Inverse Reinforcement Learning
A-Exam Presentation

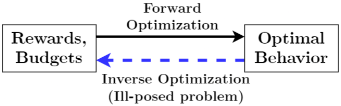
Kunal Pattanayak

Committee:
Vikram Krishnamurthy (Chair)
Jayadev Acharya
Siddhartha Banerjee

Introduction. Motivation and State-of-the-Art

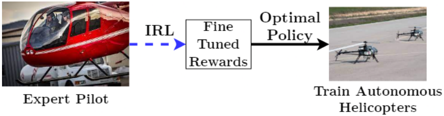
Problem: Consider a decision-making agent. Ground truth $x \rightarrow$ takes action a.
How to identify underlying strategy from behavior $p(a|x)$?

Ans: Inverse Reinforcement Learning (IRL)



Why IRL?

- **Autonomous navigation**: Learning from expert driver's actions [1]
- **Interpretable ML**: Understanding black-box classification behavior
- **Stealthy Radar Operation**: Extract adversary strategy, avoid detection



Lines of Work:

I. Traditional IRL in ML [2–4]:

- Markov Decision Process
- **Assumes** the existence of a reward that *rationalizes* agent actions

II. Behavioral/Micro- Economics [5–8] (Revealed Preference):

- Constrained Utility Maximization
- **Tests** for the existence of a rationalizing utility function (More fundamental)
- Set-valued estimation of utility function

Organization

- Revealed Preference (RP). Background and Notation

Contributions:

- **Part A:** Unifying Bayesian and non-Bayesian RP [9]
- **Part B:** Interpretable Deep Image Classification [10]
- **Part C:** Interpreting YouTube Commenting Behavior [11]
- **Part D:** Inverse Optimal Stopping [12, 13]

Revealed Preference (RP). Background and Notation

Classical RP (Single Agent) [5, 8]

Known: Sequence of budgets (probe) and consumption bundles (response) $\{g_{1:K}, \beta_{1:K}\}$, $g_k(\cdot) > 0$ and non-decreasing, $\beta_k \in \mathbb{R}_+^N$, $k = 1, 2, \ldots, K$

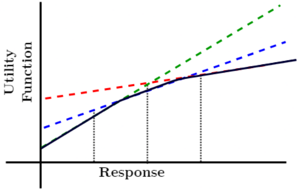
Aim: Test for budget constrained utility maximization. Estimate monotone utility function $u(\beta) > 0$ s.t.

$$\beta_k = \operatorname{argmax}_{\beta \in \mathbb{R}_+^N} u(\beta), \; g_k(\beta) \leq 0$$

Solution (Generalized Afriat's Thm. [8]):
Find positive reals u_k, λ_k s.t.

$$u_s - u_t - \lambda_t \; g_t'(\beta_s) \leq 0, \; \forall \; s, t \tag{1}$$

$$u(\beta) = \min_k \{u_k + \lambda_k g_k(\beta)\} \tag{2}$$



Bayesian RP (Multiple Agents) [7]

Known: Collection of agents $\mathcal{K} = \{1, 2, \ldots, K\}$.
Finite states $\mathcal{X}$, prior π_0, observations $\mathcal{Y}$, actions $\mathcal{A}$.
Agent k: Utility $U_k(x, a)$ (probe),
Observation likelihood $\alpha_k(y|x)$ (attention response).
Computes posterior $p_k(x|y)$ and takes action a.

Aim: Test for constrained Bayesian utility maximization (**UM**). Estimate **rational inattention** (RI) cost $C(\alpha)$ s.t.

$$\alpha_k = \operatorname{argmax}_\alpha \underbrace{\mathbb{E}_{\pi_0, \alpha}\{U_k(x, a^*(y))\}}_{J(\alpha, U_k)} - C(\alpha)$$

$$a^*(y) = \operatorname{argmax}_a \; \mathbb{E}\{U_k(x, a)|y\} \tag{3}$$

Existence (NIAS and NIAC inequalities [7]):
Find positive reals c_k s.t.

$$J(\alpha_t, U_t) - c_t \geq J(\alpha_s, U_t) - c_s \; \forall s, t \tag{4}$$

- Convex feasibility to identify utility maximization.
- Traditional IRL closely resembles NIAS inequality [7] - "Find rewards for which changing the observed policy is worse off for the agent".
- Bayesian RP is more fundamental - Does not assume the existence of C.

- Afriat's Theorem [5]: $g_k(\beta) = p_k'\beta - 1,\ p_k \in \mathbb{R}_+^M$ in (1).
- Central Idea in classical and Bayesian RP: Relative optimality suffices for global optimality.

Research Motivation:

- Piece-wise stitching of budgets to construct a utility function that rationalizes the data.
  Can it be done for Bayesian RP too? $\rightarrow$ Equivalence Result [9].
- Can the RP test be be used to understand complex black-box behavior? $\rightarrow$ [10] for Deep Image Classification, [11] for YouTube comments.
- Variation in responses due to varying probes **reveals** underlying strategy (utility).
  Extension of philosophy to stopping time problems $\rightarrow$ [12, 13]

Part A: Unifying Classical and Bayesian RP

Classical RP - 1967, Bayesian RP - 2015.
Identical Idea: Check for relative optimality.

Does there exist a formal equivalence?

Yes, but not obvious. Utility u is unknown in classical RP, and known in Bayesian RP.

Result 1. Classical RP test for **unknown** budgets $\{g(\beta) - \gamma_k \leq 0\}$, known utilities $\{u_k\}$.

$$\gamma_s - \gamma_t - \lambda_t(u_t(\beta_s) - u_t(\beta_t)) \leq 0$$
$$g(\beta) = \max_k \{\gamma_k + \lambda_k(u_k(\beta) - u_k(\beta_k))\} \quad (5)$$

Result 2. Bayesian RP test is **equivalent** to (5) on the Blackwell partial order for pmfs.

Key Idea. In classical RP, $u(x) \uparrow$ if $x \uparrow$ element-wise. Similarly, expected utility $J(\alpha, U) \uparrow$ if $\alpha \uparrow$ wrt Blackwell order ($\mathcal{B}$)[14]: Partial order on observation likelihoods.

$$\alpha \geq_{\mathcal{B}} \bar{\alpha} \implies \bar{\alpha} = \alpha Q, \; Q: \text{ row stochastic}$$

$\bar{\alpha}$ is obtained by stochastically garbling α, and hence, Blackwell dominated by α.

Parameter Mapping for Equivalence Result

Classical RP		Bayesian RP
Element-wise order	$\equiv$	**Blackwell order**
Time step k	$\equiv$	Agent index k
Consumption β_k	$\equiv$	Obs. Likelihood α_k
Budget $g(\beta) - \gamma_k$	$\equiv$	Cost $C(\alpha) - C(\alpha_k)$
Utility function $u_k(\cdot)$	$\equiv$	Exp. utility $J(\cdot, U_k)$

Result 3.

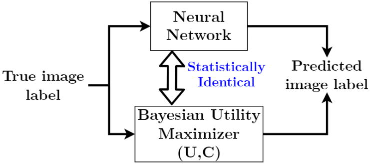
Enhancing [7]: Construction of a monotone (wrt Blackwell order) and convex cost C.

$$C(\alpha) = \max_{k \in \mathcal{K}} \{c_k + J(\alpha, U_k) - J(\alpha_k, U_k)\} \quad (6)$$

Above reconstruction follows the style of Afriat's Theorem and builds on existence conditions of [7].

Part B: Interpretable Deep Image Classification

Can neural networks' (NN) image classification be explained by Bayesian UM?



– Experiments on CIFAR-10 dataset, 200 trained NNs, 5 architectures

Main Idea.

1. Record classification performance of a trained NN by varying training parameters.
2. **Bayesian RP test for Interpretability**: Estimate **BOTH** utility and cost that rationalize NN dataset $\mathbb{D} = \{\pi_0, \{p_k(a|x), k = 1, 2, \ldots, K\}$

U - preference ordering over image classes,
C - Learning Cost (wrt training parameter).

Variable Map:

State: $X \sim \pi_0$ - true label. π_0 from CIFAR-10
Observation: $Y \sim \alpha(y|x)$ - accuracy of learned features
Action: $a = f(y)$ - predicted image label
Agent: Trained NN, $k \in \{1, \ldots, K\}$ indexes training parameter

Estimate:

Classification preference: $U_k(x, a)$
Cost of training: $C(p(a|x))$

Main Results.

1. Bayesian UM robustly fits deep image classification (dataset $\mathbb{D}$ passes Bayesian RP test with high margin).
2. Reconstructed U, C can predict NN performance without simulation (at least 94% accuracy).
3. Sparsity-enhanced version (fewer variables) of Bayesian RP test.

Robustness. *How well does NN dataset $\mathbb{D}$ pass the Bayesian RP test?*

– Vary training epochs

– Why Robustness?: Find the solution that passes Bayesian RP test with largest margin.

Robustness value $\mathcal{R}$: Distance of interior-most point from edge of feasible set.

Higher $\mathcal{R}$ $\Longrightarrow$ better fit to UM model.

$\mathcal{R} = \max_{\epsilon > 0} \epsilon, J(p_t, U_t) - c_t - J(p_s, U_t) + c_s \geq \epsilon$

Robustness results on NN dataset:

Aggregate classification performance of 20 NNs by varying **training epochs**.

Architecture	$\mathcal{R}$ $(\times 10^{-4})$
LeNet	37.97
AlexNet	40.60
VGG16	119.8
ResNet	132.3
Network-in-Network	149.1

Inference: NiN and ResNet architectures fit Bayesian UM model 4x better than less complex architectures.

Predictive Ability. *How well does interpretable model predict image classification performance?*

– Inject artificial Gaussian noise into CIFAR-10 training dataset and vary noise variance

– Use **sparsest solution** of Bayesian RP test

$\min \sum_{k=1}^{K} \|U_k\|_1, J(p_t, U_t) - c_t \geq J(p_s, U_t) - c_s$

Main Idea.

1. Estimate U for **new noise variance** by interpolating $U_{1:K}$ (from NN simulations for known noise variances).
2. Solve constrained Bayesian UM with utility U and reconstructed cost C.

Result: Predicted performance $\hat{p}(a|x)$. Compare against true performance $p(a|x)$.

Prediction Accuracy: For new noise variance.

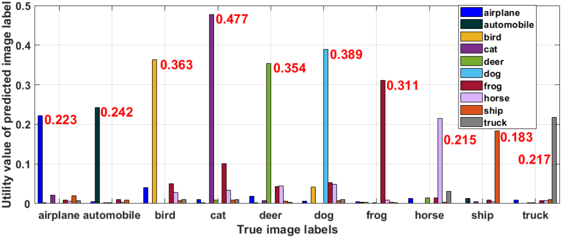
$\max_{x,a} |\hat{p}(a|x) - p(a|x)| = \mathbf{0.04}$

KL divergence between $p(a|x), \hat{p}(a|x)$:

- LeNet: 0.015
- AlexNet: 0.012
- VGG16: 0.016
- ResNet: 0.006
- NiN: 0.018.

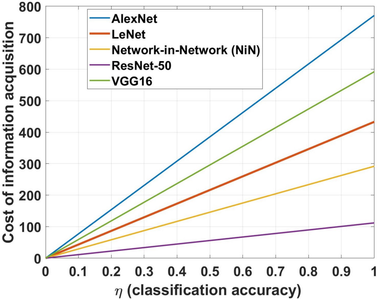
Low KL-divergence: Interpretable model is statistically similar to trained NN.

Insights: Bayesian RP on deep image classification



$U(x, a)$ for the ResNet architecture.

1. $U(x, a)$ peaks at $x = a$.
2. Higher preference given to correctly classifying a cat over a ship. (Disparity in $U(x, x)$ values over x be used for training data correction)
3. A dog is more likely to be classified as a horse than a ship. (Similarity between image labels)



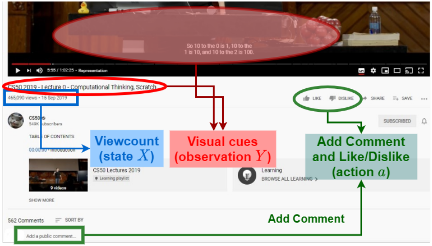
Reconstructed training cost C as a function of classification accuracy.

1. For a fixed classification accuracy, ResNet incurs least cost of training, and AlexNet incurs the most.
2. More complex and deeper networks $\Longrightarrow$ smaller cost of learning. (Prefer correct classification over cost minimization)

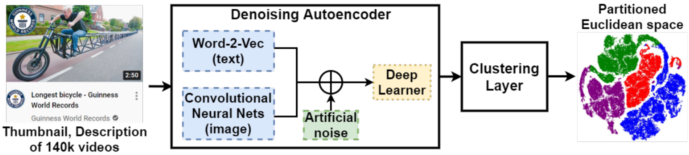
State: $X \in \{\text{low viewcount}, \text{high viewcount}\}$, $X \sim \pi_0$ (prior)
Observation: $Y \sim \alpha(y|x)$. Visual cues from thumbnail, video description
Attention function of viewer: α
Action: a. Comment count (high or low), sentiment (positive, negative or neutral)
Commenter's reputation: $U(x, a)$
Rational Inattention Cost: $C(\alpha)$
Agent: $k \in \{1, 2, \ldots, K\}$.
Video category, frame.

Aim: Given $\mathbb{D} = \{\pi_0, \{p_k(a|x)\}$ from K agents, estimate utility functions U_k and RI cost C that rationalize $\mathbb{D}$.

– **Reconstructed utility functions and cost**: Parametrize interpretable model for YouTube commenting behavior.
– Bayesian RP test (5): Pass test only if pass margin exceeds ϵ (user-defined).



Massive dataset: 140k videos, 25k channels.
Dimension Reduction. *How to group videos of specific topic with similar commenting behavior?*
(i) *User-centric*: Deep Clustering using thumbnail & description.
(ii) *Content-centric*: Video category.
Main Result: YouTube commenting is consistent with utility maximization.
Estimated utility can predict commenting behavior. (83% accuracy).



Autoencoder partitions YouTube dataset into 8 distinct clusters (agents).

How well does Bayesian Utility Maximization explain dataset?

General Rational Inattention cost: All 8 clusters pass test.

Renyi/Shannon mutual information cost: 2/8 clusters pass test.

Finer Granularity. 18 categories using topic (Gaming, Politics, Education, etc.)

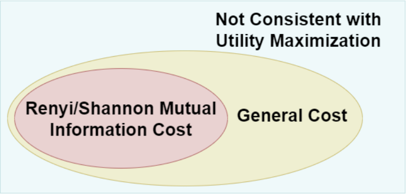
Result: 10 categories satisfy general cost, 2 categories satisfy Renyi/Shannon.

Key Insights:

- Clusters fail Renyi/Shannon by small margin $\implies$ model is robust.
- Utility (reputation) is substantially higher for popular videos.
- *Predictive Accuracy*. Given a video in a specific category, predicts comment count with 83% accuracy; sentiment with 80% accuracy.

Quantifying robustness:

- For categories that satisfy utility maximization, how far are they from failing.
- For categories that don't satisfy, how close are they to passing.



1. **For categories that fail general cost,** find min. perturbation to pass (ϵ_1).
Result: **Average** $\epsilon_1 = 1.2 \times 10^{-3}$.

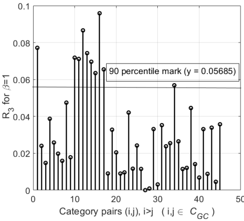
2. **For categories that satisfy general cost,** find max. perturbation to fail (ϵ_2).
Result: **Average** $\epsilon_2 = 7.01 \times 10^{-3}$.

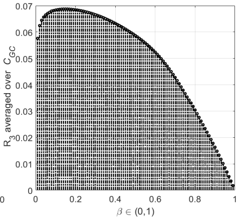
Conclusion: $\epsilon_1 / \epsilon_2 \approx 6$, hence categories are much closer to satisfying general cost than failing.

3. **For categories that satisfy general cost,** find min. perturbation to satisfy Renyi or Shannon cost.

Renyi Entropy: $H_\beta(p) = \sum_{i=1}^{n} \log(p_i^\beta)/(1-\beta)$.

Shannon cost: Renyi cost with $\beta \to 1$.





Part D: Inverse Optimal Stopping

Classical/Bayesian RP: Tests for static optimization.
– How to extend idea to detect sequential optimization, e.g. optimal stopping?
– Main Idea. Change parameters and observe change in policy (strategy)

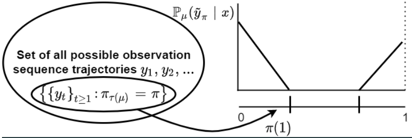
Decision Problems: $k \in \{1, 2, \dots, K\}$
Time step: $t = 1, 2, \dots$
State: $x \sim \pi_0,\ x \in \mathcal{X} = \{1, 2, \dots, X\}$
Observation: $y_t \in \mathcal{Y}, y_t \sim B(y_t|x)$
Action: $a \in \mathcal{A}$
Running cost: $\bar{c}_t = [c_t(1)\ c_t(x_2) \dots c_t(X)]$
Stationary Policy: $\mu_k : \Delta(\mathcal{X}) \to \mathcal{A} \cup \{\text{continue}\}$
Stopping Cost: $\bar{s}_k(a) = [s_k(1, a) \dots s_k(X, a)]$

Aim: Given $\{\pi_0, p_k(a|x), C(\mu_k)\}$, test if $\exists\ \bar{s}_k(a)$ s.t. μ_k minimizes expected cost, $k = 1, 2, \dots, K$:

$$\mu_k = \operatorname{argmin}_{\mu} \underbrace{\mathbb{E}_{\mu}\{\sum_{t=0}^{\tau(\mu)-1} c_t'\pi_t\}}_{C(\mu)} + \underbrace{\mathbb{E}_{\mu}\{\pi_\tau'\bar{s}_k(a)\}}_{J(\mu, \bar{s}_k)}$$

Challenges:
1. $C(\mu)$ does not have closed form expression.
2. μ_k is not known, only the surrogate action policy $p_k(a|x)$ is known.

How to tackle?
Likelihood fn. $p(y_{1:\tau}|x) \geq_{\mathcal{B}} p(a|x)$ and $J(\mu_j, \bar{s}_k) \geq J(p_j(a|x), \bar{s}_k)$. Equality when $j = k$.
– Can *at best* show relative optimality holds.



Main Results.
1. Necessary and sufficient conditions for relatively optimal stopping.
2. Examples: Optimal SHT, Bayesian Search
3. **Finite sample effects on (1)**:
Statistical tests for relative optimality, bounds on Type-I/II errors.

Conditions for Relatively Optimal Stopping:

Find positive reals $s_k(x,a)$ s.t. $\forall j,k$

$$(i)\sum_x p_k(x|a)(s_k(x,a)-s_k(x,b))\leq 0,\ \ a,b\in\mathcal{A}$$

$$(ii) J(p_k,\bar{s}_k)+C(\mu_k)\leq J(p_j,\bar{s}_k)+C(\mu_j) \qquad (7)$$

Above conditions test:

1. Optimal choice of stopping action
2. Relative optimality of policy μ_k

Ideas behind proof:

Sufficient statistic for policy μ_k: $p_{\mu_k}(y_{1:\tau}|x)$.

Necessity of (7): Uses Blackwell dominance.

$p_{\mu_k}(y_{1:\tau}|x)\geq_{\mathcal{B}} p_k(a|x)$

Sufficiency of (7): Since $\mathcal{Y}$ is unknown, assume $|\mathcal{Y}|=|\mathcal{A}|$, μ_k: injective map from $\mathcal{Y}$ to $\mathcal{A}$.

Relating optimal stopping to Bayesian UM:

$J(\mu,s)$: Only depends on stopping posterior distribution $p_\mu(x|y_{1:\tau})$. Hence,

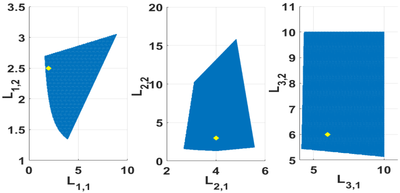
$p_\mu(y_{1:\tau}|x)\rightarrow$ attention $\alpha(y|x)$ in Bayesian RP.

$C(\mu)\rightarrow$ attention cost in Bayesian RP.

$J(\mu,\bar{s})\rightarrow -$ve of expected utility in Bayesian RP.

Example. Inverse SHT: Stopping time problem with structure. $\mathcal{X}=\mathcal{A}=\{1,2\}, s(x,x)=0$, $C(\mu)=\mathbb{E}_\mu\{\tau(\mu)\}$.

Simulation Results: 3 decision problems



— True stopping costs (yellow points) lie in the feasible set generated by (7).

— **Lower bounding expected stopping time**:

Given $p(a|x)$ for some policy μ, $\mathbb{E}_\mu(\tau)$ can be lower bounded via (7):

$\mathbb{E}_\mu\{\tau\}\geq$

$\min_{\bar{s}_{1:K}}\max_k \mathbb{E}_\mu\{\tau\}+J(\mu_k,\bar{s}_k)-J(p,\bar{s}_k)$,

where $\bar{s}_{1:K}\in$ feasible set (blue region).

(Simulation free aproximation)

Finite Sample Effects on Detecting Relative Optimality (7).

Consider Inverse SHT. Given empirical dataset $\widehat{\mathbb{D}} = \{\pi_0, \hat{p}(a|x), \widehat{\mathbb{E}_{\mu}\{\tau\}}\}$
$\widehat{\mathbb{D}}$ computed using $L_k \leq \infty$ samples for k^{th} decision problem. Denote $\mathbb{L} = \{L_k\}$.

Plug-in Test for relatively optimal stopping:

$$\sum_x \hat{p}_k(x|a)(s_k(x,a) - s_k(x,b)) \leq 0, \;\; a, b \in \mathcal{A}$$

$$J(\hat{p}_k, \bar{s}_k) + \hat{C}(\mu_k) \leq J(\hat{p}_j, \bar{s}_k) + \hat{C}(\mu_j) \qquad (8)$$

How accurate is the plug-in test (8)?

Events H_0, H_1: $\widehat{\mathbb{D}}$ generated and not generated, resp. , by relatively optimal agent policies $\{\mu_k\}$.
Hypothesis Test: Declare H_0 if (8) is feasible, otherwise H_1.

Finite Sample Result.

Bounds on Type-I/II errors of Hyp. Test:

$P(H_1|H_0) \leq \theta_{1,0}(\widehat{\mathbb{D}}, \mathbb{L}) \exp\{-\phi_{1,0}(\widehat{\mathbb{D}}, \mathbb{L})\}$, and
$P(H_0|H_1) \leq \theta_{0,1}(\widehat{\mathbb{D}}, \mathbb{L}) \exp\{-\phi_{0,1}(\widehat{\mathbb{D}}, \mathbb{L})\}$,

where $\theta_{0,1}(\cdot), \theta_{1,0}(\cdot), \phi_{0,1}(\cdot), \phi_{1,0}(\cdot) \in \mathbb{R}_+$ decrease with increasing sample size $\mathbb{L}$.

Outline of proof: Finite sample result

Pmfs $\hat{p}_k(a|x)$: Dvoretzky-Kiefer-Wolfowitz (DKW) inequality to bound error between pmfs: $\mathbb{P}(\max_a |p_k(a|x) - \hat{p}_k(a|x)| \geq \epsilon) \leq \delta_k(\epsilon)$

Empirical avg. stopping time $\widehat{\mathbb{E}_{\mu_k}\{\tau\}}$: Assume $\tau(\mu_k) \leq \tau_{\max}\ \forall\ k$ a.s. , Hoeffding's inequality to bound error from true mean:
$\mathbb{P}(|\widehat{\mathbb{E}_{\mu_k}\{\tau\}} - \mathbb{E}_{\mu_k}\{\tau\}| \geq \varepsilon) \leq \gamma_k(\tau_{\max}, \epsilon)$

Union bound: Combine DKW and Hoeffding bounds to get error bound between $\widehat{\mathbb{D}}$ and $\mathbb{D}$:

$$\mathbb{P}(|\widehat{\mathbb{D}} - \mathbb{D}| \geq \epsilon) \leq \kappa(\epsilon)$$

Compute minimum perturbation $\epsilon(\widehat{\mathbb{D}})$ such that $\widehat{\mathbb{D}} + \epsilon(\widehat{\mathbb{D}})$ fails (7), **set $\epsilon(\hat{\mathbb{D}}) \to \epsilon$ in union bound** to get Type-I error bound.

Intuition: If $\widehat{\mathbb{D}} + \epsilon(\widehat{\mathbb{D}})$ fails (8), then all datasets within $\epsilon(\widehat{\mathbb{D}})$ ball PASS the test (7).
Type-I error: Probability that true dataset lies **outside** the $\epsilon(\widehat{\mathbb{D}})$ ball.

Current Research

1. Deep Bayesian Revealed Preference: *Feature engineering for richer state space representation of real-world data*
2. Inverse Controlled Sensing: *How to detect if a sensing agent optimally switches between sensing modes based on target measurements?*
3. Inverse-Inverse Reinforcement Learning: *How to mask agent strategy? Optimal stealth-performance trade-off*
4. Structural Results: *How to exploit problem structure to reduce computation complexity of IRL conditions? Does it suffice to check relative optimality of only few pairs of agents?*

Thank You!

References

- Pieter Abbeel and Andrew Y Ng. “Apprenticeship learning via inverse reinforcement learning”. In: *Proceedings of the twenty-first international conference on Machine learning*. 2004, p. 1.
- A. Y. Ng, S. J. Russell, et al. “Algorithms for inverse reinforcement learning.”. In: *Icml*. Vol. 1. 2000, p. 2.
- Brian D Ziebart et al. “Maximum entropy inverse reinforcement learning.”. In: *Aaai*. Vol. 8. Chicago, IL, USA. 2008, pp. 1433–1438.
- Deepak Ramachandran and Eyal Amir. “Bayesian Inverse Reinforcement Learning.”. In: *IJCAI*. Vol. 7. 2007, pp. 2586–2591.
- S. N. Afriat. “The construction of utility functions from expenditure data”. In: *International economic review* 8.1 (1967), pp. 67–77.

H. R. Varian. "Non-parametric analysis of optimizing behavior with measurement error". In: *Journal of Econometrics* 30.1-2 (1985), pp. 445–458.

A. Caplin and M. Dean. "Revealed preference, rational inattention, and costly information acquisition". In: *The American Economic Review* 105.7 (2015), pp. 2183–2203.

Francoise Forges and Enrico Minelli. "Afriat's theorem for general budget sets". In: *Journal of Economic Theory* 144.1 (2009), pp. 135–145.

Kunal Pattanayak and Vikram Krishnamurthy. "Unifying Classical and Bayesian Revealed Preference". In: *arXiv preprint arXiv:2106.14486* (2021).

Kunal Pattanayak and Vikram Krishnamurthy. "Behavioral Economics Approach to Interpretable Deep Image Classification. Rationally Inattentive Utility Maximization

Explains Deep Image Classification". In: *arXiv preprint arXiv:2102.04594* (2021).

William Hoiles, Vikram Krishnamurthy, and Kunal Pattanayak. "Rationally Inattentive Inverse Reinforcement Learning Explains YouTube Commenting Behavior.". In: *J. Mach. Learn. Res.* 21.170 (2020), pp. 1–39.

Vikram Krishnamurthy and Kunal Pattanayak. "Necessary and Sufficient Conditions for Inverse Reinforcement Learning of Bayesian Stopping Time Problems". In: *arXiv preprint arXiv:2007.03481* (2020).

Kunal Pattanayak, Vikram Krishnamurthy, and Erik Blasch. "Inverse Sequential Hypothesis Testing". In: *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*. IEEE. 2020, pp. 1–7.

David Blackwell. "Equivalent comparisons of experiments". In: *The annals of mathematical statistics* (1953), pp. 265–272.