Generator	Dimension	Model Property Iteration	Our Model Plausibility	Proximity	Sparsity	D4EL Plausibility	Proximity	Sparsity
CBG-CBGW-IM	Activity	0	0.320000	4.114943	9.000000	0.160000	4.178792	11.000000
		1	0.240000	3.862351	7.840000	0.120000	3.802004	6.420000
		2	0.160000	3.791798	7.680000	0.080000	3.766728	6.340000
		3	0.180000	4.179553	9.640000	0.090000	4.211097	9.320000
		4	0.280000	4.560320	12.260000	0.140000	4.625368	12.630000
		5	0.400000	4.258332	10.720000	0.200000	4.308616	10.360000
	Resource	0	0.000000	4.251903	17.920000	0.000000	4.724028	21.460000
		1	0.000000	3.818159	14.720000	0.000000	4.358569	19.360000
		2	0.000000	3.918192	15.500000	0.000000	4.557172	20.250000
		3	0.000000	4.283800	18.340000	0.000000	4.739976	20.670000
		4	0.000000	4.681512	21.820000	0.000000	4.938832	22.410000
		5	0.000000	4.168343	17.400000	0.000000	4.682248	20.700000
ES-EGW-CBI-ES-UC3-SBM-RR-IM	Activity	0	0.000000	4.123106	6.000000	0.000000	4.182873	9.500000
		1	0.000000	3.872983	2.000000	0.000000	3.807320	3.500000
		2	0.000000	3.741657	4.000000	0.000000	3.741657	4.500000
		3	0.000000	3.875524	3.020000	0.000000	4.059082	6.010000
		4	0.000000	4.582576	8.000000	0.000000	4.636496	10.500000
		5	0.000000	4.242641	7.000000	0.000000	4.300770	8.500000
	Resource	0	0.000000	4.123106	17.000000	0.000000	4.659629	21.000000
		1	0.000000	3.492392	11.200000	0.000000	4.195686	17.600000
		2	0.000000	2.965685	8.800000	0.000000	4.080919	16.900000
		3	0.000000	4.242641	17.640000	0.000000	4.719397	20.320000
		4	0.000000	4.795832	23.000000	0.000000	4.995992	23.000000
		5	0.000000	4.000000	16.000000	0.000000	4.598076	20.000000
ES-EGW-CBI-RWS-OPC-SBM-FSR-IM	Activity	0	0.000000	4.123106	6.000000	0.000000	4.182873	9.500000
		1	0.000000	3.741657	3.000000	0.000000	3.741657	4.000000
		2	0.000000	3.741657	4.000000	0.000000	3.741657	4.500000
		3	0.000000	4.000000	4.000000	0.000000	4.121320	6.500000
		4	0.000000	4.472136	5.000000	0.000000	4.581276	9.000000
		5	0.000000	4.242641	6.000000	0.000000	4.300770	8.000000
	Resource	0	0.000000	4.000000	14.000000	0.000000	4.598076	19.500000
		1	0.000000	4.358899	16.000000	0.000000	4.628939	20.000000
		2	0.000000	2.236068	5.000000	0.000000	3.716110	15.000000
		3	0.000000	4.123106	15.000000	0.000000	4.659629	19.000000
		4	0.000000	4.582576	21.000000	0.000000	4.889364	22.000000
		5	0.000000	3.872983	15.000000	0.000000	4.534568	19.500000
RG-RGW-IM	Activity	0	0.000000	4.832195	20.960000	0.000000	4.537418	16.980000
		1	0.000000	4.822574	20.380000	0.000000	4.282115	12.690000
		2	0.000000	4.856500	20.460000	0.000000	4.299079	12.730000
		3	0.000000	4.805536	20.800000	0.000000	4.524089	14.900000
		4	0.000000	4.794502	21.740000	0.000000	4.742459	17.370000
		5	0.000000	4.731463	20.580000	0.000000	4.545181	15.290000
	Resource	0	0.000000	4.242641	18.000000	0.000000	4.719397	21.500000
		1	0.000000	3.741657	14.000000	0.000000	4.320318	19.000000
		2	0.000000	3.741657	14.000000	0.000000	4.468905	19.500000
		3	0.000000	4.242641	18.000000	0.000000	4.719397	20.500000
		4	0.000000	4.690416	22.000000	0.000000	4.943284	22.500000
		5	0.000000	4.358899	19.000000	0.000000	4.777526	21.500000

Table 1: A comparison between our model and D4EL

Table 1 shows how each model performs under the evaluation metrics chosen by Hsieh, Moreira, and Ouyang. All of them apply seperately on the sequence of resources and the sequence of activities. Each evaluation metric is the mean across all counterfactual results per model.

First, plausibility, which measures whether the sequence of activities or the sequence of resources was found in the data. Next, proximity, which is the euclidian distance between two sequences. Third, sparsity, which is computed by using the Levenshtein distance.

We see that the evolutionary models are often comparable and somethimes even better than the DiCE4EL solution by Hsieh, Moreira, and Ouyang. We see that for instance for proximity. If the proximity of our model is lower than proximity of the DiCE4EL solution, we can say that our models are on average closer to the factual. Similar holds for sparsity. We see this behaviour for both evolutionary generators. However, the Casebased-Search Generator also displays better proximity and sparsity scores than DiCE4EL. Only the Random-Search Generator appears to display worse results.