

1 Results using POM17 to guide algorithm

IGC: Implicit Goal Communication as stated in the paper using a heavily approximated version of the RG10 plan recognition algorithm.

LMG: Landmark Guided Implicit Goal communication. A planner using the novelty based algorithm in the paper except with the plan recognition performed in search being the POM17 plan recognition algorithm instead.

Q is the length of a plan prefix before the belief threshold is met. Longer prefixes therefore suggest that the goal is transmitted to the observer slower.

The data in table 1 considers the ratio of the Q values between LMG and LAMA. Smaller values indicate that the LMG planner out performed the LAMA planner for a particular observer stereotype. The data in table 2 considers the ratio between the Q values between LMG and IGC. Smaller values indicate that the LMG planner out performed the IGC planner for a particular observer stereotype.

	RG10	RG09	POM17
Blocks Word	0.799	0.961	0.719
Grid Navigation	0.872	0.864	1.031
Grid Navigation + Obs	0.992	0.992	1.009
Ticket To Ride	0.992	1.006	1.000
Campus	1.458	0.927	0.875
Intrusion Detection	0.971	1.018	1.061
Logistics	0.752	0.524	0.602
Rover	1.227	1.417	0.907

Table 1: Evaluation of LMG and LAMA. Columns RG10, RG09, POM17 are average ratios $Q(LMG)/Q(LAMA)$ for the three stereotypes considered.

	RG10	RG09	POM17
Blocks Word	1.601	1.260	0.704
Grid Navigation	2.907	2.835	1.031
Grid Navigation + Obs	2.082	2.082	1.009
Ticket To Ride	2.355	2.167	1.060
Campus	1.906	1.188	0.734
Intrusion Detection	1.326	1.355	0.594
Logistics	1.721	2.179	0.864
Rover	1.843	4.130	0.633

Table 2: Evaluation of LMG and IGC. Columns RG10, RG09, POM17 are average ratios $Q(LMG)/Q(IGC)$ for the three stereotypes considered.

The table 3 shows the amount of times for each domain that a planner (out of LAMA, IGC, LMG) had a Q value less than both of the other planners. The

table also shows the maximal difference between the Q value of the winner and the Q value of the runner up. The average is also included.

The table 4 has the same format as table 3, except that it only compares the planners LMG and IGC.

The table 5 has the same format as table 3, except that it only compares the planners LMG and IGC.

		RG10			RG09			POM17		
		F	Δ_m	Δ_a	F	Δ_m	Δ_a	F	Δ_m	Δ_a
Blocks Word	LAMA	0	-	-	1	1	1	0	-	-
	LMG	0	-	-	0	-	-	14	3	1.5
	IGC	14	5	2.214286	8	4	2.375	0	-	-
	Tie	9	0	0	14	0	0	9	0	0
Campus	LAMA	0	-	-	0	-	-	0	-	-
	LMG	0	-	-	0	-	-	4	1	1
	IGC	4	2	1.5	2	1	1	0	-	-
	Tie	12	0	0	14	0	0	12	0	0
Grid Nav	LAMA	0	-	-	0	-	-	0	-	-
	LMG	4	6	3.25	4	6	3.25	0	-	-
	IGC	12	13	4.166667	12	11	3.916667	0	-	-
	Tie	5	0	0	5	0	0	21	0	0
Grid Nav + Obs	LAMA	0	-	-	0	-	-	0	-	-
	LMG	0	-	-	0	-	-	0	-	-
	IGC	14	6	2.357143	14	6	2.357143	0	-	-
	Tie	10	0	0	10	0	0	24	0	0
Ticket To Ride	LAMA	0	-	-	0	-	-	0	-	-
	LMG	0	-	-	0	-	-	0	-	-
	IGC	20	11	2.95	19	11	2.947368	3	4	3.333333
	Tie	1	0	0	2	0	0	18	0	0
Intrusion	LAMA	0	-	-	0	-	-	2	6	3.5
	LMG	0	-	-	0	-	-	5	9	3.2
	IGC	15	4	1.4	14	4	1.357143	2	3	2
	Tie	8	0	0	9	0	0	14	0	0
Logistics	LAMA	0	-	-	0	-	-	0	-	-
	LMG	0	-	-	2	4	3.5	6	3	2
	IGC	11	6	2.545455	9	11	4.777778	3	1	1
	Tie	9	0	0	9	0	0	11	0	0
Rover	LAMA	1	1	1	2	1	1	0	-	-
	LMG	0	-	-	0	-	-	3	1	1
	IGC	13	1	1	14	9	5	1	1	1
	Tie	4	0	0	2	0	0	14	0	0

Table 3: Comparison between IGC, LMG and LAMA over stereotypes RG09, RG10 and POM17. Column F for each planner, Y , is the number of times $\forall X. Q(Y) < Q(X)$, where X is the other two planners. Columns and report maximal and average difference $Q(X2) - Q(X1)$, where $X2$ and $X1$ are the slowest and fastest planner.

		POM17			RG09			RG10		
		F	Δ_m	Δ_a	F	Δ_m	Δ_a	F	Δ_m	Δ_a
Blocks Word	LMG	16	3	1.625	2	3	2	0	-	-
	IGC	0	-	-	11	4	2.090909	14	5	2.285714
	Tie	7	0	0	10	0	0	9	0	0
Campus	LMG	8	3	1.625	0	-	-	0	-	-
	IGC	0	-	-	3	1	1	10	3	1.6
	Tie	8	0	0	13	0	0	6	0	0
Grid Nav	LMG	0	-	-	4	6	3.25	4	6	3.25
	IGC	2	3	3	14	11	4.071429	14	13	4.285714
	Tie	19	0	0	3	0	0	3	0	0
Grid Nav + Obs	LMG	0	-	-	0	-	-	0	-	-
	IGC	1	2	2	14	6	2.714286	14	6	2.714286
	Tie	23	0	0	10	0	0	10	0	0
Ticket To Ride	LMG	0	-	-	0	-	-	0	-	-
	IGC	3	4	3.333333	19	11	3	20	11	2.95
	Tie	18	0	0	2	0	0	1	0	0
Intrusion	LMG	19	9	3.894737	0	-	-	0	-	-
	IGC	3	3	1.666667	15	4	1.466667	15	4	1.533333
	Tie	1	0	0	8	0	0	8	0	0
Logistics	LMG	7	12	6.714286	2	14	13.5	0	-	-
	IGC	3	1	1	9	11	5.111111	11	7	3.181818
	Tie	10	0	0	9	0	0	9	0	0
Rover	LMG	14	5	1.785714	1	1	1	0	-	-
	IGC	1	1	1	16	9	5	14	2	1.357143
	Tie	3	0	0	1	0	0	4	0	0

Table 4: Comparison between LMG and IGC over stereotypes RG09, RG10 and POM17. Column F is the number of times $Q(LMG) < Q(IGC)$ (and vice versa). Columns and report maximal and average difference $Q(X2) - Q(X1)$, where $X2$ and $X1$ are the slowest and fastest planner.

		POM17			RG09			RG10		
		F	Δ_m	Δ_a	F	Δ_m	Δ_a	F	Δ_m	Δ_a
Blocks Word	LAMA	0	-	-	4	2	1.25	1	1	1
	LMG	14	3	1.785714	7	4	2.142857	13	4	2.230769
	Tie	9	0	0	12	0	0	9	0	0
Campus	LAMA	0	-	-	1	1	1	8	2	1.25
	LMG	4	1	1	4	2	1.75	0	-	-
	Tie	12	0	0	11	0	0	8	0	0
Grid Nav	LAMA	2	3	3	6	3	1.666667	6	3	1.666667
	LMG	0	-	-	11	8	4.363636	11	8	4.090909
	Tie	19	0	0	4	0	0	4	0	0
Grid Nav + Obs	LAMA	1	2	2	3	3	1.666667	3	3	1.666667
	LMG	0	-	-	5	5	2.4	5	5	2.4
	Tie	23	0	0	16	0	0	16	0	0
Ticket To Ride	LAMA	0	-	-	1	1	1	0	-	-
	LMG	0	-	-	0	-	-	1	1	1
	Tie	21	0	0	20	0	0	20	0	0
Intrusion	LAMA	3	6	2.666667	2	2	1.5	1	2	2
	LMG	6	9	3.5	5	1	1	5	2	1.2
	Tie	14	0	0	16	0	0	17	0	0
Logistics	LAMA	0	-	-	1	3	3	4	2	1.75
	LMG	18	17	5.777778	16	20	10.125	13	17	4.692308
	Tie	2	0	0	3	0	0	3	0	0
Rover	LAMA	0	-	-	5	4	2.4	6	2	1.166667
	LMG	3	2	1.333333	1	3	3	0	-	-
	Tie	15	0	0	12	0	0	12	0	0

Table 5: Comparison between LMG and LAMA over stereotypes RG09, RG10 and POM17. Column F is the number of times $Q(LMG) < Q(LAMA)$ (and vice versa). Columns and report maximal and average difference $Q(X2) - Q(X1)$, where $X2$ and $X1$ are the slowest and fastest planner.