Toan Q. Nguyen

Contact Information	355 Fitzpatrick Hall Notre Dame, IN 46556	+1-574-329-2791 tnguye28@nd.edu

Research Interests Neural Machine Translation (with focus on low-resource languages), Deep Learning

Education

University of Notre Dame, Notre Dame, IN — Jan 2016 to present
Ph.D., Computer Science

- Advisor: David Chiang, Ph.D

University of Southampton, Southampton, UK — Sept 2010 to June 2013

BEng, Electronic Engineering

- *First-class honours*

Research Projects

Improving lexical choice in NMT

We propose two solutions to alleviate the mistranslation for rare words issue in NMT. First, we argue that the output layer, which computes the inner product between the hidden state with all target word embeddings, rewards frequent words disproportionately. So we propose to fix all target word embedding to a certain value. Second, we integrate a simple lexical module which is jointly learned with NMT and whose output is used to bias NMT's prediction. Our experiments on 8 different language pairs show improvements of from 1.1 to 4.3 BLEUs (paper accepted to *NAACL'18*).

Transfer Learning across Low-resource, Related Languages for NMT

Using Byte-Pair-Encoding, we generate data with overlapped vocabulary between low-resource, related languages. This data is then used to train a parent model on one language pair, and the trained model is later used to initialize training on another language pair. Our method shows improvements of up to 4.3 BLEUs over a strong BPE baseline for Uyghur-English (paper accepted to *IJCNLP'17*).

Witwicky: An implementation of Transformer in PyTorch

Our reimplementation of Vaswani et al.'s **Attention Is All You Need** (Transformer) with Pytorch. We show that the successively insertion of Layer Normalization in between residual connection makes it difficult to train Transformer. By rerouting the Layer Normalization to the front of each sublayer, training is stable and we no longer need to linearly increase the learning rate (warmup) at the start of training. We further investigate other aspects of training such as regularization, gradient clipping, data preprocessing... Putting together, our implementation performs very well on both low- and high-resource translation tasks, sometimes 5 BLEU higher than other published Transformer baselines. https://github.com/tnq177/witwicky

Research Experience

Research Assistant — Jan 2016 - present
Natural Language Processing Group
Department of Computer Science and Engineering
University of Notre Dame
Supervisor: David Chiang, Ph.D

Visiting Student — Sep 2018 - Dec 2018

CILVR Group
Center for Data Science
New York University
Supervisor: Kyunghyun Cho, Ph.D

Teaching Experience

Teaching Assistant Spring 2016
CSE 40535/60535 - Computer Vision
Instructor: Adam Czajka, Ph.D
Division of Computer Science and Engineering,
University of Notre Dame

Teaching Assistant Fall 2016
CSE 40868/60868 - Introduction to Neural Networks
Instructor: Adam Czajka, Ph.D
Division of Computer Science and Engineering,
University of Notre Dame

Professional Experience

Software Developer, East Agile Mar 2015 to November 2015
Develop web-based application to convert complicated Google Analytics information to simple and meaningful reports with focus on small businesses.

Software Developer, Tripchum August 2014 to March 2015
Develop and deploy multiple experimental tools such as Hubot for Google Analytics information analysis and notification on Hipchat, Twitter Direct Message integration with in-house chatting and logging system.

Programming Skills

Proficient: Python, Pytorch, Tensorflow
Basic: C, C++, Javascript, Django, NodeJS, OpenCV

Publications

1. **Toan Q. Nguyen** and David Chiang. Improving lexical choice in neural machine translation. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 334–343. Association for Computational Linguistics, 2018

2. **Toan Q. Nguyen** and David Chiang. Transfer learning across low-resource, related languages for neural machine translation. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 296–301, Taipei, Taiwan, November 2017. Asian Federation of Natural Language Processing

3. Antonios Anastasopoulos, Alison Lui, **Toan Q. Nguyen**, and David Chiang. Neural machine translation of text from non-native speakers. In *Proc. NAACL HLT*, 2019. to appear

4. Leon Cheung, Thamme Gowda, Ulf Hermjakob, Nelson H S Liu, Jonathan May, Alexandra Mayn, Nima Pourdamghani, Michael Pust, Kevin Knight, Nikolaos Malandrakis, Pavlos Papadopoulos, Anil Ramakrishna, Karan Singla, Víctor Martínez, Colin Vaz, Doan Can, S. Narayanan, Kenton Murray, **Toan Nguyen**, David Chiang, Xiaoman Pan, Boliang Zhang, Ying Chuan Lin, Di Lu, Lifu Huang, Kevin Blissett, Tongtao Zhang, Ondrej Glembek, Murali Karthick Baskar, Santosh Kesiraju, Lukás Burget, Karel Benes, Igor Szoke, Karel Veselý, Camille Goudeseune, M. H. Johnson, Leda Sari, Wenda Chen, and Angli Liu. Elisa system description for lorehlt 2017. 2017

Awards

- Outstanding Graduate Teaching Assistant, University of Notre Dame, Department of Computer Science and Engineering, Spring 2017
- The Vietnam Education Foundation Fellowship, Cohort 2015
- GD Sims Prize, University of Southampton, School of Electronics and Computer Science, 2011
- Honourable Mention, Asian Physics Olympiad, Ulan Bator, Mongolia, 2008

References

David Chiang
Associate Professor
Department of Computer Science and Engineering
University of Notre Dame
Phone: 574-631-9441
E-mail: dchiang@nd.edu

Adam Czajka
Visiting Assistant Professor
Department of Computer Science and Engineering
University of Notre Dame
Phone: 574-631-7072
E-mail: aczajka@nd.edu