

第4回報告書

笠井淳吾

ワシントン大学（シアトル）でコンピュータサイエンスのPhDを去年の9月から始めた笠井淳吾と申します。研究分野としては、自然言語処理(NLP)に取り組んでいます。

1. 夏休み

今年の夏休みはPhDを初めて一年目ということで、大学に残って研究を続ける選択をしました。昨年インターンしたIBMで再度インターンをする話もいただきましたが、今回は企業に入って一つのプロジェクトに取り組むというより、じっくり自分がやりたいことを考える期間にしたいと考えました。夏休みのまとまった時間を裁量労働にしたことによって、イタリアフィレンツェで行われたACL（NLPの国際会議、来年はシアトルです！）で久しぶりに会議、ワークショップとも全て参加して、多くの人と再会し、話す機会に恵まれました。運悪く船井の交流会と重なってしまいましたが、収穫の多い会議になりました。来年のACLは7/5-7/10に行われるので、船井の交流会と重ならないと嬉しいです。船井のコミュニティも毎年広がっていますし、一人の都合に合わせるわけにはいきませんね笑。下はACLでの発表の様子です。



フィレンツェという場所も充実した会議になった要因の一つかと感じました。普段いるシアトルやベイエリアと違って、歩くだけで幸せな気持ちになれる街だと感じました。東京もそんな街の一つだと思います。唯一、少し気温が高かったのがシアトルっ子の甘ちゃんには若干こたえましたが。毎年膨れ上がるAI関連の国際会議の提出論文や、査読システムの破綻などに対して、現在のシステムを変えなければならないと不平をもらし

ていますが、研究のモチベーションをあげるためにもこういう場所で年に一回くらい集まるのは良いことだと思いました。

2. 秋学期

大学で過ごした夏休みを経て、秋学期はバイエリアのFacebook AIでインターンをすることに決めました。具体的には、機械翻訳の精度を保ったまま、いかに並列化して速度を上げていくか、というプロジェクトに取り組みました。言語生成を並列化することはいまだに難しい課題の一つです。2019年は、GoogleのBERTなど、GPUやTPUのもと並列計算で、大量のデータを使ってモデルをつくり、人類は言語処理において飛躍的な進歩を遂げました。このモデルは、[Google検索](#)においても、使われ始めました。一方で、言語を生成することは、一つ一つが独立事象ではなく強いDependenceがあるために、このような並列化と相性が悪いという背景があります。つまり、我々が話す言語には、文脈、そして単語間の一貫性のようなものがあるはずで、例えば、同じ単語を繰り返し発することは普通ではありません。言語生成の過程を並列化し、速度を上げるためには、このような一貫性をいかに担保するかが本質的な問題になります。今回はこの問題に挑戦するインターンシップになりました。Facebookがもつ巨大なインフラを使い、優秀な研究者たちと研究できたのは大きな経験となりました。

研究以外にも色々発見がありました。Facebook内では多くのものにディープラーニングが使われ、完全にインフラ化しています。学部三年生で初めてディープラーニングに触れた時、一時期の流行り物だと漠然と思っていましたが、数年がまさかここまでインフラ化し、浸透しているとは思っていませんでした。時代の流れは早いです。また、社内のベネフィットや環境も素晴らしく、中でも種類も豊富でレストランのような朝昼晩の社食は毎日楽しみでした。普段の学生生活とのギャップでインターン生はフルタイムの人より20%以上食糧を消費する、なんてデータがあると聞きました。耳が痛いです。楽しいところなので、もしバイエリアに行く機会があれば、ぜひ行ってみてください！映える写真もたくさん撮れますよ。



3. 今後について

1/3までベイエリアにいますが、来学期はワシントン大学に戻ります。ここ数週間、次になりをするのか、大学で何を軸に研究していきたいか考えてきました。まず、研究者としてのビジョンをはっきりさせたい、というのが当面の目標です。NLPにおいても、解決されていない問題はたくさんあると思うのですが、どこに焦点を絞り、発信していくか考えていきたいと思います。やはり多言語生成や並列化は大きなテーマだと思うので、そこに関連づけて研究していけたらと今は考えています。

4. 余談

EMNLPの[Best Paper Award](#)の第一著者はジョンスホプキンスの学部三年生だったらしいです。このような早い段階で実績を残される方もいて、身が引き締まる思いです。早いといえば、椎名林檎さんが丸の内サディスティックを作ったのは20歳の時らしいですね。とんでもない話です。報酬は入社後平行線ですが、前を向いて頑張っていきたいです。