Residuation, Structural Rules and Context Freeness

Gerhard Jäger

jaeger@ling.uni-potsdam.de

TAG+6

Venice

May 22, 2002

Outline of talk

- generative capacity of Categorial Grammars
- **NL**
- adding modalities
- context freeness of **NL**$\Diamond$-grammars
- generalizations
- comparison to Kandulski

Generative capacity of Categorial Grammars

Basic systems

	AB	CF	Bar-Hillel *et al.* (1960)
+ residuation	**NL**	CF	Kandulski (1988)
+ associativity	**L**	CF	Pentus (1993)
+ further structural rules	**LP, LC**, ...	$\subset$ CF	van Benthem (1991)

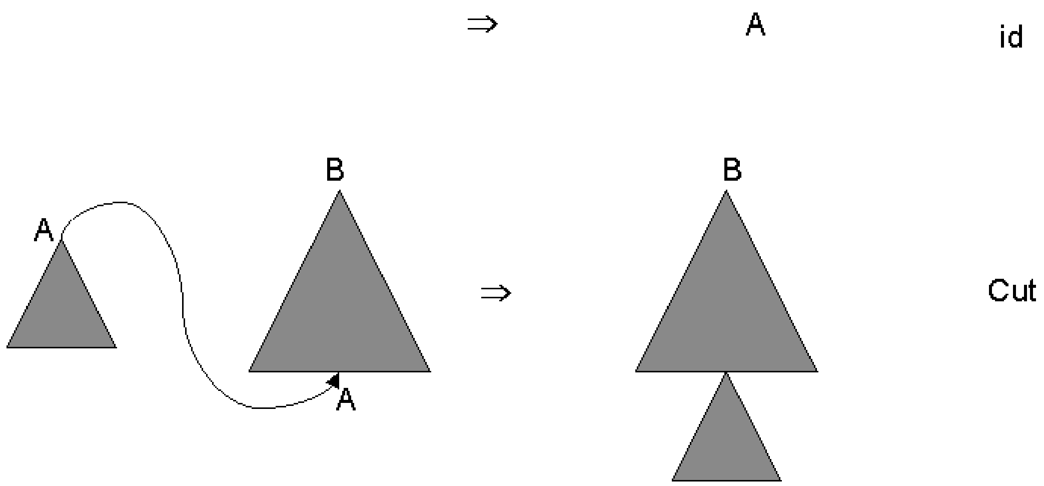
Generative capacity of Categorial Grammars

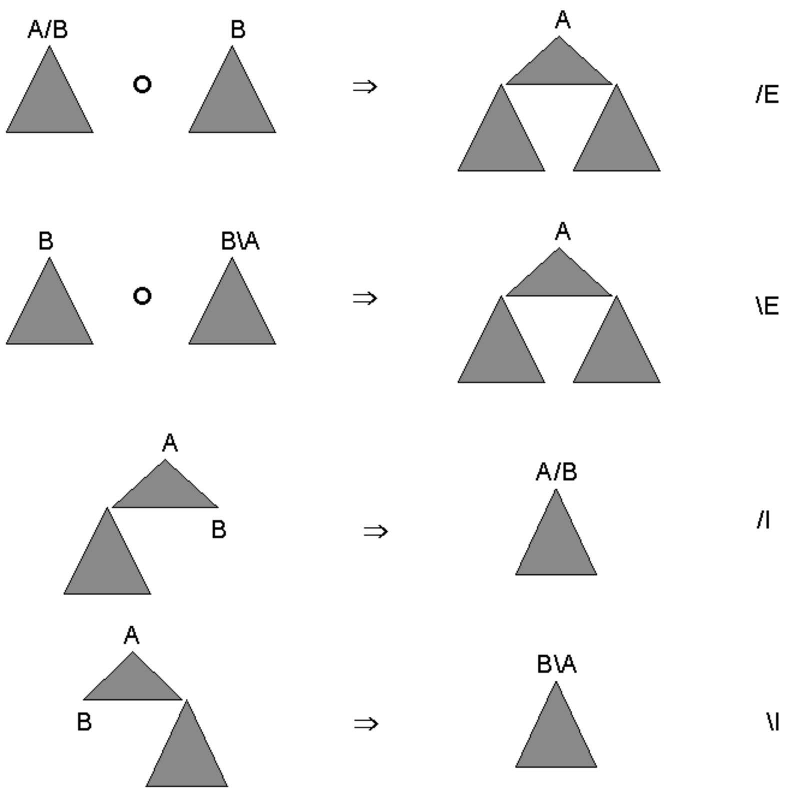
Multimodal systems

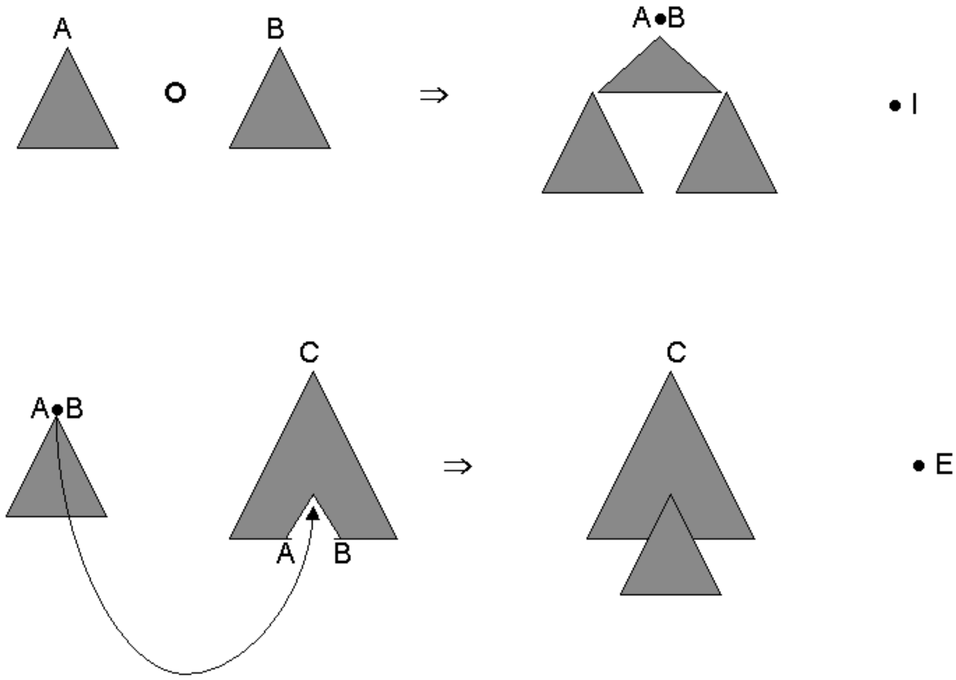
	NL $+ \Diamond_i, \Box^{\downarrow}{}_i$	CF	
+ associativity	**L** $+ \Diamond_i, \Box^{\downarrow}{}_i$	CF	Jäger (2001)
+ further structural rules	**NL** + postulates	Turing-complete	Carpenter (1999)

The non-associative Lambek Calculus

- Reasoning over binary trees
- root and leafs are labeled with formulas over $\backslash, \bullet, /$
- Natural Deduction format:

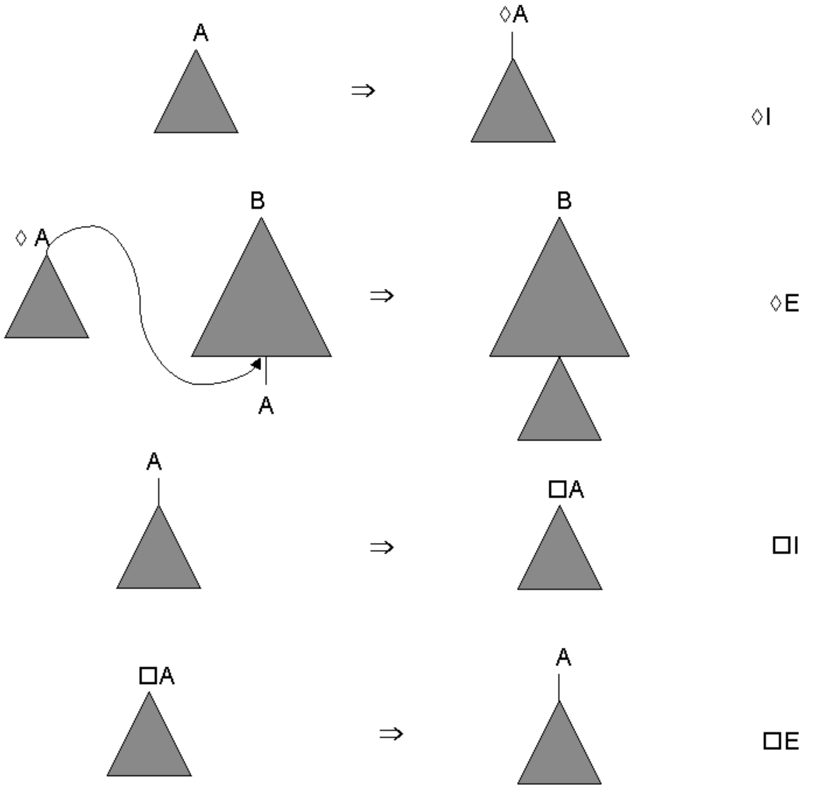






NL$\diamond$

- conservative extension of **NL**
- logical vocabulary: additionally two unary modal operators, $\diamond$ and $\Box^{\downarrow}$
- reasoning over unary/binary trees



From Logic to Grammar

- **NL**$\Diamond$-grammars consist of
  - Lexicon: finite set of elementary non-branching trees with category on top and terminal item on bottom: $\begin{array}{c} Cat \\ | \\ lex \end{array}$
  - finite set of designated categories
- A string of terminals is recognized by the grammar iff it is the yield of a tree with a designated category as root label.

Generative Capacity

Lemma 1 *Every context free language L is recognized by some* **NL**$\diamond$*-grammar.*

Proof:

- Bar-Hillel *et al.* (1960): Every context free language is recognized by some first order **AB**-grammar
- for first order grammars, the difference between **AB** and **NL**$\diamond$ doesn't matter
- basically Cohen's 1967 proof for inclusion of cfl in **L**-languages

$\dashv$

Sequent presentation of NL$\diamond$

$$\frac{}{A \Rightarrow A}\, id \qquad \frac{X \Rightarrow A \qquad Y[A] \Rightarrow B}{Y[X] \Rightarrow B}\, Cut$$

$$\frac{X[A \circ B] \Rightarrow C}{X[A \bullet B] \Rightarrow C}\, \bullet L \qquad \frac{X \Rightarrow A \qquad Y \Rightarrow B}{X \circ Y \Rightarrow A \bullet B}\, \bullet R$$

$$\frac{X \Rightarrow A \qquad Y[B] \Rightarrow C}{Y[B/A \circ X] \Rightarrow C}\, /L \qquad \frac{X \circ A \Rightarrow B}{X \Rightarrow B/A}\, /R$$

$$\frac{X \Rightarrow A \qquad Y[B] \Rightarrow C}{Y[X \circ A\backslash B] \Rightarrow C}\backslash L \qquad \frac{A \circ X \Rightarrow B}{X \Rightarrow A\backslash B}\backslash R$$

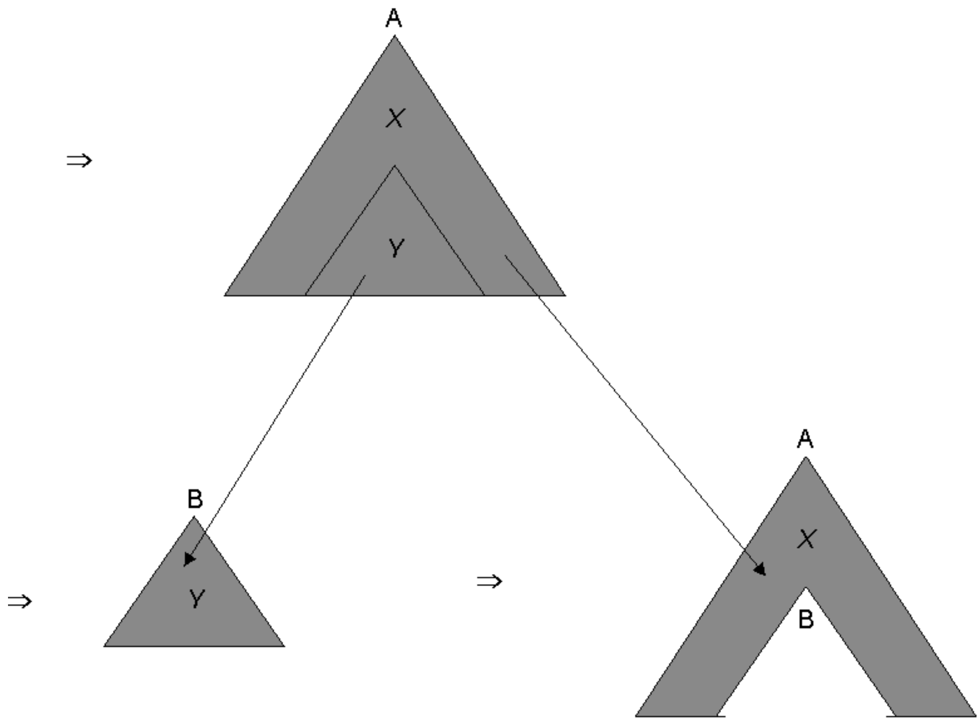
$$\frac{X \Rightarrow A}{\langle X \rangle \Rightarrow \Diamond A}\Diamond L \qquad \frac{X[\langle A \rangle] \Rightarrow B}{X[\Diamond A] \Rightarrow B}\Diamond R$$

$$\frac{X[A] \Rightarrow B}{X[\langle \Box^{\downarrow} A \rangle] \Rightarrow B}\Box^{\downarrow} R \qquad \frac{\langle X \rangle \Rightarrow A}{X \Rightarrow \Box^{\downarrow} A}\Box^{\downarrow} R$$

Lemma 2 *Let* $X[Y] \Rightarrow A$ *be a theorem of* **NL**$\Diamond$. *Then there is a type* B *such that*

1. **NL**$\Diamond \vdash Y \Rightarrow B$
2. **NL**$\Diamond \vdash X[B] \Rightarrow A$
3. *There is a type occurring in* $X[Y] \Rightarrow A$ *which contains at least as many connectives as* B.

Proof: Induction over sequent derivations. $\dashv$



Constructing a CFG from an NL◇-grammar

- **NL**◇(n): fragment of **NL**◇ where each formula contains at most n connectives
- Let n be the maximal number of connectives occurring in an **NL**◇-grammar G.
- Equivalent cfg (construction inspired by Pentus 1993):

$$\mathbf{NL}\Diamond(n) \vdash A \circ B \Rightarrow C \quad \leadsto \quad C \rightarrow A, B$$

$$\mathbf{NL}\Diamond(n) \vdash \langle A \rangle \Rightarrow B \quad \leadsto \quad B \rightarrow A$$

$$\mathbf{NL}\Diamond(n) \vdash A \Rightarrow B \quad \leadsto \quad B \rightarrow A$$

$$\begin{array}{c} Cat \\ | \\ lex \end{array} \quad \leadsto \quad Cat \rightarrow lex$$

$$A \in \mathcal{D} \quad \leadsto \quad A \rightarrow S$$

Lemma 3 *Every* **NL**$\diamond$*-recognizable language is context free.*

Proof: By lemma 2 and the construction above. $\dashv$

Theorem 1 **NL**$\diamond$*-grammars recognize exactly the context free languages.*

Proof: Immediate. $\dashv$

Generalization I: Connectives with arbitrary arity

- Both $\backslash, \bullet, /$ and $\diamond, \Box^{\downarrow}$ form families of residuated operators
- Can be generalized to products/implications of arbitrary arity
- Logic of Pure Residuation (**LPR**): Generalization of **NL** to infinite number of families of residuated connectives of any arity
- Lemma 2 holds for **LPR**

Theorem 2 **LPR***-grammars recognize exactly the context free languages.*

Generalization II: Structural Rules

- **LPR** can be strengthened by adding structural rules
- Lemma 2 remains valid if Permutation and/or Expansion are added
- Thus grammars based on these logics only recognize context free languages

$$\frac{X[Y \circ Z] \Rightarrow A}{X[Z \circ Y] \Rightarrow A}P \qquad \frac{X[Y] \Rightarrow A}{X[Y \circ Y] \Rightarrow A}E$$

- Associativity, Contraction and Weakening do not preserve lemma 2

Relation to Kandulski's work

- Kandulski (1995): Proof of context-freeness of **NLP**
- Kandulski (2002): Proof of context-freeness of **LPR**
- Differences:
  - Kandulski restricts designated categories to atomic categories
  - entirely different proof strategy: based on proof normalization in axiomatic version of **LPR**

References

Bar-Hillel, Y., Gaifman, C., and Shamir, E. (1960). On categorial and phrase structure grammars. *Bulletin of the Research Council of Israel*, **F**(9), 1–16.

Carpenter, B. (1999). The Turing-completeness of multimodal categorial grammars. Papers presented to Johan van Benthem in honor of his 50th birthday. European Summer School in Logic, Language and Information, Utrecht.

Cohen, J. M. (1967). The equivalence of two concepts of Categorial Grammar. *Information and Control*, **10**, 475–484.

Jäger, G. (2001). On the generative capacity of multimodal categorial grammars. to appear in *Journal of Language and Computation*.

Kandulski, M. (1988). The equivalence of nonassociative Lambek categorial grammars and context-free grammars. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **34**, 41–52.

Kandulski, M. (1995). On commutative and nonassociative syntactic calculi and categorial grammars. *Mathematical Logic Quarterly*, **41**, 217–135.

Kandulski, M. (2002). On generalized Ajdukiewicz and Lambek calculi and grammars. manuscript, Poznan University.

Pentus, M. (1993). Lambek grammars are context-free. In *Proceedings of the 8th Annual IEEE Symposium on Logic in Computer Science*. Montreal.

van Benthem, J. (1991). *Language in Action*. Elsevier, Amsterdam.