Research Statement

Kunal Pattanayak: Email | Google Scholar
Department of Electrical and Computer Engineering, Cornell University Ithaca, NY 14850, USA.

Keywords: Revealed Preference, Inverse Reinforcement Learning, Adversarial Machine Learning

1. Research Context

My research lies in the intersection of theoretical machine learning and information economics. More specifically, my research has primarily addressed problems in inverse optimization: *By observing the actions of a possibly expert demonstrator, how to identify if the actions are consistent with the optimal execution of a plan? If so, how to reconstruct the implicit plan of the demonstrator?* This general question has been a subject of interest in machine learning, economics and control theory for several decades. In the machine learning literature, such problems are addressed under the general theme of imitation learning (IL) (Ho and Ermon, 2016) and inverse reinforcement learning (IRL) with notable works like Ng et al. (2000); Ramachandran and Amir (2007); Ziebart et al. (2008); Choi and Kim (2011). In the economics literature, these problems are studied under revealed preference in microeconomics with notable works like Afriat (1967); Richter (1966); Diewert (1973); Varian (2006), and revealed rational inattention in information economics (Caplin and Dean, 2013, 2015; Caplin et al., 2017; Maćkowiak et al., 2021).

2. Research Contributions

In this section, I provide an overview of my research contributions that can be broadly segregated into two themes: (1) Learning from demonstration, and (2) Adversarial machine learning.

2.1 Learning from demonstration: Inverse Reinforcement Learning and Revealed Preference

Bayesian stopping problems are a special case of partially observed Markov decision processes (POMDPs) and frequently arise in modeling human decision-making.[1] In Pattanayak and Krishnamurthy (2020), we develop economics-based feasibility conditions for IRL of Bayesian stopping problems. Compared to existing IRL methods for POMDPs, our proposed approach *does not* require the model dynamics of the decision maker, nor does it require solving a POMDP. This feature is critical in practical settings such as online multimedia, healthcare records etc. where the decision maker's model dynamics are obscured from the inverse learner. In this paper, we also provide performance guarantees for our IRL algorithms in terms of sample complexity bounds. This work highlights how imposing additional structure on the underlying model dynamics allows the inverse learner to reconstruct agent utilities simply by observing its final actions without access to privileged information such as model dynamics. Our recent work Pattanayak et al. (2022a) shows how a priori knowledge of the parameter family of the decision maker's underlying utility function can speed up computation by a factor of about 10x.

Our applied works highlight the usefulness of our proposed IRL algorithm for constructing generative predictive modelling of black-box decision systems. In Hoiles et al. (2020), we show how our IRL algorithms can user engagement in YouTube videos with high accuracy. Our IRL methodology is purely data-centric and model-agnostic - without knowledge of decision maker's model dynamics, our proposed IRL algorithm reconstructs the best fitting Bayesian stopping time model that fits the data. Such model-agnostic inverse optimization algorithms lend themselves naturally to interpretable machine learning. In Pattanayak and Krishnamurthy (2021a), we show how our economics-based IRL approach approximates

1. The seminal work of Sims (2003) shows human attention is equivalent to a capacity-constrained communication channel. In other words, humans are more attentive for decisions with more lucrative payoffs and vice versa.

.

black-box neural networks trained for image classification by a human decision maker with an attention cost and a correct classification reward. We show the reconstructed human decision model accurately (low statistical distance) replicates high-level behavior of the neural network. Our interpretability result shares similar flavor to LIME (Ribeiro et al., 2016) and SHAP (Lundberg and Lee, 2017) in that our interpretability model space is the space of constrained utility maximization models parameterized by an attention cost and classification reward. However, our interpretability approach is *global* instead of *local*. We have a single interpretable model for all inputs instead of local explanations.

I now briefly discourse on a contribution in economics theory. In Pattanayak and Krishnamurthy (2021b), we unify two lines of work in economics theory, namely, identifying rational inattention in information economics (Caplin and Dean, 2015) and classical revealed preference in consumer and behavioral economics. Specifically, we show the well-known NIAC condition for identifying rational inattention is a Bayesian acyclic analog of the famous GARP condition in the celebrated Afriat's theorem for identifying utility maximization behavior, acyclic since the 'marginal utilities' in rational inattention are assumed constant across decision problems. The unification result subtly rests on the ordering rule of outcomes in the Bayesian and non-Bayesian utility maximization problems. In the classical revealed preference setup, consumption good vector 1 is said to be 'better' than vector 2 if vector 1 *element-wise dominates* vector 2 and implicitly implies yields a higher utility than vector 2. In the Bayesian decision framework (Caplin and Dean, 2015), the decision maker's *information structure*, namely, a probabilistic mapping from the ground truth to a set of perceived signals, is the analog of consumption good in revealed preference. We show that the *Blackwell partial order* (Blackwell, 1953) preserves monotonicity of the expected utility and the rational inattention cost of the decision maker, and hence, is a viable partial order to rank the decision maker's information structure.

This equivalence map from non-Bayesian decision making to Bayesian decision-making facilitates formulating the acyclic condition of NIAC (Caplin and Dean, 2015) as an Afriat inequalities-type feasibility test. Due to this equivalence result, one has a novel *Afriat-type reconstruction* of the decision maker's rational inattention cost in terms of its expected utility, and is monotone in the Blackwell partial order. This unification result is also of significance to algorithmic game theory. Hoiles et al. (2016) show that identifying play from the Nash equilibria of a potential game is equivalent to testing for the feasibility of Afriat's inequalities. Due to the equivalence result, one can now extend the above result for identifying play from Nash equilibria of *Bayesian* potential games by exploiting results from rationally inattention identification theory.

2.2 Adversarial Machine Learning for Mitigating Eavesdroppers

The previous section focused on identifying optimal decision making by observing the behavior of a decision maker. In this section, let us shift the spotlight to the forward optimizer, and assume the inverse reinforcement learner is adversarial.

In Pattanayak et al. (2022b), we address the question: how can a decision maker appear deceptively naive and mask its optimal strategy from an adversarial learner that is eavesdropping on its responses? In these works, we assume the adversarial learner uses revealed preference-based IRL techniques Afriat (1967); Caplin and Dean (2015) to reconstruct the decision maker's utility.[2] Mitigating set-valued adversarial IRL algorithms is considerably more challenging compared to point-valued IRL. We devise a *covert optimization* scheme for the decision maker for *masking* its true utility/plan from the adversarial learner. The decision maker deliberately optimizes a perturbed utility function and ensures its true utility function almost fails the revealed preference test with respect to its sub-optimal response. A smaller feasibility margin implies greater mitigation of the adversarial learner, and also, a larger perturbation in

2. In Pattanayak and Krishnamurthy (2020), we discuss how machine learning based IRL methods are equivalent to revealed preference methods in the economics literature.

its true utility function. The decision maker trades-off between minimizing the deliberate perturbation in its utility, and maximizing adversarial confusion.

In Pattanayak et al. (2022d), we build on our previous result and assume a noisy setting where the adversary has noisy measurements of the decision maker's responses and uses a revealed preference-based statistical hypothesis test for detecting optimality. In this scenario, the decision maker's adversarial mitigation objective changes from minimizing feasibility margin to maximizing the Type-I error probabilitiy of the hypothesis test. That is, maximize the probability the test detects the utility-maximizing decision maker as *not a utility maximizer*. Since the objective now has a probability term that can only be estimated empirically, the decision maker's perturbed utility is computed via a stochastic gradient algorithm. Finally, in Pattanayak et al. (2020), we provide performance guarantees of our adversarial mitigation scheme in terms of a finite sample complexity result. Specifically, we analytically compute the minimum number of time steps $\epsilon(\delta)$ the adversarial mitigation scheme needs to be executed for to ensure the error probability lies below δ. Our main observation is that in a zero-mean light-tailed noise setting, our adversarial mitigation algorithm will never fail as long as it is executed for infinitely many time steps. This can be interpreted as a law of large numbers result for privacy-preserving methods.

Although the adversarial mitigation algorithms developed in my research pertain to robust cognitive radar design in the field of electronic warfare, the philosophy translates seamlessly to privacy-preserving strategies such as differential privacy in machine learning. In an application-agnostic sense, by formalizing the trade-offs between performance and adversarial mitigation, our devised algorithms provide a principled approach to encode risk-averseness in decision systems in the face of adversarial leakage threats.

3. Research Plans

Based on my research contributions in Sec. 2, I now briefly outline my research plan that I hope to achieve in a time span of two years.

3.1 Online inverse optimization: Detecting changes in agent behavior on the fly

IRL and revealed preference typically belong to batch-wise inference methods. That is, one needs to aggregate a dataset and reconstruct the decision costs that rationalizes the dataset in an *offline* manner.[3] However, time-critical applications such as autonomous driving, medical robotics and autonomous search and rescue operations do not have the luxury of offline inference. I am keen on devising online IRL methods for such applications that achieve two critical objectives: (1) Compute set-valued estimates of a possibly evolving utility function with provable confidence bounds with computation constraints, and (2) Satisfy conditions for observed decision optimality with high probability.
Note that one cannot simultaneously do well on both objectives, hence a trade-off scheme needs to be outlined based on mission urgency. One naive approach towards achieving online revealed preference is to adopt a rolling window approach where the inverse learner only uses the last K data points to estimate the utility function. However, there is ample scope for improvement. On a related note, an essential sub-task for this research question will be to identify representative points, for example, the max-margin point in the feasible set of utilities that can be tracked easily over time and updated in an online fashion. In terms of intellectual merit, this research task will use mathematical results such as the Euler-Lagrange equation from variational calculus for semi-infinite programming, Bernstein-von Mises theorem for asymptotic statistics in Bayesian decision-making, and conformal prediction methods for provable online IRL performance guarantees.

3. Indeed, there have been preliminary works such as Aprem and Krishnamurthy (2016) for detecting jump changes in utility in multimedia platforms. However, the proposed detection procedure is yet again offline and not in real-time.

3.2 IRL and revealed preference under privacy constraints

IRL and revealed preference methods are generally sensitive to noise in dataset quality.[4] In privacy-sensitive applications such as online multimedia, recommender systems, an analyst typically has access to semi-anonymized data for performing IRL to understanding user behavior that generated the dataset where user-sensitive data attributes are abstracted away or obscured by a layer of additive Laplacian noise. From a purely inference viewpoint, the more personalized features one has at their disposal, the more precise inference on can make about the user's behavior/intent model.
I am keen on analyzing and computing performance bounds for IRL on semi-anonymized datasets. Since IRL and privacy-preserving methods can be viewed as opposing forces, this research task requires deep understanding of the percolative effects of methods like differential privacy on IRL inference. The research sub-tasks under this theme can be enumerated as follows:
(1) Investigate the application of generative adversarial networks (GANs) and variational autoencoders (VAEs) for recovering informative feature embeddings from semi-anonymized data for IRL.
(2) Provide information-theoretic identifiability bounds for IRL under privacy constraints.
(3) Devise novel persistence of excitation approaches that generate diverse yet semi-anonymized data that 'reveal' high-level behavioral changes in the user in spite of abstracting away user-sensitive features for boosting precision of privacy-constrained IRL algorithms. For example, estimating behavior variation with a target parameter via A/B testing.

References

Sydney N Afriat. The construction of utility functions from expenditure data. *International economic review*, 8(1):67–77, 1967.

Anup Aprem and Vikram Krishnamurthy. Utility change point detection in online social media: A revealed preference framework. *IEEE Transactions on Signal Processing*, 65(7):1869–1880, 2016.

David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, pages 265–272, 1953.

Andrew Caplin and Mark Dean. Behavioral implications of rational inattention with shannon entropy. Technical report, National Bureau of Economic Research, 2013.

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. Technical report, National Bureau of Economic Research, 2017.

JD Choi and Kee-Eung Kim. Inverse reinforcement learning in partially observable environments. *Journal of Machine Learning Research*, 12:691–730, 2011.

W Erwin Diewert. Afriat and revealed preference theory. *The Review of Economic Studies*, 40(3):419–425, 1973.

Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. *Advances in neural information processing systems*, 29, 2016.

William Hoiles, Vikram Krishnamurthy, and Anup Aprem. Pac algorithms for detecting nash equilibrium play in social networks: from twitter to energy markets. *IEEE Access*, 4:8147–8161, 2016.

4. Krishnamurthy and Hoiles (2012) and Pattanayak et al. (2022c) design statistical tests for detecting decision optimality with bounded Type-I error probability that can at best withstand additive noise models with known first moments.

William Hoiles, Vikram Krishnamurthy, and Kunal Pattanayak. Rationally inattentive inverse reinforcement learning explains youtube commenting behavior. *Journal of Machine Learning Research*, 21:1–39, 2020.

Vikram Krishnamurthy and William Hoiles. Afriat's test for detecting malicious agents. *IEEE Signal Processing Letters*, 19(12):801–804, 2012.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

Bartosz Maćkowiak, Filip Matějka, and Mirko Wiederholt. Rational inattention: A review. 2021.

Andrew Y Ng, Stuart Russell, et al. Algorithms for inverse reinforcement learning. In *Icml*, volume 1, page 2, 2000.

Kunal Pattanayak and Vikram Krishnamurthy. Necessary and sufficient conditions for inverse reinforcement learning of bayesian stopping time problems. *Accepted to JMLR with minor revision (arXiv preprint arXiv:2007.03481)*, 2020.

Kunal Pattanayak and Vikram Krishnamurthy. Rationally inattentive utility maximization for interpretable deep image classification. *arXiv preprint arXiv:2102.04594*, 2021a.

Kunal Pattanayak and Vikram Krishnamurthy. Unifying classical and bayesian revealed preference. *arXiv preprint arXiv:2106.14486*, 2021b.

Kunal Pattanayak, Vikram Krishnamurthy, and Erik Blasch. Inverse sequential hypothesis testing. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–7. IEEE, 2020.

Kunal Pattanayak, Shashwat Jain, Vikram Krishnamurthy, and Chris Berry. Adaptive eccm for mitigating smart jammers. *arXiv preprint arXiv:2212.02002*, 2022a.

Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry. How can a cognitive radar mask its cognition? In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5897–5901. IEEE, 2022b.

Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry. How can a radar mask its cognition? *arXiv preprint arXiv:2210.11444*, 2022c.

Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry. Meta-cognition. an inverse-inverse reinforcement learning approach for cognitive radars. *arXiv preprint arXiv:2205.01794*, 2022d.

Deepak Ramachandran and Eyal Amir. Bayesian inverse reinforcement learning. In *IJCAI*, volume 7, pages 2586–2591, 2007.

Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. Model-agnostic interpretability of machine learning. *arXiv preprint arXiv:1606.05386*, 2016.

Marcel K Richter. Revealed preference theory. *Econometrica: Journal of the Econometric Society*, pages 635–645, 1966.

Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

Hal R Varian. Revealed preference. *Samuelsonian economics and the twenty-first century*, pages 99–115, 2006.

Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, Anind K Dey, et al. Maximum entropy inverse reinforcement learning. In *Aaai*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.