

# COMPLETE RESULT TABLES

## Nonlinear Inverse Reinforcement Learning with Gaussian Processes

Anonymous Author(s)

Affiliation

Address

email

In this supplement, we present the complete results of each experiment we conducted, including results for all prior algorithms and a second highway policy. Numerical scores correspond to the expected value difference metric as described in the paper. MMPBoost used decision trees of depth 4, and both log-linear and nonlinear variants of LEARCH were tested. The nonlinear variant used decision trees for discrete features and logistic regression for continuous features.

### 1 Objectworld with Discrete Features and Varying Example Counts

Examples	Training Environments						Transfer Environments					
	4	8	16	32	64	128	4	8	16	32	64	128
GPIRL	5.8	0.5	0.5	0.4	0.2	0.1	6.0	0.5	0.4	0.4	0.3	0.1
MaxEnt	30.8	25.7	14.8	12.3	10.1	8.6	28.8	25.8	19.2	15.5	12.9	11.7
MaxEnt Laplace	15.5	13.4	11.8	10.1	9.2	8.5	15.5	13.6	13.2	12.7	11.9	11.3
MWAL	34.4	34.5	33.5	33.2	33.0	33.0	31.8	32.5	31.4	31.8	31.7	31.5
Abbeel & Ng	34.9	34.4	32.6	33.4	32.4	32.7	32.8	32.4	30.8	31.1	31.9	32.0
MMP	34.4	34.3	33.4	33.2	32.8	33.1	31.8	32.6	31.5	31.9	31.7	31.6
MMPBoost	31.1	31.1	24.8	23.4	21.4	19.6	28.9	28.0	24.6	25.3	23.9	23.5
LEARCH	22.1	20.7	17.7	15.5	12.1	10.4	21.4	22.0	19.8	19.8	15.7	12.4
LEARCH Nonlin	27.3	25.5	26.6	25.4	18.6	19.7	27.4	25.1	26.3	23.4	19.5	18.7
FIRL	16.8	11.4	8.9	5.5	3.3	2.2	16.2	11.0	9.0	7.2	5.1	3.2

### 2 Objectworld with Continuous Features and Varying Example Counts

Examples	Training Environments						Transfer Environments					
	4	8	16	32	64	128	4	8	16	32	64	128
GPIRL	10.3	1.1	0.4	0.3	0.1	0.1	10.5	1.3	0.5	0.3	0.1	0.1
MaxEnt	19.5	14.7	16.3	14.8	13.3	12.2	18.0	13.9	15.7	15.1	14.2	14.3
MaxEnt Laplace	18.2	14.4	16.2	14.8	13.3	12.2	17.0	13.9	15.6	15.0	14.2	14.3
MWAL	29.1	25.1	26.2	26.8	26.6	25.5	26.4	24.3	25.1	24.9	24.6	23.4
Abbeel & Ng	29.6	30.5	32.0	31.6	31.4	30.8	26.5	27.7	28.8	29.0	28.8	28.1
MMP	29.9	25.3	26.5	27.7	26.3	25.4	26.9	24.6	25.7	25.2	24.8	23.9
LEARCH	28.2	26.5	26.2	30.9	23.3	19.6	24.5	22.2	24.9	32.6	22.4	17.9
LEARCH Nonlin	30.3	26.3	34.5	44.7	41.0	56.0	30.5	24.8	32.1	41.3	38.2	52.4

054  
055  
056  
057  
058  
059  
060  
061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072  
073  
074  
075  
076  
077  
078  
079  
080  
081  
082  
083  
084  
085  
086  
087  
088  
089  
090  
091  
092  
093  
094  
095  
096  
097  
098  
099  
100  
101  
102  
103  
104  
105  
106  
107

### 3 Objectworld with Discrete Features and Varying Numbers of Colors

Colors	Training Environments						Transfer Environments					
	2	4	6	8	10	12	2	4	6	8	10	12
GPIRL	0.4	0.5	0.6	0.5	0.7	1.3	0.4	0.5	0.7	0.6	0.9	1.5
MaxEnt	12.3	14.6	15.7	19.6	17.3	16.6	15.5	22.5	24.6	25.1	29.9	27.6
MaxEnt Laplace	10.1	10.8	11.4	11.5	10.2	9.4	12.7	16.9	17.6	17.4	19.7	21.8
MWAL	33.2	35.7	37.6	39.1	39.6	38.6	31.8	34.0	35.7	35.4	36.3	36.4
Abbeel & Ng	33.4	36.3	38.4	40.5	40.9	39.5	31.1	33.7	34.5	35.6	35.7	35.3
MMP	33.2	36.0	37.4	39.0	39.2	38.8	31.9	34.0	35.6	35.5	36.3	36.5
MMPBoost	23.1	28.2	31.5	33.4	31.8	33.4	24.5	29.7	33.6	33.8	33.4	34.9
LEARCH	15.5	15.7	21.9	21.8	19.0	21.3	19.8	20.2	24.3	28.2	27.6	31.1
LEARCH Nonlin	25.4	17.1	23.5	27.2	22.3	25.0	23.4	22.0	27.9	31.4	31.5	33.1
FIRL	5.5	6.0	6.4	5.0	4.3	5.3	7.2	9.0	11.0	6.5	5.9	8.4

### 4 Objectworld with Continuous Features and Varying Numbers of Colors

Colors	Training Environments						Transfer Environments					
	2	4	6	8	10	12	2	4	6	8	10	12
GPIRL	0.3	0.5	0.5	0.4	0.7	1.2	0.3	0.6	0.5	0.5	0.7	1.5
MaxEnt	14.8	13.7	15.5	16.1	14.7	18.4	15.1	18.6	19.3	18.8	21.3	24.1
MaxEnt Laplace	14.8	13.2	15.8	16.2	14.6	18.6	15.0	18.5	19.1	18.7	21.1	23.7
MWAL	26.8	32.4	39.2	37.8	38.7	41.3	24.9	32.9	37.9	40.0	37.7	39.0
Abbeel & Ng	31.6	35.0	42.3	38.6	40.4	40.8	29.0	35.1	39.4	40.5	39.1	40.2
MMP	27.7	33.1	39.9	37.8	39.5	42.9	25.2	33.2	38.2	40.6	38.1	39.1
LEARCH	30.9	19.9	30.9	30.9	28.2	32.2	32.6	20.7	29.2	30.4	28.6	34.8
LEARCH Nonlin	44.7	48.1	43.7	34.7	32.2	34.0	41.3	41.5	40.5	30.1	33.0	38.1

### 5 Police Highway with Synthetic Examples and Discrete Features

Examples	Training Environments						Transfer Environments					
	2	4	8	16	32	64	2	4	8	16	32	64
GPIRL	18.4	10.2	5.7	2.2	0.5	0.4	24.3	14.7	11.3	6.5	3.4	3.0
MaxEnt	30.0	20.0	10.4	5.5	3.0	2.8	43.9	40.2	35.6	38.3	38.9	47.7
MaxEnt Laplace	23.4	16.9	9.2	5.0	3.5	2.6	32.8	27.9	25.4	25.3	27.8	30.1
MWAL	51.2	48.1	45.8	38.8	37.1	30.0	61.4	59.7	56.4	54.7	55.8	55.8
Abbeel & Ng	52.3	49.9	45.1	40.3	37.2	31.3	61.6	59.8	56.9	55.6	57.4	57.8
MMP	54.5	52.0	47.0	40.0	39.7	36.6	61.9	59.7	56.5	55.4	54.8	53.6
MMPBoost	53.5	49.4	44.0	35.4	34.2	29.3	60.2	59.5	53.9	52.6	54.3	53.6
LEARCH	51.6	50.6	50.1	48.7	46.7	43.8	70.5	70.1	70.6	69.8	70.0	71.3
LEARCH Nonlin	57.9	54.7	56.1	50.2	46.6	40.8	71.3	71.7	71.1	70.2	70.6	70.5
FIRL	20.3	14.9	8.3	4.8	3.2	2.5	36.0	27.0	26.0	19.1	15.4	16.0

### 6 Police Highway with Synthetic Examples and Continuous Features

Examples	Training Environments						Transfer Environments					
	2	4	8	16	32	64	2	4	8	16	32	64
GPIRL	19.1	11.7	5.3	3.0	0.8	0.4	25.4	16.4	12.5	9.1	4.8	2.0
MaxEnt	45.6	46.8	43.3	44.6	44.8	45.8	55.0	56.5	52.4	52.2	52.2	51.7
MaxEnt Laplace	46.7	44.2	43.1	45.3	44.8	45.9	51.4	51.0	51.8	50.2	51.7	51.5
MWAL	56.7	52.5	51.5	51.2	51.3	52.0	60.4	58.6	57.3	59.5	58.3	57.9
Abbeel & Ng	56.7	51.2	52.0	52.8	51.0	50.1	60.8	58.3	57.8	59.3	59.6	61.5
MMP	57.6	52.8	52.0	51.9	49.8	46.5	61.2	58.4	57.9	59.5	59.5	61.5
LEARCH	51.6	50.3	46.0	47.9	44.9	47.6	59.1	60.3	56.6	53.6	54.5	55.3
LEARCH Nonlin	55.1	54.0	48.8	49.1	48.9	48.1	63.4	62.8	61.6	61.4	60.3	59.3

108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161

## 7 Police Highway with Human Examples and Discrete Features

Examples	Training Environments				Transfer Environments			
	2	4	8	16	2	4	8	16
GPIRL	11.1	7.1	5.4	4.5	16.5	14.2	12.9	8.8
MaxEnt	17.9	12.3	11.1	10.3	39.8	37.9	41.3	43.2
MaxEnt Laplace	14.2	9.0	7.1	6.6	30.1	28.0	27.7	28.8
MWAL	45.9	40.5	30.1	27.3	55.8	53.2	49.8	49.5
Abbeel & Ng	47.5	41.1	32.3	30.2	56.5	54.1	51.6	51.3
MMP	50.5	43.7	40.5	39.7	56.4	54.4	52.2	52.5
MMPBoost	34.8	25.3	21.4	22.5	47.5	42.3	38.7	39.1
LEARCH	49.2	46.7	44.7	44.9	70.2	70.6	70.9	70.5
LEARCH Nonlin	61.7	53.5	52.4	63.0	71.1	71.2	71.3	71.7
FIRL	16.6	10.1	8.6	7.0	35.4	21.1	23.3	17.8

## 8 Police Highway with Human Examples and Continuous Features

Examples	Training Environments				Transfer Environments			
	2	4	8	16	2	4	8	16
GPIRL	11.4	8.0	6.2	4.3	16.9	16.6	14.7	9.7
MaxEnt	48.2	44.0	46.2	44.6	57.8	55.7	55.7	56.0
MaxEnt Laplace	48.4	45.1	45.0	44.5	53.7	53.2	53.4	53.5
MWAL	56.0	50.5	50.8	47.6	59.4	57.4	60.2	57.3
Abbeel & Ng	57.0	48.4	49.7	46.3	59.5	55.5	59.9	58.5
MMP	58.0	50.1	50.2	48.4	59.6	55.4	59.5	58.7
LEARCH	50.9	48.7	47.0	49.2	59.3	55.8	56.6	55.2
LEARCH Nonlin	49.7	53.5	49.6	50.0	61.1	62.2	61.9	61.6

## 9 Lane Highway with Synthetic Examples and Discrete Features

Examples	Training Environments						Transfer Environments					
	2	4	8	16	32	64	2	4	8	16	32	64
GPIRL	25.3	9.2	3.8	0.6	0.2	0.1	34.4	16.8	12.2	5.3	0.9	0.7
MaxEnt	33.5	16.6	7.8	5.7	2.7	2.4	62.6	50.1	42.6	47.2	44.9	54.0
MaxEnt Laplace	31.1	16.9	12.6	7.1	5.0	3.8	43.5	32.1	32.9	33.7	34.3	36.9
MWAL	60.1	55.6	49.5	41.0	33.9	26.8	69.0	68.0	64.9	61.0	60.8	60.9
Abbeel & Ng	60.7	56.2	50.4	41.1	37.5	30.7	67.6	68.9	65.8	62.2	62.6	62.8
MMP	61.9	58.5	53.6	48.3	41.6	33.9	68.4	68.0	65.8	61.1	59.5	57.4
MMPBoost	58.6	53.3	47.8	38.3	30.8	22.4	67.1	64.4	63.6	58.6	59.5	54.8
LEARCH	61.8	60.4	55.3	51.9	50.7	47.1	88.2	87.1	87.9	86.5	88.1	87.3
LEARCH Nonlin	64.9	62.3	59.8	54.7	47.9	50.1	89.7	88.5	89.0	88.0	88.4	88.1
FIRL	23.2	13.3	10.7	5.6	3.1	2.7	46.7	34.5	32.8	26.0	24.5	23.5

## 10 Lane Highway with Synthetic Examples and Continuous Features

Examples	Training Environments						Transfer Environments					
	2	4	8	16	32	64	2	4	8	16	32	64
GPIRL	21.5	4.6	1.9	0.3	0.2	0.1	27.1	10.1	3.7	0.9	0.8	0.5
MaxEnt	46.9	38.6	35.7	33.5	33.5	33.4	74.1	64.2	58.4	62.0	58.7	58.4
MaxEnt Laplace	48.3	37.5	35.7	33.4	33.7	33.6	71.0	55.9	54.0	54.4	51.2	53.5
MWAL	64.3	65.5	61.8	60.6	60.8	60.1	73.0	76.0	73.3	73.4	72.5	72.6
Abbeel & Ng	63.8	64.7	63.0	58.8	57.4	53.7	73.4	73.7	73.2	70.5	72.4	70.2
MMP	65.9	65.4	63.0	60.0	55.9	50.6	73.1	73.9	72.9	70.2	72.1	69.3
LEARCH	62.1	66.1	56.2	46.2	47.0	43.7	69.2	67.0	62.2	57.5	53.6	55.1
LEARCH Nonlin	61.6	65.9	57.7	54.6	56.9	50.2	76.6	72.4	69.2	68.9	64.6	64.5

162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215

## 11 Lane Highway with Human Examples and Discrete Features

Examples	Training Environments				Transfer Environments			
	2	4	8	16	2	4	8	16
GPIRL	21.2	14.7	11.4	9.9	26.6	24.7	22.3	14.3
MaxEnt	26.3	22.3	19.7	17.5	52.3	50.7	50.1	50.8
MaxEnt Laplace	23.5	17.9	14.9	13.7	32.7	33.7	34.2	35.3
MWAL	52.2	48.0	40.4	36.8	59.1	57.6	53.5	51.4
Abbeel & Ng	55.3	48.4	41.5	37.9	60.5	58.5	55.3	52.8
MMP	58.1	56.1	49.3	45.5	60.5	59.2	55.9	53.7
MMPBoost	35.4	28.4	30.3	30.1	42.3	42.6	42.9	38.5
LEARCH	56.8	55.4	54.9	54.3	88.5	88.4	86.5	87.6
LEARCH Nonlin	67.9	64.8	68.0	65.4	89.5	89.2	88.3	88.3
FIRL	24.7	24.1	19.0	18.7	37.2	34.5	32.6	27.2

## 12 Lane Highway with Human Examples and Continuous Features

Examples	Training Environments				Transfer Environments			
	2	4	8	16	2	4	8	16
GPIRL	20.9	11.5	9.6	9.6	21.7	13.7	9.3	9.3
MaxEnt	43.3	43.5	43.7	42.5	68.0	63.5	61.7	63.1
MaxEnt Laplace	38.2	39.9	39.0	38.7	61.7	62.3	62.3	62.2
MWAL	61.9	61.5	57.0	51.8	75.6	75.7	72.1	69.8
Abbeel & Ng	64.7	60.1	57.6	55.2	76.8	75.1	74.6	71.7
MMP	65.1	62.0	60.4	56.3	76.8	74.9	73.9	71.7
LEARCH	57.8	56.0	54.6	49.5	71.3	66.4	66.8	61.6
LEARCH Nonlin	65.7	58.8	65.3	64.3	79.7	71.3	75.0	77.2