Lambek Grammars and a New Look to Context-Free Grammars

Oleg Kiselyov

Tohoku University, Japan

AiDL
May 11, 2022

Outline

- **AB and Context-Free Grammars**
  - Grammars and Context-Free Grammars
  - AB Grammars
  - AB vs CFG
  - Dissatisfaction with AB
- Lambek Calculus and Grammars
  - Lambek Calculus
  - Lambek Grammar
- Lambek Grammars and CFG
- L Algebraically

Outline

- **AB and Context-Free Grammars**
  - Grammars and Context-Free Grammars
  - AB Grammars
  - AB vs CFG
  - Dissatisfaction with AB
- Lambek Calculus and Grammars
  - Lambek Calculus
  - Lambek Grammar
- Lambek Grammars and CFG
- L Algebraically

Grammars and Languages

Language

a set of strings (often called *sentences*), which are finite sequences of *words*

Grammar

a way to define, describe, delineate the sentences of a language: to tell which sentences belong to the language, or being well-formed

BNF

$$e ::= \mathsf{1} \mid \mathsf{x} \mid e + e \mid -e$$

Grammars and Languages

CFG: set of *productions*

$$
\begin{array}{lcl}
S & \rightarrow & \mathsf{1} \\
S & \rightarrow & \mathsf{x} \\
S & \rightarrow & S + S \\
S & \rightarrow & - S
\end{array}
$$

Grammars and Languages

CFG: set of *productions*

$$
\begin{array}{rcl}
S & \rightarrow & \mathsf{1} \\
S & \rightarrow & \mathsf{x} \\
S & \rightarrow & S + S \\
S & \rightarrow & -S
\end{array}
$$

A language (set of sentences) is specified by *generating* it

- Post System (Emil Post, 1921)
- Generative Grammar
  Noam Chomsky: The logical structure of linguistic theory, 1956

Grammars and Languages

CFG in CNF

$$
\begin{array}{rcl}
S & \rightarrow & \mathsf{1} \\
S & \rightarrow & \mathsf{x} \\
S & \rightarrow & S\ A \\
A & \rightarrow & P\ S \\
P & \rightarrow & + \\
S & \rightarrow & M\ S \\
M & \rightarrow & -
\end{array}
$$

A language (set of sentences) is specified by *generating* it

- Post System (Emil Post, 1921)
- Generative Grammar
  Noam Chomsky: The logical structure of linguistic theory, 1956

Outline

The problem of syntactic connection

> *Among these [important problems of logic] problems that of syntactic connection is of the greatest importance for logic. It is concerned with the specification of the conditions under which a word pattern constituted of meaningful words, forms an expression which itself has a unified meaning (constituted, to be sure, by the meaning of the single words belonging to it). A word pattern of this kind is called syntactically connected.*

Kazimierz Ajdukiewicz. Die syntaktische Konnexität. Studia Philosophica, 1935.

Fractions

Specification problem

How to specify that - x, - - x, etc. all belong to our language but x x does not?

Fractions

Specification problem

How to specify that - x, - - x, etc. all belong to our language but x x does not?

Index of an sentence (fragment)

- Assign to each word an index, which is a fraction

  x: 7 -: $\frac{7}{7}$

- An index of a string is the product of the indices of their constituents

x: 7	- x: 7	- - x: 7
x x: 7×7		- -: $\frac{7}{7} \times \frac{7}{7}$

Fractions

Specification problem

How to specify that - x, - - x, etc. all belong to our language but x x does not?

Index of an sentence (fragment)

- Assign to each word an index, which is a fraction

  x: 7   -: $\frac{7}{7}$

- An index of a string is the product of the indices of their constituents

  x: 7   - x: 7   - - x: 7

  x x: 7×7 - -: $\frac{7}{7} \times \frac{7}{7}$

But what about x + x vs + x x?

Non-commutative fractions

Non-commutative multiplication

$$A \times B \neq B \times A$$

Non-commutative (directional) fractions

- $A\backslash B$: A under B
- B/A: B over A

Cancellation laws

$$\begin{array}{ccc} A & \times\ A\backslash B & = B \\ B/A & \times\ A & = B \end{array}$$

Example: matrices

Grammar with directional fractional indices

Word index assignment

x: 7 1: 7 -: $7/7$ +: $(7\backslash 7)/7$

Sample sentences and their indices

- 1	:	7
1 -	:	$7 \times (7/7)$
x - 1	:	7×7
x + 1	:	7

Grammar with directional fractional indices

Word index assignment

x: 7 1: 7 -: $7/7$ +: $(7\backslash 7)/7$

Sample sentences and their indices

- 1	:	7
1 -	:	$7 \times (7/7)$
x - 1	:	7×7
x + 1	:	7

Yehoshua Bar-Hillel. A quasi arithmetical notation for syntactic description. Language, 1953.

Grammar with directional fractional indices

Word index assignment

x: s 1: s -: s/s +: $(s\backslash s)/s$

Sample sentences and their indices

- 1	:	s
1 -	:	not s
x - 1	:	not s
x + 1	:	s

Yehoshua Bar-Hillel. A quasi arithmetical notation for syntactic description. Language, 1953.

AB Grammars

Indices (Categories, Types)

Primitive types	P	::=	$s, n, np, \ldots$
Syntactic Types	A, B	::=	$P \mid A\backslash B \mid B/A$

Lexicon

Assignment of types to individual words, e.g.: $\mathsf{1}{:}s$

(Residualization, Reduction, 'Multiplication') Rules

$$\frac{u : B/A \quad v : A}{uv : B}\,/e \qquad \frac{u : A \quad v : A\backslash B}{uv : B}\,\backslash e$$

A sequence of words $w_1 w_2 \ldots w_n$ is a sentence of the language of the grammar iff its type is s

Outline

Lexicalization

AB	CFG
Lexicon	Rules

Lexicon

$$
\begin{aligned}
\mathsf{x}:&\quad s\\
\mathsf{1}:&\quad s\\
\mathsf{-}:&\quad s/s\\
+:&\quad (s\backslash s)/s
\end{aligned}
$$

Two multiplication rules

Rules

$$
\begin{aligned}
S &\rightarrow \mathsf{1}\\
S &\rightarrow \mathsf{x}\\
S &\rightarrow S\ A\\
A &\rightarrow P\ S\\
P &\rightarrow +\\
S &\rightarrow M\ S\\
M &\rightarrow -
\end{aligned}
$$

Lexicalization

AB	CFG
Lexicon	Rules

AB — Lexicon

$$\begin{array}{rl} \texttt{x}: & s \\ \texttt{1}: & s \\ \texttt{-}: & s/s \\ +: & (s\backslash s)/s \end{array}$$

Two multiplication rules

A language is defined by lexicon

CFG — Rules

$$\begin{array}{rcl} S & \rightarrow & \texttt{1} \\ S & \rightarrow & \texttt{x} \\ S & \rightarrow & S\ A \\ A & \rightarrow & P\ S \\ P & \rightarrow & + \\ S & \rightarrow & M\ S \\ M & \rightarrow & - \end{array}$$

A language is defined by rules

Lexicalization

AB	CFG
Lexicon	Rules

AB — Lexicon

$$\begin{array}{rl} \mathsf{x}: & s \\ \mathsf{1}: & s \\ \texttt{-}: & s/s \\ +: & (s\backslash s)/s \end{array}$$

Two multiplication rules

A language is defined by lexicon

Lexicalized

CFG — Rules

$$\begin{array}{rcl} S & \rightarrow & \mathsf{1} \\ S & \rightarrow & \mathsf{x} \\ S & \rightarrow & S\ A \\ A & \rightarrow & P\ S \\ P & \rightarrow & + \\ S & \rightarrow & M\ S \\ M & \rightarrow & - \end{array}$$

A language is defined by rules

Phrase-Structure

Lexicalization

AB	CFG
Lexicon	Rules
$\mathsf{x}: \quad s$ $\mathsf{1}: \quad s$ $\mathsf{-}: \quad s/s$ $+: \quad (s\backslash s)/s$	$S \rightarrow \mathsf{1}$ $S \rightarrow \mathsf{x}$ $S \rightarrow S\ A$ $A \rightarrow P\ S$ $P \rightarrow +$ $S \rightarrow M\ S$ $M \rightarrow -$
Two multiplication rules	
A language is defined by lexicon	A language is defined by rules
Lexicalized	Phrase-Structure

The two shown grammars define the same language:
they are weakly equivalent

AB and CFG

AB	CFG
$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}}{\mathsf{x}+-\mathsf{1} : s}$$	$$\begin{array}{rcl} S & \to & \mathsf{1} \\ S & \to & \mathsf{x} \\ S & \to & S\ A \\ A & \to & P\ S \\ P & \to & + \\ S & \to & M\ S \\ M & \to & - \end{array}$$

AB and CFG

AB	CFG

$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{- \mathsf{1} : s}\; S \rightarrow M S}{+ - \mathsf{1} : s\backslash s}}{\mathsf{x} + - \mathsf{1} : s}$$

$$\begin{aligned}
S &\rightarrow \mathsf{1} \\
S &\rightarrow \mathsf{x} \\
S &\rightarrow S\ A \\
A &\rightarrow P\ S \\
P &\rightarrow + \\
S &\rightarrow M\ S \\
M &\rightarrow -
\end{aligned}$$

AB and CFG

AB	CFG

AB

$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}\ A \to PS}{\mathsf{x}+-\mathsf{1} : s}$$

CFG

$$\begin{array}{rcl} S & \to & \mathsf{1} \\ S & \to & \mathsf{x} \\ S & \to & S\ A \\ \color{red}{A} & \to & \color{red}{P\ S} \\ P & \to & + \\ S & \to & M\ S \\ M & \to & - \end{array}$$

AB and CFG

AB	CFG
$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}}{\mathsf{x}+-\mathsf{1} : s}\ S \to SA$$	$S \to \mathsf{1}$ $S \to \mathsf{x}$ $S \to S\ A$ $A \to P\ S$ $P \to +$ $S \to M\ S$ $M \to -$

AB and CFG

AB

$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{- \mathsf{1} : s}}{+ - \mathsf{1} : s\backslash s}}{\mathsf{x} + - \mathsf{1} : s}$$

CFG

$$\begin{array}{rcl} S & \rightarrow & \mathsf{1} \\ S & \rightarrow & \mathsf{x} \\ S & \rightarrow & S\ A \\ A & \rightarrow & P\ S \\ P & \rightarrow & + \\ S & \rightarrow & M\ S \\ M & \rightarrow & - \end{array}$$

- Our AB and CNF grammars have the same derivation trees for any given sentence

AB and CFG

AB	CFG
$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}}{\mathsf{x}+-\mathsf{1} : s}$$	$$\begin{array}{rcl} S & \to & \mathsf{1} \\ S & \to & \mathsf{x} \\ S & \to & S\ A \\ A & \to & P\ S \\ P & \to & + \\ S & \to & M\ S \\ M & \to & - \end{array}$$

- Our AB and CNF grammars have the same derivation trees for any given sentence
- Parsing as Deduction

Grammar Equivalence

Two grammars are weakly equivalent if they define the same language

Example: For each context-free G there exists a CFG in CNF that is weakly equivalent to G

Grammar Equivalence

Two grammars are weakly equivalent if they define the same language

Example: For each context-free G there exists a CFG in CNF that is weakly equivalent to G

Two grammars are *strongly* equivalent if they produce the isomorphic parse trees

Example: For every AB grammar there exists a strongly equivalent CFG in CNF

AB are Context-Free

AB	CFG
$$\dfrac{\mathsf{x} : s \qquad \dfrac{+ : (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}}{\mathsf{x}+-\mathsf{1} : s}$$	$$\begin{array}{rcl} S & \to & \mathsf{1} \\ S & \to & \mathsf{x} \\ S & \to & S\ A \\ A & \to & P\ S \\ P & \to & + \\ S & \to & M\ S \\ M & \to & - \end{array}$$
Lexical items	Terminals
Types and *Subtypes*	Non-terminals
Instances of derivation rules	Productions

AB are Context-Free

AB	CFG
$$\dfrac{\mathsf{x} : s \qquad \dfrac{+: (s\backslash s)/s \qquad \dfrac{- : s/s \quad \mathsf{1} : s}{-\mathsf{1} : s}}{+-\mathsf{1} : s\backslash s}}{\mathsf{x}+-\mathsf{1} : s}$$	$$\begin{array}{rcl} S & \to & \mathsf{1} \\ S & \to & \mathsf{x} \\ S & \to & S\ A \\ A & \to & P\ S \\ P & \to & + \\ S & \to & M\ S \\ M & \to & - \end{array}$$
Lexical items	Terminals
Types and *Subtypes*	Non-terminals
Instances of derivation rules	Productions
The number of instances is finite	

AB and CFG

Every AB grammar is strongly equivalent to a CFG in CNF

Every ϵ-free CFG is weakly equivalent to an AB grammar

Every ϵ-free CFG in Greibach normal form is strongly equivalent to an AB grammar

Yehoshua Bar-Hillel, Chaim Gaifman, and Eli Shamir. On categorial and phrase-structure grammars. Bulletin of the research council of Israel, 1963

Outline

Dissatisfaction with AB: Algebraic

A tempting free semigroup model for AB

AB	Model
Primitive type	Set of strings
$A\backslash B$	Right semigroup action $\cdot B$
B/A	Left semigroup action $B\cdot$
Multiplication	Action

However,

$$C/B \times B/A = C/A$$

is derivable in the model but not in AB

Dissatisfaction with AB: Logical

$$\frac{u : B/A \quad v : A}{uv : B} /e \qquad \frac{u : A \quad v : A\backslash B}{uv : B} \backslash e$$

- The AB rules can be interpreted as elimination rules
  Where are the introduction rules?
- $A\backslash B$ and B/A can be interpreted as directional implications, and the AB rules as modus ponens
  But implications compose (and modus ponens can be `cut`)

Dissatisfaction with AB: Linguistic

John: np

$\times$

$$\frac{\text{likes} : (np\backslash s)/np \quad \text{cooking} : np}{\text{likes cooking} : np\backslash s}$$

$\times$

and: $((np\backslash s)\backslash(np\backslash s))/(np\backslash s)$

$\times$

$$\frac{\text{hates} : (np\backslash s)/np \quad \text{cleaning} : np}{\text{hates cleaning} : np\backslash s}$$

$=$

John likes cooking and hates cleaning: s

Dissatisfaction with AB: Linguistic

John likes and Jane hates cooking: s

Dissatisfaction with AB: Linguistic

$$\frac{\text{John} : np \quad \text{likes} : np\backslash(s/np)}{\text{John likes} : s/np}$$

$$\times$$

and: $((s/np)\backslash(s/np))/(s/np)$

$$\times$$

$$\frac{\text{Jane} : np \quad \text{hates} : np\backslash(s/np)}{\text{Jane hates} : s/np}$$

$$\times$$

cooking: np

$$=$$

John likes and Jane hates cooking: s

Outline

Outline

Lambek Calculus $\mathbf{L}$

Primitive types	P	$::=$	$s, n, np, \ldots$
Types	A, B, C	$::=$	$P \mid A\backslash B \mid B/A$
Environments	Γ, Δ	$::=$	$A_1, \ldots, A_n \quad n > 0$
Judgements	$\Gamma \vdash A$		

$$\frac{\Delta \vdash B/A \quad \Gamma \vdash A}{\Delta, \Gamma \vdash B}\, /e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}\, /i$$

$$\frac{\Gamma \vdash A \quad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\, \backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B}\, \backslash i$$

$$\frac{}{A \vdash A}\, Var$$

Natural Deduction presentation in Gentzen style

Lambek Calculus $\mathbf{L}$

Primitive types	P	$::=$	$s, n, np, \ldots$
Types	A, B, C	$::=$	$P \mid A\backslash B \mid B/A$
Environments	Γ, Δ	$::=$	$A_1, \ldots, A_n \quad n > 0$
Judgements	$\Gamma \vdash A$		

$$\frac{\Delta \vdash B/A \quad \Gamma \vdash A}{\Delta, \Gamma \vdash B} /e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A} /i$$

$$\frac{\Gamma \vdash A \quad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B} \backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B} \backslash i$$

$$\frac{}{A \vdash A} Var$$

These are *all* rules: there is no weakening or exchange

Sample derivation

$$
\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A \qquad \color{red}{\Gamma \vdash A\backslash(B/C)}}{A, \Gamma \vdash B/C}\,\backslash e \qquad C \vdash C}{A, \Gamma, C \vdash B}\,/e}{\Gamma, C \vdash (A\backslash B)}\,\backslash i}{\Gamma \vdash (A\backslash B)/C}\,/i
$$

Sample derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A \qquad \color{red}{\Gamma \vdash A\backslash(B/C)}}{A, \Gamma \vdash B/C}\ \backslash e \qquad C \vdash C}{A, \Gamma, C \vdash B}\ /e}{\Gamma, C \vdash (A\backslash B)}\ \backslash i}{\Gamma \vdash (A\backslash B)/C}\ /i$$

Likewise,

$$\Gamma \vdash C/B \;\text{ and }\; \Delta \vdash B/A \;\text{ derives }\; \Gamma, \Delta \vdash C/A$$

and

$$\Gamma \vdash A \;\text{ derives }\; \Gamma \vdash (B/A)\backslash B$$

More complicated derivation

Lの導出木の例

$$\frac{\Delta \vdash B/A \qquad \Gamma \vdash A}{\Delta, \Gamma \vdash B}\ /e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}\ /i$$

$$\frac{\Gamma \vdash A \qquad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\ \backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B}\ \backslash i$$

$rel = (n \backslash n) / (s / np)$

$$\frac{(np\backslash s)/np \vdash (np\backslash s)/np \qquad np \vdash np}{(np\backslash s)/np, np \vdash np\backslash s}\ /e$$

$$\frac{np \vdash np \qquad (np\backslash s)/np, np \vdash np\backslash s}{np, (np\backslash s)/np, np \vdash s}\ \backslash e$$

$$\frac{np, (np\backslash s)/np, np \vdash s}{np, (np\backslash s)/np \vdash s/np}\ /i$$

$$\frac{rel \vdash rel \qquad np, (np\backslash s)/np \vdash s/np}{rel, np, (np\backslash s)/np \vdash n\backslash n}\ /e$$

$$\frac{n \vdash n \qquad rel, np, (np\backslash s)/np \vdash n\backslash n}{n, rel, np, (np\backslash s)/np \vdash n}\ \backslash e$$

$$\frac{np/n \vdash np/n \qquad n, rel, np, (np\backslash s)/np \vdash n}{np/n, n, rel, np, (np\backslash s)/np \vdash np}\ /e$$

$$\frac{np/n, n, rel, np, (np\backslash s)/np \vdash np \qquad np\backslash s \vdash np\backslash s}{np/n, n, rel, np, (np\backslash s)/np, np\backslash s \vdash s}\ \backslash e$$

Outline

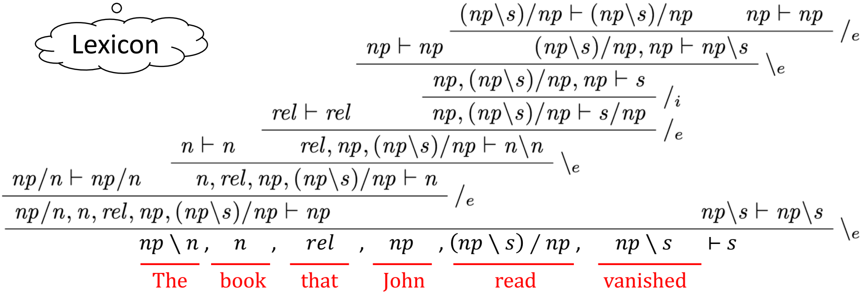
LとLambek文法LG

A_s: initial type
(文法のstart symbol)

$w_1 w_2 \dots w_n$ がLambek文法の言語 A_s に属する

$\Leftrightarrow \mathcal{L}(w_1), \mathcal{L}(w_2), \dots, \mathcal{L}(w_n) \vdash A_s$

Lexicon



$\mathbf{L}$ is the logic of resources

The environment is the sentence

$np/n,$	$n,$	$rel,$	$np,$	$(np\backslash s)/np,$	$np\backslash s$	$\vdash s$
The	book	that	John	read	vanished	

Joachim Lambek. The mathematics of sentence structure. American mathematical monthly, 1958.

Outline

LG and CFG

- LG subsumes AB, so subsumes CFG

LG and CFG

- LG subsumes AB, so subsumes CFG
- LG also supports hypothetical reasoning…

LG and CFG

- LG subsumes AB, so subsumes CFG
- LG also supports hypothetical reasoning…
- But LG is also constrained…

LG and CFG

- LG subsumes AB, so subsumes CFG
- LG also supports hypothetical reasoning...
- But LG is also constrained...

Conjecture: LG are equivalent to CFG
Noam Chomsky. Formal properties of grammars. In Handbook of Mathematical Psychology, volume 2, 1963.

Cardinality problem

Lの導出木の例

$$\frac{\Delta \vdash B/A \qquad \Gamma \vdash A}{\Delta, \Gamma \vdash B}/e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}/i$$

$$\frac{\Gamma \vdash A \qquad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B}\backslash i$$

$rel = (n \backslash n) / (s / np)$

$$\dfrac{\dfrac{np/n \vdash np/n \qquad \dfrac{n \vdash n \qquad \dfrac{rel \vdash rel \qquad \dfrac{\dfrac{np \vdash np \qquad \dfrac{(np\backslash s)/np \vdash (np\backslash s)/np \qquad np \vdash np}{(np\backslash s)/np, np \vdash np\backslash s}/e}{np, (np\backslash s)/np, np \vdash s}\backslash e}{np, (np\backslash s)/np \vdash s/np}/i}{rel, np, (np\backslash s)/np \vdash n\backslash n}/e}{n, rel, np, (np\backslash s)/np \vdash n}\backslash e}{np/n, n, rel, np, (np\backslash s)/np \vdash np}/e \qquad np\backslash s \vdash np\backslash s}{np/n, n, rel, np, (np\backslash s)/np, np\backslash s \vdash s}\backslash e$$

Cut

Sample Grammar

$$
\begin{array}{rcl}
S & \rightarrow & \mathsf{1} \\
S & \rightarrow & \mathsf{x} \\
S & \rightarrow & S + S \\
S & \rightarrow & -S
\end{array}
$$

Cut

Sample Grammar

$$\begin{aligned} S &\rightarrow \mathsf{1} \\ S &\rightarrow \mathsf{x} \\ S &\rightarrow S + S \\ S &\rightarrow -S \end{aligned}$$

Further productions (due to substitution, or cut)

$$\begin{aligned} S &\rightarrow -\mathsf{1} \\ S &\rightarrow -S + \mathsf{1} \\ S &\rightarrow -S + S + S \end{aligned}$$

Cut

Sample Grammar

$$
\begin{array}{lcl}
S & \rightarrow & \mathsf{1} \\
S & \rightarrow & \mathsf{x} \\
S & \rightarrow & S + S \\
S & \rightarrow & -S
\end{array}
$$

Further productions (due to substitution, or cut)

$$
\begin{array}{lcl}
S & \rightarrow & -\mathsf{1} \\
S & \rightarrow & -S + \mathsf{1} \\
S & \rightarrow & -S + S + S
\end{array}
$$

CFG also has arbitrary many productions

but only a finite number of *cut-free* productions

Interpolation Lemma

Let $\Gamma, \Delta, \Theta \vdash C$ where Δ is not empty be a provable judgement in L. Then there exists type I such that

1. $\Delta \vdash I$
2. $\Gamma, I, \Theta \vdash C$
3. I is 'simpler' than Δ and Γ, Θ, C

Dirk Roorda. Resource logic: proof theoretical investigations. PhD thesis, FWI, Universiteit van Amsterdam, 1991.

Thus judgements (rule instances) appearing in derivations *might* also be factored

LG are context-free

Mati Pentus. Lambek grammars are context-free. LICS, 1993.

LG are context-free

Mati Pentus. Lambek grammars are context-free. LICS, 1993.

- *Weak* equivalence of LG and CFG
- LG can be parsed in $O(n^3)$ time
- Exponential explosion in the number of productions: CFG obtained from LG are *impractical*

LG are context-free

Mati Pentus. Lambek grammars are context-free. LICS, 1993.

- *Weak* equivalence of LG and CFG
- LG can be parsed in $O(n^3)$ time
- Exponential explosion in the number of productions: CFG obtained from LG are *impractical*
- Thus, although LG are context-free in theory, they aren't in practice

Outline

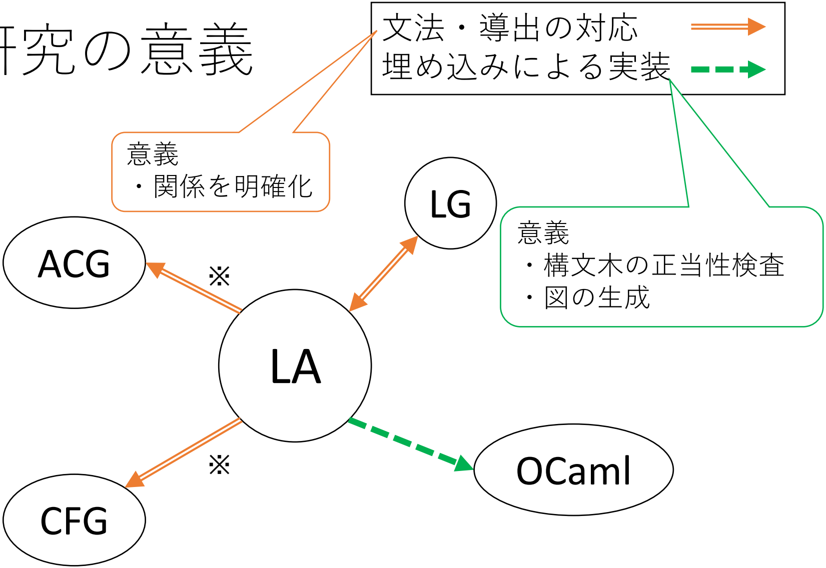
範疇文法の代数的な埋め込み手法

星野 雄哉

論文指導教員：Oleg Kiselyov 助教

東北大学大学院情報科学研究科
情報基礎科学専攻

研究の意義



※一定の制限下

計算体系LA

Primitive types	P	$::=$	$s \mid n \mid np$
Syntactic types	A, B	$::=$	$P \mid A/B \mid B\backslash A$
Environments	Γ, Δ	$::=$	$A_1, ..., A_m, \blacksquare, \blacksquare, ..., \blacksquare, B_1, ..., B_n$
Judgements	$\Gamma \vdash A$		

$$\frac{\Delta \vdash B/A \qquad \Gamma \vdash A}{\Delta, \Gamma \vdash B}\,/e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}\,/i$$

$$\frac{\Gamma \vdash A \qquad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\,\backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B}\,\backslash i$$

$$\frac{}{A \vdash A}\,\mathrm{Var}$$

■は空を表すわけではない

$$\frac{}{\blacksquare \vdash np}\,\textit{john} \qquad \frac{}{\blacksquare \vdash np/n}\,\textit{the} \qquad \frac{}{\blacksquare \vdash np}\,\textit{mary} \qquad \dots$$

LAの導出木の例

$$\frac{\Delta \vdash B/A \qquad \Gamma \vdash A}{\Delta, \Gamma \vdash B}\,/e \qquad \frac{\Gamma, A \vdash B}{\Gamma \vdash B/A}\,/i$$

$$\frac{\Gamma \vdash A \qquad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\,\backslash e \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A\backslash B}\,\backslash i$$

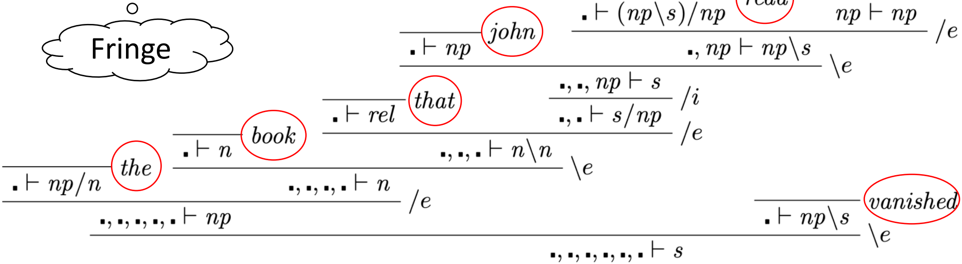
$rel = (n \backslash n) / (s / np)$

$$\frac{\dfrac{}{\blacksquare \vdash np/n}\,\textit{the} \qquad \dfrac{\dfrac{}{\blacksquare \vdash n}\,\textit{book} \qquad \dfrac{\dfrac{}{\blacksquare \vdash rel}\,\textit{that} \qquad \dfrac{\dfrac{\dfrac{}{\blacksquare \vdash np}\,\textit{john} \qquad \dfrac{\dfrac{}{\blacksquare \vdash (np\backslash s)/np}\,\textit{read} \qquad np \vdash np}{\blacksquare, np \vdash np\backslash s}\,/e}{\blacksquare, \blacksquare, np \vdash s}\,\backslash e}{\blacksquare, \blacksquare \vdash s/np}\,/i}{\blacksquare, \blacksquare, \blacksquare \vdash n\backslash n}\,/e}{\blacksquare, \blacksquare, \blacksquare, \blacksquare \vdash n}\,\backslash e}{\dfrac{\blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare \vdash np \qquad \dfrac{}{\blacksquare \vdash np\backslash s}\,\textit{vanished}}{\blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare \vdash s}\,\backslash e}\,/e$$

文法としてのLA

$w_1 w_2 \dots w_n$がLAの言語A_sに属する

$$\Leftrightarrow \mathcal{F}\left(\frac{\vdots}{\blacksquare, \blacksquare, \dots, \blacksquare \vdash A_s}\right) = w_1\, w_2 \dots w_n$$



代数としてのLA

略記		
vp	=	$np\backslash s$
tv	=	vp/np
det	=	np/n
rel	=	$(n\backslash n)/(s/np)$

シグネチャΣ_{AL}

john	:	$\langle\bullet; np\rangle$
book	:	$\langle\bullet; n\rangle$
the	:	$\langle\bullet; det\rangle$
that	:	$\langle\bullet; rel\rangle$
read	:	$\langle\bullet; tv\rangle$
vanished	:	$\langle\bullet; vp\rangle$

$\backslash e_{\langle\bullet;\bullet;np;s\rangle}$	:	$\langle\bullet; np\rangle \to \langle\bullet; vp\rangle \to \langle\bullet; s\rangle$
$\backslash e_{\langle\bullet;\bullet;n;n\rangle}$	:	$\langle\bullet; n\rangle \to \langle\bullet; pp\rangle \to \langle\bullet; n\rangle$
$/e_{\langle\bullet;\bullet;n;np\rangle}$	:	$\langle\bullet; det\rangle \to \langle\bullet; n\rangle \to \langle\bullet; np\rangle$
$/e_{\langle\bullet;\bullet;np;np\backslash s\rangle}$	:	$\langle\bullet; tv\rangle \to \langle\bullet; np\rangle \to \langle\bullet; vp\rangle$
$/e_{\langle\bullet;np;np;np\backslash s\rangle}$	:	$\langle\bullet; tv\rangle \to \langle np; np\rangle \to \langle\bullet, np; vp\rangle$
$var_{\langle np;np\rangle}$	:	$\langle np; np\rangle$
$\backslash e_{\langle\bullet;\bullet;np;s\rangle}$	:	$\langle\bullet; np\rangle \to \langle\bullet, np; vp\rangle \to \langle\bullet, np; s\rangle$
$/i_{\langle\bullet;np;s\rangle}$	:	$\langle\bullet, np; s\rangle \to \langle\bullet; s/np\rangle$
$/e_{\langle\bullet;\bullet;s/np;n\backslash n\rangle}$	:	$\langle\bullet; rel\rangle \to \langle\bullet; s/np\rangle \to \langle\bullet; pp\rangle$

⋮

項の例

$$\backslash e_{\langle\bullet;\bullet;np;s\rangle}(/e_{\langle\bullet;\bullet;n;np\rangle}(\text{the}, \backslash e_{\langle\bullet;\bullet;n;n\rangle}(\text{book}, /e_{\langle\bullet;\bullet;s/np;n\backslash n\rangle}(\text{that}, /i_{\langle\bullet;np;s\rangle}(\backslash e_{\langle\bullet;\bullet;np;s\rangle}(\text{john}, /e_{\langle\bullet;np;np;np\backslash s\rangle}(\text{read}, var_{\langle\bullet,np;np\rangle})))))), \text{vanished})$$

Conclusion

For any LG and the natural number n, there exists a CFG whose derivations are all and only LG derivations of hyp-rank n.

The LG lexicon enters CFG as is, with no duplications, let alone exponential explosions.

LG of a bounded hyp-rank are *strongly* equivalent to CFG

Real Conclusion

Lambek Grammars are Great