

CS11-731
Machine Translation and
Sequence-to-Sequence Models
Monolingual Transduction

Graham Neubig



Carnegie Mellon University

Language Technologies Institute

Site

<https://phontron.com/class/mtandseq2seq2019/>

(Slides by Graham Neubig and Antonis Anastasopoulos)

Examples

- Summarization
- Paraphrasing
- Formality

Summarization

- Granularity:
 - Sentence compression
 - Single document summarization
 - Multi-document summarization
- Type:
 - Abstractive
 - Extractive

Sentence Compression

- Reduce sentence length
- Common setting: first sentence in new article to title

Abstractive Sentence Summarization with Attentive Recurrent Neural Networks

Sumit Chopra

Facebook AI Research
spchopra@fb.com

Michael Auli

Facebook AI Research
michaেলাuli@fb.com

Alexander M. Rush

Harvard SEAS

srush@seas.harvard.edu

<https://www.aclweb.org/anthology/N16-1012.pdf>

Controlling Output Length in Neural Encoder-Decoders

Yuta Kikuchi¹

kikuchi@lr.pi.titech.ac.jp

Graham Neubig^{2*}

gneubig@cs.cmu.edu

Ryohei Sasano¹

sasano@pi.titech.ac.jp

Hiroya Takamura¹

takamura@pi.titech.ac.jp

Manabu Okumura¹

oku@pi.titecjh.ac.jp

¹Tokyo Institute of Technology, Japan

²Carnegie Mellon University, USA

<https://arxiv.org/pdf/1609.09552.pdf>

Higher-order Syntactic Attention Network for Long Sentence Compression

Hidetaka Kamigaito[♡], Katsuhiko Hayashi[◇], Tsutomu Hirao[♠] and Masaaki Nagata[♠]

[♡] Institute of Innovative Research, Tokyo Institute of Technology

[◇] Institute of Scientific and Industrial Research, Osaka University

[♠] NTT Communication Science Laboratories, NTT Corporation

`kamigaito@lr.pi.titech.ac.jp, katsuhiko-h@sanken.osaka-u.ac.jp,`

`{hirao.tsutomu,nagata.masaaki}@lab.ntt.co.jp`

<https://www.aclweb.org/anthology/N18-1155.pdf>

Single Document Summarization

- Compress a document to a sentence

Get To The Point: Summarization with Pointer-Generator Networks

Abigail See
Stanford University
abisee@stanford.edu

Peter J. Liu
Google Brain
peterjliu@google.com

Christopher D. Manning
Stanford University
manning@stanford.edu

<https://arxiv.org/abs/1704.04368>

- Adds concept of coverage

A DEEP REINFORCED MODEL FOR ABSTRACTIVE SUMMARIZATION

Romain Paulus, Caiming Xiong & Richard Socher

Salesforce Research

172 University Avenue

Palo Alto, CA 94301, USA

`{rpaulus, cxiong, rsocher}@salesforce.com`

Multi-document Summarization

- Multiple documents into a single summary

Graph-based Neural Multi-Document Summarization

**Michihiro Yasunaga¹ Rui Zhang¹ Kshitijh Meelu¹
Ayush Pareek² Krishnan Srinivasan¹ Dragomir Radev¹**

¹Department of Computer Science, Yale University

²The LNM Institute of Information Technology

`{michihiro.yasunaga,r.zhang,kshitijh.meelu}@yale.edu`

`{ayush.original}@gmail.com`

`{krishnan.srinivasan,dragomir.radev}@yale.edu`

- Extractive summary with saliency

Paraphrasing

- Microsoft research paraphrase corpus
- <https://www.aclweb.org/anthology/I05-5002.pdf>
- https://github.com/wasiahmad/paraphrase_identification/tree/master/dataset/msr-paraphrase-corpus
- https://github.com/wasiahmad/paraphrase_identification/blob/master/dataset/msr-paraphrase-corpus/msr_paraphrase_train.txt

Paraphrasing Bilingually (Neurally)

Paraphrasing Revisited with Neural Machine Translation

Jonathan Mallinson, Rico Sennrich and Mirella Lapata

Institute for Language, Cognition and Computation

School of Informatics, University of Edinburgh

10 Crichton Street, Edinburgh EH8 9AB

`J.Mallinson@ed.ac.uk, {rsennric,mlap}@inf.ed.ac.uk`

<https://www.aclweb.org/anthology/E17-1083.pdf>

Style and Attribute Transfer

- What is style? a subset of "linguistic variation"

Variation (linguistics)

From Wikipedia, the free encyclopedia

Variation is a characteristic of language: there is more than one way of saying the same thing. Speakers may vary pronunciation ([accent](#)), word choice ([lexicon](#)), or morphology and syntax (sometimes called "[grammar](#)").^[1] But while the diversity of variation is great, there seem to be boundaries on variation – speakers do not generally make drastic alterations in sentence word order or use novel sounds that are completely foreign to the language being spoken.^[2] Linguistic variation does not equate with language [ungrammaticality](#), but speakers are still (often unconsciously) sensitive to what is and is not possible in their native [lect](#).

Supervised Transfer

- Just train an MT model!

Paraphrasing for Style

Wei Xu¹ Alan Ritter² William B. Dolan³ Ralph Grishman¹ Colin Cherry⁴

(1) New York University

(2) University of Washington

(3) Microsoft Research

(4) National Research Council, Canada

`xuwei@cims.nyu.edu, aritter@cs.washington.edu, billdol@microsoft.com,
grishman@cs.nyu.edu, Colin.Cherry@nrc-cnrc.gc.ca`

<https://www.aclweb.org/anthology/C12-1177.pdf>

Adversarial Transfer

Style Transfer from Non-Parallel Text by Cross-Alignment

Tianxiao Shen¹ Tao Lei² Regina Barzilay¹ Tommi Jaakkola¹

¹MIT CSAIL ²ASAPP Inc.

¹{tianxiao, regina, tommi}@csail.mit.edu ²tao@asapp.com

<https://papers.nips.cc/paper/7259-style-transfer-from-non-parallel-text-by-cross-alignment.pdf>

Retrieval-based Transfer

Delete, Retrieve, Generate: A Simple Approach to Sentiment and Style Transfer

Juncen Li^{*1} Robin Jia² He He² Percy Liang²

¹ WeChat Search Application Department, Tencent

² Computer Science Department, Stanford University

juncenli@tencent.com

{robinjia, hehe, pliang}@cs.stanford.edu

(a) Extracting attribute markers



(b) Attribute transfer

