







# Precision\_General\_val

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	_predicate_sim	f_seq	f_seq_sim
0.795	0.763	0.782	0.792	0.801	0.774	0.795	0.699	0.721

# Precision\_General\_test

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	_predicate_sim	f_seq	f_seq_sim
0.791	0.764	0.785	0.794	0.802	0.769	0.791	0.687	0.71

# Precision\_General\_test\_unseen

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	_predicate_sim	f_seq	f_seq_sim
0.799	0.767	0.807	0.786	0.794	0.758	0.799	0.663	0.723

# Recall\_General\_val

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate f	_predicate_sim	f_seq	f_seq_sim
0.765	0.734	0.752	0.681	0.689	0.746	0.765	0.699	0.721

# Recall\_General\_test

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate 1	_predicate_sim	f_seq	f_seq_sim
0.759	0.727	0.748	0.661	0.668	0.738	0.759	0.687	0.71

# Recall\_General\_test\_unseen

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate f	_predicate_sim	f_seq	f_seq_sim
0.783	0.742	0.781	0.711	0.719	0.742	0.783	0.663	0.723

# Precision\_per\_value\_val

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.943	0.816	0.946	0.994	0.893	0.86	0.899	0.802

#### Precision\_per\_value\_test

ı	robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
	0.938	0.804	0.943	0.99	0.947	0.875	0.922	0.793

# Precision\_per\_value\_test\_unseen

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.944	0.814	0.921	1.0	0.91	0.793	0.941	0.651

# Recall\_per\_value\_val

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.842	0.689	0.804	0.893	0.888	0.751	0.872	0.901

#### Recall\_per\_value\_test

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.812	0.668	0.847	0.884	0.92	0.778	0.897	0.874

# Recall\_per\_value\_test\_unseen

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.832	0.719	0.744	0.902	0.871	0.756	0.888	0.812