EMNLP2013 参加報告（その２）
– NLP Applications 分野の論文紹介 –

江原　遥[†]

1　はじめに

本稿では，“The 2013 Conference on Empirical Methods on Natural Language Processing”（以下，EMNLP 2013 と略記）のうち，主に著者が聞いていた NLP Applications 分野の論文から，２本を紹介する．また，関連して，Andrew Ng による招待講演の様子も紹介する．

NLP Applications は，新しい自然言語処理のタスクを扱う論文が主に当てはまる分野である．前年の EMNLP 2012 では 1 セッションで口頭発表が 3 件だったのに対して[1]，今回の EMNLP 2013 では NLP Applications の口頭発表セッションが 3 セッション用意され，10 件の発表があったことは，この分野の興隆を示唆していると言える．教育応用も NLP Applications 分野に含まれる．

2　整数計画問題を用いた自動英文訂正

最初に紹介するのは，自然言語処理の教育応用分野でも注目度の高い，自動英文校正に関する論文である (Rozovskaya and Roth 2013)．この論文は，自動英文校正の最高精度を整数計画問題を使って複数の識別器を統合することにより達成した，という論文である．このようなアプローチとしては，同年の ACL 2013 でも (Wu and Ng 2013) という発表が行われているが，こちらでは，従来手法と同じ程度の精度に留まり，精度の向上は見られないという結論であった．(Rozovskaya and Roth 2013) は，同様のアプローチで早くも (Wu and Ng 2013) を凌駕する精度を達成した論文である．また，(Rozovskaya and Roth 2013) の著者らは，ACL 2013 に続いて同地で行われた CoNLL 2013 shared task で 1 位を取っている．

自動英文校正の分野では，従来，誤りの種類ごとに教師あり識別器を作成し，多値識別によっ

[†]日本学術振興会　特別研究員（PD）　受入：国立情報学研究所　宮尾研究室, JSPS Research Fellow, Miyao Lab., NII.

[1]EMNLP 2012 の会議プログラムより `http://hum.csse.unimelb.edu.au/emnlp2013/programme.html`.

て訂正先を選ぶ，という手法を取っていた．例えば冠詞誤りであれば，冠詞の “a” と “the” と ϕ（無冠詞）を教師ありの識別器を用いて多値分類するという具合である．しかし，この方法では，複数の修正を同時に行わなければならない *Interacting mistakes* と呼ばれる誤りに対しては，対応できなかった．例えば，“such situation” を直す場合，“such situations” または “such a situation“のどちらかに直さなければならない．しかし，従来手法であると，冠詞を修正する識別器と，対応する名詞（以下，NPhead[2]）の単複を修正する識別器が独立に動作していたため，“such a situations” のような訂正をしてしまう問題があった．

複数の識別器の出力を統合するため，(Rozovskaya and Roth 2013) では，式 (1) のような整数計画問題を用いて，識別器のスコアを統合している．H は可能な訂正の組み合わせの空間であり，$h \in H$ はその中の 1 つの組み合わせである．この 1 つの訂正の組み合わせ h に対応する複数の教師あり識別器のスコアの和を $score(h)$ で表しており，$score(h)$ を最大化する定式化となっている．整数計画問題の制約として，「“a” と複数」というような組み合わせがあり得ない事などを陽に入れている．さらに詳細な整数計画問題による定式化は，(Roth and Yih 2004) を参照しており，本文では数式による説明を抑え，文章で多くの説明がなされている．

$$\hat{h} = \arg\max_{h \in H} score\,(h) \tag{1}$$

実は，このアプローチ以外にも，複数の識別器を統合する簡単な手法がある．それは，複数の修正の組み合わせを陽に列挙し，その組み合わせに対して多値分類の識別器を学習する方法である．例えば，冠詞誤りと NPhead の例であれば，冠詞と NPhead の両方を入力として受け取り，{“a”, 単数 }, {“a”, 複数 }, {“the”, 単数 }, ..., {ϕ, 複数 } といった，$3 \times 2 = 6$ 通りの組み合わせを出力する識別器を構成してしまえば良い．こちらの簡単なアプローチでは，{“a”, 複数 } といったあり得ない組み合わせでも，コーパスから自然に学習される事を期待して，陽に制約を人手で入れるといった特殊な事はなされていないようである．識別器としては，Average Perceptron を用いている．

本論文の良い点は，この簡単なアプローチを *Joint learning*，整数計画問題を用いたアプローチを *Joint inference* とそれぞれ名付け，片方だけ適用した場合と，Joint learning と joint inference をさらに組み合わせた場合の精度も報告されている事である．

(Rozovskaya and Roth 2013) の Table 9 を表 1 に引用する．Illinois は，CoNLL 2013 shared task で 1 位を取った同著者の Rozovskaya らのチームによるシステムの性能である．表 1 より，Joint Learning と Joint Inference を同時に用いた場合 Joint Learning + Inf. が最も精度が良いことが分かる．

[2]名詞句 NP の head であるので．

表 1　Joint Learning, Joint Inference の精度比較.

	F1 (orig)	**F1 (Reviesed)**
Illinois	31.20	42.14
Joint Inference	32.51	43.19
Joint Learning	35.12	43.73
Joint Learning + Inf.	**35.21**	**43.74**

3　成功する本を当てる

次は，大きくタスクがかわり，成功する本を当てるタスクを解くという課題に挑戦し，84%という高い精度を達成したという論文 (Ashok, Feng, and Choi 2013) を紹介する．この論文を選択した理由は，この研究が内容的にも手法的にも，著者が昨年最先端 NLP 勉強会で紹介した，新聞の科学記事の質を様々な素性を導入して予測する，(Louis and Nenkova 2013) と近いからである．

このような新しいタスクに打って出る時は，Introduction や Related Work といったセクションを通じて，多くの参考文献をカバーし整理するサーベイを行い，実際にそのタスクが「新しい」事を示す必要がある．そのためか，この論文の Introduction は 1 ページ半以上も費やされている．Introduction によれば，「文章のスタイルと，文学的成功の間の関係を調査した初の研究」であると主張されている．

データとしては，Project Gutenberg という，歴史的な文学などを中心に，書籍の全文が手に入るサイトにおける，ダウンロード数を適当な閾値で切って，成功・不成功の 2 クラスに分けている．具体的には，τ^+,τ^- という閾値を導入し ($\tau^+ > \tau^-$)，ダウンロード数 $> \tau^+$ ならば成功，ダウンロード数 $< \tau^-$ ならば不成功，という形で 2 値分類問題に帰着させている．

手法的には，Liblinear SVM (Support Vector Machines) の L2 正則化を 5 分割交差検定で行うという，機械学習的には典型的な手法を用いている．この点は，(Louis and Nenkova 2013) でも，機械学習的な革新性は薄い手法を用いていた．このような研究では，手法よりも，使用した素性が重視される．特に，様々な仮説にもとづいて素性を入れてみて，実際に実験で効いているかを確認したりしている．素性としては，1-gram, 2-gram といった通常用いられる素性に加えて，PCFG で構文木を推定した時に使われたルールや，タグの分布，感情に関連する語をも用いている．この研究で新しく導入したと主張されている素性としては，NP (Noun Phrase) や PP (Prepositional Phrase) といった phrase tag や，clause tag の比率が挙げられる．

4 節で素性を説明した後に，5 節で，具体的にどの素性が有効かを見ている．例えば，成功した本では，前置詞句, 名詞句, WHNP (wh-noun phrases) が多いのに対して，成功していない本

では，動詞句，副詞句や Interjection（“Oh!”, “No!” など）が多いのは，面白い．

「成功する本を当てる」というのは，一見して新規なタスクに見えるが，もう少し広く「文書の読みやすさを評価する」という事になると，リーダビリティ（読みやすさ）を定義・評価した研究は多くある．例えば，古典的には，Flesch index や FOG index などが挙げられる[3]．このような指標を利用した実験も隙なく行われている．面白いことに，成功する本の方が，そうでない本より読みやすさは低いそうである（当該論文 Table 7）．

4　おわりに

本稿では，NLP Applications 分野から，2 本の論文を紹介した．その他，EMNLP では，近年スタンフォード大学から Coursera 社に移動した Andrew Ng が招待講演を行った．Coursera は，大規模公開オンライン講座 MOOC (Massively Open Online Courses) を運営するシステムの代表例である．Coursera に移動した理由として，スタンフォード大の機械学習の講義で，学生の質問に対して，Ng 自身が用意した回答よりも，同じ講義を受けている他の学生が用意した回答の方が，良質な回答になっていることに気がついたからである，と言っていたのが印象深かった．

参考文献

Ashok, V. G., Feng, S., and Choi, Y. (2013). “Success with Style: Using Writing Style to Predict the Success of Novels.” In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 1753–1764 Seattle, Washington, USA. Association for Computational Linguistics.

Louis, A. and Nenkova, A. (2013). “What Makes Writing Great? First Experiments on Article Quality Prediction in the Science Journalism Domain.” *Transactions of Association for Computational Linguistics*, **1** (July), pp. 341–352.

Roth, D. and Yih, W.-t. (2004). “A Linear Programming Formulation for Global Inference in Natural Language Tasks.” In Ng, H. T. and Riloff, E. (Eds.), *HLT-NAACL 2004 Workshop: Eighth Conference on Computational Natural Language Learning (CoNLL-2004)*, pp. 1–8 Boston, Massachusetts, USA. Association for Computational Linguistics.

Rozovskaya, A. and Roth, D. (2013). “Joint Learning and Inference for Grammatical Error Correction.” In *Proceedings of the 2013 Conference on Empirical Methods in Natural Lan-*

[3]参考文献は当該論文参照．

guage Processing, pp. 791–802 Seattle, Washington, USA. Association for Computational Linguistics.

Wu, Y. and Ng, H. T. (2013). "Grammatical Error Correction Using Integer Linear Programming." In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1456–1465 Sofia, Bulgaria. Association for Computational Linguistics.

略歴

江原　遥（正会員）：2013 年東京大学　情報理工学系研究科　数理情報学専攻博士課程を修了．博士（情報理工学）．日本学術振興会　特別研究員（DC2）を経て，現在，日本学術振興会　特別研究員（PD）　（受入：国立情報学研究所　宮尾　祐介　准教授）．自然言語処理・機械学習，特に読解支援などの教育応用の研究に従事．ACL, 言語処理学会，人工知能学会，情報処理学会，日本データベース学会，各会員．

（2013 年 11 月 13 日依頼）
（2014 年 1 月 24 日受付）