

# Alex (Oleksandr) Polozov

One Microsoft Way, Redmond, WA 98052-6339, United States of America  
[polozov@microsoft.com](mailto:polozov@microsoft.com) • +1 (425) 623-4121 • <https://www.alexpolozov.com>

## EDUCATION

---

### Ph.D. in Computer Science

2012-2017

University of Washington, Seattle, USA

Advisors: Dr. Sumit Gulwani & Prof. Zoran Popović

Ph.D. Thesis: "A Framework for Mass-Market Inductive Program Synthesis"

### B.Math. in System Analysis

2008-2012

National Technical University of Ukraine "Kyiv Polytechnic Institute"

Advisor: Prof. Yuriy Tymoshenko

Thesis: "Structure and Term Prediction for Mathematical Text"

## PUBLICATIONS

---

### Books, book chapters, journal articles, monographs

- [1] S. Gulwani, O. Polozov, and R. Singh, "Program Synthesis", in *Foundations and Trends® in Programming Languages*: Vol. 4: No. 1-2, pp 1-119, 2017.
- [2] O. Polozov. "A Framework for Mass-Market Inductive Program Synthesis." PhD dissertation, 2017.

### Peer-reviewed conference publications, talks

- [3] R. Shin, M. Allamanis, M. Brockschmidt, and O. Polozov. "Program Synthesis and Semantic Parsing with Learned Code Idioms", in *33<sup>rd</sup> Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- [4] M. Brockschmidt, M. Allamanis, A. L. Gaunt, and O. Polozov. "Generative Code Modeling with Graphs", in *7<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2019.
- [5] C. Wang, P.-S. Huang, O. Polozov, M. Brockschmidt, and R. Singh, "Execution-Guided Neural Program Decoding", in *2<sup>nd</sup> Workshop on Neural Abstract Machines & Program Induction (NAMPI)*, 2018.
- [6] S. Padhi, P. Jain, D. Perelman, O. Polozov, S. Gulwani, and T. Millstein, "FlashProfile: A Framework for Synthesizing Data Profiles", in *ACM SIGPLAN conference on Systems, Programming, Languages and Applications: Software for Humanity (SPLASH)*, 2018.
- [7] A. Kalyan, A. Mohta, O. Polozov, D. Batra, P. Jain, and S. Gulwani, "Neural-Guided Deductive Search for Real-Time Program Synthesis from Examples", in *6<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2018.
- [8] R. Rolim, G. Soares, L. D'Antoni, O. Polozov, S. Gulwani, R. Gheyi, R. Suzuki, and B. Hartmann, "Learning syntactic program transformations from examples", in *39<sup>th</sup> International Conference on Software Engineering (ICSE)*, 2017.
- [9] O. Polozov and S. Gulwani, "Program synthesis in the industrial world: inductive, incremental, interactive", in *5<sup>th</sup> Workshop on Program Synthesis (SYNT)*, 2016.
- [10] M. Mayer, G. Soares, M. Grechkin, V. Le, M. Marron, O. Polozov, R. Singh, B. Zorn, and S. Gulwani, "User interaction models for disambiguation in programming by example", in *ACM Symposium on User Interface Software and Technology (UIST)*, 2015.
- [11] O. Polozov and S. Gulwani, "FlashMeta: a framework for inductive program synthesis", in *ACM Conference on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA)*, 2015.
- [12] O. Polozov, E. O'Rourke, A. M. Smith, L. Zettlemoyer, S. Gulwani, and Z. Popović, "Personalized mathematical word problem generation", in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.
- [13] O. Polozov and S. Gulwani, "LaSEWeb: automating search strategies over semi-structured web data", in *ACM Conference on Knowledge Discovery and Data Mining (KDD)*, 2014.

## Preprints, technical reports, papers under review

- [14] B. Wang, R. Shin, X. Liu, O. Polozov, and M. Richardson. "RAT-SQL: Relation-Aware Schema Encoding and Linking for Text-to-SQL Parsers". *arXiv preprint arXiv:1911.04942*, 2019.
- [15] V. J. Hellendoorn, P. T. Devanbu, O. Polozov, and M. Marron. "Are My Invariants Valid? A Learning Approach". *arXiv preprint arXiv: 1903.06089*, 2019.
- [16] T. Shi, K. Tatwawadi, K. Chakrabarti, Y. Mao, O. Polozov, and W. Chen. "IncSQL: Training Incremental Text-to-SQL Parsers with Non-Deterministic Oracles". *arXiv preprint arXiv: 1809.05054*, 2018.
- [17] C. Wang, K. Tatwawadi, M. Brockschmidt, P.-S. Huang, Y. Mao, O. Polozov, and R. Singh. "Robust Text-to-SQL Generation with Execution-Guided Decoding". *arXiv preprint arXiv:1807.03100*, 2018.
- [18] V. Le, D. Perelman, O. Polozov, M. Raza, A. Udupa, and S. Gulwani, "Interactive Program Synthesis". *arXiv preprint arXiv:1703.03539*, 2017.
- [19] O. Polozov, S. Gulwani, and S. Rajamani, "Structure and term prediction for mathematical text". Tech. Rep. MSR-TR-2012-7, 2012.

## EMPLOYMENT

---

### Senior Researcher

*August 2017 – present*

*Microsoft Research AI, Redmond, USA*

Working on the neural program synthesis from input-output examples and natural language, intersections of machine learning and formal methods, and neuro-symbolic techniques in artificial intelligence. Overseeing various product applications of neuro-symbolic reasoning for question answering, spreadsheet assistance, software engineering, and data science.

### Research & Software Engineer – Contractor

*October 2014 – April 2017*

*Populus Group at Microsoft, Redmond, USA*

A founding member of the [Program Synthesis using Examples \(PROSE\)](#) R&D team at the Microsoft Data Group. Developed the PROSE framework for automatic synthesis of data wrangling scripts from incomplete specifications (input-output examples, constraints, demonstrations). It unified and generalized 12+ prior publications (5 years of prior work) in the field of programming by examples, allowing one to develop a by-example technology in 10 weeks instead of 10 months. Its applications are deployed in multiple Microsoft products:

- FlashFill: string transformations by example in Excel,
- ConvertFrom-String and Convert-String cmdlets in PowerShell,
- Text extraction in Azure Operational Management Suite,
- Web & email processing in Exchange and Cortana.

I have been working with a team of 10+ researchers and engineers at Microsoft concurrently with completing my Ph.D.

### Research Intern

*March 2014 – September 2014*

*Microsoft Research, Redmond, USA*

Designed and developed a modular algorithmic framework for automatic synthesis of programs in domain-specific languages from inductive specifications. Generalizes 5 years of prior work in programming by examples done by the Sumit Gulwani's group and collaborators. This work became a foundation for the Microsoft PROSE team (see above).

### Research Intern

*June 2013 – September 2013*

*Microsoft Research, Redmond, USA*

Designed a declarative language and an efficient interpreter for designing search strategies for microsegment queries based on linguistic predicates and semi-structured data on the Web.

### Research Intern

*June 2012 – September 2012*

*Microsoft Research, Redmond, USA*

Developed a language and an algorithm for 2D data visualization by example and data extraction from semi-structured images.

### Software Development & Research Practice Intern

*November 2011 – May 2012*

*Yandex, Kyiv, Ukraine*

Built a dictionary-based morphological engine with inflection prediction for Russian, Ukrainian, and English.

Microsoft, Redmond, USA

Team: Office Labs

Built a home/work location prediction with further exploration of user scenarios for a mobile digital assistant.

## SKILLS

---

<i>Programming Technologies</i>	C#, C/C++, Python, TypeScript/JavaScript, Wolfram Language, Java, Scala, F#, AnsProlog
<i>Theory</i>	TensorFlow, PyTorch, Potassco ASP toolkit, Z3 constraint solver, Rosette, Web development
<i>Languages</i>	<ul style="list-style-type: none"><li>• Deep learning; recurrent, convolutional, graph neural networks; neuro-symbolic techniques</li><li>• Natural language processing, semantic parsing, Transformers</li><li>• Program synthesis, programming languages, software engineering</li><li>• Answer set programming, SAT/SMT constraint solving, formal methods</li></ul>
<i>Languages</i>	English (fluent), Russian (native), Ukrainian (native), German (basic)

## INVITED TALKS

---

- **From Examples to Natural Language and Back**  
*“State of the Art in Program Synthesis” Workshop, 2019.*
- **Program Understanding, Synthesis, and Generalization with Graph Neural Networks**  
Invited talk at the *Learning and Reasoning with Graph-Structured Representations Workshop at ICML 2019.*
- **Neuro-Symbolic Program Synthesis**  
Invited talks at Technion University and Tel-Aviv University, Israel, 2018.
- **Program Synthesis via Neural-Guided Deductive Search**  
Invited talk at the *Machine Learning + Programming Languages Workshop 2018.*
- **Bringing Program Synthesis to the Mass Markets**  
*Approaches and Applications of Inductive Programming, Dagstuhl Seminar 2017.*
- **Data Processing Using Input-Output Examples with Microsoft PROSE SDK**  
**Creating Programming-By-Example Features in Arbitrary Domains with Microsoft PROSE SDK**  
A series of tutorials hosted as part of the *Machine Learning & Data Science Conference (MLADS) 2017.*
- **PROSE: Inductive Program Synthesis for the Mass Markets**  
Invited talk & Hackathon for graduate students at *UC Berkeley, January 2017.*
- **Automated Program Synthesis**  
Invited talk at the *Human-Like Computing Machine Intelligence Workshop (MI20-HLC), October 2016.*
- **PROSE: Growing Program Synthesis to Industrial Applications**  
*UW PLSE Research Retreat Workshop, September 2016.*
- **PROSE: Programming using Examples**  
Co-lectured with Sumit Gulwani. Invited tutorial at the *ACM SIGPLAN Conference on Programming Languages Design & Implementation (PLDI) 2016.*
- **Programming by Examples**  
Co-lectured with Sumit Gulwani at *Marktoberdorf Summer School 2015.*
- **Personalized Mathematical Word Problem Generation**  
*Approaches and Applications of Inductive Programming, Dagstuhl Seminar 2013.*

## SERVICE

---

- **PC & Review:** IJCAI-ECAI 2018, GPCE 2018, ML4PL 2018, ESEC/FSE SRC 2018, AAI 2019, ICML 2019, IJCAI 2019, NeurIPS 2019, ICLR 2020.
- **Ad-hoc reviewer:** Theoretical Computer Science Journal (2015), CAV 2018.

## PATENTS

---

- [1] Zorn, B.G., Brockschmidt, M.M.J., Choudhury, P., Polozov, O., Singh, R. and Padhi, S., Microsoft Technology Licensing LLC, 2020. *Systems, methods, and computer-readable media for improved table identification using a neural network*. U.S. Patent Application 16/034,447.
- [2] Polozov, O., Gulwani, S., Jain, P., Vijayakumar, A. K. and Mohta, A., Microsoft Technology Licensing LLC, 2019. *Neural-guided deductive search for program synthesis*. U.S. Patent Application 16/019,280.
- [3] Gulwani, S., Jain, P., Perelman, D.A., Padhi, S. and Polozov, O., Microsoft Technology Licensing LLC, 2019. *Syntactic profiling of alphanumeric strings*. U.S. Patent Application 15/663,575.
- [4] Gulwani, S., Zorn, B.G., Singh, R., Marron, M., Polozov, O., Le, V.M., Mayer, M., Soares, G.A. and Grechkin, M., Microsoft Technology Licensing LLC, 2017. *User interaction models for disambiguation in programming-by-example*. U.S. Patent 9,891,895.
- [5] Gulwani, S., Polozov, O. and Tiwary, S.K., Microsoft Technology Licensing LLC, 2015. *Control of automated tasks executed over search engine results*. U.S. Patent Application 14/619,702.
- [6] Gulwani, S. and Polozov, O., Microsoft Technology Licensing LLC, 2015. *Parsing and rendering structured images*. U.S. Patent 9,031,894.

## TEACHING EXPERIENCE

---

### CSEP 590C: Domain-Specific Languages

Spring 2016, Spring 2017

Teaching Assistant & Lecturer

Together with Prof. Rastislav Bodik and Pavel Panchekha, we co-designed and taught a graduate course on DSLs. Audience: professional software engineers with multiple years of industry experience. Course content includes foundations of compiler/interpreter development, a collection of well-known DSLs (D3.js, Mustache, Hadoop, React.js, Rx), program synthesis in PROSE, and lessons on DSL design.

### Functional Programming

2011 – 2012

Lecturer

Self-designed and taught an optional 2-semester course for undergraduate Applied Math students. Covered basic FP concepts, parallel programming with monoids and MapReduce, purely functional data structures, monads and type classes, lambda calculus.

### Algorithms and Data Structures

2009 – 2012

ACM ICPC lecturer & team coach

Taught advanced algorithms and data structures for undergraduate ACM ICPC teams, with focus on performance and programming competitions.

## REFERENCES

---

- Rastislav Bodik, professor: [bodik@cs.washington.edu](mailto:bodik@cs.washington.edu)
- Sumit Gulwani, partner research manager: [sumitg@microsoft.com](mailto:sumitg@microsoft.com)
- Rishabh Singh, senior research scientist: [rising@google.com](mailto:rising@google.com)
- Marc Brockschmidt, senior principal researcher: [mabrocks@microsoft.com](mailto:mabrocks@microsoft.com)
- Susan Dumais, technical fellow: [sdumais@microsoft.com](mailto:sdumais@microsoft.com)
- Ben Zorn, partner researcher: [zorn@microsoft.com](mailto:zorn@microsoft.com)