Semantic expressive capacity with bounded memory

Antoine Venant

Saarland University

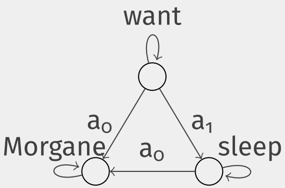
July 16, 2019

joint work with Alexander Koller

Semantic interpretation

Linguistic expression ⇒ (formal) meaning representation. Representations can be logical formulae, or graphs (AMR [Banarescu & all 2013], MRS [Copestake & all 2005]).

Morgane wants to sleep



$\text{Morgane}(x) \wedge \text{want}(e, x, \text{sleep}(e', x))$

- Consensual approach: semantic interpretation is a *compositional* process, guided by syntax.

Principle of compositionality

Statement

"The meaning of a complex expression is a **function** of the meaning of its parts and the **syntactic rule** that combines them."

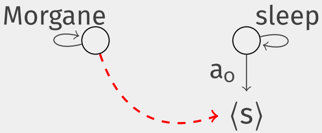
Requires:

- A syntax tree, along which semantic construction is performed in a bottom-up fashion.
- Operators for semantic composition (semantic algebra).

- Which semantic interpretation functions can we express *compositionally* using specific classes of syntax trees **and** semantic operators?

(Essentially) one job:

combine predicates with their arguments.



Two traditions

'Unification style'

Finite set of markers denoting 'holes' ($\langle s \rangle, \langle o \rangle, \langle \mathrm{mod} \rangle, \langle \mathrm{comp} \rangle$) waiting to be filled with semantic values. Markers accessible in unconstrained order [Copestake & all, 2001, Courcelles & Englefriet 2012, Groshwitz & all 2017].

'Lambda style'

Countably infinite ordered set of markers but order constrain access (variables' scope) [Montague 1977, Steedman 2001].

'Unification style'

Finite set of markers denoting 'holes' ($\langle s \rangle, \langle o \rangle, \langle \text{mod} \rangle, \langle \text{comp} \rangle$) waiting to be filled with semantic values. Markers accessible in unconstrained order [Copestake & all, 2001, Courcelles & Englefriet 2012, Groshwitz & all 2017].

→ number of 'holes' accessible at a given time of the construction process is bounded: 'bounded memory'.

'Lambda style'

Countably infinite ordered set of markers but order constrain access (variables' scope) [Montague 1977, Steedman 2001].

’bounded memory’ operators are popular

- In semantic parsing [Chiang & all 2013, Groschwitz & all 2018, Chen & all 2018].
- For the manual design of grammars [Bender 2002 *inter alia*].

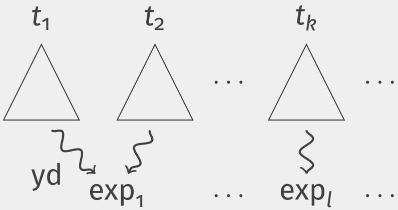
Expressive limitation due to bounded memory capacity?

- Specifically, considering long distance dependencies.
- If impossible (from distance) to combine a predicate with its argument right away → need to store argument slot until argument becomes available.

Further motivation

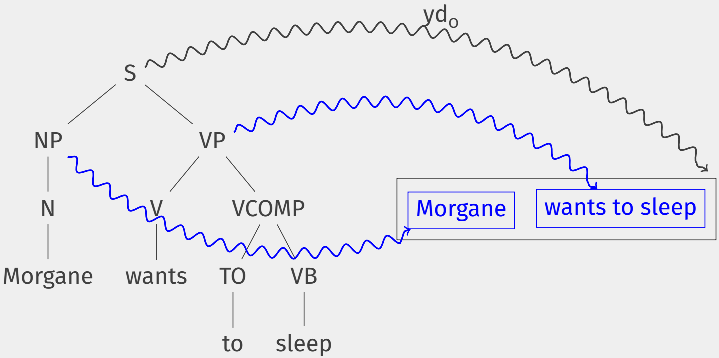
- A lot is known on expressive capacity of grammatical formalisms – in terms of languages (of words/trees).
  - *e.g.*, famous CCG/TAG/LIG [Vijay-Shanker & Weir, 1994] weak equivalence result.
- What about the joint expressivity of grammatical formalisms and specific semantic combinators in terms of *relations*?
- Do (weakly) equivalent grammatical formalisms support the same compositional interpretations?
- Inform the elaboration of semantic parsing systems

- Set of 'grammatical' syntax trees $\{t_1, t_2, \dots\}$.
- *yield* function, yd, associating each tree with its string projection (the linguistic expression for which it is a grammatical analysis).

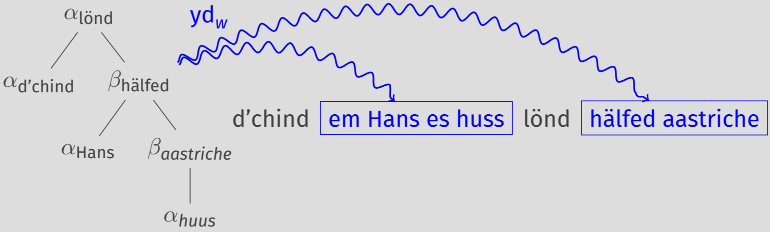


- The set $\{t_1, t_2, \dots\}$ could be given by any kind of descriptive/computing device (formal grammar, neural net,...).

- Concatenates children's yield from left to right.



Swiss-German cross-serial dependencies [Shieber 1985]



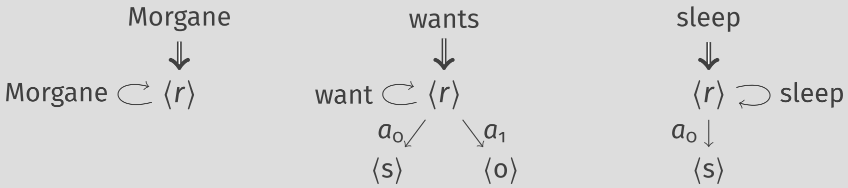
(dass)	(mer)	d' chind	em Hans	es huus	lönd
(that)	(we)	the-children-ACC	Hans-DAT	the-house-ACC	let

hälfed	aastriiche
help	paint

'(that we) let the children help Hans paint the house'

Semantic composition 1/3

Interpretation for elementary syntactic constituents

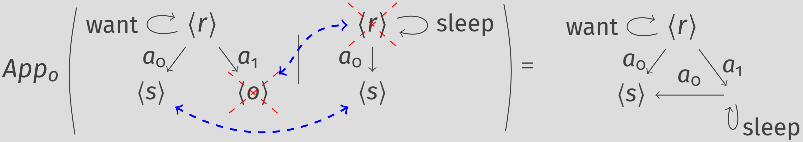


- $\langle s \rangle$, $\langle o \rangle$, $\langle r \rangle$: markers.
- $\langle s \rangle$, $\langle o \rangle$: argument placeholders ('holes'): a semantic value will eventually be substituted for them during the process of semantic composition.
- $\langle r \rangle$: root of the semantic constituant ('hook'), destined to be substituted for an argument placeholder.

Semantic composition 2/3

Semantic algebra (*i.e.* composition operators)

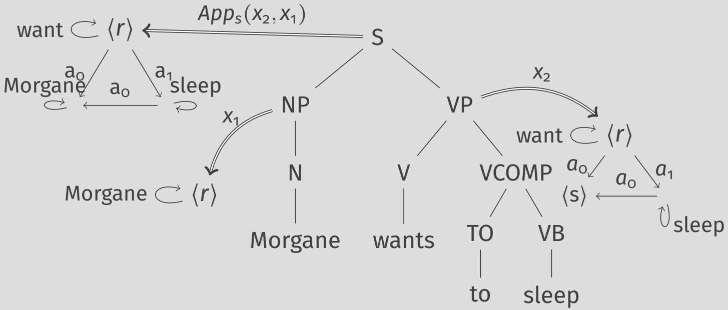
- Example with the AM algebra [Groschwitz & all 2017]



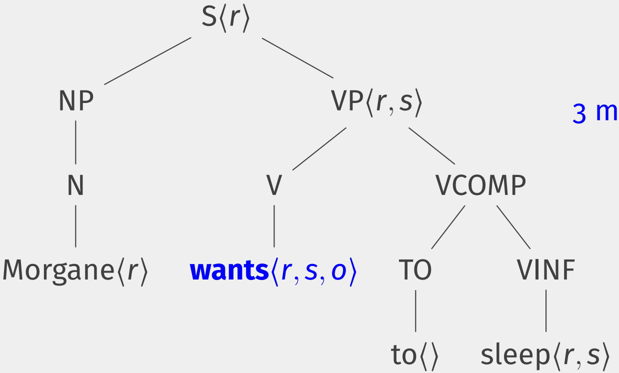
- Merge referenced marker $\langle o \rangle$ of the fonctor with the root $\langle r \rangle$ of the argument, then 'forgets' these two markers.
- Merge any other identical marker (here, $\langle s \rangle$).

Homomorphic interpretation of syntax trees

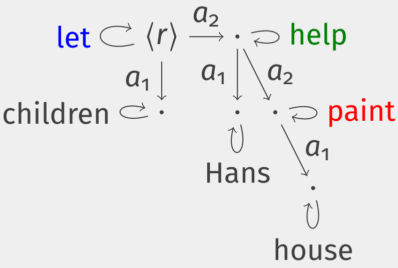
$\{VP(x_1, x_2) \rightarrow APP_o(x_1, x_2), S(x_1, x_2) \rightarrow APP_s(x_2, x_1)\}$



'Semantic' memory



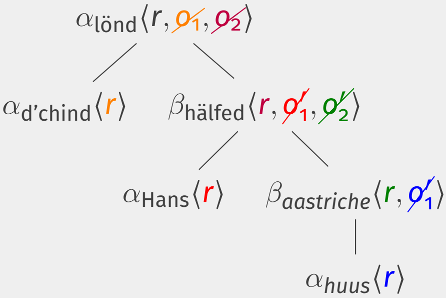
3 markers required



d'chind em Hans es huss lönd hälfed aastriche

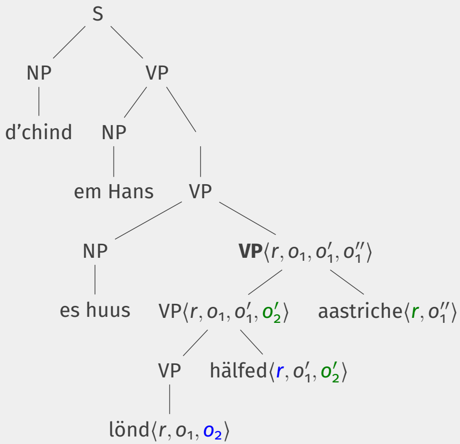
- lönd: $\langle r, o_1, o_2 \rangle$ (two objects).
- hälfed: $\langle r, o_1, o_2 \rangle$ (two objects).
- aastriche: $\langle r, o_1 \rangle$ (one object).

Non-projective analysis possible with a 3-markers capacity.



Projectivité and memory 3/3

With a projective analysis: 4 markers seem intuitively required.

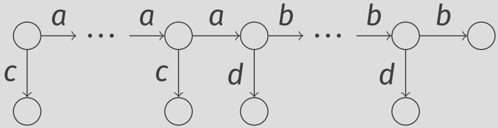


- Arbitrary long crossed-serial dependencies → infinite memory required?
- A formal relation for a mathematical proof:

CSD

Word to graph function $w \mapsto g_w$ where

- Words w are of the form: $\underbrace{a \ldots a}_{\text{n times}} \underbrace{b \ldots b}_{\text{m times}} \underbrace{c \ldots c}_{\text{n times}} \underbrace{d \ldots d}_{\text{m times}}$.
- And for each such w, g_w is:



Theorem ?

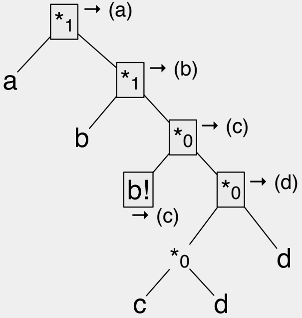
There exists no projective grammar and finite memory compositional interpretation mechanism over a projective grammar which expresses CSD.

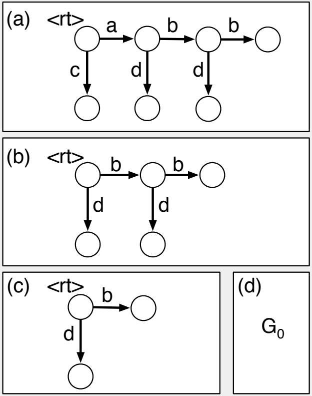
NOT A Theorem

There exists no projective grammar and finite memory compositional interpretation mechanism over a projective grammar which expresses CSD.

NOT A Theorem

There exists no projective grammar and finite memory compositional interpretation mechanism over a projective grammar which expresses CSD.





If one further impose specific alignements between elementary syntactic and semantic constituants (’*a*’ aligned with ’$\cdot \xrightarrow{a} \cdot$’, ’*b*’ aligned with ’$\cdot \xrightarrow{b} \cdot$’ ...) it can be shown:

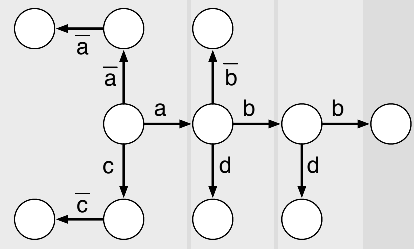
Theorem

- **there exists no projective grammar and finite memory compositional interpretation mechanism over a projective grammar** expressing CSD and respecting elementary alignments.
- There exists a **non-projective** grammar and a finite-memory compositional interpretation mechanism expressing CSD and respecting elementary alignments.
- Remark: strong assumption on alignments but no assumption on grammatical formalism.

- Whithout the alignment condition the theorem is false.
- However, weaker form of alignments can be achieved if we constrain the grammatical formalism (pumping lemma).
- Requires arbitrary complex ’arguments’ to avoid previous unnatural constructions.
- Result for Tree-Adjoining Grammars (TAG).

$\overline{\text{CSD}}$ relation

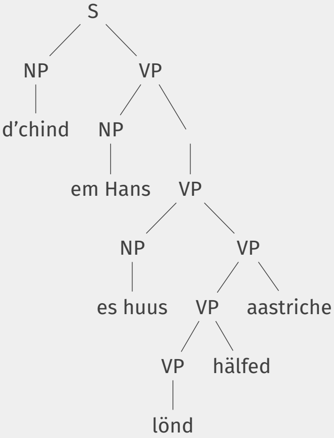
$a\overline{a}\overline{a}b\overline{b}bc\overline{c}dd$



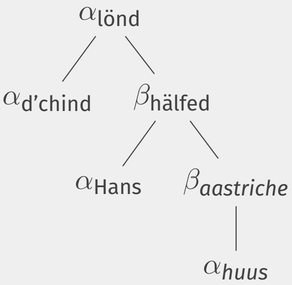
Two kinds of trees

- TAG grammars produce derivation trees and derived trees.

Derived (syntagmatic) tree

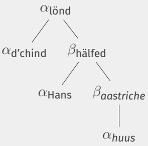


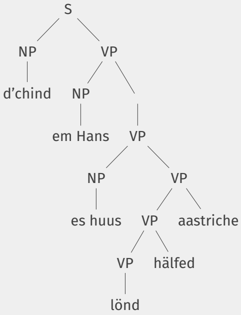
Derivation (dependency) tree



Two (weakly) equivalent grammar formalisms

- Formalism TAG: Use the **derivation trees** of some TAG grammar with a **non-projective** yield.
- Formalism PTAG: Use the **derived trees** of some TAG grammar wutg the **projective** yield.
- The two formalisms generate the same **word langages**, but not necessarily the same **relations**.





Second result

Theorem

- There exists a (non-projective) TAG grammar and a finite memory compositional interpretation mechanism expressing $\overline{\text{CSD}}$.
- There exists no (projective) PTAG grammar and finite memory compositional interpretation mechanism expressing $\overline{\text{CSD}}$.

Recap

- Theoretical result on the link between compositionality, projectivity and bounded memory capacity.
- Strong result, under strong assumption of perfect syntax/semantics aligmnents at the lexical level.
- **independent** of considered grammatical formalism.
- New light shed on the choice between derivation/derived tree as the support of semantic composition for TAG grammars.
- Do weakly equivalent grammatical formalisms support the same compositional interpretation mechanisms? → **No!**.

Conclusions and future work

- Notion of expressivity at the syntax/semantics interface.
- Theoretical study on the link between projectivity and ’semantic’ memory.
- What could we say about more restricted forms of non-projectivity? Finite increase in required memory capacity?
- Artificial non-projectivity due to imperfect aligners in semantic parsing systems.
- Locally translate from ’unification style’ to ’lambda style’ to circumvent projectivity issues?

Questions?