The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

Balancing Lexicographic Fairness and a Utilitarian Objective with Application to Kidney Exchange

Duncan C. McElfresh[†,‡]
[†]Department of Mathematics
University of Maryland
dmcelfre@math.umd.edu

John P. Dickerson[†,‡]
[‡]Department of Computer Science
University of Maryland
john@cs.umd.edu

Abstract

In this work, we close an open theoretical problem regarding the price of fairness in modern kidney exchanges. We then propose a hybrid fairness rule that balances a lexicographic preference ordering over agents, with a utilitarian objective. This rule has one parameter which controls a bound on the price of fairness. We apply this rule to real data from a large kidney exchange and show that our hybrid rule produces more reliable outcomes than other fairness rules[1].

1 Introduction

Chronic kidney disease is a worldwide problem whose societal burden is likened to that of diabetes (Neuen et al. 2013). Left untreated, it leads to end-stage renal failure and the need for a donor kidney—for which demand far outstrips supply. In the United States alone, the kidney transplant waiting list grew from 58,000 people in 2004 to over 100,000 needy patients (Hart et al. 2016).[2]

To alleviate some of this supply-demand mismatch, *kidney exchanges* (Rapaport 1986; Roth, Sönmez, and Ünver 2004) allow patients with willing *living* donors to trade donors for access to compatible or higher-quality organs. In addition to these patient-donor pairs, modern exchanges include *non-directed donors*, who enter the exchange without a patient in need of a kidney. Exchanges occur in cycle- or chain-like structures, and now account for 10% of living transplants in the United States. Yet, access to a kidney exchange does not guarantee equal access to kidneys themselves; for example, certain classes of patients may be particularly disadvantaged based on health characteristics or other logistical factors. Thus, *fairness* considerations are an active topic of theoretical and practical research in kidney exchange and the matching market community in general.

Intuitively, any enforcement of a fairness constraint or consideration may have a negative effect on overall economic efficiency. A quantification of this tradeoff is known as the *price of fairness* (Bertsimas, Farias, and Trichakis 2011), which is equal to the relative efficiency loss due to a fairness constraint. Recent work by Dickerson, Procaccia, and Sandholm (2014) adapted this concept to the kidney exchange case, and presented two fair allocation rules that strike a balance between fairness and efficiency. Yet, as we show in this work, those rules can "fail" unpredictably, yielding an arbitrarily high price of fairness.

With this as motivation, we develop a the *hybrid-lexicographic* fairness rule, which balances lexicographic fairness and a utilitarian objective. To do this, we generalize the hybrid utility function proposed by Hooker and Williams (2012), which chooses between a Rawlsian (or maximin) objective and a utilitarian objective for multiple classes of agents; instead, our method chooses between a lexicographic objective and a utilitarian objective, to favor disadvantaged classes when necessary. The hybrid-lexicographic method is parameterized by a bound on the price of fairness, as opposed to a set of parameters that may result in hard-to-predict final matching behavior, as in past work. Furthermore, we generalize the method to the case of > 2 classes of agents.

We implement our rule in a realistic mathematical programming framework and–on real data from a large, multi-center, fielded kidney exchange–show that our rule effectively balances fairness and efficiency without unwanted outlier behavior.

1.1 Our Contributions

Dickerson, Procaccia, and Sandholm (2014) finds that the theoretical price of fairness in kidney exchange is small when *only* patient-donor pairs participate in the exchange. They did not include non-directed donors (NDDs). However, in modern kidney exchanges, non-directed donors (NDDs) provide many more matches than patient-donor pairs; furthermore, NDDs create more opportunities to expand the fair matching, potentially increasing the price of fairness. Here, we prove that adding NDDs to the theoretical model actually *decreases* the price of fairness, and that—with enough NDDs—the price of fairness is zero.

Real kidney exchanges are less dense and more uncertain than the (standard) theoretical model in which we prove our results. Previous approaches to incorporating fairness into kidney exchange have neglected this fact: they have been either ad-hoc—e.g., "priority points" decided on by committee (Kidney Paired Donation Work Group 2013)—or

[1]A full version of this paper and supplemental material can be found at https://arxiv.org/abs/1702.08286

[2]https://optn.transplant.hrsa.gov/data/

brittle (Roth, Sönmez, and Ünver 2005; Dickerson, Procaccia, and Sandholm 2014), resulting in an unacceptably high price of fairness. This paper provides the first approach to incorporating fairness into kidney exchange in a way that both prioritizes disadvantaged participants, but also comes with acceptable worst-case guarantees on the price of fairness. Our method is easily applied as an objective in the mathematical-programming-based clearing methods used in today's fielded exchanges; indeed, using real data we show that this method guarantees a limit on efficiency loss.

However our method is not a "silver bullet" for enforcing fairness in kidney exchange. Just like previous methods, ours can perform either poorly or well; this depends on the particular exchange, and on the desired outcome. The question of which outcomes are most "desirable" is also important, as discussed in the following section.

1.2 Ethics and Policy Implications

Our approach is generally set in the framework of ethics and market design, as described by Li (2017). In this framework, policymakers rely on technically-minded market designers to develop the algorithms and mechanisms of a market. During the design process, market designers consider the morally-relevant implications of their designs, but remain morally neutral; policymakers ultimately decide whether or not a mechanism *should* be used. In this work we assume the role of market designer, in the application of kidney exchange.

Kidney exchange has many ethical implications, especially when some participants are marginalized or exploited by the exchange policies or each other. Policymakers may want to include several ethical criteria to identify the optimal matching–such as overall welfare, prioritization of marginalized patients, or robustness to uncertainty. Balancing these objectives is both mathematically and morally difficult. Even if policymakers can decide on the appropriate ethical criteria, writing them into an algorithm is not always easy. In this work, we study the balance of two ethical criteria in kidney exchange: overall efficiency and fairness for highly sensitized patients.

In particular, we investigate two methods for enforcing fairness: the α-lexicographic fairness method proposed by Dickerson, Procaccia, and Sandholm (2014), and the hybrid-lexicographic method proposed in this work. Each of these methods guarantees a different outcome: α-lexicographic guarantees that marginalized patients receive at least $\alpha\%$ of their maximum possible utility; hybrid-lexicographic prioritizes marginalized patients, while guaranteeing that total efficiency loss does not exceed a certain value (set by a parameter Δ).

Neither of these methods is superior to the other; indeed there are cases where either is more appropriate. For example, if the price of fairness is low, or policymakers require equality between classes of patients, α-lexicographic fairness can guarantee a fair outcome. In cases where the price of fairness is high (as in small or sparse exchanges), hybrid-lexicographic can limit total efficiency loss.

1.3 Future Work

This work addresses fairness in a single kidney exchange; however real kidney exchanges are dynamic – matching decisions made in the present naturally impact future exchanges (Anshelevich et al. 2013; Akbarpour, Li, and Gharan 2014; Anderson et al. 2015; Dickerson and Sandholm 2015). Enforcing fairness may have long-term consequences for the kidney exchange pool. For example, how does prioritizing one type of patient affect the long-term welfare of other patient types?

Another important problem is the aggregation of ethical preferences. Policymakers often disagree on the ethical criteria used to find the optimal matching. Future work should develop a method for ethical-preference elicitation and aggregation in kidney exchange. A matching algorithm should be developed to identify the optimal matching subject to these aggregated preferences, which may not be simple.

References

Akbarpour, M.; Li, S.; and Gharan, S. O. 2014. Dynamic matching market design. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 355.

Anderson, R.; Ashlagi, I.; Gamarnik, D.; and Kanoria, Y. 2015. A dynamic model of barter exchange. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 1925–1933.

Anshelevich, E.; Chhabra, M.; Das, S.; and Gerrior, M. 2013. On the social welfare of mechanisms for repeated batch matching. In *AAAI Conference on Artificial Intelligence (AAAI)*, 60–66.

Bertsimas, D.; Farias, V. F.; and Trichakis, N. 2011. The price of fairness. *Operations Research* 59(1):17–31.

Dickerson, J. P., and Sandholm, T. 2015. FutureMatch: Combining human value judgments and machine learning to match in dynamic environments. In *AAAI Conference on Artificial Intelligence (AAAI)*, 622–628.

Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2014. Price of fairness in kidney exchange. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 1013–1020.

Hart, A.; Smith, J. M.; Skeans, M. A.; Gustafson, S. K.; Stewart, D. E.; Cherikh, W. S.; Wainright, J. L.; Boyle, G.; Snyder, J. J.; Kasiske, B. L.; and Israni, A. K. 2016. Kidney. *American Journal of Transplantation (Special Issue: OPTN/SRTR Annual Data Report 2014)* 16, Issue Supplement S2:11–46.

Hooker, J. N., and Williams, H. P. 2012. Combining equity and utilitarianism in a mathematical programming model. *Management Science* 58(9):1682–1693.

Kidney Paired Donation Work Group. 2013. OPTN KPD pilot program cumulative match report (CMR) for KPD match runs: Oct 27, 2010 – Apr 15, 2013.

Li, S. 2017. Ethics and market design. In *Oxford Review of Economic Policy, forthcoming*.

Neuen, B. L.; Taylor, G. E.; Demaio, A. R.; and Perkovic, V. 2013. Global kidney disease. *The Lancet* 382(9900):1243.

Rapaport, F. T. 1986. The case for a living emotionally related international kidney donor exchange registry. *Transplantation Proceedings* 18:5–9.

Roth, A.; Sönmez, T.; and Ünver, U. 2004. Kidney exchange. *Quarterly Journal of Economics* 119(2):457–488.

Roth, A.; Sönmez, T.; and Ünver, U. 2005. Pairwise kidney exchange. *Journal of Economic Theory* 125(2):151–188.