Viacheslav Ivanov
05/28/1997

ivanov.vv@phystech.edu
github:ivanov-v-v
linkedin:vvivanov

EDUCATION

Moscow Institute Of Physics And Technology, *GPA: 9.44/10; 13th out of ≈1000 in my study year*
Bachelor of Science, Applied Mathematics & Computer Science   *08.2016 — 08.2020*
Relevant courses: Bayesian statistics, applied statistics, optimization, machine learning, deep learning, stochastic processes, algorithms and data structures, concurrent programming, distributed programming

LANGUAGES

English (C1, TOEFL iBT: 111), German (A2), Russian (native), Ukrainian (native)

RESEARCH EXPERIENCE

Research Internship, Computational Biology, DKFZ/EMBL, *01.07.2019 — 01.06.2020 (ongoing)*
Supervisors: Dr Oliver Stegle (group leader); Dr Hana Susak, Dr Nicola Casiraghi, Dr Yuanhua Huang

I am currently taking part in a big project involving 3 postdoctoral researchers. We are using **Bayesian inference** to investigate the composition and dynamics of tumor clones in **medulloblastoma**, pediatric brain tumour. My part of the project extends **"Cardelino"**, structured Bayesian model for **single-cell** data analysis which has been recently accepted in **Nature Methods**. I am working on the method and am responsible for its implementation. Currently, we jointly analyze DNA-seq and RNA-seq datasets extracted from the same patient in single-cell resolution. We construct lineage tree from **copy-number variations** inferred from DNA-seq in individual cells. To be more precise, we define a measure of **evolutionary distance** and build the minimum spanning tree. We define clonal lines as clusters in that tree. Then, we map cells with known RNA-seq profiles onto clonal lines. We currently use **Gibbs sampling** to learn for each cell the probabilities of belonging to each particular clonal line. Simulations show, that the current approach works, but it is slow, so we are going to switch to the **variational inference** — optimization-based methods in probabilistic modeling. In this project, I've already written more than **10000 lines** of code on my own, mostly in **Python** and **C++**. Given size and complexity of single-cell datasets, I have to write the code that is time- and memory-efficient. I am actively learning about the best practices of **reproducible research** and **test-driven development** and try to apply those in my project.

Mutual analysis of interaction networks and quantitative trait loci for yeast, *2018-present*:
«**Modern Combinatorics and Network Science**» **Lab** of Prof A. Raigorodsky, MIPT.
Supervisor: Yuri Pritykin, PhD (Princeton), Research Scholar at MKSCC

Quantitative trait loci or **QTLs** are genomic regions that serve as **regulatory mechanisms** that influence expression of genes. They provide a link between the genotype and phenotype. They are attracting a lot of attention in recent years due to developments in **high-throughouput sequencing** technologies. In this project we had an aim to tie together **eQTLs** — loci that regulate mRNA expression, — and **pQTLs** — same, but for protein expression — in budding yeast. Genetic architectures of eQTLs and pQTLs are very different, despite the obvious relationship between **transcription** and **translation** of the genes dictated by the central dogma of molecular biology. There is an ongoing debate on this matter in contemporary research. We decided to make a contribution by comparing genetic architectures underlying eQTLs and pQTLs linked to **functional modules** — sets of interacting genes participating in a particular cell process — by carrying out **genome-wide association studies** on **next generation sequencing** data. Experiments show, that QTLs on the level of modules are much more related than they are for the individual genes. Given that observation, we define and study **module QTLs** — genomic regions that influence the expression of a functional module as a whole. This project was my first introduction to computational biology and taught me the basics of **high performance computing** and **bioinformatics**.

- I implemented different approaches to **QTL mapping** in yeast, from basic to state-of-the-art
- I integrated **interaction networks** into QTL analysis and implemented statistical tests using igraph.
- I learned how to tackle domain-specific difficulties arising from large-scale hypothesis testing using **FDR -correction** techniques (especially qvalue from Princeton).
- I learned how to write fast and memory-efficient scientific code using **numpy**, **scipy** and **pandas**.
- I practiced **parallel programming** and **data persistence** in Python.
- I utilized MIPT **supercomputing** capabilities, learned how to use schedulers and other basic infrastructure.
- I worked with databases of **genetic ontologies** in Python and R.

HONORS

- **Governmental Scholarship for Academic Excellence** *7th term, Autumn 2019*
  Awarded to top 10% of the final-year students with proven record of academic excellence.
- **Abramov Scholarship For Academic Excellence** *2nd term − 6th term, 2016 − 2020*
  Earned by top 10% students by cumulative GPA in their academic program in academic terms 2-6.
- **Russian Government Scholarship For International Students** *2016 − 2020*
  Was selected to become one of 3 Ukrainians to receive the full-coverage scholarship to study CS at the best Russian universities and got enrolled to MIPT directly, without entrance examination.
- **All-Russian Subject Competition in English among Technical Universities** *Autumn 2019*
  Our team of three students from the same study program was awarded the first' place country-wise.
- **Future Biotech Winter Retreat "Genome function, editing and therapy"** *Winter 2019*
  Became one of 70 young researchers selected to participate in top biotech winter school in Russia sponsored by companies like AstraZeneca, Biocad, GE Healthcare etc.
- **Scholarship For Academic Excellence from the Governor of Moscow Region** *Autumn 2017*
  Awarded each term to excellent students for promising achievements in scientific activities.
- **Summer School in Bioinformatics** by Russian Bioinformatics Institute *Summer 2017*
  Became one of 50 CS majors selected to participate. Was a member of a hackathon-winning team.
- **ACM ICPC Moscow Subregional Contest (1/4 World Finals)** *2017*
  Our team ranked 11th among 301 participating teams and 7th at home university.
- **Moscow International Workshops in Competitive Programming** *Autumn 2015, Spring 2016*
  Two-times participant of the leading Russian university-level competitive programming bootcamp.
- **Winter School in Computer Science** *Winter 2015*
  Participated in one of the most prestigious bootcamps targeted on preparing the best young programmers from Russian-speaking countries for National and International subject competitions in computer science.
- **All-Ukrainian Chemistry Olympiad** — double awardee, triple winner of regional stage
- **All-Ukrainian Tournament of Young Chemists** — 2nd place

TEACHING EXPERIENCE

- **TA at «Theory and practice of concurrent programming» course at MIPT** *Spring term 2019*
  My responsibilities included weekly code reviews on GitLab and knowledge assessment sessions.
- **TA at «Mathematical statistics» and «Applied statistics» courses at MIPT** *Fall term 2019*
  Authored a primer in scientific computing for data scientists.Topics included: environment setup, Python data science ecosystem (numpy, scipy, pandas), efficient Python programming (JIT, Cython, data persistence, parallel and distributed computing), advanced visualizations (IPyWidgets, bokeh, bqplot), reproducible workflows (SnakeMake, pachyderm), cloud services (Google Colaboratory, Amazon AWS).
- **TA at «Algorithms and data structures» course at MIPT** *Fall term 2019*
  Weekly code reviews: compliance with code style, algorithm correctness assessment etc.

REFERENCES

Dr Oliver Stegle, group leader at EMBL, EMBL-EBI and DKFZ   oliver.stegle@embl.de
Dr Yuri Pritykin (thesis supervisor), research scholar at MSKCC   yuri.pritykin@gmail.com
Federal Prof. Andrei Raygorodsky, head of department at MIPT   mraigor@yandex.ru