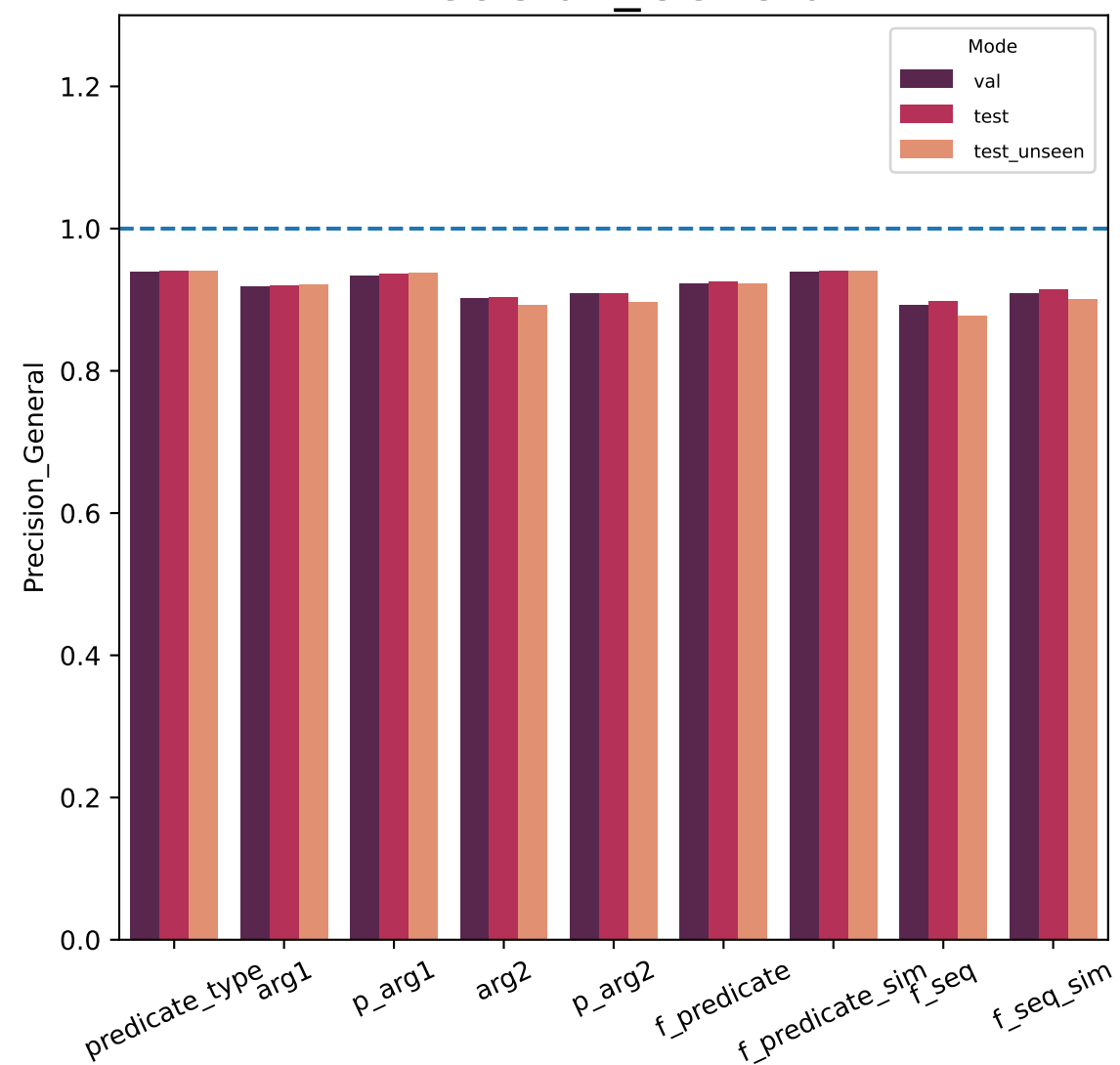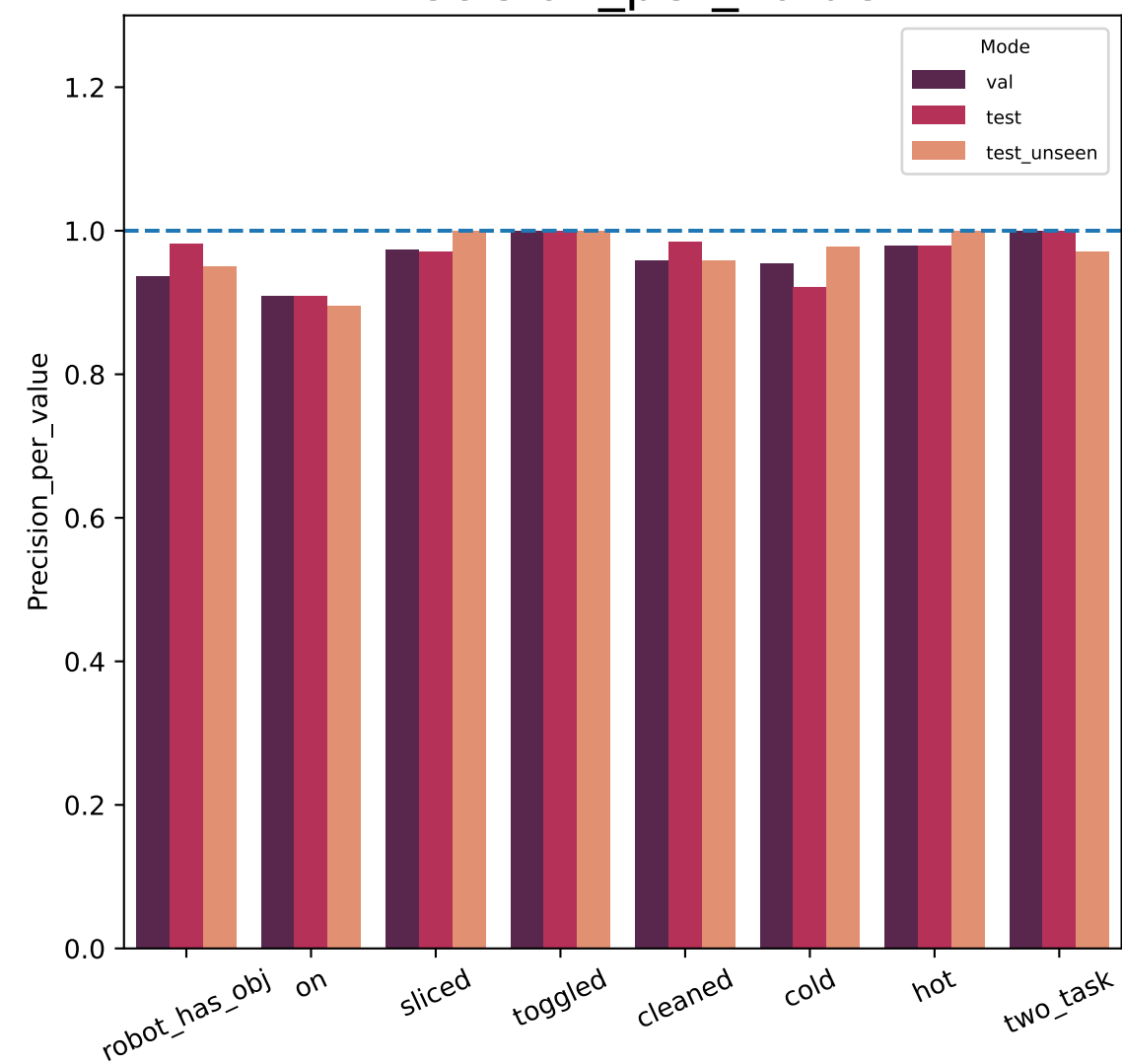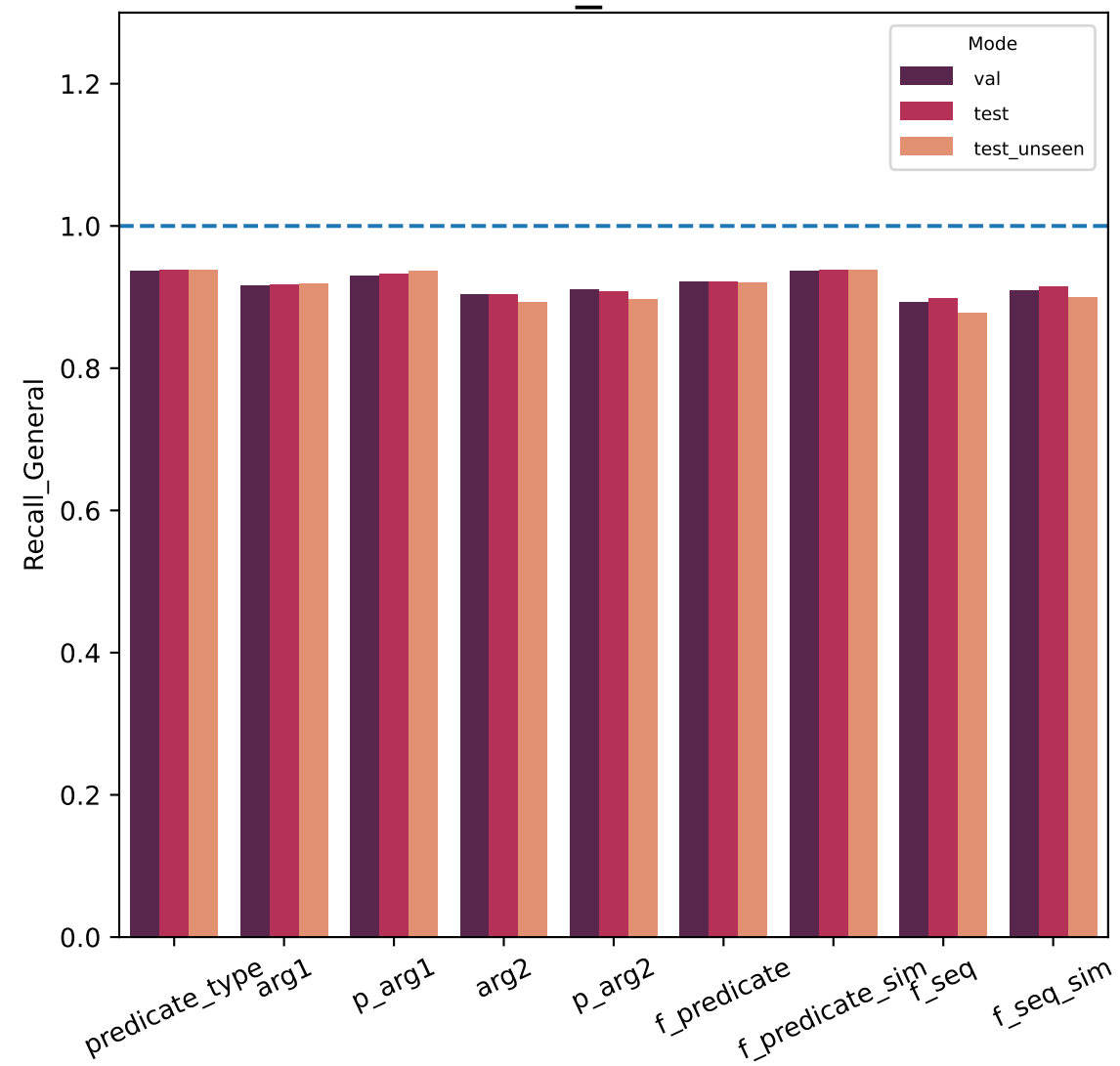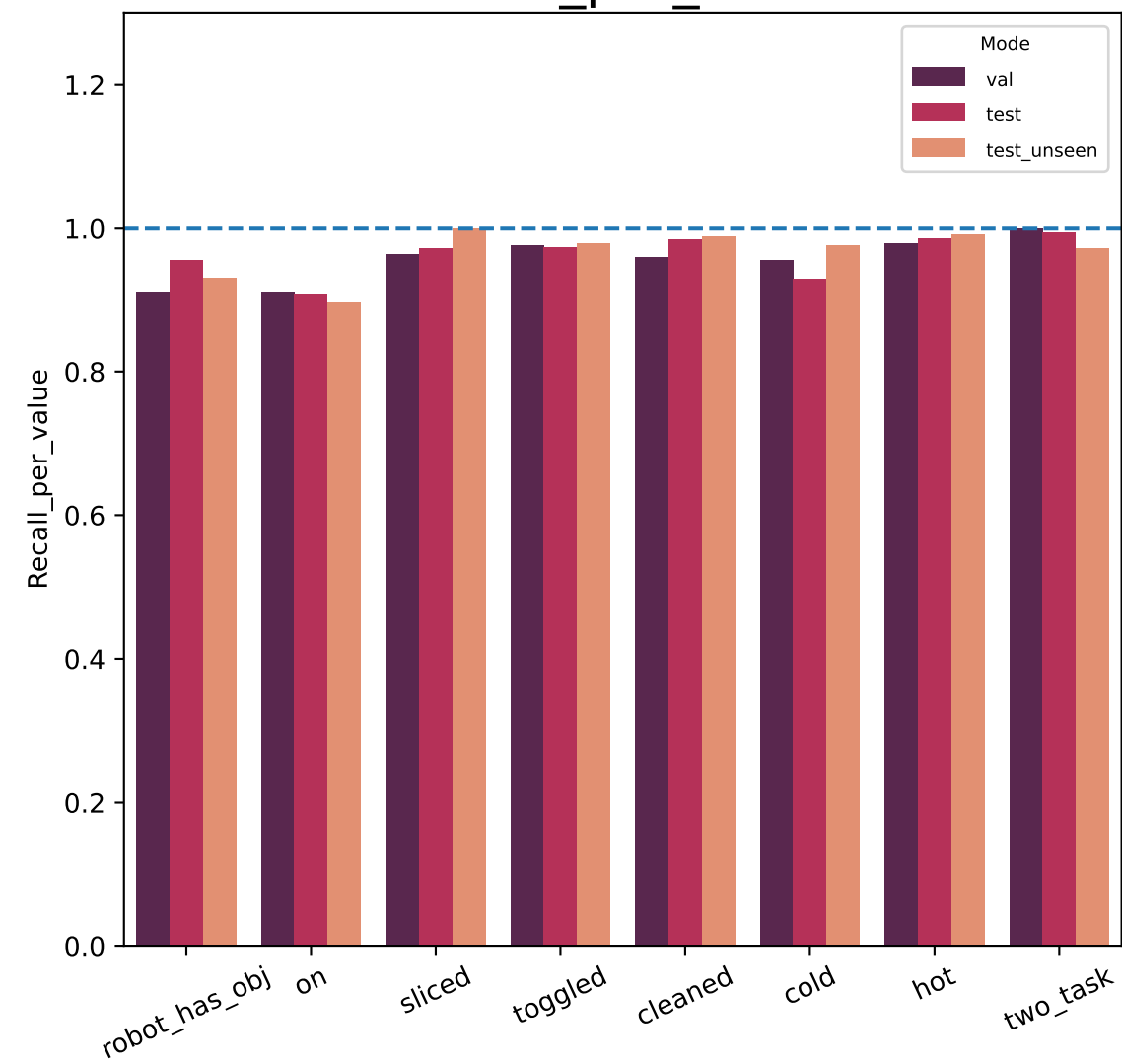Precision_General

Precision_per_value

Recall_General

Recall_per_value

Precision_General_val

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.939	0.919	0.933	0.902	0.909	0.923	0.939	0.893	0.909

Precision_General_test

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.941	0.92	0.936	0.904	0.909	0.925	0.941	0.898	0.915

Precision_General_test_unseen

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.94	0.921	0.938	0.893	0.897	0.922	0.94	0.877	0.9

Precision_per_value_val

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.936	0.909	0.973	1.0	0.959	0.954	0.979	1.0

Precision_per_value_test

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.982	0.909	0.971	1.0	0.985	0.921	0.979	1.0

Precision_per_value_test_unseen

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.95	0.895	1.0	1.0	0.958	0.977	1.0	0.971

Recall_General_val

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.937	0.916	0.93	0.903	0.91	0.921	0.937	0.893	0.909

Recall_General_test

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.938	0.918	0.933	0.904	0.908	0.922	0.938	0.898	0.915

Recall_General_test_unseen

predicate_type	arg1	p_arg1	arg2	p_arg2	f_predicate	f_predicate_sim	f_seq	f_seq_sim
0.938	0.919	0.936	0.893	0.897	0.92	0.938	0.877	0.9

Recall_per_value_val

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.91	0.91	0.962	0.977	0.959	0.954	0.979	1.0

Recall_per_value_test

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.955	0.908	0.971	0.973	0.985	0.929	0.986	0.994

Recall_per_value_test_unseen

robot_has_obj	on	sliced	toggled	cleaned	cold	hot	two_task
0.93	0.897	1.0	0.979	0.989	0.977	0.991	0.971