

# PARHAM KAZEMI

pkazemi3@gmail.com

parham-k.github.io

github.com/parham-k

**Keywords** — Bioinformatics, Algorithms, Natural Language Processing, Deep Learning  
**Skills** — C++, Python, Java, PyTorch, Keras

## EDUCATION

---

**PhD in Bioinformatics** - University of British Columbia 2021 - Present

- Supervised by Dr. Inanc Birol
- Member of the Bioinformatics Technology Lab at BC Cancer's Genome Sciences Centre

**MSc in Computer Engineering (AI)** - University of Isfahan 2019 - 2021

- Supervised by Dr. Hossein Karshenas
- Thesis: Deep Reinforcement Learning for Training Intelligent Agents in Natural Language Environments
- GPA: 18.42/20

**BSc in Computer Engineering (Software)** - University of Isfahan 2015 - 2019

- Supervised by Dr. Afsaneh Fatemi
- Thesis: Predicting Persian Twitter Users' MBTI Personality Using Text Mining Methods
- GPA: 18.25/20
- Received CEng Faculty's Distinguished Student Award
- Participated in ICPC West Asia Regionals (2016, 2017 ranked 16th)
- Participated in Robocup IranOpen 2D Soccer Simulation League 2017
- Participated in National Computer Engineering Olympiad Finals

## EXPERIENCE

---

**Graduate Research Assistant** - Canada's Michael Smith Genome Sciences Centre 2021 - Present

- Member of the Bioinformatics Technology Lab ([birollab.ca](http://birollab.ca))
- Developer of ntHash2
- Maintainer of btllib (Bioinformatics Technology Lab common code library)

**Volunteer** - Vancouver Bioinformatics User Group, VanBUG 2023 - Present

- Member of the VanBUG development group ([vanbug.org](http://vanbug.org))

**Teaching Assistant** - University of Isfahan 2016 - 2020

- Advanced Programming, Algorithm Design, Artificial Intelligence, Computer Fundamentals, Data Structures, Discrete Mathematics, Technical English

**ACM Students Chapter Member** - University of Isfahan 2017 - 2019

- Team coach at ICPC West Asia Regionals 2020
- Python course instructor

## PROJECTS

---

**ntHash2: recursive spaced seed hashing for nucleotide sequences**  
Lead developer, publication's first author

[github.com/bcgsc/ntHash](https://github.com/bcgsc/ntHash)

**University of Isfahan's Alumni Social Website**  
Backend developer

## PUBLICATIONS

---

Warren, R. L., Coombe, L., Wong, J., **Kazemi, P.**, & Birol, I. (2024). Human ancestry inference at scale, from genomic data. *bioRxiv*, 2024-03.

Coombe, L., **Kazemi, P.**, Wong, J., Birol, I., & Warren, R. L. (2024). Multi-genome synteny detection using minimizer graph mappings. *bioRxiv*, 2024-02.

Wong, J., **Kazemi, P.**, Coombe, L., Warren, R. L., & Birol, I. (11 2023). aaHash: recursive amino acid sequence hashing. *Bioinformatics Advances*, vbad162. doi:10.1093/bioadv/vbad162

Nikolić, V., **Kazemi, P.**, Coombe, L., Wong, J., Afshinfard, A., Chu, J., Warren, R. L., & Birol, I. (2022). btllib: A C++ library with Python interface for efficient genomic sequence processing. *Journal of Open Source Software*, 7(79), 4720. doi.org/10.21105/joss.04720

**Kazemi, P.**, Wong, J., Nikolić, V., Mohamadi, H., Warren, R. L., & Birol, I. (2022). ntHash2: recursive spaced seed hashing for nucleotide sequences. *Bioinformatics*, 38(20), 4812–4813. doi.org/10.1093/bioinformatics/btac564

**Kazemi, P.**, & Karshenas, H. (2021). Fuzzy Word Sense Induction and Disambiguation. *IEEE Transactions on Fuzzy Systems*, 30(9), 3918–3927. doi.org/10.1109/tfuzz.2021.3133905

## PRESENTATIONS

---

**Kazemi, P.,** Coombe, L., Warren, R. L., & Birol, I. (2024). AIEdit: Genome Assembly Polishing using Machine Learning and Spaced Seeds. Poster in BIG Research Day 24 (Vancouver, Canada)

**Kazemi, P.,** Mohamadi, H., Chu, J., Coombe, L., Warren, R. L., & Birol, I. (2023). ntHits: streaming through raw sequencing data to profile and filter k-mers with selected multiplicities. Poster in the 13th RECOMB Satellite Conference on Biological Sequence Analysis (Istanbul, Turkey)

**Kazemi, P.** (2020). Natural language processing: what, how, and why. Presentation at the University of Isfahan's Annual Research Week (Esfahan, Iran)