

gpt-4

Probabilistic Reasoning

0.8

0.9

Horizon Task

2.5

2.6

Restless Bandit

1.5

1.6

Instrumental Learning

1.5

1.6

Two Step Task

1.3

1.4

BART

0.4

0.5

claude-2

Probabilistic Reasoning

0.8

0.9

Horizon Task

1.8

1.9

Restless Bandit

0.3

0.4

Instrumental Learning

1.1

1.2

Two Step Task

0.4

0.5

BART

0.1

0.2

llama-2-7

Probabilistic Reasoning

0.4

0.5

Horizon Task

0.4

0.5

Restless Bandit

0.1

0.2

Instrumental Learning

0.1

0.2

Two Step Task

0.1

0.2

BART

0.1

0.2

Values (Normalized: Random=0, Human average=1)