Stylistic Transfer in Natural Language Generation Systems Using Recurrent Neural Networks

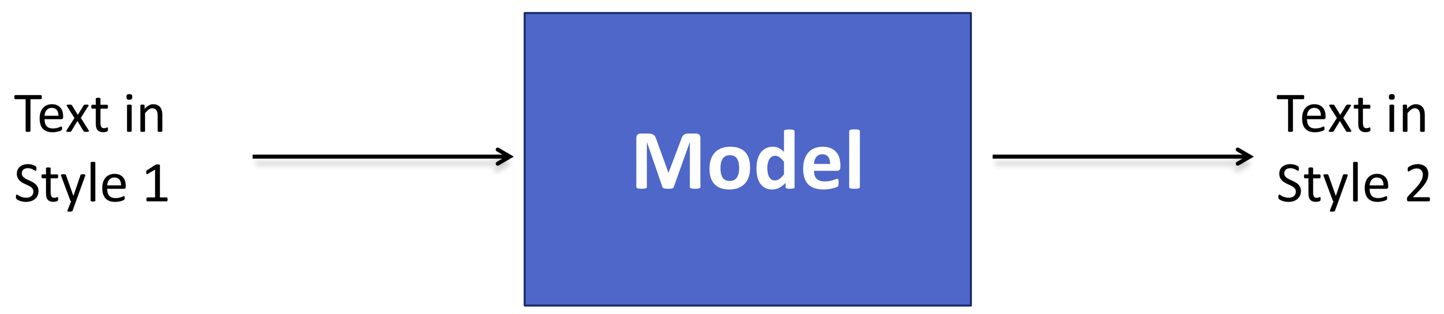
Jad Kabbara*   Jackie C.K. Cheung†

* jad@cs.mcgill.ca
† jcheung@cs.mcgill.ca

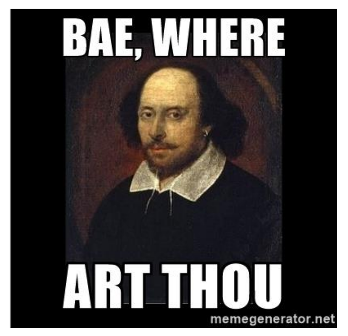
McGill School of Computer Science

Stylistic Transfer

Stylistic transfer: changing the style of a passage while preserving its meaning



Example from Shakespeare's As You Like It



Input: As I remember, Adam, it was upon this fashion bequeathed me by will but poor a thousand crowns, and, as thou sayest, charged my brother on his blessing to breed me well. And there begins my sadness.

Output: I remember, Adam, that's exactly why my father only left me a thousand crowns in his will. And as you know, my father asked my brother to make sure that I was brought up well. And that's where my sadness begins.

Applications

- Making old texts more accessible to a contemporary reader
- Reproducing technical articles to a broader audience
- Security/Privacy: Author obfuscation

Related Work

- Inkpen et al. (2004): Use list of near-synonyms
- Xu et al. (2012): Two approaches:
  (1) Translation framework
  (2) Set of rules for linguistic transfer
- Sennrich et al. (2016): adding side constraints (features) to control politeness in translation

The Learning Model

Past work requires expensive parallel data!
→ Task proposal:

- Use deep learning to tackle stylistic transfer
- Inspired by work in image processing (e.g., Cheung et al. 2014)

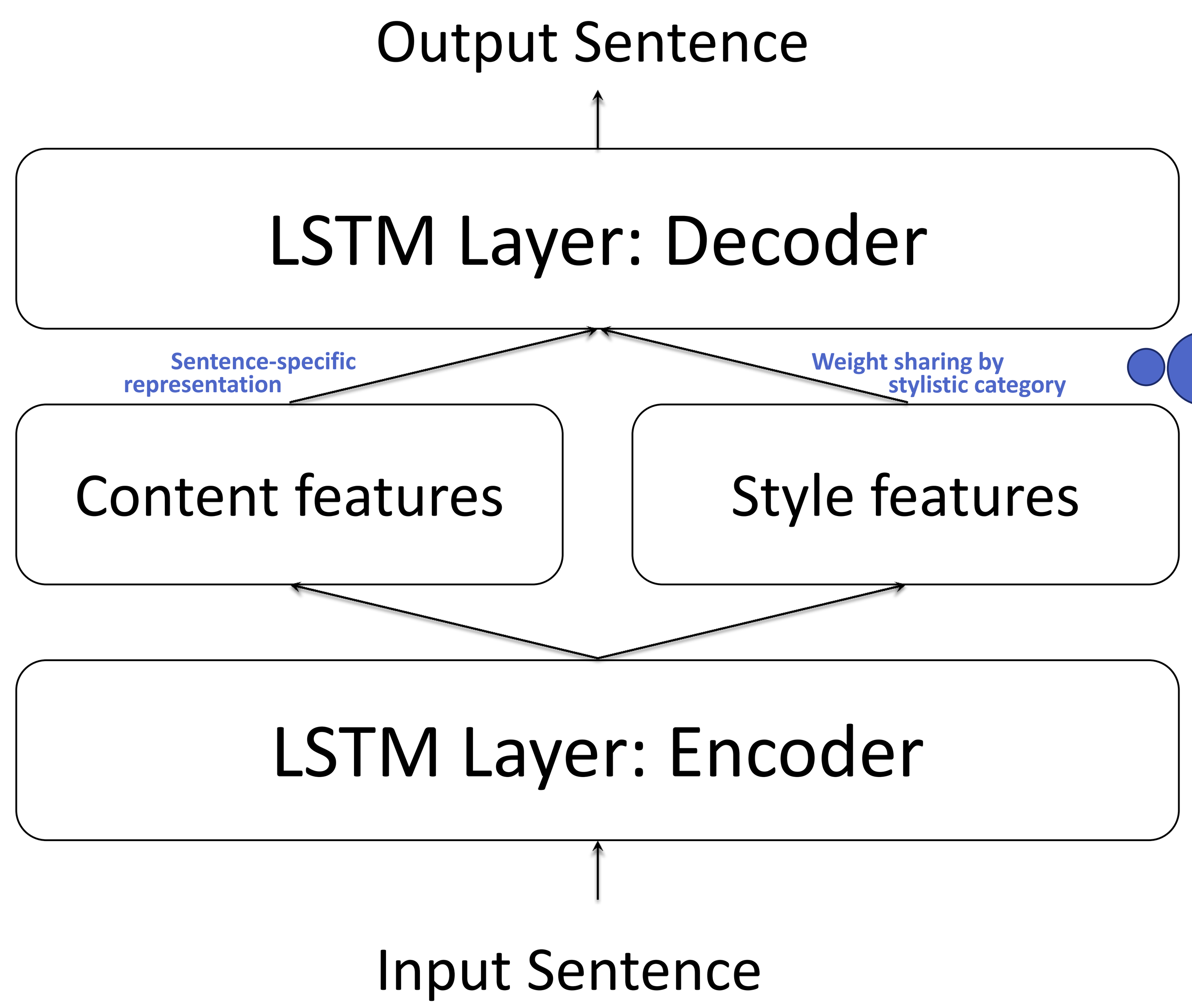
Highlights

- Recurrent Neural Network
- Encoder-Decoder Structure
- Separate content features from style features
- Cross-covariance term in cost function

Key points

- No need for time-consuming design of linguistic features (which may not generalize well anyway)
- No need for (expensive!) parallel data

The model in a nutshell



Approach

For a style transfer task from Style A to Style B:

- Collect relevant corpora for each style
- Train a model on each style to learn the stylistic features for that style
- During generation, use corresponding stylistic features for desired style

Evaluation

Three main criteria for evaluation:

- **Soundness** (textually entailed with the original version?)
- **Coherence** (free of grammatical errors? proper word usage? etc.)
- **Effectiveness** (match the desired style?)

- Human evaluation of snippets of generated text
- Automatic evaluation (e.g., ROUGE, BLEU)

Hypothesis:

- Weight sharing
- Learning across training samples all sharing the same style

➤ Stylistic latent variables would capture a high-level representation of style, complimentary to that of content variables.

- Diana Zaiu Inkpen and Graeme Hirst. 2004. Near-synonym choice in natural language generation. *Recent Advances in Natural Language Processing 2004*.
- Wei Xu, Alan Ritter, William B Dolan, Ralph Grishman, and Colin Cherry. Paraphrasing for style. *COLING 2012*.
- Rico Sennrich, Barry Haddow, and Alexandra Birch. Controlling politeness in neural machine translation via side constraints. *NAACL HLT 2016*.
- Brian Cheung, Jesse A Livezey, Arjun K Bansal, and Bruno A Olshausen. 2014. Discovering hidden factors of variation in deep networks. arXiv preprint arXiv:1412.6583