The Third Workshop of Bridging Neurons and Symbols for Natural Language Processing and Knowledge Graphs Reasoning (NeusymBridge) @AAAI 2026

Stabilizing Reinforcement Learning for Honesty Alignment in Language Models on Deductive Reasoning

Jiarui Liu, Kaustubh Dhole, Yingheng Wang, Haoyang Wen, Sarah Zhang, Haitao Mao, Gaotang Li, Neeraj Varshney, Jingguo Liu and Xiaoman Pan

HAS BEEN AWARDED THE

Best Paper Award

Yunfei Long	Yansong Feng	Xianpei Han	Shizhu He	Tiansi Dong
Organizer Co-chair	Organizer Co-chair	Organizer Co-chair	Organizer Co-chair	Organizer Co-chair

January 26, 2026