

NIKOLAY BOGOYCHEV

Research Interest: Machine Translation, GPGPU, GEMM, Low/high level performance optimisation

@ nheart@gmail.com

📍 Edinburgh, UK

🌐 www.nbogoychev.com

🐙 github.com/XapaJIaMnu

🐦 twitter.com/XapaJIaMnu

🔗 scholar.google.com/citations?user=z2hWTDMAAAAJ&hl=en

🌐 www.linkedin.com/in/nikolay-bogoychev-01831023/

EDUCATION

PhD

School of Informatics, University of Edinburgh

📅 2015 - 2019

📍 Edinburgh, United Kingdom

- Thesis: Fast machine translation on parallel and massively parallel hardware
- Supervisors: Adam Lopez, Kenneth Heafield

BSc

Artificial Intelligence and Computer science, University of Edinburgh

📅 2010 - 2014

📍 Edinburgh, United Kingdom

- First Class Honours

WORK EXPERIENCE

Postdoctoral Researcher

University of Edinburgh, School of Informatics

📅 January 2019 - Present

📍 Edinburgh, United Kingdom

Working on neural machine translation and High performance computing

- translateLocally, cross platform, fast, offline desktop translation software github.com/XapaJIaMnu/translateLocally
- marian NMT 8bit integer CPU/GPU tensorcore inference code
- CPU int8 VNNI GEMM: github.com/kpu/intgemm
- Experimental GEMM tiling: github.com/XapaJIaMnu/bfTile
- Bergamot: Privacy driven Firefox translation. For details: browser.mt

Consultant

Intel

📅 01/03/2019 - Present

📍 Edinburgh, United Kingdom

Exploring opportunities for CPU training and inference for NMT:

- BF16 and other quantisation schemes for training and inference.
- Low level hardware specific optimisation with focus on NMT.

SELECTED PUBLICATIONS

- Andreas Grivas, **Nikolay Bogoychev**, and Adam Lopez (2022). "Low-Rank Softmax Can Have Unargmaxable Classes in Theory but Rarely in Practice". In: *ACL*
- **Nikolay Bogoychev** (2021). "Not all parameters are born equal: Attention is mostly what you need". In: *BlackboxNLP*
- **Nikolay Bogoychev**, Jelmer Van der Linde, and Kenneth Heafield (2021). "TranslateLocally: Blazing-fast translation running on the local CPU". in: *EMNLP*
- **Nikolay Bogoychev** and Pinzhen Chen (2021). "The Highs and Lows of Simple Lexical Domain Adaptation Approaches for Neural Machine Translation". In: *Insights from Negative Results in NLP*
- Alham Fikri Aji, **Nikolay Bogoychev**, et al. (July 2020). "In Neural Machine Translation, What Does Transfer Learning Transfer?" In: *ACL*

SKILLS

C++ python CUDA C gdb

x86 assembly java

Machine Translation NLP

ACTIVITIES

Open Source Software

I am a firm supporter and believer in Linux and Open source software and I have contributed to a number of open source projects, among which **python**, **wine**, **archlinux**, **marian-nmt**, **moses**.

I am also the author of:

- **TranslateLocally**, a cross-platform fast and secure offline translation client.
- **ProbingPT**, a fast phrase table for statistical machine translation
- **gLM** a GPU n -gram language model
- **imageSelector** a cross-platform QT photo library organiser

Privacy driven Translation

I am the principal developer of **TranslateLocally**, a cross-platform fast and secure offline translation client.

Give it a try!

<https://nbogoychev.com/>

bringing-machine-translation-to-the-users/

Consulting business

I am hireable as a consultant, via my company **NickOptimises Ltd**

nbogoychev.com/company/

Blog

I maintain a blog about Chinese characters, research and coding.

nbogoychev.com

- **Nikolay Bogoychev**, Roman Grundkiewicz, et al. (July 2020). "Edinburgh's Submissions to the 2020 Machine Translation Efficiency Task". In: *WNGT*
- **Nikolay Bogoychev**, Marcin Junczys-Dowmunt, et al. (Oct. 2018). "Accelerating Asynchronous Stochastic Gradient Descent for Neural Machine Translation". In: *EMNLP*
- **Nikolay Bogoychev** and Adam Lopez (2016). "N-gram language models for massively parallel devices". In: *ACL*

TALKS AND TUTORIALS

Efficient Machine translation tutorial

MT Marathon

📅 11/2021

📍 Edinburgh, United Kingdom

nbogoychev.com/efficient-machine-translation/

Privacy focused machine translation in Firefox

W3C Workshop on Web and Machine Learning

📅 09/2020

📍 Online

https://www.w3.org/2020/06/machine-learning-workshop/talks/privacy_focused_machine_translation_in_firefox.html

OTHER EXPERIENCE

Consultant

myLanguage

📅 12/2020

📍 Edinburgh, United Kingdom

- Porting x86_64 software to ARM.
 - Providing knowledge about neural machine translation model training.
-

Parallelism Research intern

Intel

📅 06/06/2017 - 06/10/2017

📍 Santa Clara, California, USA

Worked on Neural machine translation training on CPUs. Produced a detailed breakdown and the communication and computation costs of each operation in neural machine translation training. Developed methods for reducing internode communication during training.

Machine Learning Intern

Amazon

📅 01/06/2016 - 30/09/2016

📍 Berlin, Germany

Optimising Spark/Scala Bayesian inference framework.

Research Assistant

University of Edinburgh

📅 01/06/2014 - 01/12/2016

📍 Edinburgh, United Kingdom

Working on Statistical Machine translation

- Multithreading optimisation for Moses.
 - Bilingual language model feature function for Moses.
-

Language Intern

SwiftKey

📅 01/06/2012 - 01/09/2012

📍 London, United Kingdom

Working on predictive typing framework for mobile phones.

HOBBIES

Languages

I am an avid language learner and I am known to speak to a degree and with varying success Bulgarian, English, Spanish, German, Mandarin Chinese, Japanese and Portuguese.