Research Statement: Clayton Thomas

I am a theorist studying mechanism and market design with an emphasis on transparency and interpretability. In this research statement, I first discuss my interdisciplinary research agenda on explaining matching mechanisms, including the economic theory paper [15] (reject and resubmit at the AER), its empirical companion [18] (in review at Econometrica), and the related computer science paper [17] (published in STOC 2024). Second, I discuss my other papers and future plans.

Strategyproofness-Exposing Mechanism Descriptions

Strategyproofness is a celebrated property of economic mechanisms that is crucial for ease of participation and equity across levels of strategic sophistication. Unfortunately, growing evidence suggests that participants frequently do not *understand* strategyproofness [20, 26].[1] To improve understanding, in [15, 18] we propose *menu descriptions*, which explain the outcome of the mechanism while making strategyproofness hold via a simple argument. Specifically, menu descriptions present a mechanism to player i using the following outline:

Step (1) uses only other players' reports to describe i's *menu* of potential outcomes.

Step (2) uses player i's report to select i's favorite outcome from her menu.

Our first main result in our theory paper [15] provides a simple and streamlined menu description of the classic stable matching mechanism, Deferred Acceptance (DA). We prove that student i's menu in (student-proposing) DA is all institutions that prefer i to their outcome in *school-proposing* DA excluding student i. This yields an intuitive menu description of DA; in contrast, prior to our work, it was not clear how to construct DA's menu, except via a trivial solution involving running the traditional description many times to separately check if different institutions are on i's menu. Our other main results in [15] investigate descriptions of the other canonical matching mechanism Top Trading Cycles (TTC), and characterize simple menu descriptions using algorithmic tools from computer science, providing a very complete picture of the concepts introduced.[2]

In our companion experimental paper [18], we test this new description of DA. We find that, like DA's traditional description, many participants can (using a novel GUI we developed) learn our new menu description and calculate its outcomes. Additionally, we find that in a suitably stripped-down form, a menu description conveys strategyproofness quite well (indeed, even outperforming a textbook description of strategyproofness), highlighting its potential for real-world use.

Our closely-related computer science paper [17] delves into many complexity questions inspired by our theory paper [15]. For one example, we show that while stability, and thus the outcome of DA, can be explained using one "cutoff" per school, explaining the outcome of TTC

[1] Prior work—including my paper [31]—has attempted to better convey strategyproofness by changing the interactive protocol used to implement the mechanism, with a particular focus on obviously strategyproof (OSP) mechanisms [23]. This literature finds that the set of OSP implementations is extremely limited, motivating other approaches.

[2] The working paper version [16] also investigates an extension of our theoretical results for auctions, and runs an experiment in a basic auction and voting setting.

requires a cutoff per *pair* of schools (formalizing the contrast between TTC and DA suggested by [5, 22]). Along with our other results, we holistically find that explaining the outcome of TTC is more complex than explaining the outcome of DA, corroborating prior concerns on TTC's complexity (in spite of the simplicity of TTC's strategyproofness which we identified in [15]).

Significant open directions remain regarding how to best explain mechanisms and algorithms to participants. Building on my work, future directions include explaining the strategyproofness of other mechanisms, or exploring equilibria beyond truthful ones. More broadly, one could develop protocols to explain properties like fairness, optimality, privacy, or accuracy, with approaches tailored to the specific settings (e.g., using the promising axiomatic explanation framework investigated for voting rules in [25, 24]). Additionally, the principled approach of my research may offer valuable insights for investigating explainability in other fields, such as AI.

Other Topics and Future Directions

Beyond explainability, my research contributes to a range of topics within economic design and theoretical computer science. I particularly value works that approach problems in principled but innovative new ways.

Auction Design for Blockchains. In [14], we study a platform (namely, a blockchain developer) designing a mechanism that an untrusted third party (namely, a blockchain miner) must actually implement. We point out that state-of-the-art protocols are vulnerable to a novel attack that—in an interesting contrast to much of the literature on untrusted auctioneers (e.g., Akbarpour and Li [1] and Roughgarden [29])—hinges around the ability of the miner to commit to a fixed strategy. Besides making an important contribution to the transaction fee mechanism literature, our framework may have implications for mechanism design questions where a platform (separate from the seller) takes a cut of the revenue.

Communication Complexity and Implementation Theory. A recurring topic in my work is communication complexity, a powerful theoretical model from computer science that takes on literal significance in mechanism design, where agents must communicate their preferences and cannot send prohibitively complex messages. In [30], we show that there can be a significant difference between what can be implemented in a dominant strategy equilibrium versus an ex-post Nash equilibrium when the mechanism is constrained to use a realistic amount of communication. Our work is the first to show that even within variants of what is traditionally called "strategyproof," the precise solution concept used can dramatically affect complexity.

In [28], we prove a new style of impossibility result in algorithmic mechanism design for which previous techniques were provably insufficient. In [8], we study a class of mechanisms related to implementation in undominated strategies, and give a novel reason why economically-reasonable mechanisms may be possible even in the presence of computational impossibility theorems.

Efficiency versus Stability. In [32], I study an important new generalization of stability known as priority-neutrality, introduced by Phil Reny in [27], which allows for (student-side) Pareto improvements to DA. I resolve an open question of [27], and show that the set of priority-neutral matchings is complex in an important way that stable matchings are simple.[3] My work gives a precise barrier which must be overcome by more tractable approaches to balance efficiency and fairness, and thus may inspire new solutions.

I also study the expected behavior of matching markets, and thus characterize the degree to which demanding stability harms students' welfare. In [7], we give simpler proofs of the main result of Ashlagi, Kanoria, and Leshno [4]—that in unbalanced matching markets, the "short side" is at a large disadvantage.[4] In [3], we study stable matchings under a novel distribution of preferences capturing the effect of ex-ante heterogeneous quality of different agents.

Additional Future Directions. I am also interested in many other areas of economics and computation, such as decision theory and information design.[5] Theoretical computer science provides powerful tools for modeling decision making, as demonstrated by [10, 9]. The models of my papers [15, 17] give novel tools (focused on memory usage rather than computation time) which should give insight in this setting and may connect to other areas of CS such as streaming algorithms. Information is another vital consideration in many settings, but is ignored by most conventional frameworks. Although recent theoretical advances are promising [2, 21], they fall short of modeling the success of real-world dynamic-offer designs such as those in [19], motivating new approaches to bridging mechanism and information design. Holistically, I believe my research is at the forefront of the economics and computation literature, with both significant existing contributions and bright potential for the future.

My Papers

[3] Itai Ashlagi, Mark Braverman, Amin Saberi, Clayton Thomas, and Geng Zhao. "Tiered Random Matching Markets: Rank Is Proportional to Popularity". In: *12th Innovations in Theoretical Computer Science Conference (ITCS)*. 2021. URL: https://arxiv.org/abs/2009.05124.

[6] Linda Cai and Clayton Thomas. "Representing All Stable Matchings by Walking a Maximal Chain". Mimeo. 2019. URL: https://arxiv.org/abs/1910.04401.

[7] Linda Cai and Clayton Thomas. "The short-side advantage in random matching markets". In: *Proceedings of the 5th SIAM Symposium on Simplicity in Algorithms (SOSA)*. 2022. URL: https://arxiv.org/abs/1910.04406.

[8] Linda Cai, Clayton Thomas, and S. Matthew Weinberg. "Implementation in Advised Strategies: Welfare Guarantees from Posted-Price Mechanisms When Demand Queries Are NP-Hard". In: *11th Innovations in Theoretical Computer Science Conference (ITCS)*. 2020. URL: https://arxiv.org/abs/1910.04342.

[3] My work extends Reny's [27] in a similar way to how [12] extends [11].

[4] Two of the authors of [4] have praised the simplicity of our arguments in [7] and/or told me they use our proof, instead of theirs, in their classes.

[5] I have two in-progress information design projects with Microsoft Research PhD interns (and with Nicole Immorlica and Brendan Lucier). With Joey Feffer, we are studying the information design problem that a school district faces when informing a population of students about the attributes of heterogeneous schools. With Ruxing Xu, we are studying the value of knowing the "source" of a piece of evidence, formalized as knowing the precise details of a signal-generating process drawn from a distribution over such processes.

[13] Venkata Gandikota, Elena Grigorescu, Clayton Thomas, and Minshen Zhu. "Maximally recoverable codes: The bounded case". In: *In Allerton Conference on Communication, Control, and Computing*. 2017. URL: https://ieeexplore.ieee.org/document/8262862.

[14] Aadityan Ganesh, Clayton Thomas, and S. Matthew Weinberg. "Revisiting the Primitives of Transaction Fee Mechanism Design". In: *Proceedings of the 25th ACM Conference on Economics and Computation (EC)*. 2024. URL: https://arxiv.org/abs/2409.18166.

[15] Yannai Gonczarowski, Ori Heffetz, and Clayton Thomas. "Strategyproofness-Exposing Descriptions of Matching Mechanisms". Reject and Resubmit at the American Economic Review. New version of "Strategyproofness-Exposing Mechanism Descriptions". 2024. URL: https://clathomasprime.github.io/papers/SeDoMM.pdf.

[16] Yannai Gonczarowski, Ori Heffetz, and Clayton Thomas. "Strategyproofness-Exposing Mechanism Descriptions". In: *Proceedings of the 24rd ACM Conference on Economics and Computation (EC)*. 2023. URL: https://arxiv.org/abs/2209.13148.

[17] Yannai Gonczarowski and Clayton Thomas. "Structural Complexities of Matching Mechanisms". In: *Proceedings of the 56th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. 2024. URL: https://arxiv.org/abs/2212.08709.

[18] Yannai A. Gonczarowski, Ori Heffetz, Guy Ishai, and Clayton Thomas. "Describing Deferred Acceptance and Strategyproofness to Participants: Experimental Analysis". In: *Proceedings of the 25rd ACM Conference on Economics and Computation (EC)*. In Review at Econometrica. 2024. URL: https://arxiv.org/abs/2410.07566.

[28] Shiri Ron, Clayton Thomas, S. Matthew Weinberg, and Qianfan Zhang. "Communication Separations for Truthful Auctions: Breaking the Two-Player Barrier". In: *65th IEEE Symposium on Foundations of Computer Science (FOCS)*. 2024. URL: https://arxiv.org/abs/2409.08241.

[30] Aviad Rubinstein, Raghuvansh R Saxena, Clayton Thomas, S Matthew Weinberg, and Junyao Zhao. "Exponential communication separations between notions of selfishness". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. 2021. URL: https://arxiv.org/abs/2012.14898.

[31] Clayton Thomas. "Classification of Priorities Such That Deferred Acceptance is OSP Implementable". In: *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*. 2021. URL: https://arxiv.org/abs/2011.12367.

[32] Clayton Thomas. "Priority-Neutral Matching Lattices Are Not Distributive". Working Paper. 2024. URL: https://arxiv.org/abs/2404.02142.

Other References

[1] Mohammad Akbarpour and Shengwu Li. "Credible Auctions: A Trilemma". In: *Econometrica* 88.2 (2020), pp. 425–467.

[2] Itai Ashlagi, Mark Braverman, Yash Kanoria, and Peng Shi. "Clearing matching markets efficiently: informative signals and match recommendations". In: *Management Science* 66.5 (2020), pp. 2163–2193.

[4] Itai Ashlagi, Yash Kanoria, and Jacob D. Leshno. "Unbalanced Random Matching Markets: The Stark Effect of Competition". In: *Journal of Political Economy* 125.1 (2017). Abstract in Proceedings of the 14th ACM Conference on Electronic Commerce (EC 2013).

[5] Eduardo M Azevedo and Jacob D Leshno. "A supply and demand framework for two-sided matching markets". In: *Journal of Political Economy* 124.5 (2016), pp. 1235–1268.

[9] Modibo K Camara. "Computationally Tractable Choice." In: *EC*. 2022.

[10] Federico Echenique, Daniel Golovin, and Adam Wierman. "A revealed preference approach to computational complexity in economics". In: *EC*. 2011.

[11] Lars Ehlers and Thayer Morrill. "(Il) legal assignments in school choice". In: *The Review of Economic Studies* 87.4 (2020), pp. 1837–1875.

[12] Yuri Faenza and Xuan Zhang. "Legal Assignments and Fast EADAM with Consent via Classic Theory of Stable Matchings". In: *Operations Research* 70.3 (May 2022).

[19] Julien Grenet, YingHua He, and Dorothea Kübler. "Preference discovery in university admissions: The case for dynamic multioffer mechanisms". In: *Journal of Political Economy* 130.6 (2022), pp. 1427–1476.

[20] Rustamdjan Hakimov and Dorothea Kübler. "Experiments on centralized school choice and college admissions: A survey". In: *Experimental Economics* 24 (2021), pp. 434–488.

[21] Nicole Immorlica, Jacob Leshno, Irene Lo, and Brendan Lucier. "Information acquisition in matching markets: The role of price discovery". Mimeo. 2020.

[22] Jacob D Leshno and Irene Lo. "The cutoff structure of top trading cycles in school choice". In: *Review of Economic Studies* 88.4 (2021), pp. 1582–1623.

[23] Shengwu Li. "Obviously strategy-proof mechanisms". In: *American Economic Review* 107.11 (2017), pp. 3257–87.

[24] Dominik Peters, Ariel D Procaccia, Alexandros Psomas, and Zixin Zhou. "Explainable voting". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 1525–1534.

[25] Ariel D Procaccia. "Axioms should explain solutions". In: *The Future of Economic Design: The Continuing Development of a Field as Envisioned by Its Researchers* (2019).

[26] Alex Rees-Jones and Ran Shorrer. "Behavioral Economics in Education Market Design: A Forward-Looking Review". In: *Journal of Political Economy Microeconomics* 1.3 (2023).

[27] Philip J Reny. "Efficient matching in the school choice problem". In: *American Economic Review* 112.6 (2022), pp. 2025–2043.

[29] Tim Roughgarden. "Transaction fee mechanism design". In: *ACM SIGecom Exchanges* 19.1 (2021), pp. 52–55.