Pranjal Agrawal

pranx.devl@gmail.com • +(91) 914 071 3253 • github.com/forumulator • linkedin.com/in/forumulator

EDUCATION

Indian Institute of Technology, Guwahati — Assam, India
Bachelor of Technology — 2014 – July 2018 (Expected)
Computer Science and Engineering — 7.95/10

PROFESSIONAL EXPERIENCE

Software Development Intern — *May 2017— July 2017*
Amazon — *Bangalore, India*

- Created an interactive debugging tool (in the form of a Java webapp) to figure out problems with product images, like incorrect images
- Linearized Amazon's image ranking algorithms and presented it in a user friendly manner to be used interactively
- Used Amazon in-house database (Sable) APIs to simulate each image reactor. Added caches in the persistence layer to reduce Sable calls
- The debugger automatically reduced and resolved around 70 issues a month, thus saving 35 developer hours (about 1 week of SDE time)

Google Summer of Code Intern — *Summer 2017*
Ceph

- Created Cdobs, a relational database storage system to store Amazon S3 style objects. Added a simple command line interface
- Designed an architecture to decouple the Ceph RGW REST interface from the Ceph storage backend, while not changing the core code
- Implemented the architecture on the main class, using Cdobs as an alternative storage backend
- Reduced the Ceph RGW build time from 2 hours to 15 minutes (because the Ceph backend got decoupled from Ceph RGW)

Google Summer of Code Intern — *Summer 2016*
Python Software Foundation

- Re-architected a tiny microprocessor called Leros, from VHDL to a python based HDL language called MyHDL. Individually coded and tested each part of the processor
- Wrote a simple assembler for the processor instruction set using ANTLR. Converted to VHDL and tested on the Xilinx Atlys FPGA

PROJECTS

Harmonize — *Ongoing*
Synchronize music across devices — *Group Project*

- Wrote a C++ program to synchronize music and other media across devices. Implemented a protocol to send control messages at the TCP layer. Current protocol uses the server client model
- Heavily multithreaded real time application in which critical delays were monitored
- Achieved a synchronization with latency of less than 0.5 ms, using a custom clock sync protocol. For music streams, no observable delays (to the human ear) were noted

Primary Entity Detection in news articles — *September 2017—November 2017*
Information Retrieval research project — *Term Project*

- Researched on how to detect primary entities in political event news articles. The entities so detected can be used to classify events.
- Used word embedding with a combination of k-Nearest Neighbors, nGrams, Centroid Similarity, Neural Network and other algorithms to classify entities as primary/secondary
- Using a combination of methods, obtained an accuracy of 75% from a test set of 140 articles, after training on 700 articles

unCap
A Captcha solver

- Developed a simple captcha solver to solve the captcha on institute webmail authentication
- Removed noise from the image using a combination of median, Gaussian, and density filters
- Programmed a segmentation heuristic to segment the clean image into characters. Applied matrix multiplication based similarity scheme to identify individual characters. Obtained an accuracy of 90% after training on 500 images

Hooli — *September 2017—October 2017*
A distributed search engine — *Group Project*

- Designed a distributed search engine to index the institute intranet. Currently indexed 100,000 pages with the ability to scale to 1,000,000
- Initially based on a single PC, with an in-memory index, then distributed crawling/indexing across nodes
- Used HBase for distributing the index and RPCs for synchronization. Used query level and index level caches to reduce query time to < 2s

ACHIEVEMENTS

Team qualified for ACM-ICPC regionals — *Ranked 75th in regionals out of 500 teams*
Ranked among the top 1000 students in IIT-JEE — *Amongst 200,000 students*

TECHNICAL COMPETENCIES

- **Experienced**: C++, C, Python, Java, HTML, SQL
- **Familiar**: Javascript, Bash, Git, Linux, Microsoft Windows, Flask, VHDL
- **Basic**: PHP, Haskell, Prolog, Neo4j, MongoDb, Hbase