

Education

- 2015 – present **University of Edinburgh**
PhD student at the Institute for Languages, Cognition and Computation
 Working on exploiting the raw computational power of GPUs to improve the performance of machine translation decoders. My supervisor is Adam Lopez. I am also a **tutor** and have been involved as a **TA** for second-year informatics courses in the areas NLP and Computer hardware.
- 2010 – 2014 **University of Edinburgh**
BSc Artificial Intelligence & Computer Science (First Class honours)
 Class rep for 4th year informatics students.
 Mentor of a group of students tasked with creating a football-playing robot

Selected publications

- 2016 Nikolay Bogoychev and Adam Lopez (2016). ***N*-gram language models for massively parallel devices**. In Proceedings of ACL, Berlin, Germany.
- 2016 Nikolay Bogoychev and Hieu Hoang (2016). **Fast and highly parallelizable phrase table for statistical machine translation**. In Proceedings of WMT, Berlin, Germany.

Related Employment

Intel, Santa Clara, USA

- June – Oct 2017 *Position:* Parallel computing lab intern
Duties included: Analyzing parallelism cost for RNNs, building theoretical computational models, improving RNN training performance in parallel setting
github: <https://github.com/marian-nmt/marian-dev>

Amazon, Berlin, Germany

- June – Oct 2016 *Position:* Machine Learning Intern
Duties included: Porting existing software to pyspark. Profiling and optimizing Scala-on-Spark stack.

The University of Edinburgh, Edinburgh, United Kingdom

- June – Dec 2014 *Position:* Research Associate
Duties included: Building an automatic performance testing environment; tracing down and fixing multithreading performance issues; GPGPU work on existing language model codebases; adding the BilingualLM feature function to Moses, the open source statistical machine translation decoder. Various Moses fixes.
github: <https://github.com/moses-smt/mosesdecoder/>

Technical Skills:

- Programming C/C++, CUDA, python, Java, Haskell, Clojure, Scala, JavaScript, Hadoop, Spark, Shell Script, Matlab, SQL, HTML, CSS
- OS Linux (Advanced knowledge and ability to set up and manage servers), Windows
- Linguistics/NLP Praat, Python NLTK library, Parsing experience, IPA Knowledge, Spectrogram reading

Notable projects

- gLM** gLM the GPU based language model is an ngram language model implementation that can

binarize an arpa encoded language model and then query it in batch manner on the GPU. It is coded in C++ and CUDA and is in active development and is eventually going to be used with a GPU based MT decoder.

github: <https://github.com/XapaJJaMnu/gLM>

ProbingPT

A C++ efficient machine translation phrase table storage and query developed as part of my honours project. Currently part of upstream Moses.

github <https://github.com/XapaJJaMnu/proj4/>

Vindication

A Haskell-based virtual court. The system is based on the Carneades argumentation framework where the user can provide a special case file with arguments and the system is going to decide whether guilt (or innocence) can be proved.

github: <https://github.com/XapaJJaMnu/vindication>

Markov Chain Monte Carlo

A Metropolis-Hastings sampler implemented in Matlab that is used for learning logistics regression model with 37 dimensions.

Code available on demand.

Plagiarism detector

A python based plagiarism detector that is able to detect both complete and near duplicates, as well as duplicated sections of otherwise different documents.

Code available on demand

Contributions to Open Source projects

wine

Several simple patches adding support for new GPUs, committed to master.

python

Add support for additional robots.txt parameters. Patch committed upstream.

issue and patch: <http://bugs.python.org/issue16099>

pyalienfx

Add support for new model, patch is used by several owners of the related hardware

issue and patch: <https://code.google.com/p/pyalienfx/issues/detail?id=30>

Moses

Implemented a Bilingual language model feature function a la *Devlin et al 2014*.

Debugged multithreading performance and committed various small fixes.

github <https://github.com/moses-smt/mosesdecoder/>

marian

Implemented various improvements catered towards multi-device training, as well as convergence with multi-device training.

github <https://github.com/marian-nmt/marian-dev>

Please check out my **github** account for more samples: <https://github.com/XapaJJaMnu>

More projects on **LinkedIn**: <http://www.linkedin.com/pub/nikolay-bogoychev/23/310/18>

Personal webpage: <http://nbogoychev.com>

Language Competences

Bulgarian (Mother Tongue), English (Fluent C2), Spanish (Fluent C1), German (B1), Mandarin Chinese (A2),

Awards:

- Edinburgh university Smart Data Hackaton – winner in the category “Best for travel” as part of a team.
- National Linguistics competition for 12th grade – national round.
- Attended (and won awards) in several translation competitions among those the “Young translators” competition organized by the European Commission; National History competition for 11th grade – 4th place.

Professional interests:

- Concurrency
- Multiprocessor programming
- GPGPU
- Machine learning
- AI, decision making and planning in particular
- Game theory

Hobbies

I have a strong passion for languages and writing systems in particular. I spend most of my spare time learning languages and looking up peculiar linguistics information. I enjoy reading, particularly Fantasy and Science fiction novels.

Sports

I enjoy playing Table tennis and Football. I have also practised Aikido.