

# COLING2012 参加報告 (その2)

## – いかにも Coling らしい発表をいくつか –

金山 博<sup>†</sup>

### 1 はじめに

Coling では、2000 年以来 1 回おきに<sup>1</sup>発表をさせていただいており、そろそろ古参の部類に分類されてしまうかもしれない。しかし、世界からの参加者の顔ぶれをみると、年輩の先生方も昔と変わらず揃っており、この中ではまだまだワカモノ<sup>2</sup>としてやっていけると感じられるような、人と出会って元気になれる場所、Coling はそんな会議だと思う。

初のインドでの開催となった Coling は、投稿・査読のプロセスやその後の案内、ビザの取得などに少なからぬ混乱が生じ、先が思いやられると感じた参加者も多かったことであろう。しかし、蓋を開けてみれば、とても快適な場所で聴講や議論ができた。Mumbai の雑然とした街中を忘れさせるほど、会場となった IITB (Indian Institute of Technology Bombay) の中は整然としており、快適に繋がる無線 LAN、潤沢な電源タップ、昼食時に振る舞われる各種のカレー、親切に案内をしてくれた大勢のボランティアの学生など、何もかもしっかりしており、準備に携わった現地の方々に感謝したい。

発表の傾向については、問題設定や精度評価において完成度が求められる ACL や EMNLP などと比べると、Coling では比較的基準が緩いと言われているが、今回もその通りの印象であった。そのぶん、主観的にしか評価できないタスクに取り組んだもの、新たな応用に繋がりそうな萌芽的な研究や、他分野との境界領域にある研究発表を楽しむことができた。興味深い言語現象や、研究者人口の少ない言語を取り上げたものが積極的に評価されるのも良い点である。

以下では、言語を社会学的な側面と結びつける試みをしている論文を中心に、5 編ほど紹介する。必ずしも高い精度やエレガントな手法が示されているわけではないものの、今後の研究のヒントとなるものが感じられた。

<sup>†</sup>日本アイ・ピー・エム株式会社 東京基礎研究所, IBM Research - Tokyo, IBM Japan, Ltd.

<sup>1</sup>すなわちオリンピックの年ごとに。

<sup>2</sup>Coling2004 (於 Genève) の夜に「ワカモノの会」と称して飲みに行ったのが、後の「NLP 若手の会 (YANS)」に発展したと勝手に思っている。

## 2 論文紹介

### 2.1 Automatic Detection of Point of View Differences in Wikipedia

一本目 (Al Khatib, Schütze, and Kantner 2012) は、同じ事物を描写する時にどれだけ異なる観点で見ているかを指標化する試みであり、Wikipedia の人物のエントリのうち、英語版・アラビア語版の対応が取れたものを題材としている。イスラエル系・アラブ系の人物に関する文書は特に、書き手の立場が変われば観点が異なる (例: 一方にとってのテロリストも、他方にとっては英雄) という点に着目し、両言語の文書の中の positive/negative の度合いを判定し、その値の差を「観点の違い」とみなした。

正解データとして、英語・アラビア語の記述のどちらが positive な態度で述べているかを 5 段階で示したものを作成した。また、クラウドソーシングにより文単位で positive/negative のタグを付けたものを学習に用いる。

実験により、イスラエルの人物は軒並み英語のほうが positive な割合が多く、アラブ諸国の人物はアラビア語のほうが positive な割合が多いという、直感通りの結果が得られた。否定文などによる sentiment の推定誤りや、対象の人物と直接の関係がない sentiment を同定してしまうのが観点の差の誤判定の原因となっていた。

この論文の良いところは、観測しづらい「観点の差」を定量化したところ、政治的に難しそうなテーマであるものの両方の立場からの著者らにより堂々と扱っているところや、クラウドソーシングの活用についての知見が述べられている点である。

一方で、sentiment の同定は bag of word や n-gram のようなナイーブな手法に頼っている点や、agreement が低くタスクとして妥当なのか疑わしい点、両言語の特性に関する議論がもう少しほしいところなどに、改善の余地がある。

### 2.2 NEER: An Unsupervised Method for Named Entity Evolution Recognition

二本目 (Tahmasebi, Gossen, Kanhabua, Holzmann, and Risse 2012) は、1995 年に都市名が Bombay から Mumbai に変わったように、社会的な理由などで一気に変化するような Named Entity を検出するというタスクであり、New York Times のコーパスを題材にしている。

手法としては、query の語の急激な頻度の上昇 (burst) を検出し、その時間範囲を変化点とみなし、共通する要素を持つ固有名詞の規則や、共通する文脈を持つ語を「時間をまたがる共参照」(temporal co-reference) の候補とし、複数の関連度の指標を用いてフィルタリングするというものである。

予め「変更された名称」の正解セットを用意しておき、それらが時系列を持った新聞記事コーパスからどれだけ再現できるかを測る。変化点 (burst) の検出精度は 66% であり、正解の burst

を用いた場合と、自動的に検出した場合を比較した。

この論文の良いところは、固有名詞の変更という社会的な現象を、言語の変化 (language evolution) の現象の一側面と位置づけている点、そして同義語とは微妙に違う単語相互の関係について考えている点である。しかし、名称の変化と、固有表現の曖昧性の問題が混ざっている (結果の中に、ミドルネームや役職名の有無のバリエーションが目立つ) など、固有表現抽出一般の問題が解決できていないのが残念である。

以上の2本は、表現と事物、言語と社会の関係を突き詰めると面白くなっていくと考えられる。特に、同義・類義表現の中にある相違点を見出して知識発見に役立てる試み (那須川, 荻野, 西山, 金山 2012) に通じるものがあり、興味深い。

### 2.3 Who's (Really) the Boss? Perception of Situational Power in Written Interactions

三本目 (Prabhakaran, Rambow, and Diab 2012) は、組織内の対話の中において、階層による力関係 (hierarchical power) とは異なる観点の、状況に応じた力関係 (situational power) がどのように出現するかを調べる研究である。

題材として、破綻した企業の社内メールを用いている。従来の研究の中には社内の上下関係 (上司と部下) をメール内の言語現象を用いて推定するものがあつた (Bramsen, Escobar-Molano, Patel, and Alonso 2011) が、業務上の命令のような力関係がどのように働くかがより重要であるという仮説のもとで、situational power を新たにタグ付けしたデータを用いて実験している。

dialog act の種別や、その他の表層情報を用いた機械学習により、situational power の有無を判定するものであるが、実験が小規模であるうえに結果の精度が低いという点、人手で付けたタグ (特に overt display of power) 自身が解こうとしている問題に極めて近い点などが残念である。しかし、評価のしやすさを重視した従来研究とは視点を変えて、社会と言語の関係を調べようとしている点で面白いと感じた。

### 2.4 Automatic Detection of Psychological Distress Indicators and Severity Assessment from Online Forum Posts

Saleem らの研究 (Saleem, Prasad, Vitaladevuni, Pacula, Crystal, Marx, Sloan, Vasterling, and Speroff 2012) は、ストレス障害や脳損傷の患者とその介護士のオンラインの会話から、自殺傾向の指標を推定するというものである。文の長さ、句読点の数や、positive/negative な語の数、そして「死にたい」「この世の終わり」のようなフレーズの有無などを素性として、SVM や Markov Logic Network で学習するというものである。素性の選択や機械学習の手法にさほどの目新しさはなく、高い精度が得られているとはいえないが、このように周辺領域からの発表をもとに行われる議論は興味深い。心理言語学で言われている、否定の数や long dependency との関連

などが考察されているとより良いと感じた。

## 2.5 Towards a Generic and Flexible Citation Classifier Based on a Faceted Classification Scheme

最後 (Jochim and Schütze 2012) は、比較的しっかりとした論文で、論文中の文献の引用を役割ごとに分類するものである。引用の分類については既にいくつかの試みがある (Teufel, Siddharthan, and Tidhar 2006) が、この論文では4種の独立な観点の分類を、他の研究で使われた手法や新しい手法を網羅しながら丁寧に実験しており、これまでの引用意図推定の研究の集大成ともいえる。

この論文で面白いのは、論文中の引用すべてに、題材となっている引用意図のタグを付与している点である。論文そのものが実例となっているとタスクの意義がわかりやすくなるし、自然言語処理の研究を自然言語で述べているのだから、他の論文でも自己完結型の例示ができるのではないか<sup>3</sup>。

## 3 おわりに

以上の論文は、教科書的に優れている論文とは限らず、改善の余地がある点も多い。しかし、新しい理論や応用の研究は、最初から評価されるものではないし、実際にこれらの論文は新たな想像をかきたてるのに充分であった。Coling は人が集まる場としても高い価値を保っており、ACL との棲み分けもできているので、このような機会を今後も大切にしていきたい。

最後に日本人研究者へ。今回、日本からの参加者が58名もいたものの、質疑応答の場面での日本人の発言はとても少ないと感じた。これだけの研究者コミュニティがあるのだから、もっと存在感を示せてもよいのではないか。あるセッションの座長は、国際会議の場では英語が母国語でない人が多いのだし、発音の違いなどは気にせずに話してほしい、ということをお話してくれて、会場でも共感を得ていた。どんな質問でも誰もが耳を傾けてくれるので、特に学生の方は、必ず会期中に数回（自分の発表で受けた質問の数など）は発言するつもりで参加するとよいと思う。

## 謝辞

現地にて会議の内容やその他諸々の情報共有をしてくださった皆様、情報処理学会 NL 研の報告会で議論をしていただいた皆様に感謝いたします。

---

<sup>3</sup>余談だが、私自身は構文解析に関する論文で、論文中のすべての文を正しく parse できるようにしようとして挫折したことがある。

## 参考文献

- Al Khatib, K., Schütze, H., and Kantner, C. (2012). “Automatic Detection of Point of View Differences in Wikipedia.” In *Proceedings of COLING 2012*, pp. 33–50.
- Bramsen, P., Escobar-Molano, M., Patel, A., and Alonso, R. (2011). “Extracting social power relationships from natural language.” In *Proceedings of ACL*, pp. 773–782.
- Jochim, C. and Schütze, H. (2012). “Towards a Generic and Flexible Citation Classifier Based on a Faceted Classification Scheme.” In *Proceedings of COLING 2012*, pp. 1343–1358.
- Prabhakaran, V., Rambow, O., and Diab, M. (2012). “Who’s (Really) the Boss? Perception of Situational Power in Written Interactions.” In *Proceedings of COLING 2012*, pp. 2259–2274.
- Saleem, S., Prasad, R., Vitaladevuni, S., Pacula, M., Crystal, M., Marx, B., Sloan, D., Vasterling, J., and Speroff, T. (2012). “Automatic Detection of Psychological Distress Indicators and Severity Assessment from Online Forum Posts.” In *Proceedings of COLING 2012*, pp. 2375–2388.
- Tahmasebi, N., Gossen, G., Kanhabua, N., Holzmann, H., and Risse, T. (2012). “NEER: An Unsupervised Method for Named Entity Evolution Recognition.” In *Proceedings of COLING 2012*, pp. 2553–2568.
- Teufel, S., Siddharthan, A., and Tidhar, D. (2006). “Automatic classification of citation function.” In *Proceedings of EMNLP*, pp. 103–110.
- 那須川, 荻野, 西山, 金山 (2012). “同義扱いされる表現を弁別する.” 言語処理学会第 18 回年次大会, pp. 167–170.

## 略歴

**金山 博 (正会員)** : 2000 年より日本アイ・ビー・エム株式会社東京基礎研究所にて、構文解析、意味解析などの自然言語処理の基礎技術、および翻訳、評判分析、文書校正、質問応答などへの応用に関する研究に従事している。博士 (情報理工学)。

(2012 年 11 月 30 日依頼)

(2013 年 1 月 21 日受付)