosophy and mathematics: when do two proofs correspond to the same argument?
- Computer science: when do two algorithms correspond to the same program?
- Linguistics: how to capture different readings of the same sentence?

...

Sequent calculi exhibit syntactically different proofs of the very same end-sequent often due to trivial permutations of inference rules.

Natural deduction calculi or proof nets are less sensitive to inference rule permutations and are taken as **benchmarks** for defining identity of proofs.

Focused sequent calculi [And92, And01, Mil04] make use of <u>syntactic restrictions</u> on the applicability of inference rules achieving three main goals:

- 1 the proof search space is reduced retaining completeness;
- 2 every cut-free proof comes in a special normal form;
- criterion for defining identity of sequent calculi proofs.

What is the mathematical underpinning of focalization?

Looking for:

- (uniform and modular structural) proof theory and
- (algebraic and categorical) semantics.

Lambek-Grishin logic (expanded with analytic structural rules closed under mutations, i.e. an equivalence relation between structural connectives: see Appendix A.):

- (heterogenous multi-type) focused display calculus fD.LG
- fully polarized algebraic semantics FP.LG

where:

- fD.LG has canonical cut-elimination, strong focalization, and is complete w.r.t. FP.LG
- **fD.LG** is complete w.r.t. LG-algebras \rightsquigarrow semantic proof of completeness of focusing (given that the standard display calculus for LG is complete w.r.t. LG-algebras)
- effective translation between **fD.LG**-proofs and **fLG**-proofs [MM12] \rightsquigarrow operational semantics (given that **fLG**-derivations are in a Curry-Howard correspondence with directional $\overline{\lambda\mu\mu}$ -terms)

General theory:

- heterogenous multi-type display calculi
- fully polarized algebras

We expect that the approach extends to every displayable logic (see Conclusions).

Basic Lambek-Grishin algebra [Moo09]:

- Poset $\mathbb{G} = (G, \leq)$
- 6 operations $\otimes, \oplus, \setminus, \otimes, /, \oslash$ s.t.

$B \leq A \setminus C$	iff	$A \otimes B \leq C$	iff	$A \leq C / B$
$C \oslash B \le A$	iff	$C \leq A \oplus B$	iff	$A \otimes C \leq B$

John		sleeps	is	a sentence
np	⊗	np \ s	_ ≤	S

(1)

- Natural generalization of Gentzen's sequent calculi;
- Display property:

$$\frac{Y \vdash X > Z}{X; Y \vdash Z}$$

display rules semantically justified by adjunction/residuation

- Multi-type: Separate syntactic types for different types of semantic objects
- Proper: Rules closed under uniform substitution (Wansing '98) within each type
- Canonical proof of cut elimination (via metatheorem)

Definition

A proper DC verifies each of the following conditions:

- structures can disappear, formulas are forever;
- **I tree-traceable** formula-occurrences, via suitably defined *congruence* relation (same shape, position, non-proliferation)
- principal = displayed
- rules are closed under uniform substitution of congruent parameters within each type (Properness!);
- **5** reduction strategy exists when cut formulas are principal.
- type-uniformity of derivable sequents;
- **strongly uniform cuts** in each/some type(s).

Theorem (Canonical!)

Cut elimination and subformula property hold for any proper m.DC.

Basic Lambek-Grishin logic 2/2

D.LG consists of the following rules (we consider only the Lambek fragment for brevity). **Axioms and cuts:**

$$p \vdash p$$
 Id $\frac{X \vdash A \quad A \vdash Y}{X \vdash Y}$ Cut

Logical rules (i.e. translation vs tonicity rules, cfr. asynchronous vs synchronous [And01]):

$$\overset{A \hat{\otimes} B \vdash X}{A \otimes B \vdash X} \quad \frac{X \vdash A \quad Y \vdash B}{X \hat{\otimes} Y \vdash A \otimes B} \otimes_{R}$$

Display postulates:

$$\hat{\otimes} + \breve{(} \frac{Y \vdash X \breve{(} Z)}{X \hat{\otimes} Y \vdash Z}$$
$$\underline{\hat{\otimes}} + \breve{(} \frac{X \hat{\otimes} Y \vdash Z}{X \vdash Z \breve{(} Y)}$$

We may expand the calculus with so-called Structural rules, e.g.:

$$\frac{(X \otimes Y) \otimes Z \vdash W}{X \otimes (Y \otimes Z) \vdash W} a_{,a^{-1}}$$

7/29

Everybody needs somebody



Everybody		needs		somebody	is	a sentence
$s / (np \setminus s)$	\otimes	$\overline{(((np \setminus s) / ((s / np) \setminus s)))}$	\otimes	$(s / np) \setminus s)$	F	S

There are 7 different sequent derivations, but only 3 different natural deduction (or proof net) derivations (in normal form).

Moving to a focused sequent system (**fLG** or **fD.LG**) we have again 3 derivations in normal form (with the bias assignment *np*:: positive, *s*:: negative).

Two derivations use associativity and correspond to the following readings:

- ∀-∃ reading: Everybody > somebody > needs
- ∃-∀ reading: Somebody > everybody > needs

The key idea relies on the following distinction.

- A focused phase is a proof-section where a formula is decomposed "as much as possible" only by means of non-invertible logical rules. This formula and all its immediate subformulas in this proof-section are said 'in focus'.
- A neutral phase is a non-focused phase, i.e. a proof section built by translation rules (applied greedily) or structural rules.

A strongly focalized proof exhibits a strict alternation between focused and neutral phases:



Definition

A sequent proof π is strongly focalized if cut-free and, for every formula A occurring in π , every **PIA subtree** of A is constructed by a proof-section of π containing only tonicity rules.

Two focalized phases:

- positive phase: only non-invertible logical rules for positive connectives are applied;
- negative phase: only non-invertible logical rules for negative connectives are applied.

How to categorize a connective as "positive" or "negative"? The usual answer has to do with the distinction: "right" versus "left" logical rules.

The mathematical underpinning is the following:

- Positive formulas: the main connective is a left-adjoint/residual (LG: ⊗, ⊘, ⊙);
- Negative formulas: the main connective is a right-adjoint/residual (LG: ⊕, \, /).

The key idea of polarization "naturally" calls for a type distinction. So, multi-type calculi seem a good candidate... but we need a further generalization.

A step back: focalization via "implicit" polarization

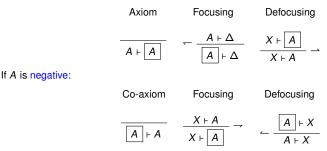
State of the art: **fLG** by Moortgat and Moot 2011 [MM12]

- Every proof is strongly focalized
- Focus implemented by a meta-linguistical marker A



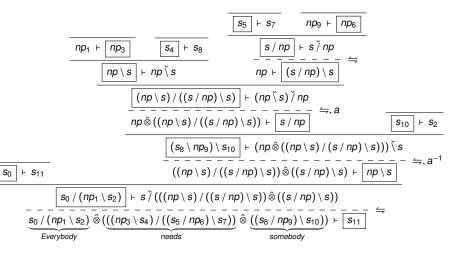
Restrictions on the applicability of rules

If A is positive:

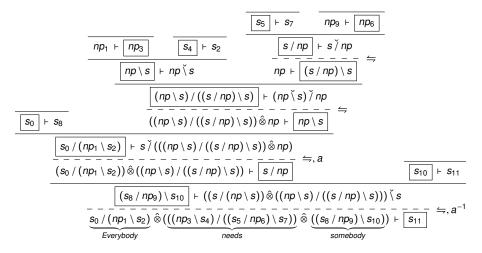


Tonicity rules have auxiliary and principal formulas in focus.

Bias assignment: *np* :: positive, *s* :: negative.



Bias assignment: *np* :: positive, *s* :: negative.



In the proof-theoretical literature, so-called shifts "operator" have been considered.

The key idea is the following:

- if A is negative, then $\downarrow A$ is positive;
- if A is positive, then $\uparrow A$ is negative.

But their status as operators is obscure.

On the other hand, in algebraic/categorical polarized semantics we have:

- ∎ ↑⊣↓;
- $\blacksquare \uparrow \downarrow \uparrow \varphi = \uparrow \varphi; \downarrow \uparrow \varphi = \varphi;$
- $\blacksquare \downarrow \uparrow \downarrow \varphi = \downarrow \varphi; \uparrow \downarrow \varphi = \varphi.$

Problem: the focusing policy could be destroyed.

The usual solution is to consider only sequents where \uparrow (resp. \downarrow) does not immediately occurr under the scope of \downarrow (resp. \uparrow).

<u>Our solution</u>: we distinguish between positive (resp. negative) **pure** formulas and positive (resp. negative) **shifted** formulas, i.e. formulas under the scope of a shift operator.

Weakening relations

W.R. are the order-theoretic equivalents of profunctors (aka distributors or bimodules) [Ben73].

W.R. are generalizations of partial orders: take $\mathcal{A} = \mathcal{B}$ and $\leq_{\mathcal{A}} = \leq_{\mathcal{B}}$.

Definition

A weakening relation is a relation $\leq \subseteq \mathcal{A} \times \mathcal{B}$ on two partially ordered set $(\mathcal{A}, \leq_{\mathcal{A}})$ and $(\mathcal{B}, \leq_{\mathcal{B}})$ that is compatible with the orders $\leq_{\mathcal{A}}$ and $\leq_{\mathcal{B}}$ in the following sense

$$\begin{array}{c|c} A' \leq_{\mathcal{R}} A & A \leq B & B \leq_{\mathcal{B}} B' \\ \hline & A' \leq B' \end{array}$$

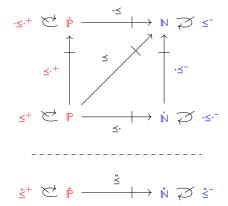
Definition

Given two w.r. $\leq_{\mathcal{R}} \subseteq \mathcal{A} \times \mathcal{A}'$ and $\leq_{\mathcal{B}} \subseteq \mathcal{B} \times \mathcal{B}'$, we say that the order-preserving functions $L : \mathcal{A} \to \mathcal{B}$ and $R : \mathcal{B}' \to \mathcal{A}'$ form a **heterogeneous adjoint pair** $L \stackrel{\leq_{\mathcal{B}}}{\underset{=}{\to}} R$ if for every $A \in \mathcal{A}$ and $B' \in \mathcal{B}'$,

$$L(A) \leq_{\mathcal{B}} B' \text{ iff } A \leq_{\mathcal{A}} R(B') \qquad \qquad \begin{array}{c} \mathcal{H}' & \sqcap & \mathcal{H}' \\ \leq_{\mathcal{A}} \uparrow & \intercal & \uparrow \\ \mathcal{H} & & \downarrow \\ \mathcal{H} & & \mathcal{H} \\ \mathcal{$$

If $\mathcal{A}' = \mathcal{A}, \, \preccurlyeq_{\mathcal{A}} = \leq_{\mathcal{A}}, \, \mathcal{B}' = \mathcal{B}$ and $\preccurlyeq_{\mathcal{B}} = \leq_{\mathcal{B}}$, we recover the usual definition of adjunction. Heterogeneous adjunctions also appear in the theory of Chu spaces.

Full polarization (1/2)

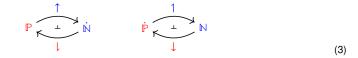


Fully polarized LG-algebras FP.LG

(Heterogeneous) operations and their residuals (we consider the Lambek fragment for brevity):

$$\otimes: \stackrel{\circ}{\mathbb{P}} \times \stackrel{\circ}{\mathbb{P}} \to \stackrel{\circ}{\mathbb{P}} \setminus : \stackrel{\circ}{\mathbb{P}}^{\partial} \times \stackrel{\circ}{\mathbb{N}} \to \mathbb{N} \qquad (2)$$
$$\stackrel{\circ}{\mathbb{Q}} \stackrel{\circ}{\leq} \stackrel{\circ}{\mathbb{P}} \setminus \stackrel{\circ}{\mathbb{N}} \quad \text{iff} \quad \stackrel{\circ}{\mathbb{P}} \otimes \stackrel{\circ}{\mathbb{Q}} \stackrel{\circ}{\leq} \stackrel{\circ}{\mathbb{N}} / \stackrel{\circ}{\mathbb{Q}}$$

Shifts:



For all $P \in \mathbb{P}$ and $N \in \mathbb{N}, \leq \mathbf{P} \times \dot{\mathbb{P}}, \leq \mathbb{P} \times \mathbb{N}$ and $\mathbf{F} = \dot{\mathbb{P}} \times \mathbb{N}$ are s.t.:

 $\uparrow P \cdot \leq^{-} N \quad \text{iff} \quad P \leq N \quad \text{iff} \quad P \leq \cdot^{+} \downarrow N \tag{4}$

i.e. \leq is the weakening relation represented by the heterogeneous adjunction $\uparrow \downarrow_{=}^{\leq^+} \downarrow$.

Collage posets: $(\mathring{\mathbb{P}}, \stackrel{\leq}{\leq}^+) := (\mathbb{P} \sqcup \stackrel{\circ}{\mathbb{P}}, \stackrel{\leq}{\leq}^+ \sqcup \stackrel{\leq}{\leq}^+ \sqcup \stackrel{\leq}{\leq}^+), (\mathring{\mathbb{N}}, \stackrel{\leq}{\leq}^-) := (\mathbb{N} \sqcup \stackrel{\circ}{\mathbb{N}}, \stackrel{\leq}{\leq}^- \sqcup \stackrel{\leq}{\leq}^-).$ Collage weakening relation: $\stackrel{\circ}{\leq} := \stackrel{\leq}{\leq} \sqcup \stackrel{\leq}{\leq} \sqcup \stackrel{\leq}{\leq} \stackrel{\circ}{\mathbb{P}} \times \mathring{\mathbb{N}}.$ Notation: $\mathring{P} \in \{P, \dot{P}\}$, resp. $\mathring{N} \in \{N, \dot{N}\}$.

Well-formed sequents (sequents in grey cells are not derivable):

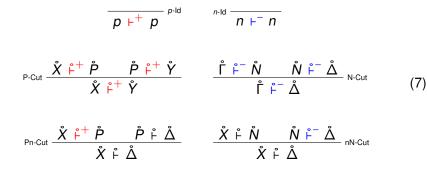
Positive sequents	X ⊦ ⁺ Y	Χ́ ₊+ Υ	X ⊦.+ Ý	X +++ Y
Negative sequents	Δ ⊦⁻ Γ	<u></u> ∆ .⊦_ L	Δ⊦₋Γ΄	<u></u> Δ΄ -⊦-⁻ Γ΄
Neutral sequents	$X \vdash \Delta$	X + Δ	X⊦·∆́	Χ́ ⋅⊦∙ Δ́

Each consequence relation is interpreted by a W.R. as follows:

t	++	+.+	·+·+	F.	·+=	·+·-	F	۰⊦	÷۰	₽ ⁺	Ê [−]	ĥ	(6)
t ^{₽₽.LG}	≤+	≤.+	·≤·+	≤-	·≤ ⁻	≤	Υ	·≤	≤∙	_≤+	°	۰×۱	(0)

18/29

(5)



$${}^{\otimes_L} \frac{\mathring{P} \hat{\otimes} \mathring{Q} \stackrel{`}{\vdash} \mathring{\Delta}}{\mathring{P} \otimes \mathring{Q} \stackrel{`}{\vdash} \mathring{\Delta}} \frac{\mathring{X} \stackrel{`}{\vdash} \stackrel{*}{P} \stackrel{`}{Y} \stackrel{`}{\vdash} \stackrel{*}{Q}}{\mathring{X} \hat{\otimes} \mathring{Y} \stackrel{`}{\vdash} \stackrel{`}{P} \otimes \mathring{Q}} {}^{\otimes_R}$$

$$\sum_{k} \frac{\mathring{X} \stackrel{h}{\models}^{+} \mathring{P}}{\mathring{P} \setminus \mathring{N} \stackrel{h}{\models}^{-} X \setminus \Delta} = \frac{\mathring{X} \stackrel{h}{\models} \mathring{P} \setminus \mathring{N}}{\mathring{X} \stackrel{h}{\models} \mathring{P} \setminus \mathring{N}} \sum_{k} \sum_{k}$$

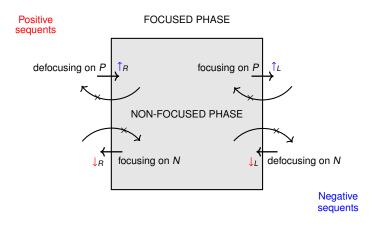
$$\downarrow_{L} \frac{N \vdash^{-} \Delta}{\downarrow N \vdash^{+} \downarrow \Delta} \frac{\mathring{X} \stackrel{+}{\vdash} \downarrow N}{\mathring{X} \stackrel{+}{\vdash} \downarrow N} \downarrow_{R} \qquad \uparrow_{L} \frac{\widehat{\uparrow} P \stackrel{+}{\vdash} \stackrel{\Delta}{\Delta}}{\uparrow P \stackrel{+}{\vdash} \stackrel{\Delta}{\Delta}} \frac{X \vdash^{+} P}{\widehat{\uparrow} X \vdash^{-} \uparrow P} \uparrow_{R}$$
(8)

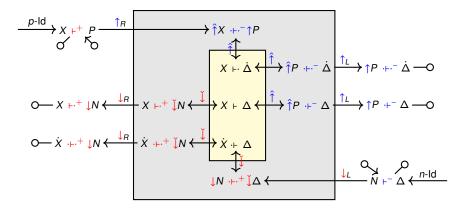
Structural rules

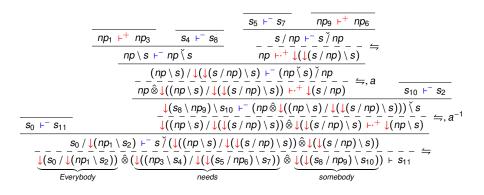
Display postulates:

Structural rules for shifts:

$$\frac{\mathring{X} \mathring{\vdash} \Delta}{\mathring{X} \mathring{\vdash}^{+} \mathring{\downarrow} \Delta} \stackrel{\uparrow}{=} \hat{\uparrow} \frac{X \mathring{\vdash} \mathring{\Delta}}{\widehat{\uparrow} X \mathring{\vdash}^{-} \mathring{\Delta}}$$
(10)







What we did.

Lambek-Grishin logic:

- (heterogenous multi-type) focused display calculus fD.LG (canonical cut-elimination and strong focalization)
- fully polarized algebraic semantics FP.LG (semantic proof of completeness of focusing)

Future work.

- We expect that the approach extends to every displayable logic. We conjecture that:
 - Every displayable logic L can be endowed with a focalized (heterogenous multi-type) display calculus **fD.L** and a fully polarized algebraic semantics **FP.L** (where **fD.L** is complete w.r.t. **FP.L**).
 - Every focalized (heterogenous multi-type) display calculus enjoys a (i) canonical semantic proof of completeness of focusing AND a (ii) canonical syntactic proof of completeness of focusing.
- We expect that the approach can be lifted at the level of <u>categories</u> (using profunctors instead of weakening relations) providing a fully-fledged semantics of proofs for a given displayable logic L.

Appendix A. Mutations

Let C be a heterogeneous multi-type calculus and let Q be the set of types of C.

 $\mathcal{S}_{\mathcal{F}} \text{ (resp. } \mathcal{S}_{\mathcal{G}} \text{) is the set of structural } \mathcal{F}\text{-connectives (resp. } \mathcal{G}\text{-connectives)}, \\ \mathcal{S} = \mathcal{S}_{\mathcal{F}} \cup \mathcal{S}_{\mathcal{G}}.$

 $\ensuremath{\mathcal{T}}$ is the set of turnstiles.

We call sort of H, sort(H) $\in Q^n$, the *n*-tuple of types that the connective takes as input.

We call **sort of** *t*, sort(*t*) $\in Q^2$, the pair of types that *t* connects.

Definition

The mutation relation of $C, \mu_C \subseteq S \times S$, is an equivalence relation between structural connectives s.t.:

if $H\mu_C H'$ then $H \in S_{\mathcal{F}}$ if and only if $H' \in S_{\mathcal{F}}$;

2 if $H\mu_C H'$ then H and H' have the same arity;

```
3 if H\mu_C H' and sort(H) = sort(H') then H = H'.
```

Informally, the mutation relation describes into which structural connectives the structural connective *H* can be mutated.

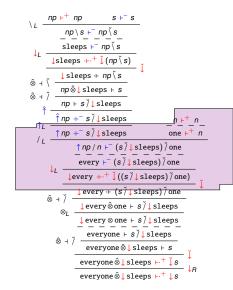
We can extend the relation to (not necessarily well typed) structures recursively on the generation tree of a structure: We say that $\Phi\mu\Psi$ if the generation trees of Φ and Ψ are identical modulo μ .

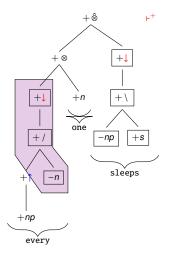
By condition 3 of Definition 4, given a structure Φ , there exists at most one well typed structure in $\mu[\Phi] = \{\Psi \mid \Phi\mu\Psi\}$, which we denote with $\mu(\Phi)$.

Appendix A. Cut-elimination

Appendix B. Focused phases and maximal PIA subtrees

The purple area is the proof-section including all the tonicity rules used to build the PIA subtree of everyone:





[And92]	Jean-Marc Andreoli. Logic programming with focusing proofs in linear logic. Journal of Logic and Computation, 2(3):297–347, 1992.	[GJL
[And01]	Jean-Marc Andreoli. Focussing and proof construction. Annals of Pure and Applied Logic, 107(1):131 – 163, 2001.	[Mil0
[Bas12]	Arno Bastenhof. Polarized Montagovian semantics for the Lambek-Grishin calculus. In P. de Groote and MJ. Nederhof, editors, <i>Formal Grammar</i> , volume 7395 of Lecture Notes in <i>Computer Science</i> . Springer, Berlin, Heidelberg, 2012.	[MM
[Ben73]	Jean Benabou. Les distributeurs: d'après le cours de questions spéciales de mathématique. Rapport n. 33 du Séminaire de Mathématique Pure. Institut de mathématique pure et appliquée, Université Catholique de Louvain, 1973.	[Moc

[GJL⊤18]	Giuseppe Greco, Peter Jipsen, Fei Liang, Alessandra Palmigiano, and Apostolos Tzimoulis. Algebraic proof theory for LE-logics, 2018.
[Mil04]	Dale Miller. An Overview of Linear Logic Programming, page 119–150. London Mathematical Society Lecture Note Series. Cambridge University Press, 2004.

MM12] Michael Moortgat and Richard Moot. Proof nets and the categorial flow of information.

In A. Baltag, D. Grossi, A. Marcoci, B. Rodenhäuser, and S. Smets, editors, *Logic and Interactive RAtionality. Yearbook 2011*, pages 270–302. ILLC, University of Amsterdam, 2012.

[Moo09] Michael Moortgat.

Symmetric categorial grammar. Journal of Philosophical Logic, 38(6):681-710, 2009.

29 / 29