

Dataset	Parameters				Evaluation		
	Batch size	Epoch	Learning rate		Train_loss	Eval_loss	Exact_match
COGS-structural	128	100	5e-5	0	0.02152565842	0.001290984452	0.0006666666667
				1	0.0220092939	0.001323060016	0.0
				2	0.0220092939	0.001323060016	0.0
				3	0.02136987701	0.0008196550771	0.0
				4	0.02153392957	0.001518002711	0.0003333333333
				5	0.02102481274	0.001736165374	0.0
				6	0.02214540267	0.001705732546	0.0
				7	0.02205265573	0.00186376716	0.0
				8	0.02293341639	0.001544175437	0.0
				9	0.02169316459	0.001099817222	0.0
				10	0.02183453109	0.001265278668	0.0
				11	0.021336415	0.00110143458	0.0
				12	0.02161602858	0.001374388579	0.0
				13	0.02108788021	0.001556925476	0.0
				14	0.02237227758	0.001625619712	0.0
				15	0.02235104603	0.001475795638	0.0003333333333
				16	0.02175871141	0.001353516476	0.0
				17	0.02069682311	0.001043781405	0.0003333333333
				18	0.02182525016	0.001097325468	0.0
				19	0.02292406146	0.001269277418	0.0003333333333