A Unified Approach to Dynamic Matching and Barter Exchange

John P. Dickerson

Monday 31st August, 2015

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Tuomas Sandholm, Chair
Avrim Blum
Ariel Procaccia
Craig Boutilier, Google
Alvin Roth, Stanford University

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Keywords: matching, barter exchange, market design, dynamic matching markets, kidney exchange

Abstract

The exchange of indivisible goods without money addresses a variety of constrained economic settings where a medium of exchange—such as money—is considered inappropriate. Participants are either matched directly with another participant or, in more complex domains, in barter cycles and chains with many other participants before exchanging their endowed goods. This thesis addresses the design, analysis, and real-world fielding of dynamic matching markets and barter exchanges.

We present new mathematical models for dynamic barter exchange that more accurately reflect reality, prove theoretical statements about the characteristics and behavior of these markets, and develop provably optimal market clearing algorithms for models of these markets that can be deployed in practice. We show that taking a holistic approach to balancing efficiency and fairness can often practically circumvent negative theoretical results. We support the theoretical claims made in this thesis with extensive experiments on data from the United Network for Organ Sharing (UNOS) Kidney Paired Donation Pilot Program, a large kidney exchange clearinghouse in the US with which we have been actively involved.

Specifically, we study three competing dimensions found in both matching markets and barter exchange: uncertainty over the existence of possible trades (represented as edges in a graph constructed from participants in the market), balancing efficiency and fairness, and inherent dynamism. For each individual dimension, we provide new theoretical insights as to the effect on market efficiency and match composition of clearing markets under models that explicitly consider those dimensions. We support each theoretical construct with new optimization models and techniques, and validate them on simulated and real kidney exchange data. In the cases of edge failure and dynamic matching, where edges and vertices arrive and depart over time, our algorithms perform substantially better than the status quo deterministic myopic matching algorithms used in practice, and also scale to larger instance sizes than prior methods. In the fairness case, we empirically quantify the loss in system efficiency under a variety of equitable matching rules.

Next, we combine all of the dimensions, along with high-level human-provided guidance, into a unified framework for learning to match in a general dynamic model. This framework, which we coin FUTUREMATCH, takes as input a high-level objective (e.g., "maximize graft survival of transplants over time") decided on by experts, then automatically (i) learns based on data how to make this objective concrete and (ii) learns the "means" to accomplish this goal—a task that, in our experience, humans handle poorly. We validate FUTUREMATCH on UNOS exchange data and make policy recommendations based on it.

Finally, we present a model for liver exchange and a model for multi-organ exchange; for the latter, we show that it theoretically and empirically will result in greater social welfare than multiple individual exchanges.

Contents

Chapter 1

Introduction

The exchange of indivisible goods without money addresses a variety of constrained economic settings where a medium of exchange—such as money—is considered repugnant [Roth, 2007] or otherwise inappropriate. One example setting is the allocation of donor organs to needy patients; while the act of organ donation is laudable, the provision of monetary payments to those who donate is often societally unacceptable and, in nearly all countries, illegal. In general, the removal of a medium of exchange forces markets operating in such settings to rely on the "double coincidence of wants" [Jevons, 1885], where two participants must both exist simultaneously and be in possession of a good that the other desires. This is typically quite restrictive, making the design and operation of efficient markets without money theoretically and computationally challenging.

This thesis addresses the design, analysis, and—when appropriate—real-world fielding of matching markets and barter exchanges through the lens of computer science. We focus on the creation of new mathematical models for these markets that more accurately reflect reality, proofs of statements about the characteristics of these markets in theory, and the development of provably optimal clearing algorithms for these models that can be deployed in practice. We show that taking a holistic approach to balancing efficiency and fairness can often practically circumvent negative theoretical results. We support the theoretical claims made in this thesis with extensive experiments on data from the United Network for Organ Sharing (UNOS) Kidney Paired Donation Pilot Program, a large kidney exchange clearinghouse in the US with which we have been actively involved.

1.1 Structure of this proposal document

In this chapter, we briefly overview completed (§1.2) and proposed (§1.3) research and formalize a thesis statement. Chapter 2 defines three competing dimensions of markets without money and addresses them individually from a theoretical and computational point of view. Throughout this document, we use *kidney exchange*—a mechanism through which patients can swap willing but medically incompatible donors—as a running example of a real-world instantiation of barter exchange; a basic mathematical model of kidney exchange is defined in Section 2.1. Following this, Section 2.2 looks at a stochastic adaptation of this model, where edges can fail after an algorithmic match occurs. Section 2.3 addresses notions of equity in matching and barter exchange, where the allocation of goods to some participants is valued unevenly. Both of these sections look at the clearing problem in a static setting; Section 2.4 extends the model to the dynamic case, where participants arrive and depart over time. In each of these sections, we provide new theoretical and algorithmic results that address the dimensions that are introduced, and support our results with extensive simulations on data from the UNOS kidney exchange.

Chapter 3 combines each of the dimensions presented in Chapter 2, along with high-level human-provided guidance, into a unified matching framework. Our framework, dubbed FUTUREMATCH, is presented in Section 3.1. Section 3.1.1 presents preliminary experimental results using FUTUREMATCH under a variety of objectives on simulated UNOS exchange data; critically, we show that it is possible to automatically learn a matching policy that strictly improves match quality in two dimensions (corresponding to efficiency and a definition of fairness) over the status quo at UNOS. Chapter 4 presents other (mostly specific to organ exchange) research we plan to include in this thesis, including multi-organ exchange (§4.1), an analysis of non-directed-donor-initiated chains in a dense model of kidney exchange (§4.2), and some exploratory application areas we are currently considering (§4.3).

Chapter 5 presents a timeline for the completion of the proposed work detailed in Chapters 2, 3, and 4, as well as the completion and defense of the thesis document. Finally, Chapter 6 briefly overviews other research projects, not to be included in our formal thesis, that we have either led or with which we have been involved. These projects may be of interest to the reader; they focus on the design of markets and/or the use of optimization to provide policy recommendations.

1.2 Completed research

This section summarizes thesis research that has already been completed. This work is organized by chapter.

1.2.1 Matching dimensions (completed research in Chapter 2)

We have already taken preliminary steps toward each of the three dimensions addressed in Chapter 2. Some of this work has been published and some is under review.

Edge failure. We show that explicitly considering edge failure probability in the clearing optimization problem can significantly increase the number of successfully matched vertices (i) in theory, in a random graph model; (ii) on real data from kidney exchange match runs between 2010 and 2014; (iii) on synthetic data generated via a model of dynamic kidney exchange. We design a branch-and-price-based optimal clearing algorithm specifically for this probabilistic exchange clearing problem and show that this new solver scales well on large simulated kidney exchange data, unlike prior clearing algorithms. We show experimentally that taking failed parts from an initial match and instantaneously *rematching* them with other vertices still in the waiting pool can result in significant gains. Finally, we show that failure-aware matching can increase overall system efficiency and simultaneously increase the expected number of transplants to *highly-sensitized* patients, in both static and dynamic models.

This work was published at EC-2013 [Dickerson et al., 2013]; an extended version that includes empirical balancing of efficiency and fairness in this static stochastic model will soon be submitted to an OR journal. Some related work was also presented at an AAMAS-2014 workshop [Dickerson et al., 2014b].

Equity. Matching under utilitarian or near-utilitarian objectives may marginalize certain vertex classes. We focus on improving access to kidneys for highly-sensitized, or hard-to-match, patients. Toward this end, we formally adapt a recently introduced measure of the tradeoff between fairness and efficiency—the *price of fairness* [Bertsimas et al., 2011, Caragiannis et al., 2009]—to the standard kidney exchange model. We show that the price of fairness in the standard theoretical model is small. We then introduce two natural definitions of fairness and formally define them in the standard *deterministic* model of kidney exchange, and in the failure-aware model of Dickerson et al. [2013] that takes post-algorithmic match failures into account. We empirically explore the tradeoff between matching more hard-to-match patients and the overall utility of a utilitarian matching, on data from the UNOS nationwide kidney exchange and simulated data from each of the standard kidney exchange distributions.

This work was published at AAMAS-2014 [Dickerson et al., 2014c].

Dynamic matching. The competitive ratios of prior-free online algorithms for dynamic matching and barter exchange are often unacceptably poor. The algorithm should take distributional information about possible futures into account in deciding what action to take now. This is typically done by drawing sample trajectories of possible futures at each time period, but may require a prohibitively large number of trajectories or prohibitive memory and/or computation to decide what action to take.

Instead, we introduce a natural, general policy parameterization approach and techniques for operationalizing it. Specifically, we propose to learn *potentials* of elements (e.g., vertices, edges, cycles, and so on) of the current problem. The potential represents an estimate of how much that element can contribute to the objective in the future. The potentials can be viewed as policy parameters to be optimized using a black box program; we learn them using parameter tuning [Hutter et al., 2009, 2011]. Then, at run time, we simply run an offline matching algorithm (like those due to Abraham et al. [2007], Li et al. [2011], Dickerson et al. [2013], Constantino et al. [2013], Dickerson et al. [2014c], Glorie et al. [2014], Klimentova et al. [2014], Anderson [2014], Anderson et al. [2015b]) at each time period, but subtracting out in the objective the potentials of the elements used up in the matching. This causes the batch optimizer—which is traditionally myopic—to take the future into account without suffering a run-time cost.

We apply the approach to kidney exchange. We theoretically compare the power of using potentials on increasingly large elements: vertices, edges, cycles, and the entire graph (optimum). Then, experiments show that by learning vertex potentials, our algorithm matches more patients than the current practice of clearing myopically. It scales to exchanges orders of magnitude beyond those handled by prior algorithms for this (unsimplified) dynamic problem.

This work was published at AAAI-2012 [Dickerson et al., 2012a].

1.2.2 Holistic matching and exchange (completed research in Chapter 3)

Drawing on the individual dimensions summarized above (§1.2.1) and discussed more formally in Chapter 2, we present FUTUREMATCH, a framework for *learning to match in a general dynamic model*. FUTUREMATCH takes as input a high-level objective (e.g., "maximize graft survival of transplants over time") decided on by experts, then automatically (i) learns based on data how to make this objective concrete and (ii) learns the "means" to accomplish this goal—a task, in our experience, that humans handle poorly. It uses data from all live kidney transplants in the US since 1987 to learn the quality of each possible match; it then learns the *potentials* of elements of the current input graph offline (e.g., potentials of pairs based on features such as donor and patient blood types), translates these to weights, and performs a computationally feasible batch matching that incorporates dynamic, failure-aware considerations through the weights. We validate FUTUREMATCH on UNOS exchange data. It results in higher values of the objective. Furthermore, even under economically inefficient objectives that enforce equity, it yields better solutions for the *efficient* objective (which does not incorporate equity) than traditional myopic matching that uses the efficiency objective.

This work was published at AAAI-2015 [Dickerson and Sandholm, 2015].

1.2.3 Other thesis research (completed research in Chapter 4)

This section overviews completed research to be included in this thesis that is organ exchange-specific, unlike the research above that is applicable to general dynamic matching and barter exchange.

Liver & multi-organ exchange. While fielded kidney exchanges see huge benefit from *altruistic* kidney donors (who give an organ without a paired needy candidate), a significantly higher medical risk to the donor deters similar altruism with livers. We explore the idea of large-scale *liver* exchange, and show on demographically accurate data that vetted kidney exchange algorithms can be adapted to clear such an exchange at the nationwide level. We then propose cross-organ donation where kidneys and livers can be bartered for each other. In two adaptations of random graph models due to Ashlagi and Roth [2011, 2014] and Ashlagi et al. [2012], we show theoretically that this *multi-organ* exchange provides linearly more transplants than running separate kidney and liver exchanges; this linear gain is a product of altruistic kidney donors creating chains that thread through the liver pool. We support this result experimentally on demographically accurate multi-organ exchanges.

This work was published at AAAI-2014 [Dickerson and Sandholm, 2014]. An extended version is under submission at J. Artificial Intelligence Research*.*

Kidney exchange chains. We provide a theoretical analysis of the efficacy of chains in the most widely used *dense* kidney exchange model due to Ashlagi and Roth [2011, 2014], proving that long chains do not help beyond chains of length of 3 in the large. This completely contradicts our real-world results gathered from the budding nationwide kidney exchange in the United States; there, solution quality improves by increasing the chain length cap to 13 or beyond (as of 2012, but this qualitative power of long chains has been shown repeatedly in practice [Ashlagi et al., 2012]). We analyze reasons for this gulf between theory and practice; in essence, a dense random graph model of kidney exchange is not appropriate.

This work was published at AAMAS-2012 [Dickerson et al., 2012b].

1.3 Proposed research

This section summarizes proposed thesis research. This work is organized by the chapter in which it will appear.

1.3.1 Matching dimensions (proposed research in Chapter 2)

We propose to continue research in each of the three dimensions addressed in Chapter 2.

Edge failure. Dickerson et al. [2013] showed that explicitly considering—in the optimization problem—the probability of failure for each edge in a graph results in large theoretical and empirical gains over status quo deterministic-style optimization. In many applications (including kidney exchange, online labor markets, etc.), additional testing can be performed on a per-edge basis to either reduce or remove the uncertainty of the true existence of that edge. Indeed, a basic form of this model has been studied in the standard matching literature as the *query-commit* problem [Chen et al., 2009, Adamczyk, 2011, Goel and Tripathi, 2012, Costello et al., 2012, Bansal et al., 2012, Blum et al., 2013], where the existence of an edge can be independently queried, but (in most of this work) any edge that is queried and exists must be included in the final matching. More related to the general stochastic barter exchange problem is our work in Blum et al. [2015], which looks at adaptively testing edges over a number of rounds and then either choosing a maximum matching or maximum k-set packing among these queried edges. Prior work assumes either no budget or a per-vertex limit on the number of incident edges to test; in this thesis, we propose to build on the query-commit literature and Dickerson et al. [2013] by considering the case where a clearinghouse has an overall budget to spend on testing edges, such that spending more of that budget on an edge increases the accuracy of an estimate of whether or not that edge exists. A related alternative that we plan to explore relaxes the hard budget constraint, instead maximizing some utility function that considers match quality minus total budget spent. This is motivated by kidney exchange, where variable levels of blood- and tissue-type testing across participating centers in a multi-center exchange lead to varying estimates of edge success probability.

Equity. We propose to extend our initial work [Dickerson et al., 2014c] on fairness in kidney exchange by generalizing the theoretical results in that paper's original framework (e.g., determining the price of fairness in exchange graphs that are dense with chains, or are sparse [Ashlagi and Roth, 2011, Ashlagi et al., 2012, Dickerson et al., 2013, Ashlagi and Roth, 2014]). Our initial work [Dickerson et al., 2014c] only considered "sensitized" and "not sensitized" patients; equity considerations could also take other features into account, like waiting time in the pool and pediatric status; we propose explore a model with a more generalized notion of fairness than our original work. We also propose to submit an invited journal version of the work from Dickerson et al. [2014c] to *Artificial Intelligence* in the next few months, hopefully with the new theoretical results on dense exchange graphs with chains included.

Dynamic matching. Our work on dynamic matching [Dickerson et al., 2012a] suffers from: high dimensionality of the state space of the potentials we are learning, non-determinism of the dynamic barter exchange-based objective we are computing given a set of potentials and their values, and the heavy computation required to even compute the value of a single run of a dynamic exchange simulation under a set of potentials. Indeed, the learning process in [Dickerson et al., 2012a] did not converge (although the best values we found did produce notable improvements over myopic matching). Furthermore, there is a tradeoff between the expressive power of the potentials (allowing potentials for larger structural elements such as edges, or cycles, or even beyond, having more expressiveness) and the computational power needed to learn the potentials (the hypothesis space being larger the more variables there are). We propose to tackle the increase in the complexity of the learning process as the space of potentials increases by making independence assumptions about classes of potentials. For example, in kidney exchange, it may be reasonable to assume that potentials based on blood type versus those based on patient or donor location are roughly independent, and thus values learned for a set of potentials in the former group can be applied repeatedly when learning the latter group's potential values. We aim to provide new theoretical results in this setting, and validate them experimentally.

1.3.2 Holistic matching and exchange (proposed research in Chapter 3)

The completion of any combination of the proposed research in the three dimensions discussed above (§1.3.1) will feed into the FUTUREMATCH framework we presented in Dickerson and Sandholm [2015]. Advances in failure-aware and fairness-aware matching will permit experimental results that more accurately mimic the reality seen in today's fielded kidney exchanges, while advances in the potentials-based dynamic matching direction described above will allow experimental results on larger graphs. This latter point will be of increasing importance for kidney exchange in the coming years, as fielded exchanges continue to grow in steady-state size. We propose to build on FUTUREMATCH with the end goal of making real-world policy recommendations to the general kidney exchange community. The idea here is to use the algorithmic and computational advances listed above, along with input from the policymakers and doctors with whom we currently interact at UNOS, to have a paper ready for submission to a medical journal (e.g., *American Journal of Transplantation*) by the time this thesis is finished. This paper will use FUTUREMATCH to provide high-level operational recommendations to those who run kidney (or other organ) exchanges.

1.3.3 Other thesis research (proposed research in Chapter 4)

This section overviews completed research to be included in this thesis that is organ exchange-specific, unlike the research above that is applicable to general dynamic matching and barter exchange. It also briefly discusses exploratory applications in new matching domains that may or may not come to fruition by the time this thesis is concluded.

Liver & multi-organ exchange. Different organs have different live transplantation characteristics. For example, a liver donation has increased risk to live donors, with very high rates of donor morbidity (24%), "near-miss" events in surgery (1.1%), and mortality (0.2%) compared to live donor kidney transplantation [Cheah et al., 2013]. Our initial work [Dickerson and Sandholm, 2014] did not address equity issues on those edges that transfer from one organ's subgraph to another organ's subgraph in matched chains or cycles that cross through multiple organ subgraphs. We propose to further generalize definitions of equity (in line with the proposed research described in Section 1.3.1 above) to this model. (This could be applicable to general barter exchanges with different goods such that swaps across agents with differently-endowed goods have vastly differentiated associated risk, but we plan to approach the problem from an organ exchange-specific point of view.)

Exploratory domains. There are a number of application areas in dynamic matching and barter exchange that we are currently exploring or plan to explore. *It is likely that work in the exploratory domains listed below will not be ready by the time this thesis is complete; fielding any of these new markets will take quite some time. Rather, we list these domains as examples of follow-up work that will use techniques described in this thesis.*

- Zoos often acquire new animals from other zoos; however, the purchase of animals is sometimes illegal (due to international trade laws) and/or is generally seen as unacceptable, so zoos currently swap animals through an offline and very ad hoc system. We would like to create a centralized mechanism, ZooSwap, through which zoos can list their inventory, express their preferences, and swap animals.
- Haluk Ergin, Tayfun Sönmez, and Utku Ünver recently begun the process of fielding a lung exchange in Japan [Ergin et al., 2014]. We would like to explore the feasibility of fielding either a liver exchange or a cross-liver-kidney exchange, in line with the preliminary work by Dickerson and Sandholm [2014].
- In Italy, workers in different geographical locations are assigned jobs, but may wish to swap those jobs with other workers either within or without their current locale. We would like to explore the fielding of a centralized swap mechanism to address this use case (joint with Nicola Gatti at Politecnico di Milano).

1.4 Thesis Statement

This thesis aims to provide a theoretically-justified and experimentally-validated holistic approach to designing and operating a fielded dynamic matching market or barter exchange. Our thesis statement is thus:

Competing dimensions—equity, efficiency, and computational tractability—in dynamic matching markets and barter exchanges can be balanced holistically through computational optimization methods and informed by random graph models.

Chapter 2

Competing dimensions in matching and barter exchange

In this section, we formalize the clearing problem in barter exchange with an explicit focus on the kidney exchange problem. We then identify and define three competing dimensions of markets without money. For each of these dimensions, we summarize the main theoretical and empirical results from our completed research and then describe proposed research.

2.1 Preliminaries

Numerous barter exchanges are currently fielded, such as:

- House exchange, where participants seek to swap homes (e.g., Intervac[1] and Best House Swap[2]);
- Room exchange, where college roommates simultaneously "trade up" to better roommates (e.g., The Room Exchange at the University of Maryland[3]);
- Book exchange, where participants swap books after reading them (e.g., Read It Swap It[4]);
- Shoe exchange, where participants who require only a single shoe or two shoes of different sizes due to injury, disease, or genetic disorder can swap shoes with similar participants (e.g., the National Odd Shoe Exchange[5]);
- Shift exchange, where nurses or other shift workers swap shifts (typically an ad hoc process); and
- General barter exchange, where participants can swap different classes of goods (e.g., Swap.com,[6] Tradeaway,[7] BarterQuest,[8] and others).

In this thesis, we focus on *kidney exchange*, a recent innovation that allows patients who suffer from terminal kidney failure, and have been fortunate enough to find a willing but incompatible kidney donor, to swap donors. Indeed, it may be the case that two donor-patient pairs are incompatible, but the first donor

[1] http://intervac-homeexchange.com/
[2] http://besthouseswap.com
[3] http://reslife.umd.edu/housing/reassignments/roomexchange/
[4] http://readitswapit.co.uk
[5] http://oddshoe.org
[6] http://swap.com
[7] http://tradeaway.com
[8] http://barterquest.com

is compatible with the second patient, and the second donor is compatible with the first patient; in this case a life-saving match is possible. As we discuss below, sequences of swaps can even take the form of long cycles or chains.

The need for successful kidney exchanges is acute because demand for kidneys is far greater than supply. Although receiving a deceased-donor kidney is a possibility, as of March 26, 2015, there are 101,663 people on the US national waiting list,[9] making the median waiting time dangerously long. The rest of this section overviews the basic kidney exchange problem and presents the mathematical foundations on which we build in the rest of this thesis.

The idea of kidney exchange was presented by Rapaport [1986], while the first organized kidney exchange, the New England Paired Kidney Exchange (NEPKE), started in 2003–2004 [Roth et al., 2004, 2005a,b, 2006]. It has since ceased operations and its pool was merged into the United Network for Organ Sharing (UNOS) kidney exchange, which started in 2010 and now includes 60% of the US transplant centers. This thesis uses data from that exchange to support its theoretical results.

2.1.1 Graph-theoretic model

The standard model for kidney exchange encodes an n-patient kidney exchange as a directed *compatibility graph* $G = (V, E)$ by constructing one vertex for each patient-donor pair. An edge e from v_i to v_j is added if the patient in v_j wants and is compatible with the donor kidney of v_i. A donor is willing to give her kidney if and only if the patient in her vertex v_i receives a kidney. The weight w_e of an edge e represents the utility to v_j of obtaining v_i's donor kidney. Indeed, this model works for general n-participant barter exchanges, where vertices trade and receive utility from items instead of kidneys.

A cycle c in the graph G represents a possible kidney swap, where each vertex in c obtains the kidney of the previous vertex. We denote by *k-cycle* a cycle with k patient-donor pairs. In fielded kidney exchange, cycles of length at most only some small constant L are allowed. All transplants in a cycle must be performed simultaneously so that no donor backs out after his patient has received a kidney but before he has donated his kidney. In most fielded kidney exchanges, including the UNOS kidney exchange, $L = 3$ (i.e., only 2- and 3-cycles are allowed).

Fielded kidney exchanges also gain great utility through the use of *chains* (see, e.g., work by Ashlagi et al. [2011], Gentry and Segev [2011], Ashlagi et al. [2012], Dickerson et al. [2012b], Gentry et al. [2009], Montgomery et al. [2006], Rees et al. [2009], Roth et al. [2006], Woodle et al. [2010], Glorie et al. [2014], Anderson et al. [2015b]). Chains start with an altruistic donor donating his kidney to a patient, whose paired donor donates her kidney to another patient, and so on.

A *matching* M is a collection of vertex-disjoint cycles and chains in the graph G. Note that the elements of the matching must be disjoint because no donor can give more than one of his kidneys. Then, given the set of all legal matchings $\mathcal{M}$, the *clearing problem* in kidney exchange (and, indeed, in any barter exchange) is to find a matching M^* that maximizes some utility function $u : \mathcal{M} \to \mathbb{R}$. Formally:

$$M^* = \arg\max_{M \in \mathcal{M}} u(M)$$

The most basic utility function solves the maximum cardinality disjoint cycle cover problem and is implemented as $u(M) = \sum_{c \in M} \sum_{e \in c} w_e$, where $w_e = 1$ for each edge $e \in E$. This thesis explores a variety of different adaptations to this utility function, as well as adaptations to the model and to the set of feasible matchings $\mathcal{M}$. Many of these adaptations stem from one of three dimensions found in real-world matching—edge failure, equity, and dynamism—which we explore in the following sections.

[9] http://optn.transplant.hrsa.gov.

2.1.2 Two random graph models for kidney exchange

Random graph models provide insight about the characteristics of "typical" graphs drawn from a probability distribution over a family of graphs. Random graph models have informed kidney exchange policy and optimization techniques since before the inception of the first fielded exchange [Roth et al., 2004, 2005a,b]. We describe the two basic random graph models of kidney exchange here. Both models build on the classical work of Erdős and Rényi [1960].

Dense random graphs

We begin by overviewing the *dense* model of kidney exchange [Roth et al., 2004, 2005a,b, Ashlagi and Roth, 2011, 2014, Toulis and Parkes, 2011, Ünver, 2010, Caragiannis et al., 2011]. This model concentrates on blood types of donors and patients. At a very high level, human blood is split into four types—O, A, B, and AB—based on the presence or absence of type A and type B proteins. Ignoring other potential complications, a type O kidney can be transplanted into any patient; type A and B kidneys can be transplanted into A and B patients respectively or an AB patient; and type AB kidneys can only be transplanted into type AB patients. Therefore, some patients are more difficult to match with a random donor than others. O-patients are the hardest to match because only O-type kidneys can be given to them. Similarly, O-donors are the easiest to match.

An *under-demanded* pair is any pair such that the donor is not ABO-compatible with the patient. If an under-demanded pair contains only type A and B blood, it is called *reciprocal*. Any pair in the pool such that the donor is ABO-compatible with the candidate is called *over-demanded*. Furthermore, if a donor and candidate share the same blood type, they are a *self-demanded* pair. Under-demanded and reciprocal pairs are intuitively "harder" to match than over-demanded and self-demanded pairs. Note that this is not necessarily the case if sensitization, the probability of matching with a random donor, is considered. For example, an A-type patient who is lowly-sensitized is typically easier to match than an O-type patient who is highly-sensitized; however, the dense model does not consider different degrees of sensitization. The dense model critically assumes that a donor and patient who are blood type compatible are tissue type incompatible with *constant* probability $\bar{p}$. This differs from the sparse model we will define below, where lowly-sensitized patients have a constant edge probability while highly-sensitized patients do not (which more closely mimics reality). It also denotes by μ_X the frequency of blood type X, and assumes $\mu_{\text{O}} < 3\mu_{\text{A}}/2$ and an ordering $\mu_{\text{O}} > \mu_{\text{A}} > \mu_{\text{B}} > \mu_{\text{AB}}$. The blood type of a *paired* patient and donor in this model are assumed to be independent; however, the blood types of patients and donors are drawn individually in accordance with the defined μ_X, for $X \in \{\text{O}, \text{A}, \text{B}, \text{AB}\}$. The United States national blood type distribution satisfies these constraints.

We prove new theoretical results in dense random graph models in Dickerson et al. [2012b,a, 2014c], Dickerson and Sandholm [2014], covered in Sections 2.3, 2.4, 4.1, and 4.2.

Sparse random graphs

Addressing some concerns about the properties of dense Erdős-Rényi graphs, Section 4.2 of Ashlagi et al. [2012] defines a sparse random graph model of kidney exchange pools with highly and non-highly sensitized patient-donor pairs. The model works with random compatibility graphs with $n + t(n)$ vertices, pertaining to n incompatible patient-donor pairs (denoted by the set P), and $t(n)$ altruistic donors (denoted by the set A) respectively. Edges between vertices represent not just blood-type compatibility, but also immunological compatibility—the *sensitization* of the patient. Given a blood type-compatible donor, let p denote the probability that an edge exists between a patient and that donor. In this sparse model, $p = \Theta(1/n)$ (or similar, Ashlagi et al. [2012] addresses variations of this); that is, p becomes smaller as the pool becomes larger, thus maintaining sparsity.

We prove new theoretical results in sparse random graph models in Dickerson et al. [2013] and in Dickerson and Sandholm [2014], covered in Sections 2.2 and 4.1.

2.1.3 Optimization model

Current kidney exchange pools are small, containing at most a few hundred patients at a time. For example, as of March 2015, the largest exchange program in the US (the National Kidney Registry) has slightly over 300 active donors, while the UNOS kidney exchange has slightly under 300 active donors. However, as kidney exchange gains traction, these pools will grow.

As discussed by Abraham et al. [2007], the estimated steady-state size of a US nationwide kidney exchange is 10,000 patients. Clearing pools of this size is a computational challenge. Abraham et al. [2007] showed that the basic clearing problem (with a maximum cardinality objective function) is NP-hard when the cycle cap $L > 2$. Note that when $L = 2$ the basic clearing problem is solvable in polynomial time as a normal matching problem using, e.g., a maximum matching algorithm due to Edmonds [1965], as discussed by Segev et al. [2005]. However, it has been shown that efficiency gains are seen in theory and in practice when using $L > 2$, so it is important to solve the more complex form of this problem [Roth et al., 2007].

One approach to solving the clearing problem is through the use of integer programming [Roth et al., 2007]. Formally, denote the set of all (uncapped length) chains and all cycles of length no greater than L by $C(L)$. Let $|c|$ represent the number of candidate-donor pairs in a cycle or chain c, and w_c the weight of a cycle or chain. Under the maximum cardinality objective, $w_c = |c|$ for cycles, and $w_c = |c| - 1$ for chains. A formulation proposed by Abraham et al. [2007] associates with each cycle or chain c in the compatibility graph a binary decision variable $x_c \in \{0, 1\} \; \forall c \in C(L)$. Then, we must solve the following integer linear program:

$$\max \sum_{c \in C(L)} w_c \, x_c \qquad s.t. \qquad \sum_{c : v \in c} x_c \leq 1 \quad \forall v \in V$$

The number of decision variables is linear in the number of cycles and chains. For constant L, there are $O(|V|^L)$ cycles in a compatibility graph, and an exponential number of possible chains. Thus, for even moderately-sized compatibility graphs, the above integer program cannot even be written in memory, much less solved. Abraham et al. [2007] developed a specialized tree search algorithm based on the branch-and-price framework for solving integer programs [Barnhart et al., 1998]; their algorithm scaled well in basic models of kidney exchange (e.g., maximum cardinality matching without chains), but is no longer scalable for present-day kidney exchanges with long chains (see discussions in Dickerson et al. [2012b], Ashlagi et al. [2012], Dickerson et al. [2013], Manlove and O'Malley [2014] and recent work by Dickerson et al. [2013], Constantino et al. [2013], Glorie et al. [2014], Klimentova et al. [2014], Anderson [2014], Anderson et al. [2015b]). A large part of this thesis focuses on *solving* more realistic adaptations of this model using novel, optimal search algorithms.

In the next sections, we describe three adaptations to the standard barter exchange model, summarize completed work in each dimension, and then briefly overview promising proposed research directions.

2.2 Dimension #1: Edge Failure

Algorithmic matches in fielded kidney exchanges do not typically result in an actual transplant. We address the problem of cycles and chains in proposed matches failing *after* the matching algorithm has committed to them. We propose to move away from the deterministic clearing model used by kidney exchanges today into a probabilistic model where the input includes failure probabilities on possible planned transplants, and the output is a transplant plan with maximum *expected* value.

The probabilistic approach has recently also been suggested by others [Chen et al., 2012, Li et al., 2011].[10] They used a general-purpose integer program solver (Gurobi) to solve their optimization models. We show that general-purpose solvers do not scale to today's real kidney exchange sizes. Then we develop a custom branch-and-price-based [Barnhart et al., 1998] integer program solver specifically for the probabilistic clearing problem, and show that it scales dramatically better. We show that failure-aware kidney exchange can significantly increase the expected number of lives saved (i) in theory, on random graph models; (ii) on real data from kidney exchange match runs between 2010 and 2014; (iii) on synthetic data generated via a model of dynamic kidney exchange. We show experimentally that taking failed parts from an initial match and instantaneously *rematching* them with other vertices still in the waiting pool can result in significant gains. Finally, we show that failure-aware matching can increase overall system efficiency and simultaneously increase the expected number of transplants to *highly-sensitized* patients, in both static and dynamic models.

2.2.1 Main results

In Dickerson et al. [2013], we build on the basic barter exchange model (§2.1). Associate with each edge $e = (v_i, v_j)$ in the graph G a value $q_e \in [0, 1]$ representing the probability that, if algorithmically matched, the patient of v_j would successfully receive a kidney from v_i's donor. We will refer to q_e as the *success probability* of the edge, and $1 - q_e$ as the *failure probability* of the edge. Using this notion of failure probability, we can define the expected (failure-discounted) utility of chains, cycles, and a full matching.

Discounted utility of a cycle. For any transplant in a k-cycle to execute, each of the k transplants in that cycle must execute. Put another way, if even one algorithmically matched transplant fails, the entire cycle fails. Thus, for a k-cycle c, define the *discounted utility* $u(c)$ of that cycle to be:

$$u(c) = \left[\sum_{e \in c} w_e\right] \cdot \left[\prod_{e \in c} q_e\right]$$

That is, the utility of a cycle is the product of the sum of its constituent weights and the probability of the cycle executing. The simplicity of this calculation is driven by the required atomicity of cycle execution—a property that is not present when considering chains.

Discounted utility of a chain. While cycles must execute entirely or not at all, chains can execute incrementally. In general, for a k-chain $c = (v_0, v_1, \ldots, v_k)$, where v_0 is an altruist, there are k possible matches (and the final match to, e.g., a deceased donor waiting list candidate). Let q_i be the probability of edge $e_i = (v_i, v_i + 1)$ leading to a successful transplant. (Here, we assume $w_e = 1$ for ease of exposition; relaxing this assumption does not complicate matters.)

[10] Some exchanges and models also handle edge failure in heuristic, ad-hoc ways (e.g., by lexicographically favoring short cycles over long cycles, by capping chains, or similar heuristic approaches [Keizer et al., 2005, Kim et al., 2007, De Klerk et al., 2010b,a, Dickerson et al., 2012b, Glorie et al., 2014, Manlove and O'Malley, 2014]).

Then, the expected utility $u(c)$ of the k-chain c is:

$$u(c) = \left[\sum_{i=1}^{k-1}(1-q_i)i\prod_{j=0}^{i-1}q_j\right] + \left[k\prod_{i=0}^{k-1}q_i\right]$$

The first portion above calculates the sum of expected utilities for the chain executing exactly $i = \{1, \ldots, k-1\}$ steps and then failing on the $(i+1)^{th}$ step. The second portion is the utility gained if the chain executes completely.

Discounted utility of a matching. The value of an individual cycle or chain hinges on the interdependencies between each specific patient and potential donor, as was formalized above. However, two cycles or chains in the same matching M fail or succeed independently. Thus, define the discounted utility of a matching M to be:

$$u(M) = \sum_{c \in M} u(c)$$

Theoretical results

Let $G(n, t(n), p)$ be a random graph with n patient-donor pairs, $t(n)$ altruistic donors, and probability $p = \Theta(1/n)$ of incoming edges (similar to the sparse model of Ashlagi et al. [2012], discussed earlier in Section 2.1.2). Such a p represents highly-sensitized patients. Let q be the probability of transplant success that we introduced, such that q is constant for each edge e. Note that for a chain of length k, the probability that $t < k$ matches execute is $q^t(1-q)$, and the probability that k matches execute is q^k. There is no chain cap (although we could impose one, which depends on q). Given a matching M, let $u_q(M)$ be its expected utility in our discounted utility model, i.e., expected number of successful transplants. Denote the set of altruistic donors by A, and denote the vertex pairs by P.

Theorem 1. *For every constants $q \in (0,1)$ and $\alpha, \beta > 0$, given a large random graph $G(n, \alpha n, \beta/n)$, with high probability there exists a matching M' such that for every maximum cardinality matching M,*

$$u_q(M') \geq u_q(M) + \Omega(n).$$

In words, this states that given any maximum matching, almost always there exists a *smaller* matching that will result in substantially more matches than the maximum matching.

Algorithmic results

Since the undiscounted clearing problem is a special case of the discounted clearing problem—that is, it is the discounted clearing problem with constant success probability $q = 1.0$—it follows that the discounted clearing problem is also NP-hard.

Proposition 1. *The discounted clearing problem is NP-hard.*

Including edge failure probabilities in the matching model increases the exchange's performance, according to Theorem 1. Unfortunately, including this extra parameter breaks the current branch-and-price-based integer programming solver (introduced in Section 2.1.3). In integer programming, a tree search that branches on each integral decision variable is used to search for an optimal solution. At each node, upper and lower bounds are computed to help prune subtrees and speed up the overall search. In practice, these bounding techniques are critical to proving optimality without exhaustively searching the space of all assignments. The current solver uses as an upper bound at each node the objective value of the unrestricted clearing problem, which is solvable in polynomial time by encoding the pool into a weighted bipartite graph and computing the maximum weighted perfect matching (see reduction by Abraham et al. [2007]). Unfortunately, Proposition 2 shows we can no longer use this:

Proposition 2. *The unrestricted discounted maximum cycle cover problem is NP-hard.*

Fortunately, the current lower bound (matching via 2-cycles only) is still tractable in this new model, as shown in Proposition 3.

Proposition 3. *The discounted clearing problem with cycle cap $L = 2$ is solvable in polynomial time.*

The current UNOS solver uses an incremental formulation called column generation to bring only some variables into the search tree at each node. The basic idea behind column generation is to start with a reduced model of the problem, and then incrementally bring in variables (and their constraints) until the solution value of this reduced model is provably the solution value of the full (implicitly represented) model. This is done by solving the *pricing problem*, which associates with each variable a real-valued price such that, if any constraint in the full model for a variable c is violated, then the price of that variable is positive. In our case, the *price* of a cycle or chain c is just the difference between the discounted utility $u(c)$ and the dual value sum of the vertices in that cycle or chain.

When no positive price cycles exist, we have proved optimality at this node in the search tree. Proving this is hard, since the solver might have to consider each cycle and chain. We now present a method for "cutting off" a chain after we know its expected utility is too low to improve the reduced problem's objective value.

We use the additional of failure probabilities into our optimization model as a way to generate *fewer* variables than in the deterministic case. An *upper bound* on the expected utility of a (possibly infinite) chain c', extended from some base k-chain $c = (v_0, v_1, \ldots, v_k)$, is given in Equation (2.1) below.

$$u(c^\infty) - u(c) =_{k \to \infty} \frac{q_{max}}{1 - q_{max}} \prod_{i=0}^{k-1} q_i \tag{2.1}$$

We are interested in using this computed value to stop extending c. This is done via Proposition 4.

Proposition 4. *Given a k-chain c, if the infinite extension c^∞ is not promising (i.e., Equation 2.2 holds), then no finite extension is promising, either.*

$$\left(\frac{q_{max}}{1 - q_{max}} \prod_{i=0}^{k-1} q_i \right) + u(c) + \ell - \left(d_{min} + \sum_{i=0}^{k} d_i \right) \leq 0 \tag{2.2}$$

Here, ℓ is the utility derived from the final donor in a chain donating his or her kidney to the deceased donor waitlist.

In the full thesis, we will discuss a variety of heuristics for generating cycles and chains and implementation details (some of which were presented by Dickerson et al. [2013]).

Experimental results

The full experimental results from Dickerson et al. [2013] will be included in the completed thesis, but we give two examples comparing our new model and solver against the status quo deterministic kidney exchange solver in practice.

	CPLEX		Ours		Ours without chain curtailing	
$\lvert V \rvert$	Solved	Time (solved)	Solved	Time (solved)	Solved	Time (solved)
10	127 / 128	0.044	128 / 128	0.027	128 / 128	0.052
25	125 / 128	0.045	128 / 128	0.023	128 / 128	0.049
50	105 / 128	0.123	128 / 128	0.046	125 / 128	0.057
75	91 / 128	0.180	126 / 128	0.072	123 / 128	0.066
100	1 / 128	1.406	121 / 128	0.075	121 / 128	0.071
150	0 / 128	–	114 / 128	0.078	95 / 128	0.098
200	0 / 128	–	113 / 128	0.135	76 / 128	0.096
250	0 / 128	–	94 / 128	0.090	48 / 128	0.133
500	0 / 128	–	107 / 128	0.264	1 / 128	0.632
700	0 / 128	–	115 / 128	1.071	0 / 128	–
900	0 / 128	–	38 / 128	2.789	0 / 128	–
1000	0 / 128	–	0 / 128	–	0 / 128	–

Table 2.1: Scaling results for our method versus CPLEX, timeout of 1 hour.

To show the scalability of our method, we compare our novel solver against CPLEX 12.2 (2010), a recent version of a state-of-the-art integer linear programming solver. Since CPLEX does not use branch-and-price, it must solve the full integer program (with one decision variable per possible cycle and chain). This comparison is shown in Table 2.1.

To demonstrate the efficacy of our method, we compare matching results for determinstic matching versus our method on two edge failure distributions (derived from real data; full details will be in the thesis [Dickerson et al., 2013]). Figure 2.1 and 2.2 show clear gains when matching under this new model, supporting the result of Theorem 1.

2.2.2 Proposed research

In the failure-aware model (§2.2.1), we explicitly consider a failure probability for each edge in the graph. In many applications (including kidney exchange), additional testing can be performed on a per-edge basis to either reduce or remove the uncertainty of the true existence of that edge. For example, in kidney exchange, blood samples from a potentially compatible pair of one patient and one donor are compared to further determine feasibility through a process called *crossmatching*. In fielded kidney exchanges, variable levels of blood- and tissue-type testing across participating centers in a multi-center exchange lead to varying estimates of edge success probability. This variance is a product of (i) different features being considered by different wet laboratories, (ii) different quality levels of the wet laboratories, and (iii) other factors like time between testing and matching and changes in the patients' or donors' underlying health status. We propose to consider *edge pre-testing* in the failure-aware model.

Indeed, a basic form of this model has been studied in the standard matching literature as the *query-commit* problem [Chen et al., 2009, Adamczyk, 2011, Goel and Tripathi, 2012, Costello et al., 2012, Bansal et al., 2012, Blum et al., 2013], where the existence of an edge can be independently queried, but (in most of this work) any edge that is queried and exists must be included in the final matching. More related to the general stochastic barter exchange problem is our work in Blum et al. [2015], which looks at adaptively testing edges over a number of rounds and then either choosing a maximum matching or maximum k-set packing among these queried edges. While we are not including the theoretical results of Blum et al.

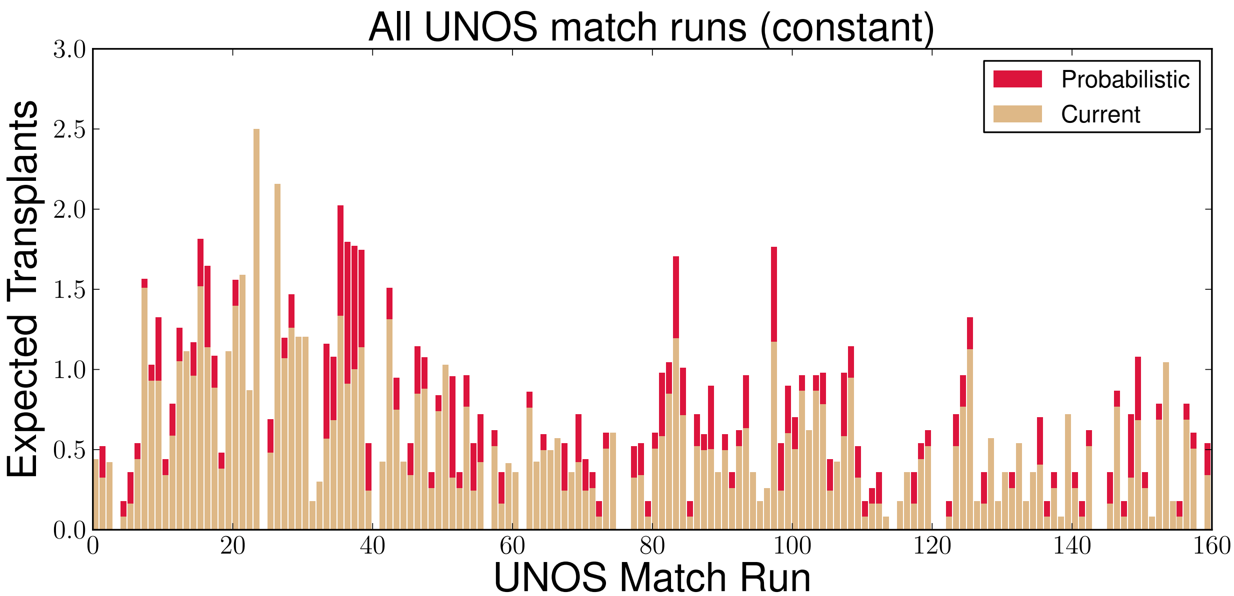


Figure 2.1: Comparison of the expected number of transplants resulting from the maximum weighted matching and discounted weighted matching methods, on 161 UNOS match runs between October 2010 and November 2014, with a constant edge success probability.

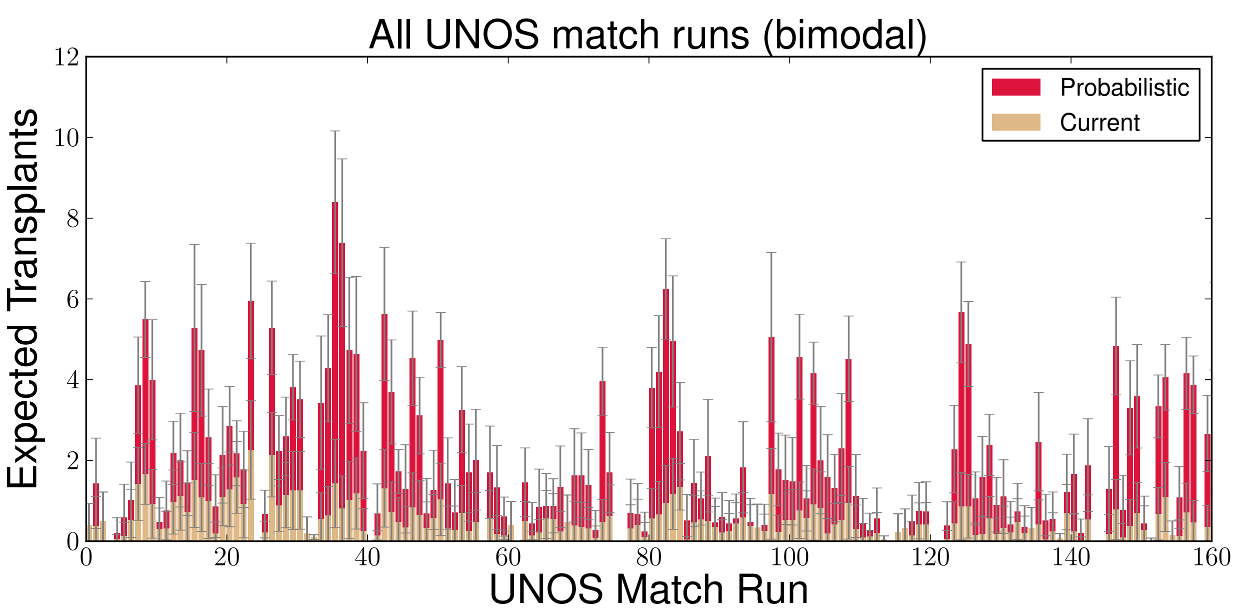


Figure 2.2: Comparison of the expected number of transplants resulting from the maximum weighted matching and discounted weighted matching methods, on 161 UNOS match runs between October 2010 and November 2014, with bimodal edge success probabilities (some very high, some very low).

[2015] in this thesis, we showed empirically that even performing R "rounds" of testing at most one incident edge per vertex in realistic kidney exchange graphs can greatly increase successful matching, as shown in Figure 2.3.

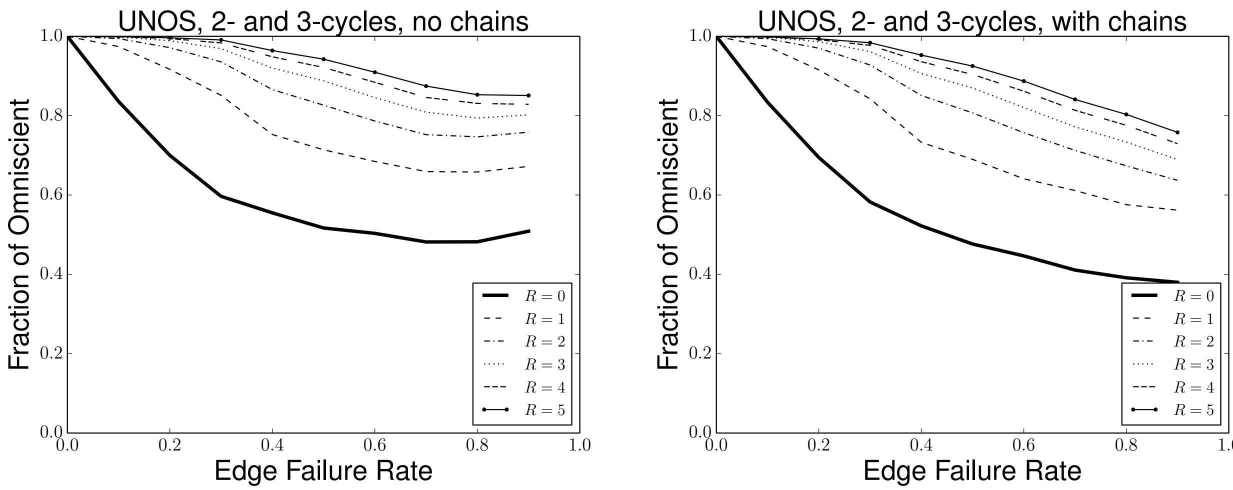


Figure 2.3: Real UNOS match runs with 2- and 3-cycles and no chains (left) and with chains (right).

Prior work assumes either no budget or a per-vertex limit on the number of incident edges to test; in this thesis, we propose to build on the query-commit literature and Dickerson et al. [2013] by considering the case where a clearinghouse has an overall budget to spend on testing edges, such that spending more of that budget on an edge increases the accuracy of an estimate of whether or not that edge exists. Indeed, in kidney exchange, it is sometimes hard to predict the probability of an edge existing at all (some preliminary work has been done by Glorie [2012], but it is still very coarse). It may be the case that spending 100% of the clearinghouse's budget on a single vertex is substantially better than spreading that budget out across multiple vertices (possibly because that single vertex has multiple incoming edges with low existence probability, but somehow triggers a large objective increase if included successfully in the final matching). We will also look at a related formulation of this problem, where the clearinghouse is assumed to have infinite budget, but the objective maximizes some function of overall match quality minus the budget spent. We propose to explore the theoretical and practical ramifications of these models.

We note that this line of thinking could also inform two-stage models like those due to Anderson [2014], Anderson et al. [2015b]. We will think about this.

2.3 Dimension #2: Equity

In general organ transplantation, some patients are *highly-sensitized*; there is a very low probability that their blood will pass a crossmatch test with a random organ. For these patients, finding a kidney is quite difficult (and median time on the waiting list jumps by a factor of three over less sensitized patients UNOS).

Roughly 17% of the adult patients on the waiting list for *deceased* donor kidneys are highly-sensitized [HHS/HRSA/HSB/DOT, 2011]. Recently, a data-driven allocation policy was designed for deceased donor kidneys that effectively balances fairness and efficiency while working within the currently fielded priority-based framework [Bertsimas et al., 2013]. The percentage of highly-sensitized patients in fielded kidney exchanges is quite high; roughly 60% of the UNOS nationwide kidney exchange is highly-sensitized, as shown in Figure 2.4.

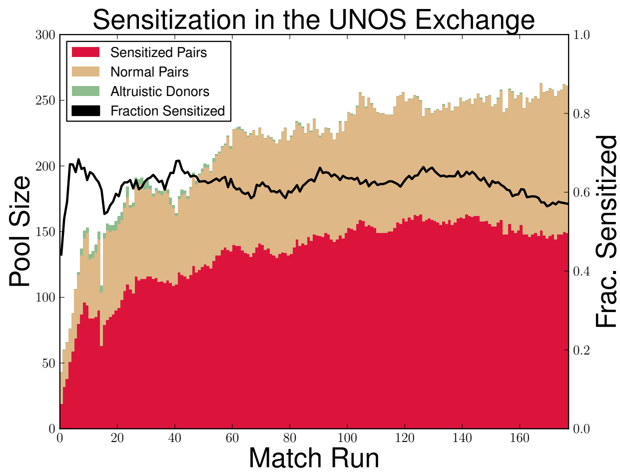


Figure 2.4: Composition of the UNOS national kidney exchange from inception through early 2015. For each of 177 match runs (x-axis), the raw number of highly-sensitized patients, non-highly-sensitized patients, and altruists are plotted (left y-axis), as well as the percentage of patients who are highly-sensitized as a percentage of the pool size (right y-axis).

Furthermore, in some exchanges, edge failure probability is correlated with level of sensitization [Ashlagi et al., 2011, Glorie, 2012]. Thus, the failure-aware matching described in Section 2.2 may further marginalize these already marginalized patients. In this section, we theoretically and experimentally explore the interactions of efficiency and fairness in kidney exchange. (The methodology is general to other barter exchanges and matching markets.) Our current results pertain almost exclusively to fairness with respect to highly-sensitized patients; part of our proposed work is to generalize those results.

2.3.1 Main results

In Dickerson et al. [2014c], we explore the *price of fairness* in kidney exchange—the relative loss in total welfare from using a "fair" matching rule, instead of an overall utility-maximizing one [Bertsimas et al., 2011, Caragiannis et al., 2009]. Theoretically, we show that the price of fairness is small in the standard theoretical kidney exchange model. We then formally define two natural definitions of fairness in the standard deterministic and in our new failure-aware (§2.2) models of kidney exchange. We quantify through computational experiments the tradeoff between efficiency and fairness on UNOS data, as well as on simulated data from the two most widely used kidney exchange distributions. We find that, on real data, prioritizing hard-to-match patients results in a price of fairness that is (often quite far from) zero. We also find that using the failure-aware model, even with strict fairness preferences, results in more transplants than fully efficient matching in the deterministic model.

Theoretical results

Let $u_f : \mathcal{M} \to \mathbb{R}$ be a *fair* utility function. Formally, a utility function is fair when its corresponding optimal match M_f^* is viewed as fair, where M_f^* is defined as:

$$M_f^* = \arg\max_{M \in \mathcal{M}} u_f(M)$$

Given a fair utility function u_f and the utilitarian utility function u, the price of fairness is defined to be:

$$\mathsf{POF}(\mathcal{M}, u_f) = \frac{u\left(M^*\right) - u\left(M_f^*\right)}{u\left(M^*\right)}$$

That is, $\mathsf{POF}(\mathcal{M}, u_f)$ is the relative loss in match efficiency (from the utilitarian point of view u) due to the maximization of a fair utility function u_f over some family of matchings $\mathcal{M}$.

We derive a theoretical upper bound on the price of fairness under the fair utility function $u_{H \succ L}$ that lexicographically ranks any highly-sensitized vertex (in the set $V_H \subseteq V$ over any lowly-sensitized vertex (in the set $V_L \subseteq V$). For any matching $M \in \mathcal{M}$, let $M_H = M \cap V_H$ be the subset of highly-sensitized vertices matched by M. Formally:

$$u_{H \succ L}(M) = \begin{cases} u(M) & \text{if } |M_H| = \max_{M' \in \mathcal{M}} |M'_H| \\ 0 & \text{otherwise} \end{cases}$$

This utility function gives nonzero weight only to those matches that include the maximum possible number of highly-sensitized patients. We informally argue that price of fairness guarantees on $u_{H \succ L}$ are upper bounds to the price of fairness of any "reasonable" fair utility function. Indeed, any utility function that does not first maximize the number of highly-sensitized pairs matched will tend to leave a thicker remaining market in which non-highly-sensitized pairs have more options for matching—and thus the resulting match will see less of an efficiency loss.

Proposition 5 gives a bound on the price of fairness in random graphs in the ABO-model (described in Section 2.1.2) parameterized in a realistic way (i.e., $\bar{p}$ mirrors that of a dense kidney exchange pool, and the blood type distribution mimics that of the US population). Figure 2.5 gives the "fair" allocation described by Proposition 5.

Proposition 5. *Assume that $\bar{p} < 2/5$, $\mu_O < 3\mu_A/2$, and $\mu_O > \mu_A > \mu_B > \mu_{AB}$, and $\lambda \geq (1 - \bar{p})$. Denote by $\mathcal{M}$ the set of matchings in $G(n)$. Then, almost surely as $n \to \infty$,*

$$\mathit{POF}(\mathcal{M}, u_{H \succ L}) \leq \frac{2}{33}.$$

(And this is achieved using only cycles of length at most 3.)

Algorithmic results

In Dickerson et al. [2014c], we present two definitions of fairness in kidney exchange—one using strict lexicographic preferences and the other using a sliding scale weighting function. Critically, we feel both definitions fit within the scope and practice of current policy in fielded exchange, a necessary consideration when fielding new technology in medicine (as noted by those who designed the recent deceased donor allocation scheme Bertsimas et al. [2013] and supported by our experience with the UNOS exchange).

In the interest of space, we leave the full definitions and methods for clearing exchanges under either rule in Dickerson et al. [2014c], but will include them in the full thesis. At a high level, the *lexicographic rule* $u^{\alpha}_{H \succ L}$ assigns 0 value to any matching M containing less than an α fraction of the absolute maximum number of highly-sensitized patients that could be included in any match, and then $u(M)$ utility to those matchings that do satisfy this global constraint (where u is the utilitarian objective defined in Section 2.1). The *weighted rule* takes as input some re-weighting function $\Delta : E \to \mathbb{R}$. Then, for any $M \in \mathcal{M}$, let M' be the matching such that every cycle $c \in M$ has augmented weight $v_\Delta(c)$, where $v_\Delta(c)$ is the value of a cycle

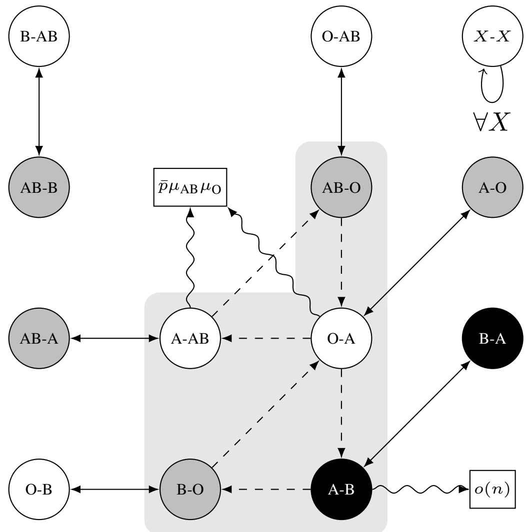


Figure 2.5: An example matching used in Proposition 5. Patient-donor pairs are ovals: under- and self-demanded pairs are white, over-demanded pairs are gray, and reciprocal pairs are black. Regular edges appear in the efficient matching, while dashed edges represent 3-cycles from the efficient matching that may be disturbed via fair matching. Efficiency loss is denoted with rectangular nodes.

or chain c such that the weight of each edge $e \in c$ is adjusted by some re-weighting function $\Delta : E \to \mathbb{R}$. Then define the weighted rule u_Δ in terms of the utilitarian rule u applied to the augmented matching M', such that $u_\Delta(M) = u(M')$.

In terms of implementation, adding lexicographic preferences into the optimization model breaks the specific branch-and-price structure on which this solver relies. While a solver could still use column generation or general branch-and-price to solve this new problem, the addition of a matching-wide constraint—that a matching must contain cycles containing some fraction of a set of marked vertices—makes solving the pricing problem (see Barnhart et al. [1998] for details) much more difficult than in the utilitarian formulation, where determining the price of a cycle not included in the current subproblem is relatively simple. Indeed, with such an allocation-wide constraint, finding a positive price cycle at a node in the search tree requires solving an integer program, whereas current solvers can use a (typically quite fast) depth-first search to find a positive price cycle in the standard kidney exchange model.

However, unlike implementing the lexicographic fairness rule $u^\alpha_{H \succ L}$, this definition of fairness does not break the branch-and-price structure on which current scalable kidney exchange solvers rely. Indeed, the u_Δ rule, for simple re-weighting functions like the multiplicative example above, can be implemented by first preprocessing a compatibility graph using Δ to determine edge weights, and then solving the maximization problem using a standard kidney exchange solver.

Experimental results

In ongoing experiments, we are comparing efficiency versus fairness in various parameterizations of both these rules, on real and simulated data. While more experiments are in the work by Dickerson et al. [2014c] and in the journal version of our EC-2013 paper [Dickerson et al., 2013], we include a sample of results here.

Figure 2.6 shows CDFs of efficiency loss for the first 73 UNOS match runs under the strict lexicographic rule for three different edge failure probability distributions. Figure 2.7 shows Pareto frontiers increasing the weight given to sensitized transplants, again under three edge failure probability distributions. It uses a multiplicative re-weighting function

$$\Delta^{\beta}(e) = \begin{cases} (1+\beta)w_e & \text{if } e \text{ ends in } V_H \\ w_e & \text{otherwise} \end{cases}$$

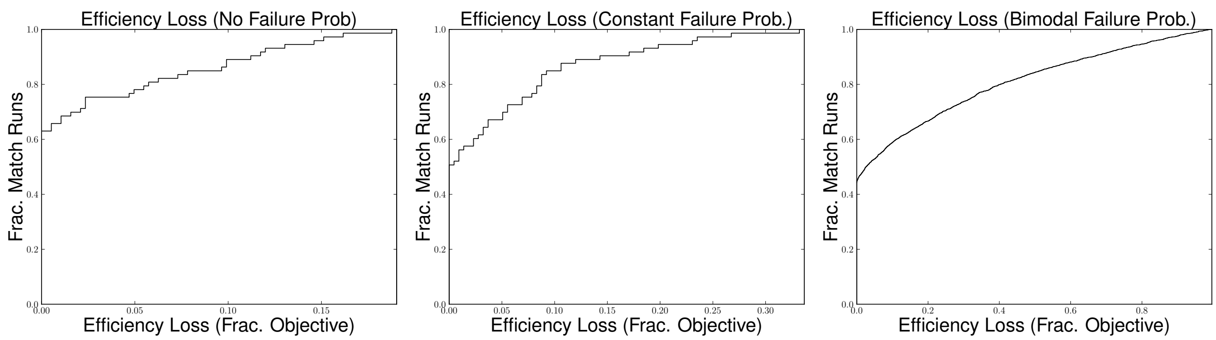


Figure 2.6: Cumulative distribution functions of the price of fairness under the lexicographic fairness rule $u^{1.0}_{H \succ L}$ according to UNOS' weighting policy, on 73 UNOS match runs since the inception of the exchange.

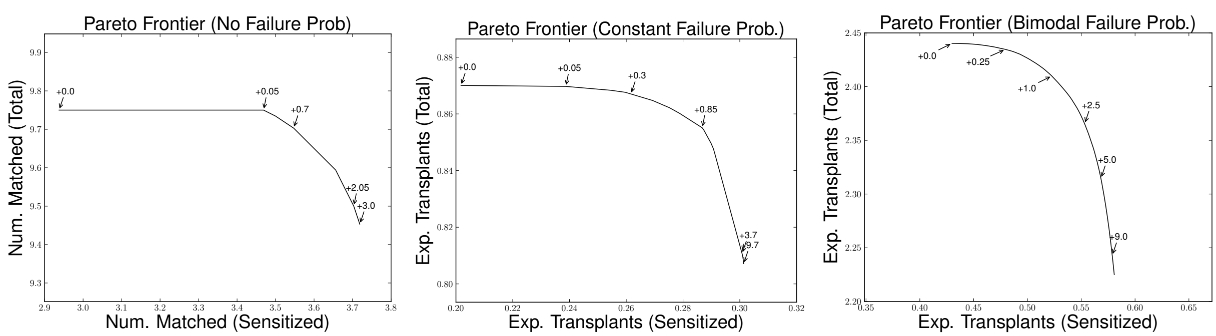


Figure 2.7: Pareto frontiers for u_{Δ} under different failure probability distributions, for $\beta \in \{0.0, 0.05, \ldots, 10.0\}$.

One might wonder, under edge failure probabilities that may or may not be correlated with vertex marginalization, how (or even if) failure-aware and fairness-aware matching can be combined in such a way that results in both more overall and more marginalized transplants than the status quo deterministic matching employed by most fielded exchanges. Many exploratory results are available in the submitted journal version of Dickerson et al. [2013], but an example result is given below as Figure 2.8. The left graph has uncorrelated failure probabilities (from published UNOS data [Kidney Paired Donation Work Group, 2013]), while the right has correlated failure probabilities (from published APD data [Ashlagi et al., 2011]).

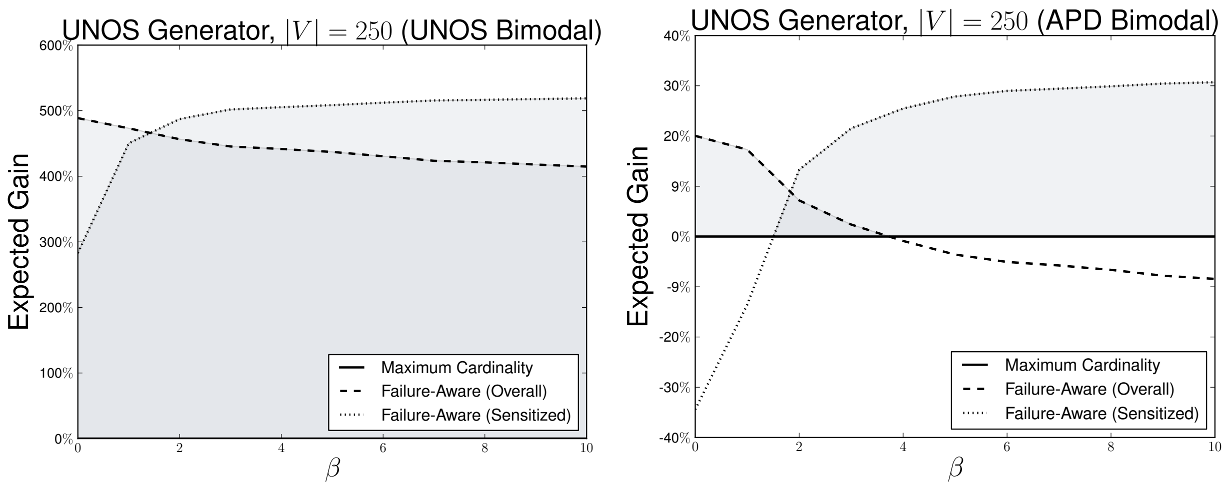


Figure 2.8: Change in the expected number of transplants on average for *generated* UNOS match runs when using failure-aware matching instead of maximum cardinality matching, assuming bimodal edge failure rates derived from UNOS (left column) and APD (right column). The x-axis varies the β fairness factor applied to the failure-aware matching algorithm.

2.3.2 Proposed research

Artificial Intelligence invited us to submit an extended version of our AAMAS-2014 paper [Dickerson et al., 2014c] to their journal. We plan to do this after extending the price of fairness result of Proposition 5 to the cases of:

- The dense ABO-model of kidney exchange, but with chains; Proposition 5 considers only the case of cycles (of any length, and shows that only those of length up to 3 are needed, as in the results of Ashlagi and Roth [2011, 2014]). We plan to do this by building off of the dense, efficient matching we presented in Dickerson et al. [2012b].
- The parameterized sparse model of kidney exchange defined in Ashlagi et al. [2012], for more than just $p = \Theta(1/n)$. Our hunch is that the price of fairness is extremely low in this model.

We would also like to extend our equity considerations beyond the simple binary function of "is this vertex sensitized or not?" Our weighted fairness function extends nicely toward this end; indeed, it could be used to mimic the current "priority points" system used at the UNOS exchange. Still, it is likely difficult—or at least complicated—to analyze theoretically; we would like to explore different definitions of fairness in theory-friendly ways that still mimic reality.

2.4 Dimension #3: Dynamism

Currently, fielded kidney exchange algorithms typically match patients to donors in a myopic fashion, maximizing the number of candidates who get kidneys (usually on a daily, weekly, monthly, or bimonthly basis) in an offline fashion. However, this is sub-optimal, since patients and donors arrive and leave the pool over time. Recent work shows that more candidates can be matched by considering the future [Awasthi and Sandholm, 2009, Ünver, 2010, Ashlagi et al., 2013, Anshelevich et al., 2013], but those approaches are either overly simplified or do not scale.[11] In this section, we explore *computationally tractable* methods for learning to match in a dynamic environment.

Note. While we focus on kidney exchange in this work, the dynamic matching problem in *deceased* organ allocation has been studied before. That problem is fundamentally different than kidney exchange, but we briefly review that work here for completeness. Su and Zenios [2005] study the online allocation problem from a theoretical point of view, where patients and kidneys have types, kidneys arrive over time, kidney types are not known ahead of time, and patients much choose to accept or decline kidney offers. Bertsimas et al. [2013] propose a data-driven allocation policy for the deceased-donor kidney wait list; they work within today's point-based allocation framework, and balance efficiency over time with various notions of fairness. Alagoz et al. [2007] formulate the organ acceptance or rejection problem for the deceased-donor liver wait list as a Markov decision process and study its properties using real data, while Akan et al. [2012] derive optimal policies for the liver wait list under a variety of settings.

2.4.1 Main results

In Dickerson et al. [2012a], we introduce a method for informing myopic optimization about the future in dynamic problems. It automatically learns *potentials* of elements of the problem (e.g., of vertices or edges in a graph) offline, and then uses these potentials to guide myopic optimization at run time. The potential represents an estimate of how much that element can contribute to the objective in the future. The potentials can be viewed as policy parameters, so they are a natural, fairly general way of parameterizing policies. They can be optimized using a black box program; we are able to leverage state-of-the-art automated algorithm configuration tools to learn their values. Then, at run time, we simply run an offline optimization algorithm at each time period, but subtracting out in the objective the potentials of the elements used up in the solution. This causes the batch optimizer—which is traditionally myopic—to take the future into account without suffering a run-time cost.

Estimating the potential—roughly, some quantification of the ability to "help" in the future—of a structural element of a problem can be viewed through the lens of the reinforcement learning/Markov decision processes (MDP) as a policy search process (e.g., [Bradtke and Barto, 1996, Lagoudakis and Parr, 2003]). For some problems (like the kidney exchange problem), however, estimating the value of a state is incredibly difficult, so we turn to different learning methods, as discussed below.

As a concrete example within the kidney exchange framework, consider potentials on vertices. In terms of ABO compatibility, an O-type donor can give to O-, A-, B-, and AB-type patients, an A-donor can give to A or AB, a B-donor can give to B or AB, and an AB-donor can only give to AB. Therefore, an O-type altruistic donor typically leads to more matched patients (i.e., has higher potential) than other types of altruist. Similarly, an altruist with a given blood type tends to have higher potential than a patient-donor vertex with that same donor type because the pair is harder to match, since it expects a kidney in return and that kidney needs to be compatible with the pair's patient. We describe how these potentials are quantified and learned later in the section.

Given the potential for a vertex, edge, or cycle type (where type is defined by the ABO blood types of the donors and the patients in the vertices), we can easily translate this information into a language the matching algorithm understands. For example, with vertex potentials and an initial maximum cardinality objective, the translation works as follows. Given vertex potentials P_X and P_Y representing the potentials of patient-donor pairs of ABO type X and Y, respectively, any edge e between vertices of type X and Y

[11] There is ongoing research in matching in theoretical models that both supports and does not support the claim that non-greedy matching can result in gains over simple greedy matching [Akbarpour et al., 2014, Anderson et al., 2015a]. In work that will not be part of this thesis, we are exploring similar theoretical models with competition between multiple exchanges.

receives weight $f(P_X, P_Y)$, for some function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (alternatively, one can think of $f : E \to \mathbb{R}$ in the vertex or edge potential cases). Cycles in the exchange are then assigned weights as usual, as the sum of their edge weights. In this way, the potentials assigned to specific elements affect the final maximum-weight exchange of disjoint cycles.

Theoretical results

There is a natural tradeoff between the expressive power of a family of potentials and the complexity of learning over that increased space. We prove bounds on the best-case benefit from associating potentials on larger elements compared to associating them with smaller elements. First, in Theorem 2 we compare the application of potentials to vertices and to edges and show that edge potentials can do notably better. Second, Theorem 3 considers allowing potentials to be applied to edges and cycles; again, we show that the finer-grained resolution of cycles can allow better overall results than just edges. Finally, Theorem 4 shows that, in certain pathological cases, even applying potentials to cycles can perform poorly compared to potentials on unlimited graph elements (which is equivalent to comparing against an omniscient algorithm with perfect foresight).

Theorem 2 (Vertex potentials vs edge/graph potentials)**.**

1. *For every $k \in \mathbb{N}$ there exists an input with cycles of length at most $2k + 4$ and no chains such that for any choice of vertex potentials the number of matched patients is at most a $(4k+4)/(6k+4)$-fraction of the optimum.*
2. *For every $k \in \mathbb{N}$ there exists an input with 2-cycles and chains of length at most $2k + 5$ such that for any choice of vertex potentials the number of matched patients is at most a $(4k+5)/(6k+5)$-fraction of the optimum.*
3. *There exists an input with cycles of length at most 2 and no chains such that for any choice of vertex potentials the match size is at most a $(5/6)$-fraction of the optimum.*

In each of the three cases, the construction is such that the optimum is achievable using edge potentials.

Theorem 3 (Edge potentials vs cycle/graph potentials)**.**

1. *For every $k \in \mathbb{N}$ there exists an input with cycles of length at most $3k + 2$ and no chains such that for any choice of edge potentials the number of matched patients is at most a $(3k + 2)/(6k + 2)$-fraction of the optimum.*
2. *For every $k \in \mathbb{N}$ there exists an input with 2-cycles and chains of length at most $3k + 3$ such that for any choice of edge potentials the number of matched patients is at most a $(3k + 3)/(6k + 3)$-fraction of the optimum.*

In both of these cases, the construction is such that the optimum is achievable using cycle potentials.

Theorem 4 (Cycle potentials vs graph potentials)**.** *Denote by L the cap on cycle length. There exists an input with cycles of length at most L (even with no altruistic donors) such that for any choice of cycle potentials the number of matched patients is at most a $1/L$-fraction of the optimum.*

Algorithmic results

In this section, we describe our technique for learning potentials for elements of kidney exchange. First, we select a set Θ of features representing different element types in the pool. Then, for each element type θ, assign some value $P_\theta \in \mathbb{R}$ that represents the expected potential usefulness of that kind of element to the pool over time.

Continuing the ABO vertex potential example given above, there are $4 \times 4 = 16$ combinations of patient and donor blood types, and 4 possible blood types of altruists. So, we have 20 different kinds of

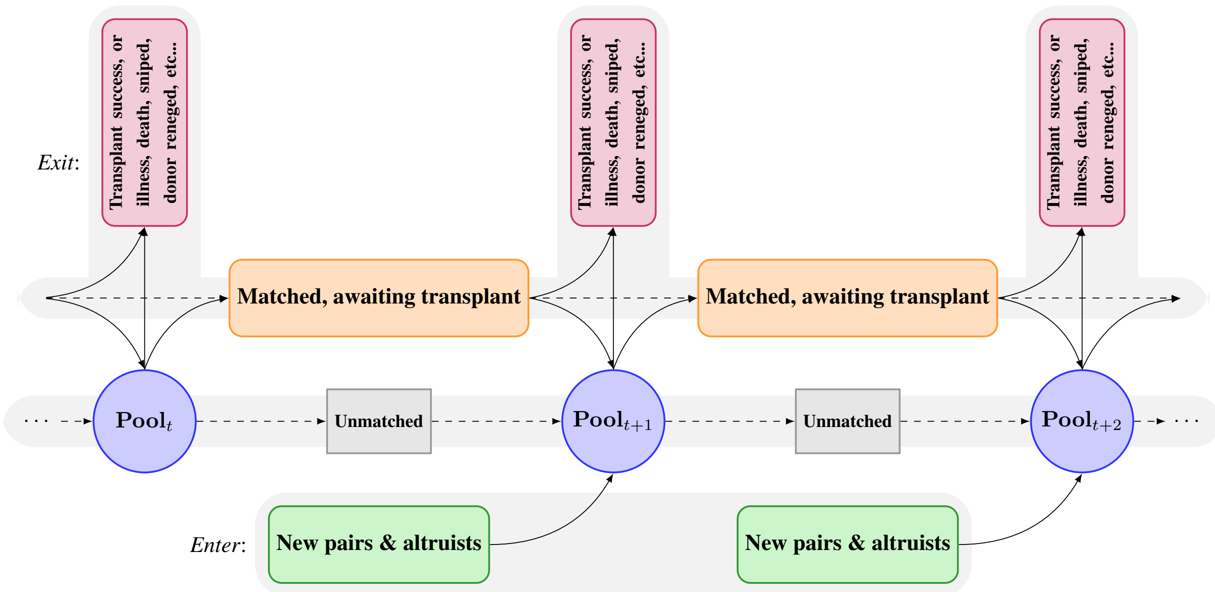


Figure 2.9: The evolution dynamics of a kidney exchange.

vertices. We want to learn a potential for each of those 20 vertex types. Formally, the types of vertex are $\Theta_{\text{ABO}} = \{\text{O-O}, \text{O-A}, \ldots, \text{AB-B}, \text{AB-AB}\} \cup \{\text{O}, \ldots, \text{AB}\}$, and we want to learn values P_θ for $\theta \in \Theta_{\text{ABO}}$.

We combine the learned potentials P_θ with a possible exogenous edge weighting function w using a function $f_w : E \to \mathbb{R}$. It balances the *myopic* value of an edge encoded by w with the *future value* of an edge encoded by potentials. The idea is that the revised weight, f_w, of an edge is its immediate value if matched minus the potentials of its vertices because if those vertices are matched now, they cannot contribute to future matches. Specifically, $f_w(e) = w(e) \cdot (1 - P_{\theta_d} - P_{\theta_p})$, where d and p are the vertices adjacent to edge e.

We use SMAC [Hutter et al., 2011], a state-of-the-art model-based algorithm configuration tool that searches through a parameter space to optimize a given objective. SMAC guides its navigation in the space of parameter vectors by constructing a model that predicts algorithm performance as a function of the parameters. It uses the model to select promising parameter vectors and tests them against the incumbent parameter vector, possibly continuing that search. In our setting, the parameter vector is the vector of potentials. At each parameter vector that SMAC navigates to, we run a large number of trials of our simulator of a kidney exchange to see how the batch matching algorithm would perform in a dynamic setting using that parameter vector. That performance number is then fed back to SMAC, and SMAC navigates to the next parameter vector to continue the search.

Experimental results

The worst-case theoretical results presented in Theorems 2, 3, and 4 are negative. However, experimentally, even learning potentials on the most restrictive class (at the vertex level) results in significant gains over myopic matching on generated kidney exchange data. First, we designed a realistic dynamic kidney exchange simulator, shown in Figure 2.9.

Figure 2.9 provides a snapshot of a compatibility graph over three points in time. The pool at time t consists of unmatched patients and donors from time $t-1$, any new pairs and altruists entering the pool, and any vertices who were waiting for a successful match, but whose match failed (due to, e.g., a positive crossmatch). Note that these patients are still formally in the pool, just marked temporarily "inactive" until the status of their pending transplant is known. At each time period t, vertices leave the pool permanently through any of many reasons, or are temporarily marked "inactive" through a pending match.

Then, we learned vertex potentials over blood types. Full experimental results are forthcoming and will be in the completed thesis, and some are already published in Dickerson et al. [2012a] and Dickerson and Sandholm [2015], but we include a sample here.

Figure 2.10 shows the improvement gained by our weighted myopic matching algorithm, relative to the difference between plain myopic matching and the optimal match size. For example, if the standard myopic algorithm matched 300 candidates, the optimal matched 360, and the weighted myopic algorithm matched 315, we report $(315 - 300)/(360 - 300) = 25\%$. Clearly, vertex potentials help. Interestingly, as the number of patient-donor pairs and altruists increases, the weighted myopic algorithm tends toward the optimal solution more quickly than standard myopic matching.

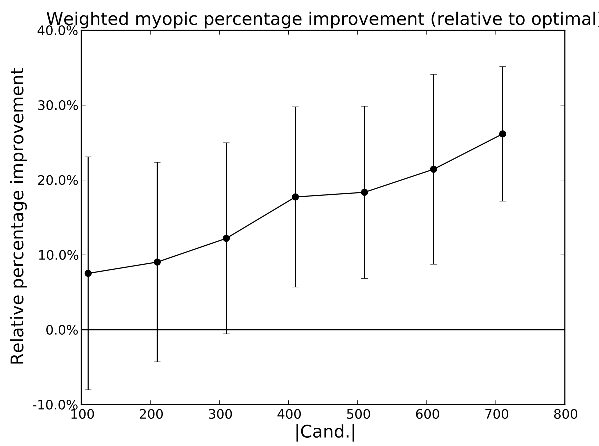


Figure 2.10: Percentage gain of weighted over unweighted myopic matching relative to the optimal match cardinality, for graphs with a variable number of patient-donor pairs and 5% as many altruistic donors. (The bars are standard deviations.)

2.4.2 Proposed research

For the full thesis, we plan to explore two future directions in the potentials framework for dynamic matching. Both are motivated by the learning method's failure to converge even on a coarse space of potentials like those induced on vertices by considering only patient and donor blood type. As the theoretical results of Section 2.4.1 showed, increasing the granularity of allowed potential types can increase the expressive power of the framework and thus, hopefully, the efficacy of the framework as well. Similarly, expanding the type of features covered by potentials in general would also increase expressive power (albeit at a computational cost to learning). The two directions of research follow.

- We propose to increase learning speed by making independence assumptions about classes of potentials. For example, in kidney exchange, it may be reasonable to assume that potentials based on blood type versus those based on patient or donor age are roughly independent, and thus values learned for a set of potentials in the former group can be applied repeatedly when learning the latter group's potential values. We aim to provide new theoretical results in this setting, and validate them experimentally.

- Presently, we do not take ordering effects within a class of potentials into account. As discussed earlier, domain knowledge may let us infer a partial ordering over potentials (e.g., O-donors are intrinsically more valuable than A- and B-donors, who are more valuable than AB-donors, in terms of the overall objective). We plan to encode ordering into the learning process. This part of the proposed research is (probably) going to be entirely experimental, but will complement the theoretical and experimental proposed work from the previous bullet point.

We are also aware of recent work by [Brown et al., 2010] on bounding values in stochastic dynamic programs. It may be possible to incorporate these into the learning process. This is exploratory work.

Chapter 3

Automatically balancing competing dimensions

In Chapter 2, we defined three dimensions dynamic matching and barter exchange—edge-failure, equity, and dynamism—that often compete with one another. There is no clearly "correct" way to balance these dimensions; rather, performance in one may be weighted more heavily than others in an application-specific manner. In this chapter, we present a matching framework that approaches the problem holistically, unifying each of these dimensions under a high-level practitioner-informed objective that is made numerically concrete by the framework. Section 3.1 presents this framework, dubbed FUTUREMATCH, as well as preliminary experimental results on UNOS exchange data (originally published in AAAI-2015 [Dickerson and Sandholm, 2015]). Section 3.2 discusses policy implications.

3.1 Combining human value judgments with optimization

We are interested in learning from demographically accurate data how to match *in the present* such that some overarching objective function is maximized over time. Scalability is important: heavy offline statistics can be computed and periodically updated, but the fielded clearing algorithm must run quickly (within minutes or at most hours). In the case of kidney exchange, the scalability of the static clearing algorithm is already a bottleneck; we designed FUTUREMATCH with this in mind.

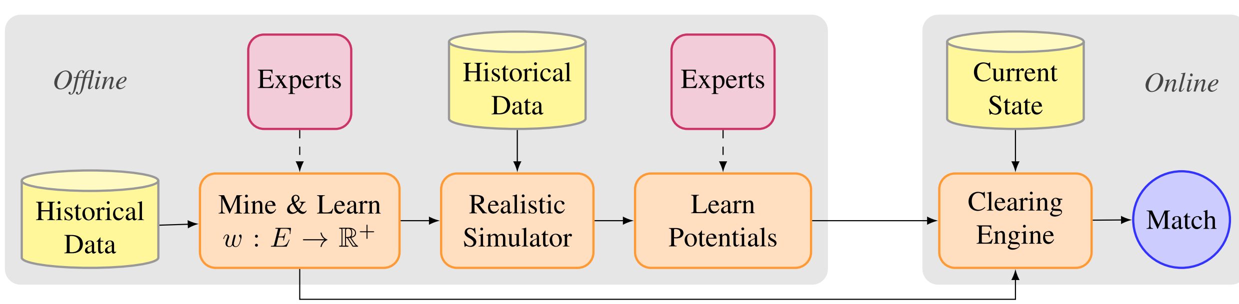


Figure 3.1: The FUTUREMATCH framework.

Figure 3.1 graphically depicts the FUTUREMATCH framework. A domain expert (e.g., a committee of medical and legal professionals) begins by describing an overall objective function for the exchange. Even *measuring* this objective can be difficult: for example, if the goal is to maximize the number of days added to patients' lives via kidney transplantation, then calculating the relative quality of a proposed match requires knowing some notion of utility for each edge—representing a potential transplant—in the compatibility graph. We propose to learn this edge weight function $w : E \to \mathbb{R}^+$ from data, and give examples for a variety of objective functions in Dickerson and Sandholm [2015], which will be included in the completed thesis.

The learned weight function w is then fed into a parameterized simulator (for our experimental results, this is a dynamic kidney exchange simulator like that described in Figure 2.9 of Section 2.4.1), calibrated by real data so that it mimics the underlying distribution. This generator in turn feeds training and test sets into a system for learning the potentials of element classes in the compatibility graph (see Section 2.4). Potentials are combined with w to quantify an edge-specific quality rating. Here, we learn potentials for the combinations of different blood types for pairs under each of the weight functions we define.

Finally, the fielded clearing algorithm incorporates the combined weight function w and set of potentials $\mathcal{P}_\Theta$ into its myopic weighted matching algorithm. This incorporation comes at very low or no cost to the runtime of the clearing algorithm; indeed, the final "potential-aware" input graph is simply a re-weighted version of the original compatibility graph, using the weights that encode the future (see Section 2.4).

In Dickerson and Sandholm [2015], we describe an in-depth implementation of FUTUREMATCH. Our goals are twofold: first, to show the general applicability and tractability of the framework, and second, to mimic a large fielded kidney exchange. Accomplishing this second goal, and leveraging our involvement with the UNOS exchange, will hopefully lead to policy recommendations, as discussed in Section 3.2.

3.1.1 Preliminary experimental results

We begin by considering two different kidney exchange models—deterministic (§2.1), where post-algorithmic match failures are not quantified in the optimization problem and failure-aware (§2.2), where they are—and three matching objectives in each of the two models:

1. MAXCARD: Maximize the total number (i.e., cardinality) of patients who are algorithmically matched (in the deterministic model) or receive transplants in expectation (in the failure-aware model);
2. MAXCARD-FAIR: Maximize the total number of patients who are algorithmically matched (in the deterministic model) or receive transplants in expectation (in the failure-aware model), where "marginalized" patients are weighted in the objective by some constant factor β more than others (§2.3); and
3. MAXLIFE: Maximize the total time algorithmically-matched (deterministic) or transplanted (failure-aware) donor organs will last in patients.

MAXCARD and MAXCARD-FAIR are formalized similarly to their respective definitions in Sections 2.1 and 2.3.1. The MAXLIFE objective is a bit more involved; we give its derivation formally below as an example of the objective creation process.

Example: Optimizing for MAXLIFE via learning to predict graft survival from data.

With the MAXLIFE objective we are interested in maximizing how long the transplanted kidneys, in aggregate, survive in the patients.[1] To do so, we must first determine an empirically sound estimate of the lifespan of a transplant as a function of donor and recipient attributes.

[1] Another objective would be to maximize aggregate increase in life duration. This would involve subtracting out the expected life duration without a transplant from the expected life duration with the transplant, and could incorporate the possibility of additional transplants after graft failure.

Delen et al. [2005] compare a variety of techniques for predicting breast cancer survivability; unlike their study, we are interested in predicting the survival *length* of a kidney graft, as opposed to whether or not a patient survives treatment at all. Data mining models are also actively being developed to predict the risk of readmission for congestive heart failure patients [Meadem et al., 2013]. Most related to our work is the Kidney Donor Profile Index (KDPI), which is currently under development by UNOS for use in the deceased donor allocation process [Kidney Transplantation Committee, 2011]. The KDPI score of a deceased donor kidney measures the estimated quality of the donor organ being allocated to the *average* recipient. In contrast, our predictor, which we will describe next, provides a unique quality score not just based on donor attributes but also based on attributes of the specific potential recipient.

We look at all 75,264 *living donor* transplant events in the US between October 1, 1987 and June 30, 2013. This data includes medical characteristics of the recipient and donor at the time of transplantation, as well as follow-up data regarding the health of the recipient and the recipient's new kidney; this follow-up data is updated at least annually.

Conditioned on a kidney graft being marked as failed in our dataset, the average graft lifetime is about 1912.7 days, or slightly over 5 years. However, due in large part to the marked increase in kidney failure since the late 1980s, nearly 75% of grafts in the dataset are not marked as failed. This occurs because either (i) the recipient is still alive with a functioning donated kidney or (ii) the recipient has died, but for a non-kidney-failure-related reason. Thus, we use *survival analysis* to estimate the lasting power of a graft.

Features of both the recipient and donor have a large effect on graft survival. For example, tissue type (HLA) testing measures the closeness of match between antigens in the cells of a donor and patient. Figure 3.2 gives a Kaplan-Meier estimator of the survival functions of (i) kidney transplants resulting from a donor and recipient being a perfect HLA match and (ii) those resulting from imperfect HLA matchings. Clearly, a kidney that is a perfect tissue type match is more desirable than an imperfectly matched one; indeed, the model estimates a median survival time of 5808 days for a perfect match compared to 4300 for an imperfect match. A log-rank test revealed that the difference between the two distributions was significant ($p \ll 0.0001$).

In our experiments, we use a Cox proportional hazards regression analysis to explore the effect of multiple features on survivability. At a high level, this method regresses the survival time of the graft against explanatory features of the donor and recipient. More specifically, define the *hazard* H at time t days after a transplant as follows:

$$H(t) = H_0(t) \times \exp(b_1 X_1 + b_2 X_2 + \ldots + b_k X_k) \tag{3.1}$$

Here, each X_i is a predictor variable corresponding to a single feature of the donor or recipient, and $H_0(t)$ is a baseline hazard rate at time t for a recipient with $X_i = 0$ for $i \in \{1, \ldots, k\}$. Then $H(t)$ represents the instantaneous risk of graft failure at time t. We want to learn this function.

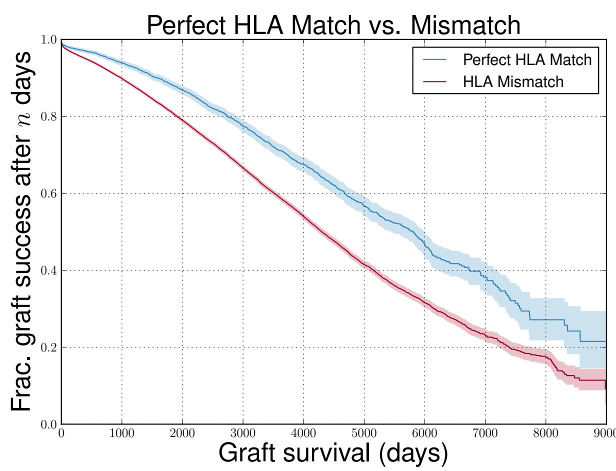


Figure 3.2: Kaplan-Meier estimator of survival functions, 95% confidence intervals.

To begin, we include the following features: recipient age, difference in donor and recipient's age, donor HLA profile, recipient HLA profile, donor and recipient blood type compatibility. The HLA profile of a donor or recipient is separated into three integral features—HLA-A, HLA-B, and HLA-DR—that can take values in $\{0, 1, 2\}$, representing 0, 1, or 2 mismatches. By separating the general HLA mismatch feature into three separate mismatch features, we complicate (but increase the power of) the model [Opelz, 1985]; this separation is motivated by evidence that mismatches at the HLA-A, -B, and -DR level have varyingly negative impact on survival.

We ran a Cox proportional hazards regression on this unpruned feature space. This used $74{,}244$ live donor transplantations during which there were $18{,}714$ graft failures (920 live donation events were dropped due to one or more missing features). Our initial regression showed that increases in the HLA-B mismatch feature did not have a significant effect on the dependent variable ($p = 0.22$). Prior research from the mid-1980s on *cadaveric* donation found a significant relationship between the combined feature of HLA-B and HLA-DR mismatches on graft outcome [Opelz, 1985]; we find that this does not hold on living donor data in the present. After selecting significant variables in this initial run—that is, all of the attributes previously discussed *except* HLA-B mismatch—we re-ran a Cox proportional hazard regression. Results are reported in Table 3.1.

feature	$\exp(b_i)$	$\mathrm{SE}(b_i)$	z	p
recipient age	1.00753	0.0008	9.715	$< 2 \times 10^{-16}$
age diff.	1.00525	0.0007	7.766	8.10×10^{-15}
HLA-A	1.05273	0.0120	4.297	1.73×10^{-5}
HLA-DR	1.08680	0.0119	6.984	2.86×10^{-12}
ABO incomp.	1.37871	0.0748	4.295	1.74×10^{-5}

Table 3.1: Learned weights via Cox regression after feature pruning for statistical significance.

Table 3.1 gives the standard error, z-score, and corresponding p-value for each of our pruned features; each clearly has a statistically significant effect on graft survival. To interpret the results, as an example first consider the HLA-DR feature. We see that $\exp(b_{\text{HLA-DR}}) \approx 1.087$; recalling Equation 3.1, a unit increase in the HLA-DR mismatch feature will result in a factor of 1.087 increase over the baseline hazard rate. Varying either recipient age or the difference in donor and recipient age was also statistically significant, with a unit increase in recipient age having a larger effect on the hazard rate. As might be expected, blood type compatibility plays the largest role in hazard rate, where an incompatible (and thus heavily immunosuppressed) transplant has a factor 1.379 increase over the base hazard rate.

Using this data, we can estimate $S_e(t)$, the survival probability at time t for a potential transplant $e \in E$ between a recipient and donor with features x_i^e, as follows: $S_e(t) = \exp\left(-H_0(t) \times \sum_i x_i^e b_i\right)$. Building on this, we define a weight function $w : E \to \mathbb{R}^+$ as $w(e) \propto \exp\left(-\sum_i x_i^e b_i\right)$. Intuitively, the weight function w assigns higher relative weight to edges with lower risk, in turn biasing the optimizer toward transplants with longer expected graft survival.

Preliminary comparison of different objectives in FUTUREMATCH

We validated FUTUREMATCH experimentally on data from the UNOS nationwide kidney exchange. We explore the effect each of the three objectives—MAXCARD, MAXCARD-FAIR, and MAXLIFE—has on a variety of metrics under FUTUREMATCH and under myopic deterministic matching, which is the fielded state of the art. The latter does not take edge failure or learned potentials into account during optimization; as described earlier (§2.1), it finds a maximum weight matching (i.e., for each chain or cycle c, $u(c) = \sum_{e \in c} w_e$) during each period separately.

	$\lvert V\rvert = 300$		$\lvert V\rvert = 400$		$\lvert V\rvert = 500$		$\lvert V\rvert = 600$		$\lvert V\rvert = 700$		$\lvert V\rvert = 800$		$\lvert V\rvert = 900$	
Total	*Gain*	p	*Gain*	p	*Gain*	p	*Gain*	p	*Gain*	p	*Gain*	p	*Gain*	p
MaxCard	+2	✓	+4	✓	+5	✓	+6	✓	+10	✓	+11	✓	+13	✓
MaxCard-Fair, $\beta = 1$	+1	✓	+4	✓	+6	✓	+8	✓	+9	✓	+11	✓	+12	✓
MaxCard-Fair, $\beta = 2$	+1		+2	✓	+3	✓	+3	✓	+5	✓	+6	✓	+10	✓
MaxCard-Fair, $\beta = 3$	+1		+0		+3	✓	+1		+1	✓	+3	✓	+2	
MaxCard-Fair, $\beta = 4$	-1		+1		+1		+1		+3	✓	+3		+2	
MaxCard-Fair, $\beta = 5$	+0		+0		+1		+1		+1		+2		+3	
MaxLife	+2	✓	+3	✓	+6	✓	+8	✓	+7	✓	+11	✓	+9	✓
Marginalized														
MaxCard	-2	✗	-2	✗	-3	✗	-4	✗	-6	✗	-7	✗	-9	✗
MaxCard-Fair, $\beta = 1$	-1	✗	-1	✗	-1	✗	-2	✗	-3	✗	-3	✗	-5	✗
MaxCard-Fair, $\beta = 2$	+0		+0		+1	✓	+1	✓	+2	✓	+1		+1	
MaxCard-Fair, $\beta = 3$	+1	✓	+1	✓	+3	✓	+3	✓	+3	✓	+5	✓	+4	✓
MaxCard-Fair, $\beta = 4$	+1	✓	+2	✓	+3	✓	+4	✓	+4	✓	+5	✓	+5	✓
MaxCard-Fair, $\beta = 5$	+1	✓	+2	✓	+3	✓	+4	✓	+5	✓	+7	✓	+5	✓
MaxLife	-1	✗	-3	✗	-3	✗	-5	✗	-6	✗	-6	✗	-9	✗

Table 3.2: Median gains in expected total number of transplants (top table) and total number of marginalized transplants (bottom table) under FutureMatch. A ✓ represents statistical significance (Wilcoxon signed-rank test, $p \ll 0.01$).

Our experiments support the following conclusions:

- FutureMatch under MaxCard and MaxCard-Fair with low $\beta = 1$ results in a significant increase in the *overall* number of transplants compared to myopic, at the cost of a smaller decrease in the number of *marginalized* transplants.
- FutureMatch under MaxCard-Fair with high β results in a significant increase in *marginalized* transplants, at *no cost* to the overall number of transplants under myopic matching.
- For a middle ground around $\beta = 2$, FutureMatch can result in both more overall expected transplants and more marginalized transplants.

3.1.2 Proposed research

The preliminary results presented as Table 3.2 above (§3.1.1) are promising; they support the notion that some additions to the standard model of (in this case) kidney exchange result in Pareto improvements to the exchange, while other additions can result in huge gains in one dimension at low or no cost to others. However, the experimental results suffer from deficiencies in each of the modules in the FutureMatch framework. For example, while the results in Table 3.2 used an improved learning method (compared to the initial method presented in Dickerson et al. [2012a]) for setting potentials, it still did not converge (even on small graphs and reasonably short horizons of 24 weeks). We plan to continue to increase the accuracy and breadth of similar experiments run using the advanced described in the proposed research in Sections 2.2, 2.3, and 2.4. This is "plug and play" proposed research; as we complete the proposed work in each of the individual dimensions of Chapter 2, we can—by design—easily drop in new optimization methods and more expressive models. We also plan to continue interacting with the UNOS exchange to get a better idea of what high-level objectives to use in FutureMatch; the hope here is to provide meaningful policy recommendations, as discussed in the next section.

3.2 Policy recommendations

Policy decisions in kidney exchange have been linked to economic and computational studies since before the first large-scale exchange was fielded in 2003–2004 [Roth et al., 2004, 2005a]. A feedback loop exists between the reality of fielded exchanges—now not only in the United States but internationally as well—and the theoretical and empirical models that inform their operation, such that the latter has grown substantially closer to accurately representing the former in recent years. That said, many gaps still exist between the mathematical models used in kidney exchange studies and the systems that actually provide matches on a day-to-day basis.

More accurate models are often not adopted quickly, if at all, by exchanges. One reason for this is complexity—and not in the computational sense. Humans—doctors, lawyers, and other policymakers who are not necessarily versed in optimization or theoretical economics and computer science—and the organizations they represent rightfully wish to understand the workings of an exchange's matching policy. We see many pieces of this thesis as easily understood, including the following.

- A commitee could use exploratory and sensitivity analysis, like we did in Chapter 2, to balance efficiency and fairness in kidney exchange under the weighted fairness rule (§2.3), which could inform prioritization schemes.
- Building (in spirit) on the work of Blum et al. [2015] and (in practice) on some of the more realistic edge testing policies we will develop in this thesis, testing some small number of promising potential matches for some subset of patient-donor pairs in a pool. As the experimental results from the Blum et al. [2015] paper show (Figure 2.3 in Section 2.2.2 of this proposal), even a *single* extra edge test per pair will produce substantially better results. Furthermore, these extra edge tests can be performed entirely in parallel if an exchange decides on a non-adaptive testing strategy (which we showed is quite effective), so the temporal cost to waiting time between a match run and transplant event would be minimal compared to the status quo.
- The FUTUREMATCH framework in general is easy to understand at a high level, and could be used to explore the effects of different high-level objectives on matching policy, where the underlying matching algorithms perform better (as supported by Sections 2.4 and 3.1.1) than the status quo algorithms.

Clearly, more extensive studies would need to be undertaken before policy recommendations could be made. These studies could take factors like the monetary cost of an extra crossmatch test or variability in testing prowess across different medical laboratories into account explicitly during the optimization process, as described in the proposed research of Section 2.2.2. We aim to submit a policy recommendation paper to a medical journal (e.g., *American Journal of Transplantation*) based on FUTUREMATCH experiments informed by input from members of the UNOS exchange and run under the improvements of Chapter 2.

Chapter 4

Other kidney exchange research & exploratory applications

In this chapter, we address work that will be included in this thesis that is specific to organ exchange, rather than to general barter exchange. We also discuss some new domains we are currently exploring that may or may not come to fruition before this thesis is completed.

4.1 Liver & multi-organ exchange

In Dickerson and Sandholm [2014], we explore the creation of living donor exchanges involving organs other than kidneys. We first explore large-scale *liver* exchange, which is similar to kidney exchange in some ways, but remains unexplored from a computational point of view.[1][2] The major difference between kidney and liver exchange rests in the increased risk to live donors, with very high rates of donor morbidity (24%), "near-miss" events in surgery (1.1%), and mortality (0.2%) compared to live donor kidney transplantation [Cheah et al., 2013]. Fielded kidney exchanges derive significant value from *altruistic* donors, who enter the exchange without a paired needy candidate and trigger long "chains" of donations within the pool. With such a high risk of complication from surgery in liver transplantation, we expect significantly fewer (or no, if deemed unethical by the medical community) altruistic donors in liver exchange.

With this in mind, we propose *multi-organ* exchange, where candidates in need of either kidneys or livers can swap donors in the same pool. We show theoretically that this combination provides linearly more transplants than running separate kidney and liver exchanges; this linear gain is a product of altruistic kidney donors creating chains that thread through the liver pool. (Figure 4.1 gives an example of a "threaded" chain.) We support this result experimentally on demographically accurate kidney, liver, and cross-organ exchanges; full experiments can be found in Dickerson and Sandholm [2014] and will be included in the completed thesis document.

[1]Recently, small-scale liver exchanges have been manually arranged by medical professionals. In Korea, 16 candidates swapped *by hand* willing living donors in a single hospital over the course of six years [Hwang et al., 2010]; similarly, in Hong Kong, 2 candidates hand-swapped willing donors [Chan et al., 2010]. This shows the feasibility of the idea at a small scale [Segev and Montgomery, 2010].

[2]Tayfun Sönmez and Utku Ünver are also looking at liver exchange; we will update this section in the thesis when their working paper is available.

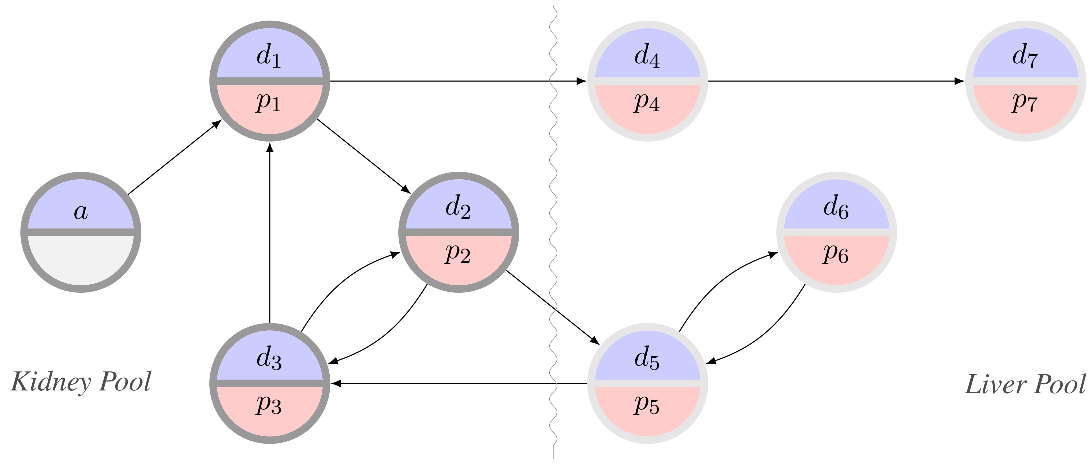


Figure 4.1: An example joint liver-kidney compatibility graph. The chain $\langle(a \to p_1), (d_1 \to p_4), (d_4 \to p_7), (d_7 \to \cdot)\rangle$ "threads" between the two pools; this type of structure drives the theoretical efficiency gains described in Section 4.1.1.

4.1.1 Main results

Theoretical results

We begin by working in an adapted version of the sparse model due to Ashlagi et al. [2012] (see Section 2.1). Recall that they build a random directed compatibility graph $D(n, \lambda, t(n), p_L, p_H)$ with n candidate-donor pairs, $t(n)$ altruistic donors, a fraction $\lambda < 1$ of the n candidate-donor pairs are lowly-sensitized with constant probability $p_L > 0$ of an incoming edge from each vertex in the pool, and a fraction $1 - \lambda > 0$ of the n candidate-donor pairs are highly-sensitized, and have probability p_H of an incoming edge from each vertex in the pool. We assume $p_H = \frac{c}{n}$ for some constant $c > 1$; thus, the graph induced by only those $1 - \lambda$ fraction of (sensitized) vertices with incoming edge probability p_H is sparse.

We assume, for kidney exchange compatibility graphs D_K, $t(n) > 0$; however, for liver exchange graphs D_L, $t(n) = 0$ (i.e., there are no altruistic liver donors). Finally, we define the graph join operator $D = join(D_K, D_L)$ between a kidney exchange graph D_K and liver exchange graph D_L as follows: add directed edges between candidate-donor pairs in both pools in accordance with each pair's associated probability (p_L or p_H); do not add edges from the $t(n)$ altruistic donors in D_K to vertices in D_L (since altruistic kidney donors are unwilling to donate a liver). For the sake of clarity, we assume that the p_L (resp. p_H) for D_K equals the p_L (resp. p_H) for D_L. This is without loss of generality; all that matters is that p_L be constant and $p_H = \frac{c}{n}$ for constant $c > 1$.

Proposition 6 assumes a linear (in the number of candidate-donor pairs) number of altruistic donors, while Proposition 7 works with just a constant number of altruistic donors.

Proposition 6. *Consider $\beta > 0$ and $\gamma > 0$, kidney compatibility graph D_K with n_K pairs and $t(n_K) = \beta n_K$ altruistic donors, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then any efficient matching on $D = join(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with probability approaching 1 as n_K approaches ∞).*

Proposition 7. *Consider $\gamma > 0$, kidney compatibility graph D_K with n_K pairs and constant $t > 0$ altruistic donors, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then there exists $\lambda' > 0$ such that for all $\lambda < \lambda'$, any efficient matching on $D = join(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of* *any such efficient matchings on D_K and D_L (with constant positive probability).*

We continue by considering a similarly-adapted version of the dense model of kidney exchange graphs (see Section 2.1) to multi-organ exchange; the full model is defined in the submitted journal version of Dickerson and Sandholm [2014] and will be included in the completed thesis. Then, in Proposition 8, we show that if there are enough altruistic kidney donors, the size of an efficient matching on D is larger by an additive linear fraction than the size of the aggregate of efficient matchings on D_L and D_K, the subgraphs induced by only the vertices consisting of pairs needing livers and kidneys, respectively. Formally, let D_K^X be the subgraph induced by only altruistic kidney donors of blood type $X \in \{\mathrm{O}, \mathrm{A}, \mathrm{B}, \mathrm{AB}\}$.

Proposition 8. *Consider $\beta^A = \mu_A\mu_{AB}$, $\beta^B = \mu_B\mu_{AB}$, and $\gamma > 0$, kidney compatibility graph D_K with n_K pairs, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. If at least one of $|D_K^A| > \beta^A n_K$ or $|D_K^B| > \beta^B n_K$, then any efficient matching on $D = join(D_K, D_L)$ matches $\Omega(n_k)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with probability approaching 1 as $n_K \to \infty$).*

Algorithmic and empirical results

We support the theoretical results defined above with liver and multi-organ exchange simulation results. One such result follows.

We simulate a demographically accurate bi-organ exchange featuring candidates in need of either a kidney or a liver who can swap donors in a *combined* candidate-donor pool. Approximately 85% of the candidates in the simulated pool need kidneys, while the other 15% need livers, as determined by OPTN waitlist data. We mimic the experiments in the previous section, with a starting pool size of $|V| = 800$ candidates who are highly sensitized and are assumed to have built up in the pool over time; we also include 100 altruistic kidney donors who enter the combined pool at an expected constant rate. We use the same exogenous transplant incompatibility parameter ($f = 0.7$) as in the previous section, and simulate candidate-donor pairs entering and exiting the pool in a similar fashion. To generate the candidates, we draw from the two different US distributions based on whether the candidate needs a kidney or a liver. Naturally, donors are drawn from the same US distribution in the two cases. We test over 24 months.

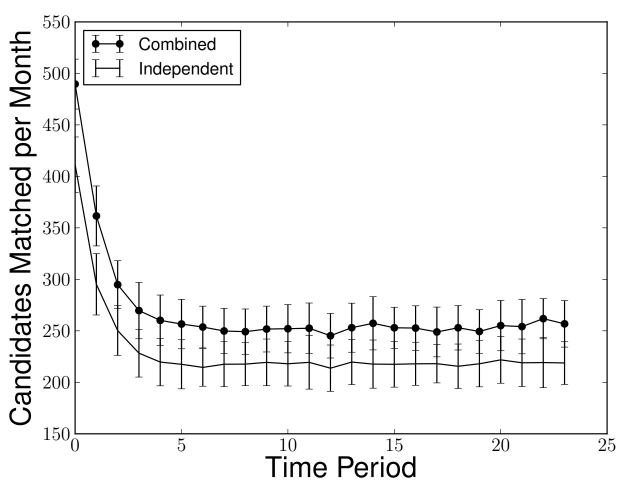


Figure 4.2: Number of matches in independent liver and kidney exchanges and a combined multi-organ exchange, per time period, in a dynamic setting over $T = 24$ months.

When we compare the total number of matches made over the entire period simulated above, the difference in lives saved between two independent pools and the combined bi-organ pool is stark. In our experiments, the combined bi-organ pool produced 16.8% more matches than the sum of the two independent organ pools. An independent samples t-test revealed that the difference between the aggregate number of lives saved using independent, simultaneous liver and kidney exchanges and using a combined multi-organ exchange was significant, $t(46) = 31.37$, $p \ll 0.0001$.

4.1.2 Proposed research

The multi-organ results of Section 4.1.1 are promising in theory and in our preliminary experimental results. Algorithmically, any advances to the optimization algorithms made in Chapter 2 will be immediately applicable to the multi-organ exchange problem as well. However, the question of equity is in many ways *far* more important in the multi-organ case. As discussed earlier in this section, the negative side effects of donation are *vastly* different for each organ. In practice, this would have to be balanced in some way; in the event that large-scale exchanges in organs beyond livers are fielded [Segev and Montgomery, 2010, Dickerson and Sandholm, 2014, Ergin et al., 2014], it is worth exploring the idea of combining pools to increase market thickness. We propose to explore this idea more extensively than was done by Dickerson and Sandholm [2014], both theoretically and through more realistic experiments. We also propose to include some of the fairness considerations of Section 2.3 in the multi-organ model, to see how quickly the gains seen from combining the two pools decrease to negligible as equity considerations are taken into account.

4.2 Analysis of kidney chains in a dense model

Earlier theoretical kidney exchange research focused on dense models of compatibility graphs, which have since been shown to not mimic the behavior of the current pools seen in fielded kidney exchanges (see, e.g., the discussions by Ashlagi et al. [2012] and Dickerson et al. [2012b]). This may change as kidney exchange matures to a steady state. We performed an analysis of chain length in a dense model of kidney exchange [Dickerson et al., 2012b].

We work in an adapted version of the dense model of kidney exchange due to Ashlagi and Roth [2011, 2014] (see Section 2.1). Here, $G(n)$ represents an n-patient kidney exchange drawn in accordance with that model, a donor and a patient who are blood type compatible are tissue type incompatible with constant probability $\bar{\gamma}$, and the frequency of each blood type X is denoted by μ_X.

Theorem 5. *Assume that $\bar{\gamma} < 2/5$, $\mu_O < 3\mu_A/2$, and $\mu_O > \mu_A > \mu_B > \mu_{AB}$. Then with high probability $G(n)$ has an efficient allocation (i.e., one that saves as many patients as possible) that uses only cycles of length at most 3 and chains of length at most 3.*

Theorem 5 states that short chains and short cycles suffice to achieve an efficient matching in a dense model of kidney exchange. The allocation builds on the allocation presented by Ashlagi and Roth [2011, 2014] (which shows a qualitatively similar result in dense kidney exchange graphs without altruistic donors), and is depicted graphically in Figure 4.3.

In Dickerson et al. [2012b], we show that the theoretical results align with simulations done in the standard generative *dense* model of kidney exchange due to Saidman et al. [2006], but that the theoretical results diverge severely from the behavior of the UNOS pool, where chains of very long length help (up to length 13 at the time of that paper's publication, while chains of length greater than 20 have been seen to help since that paper's publication).

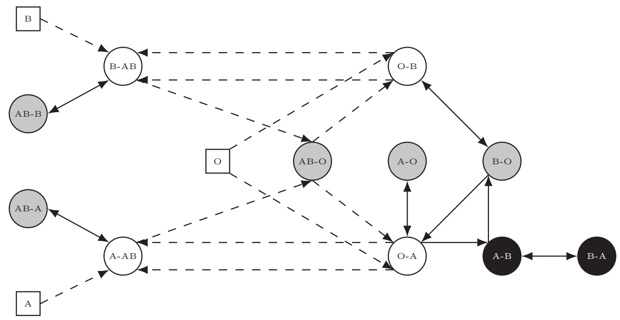


Figure 4.3: Altruists are shown as rectangles; candidate-donor pairs as ovals. Over-demanded pairs are gray, under-demanded are white, and reciprocal pairs are black.

4.3 Exploratory application areas

Ideas in this thesis are often motivated by and presented in the framework of kidney exchange; yet, the techniques are applicable to a wide variety of markets without a medium of exchange. This section describes three application areas we are currently exploring. While we make no promises about their inclusion in the final thesis,

4.3.1 ZooSwap: A centralized mechanism for zoos to exchange animals

As a byproduct of the Endangered Species Act of 1973, the acquisition of animals—especially those that are endangered, vulnerable, or threatened—by zoos and aquariums has been restricted. In response to import and export restrictions imposed by that legislation, zoos have expanded their internal breeding programs, which create animals that can then be traded with other zoos. We are exploring the idea of a centralized market that would enable zoos to trade animals.

The Association of Zoos and Aquariums keeps an online database, the Animal Exchange Database, that is accessible by zoos; here, zoos can list their inventory of animals, as well as the associated "studbooks"—or genetic histories—for those animals and other pertinent information. Zoos can then request listed animals that align with their breeding programs or inventory management plans that are also available for trade. As far as we currently know, this is done entirely by hand and only makes use of either one-for-one swaps (i.e., traditional matching) or one-for-nothing swaps, where the latter type of swap allocates some off-the-books "karma" to the zoo who donated the animal (as in a gift economy). There also exists a for-profit corporation, International Animal Exchange[3] (IAE), that specializes in the *transportation* of animals from one zoo to another, but also claims to both source and place surplus animals. This is likely done using similar ad hoc techniques.

We would like to introduce the formal ideas of barter cycles and barter chains to this market. We believe barter chains, especially, could have great impact. Zoos' breeding programs often create surplus animals that do not align with their population management programs for specific species; these surplus animals could trigger long chains of animal donations. Donor reneging, as in kidney exchange, would hopefully be low (albeit for different reasons), due to the closed nature of the zoo and aquarium communities.

[3] http://internationalanimalexchange.com/

4.3.2 General organ exchange

There has been recent work in fielding exchanges in organs other than kidneys. For example, in Korea, 16 candidates swapped willing donors in a single hospital over the course of six years [Hwang et al., 2010]; similarly, in Hong Kong, two candidates hand-swapped willing donors [Chan et al., 2010]. Similarly, the idea of lung exchange [Ergin et al., 2014]—which is different than kidney or liver exchange in that two donors are required for every one patient—was recently fielded at a small scale in Japan. As discussed in Section 4.1, combining independent exchanges that deal in different organs is feasible in theory, although significant further investigation into the legal, ethical, and logistic constraints and implications of a cross-organ swap would need to be investigated before fielding such an exchange. We plan to discuss the feasibility of fielding a single-center liver exchange (starting with local discussions at the University of Pittsburgh Medical Center), then explore future multi-center and/or multi-organ exchange directions from there.

4.3.3 Italian labor markets

In Italy, workers are assigned jobs based on a variety of constraints like location by the government. Workers then often swap their assigned jobs with other workers in an ad hoc secondary market. We are interested in designing either a centralized secondary market or a better (e.g., more expressive on the workers' side) primary market for job swapping or assignment, respectively. This will be joint work with Nicola Gatti at Politecnico di Milano.

Chapter 5

Timeline

My proposed research will roughly be completed as follows:

August–September 2015: Proposal

August–September 2015: Decomposable dynamic matching (§2.4), submit to AAAI

August–September 2015: Submit invited paper on the price of fairness to *Artificial Intelligence*

September–November 2015: Generalized equity problem (§2.3), submit to AAMAS

September 2015–February 2016: Budgeted edge testing problem (§2.2), submit to EC

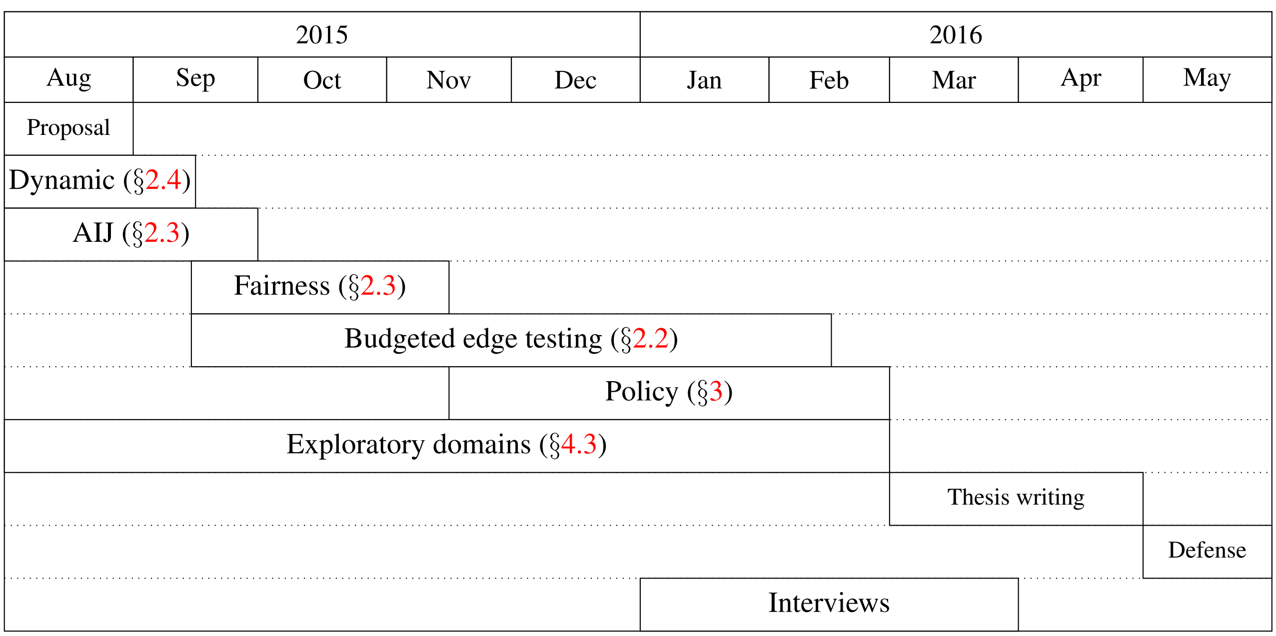
November 2015–February 2016: Policy recommendations (§3), submit to *Am J Transplant*

Ongoing: Searching for new complex matching domains (§4.3)

January–March 2016: Academic job interviews

March–April 2016: Writing my thesis document

May 2016: Thesis defense



Chapter 6

My other research (not a part of this thesis)

This chapter briefly overviews some of the research directions I have followed that are not included in my proposed thesis. A common thread in the following research projects is the design and use of new optimization methods and models to inform market design decisions or make policy recommendations. This is the general type of research I plan to continue performing as a professor.

- **Kidney exchange.** My work with Blum et al. [2015] works in models of matching and kidney exchange with post-algorithmic match edge failure (like in Dickerson et al. [2013]), but instead asks how to efficiently select a subset of edges to pre-test (possibly over multiple rounds) before finalizing the recommended match. Those algorithms result in large gains in expected number of successful transplants on generated and real data from the UNOS exchange. In ongoing work with Das et al. [2015], I am exploring the effect of competition between exchanges—specifically, exchanges with different matching policies parameterized by the cadence of match runs—on global social welfare and on the social welfare of marginalized patients. We are extending a recent model of dynamic matching within a single matching mechanism developed by Akbarpour et al. [2014].
- **Computational advertising.** With the increasing popularity of non-linear TV and the introduction of new content delivery mechanisms, monetization in the TV space is becoming more complex. With Optimized Markets, Inc.—a CMU spin-off company—I am designing a new marketplace for advertising sales that uses optimization to tame this complexity from the human's perspective—while increasing overall market efficiency from a computational economics point of view. In this context, I am interested in questions like the following. Can we use automated abstraction to deal with channel explosion in online and television advertising markets? In the case where bidders bid asynchronously and bids (possibly in the form of complex campaigns) can be accepted or rejected in an online fashion, can we tractably look to the future to yield better expected revenue for the seller or return on investment for the bidder(s)? Can we characterize the tradeoff between optimality and tractability in the context of dynamic market clearing of campaigns and abstraction of concrete channels? In January–June 2014, I served as Principal Investigator at Optimized Markets for a grant to study a related optimization problem in the context of television advertising [Dickerson, 2014].
- **Counter-terrorism & security games.** I have been actively involved in defense- and counter-terror-related research for nearly a decade, with a focus on applying game theory to real-world situations. For example, in Dickerson et al. [2010], we look at using limited resources to protect either static or dynamic assets from adversaries on a graph. For example, a central coordinator may wish to deploy a constant number of police checkpoints on roads leading to a political building (a static asset); similarly, an event planner may wish to choose a parade route for a dynamic asset such that the route

is most easily defended by deployed police forces. Deployments can be probabilistic, the strategy of the central coordinator is observed by the adversary, and the adversary is assumed to be completely rational with unbounded computational power. We give an optimal PTIME solution to the static asset protection problem and prove NP-hardness for the dynamic asset protection problem, then present a heuristic method for solving the latter. We support both methods with experiments on a 175K node representation of the US road network.

In Shakarian et al. [2012], we look at geospatial abduction, where the problem is to infer a set of hidden locations that "best explain" a given set of locations of observations. For example, given a set of observed improvised explosive device (IED) attacks in a combat zone, what set of locations for supporting weapons and materials best explain these attacks, given constraints on insurgents' movements, demographic information about the surrounding area, and other constraints? We look at this problem from a game-theoretic point of view between an agent (who searches for the explanatory locations) and an adversary (who chooses the locations of the observations and explanations), and characterize optimal strategies for both (although we also prove NP-hardness for computing either side's strategy, show that the general adversary's problem has no FPTAS unless P=NP, and show that this negative result can be removed by applying light restrictions to the reward functions). We also experimentally validate our results on real data from Baghdad, Iraq, during 21 months of the Iraq War. This research forms the basis for the Socio-Cultural Adversarial Reasoning Engine (SCARE) that predicted locations of IED weapons caches in Baghdad, Iraq, to within 700 meters; the system was delivered to a number of Department of Defense groups.

In Sawant et al. [2015], we look at a combining vector equilibria [Shapley and Rigby, 1959] and well-supported approximate equilibria [Daskalakis et al., 2006]. We give bounds on the computation of these equilibria for special cases (e.g., low rank games) and give a QPTAS for the general case when the number of players is small. This QPTAS leads to efficient algorithms that we support empirically on a five-player game between Lashkar-e-Taiba (a jihadist group), Pakistan's government, Pakistan's military, India, and the US. In Subrahmanian et al. [2012], we also focus on Lashkar-e-Taiba and build a computational model of the group using 770 variables tracked monthly since the group's inception. Subrahmanian and Dickerson [2009] also looks at computational modeling of terror groups.

- **Fair allocation of indivisible goods.** This thesis focuses on a specific case of the allocation of indivisible goods to agents who value those goods. Under a different model, Dickerson et al. [2014a] investigates the existence of envy-free allocations of indivisible goods, that is, allocations of the goods such that no player values those goods allocated to a different player more than she values her own allocation. With n agents, m goods, and additive valuations, we show that if $n = O(m/\ln m)$ then an envy-free allocation almost certainly exists, while if m is only larger than n by a linear fraction, then envy-free allocations are unlikely to exist. We also demonstrate a sharp phase transition between existence and non-existence of such allocations, and show that empirically the problem is hardest during this transition.

- **Selective sustained attention (SSA).** For the past four years, in collaboration with a group of psychology researchers and The Children's School at CMU, I have used eye tracking data to study two systems of attentional control and their relation to learning outcomes in preschoolers. Selective (e.g., paying attention to a teacher instead of a distracting poster on the wall) sustained attention (SSA) is critical for higher order cognition. I have been actively involved in the design and development of Track-It, a modular, open source software suite that measures SSA in children using eye tracking data; specifically, the suite aims to assess the contribution of exogenous and endogenous factors to SSA. This test suite is available for free at `http://www.psy.cmu.edu/~trackit/`. Our work: [Thiessen et al., 2012, Fisher et al., 2013b,a, Erickson et al., 2014, 2015a,b].

Bibliography

David Abraham, Avrim Blum, and Tuomas Sandholm. Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 295–304, 2007.

Marek Adamczyk. Improved analysis of the greedy algorithm for stochastic matching. *Information Processing Letters*, 111(15):731–737, 2011.

Mustafa Akan, Oguzhan Alagoz, Baris Ata, Fatih Safa Erenay, and Adnan Said. A broader view of designing the liver allocation system. *Operations Research*, 60(4):757–770, 2012.

Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. Dynamic matching market design. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, page 355, 2014.

Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts. Determining the acceptance of cadaveric livers using an implicit model of the waiting list. *Operations Research*, 55(1):24–36, 2007.

Ross Anderson. *Stochastic models and data driven simulations for healthcare operations*. PhD thesis, Massachusetts Institute of Technology, 2014.

Ross Anderson, Itai Ashlagi, David Gamarnik, and Yash Kanoria. A dynamic model of barter exchange. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2015a.

Ross Anderson, Itai Ashlagi, David Gamarnik, and Alvin E Roth. Finding long chains in kidney exchange using the traveling salesman problem. *Proceedings of the National Academy of Sciences*, 112(3):663–668, 2015b.

Elliot Anshelevich, Meenal Chhabra, Sanmay Das, and Matthew Gerrior. On the social welfare of mechanisms for repeated batch matching. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 60–66, 2013.

Itai Ashlagi and Alvin E. Roth. Individual rationality and participation in large scale, multi-hospital kidney exchange. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 321–322, 2011.

Itai Ashlagi and Alvin E Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9(3):817–863, 2014.

Itai Ashlagi, Duncan S. Gilchrist, Alvin E. Roth, and Michael Rees. Nonsimultaneous chains and dominos in kidney-paired donation—revisited. *American Journal of Transplantation*, 11(5):984–994, 2011.

Itai Ashlagi, David Gamarnik, Michael Rees, and Alvin E. Roth. The need for (long) chains in kidney exchange. NBER Working Paper No. 18202, July 2012.

Itai Ashlagi, Patrick Jaillet, and Vahideh H. Manshadi. Kidney exchange in dynamic sparse heterogenous pools. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 25–26, 2013.

Pranjal Awasthi and Tuomas Sandholm. Online stochastic optimization in the large: Application to kidney exchange. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 405–411, 2009.

Nikhil Bansal, Anupam Gupta, Jian Li, Julian Mestre, Viswanath Nagarajan, and Atri Rudra. When LP is the cure for your matching woes: Improved bounds for stochastic matchings. *Algorithmica*, 63(4): 733–762, 2012.

Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W. P. Savelsbergh, and Pamela H. Vance. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 46(3): 316–329, 1998.

Dimitris Bertsimas, Vivek F Farias, and Nikolaos Trichakis. The price of fairness. *Operations Research*, 59 (1):17–31, 2011.

Dimitris Bertsimas, Vivek F Farias, and Nikolaos Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, 61(1):73–87, 2013.

Avrim Blum, Anupam Gupta, Ariel D. Procaccia, and Ankit Sharma. Harnessing the power of two crossmatches. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 123–140, 2013.

Avrim Blum, John P. Dickerson, Nika Haghtalab, Ariel D. Procaccia, Tuomas Sandholm, and Ankit Sharma. Ignorance is almost bliss: Near-optimal stochastic matching with few queries. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2015.

Steven J Bradtke and Andrew G Barto. Linear least-squares algorithms for temporal difference learning. *Machine Learning*, 22(1-3):33–57, 1996.

David B Brown, James E Smith, and Peng Sun. Information relaxations and duality in stochastic dynamic programs. *Operations Research*, 58(4):785–801, 2010.

Ioannis Caragiannis, Christos Kaklamanis, Panagiotis Kanellopoulos, and Maria Kyropoulou. The efficiency of fair division. International Workshop on Internet and Network Economics (WINE), 2009.

Ioannis Caragiannis, Aris Filos-Ratsikas, and Ariel D. Procaccia. An improved 2-agent kidney exchange mechanism. International Workshop on Internet and Network Economics (WINE), 2011.

See Ching Chan, Chung Mau Lo, Boon Hun Yong, Wilson JC Tsui, Kelvin KC Ng, and Sheung Tat Fan. Paired donor interchange to avoid ABO-incompatible living donor liver transplantation. *Liver Transplantation*, 16(4):478–481, 2010.

Yee Lee Cheah, Mary Ann Simpson, James J Pomposelli, and Elizabeth A Pomfret. Incidence of death and potentially life-threatening near-miss events in living donor hepatic lobectomy: A world-wide survey. *Liver Transplantation*, 19(5):499–506, 2013.

Ning Chen, Nicole Immorlica, Anna R. Karlin, Mohammad Mahdian, and Atri Rudra. Approximating matches made in heaven. In *Proceedings of the International Conference on Automata, Languages, and Programming (ICALP)*, pages 266–278, 2009.

Yanhua Chen, Yijiang Li, John D. Kalbfleisch, Yan Zhou, Alan Leichtman, and Peter X.-K. Song. Graph-based optimization algorithm and software on kidney exchanges. *IEEE Transactions on Biomedical Engineering*, 59:1985–1991, 2012.

Miguel Constantino, Xenia Klimentova, Ana Viana, and Abdur Rais. New insights on integer-programming models for the kidney exchange problem. *European Journal of Operational Research*, 231(1):57–68, 2013.

Kevin P. Costello, Prasad Tetali, and Pushkar Tripathi. Matching with commitment. In *Proceedings of the*

International Conference on Automata, Languages, and Programming (ICALP), pages 822–833, 2012.

Sanmay Das, John P. Dickerson, Zhuoshu Li, and Tuomas Sandholm. Competing dynamic matching markets, 2015. Working paper.

Constantinos Daskalakis, Aranyak Mehta, and Christos Papadimitriou. A note on approximate nash equilibria. In *International Workshop On Internet And Network Economics (WINE)*, pages 297–306. Springer, 2006.

M De Klerk, JA Kal-van Gestel, BJ Haase-Kromwijk, FH Claas, and W Weimar. Eight years of outcomes of the Dutch living donor kidney exchange program. *Clinical Transplants*, pages 287–290, 2010a.

M De Klerk, WM Van Der Deijl, MD Witvliet, BJ Haase-Kromwijk, FH Claas, and W. Weimar. The optimal chain length for kidney paired exchanges: an analysis of the Dutch program. *Transplant International*, 23(11):1120–1125, 2010b.

Dursun Delen, Glenn Walker, and Amit Kadam. Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine*, 34(2):113–127, June 2005.

John P. Dickerson. Advertising sales and traffic optimization: Difficult customer-requested optimization constraints and scalability on real data, 2014. NSF SBIR Phase I Award #1345567.

John P. Dickerson and Tuomas Sandholm. Multi-organ exchange: The whole is greater than the sum of its parts. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 1412–1418, 2014.

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

John P. Dickerson, Gerardo I. Simari, V.S. Subrahmanian, and Sarit Kraus. A graph-theoretic approach to protect static and moving targets from adversaries. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 299–306, 2010.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Dynamic matching via weighted myopia with application to kidney exchange. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 1340–1346, 2012a.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Optimizing kidney exchange with transplant chains: Theory and reality. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 711–718, 2012b.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Failure-aware kidney exchange. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 323–340, 2013.

John P. Dickerson, Jonathan Goldman, Jeremy Karp, Ariel D. Procaccia, and Tuomas Sandholm. The computational rise and fall of fairness. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 1405–1411, 2014a.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Empirical price of fairness in failure-aware kidney exchange. In *Towards Better and more Affordable Healthcare: Incentives, Game Theory, and Artificial Intelligence (HCAGT) workshop at AAMAS-2014*, 2014b.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Price of fairness in kidney exchange. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1013–1020, 2014c.

Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965.

Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

Haluk Ergin, Tayfun Sönmez, and M Utku Ünver. Lung exchange, 2014. Working paper.

Lucy C. Erickson, Erik D. Thiessen, Karrie E. Godwin, John P. Dickerson, and Anna V. Fisher. Endogenously- but not exogenously-driven selective sustained attention is related to learning in a classroom-like setting in kindergarten children. In *Conference of the Cognitive Science Society (CogSci)*, 2014.

Lucy C. Erickson, Karrie Godwin, John P. Dickerson, Erik D. Thiessen, and Anna V. Fisher. Different mechanisms for regulating sustained attention and learning in children. In *Biennial Meeting of the Society for Research in Child Development (SRCD)*, 2015a.

Lucy C. Erickson, Karrie Godwin, John P. Dickerson, Erik D. Thiessen, and Anna V. Fisher. Endogenously- and exogenously-driven selective sustained attention: Contributions to learning in kindergarten children. Revision submitted to Journal of Experimental Child Psychology, 2015b.

Anna V. Fisher, Erik D. Thiessen, John P. Dickerson, and Lucy C. Erickson. Development of selective sustained attention: Conceptual and measurement issues. In *Biennial Meeting of the Cognitive Development Society (CDS)*, 2013a.

Anna V. Fisher, Erik D. Thiessen, Karrie Godwin, Heidi Kloos, and John P. Dickerson. Assessing selective sustained attention in 3- to 5-year-old children: Evidence from a new paradigm. *Journal of Experimental Child Psychology*, 113, 2013b.

Sommer E. Gentry and Dorry L. Segev. The honeymoon phase and studies of nonsimultaneous chains in kidney-paired donation. *American Journal of Transplantation*, 11(12):2778–2779, 2011.

Sommer E. Gentry, Robert A. Montgomery, Bruce J. Swihart, and Dorry L. Segev. The roles of dominos and nonsimultaneous chains in kidney paired donation. *American Journal of Transplantation*, 9(6):1330–1336, 2009.

Kristiaan M. Glorie. Estimating the probability of positive crossmatch after negative virtual crossmatch. Technical report, Erasmus School of Economics, 2012.

Kristiaan M. Glorie, J. Joris van de Klundert, and Albert P. M. Wagelmans. Kidney exchange with long chains: An efficient pricing algorithm for clearing barter exchanges with branch-and-price. *Manufacturing & Service Operations Management (MSOM)*, 16(4):498–512, 2014.

Gagan Goel and Pushkar Tripathi. Matching with our eyes closed. In *Symposium on the Foundations of Computer Science (FOCS)*, pages 718–727. IEEE, 2012.

HHS/HRSA/HSB/DOT. OPTN/SRTR annual data report, 2011.

Frank Hutter, Holger Hoos, Kevin Leyton-Brown, and Thomas Stützle. ParamILS: An automatic algorithm configuration framework. *Journal of Artificial Intelligence Research*, 36(1):267–306, 2009. ISSN 1076-9757.

Frank Hutter, Holger Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *Proc. of LION-5*, pages 507–523, 2011.

Shin Hwang, Sung-Gyu Lee, Deok-Bog Moon, Gi-Won Song, Chul-Soo Ahn, Ki-Hun Kim, Tae-Yong Ha, Dong-Hwan Jung, Kwan-Woo Kim, Nam-Kyu Choi, Gil-Chun Park, Young-Dong Yu, Young-Il Choi, Pyoung-Jae Park, and Hea-Seon Ha. Exchange living donor liver transplantation to overcome ABO incompatibility in adult patients. *Liver Transplantation*, 16(4):482–490, 2010. ISSN 1527-6473. doi:

10.1002/lt.22017. URL http://dx.doi.org/10.1002/lt.22017.

IBM ILOG Inc. CPLEX 12.2 User's Manual, 2010.

William Stanley Jevons. *Money and the Mechanism of Exchange*, volume 17. Appleton, London, 1885.

KM Keizer, M De Klerk, BJJM Haase-Kromwijk, and W Weimar. The Dutch algorithm for allocation in living donor kidney exchange. In *Transplantation Proceedings*, volume 37, pages 589–591. Elsevier, 2005.

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

Kidney Transplantation Committee. OPTN concepts for kidney allocation, 2011.

Beom Seok Kim, Yu Seun Kim, Soon Il Kim, Myoung Soo Kim, Ho Yung Lee, Yong-Lim Kim, Chan Duck Kim, Chul Woo Yang, Bum Soon Choi, Duck Jong Han, et al. Outcome of multipair donor kidney exchange by a web-based algorithm. *Journal of the American Society of Nephrology*, 18(3):1000–1006, 2007.

Xenia Klimentova, Filipe Alvelos, and Ana Viana. A new branch-and-price approach for the kidney exchange problem. In *Computational Science and Its Applications (ICCSA-2014)*, pages 237–252. Springer, 2014.

Michail G Lagoudakis and Ronald Parr. Least-squares policy iteration. *Journal of Machine Learning Research*, 4:1107–1149, 2003.

Yijiang Li, Jack Kalbfleisch, Peter Xuekun Song, Yan Zhou, Alan Leichtman, and Michael Rees. Optimization and simulation of an evolving kidney paired donation (KPD) program. Department of biostatistics working paper 90, University of Michigan, May 2011.

David Manlove and Gregg O'Malley. Paired and altruistic kidney donation in the UK: Algorithms and experimentation. *ACM Journal of Experimental Algorithmics*, 19(1), 2014.

Naren Meadem, Nele Verbiest, Kiyana Zolfaghar, Jayshree Agarwal, Si-Chi Chin, and Senjuti Basu Roy. Exploring preprocessing techniques for prediction of risk of readmission for congestive heart failure patients. In *Data Mining and Healthcare (DMH), at International Conference on Knowledge Discovery and Data Mining (KDD)*, 2013.

Robert Montgomery, Sommer Gentry, William H Marks, Daniel S Warren, Janet Hiller, Julie Houp, Andrea A Zachary, J Keith Melancon, Warren R Maley, Hamid Rabb, Christopher Simpkins, and Dorry L Segev. Domino paired kidney donation: a strategy to make best use of live non-directed donation. *The Lancet*, 368(9533):419–421, 2006.

Gerhard Opelz. Correlation of HLA matching with kidney graft survival in patients with or without cyclosporine treatment: for the collaborative transplant study. *Transplantation*, 40(3):240–242, 1985.

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

Michael Rees, Jonathan Kopke, Ronald Pelletier, Dorry Segev, Matthew Rutter, Alfredo Fabrega, Jeffrey Rogers, Oleh Pankewycz, Janet Hiller, Alvin Roth, Tuomas Sandholm, Utku Ünver, and Robert Montgomery. A nonsimultaneous, extended, altruistic-donor chain. *New England Journal of Medicine*, 360 (11):1096–1101, 2009.

Alvin Roth. Repugnance as a constraint on markets. *Journal of Economic Perspectives*, 21(3):37–58, 2007.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119(2): 457–488, 2004.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. A kidney exchange clearinghouse in New England. *American Economic Review*, 95(2):376–380, 2005a.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125 (2):151–188, 2005b.

Alvin Roth, Tayfun Sönmez, Utku Ünver, Frank Delmonico, and Susan L. Saidman. Utilizing list exchange and nondirected donation through 'chain' paired kidney donations. *American Journal of Transplantation*, 6:2694–2705, 2006.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. Efficient kidney exchange: Coincidence of wants in a market with compatibility-based preferences. *American Economic Review*, 97:828–851, 2007.

Susan L. Saidman, Alvin Roth, Tayfun Sönmez, Utku Ünver, and Frank Delmonico. Increasing the opportunity of live kidney donation by matching for two and three way exchanges. *Transplantation*, 81(5): 773–782, 2006.

Anshul Sawant, John P. Dickerson, Mohammad T. Hajiaghayi, and V.S. Subrahmanian. Automated generation of counter-terrorism policies using multi-expert input. *ACM Transactions on Intelligent Systems and Technology*, 2015.

Dorry Segev, Sommer Gentry, D. S. Warren, B. Reeb, and R. A. Montgomery. Kidney paired donation and optimizing the use of live donor organs. *Journal of the American Medical Association*, 293(15): 1883–1890, 2005.

Dorry L Segev and Robert A Montgomery. The application of paired donation to live donor liver transplantation. *Liver Transplantation*, 16(4):423–425, 2010.

Paulo Shakarian, John P. Dickerson, and V.S. Subrahmanian. Adversarial geospatial abduction problems. *ACM Transactions on Intelligent Systems and Technology*, 3(2):34:1–34:35, 2012.

Lloyd S Shapley and Fred D Rigby. Equilibrium points in games with vector payoffs. *Naval Research Logistics Quarterly*, 6(1):57–61, 1959.

Xuanming Su and Stefanos A. Zenios. Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations Research*, 53:443–455, 2005.

V.S. Subrahmanian and John P. Dickerson. What can virtual worlds and games do for national security? *Science*, 326(5957):1201–1202, 2009.

V.S. Subrahmanian, Aaron Mannes, Amy Sliva, Jana Shakarian, and John P. Dickerson. *Computational Analysis of Terrorist Groups: Lashkar-e-Taiba*. Springer, 2012. ISBN 978-1-4614-4768-9.

Erik D. Thiessen, John P. Dickerson, Lucy C. Erickson, and Anna V. Fisher. Eyes as the windows of cognition: The Track-It paradigm and selective attention. In *SRCD Themed Meeting on Developmental Methodology*, 2012.

Panos Toulis and David C. Parkes. A random graph model of kidney exchanges: efficiency, individual-rationality and incentives. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 323–332. ACM, 2011.

UNOS. United Network for Organ Sharing (UNOS). http://www.unos.org/.

Utku Ünver. Dynamic kidney exchange. *Review of Economic Studies*, 77(1):372–414, 2010.

Steve Woodle, John Daller, Mark Aeder, Ron Shapiro, Tuomas Sandholm, Vincent Casingal, David Goldfarb, Richard Lewis, Jens Goebel, and M Siegler. Ethical considerations for participation of nondirected living donors in kidney exchange programs. *American Journal of Transplantation*, 10:1460–1467, 2010.