

LP-Based Approaches for Goal Recognition as Planning (Supplemental Material)

Paper #6100

In this document, we present a supplemental material containing a complete comparison between our LP-based heuristics and the state-of-the-art in goal recognition as planning. We evaluated these approaches over thousands of goal recognition problems using several IPC domains models, varying the levels of observability for partial and full observable plans in 10%, 30%, 50%, 70%, and 100%, and for noisy, partial and full observable plans in 25%, 50%, 75%, and 100%. Note that the noisy datasets contain 2 noisy (spurious) actions in every observation sequence for all levels of observability. Part of the results we presented here in this document have been used to build the tables (Tables 1 and 2 in the original paper) in the original submitted paper.

Here, as in the original submitted paper, we use the following metrics for evaluation: *Accuracy*, the fraction of times that the correct hidden goal was among the goals found to be most likely; *Spread in \mathcal{G}* the average number of returned goals; and recognition time in seconds. Specifically, each row in Tables 1 and 2 express averages for the number of candidate goals $|\mathcal{G}|$; the percentage of the plan that is actually observed % Obs; the average number of observations (actions) per problem $|O|$; and for each approach, the time in seconds to recognize the goal given the observations (Time); Acc % with which the approaches correctly infer the hidden goal; and S in \mathcal{G} represents the average number of returned goals. Below the name of each domain contains the number of goal recognition problems. At the bottom of Tables 1 and 2, we provide the average for evaluated metrics we used over all domains and problems in the datasets.

We evaluate the effectiveness of our heuristics against the state-of-the-art in goal recognition as planning, and the approaches we compare against are: R&G 2009 (Ramírez and Geffner 2009), FGR 2015 (E-Martín, R.-Moreno, and Smith 2015), POM 2017 h_{gc} (Pereira, Oren, and Meneguzzi 2017), POM 2017 h_{uniq} (Pereira, Oren, and Meneguzzi 2017), and M+L 2018 (Vered et al. 2018). Since the approach of Sohrabi, Riabov, and Udrea (2016) timed out for virtually all problems in all domains (using a top-K planner, with $K=1000$, as the main author of the paper recommend to be used), even with a twenty-minute timeout, so we excluded the results from our comparison. Note that FGR 2015 timed out for all problem for DEPOTS, and had parsing issues

for some domains. When an approach had issues regarding time, we use *Timeout* in the tables. As for the parsing issues, we use the symbol †. M+L 2018 has faced some issues with respect to memory usage (we set a maximum usage limit of 4GB), so we use *Out of Memory* to indicate memory issues in the tables.

References

- E-Martín, Y.; R.-Moreno, M. D.; and Smith, D. E. 2015. A Goal Recognition Technique Based on Interaction Estimates. *IJCAI*.
- Pereira, R. F.; Oren, N.; and Meneguzzi, F. 2017. Landmark-Based Heuristics for Goal Recognition. *AAAI*.
- Ramírez, M., and Geffner, H. 2009. Plan recognition as planning. *IJCAI*.
- Sohrabi, S.; Riabov, A. V.; and Udrea, O. 2016. Plan Recognition as Planning Revisited. *IJCAI*.
- Vered, M.; Pereira, R. F.; Magnaguagno, M.; Meneguzzi, F.; and Kaminka, G. A. 2018. Towards Online Goal Recognition Combining Goal Mirroring and Landmarks. *AAMAS*.

Partial and Full Observability																									
					δ_{HC}			δ_{HCU}			R&G 2009			FGR 2015			POM 2017 h_{gc}			POM 2017 h_{uniq}			M+L 2018		
#	$ \mathcal{G} $	% Obs	$ \mathcal{O} $	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	Time	Acc %	S in \mathcal{G}	
BLOCKS (1076)	20.3	10	1.8	11.575	95.5%	7.81	11.574	95.9%	8.7	1.22	86.8%	7.84	46.563	65.9%	9.11	0.144	39.9%	1.05	0.131	31.7%	1.04	Out of Memory	Out of Memory	Out of Memory	
		30	4.9	11.57	89.4%	2.79	11.567	94.3%	5.12	1.639	87.2%	3.57	36.649	78.0%	10.53	0.156	50.6%	1.09	0.144	51.4%	1.06	Out of Memory	Out of Memory	Out of Memory	
		50	7.6	11.581	92.7%	1.78	11.587	93.9%	3.24	2.402	97.9%	2.63	34.291	81.3%	10.68	0.179	65.0%	1.09	0.168	60.1%	1.08	Out of Memory	Out of Memory	Out of Memory	
		70	11.1	11.594	98.8%	1.41	11.595	98.8%	1.8	3.785	97.5%	1.83	37.057	89.8%	8.63	0.192	84.8%	1.12	0.184	79.0%	1.14	Out of Memory	Out of Memory	Out of Memory	
		100	14.5	11.904	100.0%	1.21	11.937	100.0%	1.21	6.791	100.0%	1.46	40.405	100.0%	1.23	0.246	100.0%	1.36	0.239	100.0%	1.09	Out of Memory	Out of Memory	Out of Memory	
DEPOTS (364)	8.5	10	3.1	8.299	61.9%	2.15	8.305	71.4%	3.51	1.496	77.4%	3.99	†	†	†	0.369	35.7%	1.18	0.393	32.1%	1.1	Out of Memory	Out of Memory	Out of Memory	
		30	8.6	8.292	71.4%	1.4	8.298	88.1%	4.17	2.309	77.4%	2.39	†	†	†	0.357	58.3%	1.06	0.393	47.6%	1.07	Out of Memory	Out of Memory	Out of Memory	
		50	14.1	8.282	92.9%	1.38	8.28	98.8%	3.65	3.411	84.5%	1.92	†	†	†	0.369	76.2%	1.06	0.405	71.4%	1.02	Out of Memory	Out of Memory	Out of Memory	
		70	19.7	8.282	97.6%	1.06	8.279	98.8%	1.75	5.27	91.7%	1.68	†	†	†	0.393	89.3%	1.01	0.44	84.5%	1.01	Out of Memory	Out of Memory	Out of Memory	
		100	24.4	8.296	100.0%	1.0	8.296	100.0%	1.0	7.117	92.9%	1.46	†	†	†	0.464	100.0%	1.04	0.5	100.0%	1.04	Out of Memory	Out of Memory	Out of Memory	
DRIVERLOG (364)	10.5	10	2.6	5.159	77.4%	2.61	5.168	78.6%	3.17	1.169	96.4%	4.71	79.488	42.9%	1.92	0.333	41.7%	1.04	0.321	35.7%	1.11	Out of Memory	Out of Memory	Out of Memory	
		30	6.9	5.149	83.3%	1.69	5.163	91.7%	2.76	1.411	92.9%	3.35	60.168	70.2%	3.19	0.31	54.8%	1.13	0.31	47.6%	1.1	Out of Memory	Out of Memory	Out of Memory	
		50	11.1	5.167	92.9%	1.23	5.146	97.6%	2.0	1.694	94.0%	2.88	64.427	79.8%	4.6	0.321	72.6%	1.17	0.31	64.3%	1.14	Out of Memory	Out of Memory	Out of Memory	
		70	15.6	5.148	95.2%	1.14	5.146	95.2%	1.5	1.973	89.3%	2.46	75.085	82.1%	4.11	0.333	90.5%	1.14	0.321	90.5%	1.17	Out of Memory	Out of Memory	Out of Memory	
		100	21.7	5.166	100.0%	1.04	5.157	100.0%	1.04	2.821	89.3%	2.14	96.091	96.4%	1.11	0.321	100.0%	1.21	0.321	100.0%	1.18	Out of Memory	Out of Memory	Out of Memory	
DWR (364)	7.3	10	5.7	5.759	54.8%	2.21	5.745	95.2%	5.45	1.767	83.3%	4.21	66.496	92.9%	6.38	0.452	36.9%	1.1	0.512	33.3%	1.06	Out of Memory	Out of Memory	Out of Memory	
		30	16.0	5.745	83.3%	1.58	5.753	100.0%	4.92	2.723	81.0%	3.35	54.461	97.6%	6.56	0.452	60.7%	1.04	0.5	51.2%	1.06	Out of Memory	Out of Memory	Out of Memory	
		50	26.2	5.742	90.5%	1.21	5.734	100.0%	3.98	4.822	72.6%	2.27	56.256	98.8%	6.27	0.488	66.7%	1.0	0.548	61.9%	1.01	Out of Memory	Out of Memory	Out of Memory	
		70	36.8	5.75	97.6%	1.07	5.735	100.0%	2.26	10.914	70.2%	2.05	65.101	98.8%	6.0	0.536	89.3%	1.0	0.607	78.6%	1.05	Out of Memory	Out of Memory	Out of Memory	
		100	51.9	5.763	100.0%	1.0	5.75	100.0%	1.0	25.092	67.9%	1.68	86.46	100.0%	1.0	0.643	100.0%	1.0	0.75	96.4%	1.04	Out of Memory	Out of Memory	Out of Memory	
IPC-GRID (673)	9.0	10	2.9	6.216	92.8%	1.92	6.209	94.8%	2.32	1.091	96.1%	2.46	Timeout	Timeout	Timeout	0.248	66.7%	2.58	0.242	62.7%	2.58	0.758	58.2%	2.11	
		30	7.8	6.051	95.4%	1.29	6.044	98.0%	1.48	1.476	97.4%	1.42	Timeout	Timeout	Timeout	0.242	81.7%	1.65	0.242	83.7%	1.66	0.699	80.4%	1.62	
		50	12.7	6.123	98.7%	1.11	6.122	100.0%	1.25	1.905	100.0%	1.16	Timeout	Timeout	Timeout	0.261	90.8%	1.18	0.248	90.8%	1.18	0.667	93.5%	1.21	
		70	17.9	6.251	99.4%	1.1	6.259	100.0%	1.19	2.552	100.0%	1.05	Timeout	Timeout	Timeout	0.268	97.4%	1.07	0.268	97.4%	1.07	0.673	91.5%	1.08	
		100	24.8	5.825	100.0%	1.03	5.826	100.0%	1.03	4.057	100.0%	1.0	Timeout	Timeout	Timeout	0.262	100.0%	1.0	0.262	100.0%	1.0	0.59	83.6%	1.1	
FERRY (364)	7.5	10	2.9	4.201	100.0%	3.17	4.201	100.0%	3.2	0.491	98.8%	3.37	6.66	91.7%	6.65	0.071	58.3%	1.26	0.071	58.3%	1.18	4.548	51.2%	1.35	
		30	7.6	4.072	100.0%	1.56	4.075	100.0%	1.76	0.677	100.0%	1.76	6.801	100.0%	7.57	0.06	85.7%	1.12	0.06	83.3%	1.06	4.393	73.8%	1.13	
		50	12.3	4.141	100.0%	1.29	4.141	100.0%	1.44	0.795	100.0%	1.42	8.297	100.0%	7.57	0.06	95.2%	1.07	0.06	91.7%	1.01	4.31	81.0%	1.12	
		70	17.3	4.199	100.0%	1.1	4.197	100.0%	1.12	1.253	98.8%	1.14	10.65	100.0%	7.32	0.071	100.0%	1.01	0.071	100.0%	1.0	4.274	96.4%	1.05	
		100	24.2	4.196	100.0%	1.07	4.204	100.0%	1.07	1.632	100.0%	1.07	13.625	100.0%	1.07	0.071	100.0%	1.0	0.071	100.0%	1.0	4.214	100.0%	1.0	
LOGISTICS (673)	10.5	10	2.9	6.79	100.0%	2.5	6.789	100.0%	2.8	1.201	99.3%	2.98	†	†	†	0.641	55.6%	1.73	0.641	49.0%	1.24	3.967	51.0%	1.63	
		30	8.2	6.953	98.0%	1.3	6.944	98.0%	1.76	1.799	98.7%	1.39	†	†	†	0.621	80.4%	1.21	0.634	76.5%	1.12	3.83	77.8%	1.46	
		50	13.4	6.944	98.7%	1.13	6.959	98.7%	1.37	2.509	98.7%	1.29	†	†	†	0.641	90.2%	1.1	0.647	86.3%	1.05	5.085	88.9%	1.32	
		70	18.9	6.938	100.0%	1.08	6.95	100.0%	1.15	3.461	100.0%	1.13	†	†	†	0.667	96.7%	1.06	0.66	96.7%	1.02	3.739	95.4%	1.3	
		100	26.5	6.63	100.0%	1.0	6.633	100.0%	1.0	4.832	100.0%	1.0	†	†	†	0.607	100.0%	1.0	0.607	100.0%	1.0	3.213	100.0%	1.1	
MICRONIC (364)	6.0	10	3.9	4.905	100.0%	2.12	4.886	100.0%	2.29	0.813	100.0%	3.26	†	†	†	0.464	67.9%	1.33	0.352	54.8%	1.26	9.548	58.3%	1.54	
		30	11.1	4.897	100.0%	1.19	4.895	100.0%	1.46	1.191	100.0%	1.58	†	†	†	0.452	96.4%	1.11	0.364	90.5%	1.08	9.286	77.4%	1.26	
		50	18.1	4.901	100.0%	1.1	4.89	100.0%	1.32	1.722	100.0%	1.29	†	†	†	0.452	96.4%	1.01	0.352	96.4%	1.0	9.238	88.1%	1.21	
		70	25.3	4.892	100.0%	1.01	4.891	100.0%	1.02	2.59	100.0%	1.04	†	†	†	0.452	100.0%	1.01	0.376	100.0%	1.01	9.167	96.4%	1.12	
		100	35.6	4.894	100.0%	1.0	4.894	100.0%	1.0	5.107	100.0%	1.0	†	†	†	0.464	100.0%	1.0	0.364	100.0%	1.0	9.036	100.0%	1.0	
ROVERS (364)	6.0	10	3.0	4.863	98.8%	2.71	4.858	100.0%	2.94	0.745	98.8%	2.86	†	†	†	0.348	64.3%	1.73	0.371	51.2%	1.11	20.893	65.5%	1.85	
		30	7.9	4.881	85.7%	1.17	4.87	91.7%	1.83	1.031	100.0%	1.67	†	†	†	0.348	83.3%	1.24	0.348	69.0%	1.07	22.714	81.0%	1.42	
		50	12.7	4.87	98.8%	1.14	4.856	98.8%	1.44	1.345	100.0%	1.3	†	†	†	0.336	92.9%	1.08	0.348	85.7%	1.01	17.119	90.5%	1.13	
		70	17.9	4.858	98.8%	1.01	4.838	98.8%	1.06	1.177	100.0%	1.07	†	†	†	0.348	98.8%	1.01	0.36	91.7%	1.0	17.095	98.8%	1.07	
		100	24.9	4.823	100.0%	1.0	4.884	100.0%	1.0	2.298	100.0%	1.07	†	†	†	0.371	100.0%	1.0	0.571	100.0%	1.0	16.571	100.0%	1.04	
SATELLITE (364)	6.5	10	2.1	5.196	91.7%	2.73	5.204	92.9%	2.92	1.076	97.6%	3.42	14.82	89.3%	4.87	0.45	57.1%	1.56	0.45	47.6%	1.21	2.226	69.0%	2.05	
		30	5.4	5.178	92.9%	1.76	5.201	96.4%	2.31	1.183	97.6%	2.4	32.171	86.9%	4.21	0.45	76.2%	1.31	0.414	69.0%	1.14	2.024	83.3%	1.42	
		50	8.7	5.197	96.4%	1.32	5.194	98.8%	1.77	1.328	97.6%	1.69	51.567	88.1%	3.65	0.426	85.7%	1.1	0.414	81.0%	1.11	1.905	92.9%	1.27	
		70	12.2	5.191	97.6%	1.11	5.193	97.6%	1.21	1.841	96.4%	1.52	75.363	92.9%	2.89	0.402	97.6%	1.02	0.414	94.0%	1.04	1.893	98.8%	1.06	
		100	16.8	5.212	100.0%	1.07	5.205	100.0%	1.07	2.045	96.4%	1.32	113.381	100.0%	2.57	0.414	100.0%	1.07	0.414	100.0%	1.07	1.786	100.0%	1.04	
SOKOBAN (364)	7.3	10	3.1	7.827	72.6%	1.61	7.849	78.6%	2.39	3.153	69.0%	4.02	461.701	67.9%	2.99	0.607	53.6%	2.06	0.607	51.2%	1.86	Out of Memory	Out of Memory	Out of Memory	
		30	8.7	7.744	89.3%	1.11	7.742	95.2%	1.75	4.622	89.3%	4.17	370.413	83.3%	3.14	0.595	57.1%	1.37	0.607	56.0%	1.21	Out of Memory	Out of Memory	Out of Memory	
		50	14.1	7.701	95.2%	1.08	7.709	100.0%	1.46	7.441	89.3%	4.11	358.028	82.1%	2.27	0.595	71.4%	1.32	0.607	69.0%	1.2	Out of Memory	Out of Memory	Out of Memory	
		70																							

Noisy, Partial, and Full Observability																								
				δ_{HC}			δ_{HCU}			R&G 2009			FGR 2015			POM 2017 h_{gc}			POM 2017 h_{uniq}			M+L 2018		
#	G	% Obs	O	Time	Acc %	S in G	Time	Acc %	S in G	Time	Acc %	S in G	Time	Acc %	S in G	Time	Acc %	S in G	Time	Acc %	S in G	Time	Acc %	S in G
BLOCKS (144)	20.3	25	2.4	16.033	58.3%	6.22	16.019	75.0%	10.39	1.045	38.9%	5.39	3.044	69.4%	12.67	0.083	2.8%	1.22	0.083	8.3%	1.0	Out of Memory	Out of Memory	Out of Memory
		50	4.4	12.626	52.8%	3.31	12.675	88.9%	12.44	1.122	52.8%	4.61	2.71	80.6%	15.58	0.083	25.0%	1.19	0.083	13.9%	1.08	Out of Memory	Out of Memory	Out of Memory
		75	6.8	11.282	80.6%	2.11	11.235	91.7%	7.36	1.405	75.0%	2.72	2.647	77.8%	12.39	0.056	47.2%	1.19	0.056	38.9%	1.25	Out of Memory	Out of Memory	Out of Memory
		100	8.8	10.576	88.9%	1.92	10.629	97.2%	2.86	1.652	86.1%	2.03	2.693	88.9%	11.31	0.083	77.8%	1.36	0.056	75.0%	1.33	Out of Memory	Out of Memory	Out of Memory
DIPOTS (144)	9.3	25	4.4	11.495	52.8%	2.19	11.445	63.9%	4.47	0.284	5.6%	9.17	†	†	†	0.528	38.9%	1.64	0.528	27.8%	1.22	Out of Memory	Out of Memory	Out of Memory
		50	8.4	10.299	61.1%	1.67	10.263	83.3%	4.14	0.189	0.0%	9.33	†	†	†	0.472	52.8%	1.22	0.472	41.7%	1.19	Out of Memory	Out of Memory	Out of Memory
		75	12.7	8.671	88.9%	1.19	8.674	94.4%	2.14	0.361	11.1%	8.28	†	†	†	0.472	80.6%	1.11	0.5	75.0%	1.06	Out of Memory	Out of Memory	Out of Memory
		100	16.2	8.098	94.4%	1.19	8.075	91.7%	1.22	0.292	5.6%	8.83	†	†	†	0.472	88.9%	1.11	0.472	86.1%	1.11	Out of Memory	Out of Memory	Out of Memory
DRIVERLOG (144)	6.6	25	3.5	5.383	55.6%	2.61	5.362	83.3%	4.69	0.234	44.4%	5.89	1.732	36.1%	2.33	0.111	36.1%	1.33	0.111	25.0%	1.08	Out of Memory	Out of Memory	Out of Memory
		50	6.7	4.982	77.8%	1.72	4.97	91.7%	3.47	0.282	38.9%	4.72	1.625	55.6%	3.03	0.083	58.3%	1.28	0.083	52.8%	1.11	Out of Memory	Out of Memory	Out of Memory
		75	10.0	4.825	86.1%	1.25	4.855	94.4%	2.31	0.239	30.6%	5.47	1.902	61.1%	2.67	0.083	61.1%	1.33	0.111	52.8%	1.14	Out of Memory	Out of Memory	Out of Memory
		100	12.8	4.511	97.2%	1.31	4.511	97.2%	1.64	0.358	44.4%	4.42	2.205	75.0%	2.78	0.083	94.4%	1.47	0.083	97.2%	1.42	Out of Memory	Out of Memory	Out of Memory
DWR (144)	7.0	25	9.2	7.546	72.2%	2.0	7.614	97.2%	5.11	0.808	41.7%	5.67	37.661	94.4%	5.89	0.444	44.4%	1.14	0.5	33.3%	1.0	Out of Memory	Out of Memory	Out of Memory
		50	17.8	7.108	80.6%	1.67	7.153	94.4%	4.47	1.569	22.2%	5.39	32.228	100.0%	5.28	0.417	63.9%	1.08	0.444	50.0%	1.06	Out of Memory	Out of Memory	Out of Memory
		75	26.6	6.509	91.7%	1.22	6.482	94.4%	1.78	2.793	19.4%	5.5	34.49	94.4%	4.94	0.417	94.4%	1.06	0.472	69.4%	1.08	Out of Memory	Out of Memory	Out of Memory
		100	34.9	5.803	100.0%	1.08	5.768	97.2%	1.06	7.392	30.6%	4.42	39.713	94.4%	2.22	0.444	94.4%	1.0	0.472	94.4%	1.03	Out of Memory	Out of Memory	Out of Memory
IPC-GRID (144)	8.3	25	4.0	7.871	81.1%	1.67	7.889	85.6%	2.61	0.265	12.2%	7.56	Timeout	Timeout	Timeout	0.244	58.9%	1.78	0.233	53.3%	1.72	0.522	64.4%	1.89
		50	7.7	6.03	94.4%	1.14	6.011	94.4%	1.71	0.2	4.4%	8.07	Timeout	Timeout	Timeout	0.222	85.6%	1.33	0.211	83.3%	1.32	0.389	83.3%	1.36
		75	11.5	5.481	98.9%	1.1	5.497	97.8%	1.13	0.223	6.7%	7.89	Timeout	Timeout	Timeout	0.211	94.4%	1.09	0.211	94.4%	1.09	0.378	91.1%	1.17
		100	16.9	5.01	100.0%	1.0	4.988	90.0%	0.9	0.266	10.0%	7.77	Timeout	Timeout	Timeout	0.2	100.0%	1.0	0.2	100.0%	1.0	0.4	76.7%	1.1
FERRY (144)	7.0	25	5.8	4.745	80.6%	2.78	4.736	86.1%	5.11	0.256	75.0%	2.92	0.495	100.0%	7.0	0.028	47.2%	1.28	0.028	27.8%	1.08	0.583	50.0%	1.69
		50	11.2	4.402	97.2%	1.72	4.365	97.2%	3.83	0.375	94.4%	1.94	0.545	100.0%	7.0	0.028	88.9%	1.31	0.028	77.8%	1.06	0.5	83.3%	1.19
		75	16.6	4.236	94.4%	1.56	4.223	97.2%	2.42	0.585	88.9%	1.53	0.643	94.4%	6.64	0.028	97.2%	1.17	0.028	83.3%	1.08	0.444	83.3%	1.22
		100	21.9	4.187	97.2%	1.17	4.183	97.2%	1.17	0.906	97.2%	1.25	0.783	94.4%	4.39	0.028	100.0%	1.08	0.028	97.2%	1.06	0.472	100.0%	1.14
LOGISTICS (144)	10.0	25	4.8	7.461	91.7%	2.11	7.446	97.2%	3.42	0.203	5.6%	9.42	†	†	†	0.222	61.1%	1.56	0.222	38.9%	1.08	0.722	50.0%	1.53
		50	9.4	6.875	97.2%	1.11	6.848	97.2%	1.36	0.214	5.6%	9.33	†	†	†	0.194	83.3%	1.17	0.167	75.0%	1.06	0.611	72.2%	1.47
		75	14.1	6.253	100.0%	1.03	6.234	100.0%	1.06	0.259	13.9%	8.78	†	†	†	0.194	97.2%	1.0	0.167	100.0%	1.03	0.583	91.7%	1.42
		100	18.1	5.436	100.0%	1.08	5.438	100.0%	1.08	0.321	13.9%	8.78	†	†	†	0.194	100.0%	1.06	0.194	100.0%	1.03	0.583	94.4%	1.28
MICRONIC (144)	6.0	25	4.4	4.892	52.8%	1.97	4.908	83.3%	4.17	0.318	88.9%	2.89	†	†	†	0.111	50.0%	1.39	0.111	33.3%	1.17	0.306	50.0%	1.56
		50	8.4	4.533	80.6%	1.17	4.566	97.2%	2.17	0.387	100.0%	1.75	†	†	†	0.083	83.3%	1.11	0.083	80.6%	1.11	0.25	77.8%	1.36
		75	12.6	4.387	91.7%	1.06	4.379	100.0%	1.31	0.483	100.0%	1.19	†	†	†	0.083	97.2%	1.03	0.083	88.9%	1.03	0.25	86.1%	1.08
		100	16.3	4.294	100.0%	1.03	4.311	100.0%	1.08	0.628	100.0%	1.0	†	†	†	0.056	100.0%	1.0	0.056	100.0%	1.0	0.25	97.2%	1.06
ROVERS (144)	6.0	25	3.1	5.346	72.2%	2.22	5.359	75.0%	2.78	0.264	33.3%	4.78	†	†	†	0.083	52.8%	1.14	0.083	50.0%	1.14	0.167	52.8%	1.19
		50	5.7	4.973	80.6%	1.61	4.941	88.9%	2.92	0.296	50.0%	3.81	†	†	†	0.056	69.4%	1.31	0.056	58.3%	1.08	0.167	63.9%	1.11
		75	8.4	4.889	91.7%	1.14	4.905	100.0%	1.47	0.301	44.4%	3.92	†	†	†	0.083	86.1%	1.14	0.056	75.0%	1.08	0.167	83.3%	1.14
		100	10.8	4.513	100.0%	1.03	4.53	100.0%	1.03	0.305	38.9%	4.06	†	†	†	0.083	97.2%	1.11	0.056	86.1%	1.03	0.139	88.9%	1.03
SATELLITE (144)	6.0	25	3.3	4.51	75.0%	3.31	4.444	88.9%	4.28	0.224	58.3%	4.53	0.399	83.3%	4.94	0.056	52.8%	2.42	0.056	30.6%	1.33	0.111	58.3%	2.44
		50	5.7	4.056	72.2%	2.44	4.091	83.3%	3.92	0.255	72.2%	3.58	0.446	72.2%	4.44	0.028	72.2%	2.08	0.028	44.4%	1.31	0.111	77.8%	2.08
		75	8.4	3.915	83.3%	1.44	3.966	88.9%	2.83	0.28	77.8%	2.75	0.486	83.3%	4.17	0.028	80.6%	1.28	0.028	69.4%	1.08	0.111	80.6%	1.22
		100	10.7	3.929	94.4%	1.47	3.935	94.4%	1.86	0.298	72.2%	3.0	0.496	86.1%	3.94	0.056	94.4%	1.31	0.028	91.7%	1.19	0.083	91.7%	1.19
SOKOBAN (144)	8.6	25	5.3	13.104	36.1%	1.64	13.101	72.2%	4.69	1.953	25.0%	7.28	346.669	58.3%	4.97	0.75	41.7%	1.75	0.75	38.9%	1.56	Out of Memory	Out of Memory	Out of Memory
		50	10.3	11.319	50.0%	1.17	11.323	58.3%	1.94	2.086	19.4%	6.67	256.744	61.1%	3.56	0.667	66.7%	1.44	0.694	58.3%	1.08	Out of Memory	Out of Memory	Out of Memory
		75	15.6	9.527	36.1%	1.0	9.46	36.1%	0.56	2.121	19.4%	7.69	222.685	61.1%	3.36	0.694	80.6%	1.28	0.667	72.2%	1.03	Out of Memory	Out of Memory	Out of Memory
		100	20.1	8.946	33.3%	1.11	9.029	36.1%	0.5	5.878	33.3%	6.47	217.736	61.1%	2.75	0.694	94.4%	1.22	0.694	86.1%	1.06	Out of Memory	Out of Memory	Out of Memory
ZENO (144)	6.6	25	3.0	8.209	44.4%	2.72	8.27	80.6%	5.25	0.946	72.2%	3.92	30.776	44.4%	2.25	0.417	55.6%	1.92	0.417	33.3%	1.03	3.361	47.2%	1.67
		50	5.8	7.712	91.7%	1.61	7.653	97.2%	4.03	1.025	88.9%	1.78	27.018	72.2%	3.53	0.361	77.8%	1.67	0.361	61.1%	1.08	3.639	77.8%	1.56
		75	8.8	6.785	91.7%	1.08	6.766	100.0%	2.44	1.167	100.0%	1.22	27.262	97.2%	2.81	0.361	88.9%	1.25	0.361	77.8%	1.0	3.083	94.4%	1.17
		100	11.3	6.177	97.2%	1.0	6.141	100.0%	1.42	1.289	97.2%	1.08	30.361	77.8%	2.08	0.361	97.2%	1.08	0.333	97.2%	1.03	3.083	97.2%	1.08
Average				6.870	80.14%	1.69	6.868	88.68%	3.04	0.929	47.22%	5.01	41.444	67.97%	4.65	0.234	73.21%	1.30	0.234	64.71%	1.13	0.767	77.34%	1.37

Table 2: Results comparing the state-of-the-art in goal recognition as planning for noisy, partial, and full observability.