

40th AAAI Conference on Artificial Intelligence (AAAI-26), 38th Conference on Innovative Applications of Artificial Intelligence (IAAI-26), and 16th Symposium on Educational Advances in Artificial Intelligence (EAAI-26)

Volume 40: AAAI Technical Tracks

- Natural Language Processing V

Singapore
20-27 January 2026

ISBN: 979-8-3313-3414-7

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© (2026) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact Association for the Advancement of Artificial Intelligence
at the address below.

Association for the Advancement of Artificial Intelligence
2275 East Bayshore Road
Suite 160
Palo Alto, California 94303
USA

Phone: 1-650-328-3123
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

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

AAAI TECHNICAL TRACK ON NATURAL LANGUAGE PROCESSING V

Collaborative Enhancement of Large and Small Models for Question Answering Via Dual Knowledge Transfer	33630
<i>Shaofei Wang, Yunan Liu, Xiaolan Tang, Wenlong Chen</i>	
Scaling and Transferability of Annealing Strategies in Large Language Model Training	33639
<i>Siqi Wang, Zhengyu Chen, Teng Xiao, Zheqi Lv, Jinluan Yang, Xunliang Cai, Jingang Wang, Xiaomeng Li</i>	
Ghost in the Transformer: Detecting Model Reuse with Invariant Spectral Signatures	33648
<i>Suqing Wang, Ziyang Ma, Li Xinyi, Zuchao Li</i>	
OmniBench: A Comprehensive Benchmark Integrating Real-World, Time-Sensitive, and Multi-Hop Questions with a Multi-Dimensional Hybrid Evaluation Framework	33657
<i>Wenjie Wang, Yufeng Jiang, Ge Sun, Chenghang Dong, Zheng Jun, Li Mengjie, Lixin Chen, Huan Wang, Haoyu Wang, Bin Chen</i>	
A Rolling Stone Gathers No Moss: Adaptive Policy Optimization for Stable Self-Evaluation in Large Multimodal Models	33666
<i>Wenkai Wang, Hongcan Guo, Zheqi Lv, Shengyu Zhang</i>	
Hallucination as a Computational Boundary: A Hierarchy of Inevitability and the Oracle Escape.....	33675
<i>Xi Wang, Quan Shi, Zenghui Ding, Jianqing Gao, Xianjun Yang</i>	
DPRM: A Dual Implicit Process Reward Model in Multi-Hop Question Answering	33683
<i>Xinyi Wang, Yiping Song, Zhiliang Tian, Bo Liu, Tingjin Luo, Minlie Huang</i>	
Efficient Post-Training Refinement of Latent Reasoning in Large Language Models.....	33692
<i>Xinyuan Wang, Dongjie Wang, Wangyang Ying, Haoyue Bai, Nanxu Gong, Sixun Dong, Kunpeng Liu, Yanjie Fu</i>	
REFO: Reinforced Evolutionary Faithfulness Optimization for Large Language Models.....	33701
<i>Yi Wang, Xiaqiang Tang, Keyu Hu, Haojie Lu, Sihong Xie</i>	
OptScale: Probabilistic Optimality for Inference-Time Scaling.....	33710
<i>Youkang Wang, Jian Wang, Rubing Chen, Xiao-Yong Wei</i>	
Beyond N-Grams: A Hierarchical Reward Learning Framework for Clinically-Aware Medical Report Generation	33719
<i>Yuan Wang, Shujian Gao, Jiaxiang Liu, Songtao Jiang, Xia Haoxiang, Xiaotian Zhang, Zhaolu Kang, Yemin Wang, Zuozhu Liu</i>	
DualSpeechLM: Towards Unified Speech Understanding and Generation Via Dual Speech Token Modeling with Large Language Models.....	33728
<i>Yuan Yuan Wang, Dongchao Yang, Yiwen Shao, Hangting Chen, Jiankun Zhao, Zhiyong Wu, Helen Meng, Xixin Wu</i>	
BEE-RAG: Balanced Entropy Engineering for Retrieval-Augmented Generation	33737
<i>Yuhao Wang, Ruiyang Ren, Yucheng Wang, Jing Liu, Xin Zhao, Hua Wu, Haifeng Wang</i>	

Multi-Level Style Preference Optimization: An Adaptive Detection Framework for Human-Machine Hybrid Text.....	33746
<i>Zehao Wang, Lianwei Wu, Wenbo An, Hang Zhang, Yaxiong Wang</i>	
CP-Search: A Chain Progressive Search Training Framework Incentivizing the Cognitive Behaviors for Searching in LLMs	33755
<i>Zehua Wang, Shipeng Li, Buzhou Tang</i>	
COIN: Uncertainty-Guarding Selective Question Answering for Foundation Models with Provable Risk Guarantees.....	33764
<i>Zhiyuan Wang, Jinhao Duan, Qingni Wang, Xiaofeng Zhu, Tianlong Chen, Xiaoshuang Shi, Kaidi Xu</i>	
MetaEval: Measuring the Discrimination of Benchmarks for Efficient LLM Evaluation.....	33773
<i>Zhuo Wang, Wen Wu, Guoqing Wang, Guangze Ye, Zhenxiao Cheng</i>	
Eliciting Chain-Of-Thought in Base LLMs Via Gradient-Based Representation Optimization.....	33782
<i>Zijian Wang, Yanxiang Ma, Chang Xu</i>	
Knowledge Boundary Discovery for Large Language Models	33791
<i>Ziquan Wang, Zhongqi Lu</i>	
SheetBrain: A Neuro-Symbolic Agent for Accurate Reasoning Over Complex and Large Spreadsheets	33800
<i>Ziwei Wang, Jiayuan Su, Mengyu Zhou, Huaxing Zeng, Mengni Jia, Xiao Lv, Haoyu Dong, Xiaojun Ma, Shi Han, Dongmei Zhang</i>	
AutoLink: Autonomous Schema Exploration and Expansion for Scalable Schema Linking in Text-To-SQL at Scale.....	33809
<i>Ziyang Wang, Yuanlei Zheng, Zhenbiao Cao, Xiaojin Zhang, Zhongyu Wei, Pei Fu, Zhenbo Luo, Wei Chen, Xiang Bai</i>	
MIRAGE: Scaling Test-Time Inference with Parallel Graph-Retrieval-Augmented Reasoning Chains.....	33818
<i>Kaiwen Wei, Rui Shan, Dongsheng Zou, Jianzhong Yang, Bi Zhao, Junnan Zhu, Jiang Zhong</i>	
MovieGraph-ToM: Evaluating Long-Range Theory of Mind in Large Language Models Via Implicit Social-Causal Graphs.....	33827
<i>Tingjiang Wei, Qin Ni, Rong Gao, Yingying Wang, Liang He</i>	
Thinking Forward and Backward: Multi-Objective Reinforcement Learning for Retrieval-Augmented Reasoning.....	33836
<i>Wenda Wei, Yu-An Liu, Ruqing Zhang, Jiafeng Guo, Lixin Su, Shuaiqiang Wang, Dawei Yin, Maarten De Rijke, Xueqi Cheng</i>	
Beyond ReAct: A Planner-Centric Framework for Complex Tool-Augmented LLM Reasoning	33845
<i>Xiaolong Wei, Yuehu Dong, Xingliang Wang, Xingyu Zhang, Zhejun Zhao, Dongdong Shen, Long Xia, Dawei Yin</i>	
Mixture-Of-Trees: Learning to Select and Weigh Reasoning Paths for Efficient LLM Inference	33854
<i>Yangbo Wei, Zhen Huang, Shaoqiang Lu, Junhong Qian, Dongge Qin, Ting Jung Lin, Wei W. Xing, Chen Wu, Lei He</i>	
AdaMCoT: Rethinking Cross-Lingual Factual Reasoning Through Adaptive Multilingual Chain-Of-Thought.....	33863
<i>Zheng Weihua, Xin Huang, Zhengyuan Liu, Tarun Kumar Vangani, Bowei Zou, Xiyao Tao, Yuhao Wu, Aiti Aw, Nancy F. Chen, Roy Ka-Wei Lee</i>	

MultiMedBench: A Scenario-Aware Benchmark for Evaluating Knowledge Editing in Medical VQA	33872
<i>Shengtao Wen, Haodong Chen, Yadong Wang, Zhongying Pan, Xiang Chen, Yu Tian, Bo Qian, Dong Liang, Sheng-Jun Huang</i>	
Reinforcement Learning Enhanced Muti-Hop Reasoning for Temporal Knowledge Question Answering	33881
<i>Wuzhenghong Wen, Chao Xue, Su Pan, Yrwei Sun, Minlong Peng</i>	
Vision-Language Reasoning for Geolocalization: A Reinforcement Learning Approach	33890
<i>Biao Wu, Meng Fang, Ling Chen, Ke Xu, Tao Cheng, Jun Wang</i>	
Introducing Visual Scenes and Reasoning: A More Realistic Benchmark for Spoken Language Understanding	33899
<i>Di Wu, Liting Jiang, Ruiyu Fang, Bianjing , Hongyan Xie, Haoxiang Su, Hao Huang, Zhongjiang He, Shuangyong Song, Xuelong Li</i>	
ReCode: Updating Code API Knowledge with Reinforcement Learning	33908
<i>Haoze Wu, Yunzhi Yao, Wenhao Yu, Ningyu Zhang</i>	
LAMDAS: LLM as an Implicit Classifier for Domain-Specific Data Selection.....	33917
<i>Jian Wu, Hang Yu, Bingchang Liu, Yang Wenjie, Peng Di, Jianguo Li, Yue Zhang</i>	
AStar: Boosting Multimodal Reasoning with Automated Structured Thinking	33926
<i>Jinyang Wu, Mingkuan Feng, Guocheng Zhai, Shuai Zhang, Zheng Lian, Fangrui Lv, Pengpeng Shao, Ruihan Jin, Zhengqi Wen, Jianhua Tao</i>	
Cognitive Enhancement Chain-Of-Thought Towards Enhancing Style Learning and Content Preservation for Long Style Transfer.....	33935
<i>Lianwei Wu, Botao Wang, Wenbo An, Tiejiao Li, Xianghua Li</i>	
Reasoning Or Memorization? Unreliable Results of Reinforcement Learning Due to Data Contamination	33944
<i>Mingqi Wu, Zhihao Zhang, Qiaole Dong, Zhiheng Xi, Jun Zhao, Senjie Jin, Xiaoran Fan, Yuhao Zhou, Huijie Lv, Ming Zhang, Yanwei Fu, Qin Liu, Songyang Zhang, Qi Zhang</i>	
A Disease-Aware Dual-Stage Framework for Chest X-Ray Report Generation.....	33953
<i>Puzhen Wu, Hexin Dong, Yi Lin, Yihao Ding, Yifan Peng</i>	
Audio-Thinker: Guiding Large Audio Language Model When and How to Think Via Reinforcement Learning	33962
<i>Shu Wu, Chenxing Li, Wenfu Wang, Hao Zhang, Hualei Wang, Meng Yu, Dong Yu</i>	
Finding the Translation Switch: Discovering and Exploiting the Task-Initiation Features in LLMs	33971
<i>Xinwei Wu, Heng Liu, Xiaohu Zhao, Yuqi Ren, Linlong Xu, Longyue Wang, Deyi Xiong, Weihua Luo, Kaifu Zhang</i>	
StepFun-Formalizer: Unlocking the Autoformalization Potential of LLMs Through Knowledge-Reasoning Fusion	33980
<i>Yutong Wu, Di Huang, Ruosi Wan, Yue Peng, Shijie Shang, Chenrui Cao, Lei Qi, Rui Zhang, Xishan Zhang, Zidong Du, Jie Yan, Xing Hu</i>	
GEM: Gaussian Embedding Modeling for Out-Of-Distribution Detection in GUI Agents	33989
<i>Zheng Wu, Pengzhou Cheng, Zongru Wu, Lingzhong Dong, Zhuosheng Zhang</i>	

GlitchMiner: Mining Glitch Tokens in Large Language Models Via Gradient-Based Discrete Optimization.....	33998
<i>Zihui Wu, Haichang Gao, Ping Wang, Shudong Zhang, Zhaoxiang Liu, Shiguo Lian</i>	
MetaAct-RL: Training Language Models for Reasoning Through Meta-Action-Based Reinforcement Learning.....	34006
<i>Zhiheng Xi, Yuhui Wang, Yiwen Ding, Guanyu Li, Senjie Jin, Shichun Liu, Jixuan Huang, Dingwen Yang, Jiafu Tang, Boyang Hong, Junjie Ye, Shihan Dou, Ming Zhang, Jian Guan, Wei Wu, Rui Zheng, Tao Gui, Qi Zhang, Xuanjing Huang</i>	
KALL-E: Autoregressive Speech Synthesis with Next-Distribution Prediction	34016
<i>Kangxiang Xia, Xinfu Zhu, Jixun Yao, Wenjie Tian, Wenhao Li, Lei Xie</i>	
SDA: Steering-Driven Distribution Alignment for Open LLMs Without Fine-Tuning.....	34025
<i>Wei Xia, Zhi-Hong Deng</i>	
SCALE: Selective Resource Allocation for Overcoming Performance Bottlenecks in Mathematical Test-Time Scaling.....	34034
<i>Yang Xiao, Chunpu Xu, Ruifeng Yuan, Jessie Wang, Wenjie Li, Pengfei Liu</i>	
Enhancing Uncertainty Estimation in LLMs with Expectation of Aggregated Internal Belief	34043
<i>Zeguan Xiao, Diyang Dou, Boya Xiong, Yun Chen, Guanhua Chen</i>	
DeepOR: A Deep Reasoning Foundation Model for Optimization Modeling.....	34052
<i>Ziyang Xiao, Yuan Jessica Wang, Xiongwei Han, Shisi Guan, Jingyan Zhu, Jingrong Xie, Lilin Xu, Han Wu, Wing Yin Yu, Zehua Liu, Xiaojin Fu, Gang Chen, Dongxiang Zhang</i>	
CroPS: Improving Dense Retrieval with Cross-Perspective Positive Samples in Short-Video Search	34061
<i>Ao Xie, Jiahui Chen, Quanzhi Zhu, Xiaoze Jiang, Zhiheng Qin, Enyun Yu, Han Li</i>	
LLM-Oriented Token-Adaptive Knowledge Distillation	34070
<i>Xurong Xie, Zhucun Xue, Jiafu Wu, Jian Li, Yabiao Wang, Xiaobin Hu, Yong Liu, Jiangning Zhang</i>	
Advanced Black-Box Tuning of Large Language Models with Limited API Calls	34079
<i>Zhikang Xie, Weilin Wan, Peizhu Gong, Weizhong Zhang, Cheng Jin</i>	
OPLoRA: Orthogonal Projection LoRA Prevents Catastrophic Forgetting During Parameter-Efficient Fine-Tuning	34088
<i>Yifeng Xiong, Xiaohui Xie</i>	
Editing is a Bargaining Game: Balanced Knowledge Editing in Large Language Models	34097
<i>Chenghao Xu, Jiexi Yan, Muli Yang, Fen Fang, Huilin Chen, Cheng Deng</i>	
MedMKEB: A Comprehensive Knowledge Editing Benchmark for Medical Multimodal Large Language Models	34106
<i>Dexuan Xu, Jiayi Wang, Zhongyan Chai, Yongzhi Cao, Hanpin Wang, Huamin Zhang, Yu Huang</i>	
ComLQ: Benchmarking Complex Logical Queries in Information Retrieval.....	34115
<i>Ganlin Xu, Zhitao Yin, Linghao Zhang, Jiaqing Liang, Weijia Lu, Xiaodong Zhang, Zhifei Yang, Sihang Jiang, Deqing Yang</i>	
Multiplicative Orthogonal Sequential Editing for Language Models.....	34124
<i>Hao-Xiang Xu, Jun-Yu Ma, Ziqi Peng, Yuhao Sun, Zhen-Hua Ling, Jia-Chen Gu</i>	

Multi-Value Alignment for LLMs Via Value Decorrelation and Extrapolation.....	34133
<i>Hefei Xu, Le Wu, Chen Cheng, Hao Liu</i>	
Thinker: Training LLMs in Hierarchical Thinking for Deep Search Via Multi-Turn Interaction.....	34142
<i>Jun Xu, Xinkai Du, Yu Ao, Peilong Zhao, Yang Li, Ling Zhong, Lin Yuan, Zhongpu Bo, Xiaorui Wang, Mengshu Sun, Zhengke Gui, Dalong Zhang, Zhaoyang Wang, Wang Qiwei, Yangyang Hou, Zhiying Yin, Haofen Wang, Huajun Chen, Lei Liang, Jun Zhou</i>	
RLKD: Distilling LLMs' Reasoning Via Reinforcement Learning.....	34151
<i>Shicheng Xu, Liang Pang, Yunchang Zhu, Jia Gu, Zihao Wei, Jingcheng Deng, Feiyang Pan, Huawei Shen, Xueqi Cheng</i>	
DeepPhy: Benchmarking Agentic VLMs on Physical Reasoning.....	34160
<i>Xinrun Xu, Pi Bu, Ye Wang, Börje F. Karlsson, Ziming Wang, Tengtao Song, Qi Zhu, Jun Song, Zhiming Ding, Bo Zheng</i>	
Causal-HalBench: Uncovering LVLMS Object Hallucinations Through Causal Intervention.....	34169
<i>Zhe Xu, Zhicai Wang, Junkang Wu, Jinda Lu, Xiang Wang</i>	
ST-LLM: Spatial Transcriptomics Embedding with Large Language Models.....	34178
<i>Zetao Xu, Xiaohua Wan, Le Li, Shuang Feng, Yiming Zhang, Fa Zhang, Bin Hu</i>	
Benchmarking and Enhancing Rule Knowledge-Driven Reasoning of Large Language Models.....	34187
<i>Zijie Xu, Wenjun Ke, Peng Wang, Guozheng Li, Qingjian Ni, Jiajun Liu, Ziyu Shang, Jing Zhou</i>	
MMBERT: Scaled Mixture-Of-Experts Multimodal BERT for Robust Chinese Hate Speech Detection Under Cloaking Perturbations.....	34196
<i>Qiyao Xue, Yuchen Dou, Zheyuan Ryan Shi, Xiang Lorraine Li, Wei Gao</i>	
AgriEval: A Comprehensive Chinese Agricultural Benchmark for Large Language Models.....	34205
<i>Lian Yan, Haotian Wang, Chen Tang, Haifeng Liu, Tianyang Sun, Liangliang Liu, Yi Guan, Jingchi Jiang</i>	
From Imitation to Discrimination: Toward a Generalized Curriculum Advantage Mechanism Enhancing Cross-Domain Reasoning Tasks.....	34214
<i>Changpeng Yang, Jinyang Wu, Yuchen Liu, Shuai Zhang, Yang Li, Qiliang Liang, Hongzhen Wang, Shuai Nie, Jiaming Xu, Runyu Shi, Ying Huang, Guoquan Zhang</i>	
Test-Time Prompt Intervention.....	34223
<i>Chenxu Yang, Qingyi Si, Mz Dai, Dingyu Yao, Mingyu Zheng, Minghui Chen, Zheng Lin, Weiping Wang</i>	
Breaking the Trade-Off Between Faithfulness and Expressiveness for Large Language Models.....	34232
<i>Chenxu Yang, Qingyi Si, Lanrui Wang, Zheng Lin</i>	
PrivSV: Differentially Private Steering Vector for Large Language Models.....	34241
<i>Haocheng Yang, Xiang Cheng, Chenhao Sun, Pengfei Zhang, Sen Su</i>	
HeartLLM: Discretized ECG Tokenization for LLM-Based Diagnostic Reasoning.....	34250
<i>Jinning Yang, Wenjie Sun, Wen Shi</i>	
RCP-Merging: Merging Long Chain-Of-Thought Models with Domain-Specific Models by Considering Reasoning Capability as Prior.....	34259
<i>Junyao Yang, Jianwei Wang, Huiping Zhuang, Cen Chen, Ziqian Zeng</i>	

Decoupling Knowledge and Reasoning in LLMs: An Exploration Using Cognitive Dual-System Theory	34268
<i>Mutian Yang, Jiandong Gao, Ji Wu</i>	
OncoCoT: A Temporal-Causal Chain-Of-Thought Dataset for Oncologic Decision-Making	34277
<i>Peiru Yang, Yudong Li, Shiting Wang, Xinyi Liu, Haotian Gan, Xintian Li, Qingyu Gao, Yongfeng Huang</i>	
ShieldRAG: Safeguarding Retrieval-Augmented Generation from Untrusted Knowledge Bases	34286
<i>Peiru Yang, Haoran Zheng, Yi Luo, Xinyi Liu, Jinrui Wang, Huili Wang, Xintian Li, Yongfeng Huang, Tao Qi</i>	
MrM: Black-Box Membership Inference Attacks Against Multimodal RAG Systems	34295
<i>Peiru Yang, Jinhua Yin, Haoran Zheng, Xueying Bai, Huili Wang, Yufei Sun, Xintian Li, Songwei Pei, Yongfeng Huang, Tao Qi</i>	
HyCoRA: Hyper-Contrastive Role-Adaptive Learning for Role-Playing	34304
<i>Shihao Yang, Zhicong Lu, Yong Yang, Bo Lv, Yang Shen, Nayu Liu</i>	
RegionMarker: A Region-Triggered Semantic Watermarking Framework for Embedding-As-A-Service Copyright Protection.....	34313
<i>Shufan Yang, Zifeng Cheng, Zhiwei Jiang, Yafeng Yin, Cong Wang, Shiping Ge, Yuchen Fu, Qing Gu</i>	
AsFT: Anchoring Safety During LLM Fine-Tuning Within Narrow Safety Basin.....	34322
<i>Shuo Yang, Qihui Zhang, Yuyang Liu, Yue Huang, Xiaojun Jia, Kun-Peng Ning, Jia-Yu Yao, Jigang Wang, Dai Hailiang, Yibing Song, Li Yuan</i>	
Talk2Code: A Multi-Turn Interaction Benchmark with Dual-Track Evaluation for Code Generation.....	34331
<i>Weibin Yang, Liangru Xie, Jieryun Cai, Yuxiang Yan, Hong-Ning Dai, Hao Wang</i>	
EHRStruct: A Comprehensive Benchmark Framework for Evaluating Large Language Models on Structured Electronic Health Record Tasks	34340
<i>Xiao Yang, Xuejiao Zhao, Zhiqi Shen</i>	
RICo: Refined In-Context Contribution for Automatic Instruction-Tuning Data Selection.....	34349
<i>Yixin Yang, Qingxiu Dong, Linli Yao, Fangwei Zhu, Weilin Luo, Bin Wang, Zhifang Sui</i>	
SASST: Leveraging Syntax-Aware Chunking and LLMs for Simultaneous Speech Translation.....	34358
<i>Zeyu Yang, Lai Wei, Roman Koshkin, Xi Chen, Satoshi Nakamura</i>	
Multimodal DeepResearcher: Generating Text-Chart Interleaved Reports from Scratch with Agentic Framework	34368
<i>Zhaorui Yang, Bo Pan, Han Wang, Yiyao Wang, Xingyu Liu, Luoxuan Weng, Yingchaojie Feng, Haozhe Feng, Minfeng Zhu, Bo Zhang, Wei Chen</i>	
ORACLE: Optimizing Reasoning Abilities of Large Language Models Via Constraint-Led Synthetic Data Elicitation.....	34378
<i>Zhuojie Yang, Wentao Wan, Keze Wang</i>	
Conversational Learning Diagnosis Via Reasoning Multi-Turn Interactive Learning	34387
<i>Fangzhou Yao, Sheng Chang, Weibo Gao, Qi Liu</i>	
V-Pruner: A Fast and Globally-Informed Token Pruning Framework for Vision Transformer	34396
<i>Guangzhen Yao, Jiayun Zheng, Zezhou Wang, Wenxin Zhang, Renda Han, Chuangxin Zhao, Zeyu Zhang, Runhao Liu</i>	

Incorporating Self-Rewriting into Large Language Model Reasoning Reinforcement	34405
<i>Jiashu Yao, Heyan Huang, Shuang Zeng, Chuwei Luo, Wangjie You, Jie Tang, Qingsong Liu, Yuhang Guo, Yangyang Kang</i>	
Interpreting FedSpeak with Confidence: A LLM-Based Uncertainty-Aware Framework Guided by Monetary Policy Transmission Paths	34414
<i>Rui Yao, Qi Chai, Jinhai Yao, Siyuan Li, Junhao Chen, Qi Zhang, Hao Wang</i>	
Seeing Through the Conflict: Transparent Knowledge Conflict Handling in Retrieval-Augmented Generation	34423
<i>Hua Ye, Siyuan Chen, Ziqi Zhong, Canran Xiao, Haoliang Zhang, Yuhan Wu, Fei Shen</i>	
Themis: Automated Constraint-Aware Test Synthesis Framework for Code Reinforcement Learning	34432
<i>Shengyu Ye, Qi Liu, Hao Jiang, Zheng Zhang, Heng Yu, Zhenya Huang</i>	
RealWebAssist: A Benchmark for Long-Horizon Web Assistance with Real-World Users	34441
<i>Suyu Ye, Haojun Shi, Darren Shih, Hyokun Yun, Tanya G. Roosta, Tianmin Shu</i>	
LangGPS: Language Separability Guided Data Pre-Selection for Joint Multilingual Instruction Tuning	34450
<i>Yangfan Ye, Xiaocheng Feng, Xiachong Feng, Lei Huang, Weitao Ma, Qichen Hong, Yunfei Lu, Duyu Tang, Dandan Tu, Bing Qin</i>	
K-12EduBench: A Benchmark for Evaluating Large Language Models' Knowledge, Problem-Solving, and Educational Goal Cognition in K-12 Education	34459
<i>Yuqing Ye, Xuan Zhou, Zhifu Chen, Dandan Li, Hengnian Gu, Jin Peng Zhou, Dongdai Zhou</i>	
SpeakerLM: End-To-End Versatile Speaker Diarization and Recognition with Multimodal Large Language Models	34467
<i>Han Yin, Yafeng Chen, Chong Deng, Luyao Cheng, Hui Wang, Chao-Hong Tan, Qian Chen, Wen Wang, Xiangang Li</i>	

Author Index