THINKING DEEPLY WITH RECURRENCE: GENERALIZING FROM EASY TO HARD SEQUENTIAL REASONING PROBLEMS

Avi Schwarzschild [*], **Arjun Gupta** [*], **Micah Goldblum, & Tom Goldstein**
University of Maryland
College Park, MD, USA
{avi1, arjung15, goldblum, tomg}@umd.edu

ABSTRACT

Deep neural networks are powerful machines for visual pattern recognition, but reasoning tasks that are easy for humans may still be difficult for neural models. Humans can extrapolate simple reasoning strategies to solve difficult problems using long sequences of abstract manipulations, i.e., harder problems are solved by thinking for longer. In contrast, the sequential computing budget of feed-forward networks is limited by their depth, and networks trained on simple problems have no way of extending their reasoning capabilities without retraining. In this work, we observe that recurrent networks have the uncanny ability to closely emulate the behavior of non-recurrent deep models, often doing so with far fewer parameters, on both image classification and maze solving tasks. We also explore whether recurrent networks can make the generalization leap from simple problems to hard problems simply by increasing the number of recurrent iterations used as test time. To this end, we show that recurrent networks that are trained to solve simple mazes with few recurrent steps can indeed solve much more complex problems simply by performing additional recurrences during inference.

[*]Equal contribution.