



Allowing agent interactions align with various simulation objectives

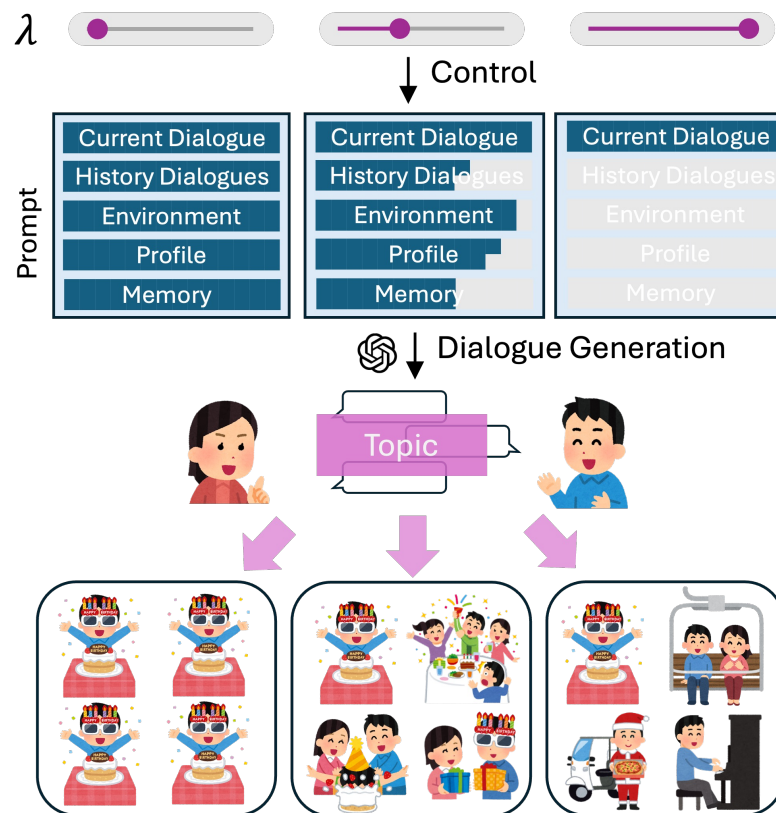
Diversity ↓: stable, stick to main story



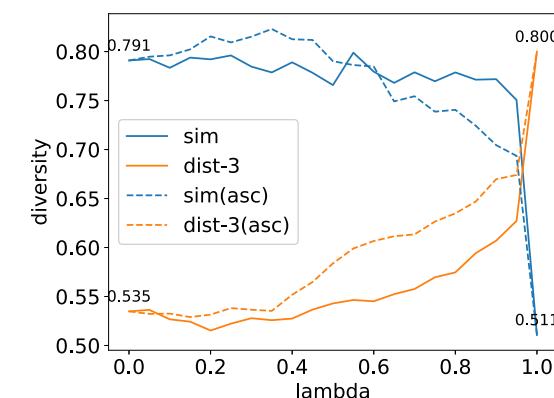
Diversity ↑: varied, engaging experience



Adaptive Prompt Pruning:  
control diversity by  $\lambda$  based on attention score



More effective than  
(and compatible with)  
temperature and top-p



	config	sim (↓)	dist-1	dist-2	dist-3 (↑)	len
Full	default	0.791	0.095	0.350	0.535	39.9
APP	default	<b>0.771</b>	<b>0.107</b>	<b>0.393</b>	<b>0.594</b>	38.4
Full	T=1.0	0.791	0.103	0.381	0.578	40.1
APP	T=1.0	<b>0.778</b>	<b>0.113</b>	<b>0.419</b>	<b>0.634</b>	38.7
Full	p=0.99	0.800	0.102	0.375	0.569	40.0
APP	p=0.99	<b>0.776</b>	<b>0.111</b>	<b>0.414</b>	<b>0.624</b>	38.4
Full	sequential	<b>0.634</b>	0.197	0.524	0.695	21.9
APP	sequential	0.645	<b>0.216</b>	<b>0.563</b>	<b>0.740</b>	21.3