On the Computational Complexity of Dominance Links in Grammatical Formalisms

Sylvain Schmitz

LSV, ENS Cachan & CNRS

LIFO Seminars, June 28, 2010

Linguistic Motivations

Computational Motivations

Complexity Survey

References

```
dass Peter den Kühlschrank zu versuchen zu reparieren verspricht
that Peter_{nom} the fridge<sub>acc</sub>
                                               to repair
                                to try
                                                             promises
```

```
dass [Peter] [den Kühlschrank] zu versuchen zu reparieren verspricht
that Peter_{nom} the fridge<sub>acc</sub>
                                                  to repair
                                   to try
                                                                 promises
dass [den Kühlschrank] [Peter] zu versuchen zu reparieren verspricht
that the fridge<sub>acc</sub>
                         Peter<sub>nom</sub> to try
                                                  to repair
                                                                 promises
```

```
dass Peter den Kühlschrank zu versuchen zu reparieren verspricht
 that Peter_{nom} the fridge<sub>acc</sub>
                                                  to repair
                              to try
                                                                 promises
dass den Kühlschrank Peter [zu versuchen] [zu reparieren] verspricht
that the fridge<sub>acc</sub> Peter<sub>nom</sub> to try</sub>
                                                  to repair
                                                                   promises
dass den Kühlschrank Peter [zu reparieren] [zu versuchen] verspricht
that the fridge<sub>acc</sub> Peter<sub>nom</sub> to repair</sub>
                                                  to try
                                                                   promises
```

(Becker et al., 1991; Rambow, 1994a; Lichte, 2007, ...)

```
dass Peter den Kühlschrank zu versuchen zu reparieren verspricht
that Peter_{nom} the fridge<sub>acc</sub>
                                   to try
                                                   to repair
                                                                  promises
dass den Kühlschrank Peter zu versuchen zu reparieren verspricht
that the fridge<sub>acc</sub>
                    Peter<sub>nom</sub> to try
                                                   to repair
                                                                  promises
dass den Kühlschrank [Peter] [zu reparieren] zu versuchen verspricht
that the fridge<sub>acc</sub>
                        Peter<sub>nom</sub> to repair
                                                                   promises
                                                   to try
dass den Kühlschrank [zu reparieren] [Peter] zu versuchen verspricht
that the fridge<sub>acc</sub>
                         to repair Peter<sub>nom</sub> to try
                                                                   promises
```

* dass den Kühlschrank zu versuchen Peter zu reparieren verspricht that the fridge_{acc} to try Peter_{nom} to repair promises

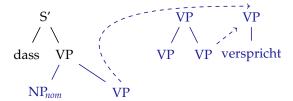
```
dass Peter den Kühlschrank zu versuchen zu reparieren verspricht
 that Peter_{nom} the fridge<sub>acc</sub>
                                   to try
                                                   to repair
                                                                  promises
 dass den Kühlschrank Peter zu versuchen zu reparieren verspricht
 that the fridge<sub>acc</sub>
                    Peter<sub>nom</sub> to try
                                                   to repair
                                                                  promises
 dass den Kühlschrank Peter zu reparieren zu versuchen verspricht
 that the fridge<sub>acc</sub>
                         Peter<sub>nom</sub> to repair
                                                   to try
                                                                  promises
 dass den Kühlschrank zu reparieren Peter zu versuchen verspricht
 that the fridge<sub>acc</sub>
                          to repair Peter<sub>nom</sub> to try
                                                                  promises
* dass den Kühlschrank zu versuchen Peter zu reparieren verspricht
  that the fridge<sub>acc</sub>
                          to try Peter<sub>nom</sub> to repair
                                                                   promises
```

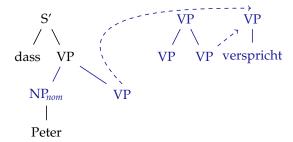
Unordered Vector Grammars with Dominance Links

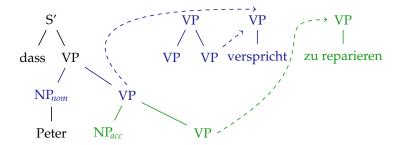
UVG-dls (Rambow, 1994a,b)

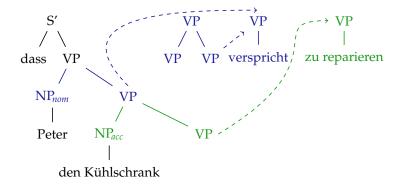
$$\left\{ \begin{array}{c|cccc} VP & & VP & & VP \\ / & & / & & / & & | \\ NP_{nom} & VP & & VP & VP & verspricht \end{array} \right\}$$

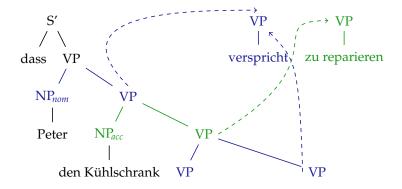


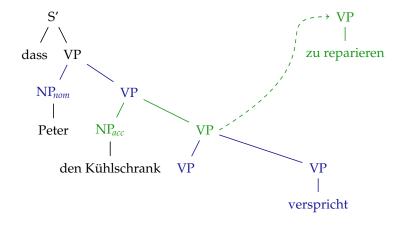


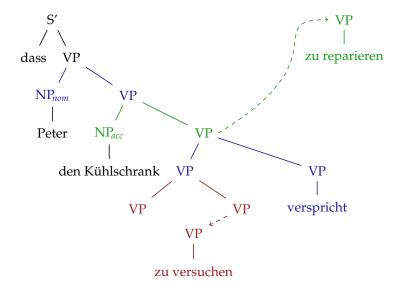


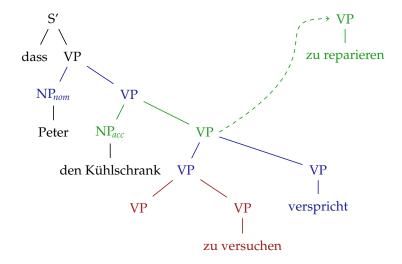


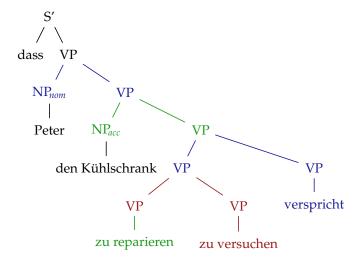












Directly Related Formalisms

- ► D-tree substitution grammars (Rambow et al., 1995, 2001),
- graph-driven adjunction grammars (Candito and Kahane, 1998),
- tree description grammars (Kallmeyer, 2001),
- interaction grammars (Guillaume and Perrier, 2010),
- vector tree adjoining grammars (Becker et al., 1991; Rambow, 1994a),

. . . .

Multiset-valued Linear Indexed Grammars

```
MLIGs (Rambow, 1994a,b)
```

Tuples $\mathcal{G} = \langle N, \Sigma, P, (S, \overline{x}_0) \rangle$:

- N: a finite set of *nonterminal* symbols,
- \triangleright Σ : a finite *terminal* alphabet, disjoint from N,
- $V = (N \times \mathbb{N}^n) \oplus \Sigma$: the vocabulary,
- ▶ P: a finite set of *productions* in $(N \times \mathbb{N}^n) \times V^*$,
- $(S, \overline{x}_0) \in \mathbb{N} \times \mathbb{N}^n$: the axiom.

Productions are more easily written as

$$(A,\overline{x}) \rightarrow \mathfrak{u}_0(B_1,\overline{x}_1)\mathfrak{u}_1\cdots\mathfrak{u}_{\mathfrak{m}}(B_{\mathfrak{m}},\overline{x}_{\mathfrak{m}})\mathfrak{u}_{\mathfrak{m}+1} \qquad (\star)$$

with each u_i in Σ^* and each (B_i, \overline{x}_i) in $N \times \mathbb{N}^n$.

Multiset-valued Linear Indexed Grammars

MLIGs (Rambow, 1994a,b)

$$(A,\overline{x}) \rightarrow \mathfrak{u}_0(B_1,\overline{x}_1)\mathfrak{u}_1 \cdots \mathfrak{u}_{\mathfrak{m}}(B_{\mathfrak{m}},\overline{x}_{\mathfrak{m}})\mathfrak{u}_{\mathfrak{m}+1} \qquad (\star)$$

The *derivation* relation $\Rightarrow \subseteq V^* \times V^*$:

$$\delta(A, \overline{y})\delta' \Rightarrow \delta u_0(B_1, \overline{y}_1)u_1 \cdots u_m(B_m, \overline{y}_m)u_{m+1}\delta'$$

for some $\delta, \delta' \in V^*$, if

- 1. $\overline{X} \leq \overline{V}$
- 2. $\forall 1 \leq i \leq m, \overline{x}_i \leq \overline{y}_i$
- 3. $\overline{y} \overline{x} = \sum_{i=1}^{m} \overline{y}_i \overline{x}_i$.

Complexity Survey

Multiset-valued Linear Indexed Grammars

MLIGs (Rambow, 1994a,b)

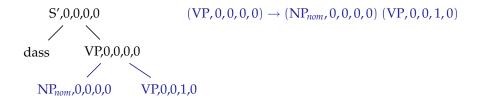
The *language* of a MLIG is the set of terminal strings derived from (S, \overline{X}_0) , i.e.

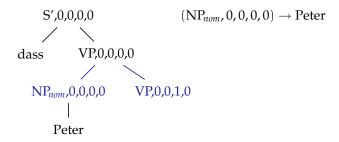
$$L(\mathfrak{G}) = \{ w \in \Sigma^* \mid (S, \overline{X}_0) \Rightarrow^* w \}$$

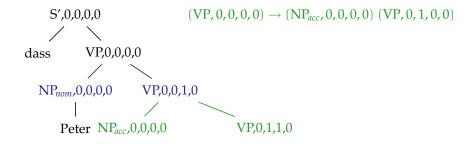
A 4-dimensional MLIG with productions

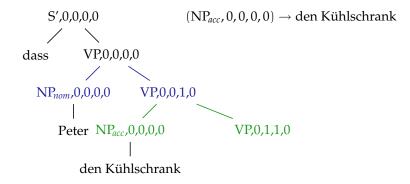
```
(S', 0, 0, 0, 0) \rightarrow dass (VP, 0, 0, 0, 0)
 (NP_{nom}, 0, 0, 0, 0) \rightarrow Peter
  (NP_{acc}, 0, 0, 0, 0) \rightarrow den Kühlschrank
     (VP.0.0.0.0) \rightarrow (VP.0.0.0.0) (VP.1.0.0.0)
  (VP, -1, 0, 0, 0) \rightarrow zu \text{ versuchen}
     (VP, 0, 0, 0, 0) \rightarrow (NP_{acc}, 0, 0, 0, 0) (VP, 0, 1, 0, 0)
  (VP, 0, -1, 0, 0) \rightarrow zu reparieren
     (VP, 0, 0, 0, 0) \rightarrow (NP_{nom}, 0, 0, 0, 0) (VP, 0, 0, 1, 0)
     (VP, 0, 0, 0, 0) \rightarrow (VP, 0, 0, 0, 0) (VP, 0, 0, 0, 1)
(VP, 0, 0, -1, -1) \rightarrow verspricht
```

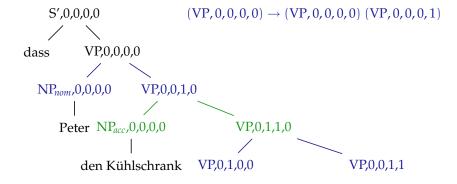
$$(S', 0, 0, 0, 0) \rightarrow dass (VP, 0, 0, 0, 0)$$

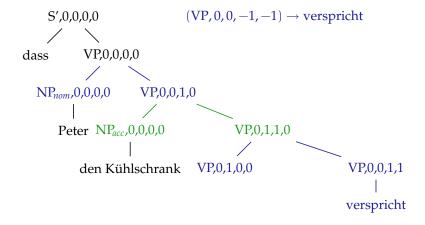


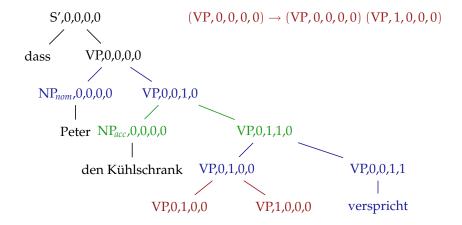


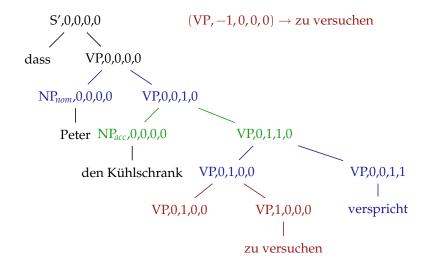


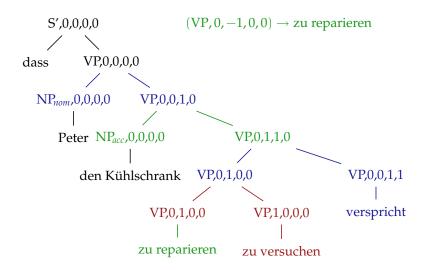












Equivalence with UVG-dls

Theorem (Rambow, 1994b)

Every MLIG can be transformed into an equivalent UVG-dl in logarithmic space, and conversely.

Proof idea: from UVG-dls.

- 1. Convert the UVG-dl to *strict* form, with vectors of productions connected by the dominance links,
- 2. each coordinate in \mathbb{N}^n encodes a dominance link.

Equivalence with UVG-dls

Proof idea: to UVG-dls.

- 1. Convert the MLIG to *ordinary form*: all coordinates in vectors hold values in {0, 1},
- 2. convert now to restricted index normal form:
 - $(A,\overline{0}) \rightarrow \alpha, \alpha \in (\Sigma \cup (N \times {\overline{0}}))^*,$
 - $(A,\overline{0}) \rightarrow (B,\overline{e}_i)$, or
 - $(A,\overline{e}_i) \rightarrow (B,\overline{0});$
- 3.
 pair in UVG-dl vectors productions of forms $(A,\overline{0}) \to (B,\overline{e}_i)$ and $(C,\overline{e}_i) \to (D,\overline{0})$;
 - ► productions of form $(A,\overline{0}) \rightarrow \alpha$ result in singleton vectors.

Related Formalisms (1)

MLIGs are *exactly* the same as

- vector addition tree automata (de Groote et al., 2004), and
- branching vector addition systems with states (Verma and Goubault-Larrecq, 2005).
- ► They are indeed a natural generalization of Petri nets/vector addition systems.

Related Formalisms (1)

MLIGs are *exactly* the same as

- vector addition tree automata (de Groote et al., 2004), and
- branching vector addition systems with states (Verma and Goubault-Larrecq, 2005).
- ► They are indeed a natural generalization of Petri nets/vector addition systems.

Related Formalisms (2)

MLIG emptiness and membership reduce to

- emptiness and membership in UVG-dls (and related formalisms),
- provability in multiplicative exponential linear logic (de Groote et al., 2004),
- emptiness and membership of abstract categorial grammars (de Groote et al., 2004; Yoshinaka and Kanazawa, 2005),

Related Formalisms (3)

MLIG emptiness and membership reduce to

- emptiness and membership of Stabler (1997)'s minimalist grammars without shortest move constraint (Salvati, 2010),
- satisfiability of first-order logic on data trees (Bojańczyk et al., 2009).
- ► Their decidability is a central open problem.

Related Formalisms (3)

MLIG emptiness and membership reduce to

- emptiness and membership of Stabler (1997)'s minimalist grammars without shortest move constraint (Salvati, 2010),
- satisfiability of first-order logic on data trees (Bojańczyk et al., 2009).
- ► Their decidability is a central open problem.

Problem	Lower bound	Upper bound
MLIG Emptiness,	2ExpSpace (Lazić,	Not known to be
MLIG Membership	2010)	decidable
kb-MLIG Emptiness,	ExpTime (this talk)	ExpTime (this talk)
kb-MLIG Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTime (Koller and	NPTIME (trivial)
Membership	Rambow, 2007)	
kr-MLIG (Emptiness,	РТіме (Jones and	PTIME (this talk)
Membership}	Laaser, 1976)	
MLIG Boundedness	2ExpTime (Demri	2ExpTime (Demri
	et al., 2009)	et al., 2009)
MLIG k-Boundedness	ExpTime (this talk)	ExpTime (this talk)

k-Ranked MLIGs

A MLIG derivation

$$\alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_p$$

is of *rank* k for some $k \ge 0$ if no vector with a *sum* of components larger than k can appear in any α_j .

A MLIG is k-ranked (noted kr-MLIG) if any derivation starting with $\alpha_0 = (S, \overline{x}_0)$ is of rank k.

k-Ranked MLIGs

Lemma

Any n-dimensional k-ranked MLIG 9 can be transformed into an equivalent CFG \mathfrak{G}' in time $O(|\mathfrak{G}| \cdot (\mathfrak{n}+1)^{k^3})$.

Proof idea:

1. Convert to extended two form:

terminal
$$(A, \overline{0}) \rightarrow \alpha$$
, $(A, \overline{0}) \rightarrow \varepsilon$, nonterminal $(A, \overline{x}) \rightarrow (B_1, \overline{x}_1)(B_2, \overline{x}_2)$, $(A, \overline{x}) \rightarrow (B_1, \overline{x}_1)$,

- 2. at most $|N| \cdot (n+1)^k$ nonterminals (A, \overline{y}) in $N' \subseteq N \times \mathbb{N}^n$ with $\sum_{i=1}^n \overline{y}(i) \leq k$,
- 3. at most $(n + 1)^{k^3}$ choices of nonterminals in N' for a production $(A, \overline{x}) \rightarrow (B_1, \overline{x}_1)(B_2, \overline{x}_2)$.

Problem	Lower bound	Upper bound
MLIG Emptiness,	2ExpSpace (Lazić,	Not known to be
MLIG Membership	2010)	decidable
kb-MLIG Emptiness,	ExpTime (this talk)	ExpTime (this talk)
kb-MLIG Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTime (Koller and	NPTIME (trivial)
Membership	Rambow, 2007)	
kr-MLIG {Emptiness,	РТіме (Jones and	РТіме (this talk)
Membership}	Laaser, 1976)	
MLIG Boundedness	2ExpTime (Demri	2ExpTime (Demri
	et al., 2009)	et al., 2009)
MLIG k-Boundedness	ExpTime (this talk)	ExpTime (this talk)

k-Bounded MLIGs

A MLIG derivation

$$\alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_p$$

is k-bounded for some $k \ge 0$ if no vector with a coordinate larger than k can appear in any α_j .

A MLIG is k-bounded (noted kb-MLIG) if any derivation starting with $\alpha_0 = (S, \overline{x}_0)$ is k-bounded.

k-Bounded MLIGs

Lemma

Any n-dimensional k-bounded MLIG 9 can be transformed into an equivalent CFG 9' in time $O(|G| \cdot (k+1)^{n^2}).$

Proof idea:

- 1. Convert to extended two form,
- 2. $N' \subset N \times \{0, ..., k\}^n, |N'| \leq |N| \cdot (k+1)^n,$
- 3. $(A, \overline{x}) \rightarrow (B_1, \overline{x}_1)(B_2, \overline{x}_2)$ • $(A, \overline{V}) \in N'$, • $0 < i \le n$.

result in $\leq k + 1$ ways to split $(\overline{y}(i) - \overline{x}(i)) \leq k$ into $\overline{y}_1(i) + \overline{y}_2(i)$ in production $(A, \overline{y}) \rightarrow (B_1, \overline{x}_1 + \overline{y}_1)(B_2, \overline{x}_2 + \overline{y}_2).$

k-Bounded MLIGs

Theorem

Emptiness and membership for k-bounded MLIGs are ExpTime-complete, even for fixed $k \ge 1$.

- lower bound: by encoding computations of alternating Turing machines running in polynomial space,
- upper bound: by the previous lemma.

Problem	Lower bound	Upper bound
MLIG Emptiness,	2ExpSpace (Lazić,	Not known to be
MLIG Membership	2010)	decidable
kb-MLIG Emptiness,	ExpТіме (this talk)	ExpТіме (this talk)
kb-MLIG Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTime (Koller and	NPTime (trivial)
Membership	Rambow, 2007)	
kr-MLIG {Emptiness,	РТіме (Jones and	РТіме (this talk)
Membership}	Laaser, 1976)	
MLIG Boundedness	2ExpTime (Demri	2ExpTime (Demri
	et al., 2009)	et al., 2009)
MLIG k-Boundedness		

Boundedness

A MLIG is

- ▶ *bounded* if \exists k s.t. it is k-bounded,
- ► *ranked* if ∃k s.t. it is k-ranked:

bounded ⇔ ranked

Theorem (Demri et al., 2009)

Boundedness for MLIGs is 2ExpTime-complete.

Problem	Lower bound	Upper bound
MLIG Emptiness,	2ExpSpace (Lazić,	Not known to be
MLIG Membership	2010)	decidable
kb-MLIG Emptiness,	ExpТіме (this talk)	ExpТіме (this talk)
kb-MLIG Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTime (Koller and	NPTime (trivial)
Membership	Rambow, 2007)	
kr-MLIG {Emptiness,	РТіме (Jones and	РТіме (this talk)
Membership}	Laaser, 1976)	
MLIG Boundedness	2ExpTime (Demri	2ExpTime (Demri
	et al., 2009)	et al., 2009)
MLIG k-Boundedness	ExpTime (this talk)	ExpTime (this talk)

k-Boundedness

Corollary

Let $k \ge 1$; k-boundedness for MLIGs is ExpTime-complete.

- ▶ lower bound: by the ExpTime-hardness of emptiness and membership in 1-bounded MLIGS,
- upper bound: by converting into a CFG.

Problem	Lower bound	Upper bound
MLIG Emptiness,	2ExpSpace (Lazić,	Not known to be
MLIG Membership	2010)	decidable
kb-MLIG Emptiness,	ExpТіме (this talk)	ExpТіме (this talk)
kb-MLIG Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTime (Koller and	NPTime (trivial)
Membership	Rambow, 2007)	
kr-MLIG {Emptiness,	РТіме (Jones and	РТіме (this talk)
Membership}	Laaser, 1976)	
MLIG Boundedness	2ExpTime (Demri	2ExpTime (Demri
	et al., 2009)	et al., 2009)
MLIG k-Boundedness	ExpТime (this talk)	ExpTime (this talk)

Lexicalized MLIGs

A terminal MLIG derivation

$$\alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_p = w$$

is c-lexicalized for some $c \ge 0$ if $p \le c \cdot |w|$.

A MLIG is *lexicalized* if there exists c such that any terminal derivation starting from (S, \overline{X}_0) is c-lexicalized.

Note: This captures lexicalization in UVG-dls.

Lexicalized MLIGs

Theorem (Koller and Rambow, 2007)

Uniform membership of $\langle \mathfrak{G}, w \rangle$ for \mathfrak{G} a 1-bounded, lexicalized, UVG-dl with finite language is NPTIME-hard, even for |w| = 1.

By a reduction from the *normal dominance graph* configurability problem (Althaus et al., 2003).

Problem	Lower bound	Upper bound
{MLIG, MLIG _ℓ }	2ExpSpace (Lazić,	Not known to be
Emptiness, MLIG	2010)	decidable
Membership		
$\{\text{kb-MLIG, kb-MLIG}_{\ell}\}$	ExpТіме (this talk)	ExpТіме (this talk)
Emptiness, kb-MLIG		
Membership		
$\{MLIG_{\ell}, kb-MLIG_{\ell}\}$	NPTIME (Koller and	NPTime (trivial)
Membership	Rambow, 2007)	
kr-MLIG {Emptiness,	РТіме (Jones and	РТіме (this talk)
Membership}	Laaser, 1976)	
{MLIG, MLIG _ℓ }	2ExpTime (Demri	2ExpTime (Demri
Boundedness	et al., 2009)	et al., 2009)
{MLIG, MLIG _ℓ }	ExpТіме (this talk)	ExpТіме (this talk)
k-Boundedness		

Concluding Remarks

About *mild context-sensitivity*:

- 1. support for *limited cross-serial dependencies*: seems doubtful, Rambow (1994a) conjectures $L_{copy} = \{ww \mid w \in \{a, b\}^*\}$ is not derivable,
- 2. *semilinearity*: does not hold, from the Petri net literature,
- 3. *polynomial recognition*: only for k-ranked grammars.

But all 3 hold e.g. for k-ranked V-TAG.

- Althaus, E., Duchier, D., Koller, A., Mehlhorn, K., Niehren, J., and Thiel, S., 2003. An efficient graph algorithm for dominance constraints. *Journal of Algorithms*, 48(1):194–219. doi:10.1016/S0196-6774(03)00050-6.
- Becker, T., Joshi, A.K., and Rambow, O., 1991. Long-distance scrambling and tree adjoining grammars. In *EACL'91*, pages 21–26. ACL Press. doi:10.3115/977180.977185.
- Bojańczyk, M., Muscholl, A., Schwentick, T., and Segoufin, L., 2009. Two-variable logic on data trees and XML reasoning. *Journal of the ACM*, 56(3):1–48. doi:10.1145/1516512.1516515.
- Candito, M.H. and Kahane, S., 1998. Defining DTG derivations to get semantic graphs. In *TAG+4*, pages 25–28.
- de Groote, P., Guillaume, B., and Salvati, S., 2004. Vector addition tree automata. In *LICS* 2004, pages 64–73. IEEE Computer Society. ISBN 0-7695-2192-4. doi:10.1109/LICS.2004.51.

- Demri, S., Jurdziński, M., Lachish, O., and Lazić, R., 2009. The covering and boundedness problems for branching vector addition systems. In Kannan, R. and Narayan Kumar, K., editors, FSTTCS 2009, volume 4 of LIPIcs, pages 181–192. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.FSTTCS.2009.2317.
- Guillaume, B. and Perrier, G., 2010. Interaction grammars. Research on Language and Computation. doi:10.1007/s11168-010-9066-x. To appear.
- Jones, N.D. and Laaser, W.T., 1976. Complete problems for deterministic polynomial time. Theor. Comput. Sci., 3(1):105–117. doi:10.1016/0304-3975(76)90068-2.
- Kallmeyer, L., 2001. Local tree description grammars. *Grammars*, 4(2): 85-137. doi:10.1023/A:1011431526022.
- Koller, A. and Rambow, O., 2007. Relating dominance formalisms. In FG 2007.
- Lazić, R., 2010. The reachability problem for branching vector addition systems requires doubly-exponential space. *Information Processing Letters*. doi:10.1016/j.ipl.2010.06.008. To appear.
- Lichte, T., 2007. An MCTAG with tuples for coherent constructions in German. In FG 2007.
- Rambow, O., 1994a. Formal and Computational Aspects of Natural Language Syntax. PhD thesis, University of Pennsylvania.

Rambow, O., 1994b. Multiset-valued linear index grammars: imposing dominance constraints on derivations. In ACL'94, pages 263–270. ACL Press. doi:10.3115/981732.981768.

Complexity Survey

- Rambow, O., Vijay-Shanker, K., and Weir, D., 1995. D-tree grammars. In ACL'95, pages 151–158. ACL Press. doi:10.3115/981658.981679.
- Rambow, O., Weir, D., and Vijay-Shanker, K., 2001. D-tree substitution grammars. Comput. Linguist., 27(1):89–121. doi:10.1162/089120101300346813.
- Salvati, S., 2010. Minimalist grammars in the light of logic. Manuscript.
- Stabler, E.P., 1997. Derivational minimalism. In Retoré, C., editor, LACL'96. volume 1328 of LNCS, pages 68–95. Springer. doi:10.1007/BFb0052147.
- Verma, K.N. and Goubault-Larrecq, J., 2005. Karp-Miller trees for a branching extension of VASS. Discrete Mathematics and Theoretical *Computer Science*, 7(1):217–230. http://www.dmtcs.org/volumes/abstracts/dm070113.abs.html.
- Yoshinaka, R. and Kanazawa, M., 2005. The complexity and generative capacity of lexicalized abstract categorial grammars. In Blache, P., Stabler, E., Busquets, J., and Moot, R., editors, LACL 2005, volume 3492 of LNCS, pages 330-346. Springer. doi:10.1007/11422532_22.