Introduction

Welcome to the proceedings of the twentieth China National Conference on Computational Linguistics (20th CCL). The conference and symposium were hosted and co-organized by Inner Mongolia University, China.

CCL is an annual conference (bi-annual before 2013) that started in 1991. It is the flagship conference of the Chinese Information Processing Society of China (CIPS), which is the largest NLP scholar and expert community in China. CCL is a premier nation-wide forum for disseminating new scholarly and technological work in computational linguistics, with a major emphasis on computer processing of the languages in China such as Mandarin, Tibetan, Mongolian, and Uyghur.

The Program Committee selected 108 papers (77 Chinese papers and 31 English papers) out of 303 submissions for publication. The acceptance rate is 35.64%. The 108 papers cover the following topics:

– Machine Translation and Multilingual Information Processing (7)
– Fundamental Theory and Methods of Computational Linguistics (8)
– Minority Language Information Processing (11)
– Social Computing and Sentiment Analysis (5)
– Text Generation and Summarization (4)
– Information Retrieval, Dialogue and Question Answering (12)
– Linguistics and Cognitive Science (15)
– Language Resource and Evaluation (8)
– Knowledge Graph and Information Extraction (13)
– NLP Applications (25)

The final program for the 20th CCL was the result of intense work by many dedicated colleagues. We want to thank, first of all, the authors who submitted their papers, contributing to the creation of the high-quality program. We are deeply indebted to all the Program Committee members for providing high-quality and insightful reviews under a tight schedule, and extremely grateful to the sponsors of the conference. Finally, we extend a special word of thanks to all the colleagues of the Organizing Committee and secretariat for their hard work in organizing the conference, and to ACL Anthology for their assistance in publishing the proceedings in due time.

We thank the Program and Organizing Committees for helping to make the conference successful, and we hope all the participants enjoyed the first online CCL conference.

July 2021

Sheng Li, Maosong Sun, Yang Liu
Hua Wu, Kang Liu, Wanxiang Che
Organizers

**Conference Chairs**

Sheng Li  
Harbin Institute of Technology

Maosong Sun  
Tsinghua University, China

Yang Liu  
Tsinghua University, China

**Program Chairs**

Hua Wu  
Baidu, Inc.

Kang Liu  
Institute of Automation, Chinese Academy of Sciences, China

Wanxiang Che  
Harbin Institute of Technology, China

**Area Co-Chairs**

Renfen Hu  
Beijing Normal University, China

Xiaofei Lu  
The Pennsylvania State University, USA

Zhenghua Li  
Soochow University, China

Pengfei Liu  
Carnegie Mellon University, USA

Jiafeng Guo  
Institute of Computing Technology, CAS, China

Qingyao Ai  
The university of Utah, USA

Rui Yan  
Renmin University of China, China

Jing Li  
The Hong Kong Polytechnic University

Wenliang Chen  
Soochow University, China

Bowei Zou  
Institute for Infocomm Research, Singapore

Shujian Huang  
Nanjing University, China

Hui Huang  
University of Macau, China

Yating Yang  
Xinjiang Physics and Chemistry Institute, CAS, China

Chenhui Chu  
Kyoto University, Japan

Junsheng Zhou  
Nanjing Normal University, China

Wei Lu  
Singapore University of Technology and Design, Singapore

Rui Xia  
Nanjing University of Science and Technology, China

Lin Gui  
The University of Warwick, England

Chenliang Li  
Wuhan University, China

Jiashu Zhao  
Wilfrid Laurier University, Canada

**Local Arrangement Chairs**

Guanglai Gao  
Inner Mongolia University, China

Hongxu Hou  
Inner Mongolia University, China

**Evaluation Chairs**
Hongfei Lin  Dalian University of Technology, China
Wei Song  Capital Normal University, China

Publications Chairs

Shizhu He  Institute of Automation, CAS, China
Gaoqi Rao  Beijing Language and Culture University, China

Workshop Chairs

Xianpei Han  Institute of Software, CAS, China
Yanyan Lan  Tsinghua University, China

Sponsorship Chairs

Zhongyu Wei  Fudan University, China
Yang Feng  Institute of Computing Technology, CAS, China

Publicity Chair

Xipeng Qiu  Fudan University, China

Website Chair

Liang Yang  Dalian University of Technology, China

System Demonstration Chairs

Jiajun Zhang  Institute of Automation, CAS, China
Haofen Wang  Tongji University, China

Student Counseling Chair

Zhiyuan Liu  Tsinghua University, China

Student Seminar Chairs

Xin Zhao  Renmin University of China, China
Libo Qin  Harbin Institute of Technology, China
Finance Chair

Yuxing Wang  Tsinghua University, China
# Table of Content

融合零指代识别的篇章级机器翻译  
汪浩, 李军辉, 贡正仙 ................................................. 1

融合 XLM 词语表示的神经机器译文自动评价方法  
胡纬, 李茂西, 袁白莲, 王明文 ........................................ 13

利用语义关联增强的跨语言预训练模型的译文质量评估  
叶恒, 贡正仙 .......................................................... 23

基于模型不确定性约束的半监督汉缅神经机器翻译  
王琳钦, 余正涛, 毛存礼, 高盛祥, 满志博, 王振晗 .................................................................. 35

基于枢轴语言系统融合的词汇混淆网络神经机器翻译  
赵小兵, 金波, 孙媛 ...................................................... 46

基于义原表示学习的词向量表示方法  
于宁, 王江萍, 王立, 刘建毅 .......................................... 57

一种新的处理汉语动名超常搭配的方法  
汪梦翔 ........................................................................ 66

基于双编码器的医学文本中文分词  
宗源, 常宝宝 .......................................................... 76

中文连动句语义关系识别研究  
孙超, 曲维光, 魏庭新, 顾彦慧, 李斌, 周俊生 .................................................. 86

中文词语离合现象识别研究  
周露, 曲维光, 魏庭新, 周俊生, 李斌, 顾彦慧 .................................................. 97

先秦词网构建及楚汉对比研究  
卢雪晖, 徐会丹, 陈思瑜, 李斌 .......................................... 108

基于 BERT 的意图分类与槽填充联合方法  
覃俊, 马天宇, 刘晶, 帖军, 后琦 .......................................... 121

基于依存语法的偷抢类动词研究  
王珊, 刘晓骏 .......................................................... 130

基于中文信息与越南语句法指导的越南语事件检测  
陈龙, 郭军军, 张亚飞, 高盛祥, 余正涛 .......................................... 143

基于多层次预训练策略和多任务学习的端到端蒙汉语音翻译
王宁宁，飞龙，张晖

基于层间知识蒸馏的神经机器翻译
金畅，段仁翀，肖妮妮，段湘熠

融合多层语义特征图的缅甸语图像文本识别方法
刘福浩，毛存礼，余正涛，高盛祥，王琳钦，谢旭阳

JCapsR：一种联合胶囊神经网络的藏语知识图谱表示学习模型
孙媛，梁家亚，陈安东，赵小兵

基于阅读理解的汉越跨语言新闻事件要素抽取方法
朱恩昌，余正涛，高盛祥，黄宇欣，郭军军

面向机器阅读理解的高质量藏语数据集构建
孙媛，刘思思，陈超凡，旦正错，赵小兵

Ti-Reader：基于注意力机制的藏文机器阅读理解端到端网络模型
孙媛，陈超凡，刘思思，赵小兵

藏文文本校对评测集构建
三毛措，才智杰，道吉扎西

结合标签转移关系的多任务笑点识别方法
张童越，张绍武，徐博，杨亮，林鸿飞

基于时间注意力胶囊网络的维吾尔语情感分类模型
罗涵天，杨雅婷，董瑞，马博

基于HowNet的无监督汉语动词隐喻识别方法
张明昊，张冬瑜，林鸿飞

基于风格化嵌入的中文文本风格迁移
王晨光，林鸿飞，杨亮

基于双星型自注意力网络的搜索结果多样化方法
秦绪博，窦志成，朱余韬，文继荣

基于迭代信息传递和滑动窗口注意力的问题生成模型研究
陈千，高晓影，王素格，郭鑫

面向中文口语理解的基于依赖引导的字特征槽填充模型
朱展标，黄沛杰，张业兴，刘树东，张华林，黄均曜

利用图像描述与知识图谱增强表示的视觉问答
王屹超，王华林，许晨，张琰，王会珍，朱靖波

面向对话文本的实体关系抽取
结合边界预测和动态模板方法的槽填充模型
朱展标, 黄沛杰, 张业兴, 刘树东, 张华林, 黄均曜

近十年来澳门的词汇增长
王珊, 陈钊, 张昊迪

基于结构树库的状位动词语义分类及搭配库构建
邵田, 翟世权, 饶高琦, 肆恩东

基于序列到序列的中文 AMR 解析
黄子怡, 李军辉, 贡正仙

基于词信息嵌入的汉语构词结构识别研究
郑婳, 殷雅琦, 王悦, 代达劢, 刘扬

汉语语体特征的计量与分类研究
邰沁清, 饶高琦

多模态表述视域下的小学数学课堂语言计量初探
郑泽芝, 赵骞

替换类动词的句法语义分析
王珊, 吴乐

基于自动识别的委婉语历时性发展变化与社会共变研究
张辰麟, 王明文, 谭亦鸣, 尹明, 张心怡

基于篇章结构攻击的阅读理解任务探究
马树楷, 邹家杰, 丁鼐

中美学者学术英语写作中词汇难度特征比较研究 —— 以计算语言学领域论文为例
谢永慧, 刘洋, 杨尔弘, 杨麟儿

基于词汇链强化表征的篇章修辞结构分析研究
王金锋, 孔芳

不同类型噪声环境下言语理解的脑机制研究
耿立波, 薛紫炫, 杨亦鸣

回避类动词的句法语义
王珊, 刘晓骏

欺骗类动词的句法语义研究
王珊, 周洁

基于结构检索的汉语介动搭配知识库构建
Xiaoning Jia, Hongxu Hou, Nier Wu, Haoran Li and Xin Chang ........................................ 886

Incorporating translation quality estimation into Chinese-Korean neural machine translation
Feiyu Li, Yahui Zhao, Feiyang Yang and Rongyi Cui ......................................................... 906

Emotion Classification of COVID-19 Chinese Microblogs based on the Emotion Category Description
Xianwei Guo, Hua Lai, Yan Xiang, Zhengtao Yu and Yuxin Huang ..................................... 916

Multi-level Emotion Cause Analysis by Multi-head Attention Based Multi-task Learning
Xiangju Li, Shi Feng, Yifei Zhang and Daling Wang ......................................................... 928

Using Query Expansion in Manifold Ranking for Query-Oriented Multi-Document Summarization
Quanye Jia, Rui Liu and Jianying Lin ................................................................. 940

Jointly Learning Salience and Redundancy by Adaptive Sentence Reranking for Extractive Summarization
Ximing Zhang and Ruifang Liu ....................................................................................... 952

Incorporating Commonsense Knowledge into Abstractive Dialogue Summarization via Heterogeneous Graph Networks
Xiachong Feng, Xiaocheng Feng and Bing Qin ............................................................... 964

Enhancing Question Generation with Commonsense Knowledge
Xin Jia, Hao Wang, Dawei Yin and Yunfang Wu ............................................................. 976

Topic Knowledge Acquisition and Utilization for Machine Reading Comprehension in Social Media Domain
Zhixing Tian, Yuanzhe Zhang, Kang Liu and Jun Zhao ..................................................... 988

Category-Based Strategy-Driven Question Generator for Visual Dialogue
Yanan Shi, Yanxin Tan, Fangxiang Feng, Chunping Zheng and Xiaojie Wang ...................... 1000

From Learning-to-Match to Learning-to-Discriminate: Global Prototype Learning for Few-shot Relation Classification
Fangchao Liu, Xinyan Xiao, Lingyong Yan, Hongyu Lin, Xianpei Han, Dai Dai, Hua Wu and Le Sun ................................................................. 1012

Multi-Strategy Knowledge Distillation Based Teacher-Student Framework for Machine Reading Comprehension
Xiaoyan Yu, Qingbin Liu, Shizhu He, Kang Liu, Shengping Liu, Jun Zhao and Yongbin Zhou ......................................................................................... 1024

LRRA: A Transparent Neural-Symbolic Reasoning Framework for Real-World Visual Question Answering
Zhang Wan, Keming Chen, Yujie Zhang, Jinan Xu and Yufeng Chen .................................. 1037

Meaningfulness and unit of Zipf’s law: evidence from danmu comments

xii
Yihan Zhou

Unifying Discourse Resources with Dependency Framework
Yi Cheng, Sujian Li and Yueyuan Li

A Chinese Machine Reading Comprehension Dataset Automatic Generated Based on Knowledge Graph
Hanyu Zhao, Sha Yuan, Jiahong Leng, Xiang Pan, Zhao Xue, Quanyue Ma and Yangxiao Liang

Morphological Analysis Corpus Construction of Uyghur
Gulinigeer Abudouwaili, Kahaerjiang Abiderexiti, Jiamila Wushouer, Yunfei Shen, Turenisha Maimaitimin and Tuergen Yibulayin

Improving Entity Linking by Encoding Type Information into Entity Embeddings
Tianran Li, Erguang Yang, Yujie Zhang, Yufeng Chen and Jinan Xu

A Unified Representation Learning Strategy for Open Relation Extraction with Ranked List Loss
Renze Lou, Fan Zhang, Xiaowei Zhou, Yutong Wang, Minghui Wu and Lin Sun

NS-Hunter: BERT-Cloze Based Semantic Denoising for Distantly Supervised Relation Classification
Tielin Shen, Daling Wang, Shi Feng and Yifei Zhang

A Trigger-Aware Multi-Task Learning for Chinese Event Entity Recognition
Yangxiao Xiang and Chenliang Li

Improving Low-Resource Named Entity Recognition via Label-Aware Data Augmentation and Curriculum Denoising
Wenjing Zhu, Jian Liu, Jinan Xu, Yufeng Chen and Yujie Zhang

Global entity alignment with Gated Latent Space Neighborhood Aggregation
Wei Chen, Xiaoying Chen and Shengwu Xiong

Few-Shot Charge Prediction with Multi-Grained Features and Mutual Information
Han Zhang, Yutao Zhu, Zhicheng Dou and Ji-Rong Wen

Sketchy Scene Captioning: Learning Multi-Level Semantic Information from Sparse Visual Scene Cues
Lian Zhou, Yangdong Chen and Yuejie Zhang

BDCN: Semantic Embedding Self-explanatory Breast Diagnostic Capsules Network
Dehua Chen, Keting Zhong and Jianrong He

GCN with External Knowledge for Clinical Event Detection
Dan Liu, Zhichang Zhang, Hui Peng and Ruirui Han

A Prompt-independent and Interpretable Automated Essay Scoring Method for Chinese Second Language Writing
Yupei Wang and Renfen Hu

A Robustly Optimized BERT Pre-training Approach with Post-training
Zhuang Liu, Wayne Lin, Ya Shi, Jun Zhao