COLING 2018

August 20-26, 2018, Santa Fe, New Mexico, USA

COLING 2018 Best papers

Posted on **June 14, 2018** by **Leon Derczynski**

There are multiple categories of award at COLING 2018, as we laid out in an earlier blog post. We received 44 nominations for best papers over ten categories, and conferred best paper awards in the categories as follows:

- **Best error analysis:** SGM: Sequence Generation Model for Multi-label Classification, by Pengcheng Yang, Xu Sun, Wei Li, Shuming Ma, Wei Wu and Houfeng Wang.
- **Best evaluation:** SGM: Sequence Generation Model for Multi-label Classification, by Pengcheng Yang, Xu Sun, Wei Li, Shuming Ma, Wei Wu and Houfeng Wang.
- **Best linguistic analysis:** Distinguishing affixoid formations from compounds, by Josef Ruppenhofer, Michael Wiegand, Rebecca Wilm and Katja Markert
- **Best NLP engineering experiment:** Authorless Topic Models: Biasing Models Away from Known Structure, by Laure Thompson and David Mimno
- **Best position paper:** Arguments and Adjuncts in Universal Dependencies, by Adam Przepiórkowski and Agnieszka Patejuk
- **Best reproduction paper:** Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering, by Wuwei Lan and Wei Xu
- **Best resource paper:** AnlamVer: Semantic Model Evaluation Dataset for Turkish – Word Similarity and Relatedness, by Gökhan Ercan and Olcay Taner Yıldız
- **Best survey paper:** A Survey on Open Information Extraction, by Christina Niklaus, Matthias Cetto, André Freitas and Siegfried Handschuh
- **Most reproducible:** Design Challenges and Misconceptions in Neural Sequence Labeling, by Jie Yang, Shuailong Liang and Yue Zhang

Note that, as announced last year, for open science & reproducibility COLING 2018 did not confer best paper awards to paper that could not make the code/resources publicly

available by camera ready time. This means you can ask the best paper authors for associated data and programs right now, and they should be able to provide you with a link.

In addition, we would like to note the following papers as "Area Chair Favorites", which were nominated by reviewers and recognised as excellent by chairs.

- Visual Question Answering Dataset for Bilingual Image Understanding: A study of cross-lingual transfer using attention maps. Nobuyuki Shimizu, Na Rong and Takashi Miyazaki
- Using J-K-fold Cross Validation To Reduce Variance When Tuning NLP Models. Henry Moss, David Leslie and Paul Rayson
- Measuring the Diversity of Automatic Image Descriptions. Emiel van Miltenburg, Desmond Elliott and Piek Vossen
- Reading Comprehension with Graph-based Temporal-Causal Reasoning. Yawei Sun, Gong Cheng and Yuzhong Qu
- Diachronic word embeddings and semantic shifts: a survey. Andrey Kutuzov, Lilja Øvrelid, Terrence Szymanski and Erik Velldal
- Transfer Learning for Entity Recognition of Novel Classes. Juan Diego Rodriguez, Adam Caldwell and Alexander Liu
- Joint Modeling of Structure Identification and Nuclearity Recognition in Macro Chinese Discourse Treebank. Xiaomin Chu, Feng Jiang, Yi Zhou, Guodong Zhou and Qiaoming Zhu
- Unsupervised Morphology Learning with Statistical Paradigms. Hongzhi Xu, Mitchell Marcus, Charles Yang and Lyle Ungar
- Challenges of language technologies for the Americas indigenous languages. Manuel Mager, Ximena Gutierrez-Vasques, Gerardo Sierra and Ivan Meza-Ruiz
- A Lexicon-Based Supervised Attention Model for Neural Sentiment Analysis. Yicheng Zou, Tao Gui, Qi Zhang and Xuanjing Huang
- From Text to Lexicon: Bridging the Gap between Word Embeddings and Lexical Resources. Ilia Kuznetsov and Iryna Gurevych
- The Road to Success: Assessing the Fate of Linguistic Innovations in Online Communities. Marco Del Tredici and Raquel Fernández
- Relation Induction in Word Embeddings Revisited. Zied Bouraoui, Shoaib Jameel and Steven Schockaert
- Learning with Noise-Contrastive Estimation: Easing training by learning to scale. Matthieu Labeau and Alexandre Allauzen
- Stress Test Evaluation for Natural Language Inference. Aakanksha Naik, Abhilasha Ravichander, Norman Sadeh, Carolyn Rose and Graham Neubig
- Recurrent One-Hop Predictions for Reasoning over Knowledge Graphs. Wenpeng Yin, Yadollah Yaghoobzadeh and Hinrich Schütze

- SMHD: a Large-Scale Resource for Exploring Online Language Usage for Multiple Mental Health Conditions. Arman Cohan, Bart Desmet, Andrew Yates, Luca Soldaini, Sean MacAvaney and Nazli Goharian
- Automatically Extracting Qualia Relations for the Rich Event Ontology. Ghazaleh Kazeminejad, Claire Bonial, Susan Windisch Brown and Martha Palmer
- What represents "style" in authorship attribution?. Kalaivani Sundararajan and Damon Woodard
- SeVeN: Augmenting Word Embeddings with Unsupervised Relation Vectors. Luis Espinosa Anke and Steven Schockaert
- GenSense: A Generalized Sense Retrofitting Model. Yang-Yin Lee, Ting-Yu Yen, Hen-Hsen Huang, Yow-Ting Shiue and Hsin-Hsi Chen
- A Multi-Attention based Neural Network with External Knowledge for Story Ending Predicting Task. Qian Li, Ziwei Li, Jin-Mao Wei, Yanhui Gu, Adam Jatowt and Zhenglu Yang
- Abstract Meaning Representation for Multi-Document Summarization. Kexin Liao, Logan Lebanoff and Fei Liu
- Cooperative Denoising for Distantly Supervised Relation Extraction. Kai Lei, Daoyuan Chen, Yaliang Li, Nan Du, Min Yang, Wei Fan and Ying Shen
- Dialogue Act Driven Conversation Model: An Experimental Study. Harshit Kumar, Arvind Agarwal and Sachindra Joshi
- Dynamic Multi-Level, Multi-Task Learning for Sentence Simplification. Han Guo, Ramakanth Pasunuru and Mohit Bansal
- A Knowledge-Augmented Neural Network Model for Implicit Discourse Relation Classification. Yudai Kishimoto, Yugo Murawaki and Sadao Kurohashi
- Abstractive Multi-Document Summarization using Paraphrastic Sentence Fusion. Mir Tafseer Nayeem, Tanvir Ahmed Fuad and Yllias Chali
- They Exist! Introducing Plural Mentions to Coreference Resolution and Entity Linking. Ethan Zhou and Jinho D. Choi
- A Comparison of Transformer and Recurrent Neural Networks on Multilingual NMT. Surafel Melaku Lakew, Mauro Cettolo and Marcello Federico
- Expressively vulgar: The socio-dynamics of vulgarity and its effects on sentiment analysis in social media. Isabel Cachola, Eric Holgate, Daniel Preoţiuc-Pietro and Junyi Jessy Li
- On Adversarial Examples for Character-Level Neural Machine Translation. Javid Ebrahimi, Daniel Lowd and Dejing Dou
- Neural Transition-based String Transduction for Limited-Resource Setting in Morphology. Peter Makarov and Simon Clematide
- Structured Dialogue Policy with Graph Neural Networks. Lu Chen, Bowen Tan, Sishan Long and Kai Yu

We would like to recognise with exceptional thanks our best paper committee.

This entry was posted in **PC Blog** by **Leon Derczynski**. Bookmark the **permalink [https://coling2018.org/coling-2018-best-papers/]** .

Jason Eisner
on **June 16, 2018 at 2:42 pm** said:

Thanks – I like the use of multiple categories. I can't help noticing that this year's awards (at least in name) seem to focus on the evaluation or facilitation of existing ideas. Some award categories you might consider in the future:

* Best mathematical model [generative probability model, grammar formalism, etc.]
* Best theoretical result
* Best new algorithm
* Best new problem

as well as other kinds of insight:

* Best generalization, synthesis, or illumination of previous work
* Best exposition of a difficult concept

Followup: The award names may have been misleading (to me, anyway). One thing that I liked about the new awards is that they seemed to be calling out particular noteworthy *elements* of the paper, such as whether the linguistic analysis of the data was exemplary (regardless of the rest of the paper). However, I'm now doubting this interpretation, since I see from https://coling2018.org/best-paper-categories-and-requirements/ that several of the 2018 awards were actually named after the paper types.

* So "Best linguistic analysis" probably just means "Best paper in the category 'Computationally aided linguistic analysis'". So the paper's key contribution might actually have been a new mathematical model. But the award title sounds like it's rewarding a good manual analysis of linguistic data (as in a linguistics journal) or a good linguistic error analysis of system output.
* Similarly, "Best NLP engineering experiment" probably just means "Best paper in the category 'NLP engineering experiment paper'". So the paper's key contribution might have been a new task or a new algorithm. But the award title sounds like the award is given for carefully comparing the performance of two methods in an engineering setting.

* It's a pity that there are no categories for theoretical papers (formal language theory, computational complexity, algorithms, etc.), as some commenters already noted at http://coling2018.org/call-for-input-paper-types-and-associated-review-forms/ . One of the NAACL 2018 outstanding paper awards was given to such a paper.

Leon Derczynski
on **July 2, 2018 at 9:04 am** said:

Thanks Jason. We tried to capture community feedback last year when structuring the awards, as you link to, and this did impact the final configuration. I share your impression about the theoretical papers – though based on the discussion in comments on the page about paper types, as well as discussions on mailing lists and direct emails, it quickly became evident that interpretation of this category was wildly diverse, and so we couldn't create recognition under this label and then award it in a consistent manner.

While awards are named after paper types, ACs were briefed to award them as they saw fit. The unconscious bias given by the naming suited well – recognition must be broader than just NLPEE papers.

Regarding changes at the next COLING – the best bet for implementing these suggestions is to run that event!

Ravin
on **June 21, 2018 at 2:17 am** said:

When will we know that whether the paper is oral or poster？

Leon Derczynski
on **June 28, 2018 at 9:55 pm** said:

The notifications were sent out today.