

Tobin Yehle

EXPERIENCE

WORK **3M HIS Data Science Lab – Software Developer (2017-Now)**

Medical Record Parsing – Built a scalable tool for medical record extraction and transformation using Scala and Spark. The tool uses functional programming techniques to ensure distributed failures are handled properly. The tool allowed the data science lab, for the first time, to access production data for training models.

Predictive Model Deployment – Designed the first deployment of a machine learning model from the data science lab. Requests are served using an AWS lambda function written in Python and a model in S3.

Fusion-io – Software Developer (2012-2014)

Build and QA automation. Contributed to a refactor of the test infrastructure.

CONSULTING **U of U Department of Education – Freelance Software Developer (2018-Now)**

Built a web app for online therapy allowing better teaching tools and faster responses for patients.

EDUCATION **Recurse Center – Self-directed programming retreat (2018)**

Compiler using LLVM – Built a compiler in Haskell using LLVM that compiles a non-trivial language to x86. The compiler implements closure conversion without garbage collection and converts all expressions to SSA.

University of Utah – Student & Research Assistant (2011-2016)

Parsing with Derivatives – Senior thesis research under Dr. Vivek Srikumar to extend the derivative parsing algorithm to English. The derivative parser is left to right, which we leveraged to save a state after parsing each token. This state could then be loaded from a cache if a matching sentence prefix was seen in the future, enabling increased performance on large datasets.

Python Compiler – A compiler for all of Python 3 written in Racket. It lexed and parsed all of Python 3 and implemented two desugaring passes, eliminating most syntactic constructs. The resulting IR needed one more desugaring pass before code generation in assembly language.

Clustering of Suicide Cases – Project with the Department of Psychiatry at the U of U to study familial clustering of suicide cases. I used network clustering algorithms in Python to find subsets of familial suicide cases with distinct demographic and/or diagnostic characteristics. The research group is using the clusters for further analysis.

Florida Institute of Technology – Research Assistant (2014)

Spatial Structure of Crime – Competitive NSF-funded program hosted by FIT where I worked with a team to use complex networks on police data to uncover structure in the timing and location events. We built networks with links between spatially or temporally close events, and used network clustering algorithms to find interesting regions, and allow new types of visualizations. There are two publications for this work, White et al. and Oliveira et al.

Publications

From Criminal Spheres of Familiarity to Crime Networks: Marcos A. C. Oliveira, Hugo Serrano Barbosa Filho, Tobin Yehle, Sarah White, Ronaldo Menezes (2015)

The Spatial Structure of Crime in Urban Environments: Sarah White, Tobin Yehle, Hugo Serrano Barbosa Filho, Marcos A. C. Oliveira, Ronaldo Menezes (2014)

OTHER **Open Source Contributions**

Three patches to mypy, the Python type checker. Two merged, one in review. [#2965](#) [#3369](#) [#5617](#)

Type annotations for a Python package emulating Rust's `Result<T, E>` structure. [#4](#)

Misc Projects

Compiler targeting the lambda calculus – Written in Haskell, and compiles any Sum of Products data type.

Sherlock – A question answering system that used classic NLP techniques and Stanford's NLP tools.

Dominion AI – Class project written in Haskell that won our class tournament.

Sudoku Solver – Compiled constraints to SAT and used Z3 to solve the puzzle.

Sheet-music optical character recognition – Used classic image recognition techniques in Python.

Genetic algorithm convergence – Project to get faster convergence of genetic algorithms using momentum.

Physical simulations – Two simulations; one of [charged particles](#) and the other of [dynamical friction](#).

More – On [github](#) and [bitbucket](#)

EDUCATION

2011 – 2016 **University of Utah – Honors BS Computer Science, Magna cum Laude**

Undergrad Research Scholar

Completed tracks for Programming Languages, Artificial Intelligence, Information in Data, and Theory

Minors in Music & Astronomy

INTERESTS

PROGRAMMING **Functional Programming & Compilers** – Haskell, Scala, Racket, LLVM, λ -calculus
Distributed Systems & Machine Learning – Spark, Hadoop, Python, SQL

OTHER Classical Music (trumpet), Climbing, Pottery, Skiing, Biking, Hiking