Tobin Yehle

Experience

Work

3M HIS Data Science Lab – Senior Software Developer (May 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 gracefully. The tool can scale out to hundreds of millions of records, and for the first time allowed the data science lab access production data for training models.

Predictive Model Deployments – Designed the deployment of HIPAA compliant machine learning models for the data science lab. Older versions use AWS Lambda running SKLearn models stored in S3, and newer versions use AWS Sagemaker containers. These microservices run hundreds of thousands of predictions every day.

Data Science Infrastructure – Tools and security policy to allow the data scientists I work with to build clusters and train models while ensuring HIPAA compliance and secure access to protected data. This includes automatic generation of policy and templates for different data access permissions and different training needs (clusters, GPU instances, and/or notebooks).

Fusion-io – Software Developer (May 2012 – May 2014)

Build and QA automation. Refactored test infrastructure to run more smoothly in Jenkins.

Consulting

U of U Department of Education – Software Developer (March 2018 – Now)

Built a web app for online therapy supported by the NSF. This tool provides a better way to teach student therapists and will allow faster response times for patients.

Education

Recurse Center – Self-directed programming retreat (Jan 2018)

Compiler using LLVM – Built a compiler in Haskell using LLVM that compiles a non-trivial language to x86. The compiler implements closure conversion, garbage collection, and tail call optimization.

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

Parsing with Derivatives – Senior thesis research under Dr. Vivek Srikumar to extend the derivative parsing algorithm to English. This algorithm allows caching for increased performanc 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.

Florida Institute of Technology – Research Assistant (June 2014 – Aug 2014)

Competitive NSF-funded program hosted by FIT where we built networks of temporally and spatially related 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* (2015)**

Marcos A. C. Oliveira, Hugo Serrano Barbosa Filho, Tobin Yehle, Sarah White, Ronaldo Menezes

***The Spatial Structure of Crime in Urban Environments* (2014)**

Sarah White, Tobin Yehle, Hugo Serrano Barbosa Filho, Marcos A. C. Oliveira, Ronaldo Menezes

Open Source

Patches

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

Projects

Contributor on Aft – A brute force dynamic type inference engine for Python program fragments.

Control Systems Demo – An implementation of PID control in two dimensions.

Hardware Sunrise Alarm Clock – Hardware and firmware for a sunrise alarm clock.

3D Engine – In Scala and Rust that preserves angles and allows a full 360 degree field of view.

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

Machine Learning - Question answering system, genetic algorithm convergence, and sheet music OCR.

Physics simulations – One of charged particles and another of dynamical friction.

Qualifications

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

Hardware & Systems – Rust, C, Control Theory, Operating Systems

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

tobinyehle@gmail tobin.yehle.us tyehle