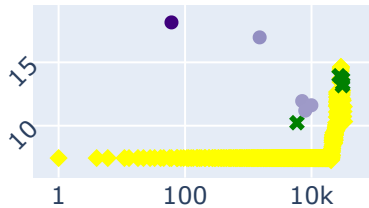
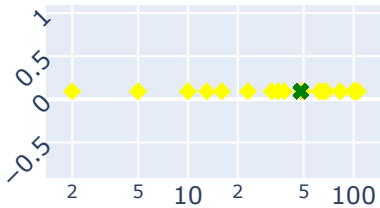


KL(model, ablated)

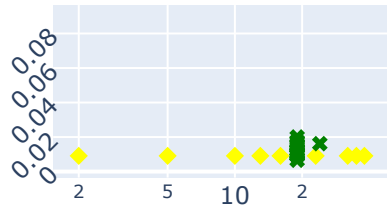
Circuit Recovery (IOI)



Tracr (Reverse)

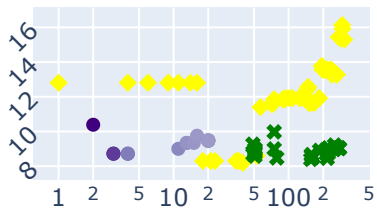


Tracr (Proportion)

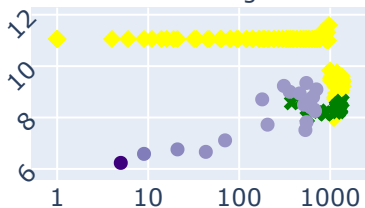


◆ HISP
✕ SP
● ACDC

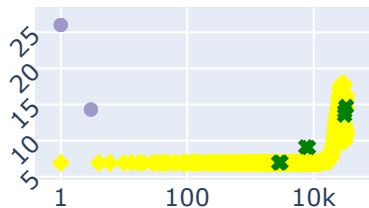
Induction



Docstring



Greater Than



Number of edges