Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)

Languages of Latin America

Chapter 2: Short Papers

Mexico City, Mexico
16-21 June 2024

ISBN: 978-1-7138-9895-5

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2024) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Revisiting Zero-Shot Abstractive Summarization in the Era of Large Language Models from the Perspective of Position Bias
Anshuman Chhabra, Hadi Askari and Prasant Mohapatra 1

Struc-Bench: Are Large Language Models Good at Generating Complex Structured Tabular Data?
Xiangru Tang, Yiming Zong, Jason Phang, Yilun Zhao, Wangchunshu Zhou, Arman Cohan and Mark Gerstein 12

Improving Toponym Resolution by Predicting Attributes to Constrain Geographical Ontology Entries
Zeyu Zhang, Egoitz Laparra and Steven Bethard 35

Advancing Regular Language Reasoning in Linear Recurrent Neural Networks
Ting-Han Fan, Ta-Chung Chi and Alexander Rudnicky 45

Extracting Lexical Features from Dialects via Interpretable Dialect Classifiers
Roy Xie, Orevaoghene Ahia, Yulia Tsvetkov and Antonios Anastasopoulos 54

Clear Up Confusion: Advancing Cross-Domain Few-Shot Relation Extraction through Relation-Aware Prompt Learning
Ge Bai, Chenji Lu, Daichi Guo, Shilong Li, Ying Liu, Zhang Zhang, Guanting Dong, Ruifang Liu and Sun Yong 70

Fusion Makes Perfection: An Efficient Multi-Grained Matching Approach for Zero-Shot Relation Extraction
Shilong Li, Ge Bai, Zhang Zhang, Ying Liu, Chenji Lu, Daichi Guo, Ruifang Liu and Sun Yong 79

Personalized Review Recommendation based on Implicit dimension mining
Bei Xu and Yifan xu 86

Unlocking Structure Measuring: Introducing PDD, an Automatic Metric for Positional Discourse Coherence
Yinhong Liu, Yixuan Su, Ehsan Shareghi and Nigel Collier 92

Returning to the Start: Generating Narratives with Related Endpoints
Anneliese Brei, Chao Zhao and Snigdha Chaturvedi 101

Unified Examination of Entity Linking in Absence of Candidate Sets
Nicolas Ong, Hassan Shavarani and Anoop Sarkar 113

MultiParaDetox: Extending Text Detoxification with Parallel Data to New Languages
Daryna Dementieva, Nikolay Babakov and Alexander Panchenko 124

SKICSE: Sentence Knowable Information Prompted by LLMs Improves Contrastive Sentence Embeddings
Fangwei Ou and Jinan Xu 141

A Multi-Aspect Framework for Counter Narrative Evaluation using Large Language Models
Jaylen Jones, Lingbo Mo, Eric Fosler-Lussier and Huan Sun 147

How does Multi-Task Training Affect Transformer In-Context Capabilities? Investigations with Function Classes
Harmon Bhasin, Timothy Ossowski, Yiqiao Zhong and Junjie Hu 169

CELI: Simple yet Effective Approach to Enhance Out-of-Domain Generalization of Cross-Encoders.
Crystina Zhang, Minghan Li and Jimmy Lin 188

ContrastiveMix: Overcoming Code-Mixing Dilemma in Cross-Lingual Transfer for Information Retrieval
Junggeun Do, Jaeseong Lee and seung-won hwang 197

SLIDE: Reference-free Evaluation for Machine Translation using a Sliding Document Window
Vikas Raunak, Tom Kocmi and Matt Post 205

Separately Parameterizing Singleton Detection Improves End-to-end Neural Coreference Resolution
Xiyuan Zou, Yiran Li, Ian Porada and Jackie CK Cheung 212

Unveiling Divergent Inductive Biases of LLMs on Temporal Data
Sindhu Kishore and Hangfeng He 220

On Retrieval Augmentation and the Limitations of Language Model Training
Ting-Rui Chiang, Xinyan Velocity Yu, Joshua Robinson, Ollie Liu, Isabelle Lee and Dani Yogatama 229

GenDecider: Integrating "None of the Candidates" Judgments in Zero-Shot Entity Linking Re-ranking
Kang Zhou, Yuepei Li, Qing Wang, Qiao Qiao and Qi Li 239

Advancing the Robustness of Large Language Models through Self-Denoised Smoothing
Jiabao Ji, Bairu Hou, Zhen Zhang, Guanhua Zhang, Wenqi Fan, Qing Li, Yang Zhang, Gaowen Liu, Sijia Liu and Shiyu Chang 246

Can LLM's Generate Human-Like Wayfinding Instructions? Towards Platform-Agnostic Embodied Instruction Synthesis
Vishnu Sashank Dorbala, Sanjoy Chowdhury and Dinesh Manocha 258

On the Role of Summary Content Units in Text Summarization Evaluation
Marcel Nawrath, Agnieszka Wiktoria Nowak, Tristan Ratz, Danilo Constantin Walenta, Juri Opitz, Leonardo F. R. Ribeiro, João Sedoc, Daniel Deutsch, Simon Mille, Yixin Liu, Sebastian Gehrmann, Lining Zhang, Saad Mahamood, Miruna Clinciu, Khyathi Chandu and Yufang Hou 272

More room for language: Investigating the effect of retrieval on language models
David Samuel, Lucas Georges Gabriel Charpentier and Sondre Wold 282

Discourse-Aware In-Context Learning for Temporal Expression Normalization
Akash Kumar Gautam, Lukas Lange and Jannik Strötgen 306

Contextualizing Argument Quality Assessment with Relevant Knowledge
Darshan Girish Deshpande, Zhivar Sourati, Filip Ilievski and Fred Morstatter 316

Selective Perception: Learning Concise State Descriptions for Language Model Actors
Kolby Nottingham, Yasaman Razeghi, Kyungmin Kim, JB Lanier, Pierre Baldi, Roy Fox and Sameer Singh 327

ALOHa: A New Measure for Hallucination in Captioning Models
Suzanne Petryk, David Chan, Anish Kachinthaya, Haodi Zou, John Canny, Joseph E. Gonzalez and Trevor Darrell 342

Beyond Yes and No: Improving Zero-Shot LLM Rankers via Scoring Fine-Grained Relevance Labels
Honglei Zhuang, Zhen Qin, Kai Hui, Junru Wu, Le Yan, Xuanhui Wang and Michael Bendersky
358

LLM-Driven Knowledge Injection Advances Zero-Shot and Cross-Target Stance Detection
Zhao Zhang, Yiming Li, Jin Zhang and Hui Xu 371

Leveraging Prototypical Representations for Mitigating Social Bias without Demographic Information
Shadi Iskander, Kira Radinsky and Yonatan Belinkov 379

Direct Preference Optimization for Neural Machine Translation with Minimum Bayes Risk Decoding
Guangyu Yang, Jinghong Chen, Weizhe Lin and Bill Byrne 391

EchoPrompt: Instructing the Model to Rephrase Queries for Improved In-context Learning
Raja Sekhar Reddy Mekala, Yasaman Razeghi and Sameer Singh 399

LEAF: Language Learners' English Essays and Feedback Corpus
Shabnam Behzad, Omid Kashefi and Swapna Somasundaran 433

Zero-Shot vs. Translation-Based Cross-Lingual Transfer: The Case of Lexical Gaps
Abteen Ebrahimi and Katharina von der Wense 443

On the True Distribution Approximation of Minimum Bayes-Risk Decoding
Atsumoto Ohashi, Ukyo Honda, Tetsuro Morimura and Yuu Jinnai 459

Rehearsal-Free Modular and Compositional Continual Learning for Language Models
Mingyang Wang, Heike Adel, Lukas Lange, Jannik Strötgen and Hinrich Schuetze 469

Llama meets EU: Investigating the European political spectrum through the lens of LLMs
Ilias Chalkidis and Stephanie Brandl 481

M3T: A New Benchmark Dataset for Multi-Modal Document-Level Machine Translation
Benjamin Hsu, Xiaoyu Liu, Huayang Li, Yoshinari Fujinuma, Maria Nadejde, Xing Niu, Ron Litman, Yair Kittenplon and Raghavendra Pappagari 499

Control-DAG: Constrained Decoding for Non-Autoregressive Directed Acyclic T5 using Weighted Finite State Automata
Jinghong Chen, Weizhe Lin, Jingbiao Mei and Bill Byrne 508

Do Vision-Language Models Understand Compound Nouns?
Sonal Kumar, Sreyan Ghosh, S Sakshi, Utkarsh Tyagi and Dinesh Manocha 519

Is Prompt Transfer Always Effective? An Empirical Study of Prompt Transfer for Question Answering
Minji Jung, Soyeon Park, Jeewoo Sul and Yong Suk Choi 528

Lost in Space: Probing Fine-grained Spatial Understanding in Vision and Language Resamplers
Georgios Pantazopoulos, Alessandro Suglia, Oliver Lemon and Arash Eshghi 540

Do Multilingual Language Models Think Better in English?
Julen Etxaniz, Gorka Azkune, Aitor Soroa, Oier Lopez de Lacalle and Mikel Artetxe 550

A Continued Pretrained LLM Approach for Automatic Medical Note Generation
Dong Yuan, Eti Rastogi, Gautam Naik, Sree Prasanna Rajagopal, Sagar Goyal, Fen Zhao, Bharath Chintagunta and Jeffrey M. Ward 565

Lost in Translation? Translation Errors and Challenges for Fair Assessment of Text-to-Image Models on Multilingual Concepts
Michael Saxon, Yiran Lawrence Luo, Sharon Levy, Chitta Baral, Yezhou Yang and William Yang Wang 572

Self-Improving for Zero-Shot Named Entity Recognition with Large Language Models
Tingyu Xie, Qi Li, Yan Zhang, Zuozhu Liu and Hongwei Wang 583

Lifelong Event Detection with Embedding Space Separation and Compaction
Chengwei Qin, Ruirui Chen, Ruochen Zhao, Wenhan Xia and Shafiq Joty 594

Language Models (Mostly) Do Not Consider Emotion Triggers When Predicting Emotion
Smriti Singh, Cornelia Caragea and Junyi Jessy Li 603

CPopQA: Ranking Cultural Concept Popularity by LLMs
Ming Jiang and Mansi Joshi 615

The Impact of Language on Arithmetic Proficiency: A Multilingual Investigation with Cross-Agent Checking Computation
Chung-Chi Chen, Hiroya Takamura, Ichiro Kobayashi and Yusuke Miyao 631

Efficient Information Extraction in Few-Shot Relation Classification through Contrastive Representation Learning
Philipp Borchert, Jochen De Weerdt and Marie-Francine Moens 638

A diverse Multilingual News Headlines Dataset from around the World
Felix Leeb and Bernhard Schölkopf 647

The Unreasonable Effectiveness of Random Target Embeddings for Continuous-Output Neural Machine Translation
Evgeniia Tokarchuk and Vlad Niculae 653

Efficient Sample-Specific Encoder Perturbations
Yassir Fathullah and Mark Gales 663

Diverse Perspectives, Divergent Models: Cross-Cultural Evaluation of Depression Detection on Twitter
Nuredin Ali Abdelkadir, Charles Chuankai Zhang, Ned Mayo and Stevie Chancellor 672

Removing RLHF Protections in GPT-4 via Fine-Tuning
Qiusi Zhan, Richard Fang, Rohan Bindu, Akul Gupta, Tatsunori Hashimoto and Daniel Kang 681

LifeTox: Unveiling Implicit Toxicity in Life Advice
Minbeom Kim, Jahyun Koo, Hwanhee Lee, Joonsuk Park, Hwaran Lee and Kyomin Jung . . . 688

Arithmetic Reasoning with LLM: Prolog Generation & Permutation
Xiaocheng Yang, Bingsen Chen and Yik-Cheung Tam 699

Verifying Claims About Metaphors with Large-Scale Automatic Metaphor Identification
Kotaro Aono, Ryohei Sasano and Koichi Takeda 711

InstructABSA: Instruction Learning for Aspect Based Sentiment Analysis
Kevin Scaria, Himanshu Gupta, Siddharth Goyal, Saurabh Arjun Sawant, Swaroop Mishra and Chitta Baral 720

MEMORY-VQ: Compression for Tractable Internet-Scale Memory
Yury Zemlyanskiy, Michiel de Jong, Luke Vilnis, Santiago Ontanon, William W. Cohen, Sumit Sanghai and Joshua Ainslie 737

Unveiling the Magic: Investigating Attention Distillation in Retrieval-Augmented Generation
Zizhong Li, Haopeng Zhang and Jiawei Zhang 745

Improving Factuality in Clinical Abstractive Multi-Document Summarization by Guided Continued Pretraining
Ahmed Salem Elhady, Khaled Mostafa Elsayed, Eneko Agirre and Mikel Artetxe 755

MuLan: A Study of Fact Mutability in Language Models
Constanza Fierro, Nicolas Garneau, Emanuele Bugliarello, Yova Kementchedjhieva and Anders Søgaard 762

Language-Independent Representations Improve Zero-Shot Summarization
Vladimir Solovyev, Danni Liu and Jan Niehues 772

Trusting Your Evidence: Hallucinate Less with Context-aware Decoding
Weijia Shi, Xiaochuang Han, Mike Lewis, Yulia Tsvetkov, Luke Zettlemoyer and Wen-tau Yih 783

GuyLingo: The Republic of Guyana Creole Corpora
Christopher Clarke, Roland Daynauth, Jason Mars, Charlene Wilkinson and Hubert Devonish792

DoubleLingo: Causal Estimation with Large Language Models
Marko Veljanovski and Zach Wood-Doughty 799

Improved Text Emotion Prediction Using Combined Valence and Arousal Ordinal Classification
Michail Mitsios, Georgios Vamvoukakis, Georgia Maniati, Nikolaos Ellinas, Georgios Dimitriou, Konstantinos Markopoulos, Panos Kakoulidis, Alexandra Vioni, Myrsini Christidou, Junkwang Oh, Gunu Jho, Inchul Hwang, Georgios Vardaxoglou, Aimilios Chalamandaris, Pirros Tsiakoulis and Spyros Raptis 808

On Narrative Question Answering Skills
Emil Kalbaliyev and Kairit Sirts 814

Order-Based Pre-training Strategies for Procedural Text Understanding
Abhilash Nandy, Yash Kulkarni, Pawan Goyal and Niloy Ganguly 821

Breaking the Language Barrier: Can Direct Inference Outperform Pre-Translation in Multilingual LLM Applications?
Yotam Intrator, Matan Halfon, Roman Goldenberg, Reut Tsarfaty, Matan Eyal, Ehud Rivlin, Yossi Matias and Natalia Aizenberg 829