Neurosymbolic Proof Search for Linguistics

Konstantinos Kogkalidis
Utrecht Institute of Linguistics OTS, Utrecht University

Logic for the AI Spring
September 2022, Como

A composition calculus for
vector-based semantic modelling
with a localization for Dutch

NWO 360-89-070, 2017-2022

The Holy Trinity

Language	**Logic**	**Computation**
grammars	substructural logics	λ-calculi
empirical linguistic rules	logical inference rules	computation steps
grammaticality	provability	type inhabitation
sentence	proof	program

Lexicon: a look-up table from words to *types* and *meanings*

The **type-logical** perspective:

- focus on *surface* syntax
- use LC/NL as core logic (ordered & linear types)
- $\Diamond$, $\Box$ modalities for structural control
- applicative derivational semantics via proof/term morphisms

The Holy Trinity

Language	Logic	Computation
grammars	substructural logics	λ-calculi
empirical linguistic rules	logical inference rules	computation steps
grammaticality	provability	type inhabitation
sentence	proof	program

Lexicon: a look-up table from words to *types* and *meanings*

The **type-logical** perspective:

- focus on *surface* syntax
- use LC/NL as core logic (ordered & linear types)
- $\diamondsuit$, $\Box$ modalities for structural control
- applicative derivational semantics via proof/term morphisms

The Holy Trinity

Language	Logic	Computation
grammars	substructural logics	λ-calculi
empirical linguistic rules	logical inference rules	computation steps
grammaticality	provability	type inhabitation
sentence	proof	program

Lexicon: a look-up table from words to *types* and *meanings*

The **type-logical** perspective:

- focus on *surface* syntax
- use LC/NL as core logic (ordered & linear types)
- $\diamondsuit$, $\Box$ modalities for structural control
- applicative derivational semantics via proof/term morphisms

The Holy Trinity

Language	Logic	Computation
grammars	substructural logics	λ-calculi
empirical linguistic rules	logical inference rules	computation steps
grammaticality	provability	type inhabitation
sentence	proof	program

Lexicon: a look-up table from words to *types* and *meanings*

A **heretical** perspective:

- focus on *deep* (discontinuous) syntax
- use ILL as core logic (linear types)
- $\diamondsuit$, $\Box$ modalities for dependency annotation
- extra semantic dimension: adjunct/complement distinction

Types

$\text{ILL}_{\multimap}$ plus $\diamondsuit$, $\Box$ modalities for *dependency domain demarcation*.

$\mathbb{T}$ inductively defined as:

$$\mathbb{T} := A \mid T \multimap T \mid \diamondsuit^d T \mid \Box^d T \qquad A \in \mathbb{A}, T \in \mathbb{T}$$

- $\mathbb{A}$ – closed set of base types
- $\multimap$ – linear function builder
- $\diamondsuit$ – reserved for "necessary arguments", i.e. complements
- $\Box$ – reserved for "optional functors", i.e. adjuncts

Types

ILL$_{\multimap}$ plus $\diamondsuit, \Box$ modalities for *dependency domain demarcation*.

$\mathbb{T}$ inductively defined as:

$$\mathbb{T} := A \mid T \multimap T \mid \diamondsuit^d T \mid \Box^d T \qquad A \in \mathbb{A}, T \in \mathbb{T}$$

- $\mathbb{A}$ – closed set of base types
- $\multimap$ – linear function builder
- $\diamondsuit$ – reserved for "necessary arguments", i.e. complements
- $\Box$ – reserved for "optional functors", i.e. adjuncts

Types

ILL$_{\multimap}$ plus $\Diamond$, $\Box$ modalities for *dependency domain demarcation*.

$\mathbb{T}$ inductively defined as:

$$\mathbb{T} := A \mid T \multimap T \mid \Diamond^{d} T \mid \Box^{d} T \qquad A \in \mathbb{A}, T \in \mathbb{T}$$

- $\mathbb{A}$ – closed set of base types
- $\multimap$ – linear function builder
- $\Diamond$ – reserved for "necessary arguments", i.e. complements
- $\Box$ – reserved for "optional functors", i.e. adjuncts

Types

ILL$_{\multimap}$ plus $\diamond$, $\Box$ modalities for *dependency domain demarcation*.

$\mathbb{T}$ inductively defined as:

$$\mathbb{T} := A \mid T \multimap T \mid \diamond^d T \mid \Box^d T \qquad A \in \mathbb{A}, T \in \mathbb{T}$$

- $\mathbb{A}$ – closed set of base types
- $\multimap$ – linear function builder
- $\diamond$ – reserved for ”necessary arguments”, i.e. complements
- $\Box$ – reserved for ”optional functors”, i.e. adjuncts

Terms

$$\frac{}{\mathsf{c} : T \vdash \mathsf{c} : T}\ Lex$$

$$\frac{\Gamma \vdash \mathsf{s} : T_1 \multimap T_2 \quad \Delta \vdash \mathsf{t} : T_1}{\Gamma, \Delta \vdash \mathsf{s}\ \mathsf{t} : T_2}\ \multimap E$$

$$\frac{\Gamma \vdash \mathsf{t} : T}{\langle \Gamma \rangle^d \vdash \vartriangle^d\ \mathsf{t} : \Diamond^d T}\ \Diamond^d I$$

$$\frac{\Gamma \vdash \mathsf{s} : \Box^d T}{\langle \Gamma \rangle^d \vdash \blacktriangledown^d \mathsf{s} : T}\ \Box^d E$$

$$\frac{}{\mathsf{x} : T \vdash \mathsf{x} : T}\ Ax$$

$$\frac{\Gamma, \mathsf{x} : T_1 \vdash \mathsf{s} : T_2}{\Gamma \vdash \lambda \mathsf{x}.\mathsf{s} : T_1 \multimap T_2}\ \multimap I$$

$$\frac{\langle \Gamma \rangle^d \vdash \mathsf{s} : T}{\Gamma \vdash \blacktriangle^d \mathsf{s} : \Box^d T}\ \Box^d I$$

$$\frac{\Gamma[\langle \mathsf{x} : T_1 \rangle^d] \vdash \mathsf{t} : T_2 \qquad \Delta \vdash \mathsf{s} : \Diamond^d T_1}{\Gamma[\Delta] \vdash \mathsf{t}[\mathsf{x} \mapsto \triangledown^d \mathsf{s}] : T_2}\ \Diamond^d E$$

$$\frac{\langle \Gamma, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}}, \Delta \rangle^d \vdash \mathsf{s} : T_2}{\langle \Gamma, \Delta \rangle^d, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}} \vdash \mathsf{s} : T_2}\ \mathsf{X}$$

Terms

$$\frac{}{\mathsf{c} : T \vdash \mathsf{c} : T}\ Lex \qquad \frac{\Gamma \vdash \mathsf{s} : T_1 \multimap T_2 \quad \Delta \vdash \mathsf{t} : T_1}{\Gamma, \Delta \vdash \mathsf{s\ t} : T_2}\ \multimap E$$

$$\frac{\Gamma \vdash \mathsf{t} : T}{\langle \Gamma \rangle^d \vdash_{\vartriangle^d} \mathsf{t} : \Diamond^d T}\ \Diamond^d I \qquad \frac{\Gamma \vdash \mathsf{s} : \Box^d T}{\langle \Gamma \rangle^d \vdash \blacktriangledown^d \mathsf{s} : T}\ \Box^d E$$

$$\frac{}{\mathsf{x} : T \vdash \mathsf{x} : T}\ Ax \qquad \frac{\Gamma, \mathsf{x} : T_1 \vdash \mathsf{s} : T_2}{\Gamma \vdash \lambda \mathsf{x}.\mathsf{s} : T_1 \multimap T_2}\ \multimap I$$

$$\frac{\langle \Gamma \rangle^d \vdash \mathsf{s} : T}{\Gamma \vdash \blacktriangle^d \mathsf{s} : \Box^d T}\ \Box^d I \qquad \frac{\Gamma[\langle \mathsf{x} : T_1 \rangle^d] \vdash \mathsf{t} : T_2 \qquad \Delta \vdash \mathsf{s} : \Diamond^d T_1}{\Gamma[\Delta] \vdash \mathsf{t}[\mathsf{x} \mapsto \triangledown^d \mathsf{s}] : T_2}\ \Diamond^d E$$

$$\frac{\langle \Gamma, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}}, \Delta \rangle^d \vdash \mathsf{s} : T_2}{\langle \Gamma, \Delta \rangle^d, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}} \vdash \mathsf{s} : T_2}\ \mathsf{X}$$

Terms

$$\frac{}{\mathsf{c} : T \vdash \mathsf{c} : T}\ Lex \qquad \frac{\Gamma \vdash \mathsf{s} : T_1 \multimap T_2 \quad \Delta \vdash \mathsf{t} : T_1}{\Gamma, \Delta \vdash \mathsf{s}\ \mathsf{t} : T_2}\ \multimap E$$

$$\frac{\Gamma \vdash \mathsf{t} : T}{\langle \Gamma \rangle^d \vdash_{\triangle^d} \mathsf{t} : \Diamond^d T}\ \Diamond^d I \qquad \frac{\Gamma \vdash \mathsf{s} : \Box^d T}{\langle \Gamma \rangle^d \vdash \blacktriangledown^d \mathsf{s} : T}\ \Box^d E$$

$$\frac{}{\mathsf{x} : T \vdash \mathsf{x} : T}\ Ax \qquad \frac{\Gamma, \mathsf{x} : T_1 \vdash \mathsf{s} : T_2}{\Gamma \vdash \lambda \mathsf{x}.\mathsf{s} : T_1 \multimap T_2}\ \multimap I$$

$$\frac{\langle \Gamma \rangle^d \vdash \mathsf{s} : T}{\Gamma \vdash \blacktriangle^d \mathsf{s} : \Box^d T}\ \Box^d I \qquad \frac{\Gamma[\langle \mathsf{x} : T_1 \rangle^d] \vdash \mathsf{t} : T_2 \qquad \Delta \vdash \mathsf{s} : \Diamond^d T_1}{\Gamma[\Delta] \vdash \mathsf{t}[\mathsf{x} \mapsto \triangledown^d \mathsf{s}] : T_2}\ \Diamond^d E$$

$$\frac{\langle \Gamma, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}}, \Delta \rangle^d \vdash \mathsf{s} : T_2}{\langle \Gamma, \Delta \rangle^d, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}} \vdash \mathsf{s} : T_2}\ \mathsf{X}$$

Terms

$$\frac{}{\mathsf{c} : T \vdash \mathsf{c} : T}\ Lex \qquad \frac{\Gamma \vdash \mathsf{s} : T_1 \multimap T_2 \quad \Delta \vdash \mathsf{t} : T_1}{\Gamma, \Delta \vdash \mathsf{s}\ \mathsf{t} : T_2}\ \multimap E$$

$$\frac{\Gamma \vdash \mathsf{t} : T}{\langle \Gamma \rangle^d \vdash_{\triangle^d} \mathsf{t} : \Diamond^d T}\ \Diamond^d I \qquad \frac{\Gamma \vdash \mathsf{s} : \Box^d T}{\langle \Gamma \rangle^d \vdash \blacktriangledown^d \mathsf{s} : T}\ \Box^d E$$

$$\frac{}{\mathsf{x} : T \vdash \mathsf{x} : T}\ Ax \qquad \frac{\Gamma, \mathsf{x} : T_1 \vdash \mathsf{s} : T_2}{\Gamma \vdash \lambda \mathsf{x}.\mathsf{s} : T_1 \multimap T_2}\ \multimap I$$

$$\frac{\langle \Gamma \rangle^d \vdash \mathsf{s} : T}{\Gamma \vdash \blacktriangle^d \mathsf{s} : \Box^d T}\ \Box^d I \qquad \frac{\Gamma[\langle \mathsf{x} : T_1 \rangle^d] \vdash \mathsf{t} : T_2 \qquad \Delta \vdash \mathsf{s} : \Diamond^d T_1}{\Gamma[\Delta] \vdash \mathsf{t}[\mathsf{x} \mapsto \triangledown^d \mathsf{s}] : T_2}\ \Diamond^d E$$

$$\frac{\langle \Gamma, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}}, \Delta \rangle^d \vdash \mathsf{s} : T_2}{\langle \Gamma, \Delta \rangle^d, \langle \mathsf{x} : T_1 \rangle^{\mathsf{X}} \vdash \mathsf{s} : T_2}\ \mathsf{X}$$

Example

"Meloni e Salvini in Piazzale Loreto"

$$\dfrac{\dfrac{\dfrac{\mathsf{c}_3 : \mathsf{in}}{\mathsf{c}_3 \vdash \Diamond^{obj} np \multimap \Box^{mod}(np \multimap np)} \quad \dfrac{\dfrac{\dfrac{\dfrac{\mathsf{c}_5 : \mathsf{Loreto}}{\mathsf{c}_5 \vdash \Box^{app}(np \multimap np)}}{\langle \mathsf{c}_5 \rangle^{app} \vdash np \multimap np}\,\Box E \quad \dfrac{\mathsf{c}_4 : \mathsf{Piazzale}}{\mathsf{c}_4 \vdash np}}{\langle \mathsf{c}_5 \rangle^{app}, \mathsf{c}_4 \vdash np}\multimap E}{\langle \langle \mathsf{c}_5 \rangle^{app}, \mathsf{c}_4 \rangle^{obj} \vdash \Diamond^{obj} np}\Diamond I}{\dfrac{\mathsf{c}_3, \langle \langle \mathsf{c}_5 \rangle^{app}, \mathsf{c}_4 \rangle^{obj} \vdash \Box^{mod}(np \multimap np)}{\langle \mathsf{c}_3, \langle \langle \mathsf{c}_5 \rangle^{app}, \mathsf{c}_4 \rangle^{obj} \rangle^{mod} \vdash np \multimap np}\Box E}\multimap E \quad \dfrac{\dfrac{\mathsf{c}_1 : \mathsf{e}}{\mathsf{c}_1 \vdash \Diamond^{cnj} np \multimap \Diamond^{cnj} np \multimap np} \quad \dfrac{\dfrac{\mathsf{c}_0 : \mathsf{Meloni}}{\mathsf{c}_0 \vdash np}}{\langle \mathsf{c}_0 \rangle^{cnj} \vdash \Diamond^{cnj} np}\Diamond I \quad \dfrac{\dfrac{\mathsf{c}_2 : \mathsf{Salvini}}{\mathsf{c}_0 \vdash np}}{\langle \mathsf{c}_2 \rangle^{cnj} \vdash \Diamond^{cnj} np}\Diamond I}{\mathsf{c}_1, \langle \mathsf{c}_0 \rangle^{cnj}, \langle \mathsf{c}_2 \rangle^{cnj} \vdash np}\multimap E}{\langle \mathsf{c}_3, \langle \langle \mathsf{c}_5 \rangle^{app}, \mathsf{c}_4 \rangle^{obj} \rangle^{mod}, \mathsf{c}_1, \langle \mathsf{c}_0 \rangle^{cnj}, \langle \mathsf{c}_2 \rangle^{cnj} \vdash np}\multimap E$$

λ read-off:

$$\blacktriangledown^{mod}(\mathsf{in}\ (\blacktriangledown^{app}(\mathsf{Loreto})\ \mathsf{Piazzale}))\ (\mathsf{e}\ \ \triangle^{cnj}(\mathsf{Meloni})\ \ \triangle^{cnj}(\mathsf{Salvini}))$$

Resource: æthel ~70 000 proofs (in Dutch)
see: abs/1912.12635

Towards Wide Coverage

Lexical Types on demand?

Word $\mapsto$ Type is not 1-1 $\Longrightarrow$ how to choose right type?
Corpus $\subset$ language $\Longrightarrow$ new words?
Inductive type universe $\Longrightarrow$ new types?

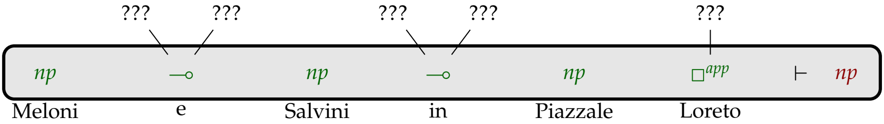
Proof Search?

Non-directional types $\Longrightarrow$ all permutations must be considered
Hypothetical reasoning $\Longrightarrow$ Prawitz problem

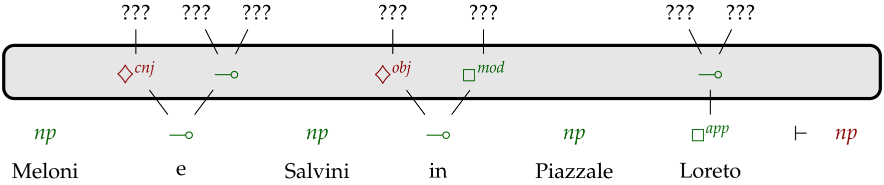
Contextual Dynamic Types

???	???	???	???	???	???	$\vdash$???
Meloni	e	Salvini	in	Piazzale	Loreto	

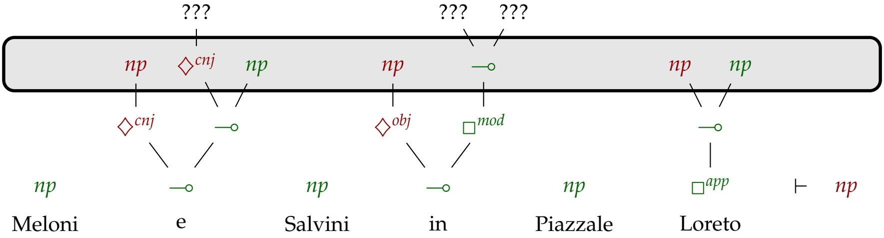
Contextual Dynamic Types



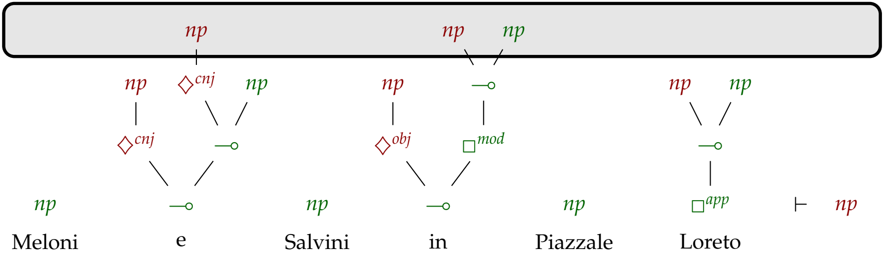
Contextual Dynamic Types



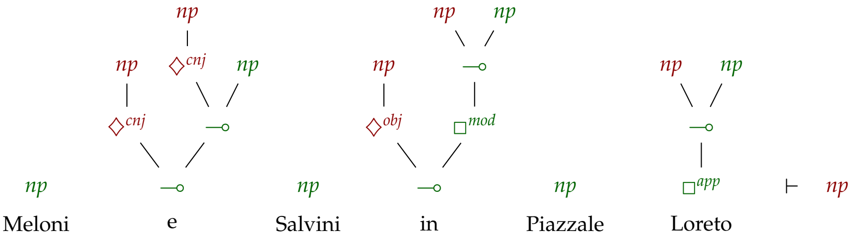
Contextual Dynamic Types



Contextual Dynamic Types



Contextual Dynamic Types

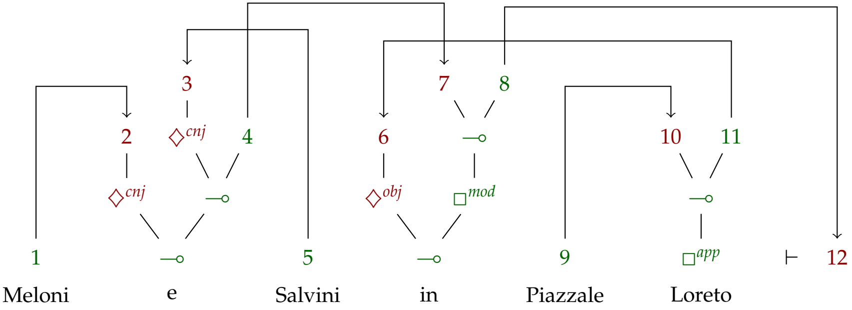


gray box := heterogenous graph convolution kernel
see: abs/2203.12235

Proofs as Graphs

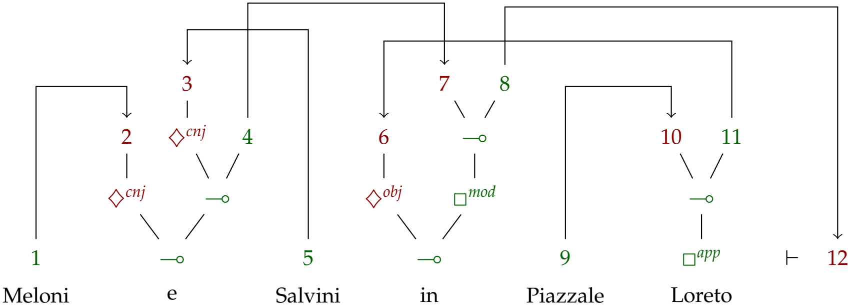
From N.D. to Proof Nets

- proof structure : tree → graph
- proof construction : depth induction → parallel assignment
- proof search : backward/forward chaining → factorial combinatorics



Proofs as Neural Graphs

The problem
Find bijection $\{p : A^+\} \longleftrightarrow \{n : A^-\}\ \forall A \in \mathbb{A}$ in sentence
here $\{1 \mapsto 2,\ 4 \mapsto 7,\ 5 \mapsto 3,\ 8 \mapsto 12,\ 9 \mapsto 10,\ 11 \mapsto 6\}$



Proofs as Neural Graphs

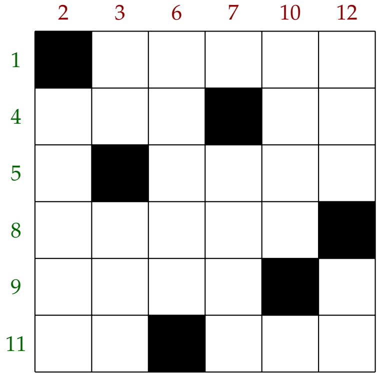
The problem

Find bijection $\{p : A^+\} \longleftrightarrow \{n : A^-\}\ \forall\ A \in \mathbb{A}$ in sentence

here $\{1 \mapsto 2,\ 4 \mapsto 7,\ 5 \mapsto 3,\ 8 \mapsto 12,\ 9 \mapsto 10,\ 11 \mapsto 6\}$

The solution

compute & normalize $A^+ \times A^-$, train against underlying permutation matrix



see: abs/2009.12702

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?
- adaptation to other languages/frameworks?

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?
- adaptation to other languages/frameworks?
- thank audience

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?
- adaptation to other languages/frameworks?
- thank audience

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?
- adaptation to other languages/frameworks?
- thank audience

Final words

done

- a logic for dependency control
- a corpus of linguistically grounded theorems
- a hyper-efficient neurosymbolic parser

todo

- derivational ambiguity?
- neural integration of proof verification?
- neural proof semantics?
- adaptation to other languages/frameworks?
- thank audience