Dataset	Metric	Llama3-8B	Claude-3.5	gpt-4o-mini	glm-4-air	qwen2.5	Mistral-8X7B	Mistral-7B
Shopping	Pt	–	-0.0044	0.0669	0.105	-0.0177	-0.0268	-0.0142
	Rt	–	0.0143	-0.0125	0.0702	0.0154	0.0507	0.0000
	At	–	0.0512	0.0664	-0.0602	0.1686	0.1066	-0.0145
	Ft	–	-0.006	-0.0075	-0.0061	-0.004	0.0105	-0.0014
	Acc (%)	26.3	32.4	**37.4**	**37.5**	37.2	31.7	5.2
	Δ Acc (%)	–	+6.2	**+11.2**	**+11.2**	+10.9	+5.4	-3.7
Math (Algebra)	Pt	–	0.029	0.078	0.052	0.067	0.009	-0.065
	Rt	–	0.176	0.038	0.065	0.150	-0.010	-0.014
	At	–	0.397	0.318	0.343	0.444	0.190	-0.051
	Ft	–	0.031	0.032	-0.004	0.015	-0.008	-0.005
	Acc (%)	21.6	84.4	67.2	68.8	**86.8**	39.2	8.0
	Δ Acc (%)	–	+62.8	+45.6	+47.2	**+65.2**	+17.6	-13.6
Math (Geometry)	Pt	–	0.056	0.034	0.054	0.074	0.006	-0.060
	Rt	–	0.077	0.040	0.018	0.064	0.018	0.011
	At	–	0.471	0.526	0.346	0.530	0.142	-0.038
	Ft	–	0.039	0.019	0.006	0.044	-0.022	-0.007
	Acc (%)	14.4	82.4	78.0	57.6	**86.4**	28.4	6.4
	Δ Acc (%)	–	+68.0	+63.6	+43.2	**+72.0**	+14.0	-8.0
ATP (Coq)	Pt	–	-0.011	0.028	0.013	0.008	0.010	0.011
	Rt	–	0.054	-0.026	0.016	0.024	0.051	0.007
	At	–	0.775	0.388	0.115	0.611	0.118	0.011
	Ft	–	0.011	0.017	0.031	0.021	0.006	-0.020
	Acc (%)	6.4	**96.4**	49.6	24.3	74.8	27.0	10.8
	Δ Acc (%)	–	**+90.0**	+43.2	+17.9	+68.4	+20.6	+4.4
ATP (Lean 4)	Pt	–	0.007	0.003	-0.006	-0.012	-0.000	0.011
	Rt	–	0.059	-0.010	0.001	0.012	-0.016	0.005
	At	–	0.667	0.404	0.193	0.474	0.061	0.005
	Ft	–	0.106	0.024	0.015	0.044	-0.018	0.007
	Acc (%)	2.7	**84.7**	39.6	23.4	57.7	6.3	8.1
	Δ Acc (%)	–	**+82.0**	+36.9	+20.7	+55.0	+3.6	+5.4
ATP (Isabelle)	Pt	–	0.024	0.027	0.005	0.050	0.060	-0.007
	Rt	–	0.046	-0.009	-0.002	0.047	0.021	0.009
	At	–	0.512	0.243	0.178	0.439	-0.064	-0.070
	Ft	–	0.076	0.018	0.014	0.036	-0.024	0.005
	Acc (%)	7.2	**74.8**	36.0	25.2	63.1	4.5	0.0
	Δ Acc (%)	–	**+67.6**	+28.8	+18.0	+55.9	-2.7	-7.2
Robot Cooperation	Pt	–	0.1140	0.0730	-0.0233	0.0894	-0.0054	-0.0142
	Rt	–	0.3879	0.1868	0.1140	0.2683	0.0324	0.0000
	At	–	0.2161	0.1224	0.0086	0.1614	0.0383	-0.0145
	Ft	–	0.0172	0.0005	-0.0133	0.0028	0.0035	-0.0014
	Reward (%)	8.9	**92.6**	54.3	17.4	72.6	17.1	5.2
	Δ Reward (%)	–	**+83.7**	+45.5	+8.6	+63.7	+8.2	-3.7