

船井情報科学振興財団第十回留学報告書

大谷直樹

2021 年 12 月

2017 年秋よりカーネギーメロン大学 (Carnegie Mellon University) の Language Technologies Institute (LTI) に在学中の大谷直樹です。自然言語処理 (NLP) の研究をしています。

1 近況報告

トップニュースとして、パンデミック前に中国に帰ったっきりアメリカに戻れなくなっていた奥さんが九月に帰ってきました。まだ中国から直接入国することは許されていなかったため、シンガポールに二週間滞在してからアメリカに入国しました。以来二人で楽しく生活しています。

奥さんは旅行好きなので、何回か車で小旅行をしています。10 月末にはペンシルベニア州のランカスターという街に行きました。この地域はアーミッシュ (Amish) という、昔の様式 (e.g., 電気などを極力使わない) を大切に守って生活している人々がいることで有名です。彼らは車を使わず、代わりに馬車で公道を走ります。ランカスターではアーミッシュの生活環境を再現した施設に行って文化を学びましたが、信仰って大変だな ...¹ という身も蓋もない感想を抱いた一方、電力をひたすら消費して大して社会の役にも立たない機械学習モデルを訓練している自分を小さく感じました。私も数年後はコンピューターを捨てて馬車に乗ってるかもしれません。

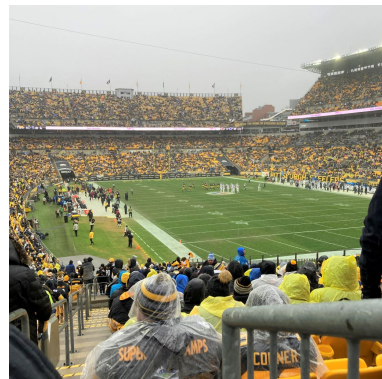
¹例えば、娯楽のための読書は禁止。宗教行事での楽器の演奏も禁止。



(a) アーミッシュの馬車



(b) ランカスターの帰りに寄ったゲティスバーグの古戦場



(c) ピッツバーグ生活 5 年目にして初めてアメフト観戦に行きました。寒かった。

2 学業

講義はすべて履修し終わったので、今学期は残り一回の TA 要件を満たすことと、研究に専念しました。

TA: 論文サーベイの練習をする修士向け講義の TA を担当しました。90 人近い受講者が提出する論文のレビューを 6 人の TA で採点するのがメインのお仕事でした。最初は面白い講義だと思ったのですが、次第に提出されるレビューのクオリティが下がってきて、TA からタイムリーにフィードバックを与えられず、あまり良い学習機会にはならなかったと思います。私も過去に別講義の受講生の立場で感じたことがあるのですが、とにかく宿題が多すぎるし、デザインが雑です。宿題自体にエラーがあったり、インストラクターから役に立つフィードバックがほぼありません (後者は私も反省)。学習課題というより、ほぼただの苦行になってしまっていました。TA も厳しく選別されているとは限りません。今回の TA の中には入学直後の修士 1 年がいて、NLP に関する知識が怪しかったり、仕事が遅かったりしました。留学前はアメリカの大学の講義はすべてレベルが高いと思っていましたが、そういうわけでもないことがわかってきました。

研究 1: 前の報告書に書いた通り、夏はマイクロソフトの研究所 (Microsoft Research; MSR) でインターンシップをしました。To Do 管理アプリの Microsoft To Do に様々な知的アシスタント機能を実装するための基幹技術の研究に取り組みました。現在、結果をまとめた論文を国際会議に投稿中です。プロジェクトでは実際のユーザーから集めたログに触れることができました。実データはだいたい「汚い」(スペルミスや解釈不能な単語列がたくさん) ので、理論的に動きそうな言語処理システムが予想に反して動かなかったりします。私の研究は、様々なタスクで有効性が示されている大規模な統計モデル (事前学習済み言語モデル) が To Do テキストに対してうまく動かないことを示し、それを解決するための手法を考案しました。

研究 2: インターンシップから戻った後は元奨学生の荒木さん + Bosch の研究者と行っている共同研究を進めました。こちらは会話型アシスタントに関するもので、会話の状況やユーザーが達成したいゴールを認識した上で、ユーザーのリクエストに対して「気の利いた」返答をしてあげる技術を研究しています。具体的にはこのような返答することを目標としています:

ユーザーのリクエスト (入力): エアコンをつけて。

会話の状況 (入力): ユーザーは帰宅したばかり。部屋の温度は低い。

ユーザーのゴール (推定): 身体を温めたい。

「気の利いた」返答 (出力): わかりました。お風呂の準備もしましょうか？

この問題に対して、私たちは人間の常識的知識を機械学習モデルに入れる方法を考えています。秋ごろにある程度知見が溜まったのでひとまず論文にまとめて国際会議に投稿しました。まだまだ改善の余地が大きいので、もう少しこのプロジェクトに取り組む予定です。

3 おわりに

今学期から博士論文の準備に本格的に取り組み始めました。あと数か月でプロポーザルを終えて、卒業までのタイムラインが明確になるはずです(たぶん...)。身体的にも精神的にも健康を損なわずに走り切れるように適度に息抜きしながらがんばります。