

COLING2014 参加報告 (その 1)

— 概要 —

網川 隆司[†]

1 はじめに

本稿では、2014 年 8 月 23 日から 29 日にかけてアイルランド・ダブリンの Dublin City University にて開催された国際会議 The 25th International Conference on Computational Linguistics (COLING 2014) について報告する。COLING は、The International Committee on Computational Linguistics (ICCL) の後援により 1960 年代 から 2 年に 1 度開催されている、計算言語学・自然言語処理分野に関する代表的な査読付き国際会議の一つである。8 月 23 日から 24 日にかけては、SemEval-2014 を含む 18 件の併催ワークショップ、および 6 件のチュートリアルが開催され、25 日から 29 日まで本会議が開催された。本稿では特に、採択率、参加者数、論文の分野の分布等の概要について報告する。

2 投稿数と採択率

COLING 2014 には口頭発表、ポスター発表を合わせて 691 件の投稿があり、うち、採択されたものは 218 件で、全体の採択率は 31.5%であった。最近 5 回の投稿数、採択数、採択率の推移を表 1 に示した¹。両者を合わせた採択率は 25%から 43%の間で幅があるが、口頭発表に限っては概ね 20%前後で安定している。2012 年は例外的に投稿数が 1000 件を超えていたが、今回の投稿数は例年並みであった。

国別の採択数の概要を図 1 に、国別の採択率を図 2 に示した。採択率は投稿数が 10 以上の国のみを示した。欧州開催の学会にもかかわらず、中国の採択数が米国を上回って最も多く、大きな存在感を示している。日本はドイツに次ぎ 4 番目に多かったが、一方で採択率においては平均を下回った。

原稿の投稿形式について、前回、モバイル端末での可読性向上のために B5 の 1 段組の形式が

[†] 静岡大学大学院情報学研究所, Graduate School of Informatics, Shizuoka University

¹ データは各年度の Proceedings の Front Matter を参照して記した。

表 1 最近の COLING の投稿数, 採択数, 採択率の推移

開催年	開催地	投稿数	採択数 (口頭+ポスター)	採択率 (口頭+ポスター)
2006 (ACL 共催)	シドニー	630	272 (147 + 125)	43% (23% + 20%)
2008	マンチェスター	600	180 (145 + 35)	30% (24% + 6%)
2010	北京	815	334 (155 + 179)	41% (19% + 22%)
2012	ムンバイ	-	333 (195 + 138)	~25%
2014	ダブリン	685	217 (138 + 79)	32% (20% + 12%)

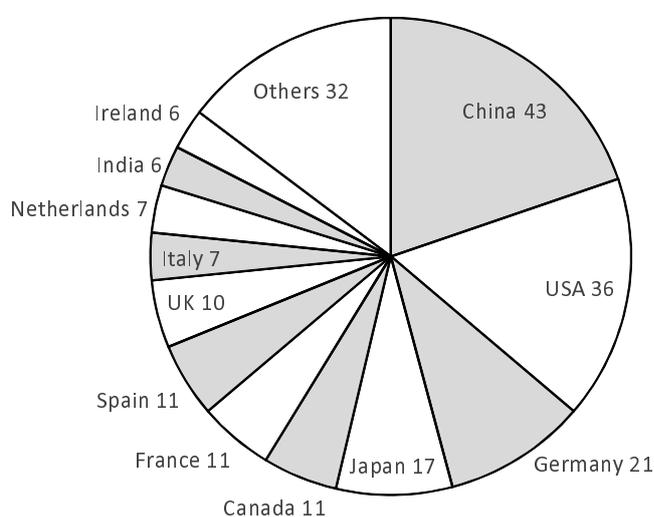


図 1 国別の採択数

採用されたが、今回は A4 の 1 段組の形式に変更された。また、第二言語によるタイトル・概要の記述、およびキーワードの記述は今回は採用されなかった。

3 参加者の傾向

開会直後時点での参加者数を表 2 にまとめる。参加者総数は 800 名を超す規模となった。国別では、開催国アイルランドとその隣国の英国からの参加者が多く、欧州からの参加者が半数以上を占めている。一方、前回はビザの問題で少なかったと思われる中国からの参加者が増加し、日本は前回より微減となった。今後も日本を拠点とする研究者の積極的な参加が望まれる。

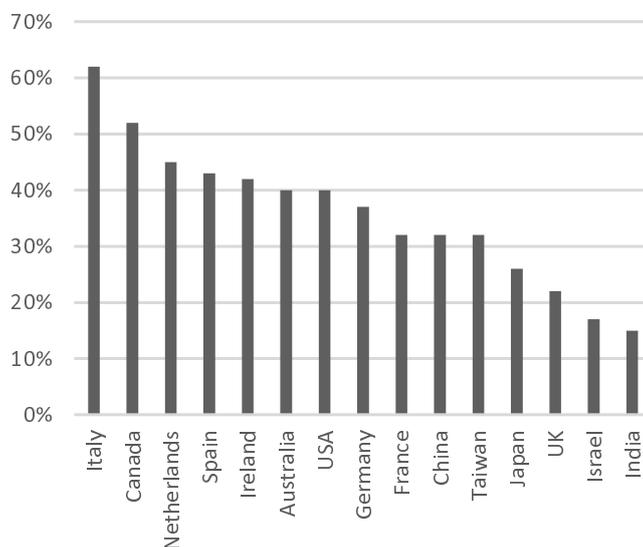


図 2 国別の採択率 (投稿数 10 以上の国のみ)

表 2 国別の参加者数

国籍	参加者数	割合	国籍	参加者数	割合
米国	102	12%	イタリア	32	4%
ドイツ	89	11%	スペイン	30	4%
英国 (UK)	88	11%	オランダ	21	3%
アイルランド	73	9%	カナダ	20	2%
中国	67	8%	チェコ	16	2%
フランス	57	7%	インド	16	2%
日本	45	5%			

4 採択論文の分野

会議に採択された論文のうち、口頭発表の分野について、発表セッションごとに表 3 にまとめた。例年通り、各分野で多様な論文が採択されているが、特に、談話モデル、評判分析、意見マイニング、情報抽出といった分野の発表が増加した。また、分布仮説に基づく意味論 (Distributional Semantics) や語彙意味論に関する発表も多い印象を受けた。ニューラルネットワークによる単語の新たなベクトル表現として昨年から注目を集める、word2vec (Mikolov, Chen, Corrado, and Dean 2013) を参照する論文が少なくとも 10 件あり、評判分析や構文解析への応用もみられ、今

表 3 口頭発表の分野

セッション	発表数	割合
Modeling of Discourse and Dialogue, and Multimodal Processing	16	12%
Sentiment Analysis, Opinion Mining and Social Media	12	9%
Syntax, Grammar Induction, Syntactic and Semantic Parsing	12	9%
Machine Translation	11	8%
Semantic Processing, Distributional Semantics and Compositional Semantics	11	8%
Lexical Semantics and Ontologies	11	8%
Information Extraction/Database Linking	11	8%
Applications	8	6%
Natural Language Generation and Summarisation and Paraphrasing	8	6%
Linguistic and Cognitive Issues in CL and NLP	8	6%
Morphology, Word Segmentation, Tagging and Chunking	8	6%
Information Retrieval and Question Answering	4	3%
Machine Learning for CL and NLP	4	3%
Software, Tools	4	3%
Language Resources	4	3%
Speech Recognition, Text-to-Speech, Spoken Language Understanding	3	2%
Lesser Resourced Languages	3	2%

後の発展が期待される。

5 Best Paper Awards

会議の最後に以下の2つの論文がベストペーパーとして表彰された。“Relation Classification via Convolutional Deep Neural Network” (Zeng, Liu, Lai, Zhou, and Zhao 2014) では、関係分類における素性学習において、畳み込みニューラルネットワークを用いることで、既存のNLPツール（品詞タグ付け、固有表現抽出、構文解析等）を用いることなく従来と同等以上の性能を達成することを示した。“A context-based model for Sentiment Analysis in Twitter” (Vanzo, Croce, and Basili 2014) は、Twitterにおける評判分析において、ツイートのトピックといった、より広い文脈を扱うことを狙いとしている。ツイートの列を逐次的に分類するSVMのマルコフ定式化モデルを提案し、 F_1 値で20%の性能向上を示した。

6 おわりに

本稿では国際会議 COLING 2014 の概要について紹介した。会議の運営や会場については大きな問題もなく、ポスター発表では、比較的広いスペースがあったことで議論がしやすかったことが印象的であった。今回の COLING は全体として、成功裏に終了したといえる。次回の COLING は、2016 年に大阪にて開催予定である。日本での開催は 1980 年の東京以来 36 年ぶりとなり、関連する国際会議としては名古屋で開催された IJCNLP 2013 から 3 年ぶりとなる。近年、「ビッグデータ」が流行語としてノミネートされるなど、日本における自然言語処理分野への認知度は昔に比べて高まっている。次回は日本における活発な研究活動をアピールする絶好の機会の一つであり、数多くの完成度の高い研究が発表されることを期待したい。

参考文献

- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). “Efficient Estimation of Word Representations in Vector Space.” *CoRR*, **abs/1301.3781**.
- Vanzo, A., Croce, D., and Basili, R. (2014). “A context-based model for Sentiment Analysis in Twitter.” In *Proceedings of the 25th International Conference on Computational Linguistics*, pp. 2345–2354.
- Zeng, D., Liu, K., Lai, S., Zhou, G., and Zhao, J. (2014). “Relation Classification via Convolutional Deep Neural Network.” In *Proceedings of the 25th International Conference on Computational Linguistics*, pp. 2335–2344.

略歴

綱川 隆司（正会員）：2008年 東京大学大学院情報理工学系研究科コンピュータ科学専攻博士課程を単位取得後退学。博士（情報理工学）（東京大学）。同特任研究員，静岡大学情報学部学術研究員を経て，現在，静岡大学大学院情報学研究科 助教。多言語を扱う自然言語処理に関する研究に従事。言語処理学会会員。

(20xx年x月xx日依頼)

(20xx年x月xx日受付)