

船井情報科学振興財団 第5回留学報告書

オフィスで飲む緑茶が美味しい

長谷川公大

kimihiro@andrew.cmu.edu

2024年6月

2022年5月から米国カーネギーメロン大学計算機科学大学院言語技術研究所^{*1}の博士課程に在籍しています、長谷川公大と申します。前回の報告（2023年12月）から2024年6月現在に至るまでの留學生活について、簡単にではありますがご報告させていただきます。^{*2}

学業

昨年の秋学期に引き続き、春学期に入ってからピッツバーグで黙々と研究に励む日々を送りました。

リジェクトの話

前回の報告書でざっくりとお話をしたプロジェクトの論文に関してですが、春学期の間に絶賛2度連続してリジェクトをもらい苦戦しています。内容は文章中から出来事情報を取り出し文脈を考慮して時系列順に並べ替えてタイムラインを作るというタスクにおいて、個別の既存手法を大きく分類分けし分類自体のメタな比較を行うかつ、そのメタ比較のための新しい評価用のデータセットを提案するというものです^{*3}。一度目は比較実験の部分を推しに二度目はデータセットの部分を推しにして投稿を行ったのですが、査読者からあまり良い評価を得ることができませんでした。比較実験にもデータセットにも新規性及び分野への貢献があると思うのですが、おそらく新規性のある部分を伝えなくてはいけないという気持ちが強くなりすぎたあまり、論文の“型”みたいなものを疎かにしてしまって、推しの部分をわかりやすく一貫したストーリーでまとめることができていなかったのだらうと思います。しっかりした伝える技術があれば、より自由な形で自分の書きたいようにかいて伝えることも可能なのかもしれません^{*4}、私の場合は型が定まっていないのに好きに書こうとして、型破りではなく型無し^{*5}になってしまっていたのかななどと、とりとめもなく考えたりしていました。自分の頭の中での整理も兼ねて少し時間をおいたり、ラボメイトに相談に乗ってもらったりしながら、論文としての“型”を意識し、よりわかりやすい形で伝えられるようにストーリーを書きかえ、その過程で不足に気づいた実験を追加で取り組んでいます。とりあえず打席に立たなければ打てないので、根気よく頑張ろうと思います。

^{*1} Carnegie Mellon University, School of Computer Science, Language Technologies Institute の訳です。

^{*2} 副題は本文と関係ありません。おすすめは一保堂という京都の老舗のお茶屋さんです：<https://www.ippodo-tea.co.jp/>

^{*3} <https://arxiv.org/abs/2403.00990>

^{*4} 広義では FOS2018 の笠井さんが恩師へのお悔やみの言葉を論文という形で投稿したものなどが思い当たります <https://arxiv.org/abs/2303.18027>

^{*5} by. 18代目中村勘三郎が無着成恭という僧侶の発言を引用とのこと。

ちなみにお恥ずかしながら査読結果が返ってきて次の日くらいまでは、“なんで理解ってくれへんねん、” “悪いところやなくてええとこ見てや” などと駄々をこねたり、SNS 上での査読への不満・運要素を嘆く声を見て溜飲を下げたりしたくなります（しています）。ただ、そういった誘惑と折り合いをつけながら、己が論文の不出来を見つめ直す機会として活用しようと心がけています。

マルチタスクの話

春学期から新しくマルチモーダル^{*6}を対象とした質問応答のプロジェクトに取り組んでいます。今まで主にテキストを対象にした研究を行ってきたので、新しいことを学ぶ苦労はあるものの新鮮な気持ちで楽しみながら研究に取り組んでいます。ただ一点困っているのが、上記“リジェクトの話”のプロジェクトとともにこちらのプロジェクトにも取り組む必要が出てきたため、自分が100%の力で取り組んだとして一定期間必要になるものを複数同時に行う、マルチタスクの難しさに面し四苦八苦しています。これまで、プロジェクトはなるべく一つに集中した方が良いかなと考えて過ごしてきたのですが、一方で、マルチタスクは技術でありこれから必要になる機会は増えてくるから、せっかくなら博士課程の間にも練習すると良いものなどの意見を耳にすることもあります。洗濯物を回している間にトイレの掃除をすとか、ラジオを聴きながらランニングすとかであれば特に問題はないのですが、私が研究での作業を行うときは、前回の作業時の内容を思い出した上で数時間かたまった時間取り組んでその後半ようやく頭が回ってきたり、机に向かっていない帰宅中にいい感じに頭が整理される、そもそもプロジェクト間を移動するのがもどかしくて結局片方のプロジェクトにのみ時間を費やしてしまったりします。作業内容のメモを残しておくとかしてタスク間の移動をなるべくスムーズにできるよう心がけてはいるのですが、どうも相性が悪いなと感じています^{*7}。また抱えているタスクの数に比例して作業時間もしくは作業効率を定数倍にしない限りは、基本的にはそれぞれのタスクの進捗は一つに集中した時に比べて緩やかになってしまうため、達成感みたいなものを感じづらくなり、モチベーションの低下&フラストレーションの増加にもつながるなと感じました。かといって作業時間を増やして補おうとすると疲労が溜まって作業効率が落ちてしまったりもするので、ままなりません。現状必要であり、身につけられたら大変有用なものであるとは思われるのですが、論文がリジェクト続きの時に、練習だからと言ってより進捗が出にくくなるものを頑張るのは精神衛生上も良くないなと感じています。ということで、最近は数日もしくは週といった、一つのプロジェクトである程度の進捗を生めるだけのまとまった時間をとった上で、ゆったりとプロジェクト間の行ったり来たりをしています^{*8}。

その他

CMU の Spring Carnival と呼ばれる、文化祭みたいなものに今年初めちゃんといってみました。過去存在は知っていても横目で見ただけで、あまりよくわかっていなかったのですが、簡単な遊園地みたいなものが設置されたり、学生主体の出し物があったりと楽しそうな雰囲気を醸し出していました。学生の出し物の中には、CS の強い大学らしく自作のゲームを体験できたりするものもあり、面白かったです。全体的な雰囲気は高校の文化祭のようなものを感じであり、少し懐かしくなりました^{*9}

^{*6} テキストと画像など複数の異なる種類のデータを同時に扱う

^{*7} 自分がキャッシュメモリに頼った作業の仕方をしていることやそもそもの自分の手の速さが足りてないことも理由に考えられます。

^{*8} 研究だけに集中できているからこそその悩みでもあるので、贅沢な悩みだとも思います。

^{*9} 私の高校は文化祭がなかったので、私自身文化祭というものを経験はしていません。

これ以外で特段、話の種になるような活動をあまりできなかったのですが、近くの公園*10を散歩したり、軽くジョギングしたりして過ごしていました*11。中高と運動部に所属していたおかげか、大学に入ってから特段ダイエットなど考えなくても体型で気になることはあまりなかったのですが、最近は食べている量は以前と変わらないのにも関わらずお腹周りの贅肉が気になり始めてきたため、運動の習慣をつけないとなと思い始めています。普段から座って作業する時間が長いので、加齢も相まって代謝が落ちてきているのだらうと思われれます。研究を頑張るためにも身体は大事なので、健康にも気を配って過ごしていきたいと思います。

また最近家で家事等をしている際に、お笑いのラジオを聴いています。目も手も空いた状態で他の作業をしながら楽しめるので、今更ながら良いなと感じています。芸人さんの純粋な掛け合いだけで人を笑かせる話術やリスナーのよくそんなこと思いつくなという投稿にいつも楽しませてもらっています*12。

他には、FOS2019のファミさんと定期的に一緒にランチや散歩をしたり、FOS2020の荒川さん企画のBBQに参加したり、FOS2023の金村さんに誘ってもらって大学の体育館でバスケットボールのpickup gameに加わったりしていました。今年の上半期は包丁での怪我をせずに過ごせたのでよかったです*13。ただ年末に食中毒になって以来、お腹が弱くなっているみたいで何度かお腹を壊しかけました。鶏肉は十分火を通して食べないといけないなと身を持って学びました。

謝辞

2年間にわたる船井財団からのご支援改めてありがとうございます。絶賛諸々苦戦中ではありますが、競争的な環境で研究に集中できているからこそその悩みであり、そのような機会を得られたのは船井財団からのご支援あってのことだと思っています。心から感謝申し上げます。今後も引き続き学位取得に向かって、分野へそして社会に貢献・還元できるような研究に勤しんでまいりたいと思います。

*10 最近、夕方～夜になると蛍が飛んでいるのを見かけました。勝手に日本固有の虫だと思っていたので驚きましたが、ペンシルベニア州の州の虫になるくらい一般的なようです。

*11 ピクミンブルームというアプリでピクミンを育てています。ポケモンGoの類似アプリのようなものです。

*12 お気に入りの中川家 ザ・ラジオショーと霜降り明星のオールナイトニッポン、金属バットの社会の窓です。下ネタも多いので（特に後二つ）、あまり大きな声で人様に勧めするのは憚りますが。

*13 先学期包丁で指を切り4針縫っています。詳しくは前回のレポートをご参照ください



(a) 大学のすぐ近くの Schenley Park。この写真のところあたりで、夕暮れから蛍が飛んでいたります。



(b) 家の近所の Frick Park。ドッグランもあるためか犬の散歩をしている人が多いです。



(c) Carnival の学生の出し物。このような木造の箱にデコレーションしたものが10以上並んでいます。大雨でした。



(d) Carnival の観覧車。乗る部分は囲いが無いみたいでなかなかスリルがありそうでした。

用語

厳密にはリジェクトではなく、査読で自信がつくような点数を得られなかったなどと表現した方が正しいかもしれません。自然言語処理分野ではACLと略称される学会を中心に、数年前から査読と学会での発表の可否を分ける運用がなされています。投稿者はまず査読フェーズに投稿し、その査読結果を得た後、実際に学会へ投稿するかどうかを決めることができます。そして、それぞれの学会が最終的に発表の機会を与えるか否かを判断するという形です。査読フェーズには2ヶ月に一度投稿が可能なのに対し、学会は主に秋から春にかけて投稿の締め切りがあるので、投稿者は学会への締め切りまでに査読フェーズで好評価の査読結果を手に入れることが求められます。どのような査読結果であれ学会には提出できるのですが、大体のボーダーラインのようなものが存在すると言われていて、私は過去2回そのボーダーラインだと思われるもの以下の評価しか手に入れられなかったので、学会への投稿を見送ったという形です。ただそれぞれの学会での acceptance rate は私のように学会への投稿を見送ったものが考慮されるため（詳細略）、直接学会が査読と発表の可否を判断する場合のリジェクトとだいたい似たようなものになっているかなと思い、リジェクトという言葉を使いました。