Aleksandr Drozd

curriculum vitae

Tokyo, Japan
Shinagawa ku
mobile +81-80-4332-5304
alex@blackbird.pw
blackbird.pw

Brief Profile

My research interests lie at the intersections of **artificial intelligence**, especially areas like natural language processing and artificial life, and **high performance computing**. In addition to academic background I have strong software development skills and experience.

Education

2010-2014	Ph.D., Tokyo Institute of Technology, Graduate School of Information Science and Technology, Tokyo.
Thesis title:	"Memory-Conscious Optimizations for Sorting and Sequence Alignment for Massively Parallel Heterogeneous Architectures."
2000-2005	Specialist degree (M.Sc. equivalent), Moscow State University.
Thesis title:	"Semantic Pseudo-Code: Approach to Meaning-Base Search."

Work Experience

2018.05-onwards	**Invited Researcher** at *AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory (RWBC-OIL)*
2018.04-onwards	**Researcher** at *Tokyo Institute of Technology, School of Computing, Department of Mathematical and Computing Science*
Responsibilities:	scientific research in areas of artificial intelligence and high performance computing.
2014.04-2018.03	**Researcher** at *Tokyo Institute of Technology*, Global Scientific Information and Computing Center.
Responsibilities:	developing algorithms for extreme-scale data-intensive computing.
2005.06-2010.05	**Lecturer / Senior Lecturer** (from 2008) at *Moscow State University* (Sevastopol Branch, http://www.msusevastopol.net/), Programming Department.
Responsibilities:	teaching courses on Parallel Data Processing, Operating Systems and Computer Graphics.
2006.09-2009.06	**Software Architect and Developer** at *Outsourcing Ukraine* (http://www.outsourcing-ukraine.com/), Sevastopol
Responsibilities:	design and implementation of commercial software ($C++/C\#$)

Fellowships and Grants

- JSPS KAKENHI Grant number JP17K12739 adopted FY 2017: "Corpora on Demand: Scalable Methods of Obtaining Linguistic Data".
- Japanese Government (Monbukagakusho) scholarship for conducting PhD research 2010-2014.

Research Interests

My current focus is on the intersection of the high performance computing (**HPC**) and **intelligent data processing** in various applied tasks. The areas in which I have worked (both by myself and in collaboration with experts in these areas) include:

- *Natural language processing and computational linguistics*: work in the context of vector space models framework from engineering high-performance construction of word embedding to applying various deep learning method for text understanding.
- *Other aspects of artificial intelligence*: large scale deep learning for video recognition, artificial life modeling, swarming behaviour, social simulations.

In the past I have also been working in such areas as

- *Core HPC methods*: algorithmic kernels for large scale data processing, e.g. distributed sorting
- *Computational biology*: high-performance processing of big genomic data on accelerators

Software Development and Other Relevant Skills

I am a passionate programmer who does a fair amount of coding for research and sometimes for fun.

- **Coding/Software Development**: My experience as a developer of commercial software gave me such skills as object oriented design, patterns and development processes.
- **C, C++** (including C++14 standard), along with such libraries and tools for parallel programming as **CUDA, OpenMP, MPI, OpenCL, TBB**, etc for performance-critical parts.
- **Python** for everything else: high level scripting, quick prototyping and such. Being open-source enthusiast I'm trying to contribute back to the Python ecosystem.
- I have experience with databases (SQL and noSQL), web technologies and version control systems, computer algebra and publishing systems.
- I use **machine learning** extensively - from basic statistical analysis methods to artificial neural networks.

Other Related Activities

Organized workshops and tutorials:

- "Deep Learning from HPC Perspectives: Opportunities and Challenges" Mini-Symposium at SIAM PP 2018 conference.
- "Distributional Compositional Semantics in the Age of Word Embeddings: Tasks, Resources and Methodology". Tutorial 4 at LREC 2018 conference.

Other activities:

- I have served as a program committee member of a number of conferences and workshops, including NAACL, *SEM, SC, ISC, PARCO among others.
- From 2017 I serve as one of the organizers of "Tokyo Machine Learning Gym" meetup.

Languages

Russian	native
English	fluent
Japanese	conversational
Ukrainian	conversational

Interests

Music	I play cello, guitar, and sing
Photography	Taking pictures of people and events http://nightwind.in
Sport	Hiking, mountaineering, martial arts

Selected Publications

- Marzena Karpinska, Bofang Li, Anna Rogers and Aleksandr Drozd. Subcharacter Information in Japanese Embeddings: When Is It Worth It?. *In Proceedings of the Workshop on Relevance of Linguistic Structure in Neural Architectures for NLP (RELNLP) 2018 at ACL 2018*. Melbourne, Australia. To appear.
- Bofang Li and Aleksandr Drozd. Subword-Level Composition Functions for Learning Word Embeddings. *Proceedings of The 2nd Workshop on Subword and Character level models in NLP (SCLeM)* at NAACL 2018. To appear.
- Bofang Li, Tao Liu, Zhe Zhao, Buzhou Tang, Aleksandr Drozd, Anna Rogers and Xiaoyong Du. Investigating Different Syntactic Context Types and Context Representations for Learning Word Embeddings.*Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. pp 2421–2431.
- Anna Rogers, Aleksandr Drozd and Bofang Li. The (too Many) Problems of Analogical Reasoning with Word Vectors. *In Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)*, Association for Computational Linguistics, pp 135–148, Vancouver, Canada.
- Aleksandr Drozd, Anna Gladkova, Satoshi Matsuoka. Word Embeddings, Analogies, and Machine Learning: Beyond King - Man + Woman = Queen. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp 3519–3530, Osaka, Japan, December 11-17 2016*
- Mateusz Bysiek, Aleksandr Drozd and Satoshi Matsuoka. Migrating Legacy Fortran to Python While Retaining Fortran-Level Performance through Transpilation and Type Hints. Proceedings of PyHPC 16: the 6th Workshop on Python for High-Performance and Scientific Computing. pp 9-18.
- Aleksandr Drozd, Anna Gladkova, Satoshi Matsuoka. Discovering Aspectual Classes of Russian Verbs in Untagged Large Corpora. *The 2015 IEEE International Conference on Data Science and Data Intensive Systems (DSDIS 2015), At Sydney, Australia, Dec 2015, pp 61 - 68.*
- Aleksandr Drozd, Anna Gladkova, Satoshi Matsuoka. Python, Performance and Natural Language Processing. *5th Workshop on Python for High-Performance and Scientific Computing, at Austin, Texas, USA, Nov 2015 in conjunction with SC15, pp 1-10.*
- Aleksandr Drozd, Olaf Witkowski, Satoshi Matsuoka, Takashi Ikegami. Signal-Driven Swarming: A Parallel Implementation of Evolved Autonomous Agents to Perform A Foraging Task *Proceedings of SWARM 2015 - The First International Symposium on Swarm Behavior and Bio-Inspired Robotics, Kyoto, Oct 2015.*
- Aleksandr Drozd, Satoshi Matsuoka. HPC and Interactive Big Data Analytics: Case Study of Distributional Semantics. *Proceedings of IPSJ SIG Technical Reports 2014-HPC-146, Naha, Oct 2014.*
- Hideyuki Shamoto, Koichi Shirahata, Aleksandr Drozd, Hitoshi Sato, Satoshi Matsuoka. Large-scale Distributed Sorting for GPU-based Heterogeneous Supercomputers. *Proceedings of 2014 IEEE Conference of Big Data, October 2014, pp 510 - 518.*
- Aleksandr Drozd, Miquel Pericàs, Satoshi Matsuoka. Efficient String Sorting on Multi- and Many-Core Architectures *in Proceedings of IEEE 3rd International Congress on Big Data (2014), Anchorage, AK, pp 637 - 644.*
- Aleksandr Drozd, Naoya Maruyama, Satoshi Matsuoka. Sequence Alignment on Massively Parallel Heterogeneous Systems in *Proceedings of IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum (2012), pp. 2498 - 2501, ISBN 978-1-4673-0974-5*

- Aleksandr Drozd, Naoya Maruyama, Satoshi Matsuoka. A Multi GPU Read Alignment Algorithm with Model-Based Performance Optimization, *Springer's Lecture Notes in Computer Science N7851 (2012), pages 270-277.*
- Aleksandr Drozd, Naoya Maruyama, Satoshi Matsuoka. Fast GPU Read Alignment with Burrows Wheeler Transform Based Index, *In Companion Proceeding of SC'11 Conference on High Performance Computing Networking, Storage and Analysis, 2011, Pages 21-22 .*
- Aleksandr Drozd, Naoya Maruyama, Satoshi Matsuoka. Fast Read Alignment with Burrows Wheeler Transform: the GPU Perspective, In Proceedings of the 24th Summer United Workshops on Parallel, Distributed, and Cooperative Processing (SWoPP 2011) , August 2011.
- Anna Gladkova and Aleksandr Drozd. Towards Easier Querying of XML-based Linguistic Corpora, *Taurida Bulletin of Mathematics and Informatics. #2, 2009, pages 71-77*