



会議レポート

EMNLP 2017 参加報告

2017年9月、自然言語処理分野におけるトップ国際会議の1つであるEMNLP2017がデンマークのコペンハーゲンで開催された(図-1)。EMNLPとは、Empirical Methods in Natural Language Processingの略称であり、自然言語処理における「経験的(empirical)」アプローチを主として扱う会議である。この「経験的」という用語は、自然言語処理の専門外の方々にはやや奇妙に聞こえるかもしれないが、いわゆる機械学習などのデータ駆動型のアプローチ全般を指すと考えていただければ問題ないと思う。他分野ではこのようなアプローチを「経験的」と呼称することは珍しいかもしれない。これには、かつて自然言語処理には「合理主義的(人間の言語理解の過程に、ある種の計算モデルを仮定し、これにより自然言語を科学的に探求しようとする考え方)」アプローチが主流となっていた時代があって、EMNLPはこの「合理主義」に対するカウンターパートとして命名されたという歴史的経緯がある。ただし、2017年現在、自然言語処理における主流はデータ駆動型アプローチへと完全に移行しており、今ではほぼすべての研究が「経験的」アプローチを採用している。そのためempiricalという看板はあまり意味をなさなくなっている。

私は発表者の1人としてEMNLP 2017に参加してきたが、昨今の人工知能ブームの追い風もあって会議は非常に盛況であった。参加登録者数は初日の時点でおおよそ1,100人であり、論文投稿数は、ロングペーパーとショートペーパーを合わせて1,418本(それぞれ836本と582本)であり過去最多だったそうである。ちなみに、私が初めてEMNLPで発表をしたのは、これを執筆している2017年からちょうど10年前の2007年であった。当時のEMNLPはトップ国際会議の一角として認識されつつあったが、まだ参加者数はそれほど多くなく、ACL(自然言語処理における最も権威ある国際会議)との連続開催が通例となっているという状況であった。当時の参加者数に関する詳細な数字は分からないが、論文投稿数を調べてみたところ398本と今年の半分以下であったよ



図-1 本会議初日のOpening Remarksの様子

うだ(当時はロングペーパーのみの募集)^{☆1}。この数字から、過去10年間での自然言語処理の成長具合を感じ取ることができ、長年この分野にかかわってきた研究者の1人として感慨深いものがある。

10年前との比較という視点から今年のEMNLPを振り返ってみると、まず感じることは、自然言語処理と社会の接点が大きく広がっているということである。このように感じる理由として、まず会議の参加者層が大きく変化したことが挙げられる。10年前のEMNLPでは、日本からの参加者は大学や公的機関の研究者ばかりであったと記憶している。しかし今年は、いわゆるIT企業からだけでなく、新聞社のように従来は自然言語処理から縁遠いと思われていた(自然言語処理の研究者が在籍しているというイメージがない)企業からも、論文発表や情報収集のために参加している方が多かったように思う。このように幅広い業界からの参加者が増えたことに加え、発表内容にも社会との密接なつながりを感じさせるものが増えてきたことも印象的であった。個人的に最も記憶に残っているのはStanford大学のDan Jurafsky教授による招待講演である。自然言語処理技術を駆使して警察官と市民の対話ログ(交通違反などの軽微な犯罪にかかわる職務質問のやりとりを録音したものであったと記憶している)を分析するプロジェクトに関する講演であり、大変興味深いものであった。また、会議と併設する形でNatural Language Processing meets Journalismというワークショップが開催されていたことも、自然言語処理技術に対する社会の関心の高まりを強く感じた。

もちろん、この10年間における自然言語処理技術の変革、進化にも目を見張るものがある。もはや取り立てて言うべきことではなくなってしまっているが、EMNLP

^{☆1} https://aclweb.org/aclwiki/Conference_acceptance_rates

でもディープラーニングを利用した研究が多数派を占めるようになり、10年前からはすっかりパラダイムが変わってしまった感がある。これは機械翻訳の分野においてとりわけ顕著であり、10年前に隆盛を誇っていた統計的機械翻訳に関する研究はもはやほとんど見られなくなっている。統計的機械翻訳というのは、大雑把に言うと、文を語句単位で翻訳する方式のことである。数年前からこれをニューラルネットというブラックボックスな処理で置き換える動きが急速に進み、現在ではこちらが完全に主流になっている。2007年当時、10年後にこのような日が来ることは誰も想像できなかったのではない

か。10年先を見据えた研究をすることの難しさについて身を持って感じることはできなかったのは、EMNLPに限らず、最近国際会議に参加して思うことの1つである。

以上、10年前のことを思い出しながらEMNLP 2017に参加した雑感を書きつづけた。今から10年後にEMNLP 2027が開催されるころ、自然言語処理を取り巻く環境がどうなっているかは分からないが、まずは10年後も筆頭著者として論文発表することを目標にしつつ日々の研究に励もうと思う。

(鍛冶伸裕/ヤフー (株))



会議レポート

SPLC 2017 参加報告 ～ソフトウェアプロダクトラインに関する トップカンファレンス～

会議の概要

ソフトウェアプロダクトラインとは、共通の特徴を持ち、特定のマーケットやミッションのために共通の再利用資産から作られるソフトウェア群のことであり、その開発技術に関する研究はソフトウェア工学において非常に重要な分野の1つである。The 21st International Systems and Software Product Line Conference (ソフトウェアプロダクトライン国際会議, SPLC) は、ソフトウェアプロダクトライン開発に関するトップカンファレンスであり、2017年9月25日から29日にスペインのセビリヤにおいて、第21回目が開催された。最新の研究成果を発表する場であるとともに、現場の開発で得られた知見や課題を共有する場でもあり、研究者と開発現場の技術者が一同に会する場となっている。

この会議は、90年代より欧州で行われていたPFE (Program Family Engineering) という会議と、2000年から米国でスタートしたSPLC (Software Product Line Conference, SPLC) とが2005年にマージされたものである。この際PFEとSPLCの開催回数もマージされ、今

回が21回目というカウントになっている。ソフトウェアだけでなくソフトウェアの搭載されたシステムに関する議論も増えつつあり、近年はシステムプロダクトラインを対象としたトラックなどが設けられていたが2016年より、会議名自体がSystems and Software Product Line Conference (略称はSPLCのまま)に変更されている。

会議プログラムの構成は、ほかの国際会議と大きくは変わらず、基調講演や論文発表からなる3日間の本会議、バンケット等のソーシャルイベント、本会議に先立って行われる2日間のワークショップとチュートリアルというものである。論文は、例年、研究論文と経験論文のカテゴリで募集され、さらに昨年からは将来のビジョンについて論じるビジョン論文も募集されるようになった。今年の採択率は、研究論文については49編の投稿中15編が採録で採択率31%となっている。

基調講演

2017年は3つの基調講演があった。

(1) ビジネスプロセスとの関連

ソフトウェアプロダクトライン開発では、プロダクト群に共通する特徴である共通性と、個々のプロダクトにおける差異である可変性を明確に認識し、管理することが重要である。この共通性と可変性について、他分野の研究状況を学ぼうというのが1つ目の基調講演である。

Marcello La Rosa (クイーンズランド工科大学) が、ビジネスプロセスにおける共通性や可変性、それに基づくカスタマイズについて講演した。既存のビジネスプロセスモデリングの手法を対象に、それらがノード構成やアクティビティ (タスクの集合) などをどのようにカスタマイズできるかの調査結果を紹介し、また手法の選択における基準やこの分野の研究のチャレンジなどについて触れた。