The Defeat of the Winograd Schema Challenge (Abstract Reprint)

Vid Kocijan[1], Ernest Davis[2], Thomas Lukasiewicz[3,4], Gary Marcus[5], Leora Morgenstern[6]

[1]Kumo.ai, 357 Castro Street, Suite 200 Mountain View, CA 94041, United States
[2]New York University, Department of Computer Science, 251 Mercer St, NY 10012, United States
[3]Institute of Logic and Computation, Vienna University of Technology, Austria
[4]Department of Computer Science, University of Oxford, UK
[5]New York University, New York, NY 10012, United States
[6]Palo Alto Research Center, part of SRI International, 3333 Coyote Hill Rd, Palo Alto, CA 94304, United States

Abstract Reprint. This is an abstract reprint of a journal article by Kocijan, Davis, Lukasiewicz, Marcus, and Morgenstern (2023).

Abstract

The Winograd Schema Challenge—a set of twin sentences involving pronoun reference disambiguation that seem to require the use of commonsense knowledge—was proposed by Hector Levesque in 2011. By 2019, a number of AI systems, based on large pre-trained transformer-based language models and fine-tuned on these kinds of problems, achieved better than 90% accuracy. In this paper, we review the history of the Winograd Schema Challenge and discuss the lasting contributions of the flurry of research that has taken place on the WSC in the last decade. We discuss the significance of various datasets developed for WSC, and the research community's deeper understanding of the role of surrogate tasks in assessing the intelligence of an AI system.

References

Kocijan, V.; Davis, E.; Lukasiewicz, T.; Marcus, G.; and Morgenstern, L. 2023. The defeat of the Winograd Schema Challenge. *Artificial Intelligence*, 325: 103971.