

10/14/2019 - 10/20/2019

E E/Cpr E/S E 491 Weekly Report 4

Intelligent Code Editor

Client & Advisor: Ali Jannesari

sdmay20-46

John Jago – Software Lead

Keaton Johnson – Systems Lead

Jon Novak – Machine Learning Lead

Matthew Orth – Meeting Facilitator

Garet Phelps – Report Manager

Isaac Spanier – Test Lead

Weekly summary

This week was focused on getting started with the dataset, finalizing the format for the UI, and optimizing the OpenNMT-py.

Past week accomplishments

John Jago

- Implemented the first real iteration of the plugin UI assuming an OpenNMT-py backend
 - See individual contributions below for details

Keaton Johnson

- Contributed to the System.out.println dataset

Jon Novak

- Contributed to the System.out.println dataset

Matthew Orth

- Research NMT Implementation and Dataset Creation
 - Research NMT implementation architectures, hyperparameters, and optimization
 - Determine considerations and further methods for creating simple dataset for the Java print function
 - Research and run various already created datasets through OpenNMT system to observe results as a baseline
- GPU Resource Research:
 - Determine what dedicated GPU options we have to train and test our model on
 - Optimize the GPU memory usage to resolve the memory issue
- Google Colab Configuration and Documentation:
 - Setup training environment in Google Colab's free environment
 - Train Python Code Corpora using Google Colab
 - Create document outlining OpenNMT-py and Google Colab integration

Garet Phelps

- Contributed to the System.out.println() dataset
- Worked on finalizing how the english statements should be phrased.

Isaac Spanier

- Began experimenting with Google's BigQuery

Individual contributions

Name	Contributions	Hours this week	Hours cumulative
John Jago	<ul style="list-style-type: none">• Updated the plugin UI and logic to handle a list of possible choices for translation• Improved user experience in the editor by adding keyboard shortcut and error messages	5	15
Keaton Johnson	<ul style="list-style-type: none">• Worked on the system.out.println dataset	1	10
Jon Novak	<ul style="list-style-type: none">• Worked on the system.out.println dataset	1	9
Matthew Orth	<ul style="list-style-type: none">• Research NMT Implementation and Dataset Creation• GPU Resource Research• Google Colab Configuration and Documentation	10	53
Garet Phelps	<ul style="list-style-type: none">• Contributed to the System.out.println() dataset	2	11
Isaac Spanier	<ul style="list-style-type: none">• Began experimenting with Google's BigQuery	1	8

Plans for the upcoming week

John Jago

- Contribute to realistic System.out.println() dataset using GitHub repos as a corpus
- Figure out a way to get the surrounding context in the editor to help ensure translation accuracy

Keaton Johnson

- Contribute to the System.out.println dataset

Jon Novak

- Contribute to the System.out.println dataset

Matthew Orth

- Continue experimenting with OpenNMT-py optimization
- Run Java print dataset through system

Garet Phelps:

- Contribute to the System.out.println dataset

Isaac Spanier

- Contribute to the System.out.println dataset

Summary of weekly client/advisor meeting

We had a meeting with Hung on 10/18 at 4pm. 4 team members were present. We discussed the progress on the UI, as well as how we should be making our dataset. We established that we should be putting spaces in our outputs so they are interpreted as more tokens in the output of the NMT. Goal for next week is to work on database, and resolve some UI issues with getting multiple results from OpenNMT.