

留学報告書

安永迪弘

2020年12月

2019年9月より Stanford大学にてコンピュータサイエンス(CS)の博士学生をしております安永迪弘と申します。今回は1年目夏~2年目秋学期の生活を振り返りたいと思います。

1. 学会発表

7月に International Conference on Machine Learning (ICML) という機械学習の国際会議にて、PhD 1年目に行ってきたプログラミングのデバッグを自動化する研究を発表しました:

Michihiro Yasunaga and Percy Liang. "Graph-based, Self-Supervised Program Repair from Diagnostic Feedback" (<https://arxiv.org/abs/2005.10636>)

今年は完全にオンラインの学会になっており、事前に15分のプレゼンを録画してアップロードし、当日はZOOMで2時間ほどのQ&A・交流セッションを行う形です。機械学習を用いてプログラミングやソフトウェア開発を効率化するという課題は近年インダストリー(IT企業)をはじめ大変重要視されてきており、自分のQ&Aセッションにも企業の研究者が参加してくださいました。論文の知見を企業で行われている大規模なプロジェクトに適用するにはどうすればよいかなどの議論ができ、インダストリーで実際に役立つ技術について学ぶ機会になりました。またありがたいことに論文が stateof.ai というメディアの記事 [State of AI Report 2020](https://stateof.ai/state-of-ai-report-2020) にも取り上げられ、この研究分野の注目度が高まっていることを嬉しく感じています。

2. 研究

秋学期は自分が興味を持っているもう一つのトピックであるReasoning (推論)、特に知識推論についての研究を行いました。例えば「モナ・リザの作者の出身はどこ?」という質問について、私たちは「作者はレオナルド・ダ・ヴィンチで、彼はイタリア出身」といったように、背景知識を使いステップを踏んで推論をします。またこういった質問に限らず、私たちは日々頭にある知識や調べ物で得られた情報を使って推論し意思決定をしており、推論はAIにおいても重要な課題だと考えています。現在のAI・自然言語処理において、「知識」の情報源として

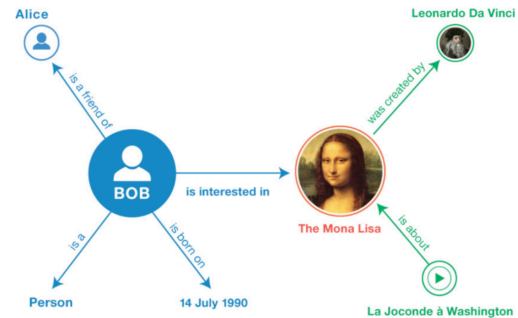
(1) 大量のテキストデータ(Wikipediaなど)で事前学習された言語モデル

(2) 知識概念を構造化してまとめたデータベースである知識グラフ

の2つがよく使われています。前者の言語モデルは大量の知識を学習しているけれどもステップを踏んだ論理的推論は不得意であり、後者の知識グラフは構造化・可視化されていて論理的推論に向いているが知識の網羅性に弱い、という長所短所があります。そのため、この2つの情報源の長所をどう効果的に組み合わせて推論問題を解くかという研究を現在行っています。



事前学習された言語モデル



知識グラフ

3. 生活

引き続きリモートワークが続くなか積極的にラボに関わっていきたいという思いもあり、ラボのorganizer役を今学期やってみることにしました。主な仕事は、毎週のラボミーティングをセットアップして議論を進めたり、social (食事や交流イベント) を行うことです。Social もリモートなので、皆でZOOMしながらオンラインゲームをすることにしました。アメリカでは Among Us や Avalon といった人狼系のゲームが人気で自分も好きになりました。またラボミーティングを主導する役割も、ラボメイトの研究や他分野をよりよく理解したり、ラボ全体のこと(共同研究や新入生の手助け)を考えようとする姿勢がつくようになり、大変でしたがとても学びとやりがいがありました。



Among Us: iOS, Android, Windowsのアプリを使って皆で繋がれてゲームできる

4. 最後に

2年目になってラボメイトや学科のメンバーとの交流も増え、自分の居場所をより確立できつつある学期になったと感じています。船井財団にはいつもサポートしていただき、本当に感謝しております。引き続き充実した留学生活が送れるよう頑張ります。