Data-to-text generation **across domains**?

Take trivial templates & **fuse** them one-by-one!

Data-to-Text Generation with Iterative Text Editing

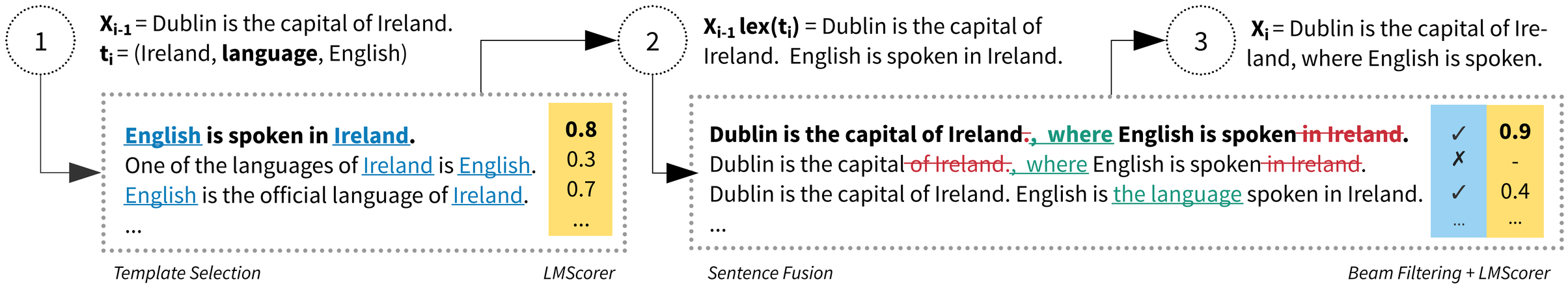
Zdeněk Kasner and Ondřej Dušek
{kasner,odusek}@ufal.mff.cuni.cz
Charles University, Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Prague, Czech Republic

Summary

- We generate **text** from **RDF triples** by iteratively fusing simple templates for each predicate with a neural model.
- Our approach maximizes **semantic accuracy** of the text with strict entity matching and **limited vocabulary** of the model.
- Text generation is possible even without any in-domain examples: **zero-shot domain adaptation**.

Data & Templates

- **Datasets**
  - **WebNLG** (Gardent et al., 2017) – DBpedia
  - **E2E** (Dušek et al. 2019) – restaurants
  - \+ **DiscoFuse** (Geva et al. 2019) – sentence fusion
- **Training data**: *lex*(*n* triples) + template → *lex*(*n+1* triples)
- **Templates:** automatically extracted from the datasets (+ fallback)



Approach

- **LaserTagger** (Malmi et al., 2019): a BERT-based text-editing model trained for sentence fusion
- **LMScorer**: a pretrained GPT-2 language model (Radford et al., 2019) used for scoring the sentences
- **Decoding algorithm**

for each triple t_i **do**
  X'_i = **concatenate** the text X_{i-1} and a template for the triple t_i
  apply the **sentence fusion** model (LaserTagger) on X'_i
  filter the fusion hypotheses in the beam with entity matching
  X_i = **select the best** fusion hypothesis with LMScorer

Results & Future Work

- **Results**
  - the model beats the baseline (∼5-8 BLEU), but not SOTA
  - a fused sentence with no entities missing is generated in 50-70% of steps
  - otherwise a fallback is used → **entities preserved in all cases**
  - the model trained on DiscoFuse is able to perform simple sentence fusion on both WebNLG and E2E datasets
- **Future work**: improving the sentence fusion model; flexible sentence ordering; better entity matching

Triples	*(Albert Jennings Fountain, deathPlace, New Mexico Territory); (Albert Jennings Fountain, birthPlace, New York City); (Albert Jennings Fountain, birthPlace, Staten Island)*
Text X_0	Albert Jennings Fountain died in New Mexico Territory.
Text X_1	Albert Jennings Fountain, who died in New Mexico Territory, was born in New York City.
Text X_2	Albert Jennings Fountain, who died in New Mexico Territory, was born in New York City, Staten Island.
Reference	Albert Jennings Fountain was born in Staten Island, New York City and died in the New Mexico Territory.

Presented at INLG, online, 15 Dec 2020.
Supported by the Charles University grant No. 140320, the SVV project No. 260575, and the Charles University project PRIMUS/19/SCI/10.