# The Concordia NLG Surface Realizer at SR'19


Farhood Farahnak, Laya Rafiee, Leila Kosseim and Thomas Fevens
Dept. of Computer Science and Software Engineering
Concordia University
Montreal, Canada
firstname.lastname@concordia.ca


## Objective

We participated in the SR'19 [1] shallow track only. The model takes as input an unordered and lemmatized list of words, and finds the correct word order and inflected words.

## Dataset

| # | Training set | Number of Sentences |
|---|---|---|
| 1 | en_ewt-ud-train | 12,543 |
| 2 | en_gum-ud-train | 2,914 |
| 3 | en_lines-ud-train | 2,738 |
| 4 | en_partut-ud-train | 1,781 |

Table 1: Training data which are taken from Universal Dependency (UD) datasets [2]

## Methodology

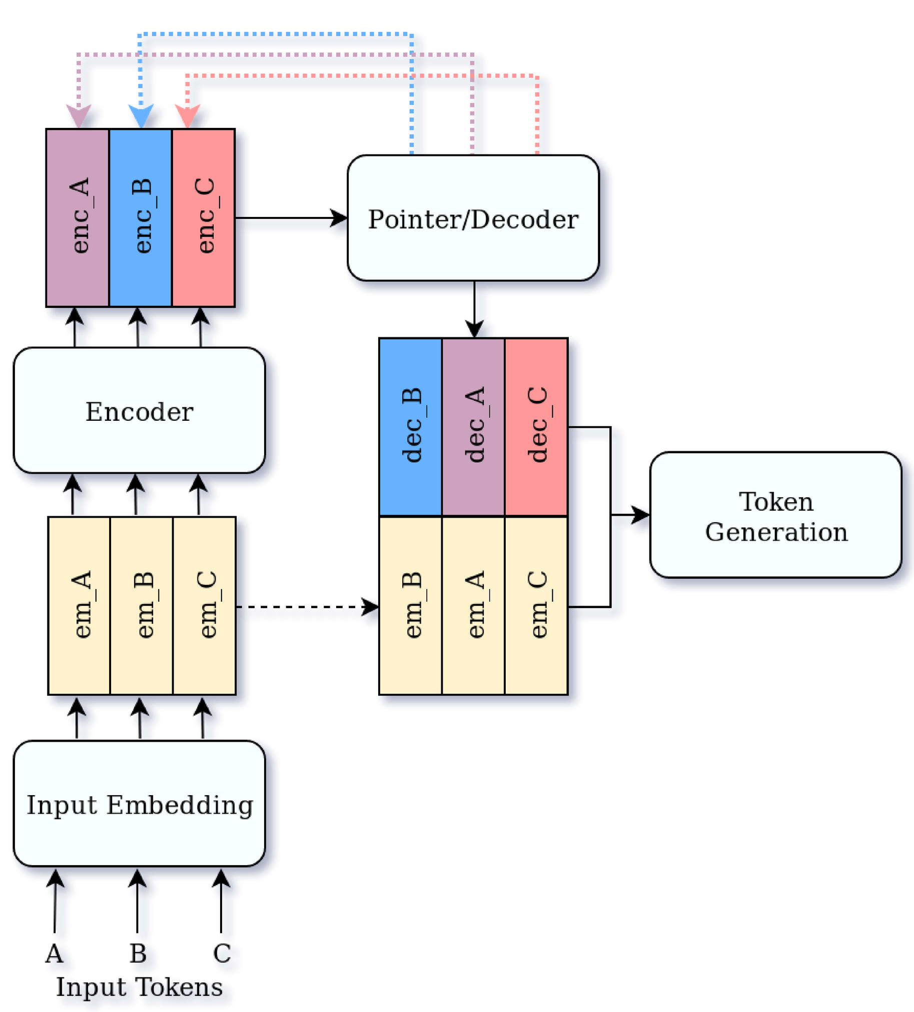



Figure 1: The model architecture used for the shallow track at SR'19

Our model (Figure 1) consists of five main sub-modules.

- **Input Embedding:** embeds each token $X$ alongside its features to an embedded token ($em\_X$)
- **Encoder:** encodes each embedded token ($enc\_X$) [1]
- **Decoder:** generates the query for the Pointer ($dec\_X$) given previously selected tokens and encoded tokens from the Encoder [1]
- **Pointer:** is an attention mechanism that attends over the encoded tokens and uses $dec\_X$ as its query [4]
- **Token Generation:** generates the inflected form of tokens using the concatenation of embedded token ($em\_X$) and decoded token ($dec\_X$)

## Results and Analysis

Our model achieved the average scores of 48.1 and 60.9 for the Readability/Quality and Meaning Similarity on the English datasets. Both the automatic and the human evaluations show that our system performance was below the median.

| # | | Test sets | BLEU |
|---|---|---|---|
| 1 | In-domain | en_ewt-ud-test | 22.08 |
| 2 | | en_gum-ud-test | 15.32 |
| 3 | | en_lines-ud-test | 15.30 |
| 4 | | en_partut-ud-test | 10.07 |
| 5 | Out-of-domain | en_pud-ud-test | 12.36 |
| 6 | Predicted | en_ewt-Pred-HIT-edit | 21.21 |
| 7 | | en_pud-Pred-LATTICE | 12.89 |

Table 2: BLEU Scores of our submission in SR'19

[1] We employed transformer encoder and decoder [3]

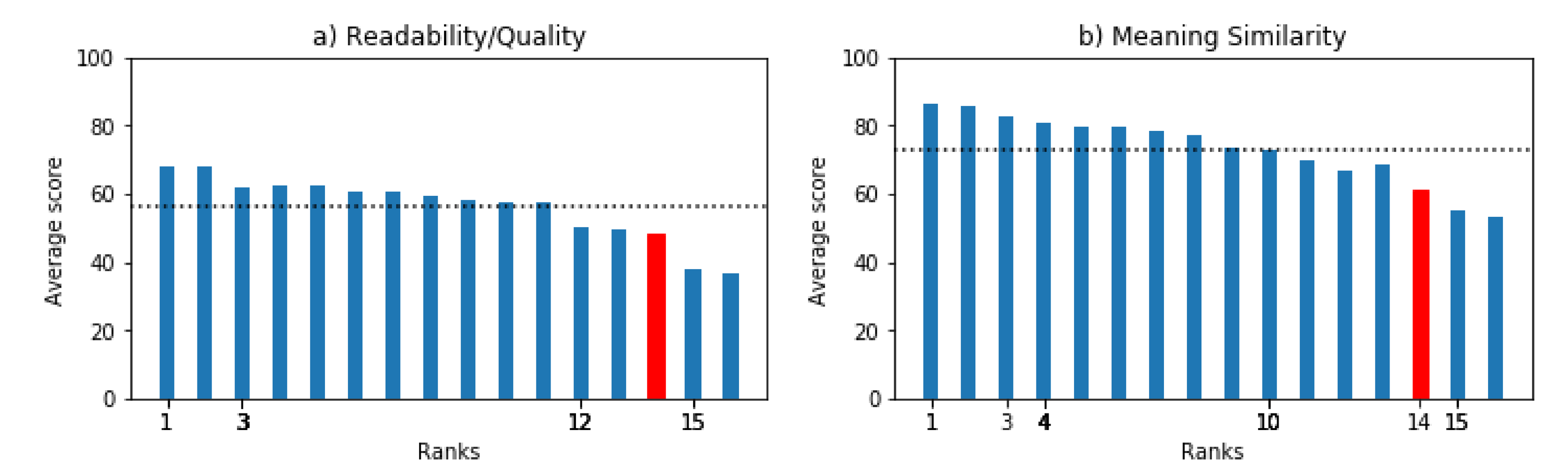
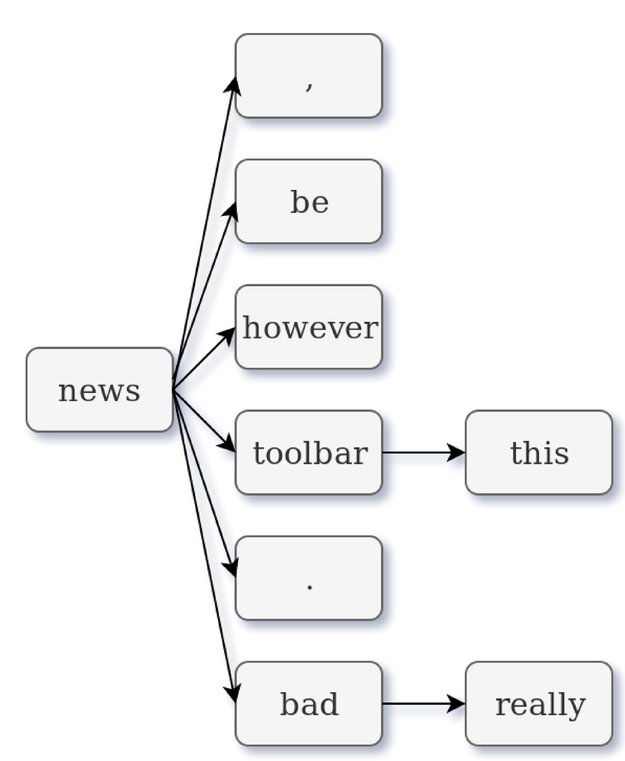



Figure 2: Human evaluation results compared to all participants at SR'19

**Analysis:** An analysis of a few generated outputs of our model showed that the low performance is mainly due to the poor performance of the Token generation module. An example is provided in Figure 3.




- Generated Text: *however, this hacaoglu is really bad news.*
- Ground Truth: *however, this toolbar is really bad news.*

Figure 3: Example of a generated text

## Future Work

The proposed system is composed of a pointer network where its encoder and decoder modules borrowed from transformer, aim to reconstruct the tokens' order and inflection.
For the future work, it would be interesting to

- Investigate the model sensitivity to the training size
- Utilize more features provided by the universal dependency structure
- The possibility of using pretrained language models

## Contact

For further information please visit:
https://github.com/farhoodf/SR19

## References


[1] Simon Mille, Anja Belz, Bernd Bohnet, Yvette Graham, and Leo Wanner. The Second Multilingual Surface Realisation Shared Task (SR'19): Overview and Evaluation Results. In *Proceedings of the 2nd Workshop on Multilingual Surface Realisation (MSR), 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Hong Kong, China, 2019.

[2] Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher D. Manning. Universal Stanford dependencies: A cross-linguistic typology. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, pages 4585–4592, Reykjavik, Iceland, May 2014. European Languages Resources Association (ELRA).

[3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, pages 5998–6008. Curran Associates, Inc., 2017.

[4] Oriol Vinyals, Meire Fortunato, and Navdeep Jaitly. Pointer Networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, pages 2692–2700. Curran Associates, Inc., 2015.