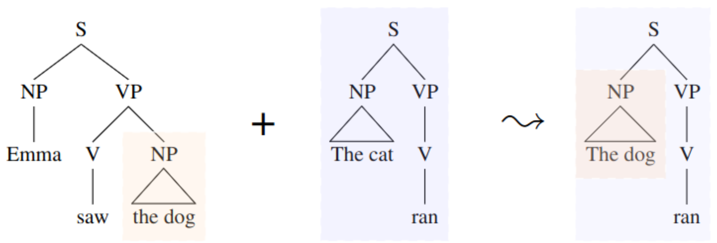
Kickoff ANR COMPO
Structural Biases
for Semantic Prediction

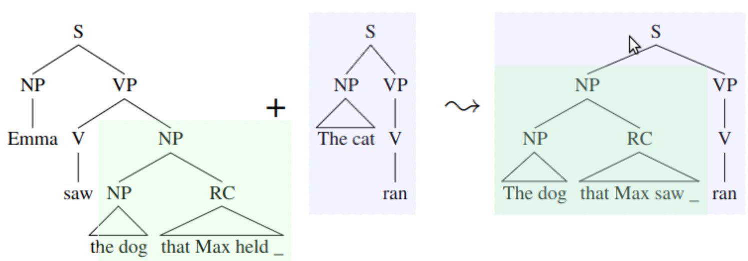
25 Janvier 2024

Datasets for compositional generalization

- COGS (Kim et al 2020) / SLOG (Li et al 2023): artificial semantic parsing datasets constucted to assess the compositional generalization abilities of NLP models
- Systematic shifts / gaps between train and test



(a) Lexical generalization: object → subject (COGS)



(b) Structural generalization: RC object→RC subject (SLOG)

- Output: logical form
  - conjunction of predicates with thematic roles
- Emma$_0$ saw$_1$ the$_2$ dog$_3$.
  *dog(x3); see.agent(x1 ,Emma)
  ∧ see.theme(x1 , x3)
- **Also exists in a variable-free version:**
  - see (agent = Emma , theme = * dog)

Datasets for compositional generalization : COGS Generalization Examples

S.3.2. Novel Combination Modified Phrases and Grammatical Roles		
Object modification → Subject modification	Noah ate **the cake on the plate.**	**The cake on the table** burned.
S.3.3. Deeper Recursion		
Depth generalization: Sentential complements	Emma said **that** Noah knew **that** the cat danced.	Emma said **that** Noah knew **that** Lucas saw **that** the cat danced.
Depth generalization: PP modifiers	Ava saw the ball **in the bottle on the table**.	Ava saw the ball **in the bottle on the table on the floor**.
S.3.4. Verb Argument Structure Alternation		
Active → Passive	The crocodile **blessed** William.	A muffin **was blessed**.
Passive → Active	The book **was squeezed**.	The girl **squeezed** the strawberry.
Object-omitted transitive → Transitive	Emily **baked**.	The giraffe **baked a cake**.
Unaccusative → Transitive	The glass **shattered**.	Liam **shatterd** the jigsaw.
Double object dative → PP dative	The girl **teleported** Liam the cookie.	Benjamin **teleported** the cake **to** Isabella.
PP dative → Double Object Dative	Jane shipped the cake to John.	Jane shipped John the cake.

Datasets for compositional generalization

- COGS :
  - Lexical generalization 80 % of the corpus
  - Structural generalization : 20 %
  - seq2seq models at around 80 % accuracy (100 on lexical, ~0 on stuctural)
- **SLOG : focus on structural generalization**
  - larger coverage of syntactic structures
    - center embedding, wh-questions, relative clauses
  - larger coverage of generalization types :
    - generalization to shorter or deeper recursion
    - distribution of semantic roles

COGS and syntax : Yao and Keller 2022

- Structural generalization is hard for seq2seq models
  - BART / T5 fail on structural generalisation
  - even when given gold syntactic trees as input
- What happens if you simplify the task ?
  - 1. replace Logical Form by gold syntactic tree
  - 2. replace Logical Form by sequence of POS
  - → BART / T5 still fail (a bit less miserably)
- In contrast, Berkeley Neural parser is 84-99 % (depending of types of syntactic generalization)

Results fom Yao & Keller 2022

	Model Class	Model	STRUCT Obj to Subj PP	CP recursion	PP recursion	LEX all 18 other types	Overall
semantics	seq2seq	BART	0	0	12	91	79
		BART+syn	0	5	8	93	80
		T5	0	0	9	97	83
		Kim and Linzen 2020	0	0	0	73	63
		Akyürek and Andreas 2021	0	0	1	96	82
		Zheng and Lapata 2022	0	12	39	99	89
		Conklin et al. 2021	0	0	0	88	75
		Csordás et al. 2021	0	0	0	95	81
		Qiu et al. 2021 *	100	100	100	100	100
	structure-aware	Liu et al. 2021	93	100	99	99	99
		Weißenhorn et al. 2022	78	100	99	100	98
syntax	seq2seq	BART	0	9	22	99	87
		T5	5	7	9	99	86
	structure-aware	Neural Berkeley Parser	84	95	98	100	99
POS tags	seq2seq	BART	0	6	19	98	85
		T5	0	4	4	98	85
	structure-aware	most frequent POS	92	98	100	92	93

COGS and syntax : Lessons Learned

- Seq2seq models (in their current form) are not appropriate for structured prediction
  - Yao & Keller hypothesis : they have access to syntactic information but can't make anything out of it
  - Recall that the early seq2seq parser (Vinyals et al 2014) needed 11M sentences to compare to sota parsers trained on PTB
- On-the-shelf parsers are much better at structural generalisation
  - Weissenhorn & al 2022 : concept prediction + graph parsing
  - Liu et al 2021 : syntactic module (unlaballed structure) + semantic module (that predicts the nature of semantic links)

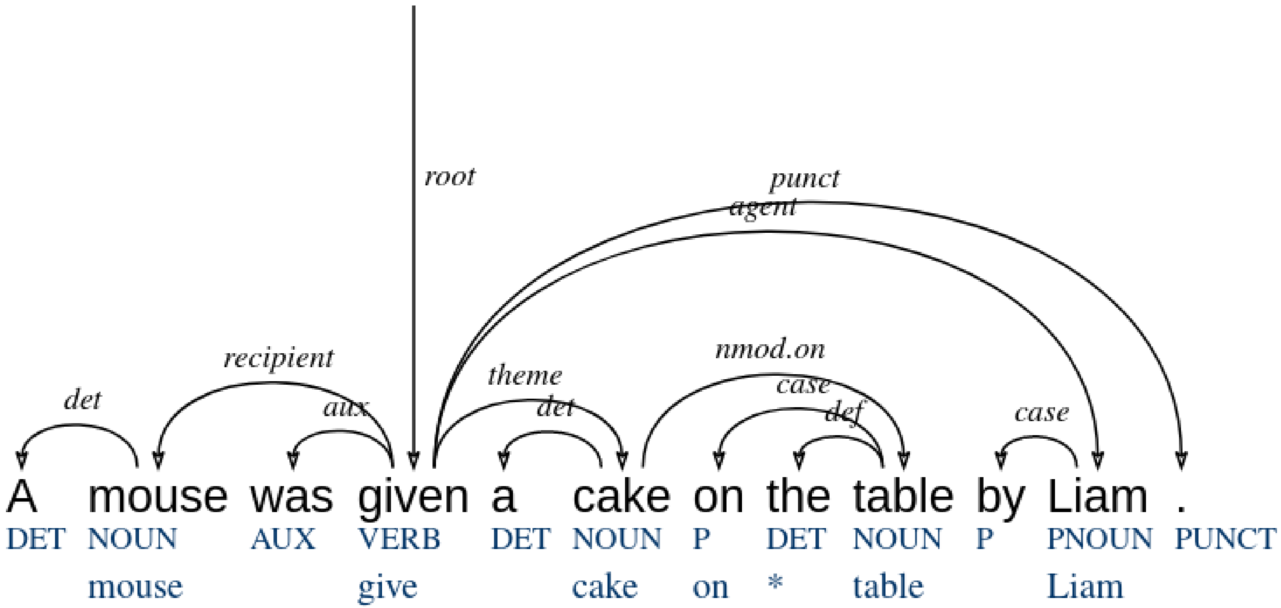
Proposals : baselines based on dependency parsing

Some observations:

- The semantic formalism used in COGS / SLOG is **anchored**
  - There is almost always a direct connection between a token and the corresponding concept / predicate
- Besides structures with control verbs, every semantic graph is in fact a **tree**
- The agent arc between a control verb and its semantic subject is redundant (can be inferred from **xcomp** relation)

Proposal : treat the task as dependency parsing

- A mouse was given a cake on the table by Liam .
- give (recipient = mouse , theme = cake (nmod . on = * table) , agent = Liam)



Only requirement of the method:

Reversible algorithm to transform logical form to dep tree

Proposal : treat the task as dependency parsing

- The penguin dreamed that Emma wanted to paint .
- dream (agent = * penguin , ccomp = want (agent = Emma , xcomp = paint (agent = Emma)))