

The 33rd

ROCLING 2021

第三十三屆自然語言與語音處理研討會

October 15-16, 2021, Taoyuan, Taiwan, R.O.C.

Proceedings of the Thirty-third Conference on Computational
Linguistics and Speech Processing

**ROCLING 2021: The 33rd Conference on
Computational Linguistics and Speech Processing**

第三十三屆自然語言與語音處理研討會

October 15-16, 2021

National Central University, Taoyuan, Taiwan, R.O.C.

主辦單位：

國立中央大學、國立臺灣科技大學、中華民國計算語言學學會

共同主辦單位：

科技部人工智慧技術暨全幅健康照護聯合研究中心、人工智慧普適研究中心、科技部人工智慧生技醫療創新研究中心

協辦單位：

科技部、教育部

贊助單位：

玉山金控、賽微科技股份有限公司、中華電信研究院

First Published in September 2021

By The Association for Computational Linguistics and Chinese Language Processing
(ACLCLP)

Copyright©2021 the Association for Computational Linguistics and Chinese Language
Processing (ACLCLP), Authors of Papers

Each of the authors grants a non-exclusive license to the ACLCLP to publish the paper
in printed form. Any other usage is prohibited without the express permission of the
author who may also retain the on-line version at a location to be selected by him/her.

Lung-Hao Lee, Chia-Hui Chang, Kuan-Yu Chen, Yung-Chun Chang, Yi-Chin Huang,
Hung-Yi Lee, Jheng-Long Wu, Chun-Hsien Hsu, Liang-Chih Yu (eds.)

Proceedings of the 33rd Conference on Computational Linguistics and Speech
Processing (ROCLING XXXIII)

2021-10-15—2021-10-16

ACLCLP

October 2021

ISBN: 978-986-95769-4-9

Organizing Committee

Honorary Chair

Jing-Yang Jou, National Central University

Conference Chairs

Lung-Hao Lee, National Central University

Chia-Hui Chang, National Central University

Kuan-Yu Chen, National Taiwan University of Science and Technology

Program Chairs

Yung-Chun Chang, Taipei Medical University

Yi-Chin Huang, National Pingtung University

Tutorial Chair

Hung-Yi Lee, National Taiwan University

Publication Chair

Jheng-Long Wu, Soochow University

Special Session Chair

Chun-Hsien Hsu, National Central University

Shared Task Chair

Liang-Chih Yu, Yuan Ze University

Program Committee

Jia-Wei Chang (張家瑋), National Taichung University of Science
Ru-Yng Chang (張如瑩), AI clerk international co., ltd.
Chung-Chi Chen (陳重吉), National Taiwan University
Yun-Nung Chen (陳縉儂), National Taiwan University
Yu-Tai Chien (簡宇泰), National Taipei University of Business
Hong-Jie Dai (戴鴻傑), National Kaohsiung University of Science and Technology
Min-Yuh Day (戴敏育), National Taipei University
Yu-Lun Hsieh (謝育倫), CloudMile
Wen-Lian Hsu (許聞廉), Asia University
Hen-Hsen Huang (黃瀚萱), Academia Sinica
Jieh-weih Hung (洪志偉), National Chi Nan University
Chih-Hao Ku (顧值豪), Cleveland State University
Ying-Hui Lai (賴穎暉), National Yang Ming Chiao Tung University
Cheng-Te Li (李政德), National Cheng Kung University
Chun-Yen Lin (林君彥), Taipei Medical University
Jen-Chun Lin (林仁俊), Academia Sinica
Szu-Yin Lin (林斯寅), National Ilan University
Shih-Hung Liu (劉士弘), Digiwin
Chao-Lin Liu (劉昭麟), National Chengchi University
Jenn-Long Liu (劉振隆), I-Shou University
Yi-Fen Liu (劉怡芬), Feng Chia University
Wen-Hsiang Lu (盧文祥), National Cheng Kung University
Shang-Pin Ma (馬尚彬), National Taiwan Ocean University
Emily Chia-Yu Su (蘇家玉), Taipei Medical University
Ming-Hsiang Su (蘇明祥), Soochow University
Richard Tzong-Han Tsai (蔡宗翰), National Central University
Chun-Wei Tung (童俊維), National Health Research Institutes
Hsin-Min Wang (王新民), Academia Sinica

Jenq-Haur Wang (王正豪), National Taipei University of Technology

Yu-Cheng Wang (王昱晟), Lunghwa University of Science and Technology

Jheng-Long Wu (吳政隆), Soochow University

Shih-Hung Wu (吳世弘), Chaoyang University of Technology

Jui-Feng Yeh (葉瑞峰), National Chiayi University

Liang-Chih Yu (禹良治), Yuan Ze University

Messages from Conference Chairs

As the Conference Chairs, we welcome you to the 33rd Conference on Computational Linguistics and Speech Processing (ROCLING 2021) in Taoyuan, Taiwan, during October 15-16, 2021. ROCLING 2021 is hosted by National Central University (NCU), National Taiwan University of Science and Technology (NTUST), and the Association for Computational Linguistics and Chinese Language Processing (ACLCLP) and co-organized by MOST Joint Research Center for AI Technology and All Vista Healthcare (AINTU), Pervasive Artificial Intelligence Research (PAIR) Labs, and MOST Artificial Intelligence Biomedical Research Center (AIBMRC).

We would like to thank the Program Chairs Prof. Yung-Chun Chang and Prof. Yi-Ching Huang, Tutorial Chair Prof. Hung-Yi Lee, Special Session Chair Prof. Chun-Hsien Hsu and Shared Task Chair Prof. Liang-Chih Yu for their hard work in coordinating the review process allowing for top quality papers and inspiring talks to be presented at the conference. We also thank Prof. Jheng-Long Wu for the publication of conference proceedings. The conference proceedings will be published from ACL Anthology.

Last but not least, we would like to thank all authors for submitting high-quality research papers, and all attendees for making the journey. Hope you all enjoy the conference program.

Lung-Hao Lee, National Central University

Chia-Hui Chang, National Central University

Kuan-Yu Chen, National Taiwan University of Science and Technology

ROCLING 2021 Conference Chairs

Messages from Program Chairs

The excellent program and activities of ROCLING 2021 are the result of collaborative efforts of more than 50 program committee members and conference organizers. Each paper has been reviewed by 2 to 3 PC members, and we thank all of them for their insightful reviews, from which we can build an outstanding technical program. We would also like to thank the Tutorial Chair, Prof. Hung-Yi Lee of National Taiwan University, for coordinating three excellent tutorials. We are very grateful to the Publication Chair, Prof. Jheng-Long Wu of the Soochow University, for editing the conference proceedings. We would also like to express our gratitude to the Special Session Chair, Prof. Chun-Hsien Hsu of National Central University, and Shared Task Chair, Prof. Liang-Chih Yu of Yuan Ze University, for organizing the special session and shared task that enable the outreach of conference events to many important communities. Last but not least, we appreciate the contributions of Conference Co-chairs, Prof. Lung-Hao Lee of National Central University, Prof. Chia-Hui Chang of National Central University, and Prof. Kuan-Yu Chen of National Taiwan University of Science and Technology, to the construction of the conference website and event coordination.

Yung-Chun Chang, Tapiei Medical University
Yi-Chin Huang, National Pingtung University
ROCLING 2021 Program Chairs

NLP Keynote by Prof. Vincent Ng



Event Coreference Resolution: Successes and Future Challenges

Speaker: Prof. Vincent Ng

Professor, The University of Texas at Dallas

Time: Friday, October 15, 2021, 09:10 - 10:10

Biography

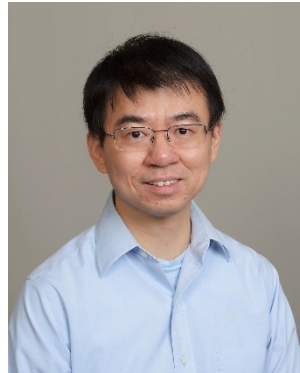
Vincent Ng is a Professor in the Computer Science Department at the University of Texas at Dallas. He is also the director of the Machine Learning and Language Processing Laboratory in the Human Language Technology Research Institute at UT Dallas. He obtained his B.S. from Carnegie Mellon University and his Ph.D. from Cornell University. His research is in the area of Natural Language Processing, focusing on the development of computational methods for addressing key tasks in information extraction and discourse processing.

Abstract

Recent years have seen a gradual shift of focus from entity-based tasks to event-based tasks in information extraction research. This talk will focus on event coreference resolution, the event-based counterpart of the notoriously difficult entity coreference resolution task. Specifically, I will examine the major milestones made in event

coreference research since its inception more than two decades ago, including the recent successes of neural event coreference models and their limitations, and discuss possible ways to bring these models to the next level of performance.

Speech Keynote by Dr. Jinyu Li



Advancing end-to-end automatic speech recognition

Speaker: Dr. Jinyu Li

Partner Applied Scientist and Technical Lead,
Microsoft Corporation, Redmond, USA

Time: Saturday, October 16, 2021, 09:00 - 10:00

Biography

Jinyu Li received the Ph.D. degree from Georgia Institute of Technology, Atlanta, in 2008. From 2000 to 2003, he was a Researcher in the Intel China Research Center and Research Manager in iFlytek, China. Currently, he is a Partner Applied Scientist and Technical Lead in Microsoft Corporation, Redmond, USA. He leads a team to design and improve speech modeling algorithms and technologies that ensure industry state-of-the-art speech recognition accuracy for Microsoft. His major research interests cover several topics in speech recognition, including end-to-end modeling, deep learning, noise robustness, etc. He is the leading author of the book "Robust Automatic Speech Recognition -- A Bridge to Practical Applications", Academic Press, Oct, 2015. He is the member of IEEE Speech and Language Processing Technical Committee since 2017. He also served as the associate editor of IEEE/ACM Transactions on Audio, Speech and Language Processing from 2015 to 2020.

Abstract

Recently, the speech community is seeing a significant trend of moving from deep neural network based hybrid modeling to end-to-end (E2E) modeling for automatic

speech recognition (ASR). While E2E models achieve the state-of-the-art results in most benchmarks in terms of ASR accuracy, hybrid models still dominate the commercial ASR systems at current time. There are lots of practical factors that affect the production model deployment decision. Traditional hybrid models, being optimized for production for decades, are usually good at these factors. Without providing excellent solutions to all these factors, it is hard for E2E models to be widely commercialized. In this talk, I will overview the recent advances in E2E models with the focus on technologies addressing those challenges from the perspective of industry. Specifically, I will describe methods of 1) building high-accuracy low-latency E2E models, 2) building a single E2E model to serve all multilingual users, 3) customizing and adapting E2E models to a new domain 4) extending E2E models for multi-talker ASR etc. Finally, I will conclude the talk with some challenges we should address in the future.

Table of Contents

Universal Recurrent Neural Network Grammar	1
A Study on Using Transfer Learning to Improve BERT Model for Emotional Classification of Chinese Lyrics	13
Nested Named Entity Recognition for Chinese Electronic Health Records with QA-based Sequence Labeling	18
AI Clerk Platform : Information Extraction DIY Platform	26
A Flexible and Extensible Framework for Multiple Answer Modes Question Answering	33
Speech Emotion Recognition Based on CNN+LSTM Model	43
A Study on Contextualized Language Modeling for Machine Reading Comprehension	48
Discussion on the relationship between elders' daily conversations and cognitive executive function: using word vectors and regression models	58
Chinese Medical Speech Recognition with Punctuated Hypothesis	63
Data centric approach to Chinese Medical Speech Recognition	72
Exploiting Low-Resource Code-Switching Data to Mandarin-English Speech Recognition Systems	81

Discussion on domain generalization in the cross-device speaker verification system	87
Integrated Semantic and Phonetic Post-correction for Chinese Speech Recognition	95
A Preliminary Study on Environmental Sound Classification Leveraging Large-Scale Pretrained Model and Semi-Supervised Learning	103
Mining Commonsense and Domain Knowledge from Math Word Problems	111
RCRNN-based Sound Event Detection System with Specific Speech Resolution	118
Exploring the Integration of E2E ASR and Pronunciation Modeling for English Mispronunciation Detection	124
Predicting elders' cognitive flexibility from their language use	132
Improve Chit-Chat and QA Sentence Classification in User Messages of Dialogue System using Dialogue Act Embedding	138
Data Augmentation Technology for Dysarthria Assistive Systems	144
A Survey of Approaches to Automatic Question Generation from 2019 to Early 2021	151
A BERT-based Siamese-structured Retrieval Model	163
MMTL: The Meta Multi-Task Learning for Aspect Category Sentiment Analysis	173

Home Appliance Review Research Via Adversarial Reptile	183
Unsupervised Multi-document Summarization for News Corpus with Key Synonyms and Contextual Embeddings	192
Aspect-Based Sentiment Analysis and Singer Name Entity Recognition using Parameter Generation Network Based Transfer Learning	202
Using Valence and Arousal-infused Bi-LSTM for Sentiment Analysis in Social Media Product Reviews	210
Numerical Relation Detection in Financial Tweets using Dependency-aware Deep Neural Network	218
Aggregating User-Centric and Post-Centric Sentiments from Social Media for Topical Stance Prediction	226
Employing low-pass filtered temporal speech features for the training of ideal ratio mask in speech enhancement	236
Hidden Advertorial Detection on Social Media in Chinese	243
Automatic Extraction of English Grammar Pattern Correction Rules	252
Multi-Label Classification of Chinese Humor Texts Using Hypergraph Attention Networks	257
Incorporating Domain Knowledge into Language Transformers for Multi-Label Classification of Chinese Medical Questions	265
What confuses BERT? Linguistic Evaluation of Sentiment Analysis on Telecom Customer Opinion	271

Generative Adversarial Networks based on Mixed-Attentions for Citation Intent Classification in Scientific Publications	280
A Corpus for Dimensional Sentiment Classification on YouTube Streaming Service	286
Improved Text Classification of Long-term Care Materials	294
Learning to Find Translation of Grammar Patterns in Parallel Corpus	301
Keyword-centered Collocating Topic Analysis	310
Confiscation Detection of Criminal Judgment Using Text Classification Approach	318
Incorporating speaker embedding and post-filter network for improving speaker similarity of personalized speech synthesis system	326
Identify Bilingual Patterns and Phrases from a Bilingual Sentence Pair	333
Extracting Academic Senses: Towards An Academic Writer's Dictionary	339
SCUDS at ROCLING-2021 Shared Task: Using Pretrained Model for Dimensional Sentiment Analysis Based on Sample Expansion Method	346
ntust-nlp-1 at ROCLING-2021 Shared Task: Educational Texts Dimensional Sentiment Analysis using Pretrained Language Models	354
ntust-nlp-2 at ROCLING-2021 Shared Task: BERT-based semantic analyzer with word-level information	360

CYUT at ROCLING-2021 Shared Task: Based on BERT and MacBERT	367
SoochowDS at ROCLING-2021 Shared Task: Text Sentiment Analysis Using BERT and LSTM	375
NCU-NLP at ROCLING-2021 Shared Task: Using MacBERT Transformers for Dimensional Sentiment Analysis	380
ROCLING-2021 Shared Task: Dimensional Sentiment Analysis for Educational Texts	385