

自然言語処理技術の現状と展望

—エラー分析プロジェクトを通して—

編集にあたって

関根 聡（ニューヨーク大学） 乾 健太郎（東北大学）

ここ数十年にわたる機械学習研究の進展と、計算機の大規模化、テキストや音声による言語データ量の爆発的な増大により、自然言語処理の技術も大きな進展を遂げた。また、巷では人工知能が一種のブームになり、自然言語処理にも期待が高まっている。そこで、本特集では自然言語処理の技術の現状を研究者自身がしっかりと見つめることを目標に行われた Project Next NLP の話題を中心に、自然言語処理のさまざまな技術の内容と課題と今後の見通しについての解説を行う。自然言語処理において機械学習の応用は今や欠かせない要素技術となっているが、機械学習を単に与えられたタスクで応用しシステムを構築するだけではなく、システムによるエラーの分析を通じて自然言語処理が持つ本質的な課題を地道にしっかりと見つめることが、Project Next NLP の目標である。総勢 200 人以上の自然言語処理の研究者がなんらかの形で参加したこのプロジェクトを通し、彼らの真摯な研究に対する姿勢を感じていただきながら、本特集を楽しんでいただければ幸いである。

本特集は、まず最初に Project Next NLP の紹介と自然言語処理の技術に馴染みのない方に対して、自然言語処理の導入的な解説を行う。その後、Project Next NLP で設定されたタスクのリーダーが、各タスクの紹介、課題、今後について説明するという形式をとっている。読者は、興味あるタスクについて、順序は関係なく読んでいただいても楽しめるような構成になっている。Project Next NLP では、それぞれのタスクに取り組んでいる大学や企業の研究者が横断的に集い、研究対象を客観的に見つめなおすという作業が行われた。この協調的な作業が本プロジェクトの核心であり、それを通じて参加者が今後の研究についてのヒントを得られたことが大きな成果である。すべてのタスクにおいてそれぞれ素晴らしい成果が得られたことはひとえに各タスクのリーダーの献身的努力と参加者の情熱に負っている。リーダー、および参加者に心から感謝するとともに、このプロジェクトの企画、全体調整役として働けたことに感謝したい。

(2015 年 11 月 16 日)