

Effect of depth order on iterative nested named entity recognition models

Perceval Wajsbürt¹, Yoann Taille^{1,2}, Xavier Tannier¹

¹ LIMICS, Sorbonne Université, Paris, ² SCAI, Sorbonne Université
firstname.lastname@sorbonne-universite.fr

Objective & context

We study the effect of the order of depth of mention on nested named entity recognition (NER) models. Classic NER systems only predict non overlapping entities. Iterative models for nested NER use multiple predictions to enumerate all entities, imposing a predefined order from largest to smallest or smallest to largest. We design an order-agnostic auto-regressive model and evaluate the effet of training depth order through multiple strategies.

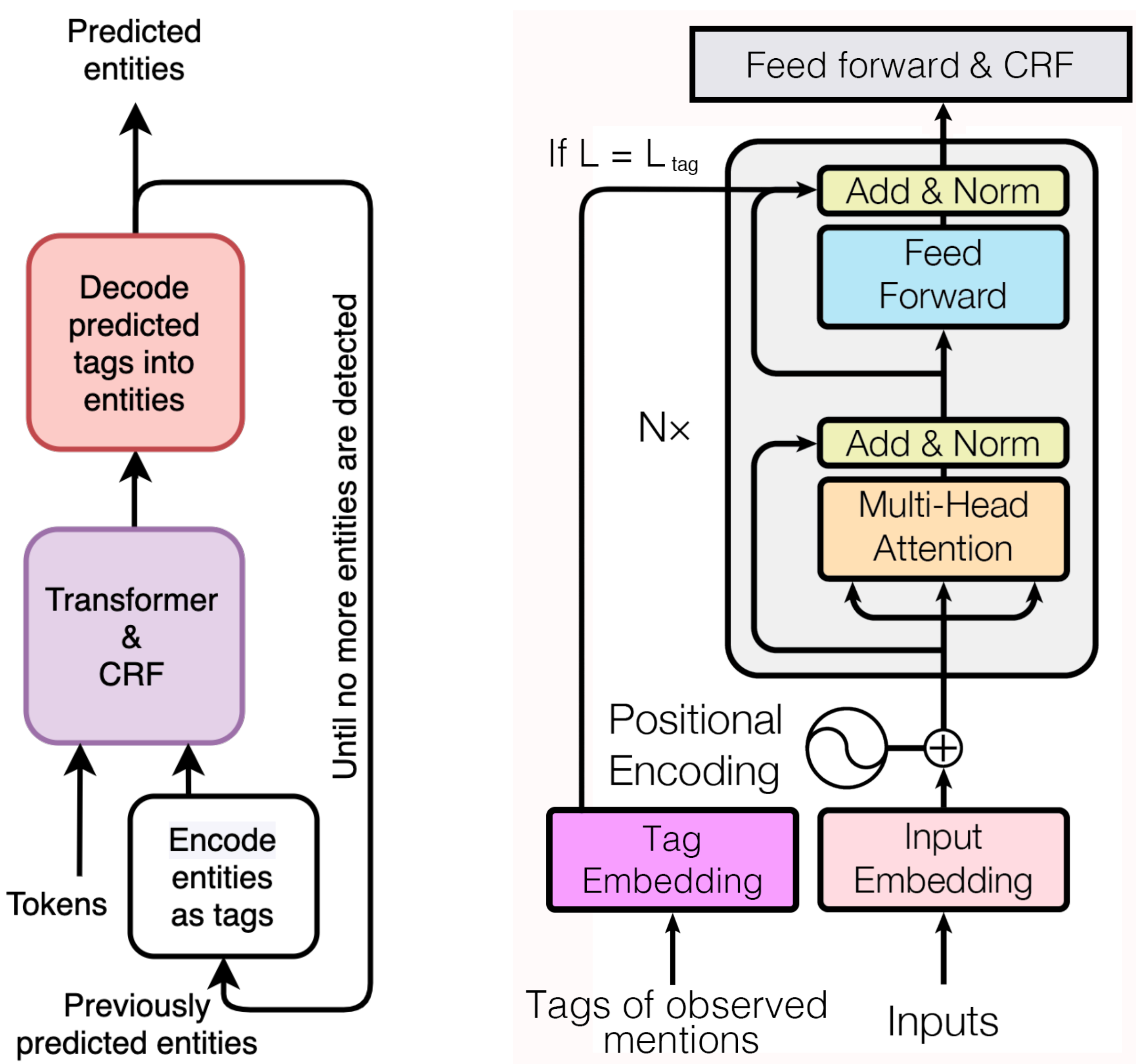
Data

	GENIA (English)			DEFT (French)		
	train	dev	test	train	dev	test
sentences	15022	1669	1855	1481	365	1024
documents	1599	190	213	82	18	67
mentions	47027	4469	5600	6439	1498	4791
mentions D_0	42965	4072	5007	5098	1226	3538
mentions D_1	3959	394	1282	1282	261	1163
mentions D_2	102	3	59	59	12	90
mentions D_3	1	0	0	0	0	0

Autoregressive training procedure

- 1: Featurize**
Encode the previously extracted entities into tags embeddings and add them to the L^{th} layer's output in BERT
 - 2: Extract**
Tag words using a BERT + CRF and decode the named entities
 - 3: Match**
When overlap between the prediction and multiple possible remaining target entities, choose one according to the order strategy and compute token cross-entropy.
- Repeat while there are entities

System & model overview



Training order strategies

Gold: *the patient has breast cancer*
Prediction: *the patient has breast cancer*

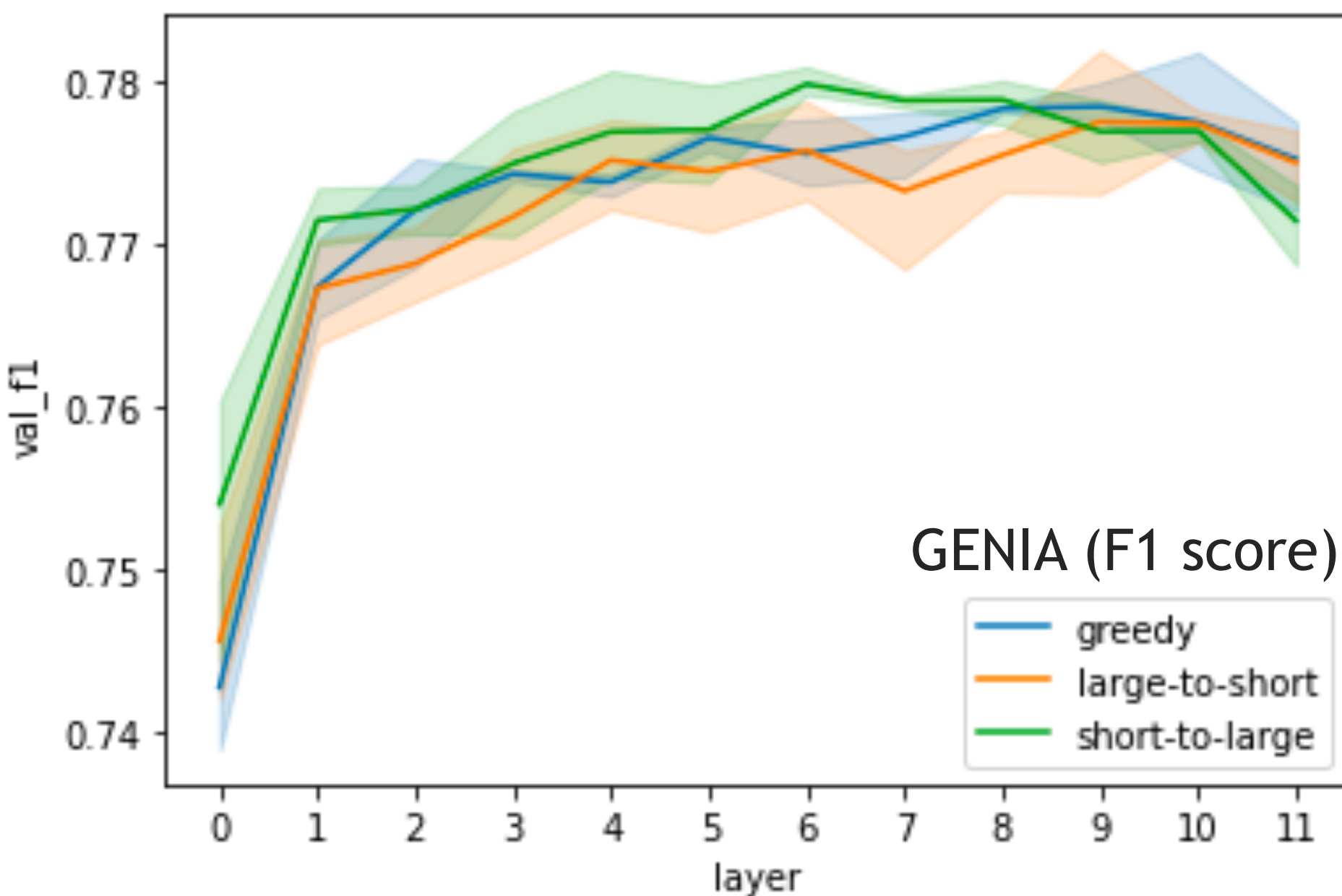
Short → large target : *cancer*
Large → short target : *breast cancer*
Greedy target: *breast cancer* (best overlap)

Encoding & decoding schemes

cancer of breast
encode ↓ ↑ decode
BIO scheme: B-dis I-dis I-dis
BIOUL scheme: B-dis I-dis L-dis

DEFT F1 score	BIO decoding	BIOUL decoding
BIO encoding	0.7221 ± 2.98e ⁻³	0.7341 ± 6.09e ⁻³
BIOUL decoding	0.7261 ± 5.40e ⁻³	0.7368 ± 4.28e ⁻³

Tag insertion layer index



Results

GENIA	GENIA			DEFT task 3.1			DEFT task 3.2			DEFT
	P	R	F1	P	R	F1	P	R	F1	
large→short	0.8016	0.7184	0.7577	0.6257	0.6062	0.6158	0.7407	0.7470	0.7439	0.7082
short→large	0.8028	0.7336	0.7666	0.6105	0.6194	0.6149	0.7564	0.7449	0.7506	0.7120
greedy	0.8126	0.7211	0.7641	0.6263	0.6090	0.6175	0.7615	0.7416	0.7514	0.7134