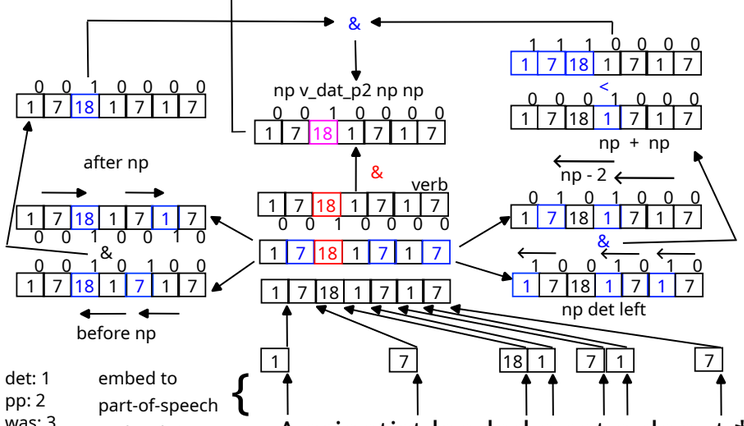
Encoder



Encoder attempts 19 flat pattern matches at once
(matching v_dat_p2 shown here)

det: 1
pp: 2
was: 3
by: 4
to: 5
that: 6
common_noun: 7
proper_noun: 8
v_trans_omissible: 9
v_trans_omissible_pp: 10
v_trans_not_omissible: 11
v_trans_not_omissible_pp: 12
v_cp_taking: 13
v_inf_taking: 14
v_unacc: 15
v_unerg: 16
v_inf: 17
v_dat: 18
v_dat_pp: 19
v_unacc_pp: 20

embed to part-of-speech and verb-type

Encoder provides words sorted into noun,verb,other order

(Transformers can derive a sorted sequence from a sequence and it is natively supported in RASP see (Weiss et al 2021))

np_det_seq = [0, 1, 0, 0, 1, 0, 1]
np_prop_seq = [0, 0, 0, 0, 0, 0, 0]
v_dat_seq = [0, 0, 1, 0, 0, 0, 0]
np_det_left_seq = [1, 0, 0, 1, 0, 1, 0]
np_two_before_seq = [0, 1, 0, 1, 0, 0, 0]
np_before_seq = [0, 0, 1, 0, 1, 0, 0]
np_after_seq = [0, 0, 1, 0, 0, 1, 0]

In bidirectional Encoder of Encoder-Decoder Transformer equivalent model

(A) = (np_det_left_seq & np_two_before_seq) = [0, 0, 0, 1, 0, 0, 0]
(B) = (np_prop_seq & np_before_seq) = [0, 0, 0, 0, 0, 0, 0]
np_np_seq = (A or B) = [0, 0, 0, 1, 0, 0, 0]
np_np_any_before_seq = [1, 1, 1, 0, 0, 0, 0]
np_v_dat_p_np_np = np_after_seq & v_dat_seq & np_before_seq & np_np_any_before_seq = [0, 0, 1, 0, 0, 0, 0]

*COGS official training data uses "lended", instead of "lent"

POS/verb-type sequence matched	Noun-verb relationship template	POS/verb-type sequence matched	Noun-verb relationship template
((det common)|proper) v_dat_p2 ((det common)|proper) ((det common)|proper)	agent, recipient, theme	((det common)|proper) v_unacc_p1 ((det common)|proper)	agent, theme
((det common)|proper) v_dat_p1 ((det common)|proper) to ((det common)|proper)	agent, theme, recipient	((det common)|proper) was v_unacc_pp_p1	theme
		((det common)|proper) was v_unacc_pp_p2 by ((det common)|proper)	theme, agent
((det common)|proper) was v_trans_omissible_p1	theme	((det common)|proper) v_inf_taking to v_inf	agent
((det common)|proper) v_trans_omissible_p1	agent	((det common)|proper) v_unerg	agent
((det common)|proper) v_trans_omissible_p2 ((det common)|proper)	agent, theme	((det common)|proper) v_unacc_p2	theme
((det common)|proper) was v_trans_omissible_pp_p2 by ((det common)|proper)	theme, agent	((det common)|proper) was v_dat_pp_p3 ((det common)|proper)	recipient, theme
((det common)|proper) v_trans_not_omissible ((det common)|proper)	agent, theme	((det common)|proper) was v_dat_pp_p4 ((det common)|proper) by ((det common)|proper)	recipient, theme, agent
((det common)|proper) was v_trans_not_omissible_pp_p1	theme	((det common)|proper) was v_dat_pp_p2 to ((det common)|proper) by ((det common)|proper)	theme, recipient, agent
((det common)|proper) was v_trans_not_omissible_pp_p2 by ((det common)|proper)	theme, agent	((det common)|proper) was v_dat_pp_p1 to ((det common)|proper)	theme, recipient

Decoder

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer | scientist (1) ; cat (_

1 4 6 9 2 0 3 5 7 8
scientist cat donut computer lended A a a beside a
lend
(stemmed verb)

} noun and verb introduction phase
number of nouns,verbs in output matches input?

v_dat_p2 (index 2) template:
agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6)

1 4 6
scientist cat donut

~~A~~ scientist ~~lended a~~ cat ~~a~~ donut beside ~~a computer~~
(extract non-pp nouns)
nouns preceded by preposition ignored in this phase

beside (6 , 9)

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer

} verb relationship phase
number of relationships in output matches template size for matched verb-centered grammar pattern?

} nmod handling phase
number of nmods in output matches number of prepositions in input?

---

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer | scientist (1) ; cat (4) ; donut (6) ; computer (9) ; lend (2) AND agent (2 , 1) AND_

1 4 6 9 2 0 3 5 7 8
scientist cat donut computer lended A a a beside a
lend
(stemmed verb)

} noun and verb introduction phase
number of nouns,verbs in output matches input?

v_dat_p2 (index 2) template:
agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6)

1 4 6
scientist cat donut

~~A~~ scientist ~~lended a~~ cat ~~a~~ donut beside ~~a computer~~
(extract non-pp nouns)
nouns preceded by preposition ignored in this phase

beside (6 , 9)

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer

} verb relationship phase
number of relationships in output matches template size for matched verb-centered grammar pattern?

} nmod handling phase
number of nmods in output matches number of prepositions in input?

---

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer | scientist (1) ; cat (4) ; donut (6) ; computer (9) ; lend (2) AND agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6) AND nmod . beside (_

1 4 6 9 2 0 3 5 7 8
scientist cat donut computer lended A a a beside a
lend
(stemmed verb)

} noun and verb introduction phase
number of nouns,verbs in output matches input?

v_dat_p2 (index 2) template:
agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6)

1 4 6
scientist cat donut

~~A~~ scientist ~~lended a~~ cat ~~a~~ donut beside ~~a computer~~
(extract non-pp nouns)
nouns preceded by preposition ignored in this phase

beside (6 , 9)

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer

} verb relationship phase
number of relationships in output matches template size for matched verb-centered grammar pattern?

} nmod handling phase
number of nmods in output matches number of prepositions in input?

---

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer | scientist (1) ; cat (4) ; donut (6) ; computer (9) ; lend (2) AND agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6) AND nmod . beside (6 , 9)

1 4 6 9 2 0 3 5 7 8
scientist cat donut computer lended A a a beside a
lend
(stemmed verb)

} noun and verb introduction phase
number of nouns,verbs in output matches input?

v_dat_p2 (index 2) template:
agent (2 , 1) AND recipient (2 , 4) AND theme (2 , 6)

1 4 6
scientist cat donut

~~A~~ scientist ~~lended a~~ cat ~~a~~ donut beside ~~a computer~~
(extract non-pp nouns)
nouns preceded by preposition ignored in this phase

beside (6 , 9)

0 1 2 3 4 5 6 7 8 9
A scientist lended a cat a donut beside a computer

} verb relationship phase
number of relationships in output matches template size for matched verb-centered grammar pattern?

} nmod handling phase
number of nmods in output matches number of prepositions in input?