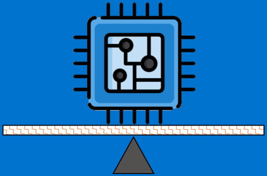
UNIVERSITY OF CAMBRIDGE

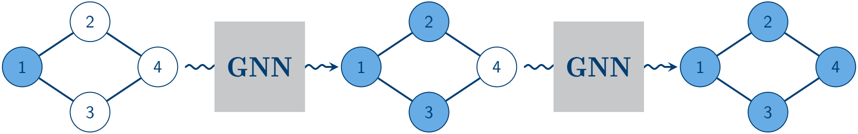


Deep Equilibrium Algorithmic Reasoning

Dobrik Georgiev[1], JJ Wilson, Davide Buffelli[2], Pietro Liò[1]

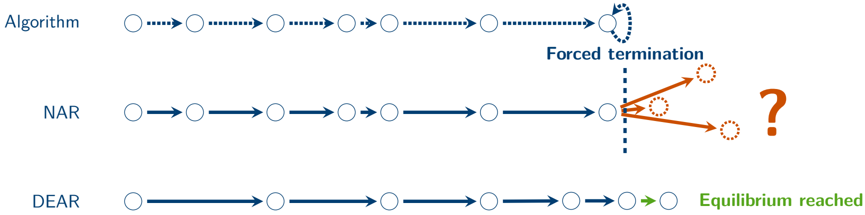
[1]Department of Computer Science and Technology, University of Cambridge, UK
[2]MediaTek Research

A quick primer on neural algorithmic reasoning (NAR)



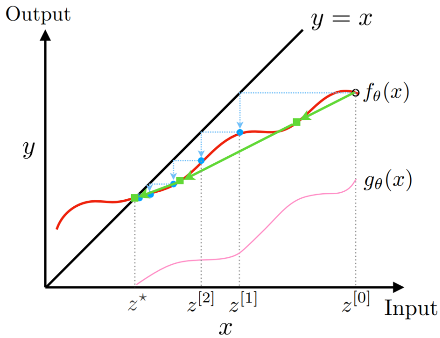
- A Graph Neural Network (GNN) neurally executing BFS
- Robust performance is achieved through alignment

Equilibrium is a missed alignment



- Termination was “forced” (also assumes test-time information)
- Denotations of `for` loops require finding a fixed point

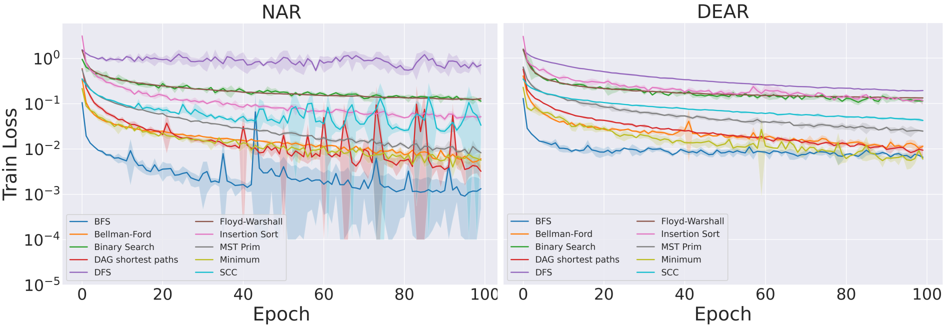
How? DEQs [Bai et al., 2019]



In order to find $\boldsymbol{x}$, s.t.
$f_\theta(\boldsymbol{x}) = \boldsymbol{x}$
find the roots of
$g_\theta(\boldsymbol{x}) = f_\theta(\boldsymbol{x}) - \boldsymbol{x}$

- Extra inductive bias
- Does not require termination signal
- Any choice of root finding

Can you really train them this way?



And are they as accurate?

Algorithm	**NAR**♦	**NAR**♦ (Triplet-MPNN)	**NAR**◊ (LT)	**DEAR** (ours)
Weighted graphs				
Bellman-F.	$97.06\% \pm 0.40$	$97.23\% \pm 0.15$	$95.39\% \pm 1.42$	$96.78\% \pm 0.43$
Floyd-W.	$52.53\% \pm 0.98$	$61.86\% \pm 1.57$	$49.30\% \pm 0.53$	$55.75\% \pm 2.20$
DSP	$94.21\% \pm 1.77$	$93.32\% \pm 1.60$	$88.30\% \pm 1.04$	$89.81\% \pm 0.14$
MST Prim	$93.56\% \pm 0.77$	$92.01\% \pm 1.50$	$87.69\% \pm 1.17$	$88.67\% \pm 0.74$
Unweighted graphs				
BFS	$99.85\% \pm 0.09$	$99.69\% \pm 0.29$	$99.51\% \pm 0.06$	$98.73\% \pm 0.37$
DFS	$16.89\% \pm 5.73$	$31.20\% \pm 4.02$	$29.07\% \pm 2.32$	$40.62\% \pm 0.44$
SCC	$40.70\% \pm 1.39$	$46.84\% \pm 1.70$	$39.33\% \pm 1.52$	$43.63\% \pm 1.19$
Arrays				
Search (Binary)	$94.67\% \pm 2.31$	$93.33\% \pm 2.31$	$84.33\% \pm 8.33$	$59.00\% \pm 12.3$
Minimum	$97.67\% \pm 5.73$	$96.67\% \pm 2.31$	$94.00\% \pm 2.00$	$97.22\% \pm 3.82$
Sort (Ins.)	$27.07\% \pm 10.3$	$63.67\% \pm 39.97$	$33.8\% \pm 12.04$	$86.93\% \pm 3.87$

♦ – termination given at train+test time ◊ – termination given at train time only

And are they as accurate?

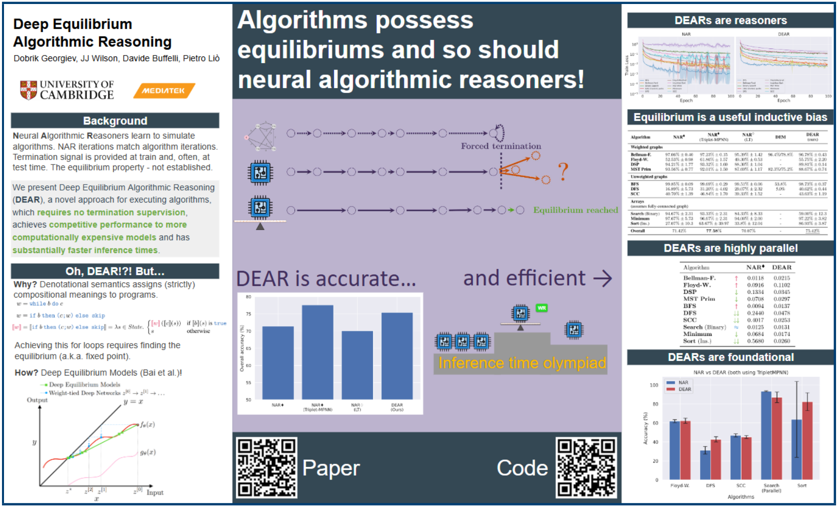
Algorithm	NAR$^\blacklozenge$	NAR$^\blacklozenge$ (Triplet-MPNN)	NAR$^\lozenge$ (LT)	DEAR (ours)
Overall	71.42%	$\mathbf{77.58\%}$	70.07%	$\underline{75.42\%}$

$\blacklozenge$ – termination given at train+test time $\lozenge$ – termination given at train time only

They cannot be fast as well

Algorithm		NAR♦	DEAR
Bellman-F.	$\uparrow$	0.0118	0.0215
Floyd-W.	$\uparrow$	0.0916	0.1102
DSP	$\downarrow$	0.1334	0.0345
MST Prim	$\downarrow$	0.0708	0.0297
BFS	$\uparrow$	0.0094	0.0137
DFS	$\downarrow\downarrow$	0.2440	0.0478
SCC	$\downarrow\downarrow$	0.4017	0.0253
Search (Binary)	$\approx$	0.0125	0.0131
Minimum	$\downarrow$	0.0684	0.0174
Sort (Ins.)	$\downarrow\downarrow$	0.5680	0.0260

See you at the poster session!



Bai, S., Kolter, J. Z., and Koltun, V. (2019). Deep equilibrium models. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., editors, Advances in Neural Information Processing Systems, volume 32. Curran Associates, Inc.