

Toan Q. Nguyen

CONTACT INFORMATION	355 Fitzpatrick Hall Notre Dame, IN 46556	+1-574-329-2791 tnguye28@nd.edu tnq177.github.io
RESEARCH INTERESTS	Neural Machine Translation (with focus on low-resource languages), Deep Learning	
EDUCATION	University of Notre Dame , Notre Dame, IN Ph.D., Computer Science • Advisor: David Chiang, Ph.D University of Southampton , Southampton, UK BEng, Electronic Engineering • <i>First-class honours</i>	Jan 2016 to present Sept 2010 to June 2013
RESEARCH PROJECTS	Transformers without Tears: Improving the Normalization of Self-Attention We evaluate three simple, normalization-centric changes to improve Transformer training. Our changes include proper routing of the residual connection, replacing <i>every</i> LayerNormalization with ℓ_2 normalization with a single scale parameter, and fixing the word embedding norm to 1. Putting together, these modifications allow sharper performance curve, stable training even with warm-up free, large learning rate. For low-resource, we achieve an average of +1.1 BLEU over SOTA baselines. For high-resource, our model performs comparable to the baseline. Additionally, our ℓ_2 -based normalization uses less computation than LayerNormalization so training is faster (roughly 5%) (paper accepted to <i>IWSLT'19</i>). 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 <i>NAACL'18</i>). 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 <i>IJCNLP'17</i>).	
RESEARCH EXPERIENCE	Research Assistant Natural Language Processing Group Department of Computer Science and Engineering University of Notre Dame Supervisor: David Chiang, Ph.D	Jan 2016 - present

	Applied Scientist Intern Amazon AWS AI Transcribe Science team Mentor: Julian Salazar Manager: Katrin Kirchhoff	Jun 2019 - Sept 2019
	Visiting Student CILVR Group Center for Data Science New York University Supervisor: Kyunghyun Cho, Ph.D	Sep 2018 - Dec 2018
TEACHING EXPERIENCE	Teaching Assistant CSE 40868/60868 - Introduction to Neural Networks Instructor: Adam Czajka, Ph.D Division of Computer Science and Engineering, University of Notre Dame	Fall 2016
	Teaching Assistant CSE 40535/60535 - Computer Vision Instructor: Adam Czajka, Ph.D Division of Computer Science and Engineering, University of Notre Dame	Spring 2016
PROFESSIONAL EXPERIENCE	Software Developer , East Agile Develop web-based application to convert complicated Google Analytics information to simple and meaningful reports with focus on small businesses.	Mar 2015 to November 2015
	Software Developer , Tripchum 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.	August 2014 to March 2015
PROGRAMMING SKILLS	Proficient: Python, Pytorch, Tensorflow Basic: C, C++, Javascript, Django, NodeJS, OpenCV	
PUBLICATIONS	<ol style="list-style-type: none"> Toan Q. Nguyen and Julian Salazar. Transformers without tears: Improving the normalization of self-attention. In <i>Proceedings of the International Workshop on Spoken Language Translation</i>, 2019 Kenton Murray, Jeffery Kinnison, Toan Q. Nguyen, Walter Scheirer, and David Chiang. Auto-sizing the Transformer network: Improving speed, efficiency, and performance for low-resource machine translation. In <i>Proc. Workshop on Neural Text Generation</i>, 2019. To appear Antonios Anastasopoulos, Alison Lui, Toan Q. Nguyen, and David Chiang. Neural machine translation of text from non-native speakers. In <i>Proc. NAACL HLT</i>, 2019 Toan Q. Nguyen and David Chiang. Improving lexical choice in neural machine translation. In <i>Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)</i>, pages 334–343. Association for Computational Linguistics, 2018 Toan Q. Nguyen and David Chiang. Transfer learning across low-resource, related languages for neural machine translation. In <i>Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)</i>, pages 296–301, Taipei, Taiwan, November 2017. Asian Federation of Natural Language Processing 	

6. 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