

{NeurIPS, EMNLP, LoG}

Self-play, retrieval, steering, evaluation,
and relational generalisation.