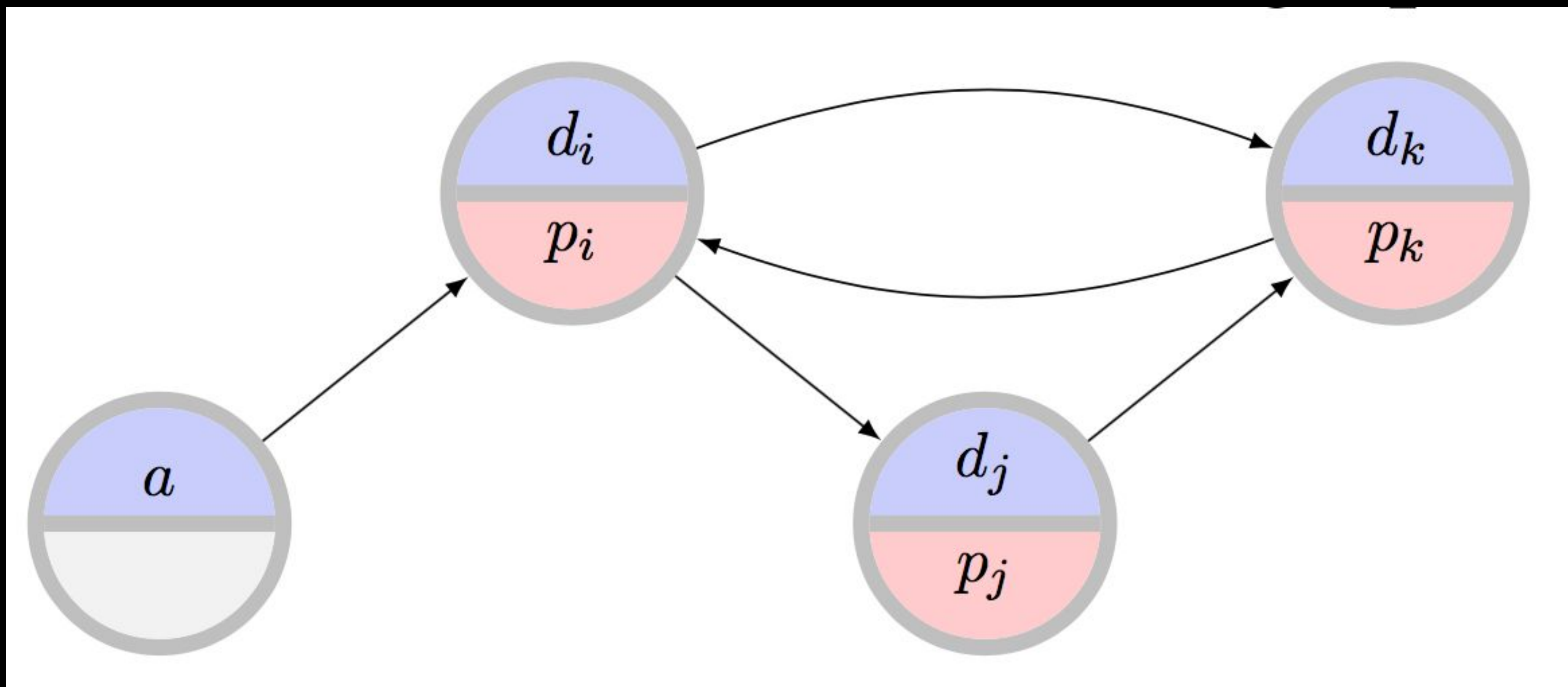
FutureMatch: Combining Human Value Judgments and Machine Learning to Match in Dynamic Environments

By John Dickerson and Tuomas Sandholm

Presentation by Ashton Webster

Background: Kidney Exchanges and Our Favorite Diagram



- Patient-Donor Pairs enter exchange
- Edges represent compatibility
- Only going to consider live transplants (not deceased) in this talk

Background: State of the Art In Practice*

- Myopic matching
  - Current UNOS matching attempts to maximize matches weighted by priority points
- Sensitization Priority
  - Current UNOS matching prioritizes matching for highly sensitized patients and patients under 18

* as of 2014

Background: State of the Art In Theory*

- Potentials:
  - Dickerson et al. [1]: Considering future value of edges in kidney exchanges can provide many benefits
- Edge Failure and Failure Aware Models
  - Alshagi et al. [2]: Use reported failure rate for ADP data
  - Dickerson et al. [3]: Show that failure aware models outperform others

* as of 2014

Methodology overview

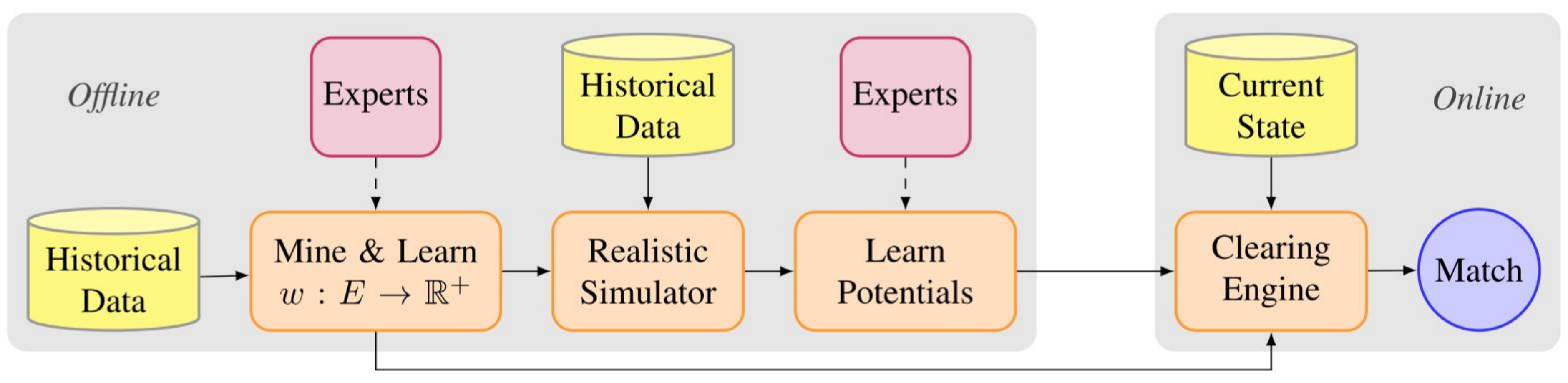


Figure 2: The FUTUREMATCH framework.

- Combine 3 main factors:
  - Edge weights - representing objective function values
  - Potentials - representing future value of pair
  - Edge Failure - Utility discounting based on probability of edge exiting the pool

Separating the Ends and the Means

- Current discussion by experts often confounds the desired objectives (ends) and (means)
  - Example: “We should seek to increase total matches by preferring matching highly sensitized matches, because highly sensitized patients will be difficult to match in the future”
- Experts should only be discussing the Ends, model should handle the Means
  - Ends roughly correspond to edge weights
  - Means roughly correspond to potentials

Objective Functions

	Deterministic	Failure Aware
MaxCard	Maximize the total number of patients who are algorithmically matched	Maximize the total number of patients who receive transplants in expectation
MaxCard-Fair	Maximize the total number of patients who are algorithmically matched, where “marginalized” patients are weighted in the objective by some constant factor β	Maximize the total number of patients who receive transplants in expectation, where “marginalized” patients are weighted in the objective by some constant factor β
MaxLife	Maximize the total time algorithmically matched donor organs last in patients	Maximize the total time transplanted donor organs will last in patients in expectation

Objective Functions: MaxCard and MaxCard-Fair

- MaxCard-Fair is a generalization of MaxCard
  - Think of MaxCard as MaxCard-Fair with no vertices receiving the β increase

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

- $V_p \subseteq V$ represents preferred edges
- β is positive (in this case in the set {1,2,3,4,5})
- w_e is original edge weight

MaxCard-Fair: Preferred (Marginalized) Vertices

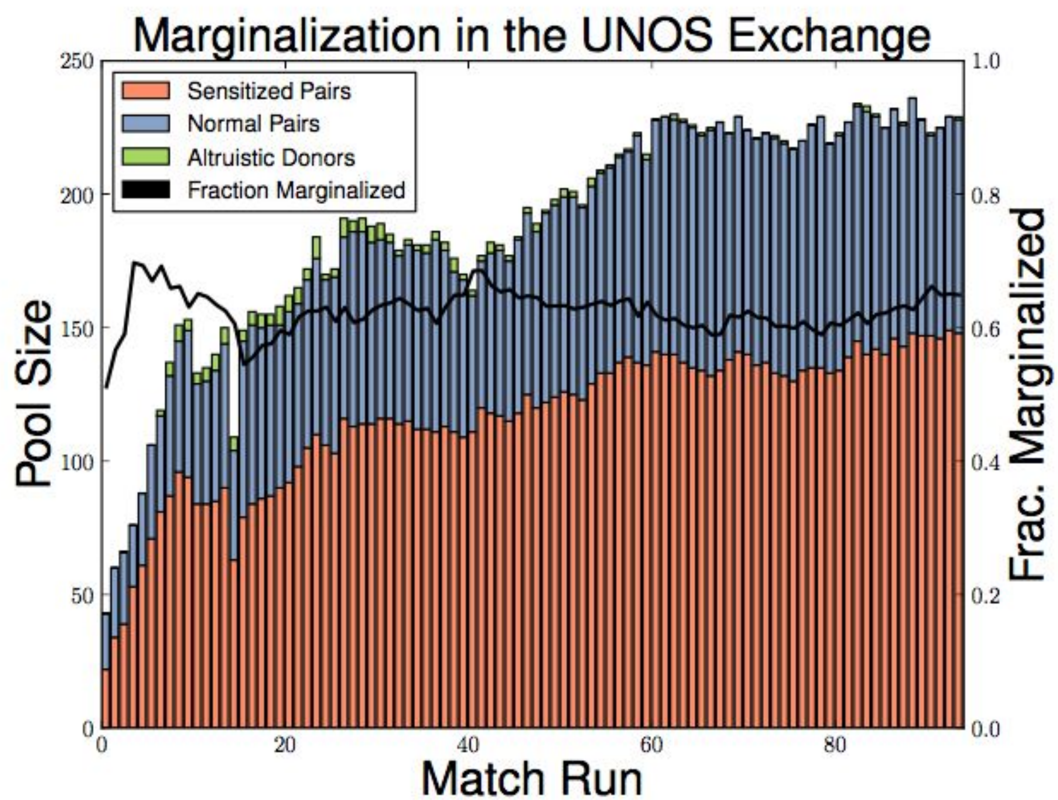
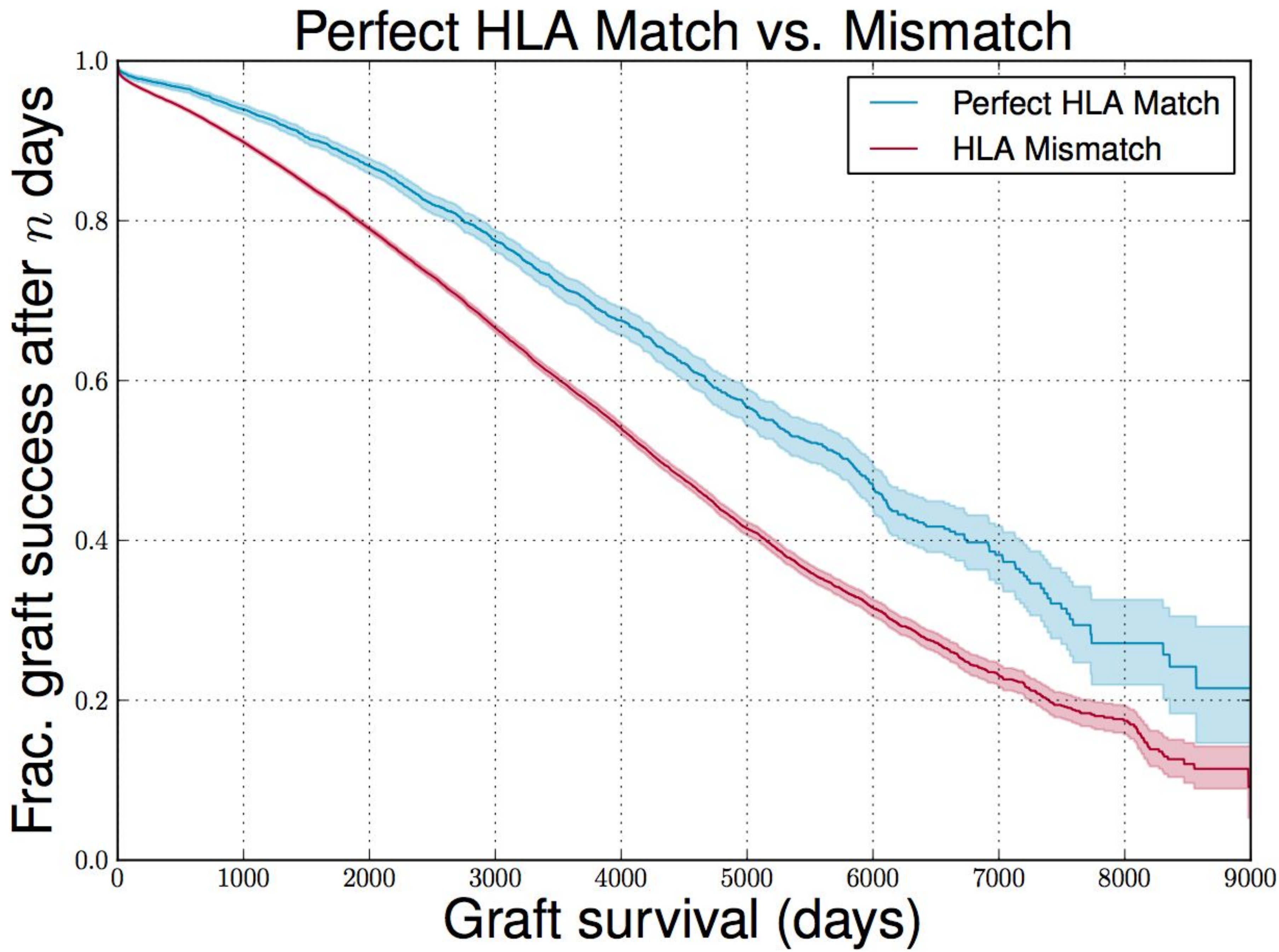


Figure 4: Evolution of the UNOS national kidney exchange. For each of 94 match runs (x-axis), the number of highly-sensitized or underage patients, non-highly-sensitized patients, and altruists are plotted (left y-axis), as well as the percentage of patients who are highly sensitized or underage as a percentage of the pool size (right y-axis).

Objective Functions: MaxLife

- How long does the Kidney survive in its new host?
- Available data:
  - 75,264 living donor transplant events between 11/1/1987 and 6/30/2013
  - 25% are failed, 75% are not marked as failed
  - Human Leukocyte Antigen (HLA) test results (tissue types) for patient and donors


Perfect HLA Match vs. Mismatch
Perfect HLA Match
HLA Mismatch
Frac. graft success after n days
Graft survival (days)
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000

MaxLife: How long does a Kidney last?

- Use a Cox Proportional Hazards Regression:

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

- Features (X):
  - recipient age
  - donor age - recipient age
  - recipient and donor HLA profile (3 components each: HLA-A, HLA-B, HLA-DR)
  - donor and recipient blood type compatibility

MaxLife: Modeling how long a Kidney Lasts

feature	$\exp(b_i)$	$SE(b_i)$	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 1: Learned weights via Cox regression after feature pruning for statistical significance.

- omitting HLA-B because it does not have a significant impact on hazard rate
- Example: unit increase in HLA-DR mismatch feature will result in a 1.087 * baseline hazard rate

MaxLife: Using the learned model to get edge weights

$$w(e) \propto \exp\left(-\sum_i x_i^e b_i\right)$$

- x^e is features for donor pair associated with edge e
- b_i are learned Cox regression weights
- Intuitively: lower risk edges have higher weights

What are Potentials?

- How much do we expect this pair to contribute *in the future?*
- Let Θ be each blood type combination:
  - 4 x 4 = 16 blood type pairings for pairs ({O-A, O-B, O-AB … AB-AB})
  - 4 blood types for altruists ({O, A, B, AB})
- For each type $\theta \in \Theta$, determine potential P_θ

Learning Potentials

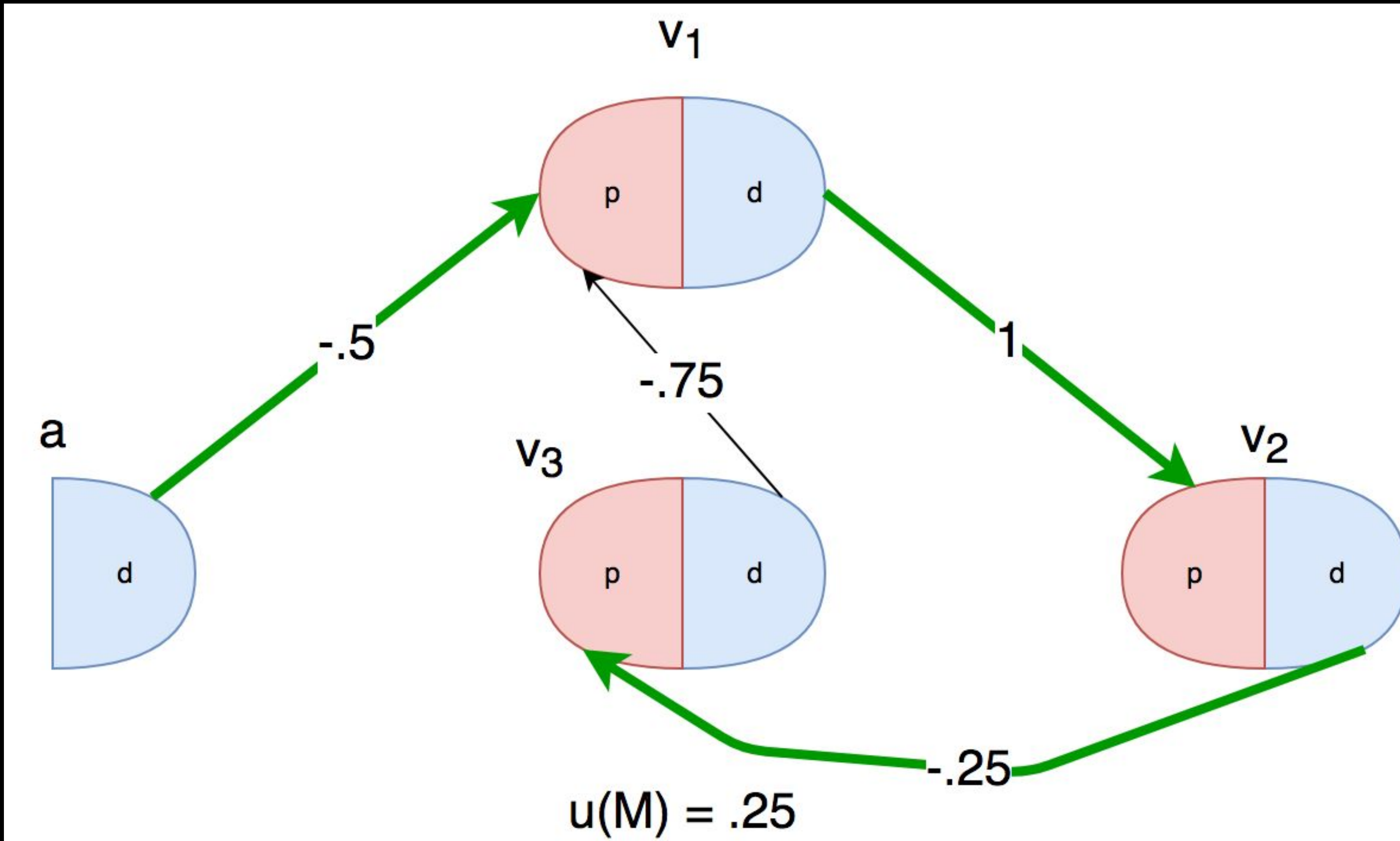
- Sequential Model-Based Algorithm Configuration (SMAC) used to derive potentials for each type [4]
- Process:
  - Select Potentials ("parameters")
  - Run simulation and calculate performance metric
  - Feed back performance metric
  - SMAC updates potentials accordingly
  - Repeat

Combining Edge Weights and Potentials

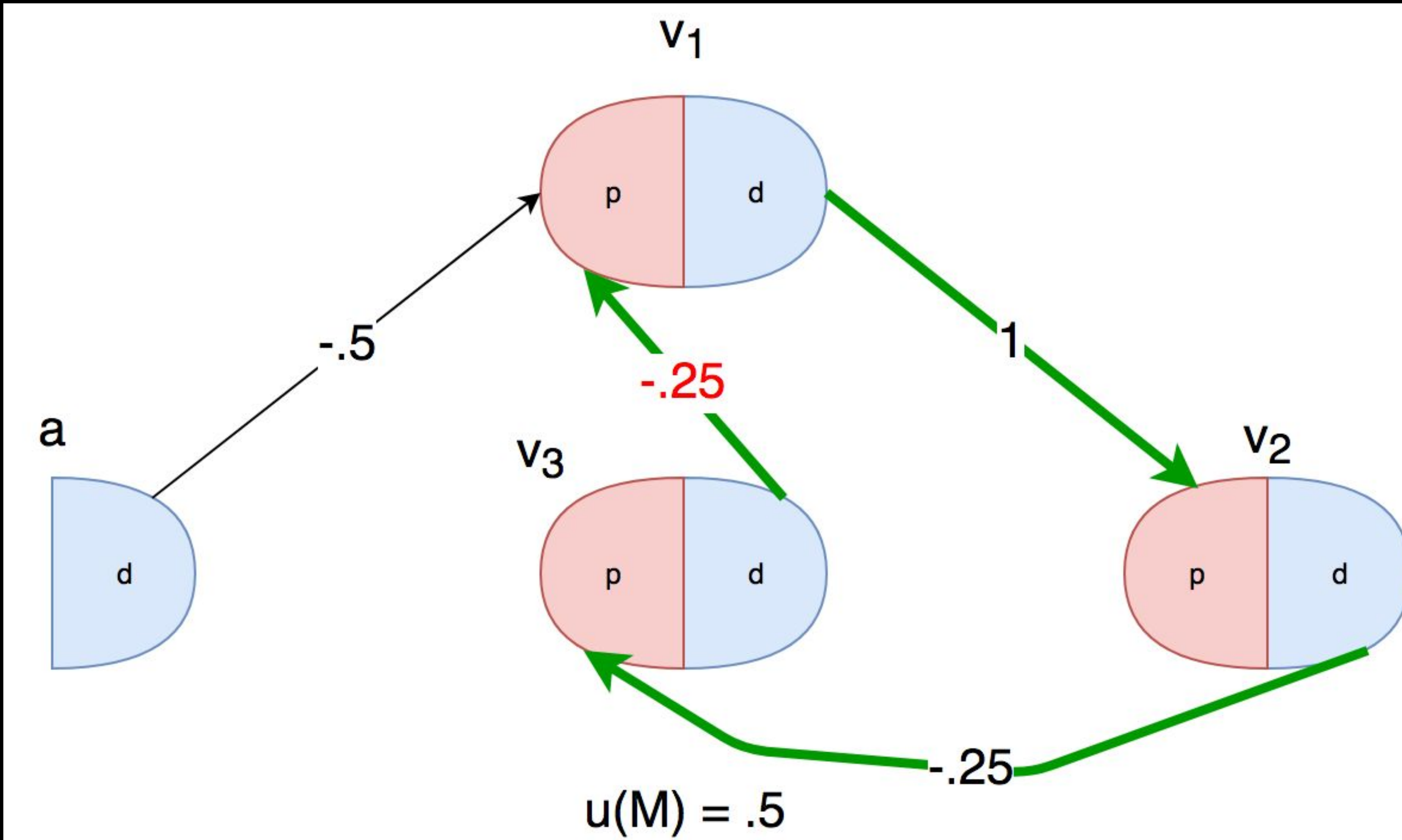
$$f_w(e) = w(e) \cdot (1 - P_{\theta_d} - P_{\theta_p})$$

- f_w (e) is updated edge weight
- w(e) is original edge weight
- $P_{\theta d}$ and $P_{\theta p}$ are donor and patient potentials, respectively

Example: High Potential Pair



Example: High Potential Altruists



Realistic Dynamic Failure-Aware Model

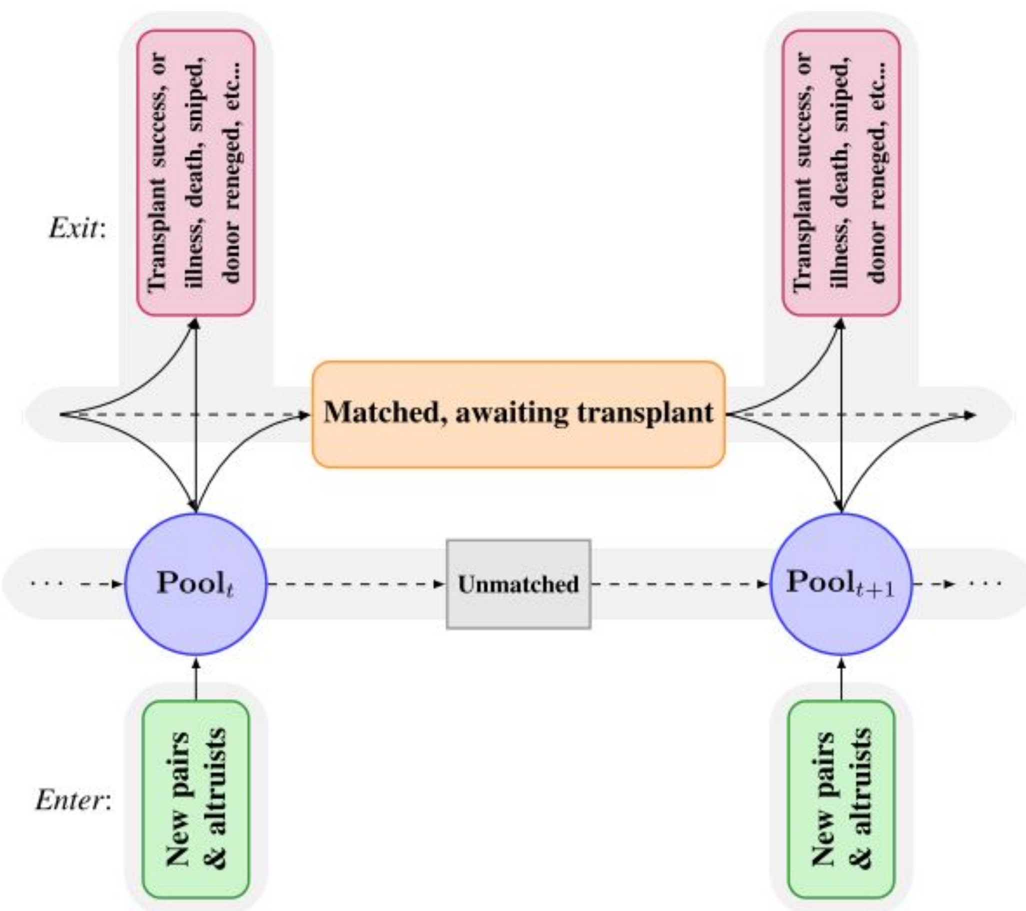
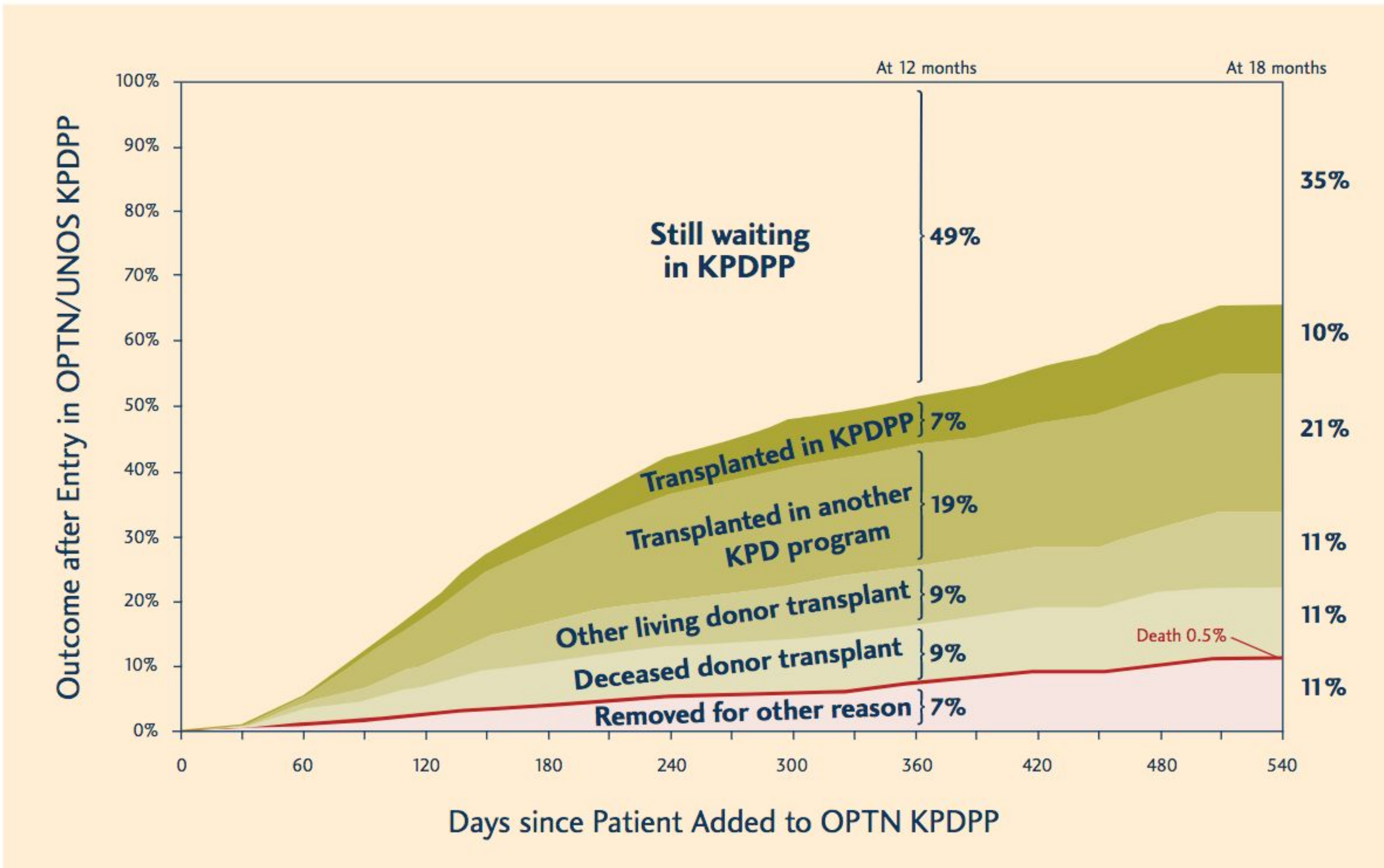


Figure 6: Dynamic kidney exchange.

- Patient/Donor Pairs enter the pool due to self/family member diagnoses or altruism
- Patients may be matched
  - If matched, they may exit if they die, switch to another exchange, or successfully receive a kidney
  - If matched, the may return if their matched partner exits
- Unmatched patients return to the pool

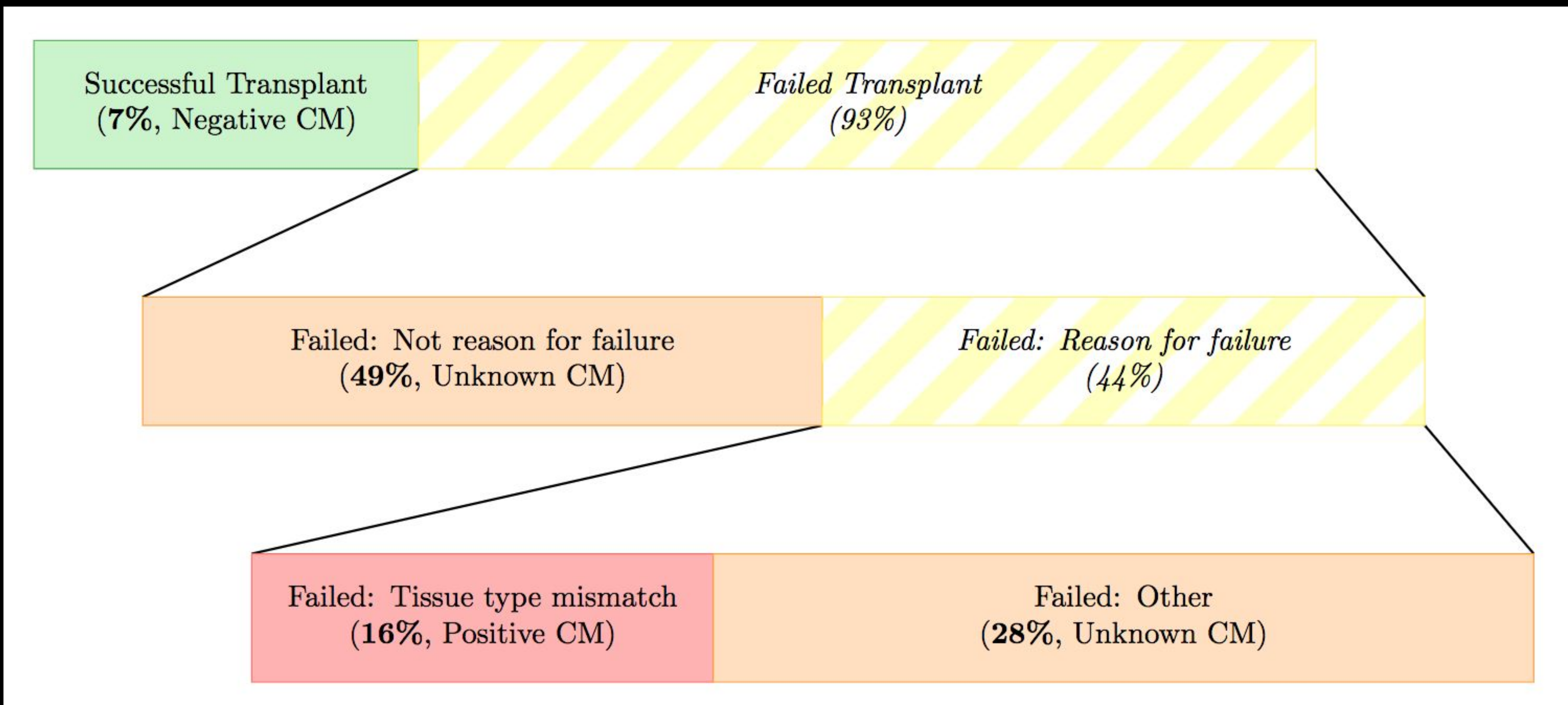
Realistic Dynamic Model: Informing Transition Probabilities

Figure 10A: Time to Transplant (or Other Outcome[10]) for Candidates Added to the KPDPP
Includes Match Run-Eligible Patients Added to the KPDPP[11] from Jan 1, 2012 – Nov 22, 2013



https://optn.transplant.hrsa.gov/media/1827/kpd_report.pdf

Edge Failure: Most Matches Fail



https://www.cs.cmu.edu/~sandholm/failure-aware%20kidney%20exchange.ec13.pdf

Failure Aware Model: Modeling Edge Failure

- How can we make utility take into account edge failure?
  - Can't just encode in edge weights: not independent
- Updated cycle utility

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

- Updated chain utility

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

Experiments

- All models use edge weights learned for objective functions MaxCard, MaxCard-Fair, and MaxLife
- 24 time steps representing 1 week intervals of matching
- Graphs randomly sampled from *real* UNOS data, 140 random graphs per configuration
- Baseline: Deterministic (not failure aware) and myopic (no potentials)
- FutureMatch: failure aware and potentials

Results: Median expected Gain/Loss

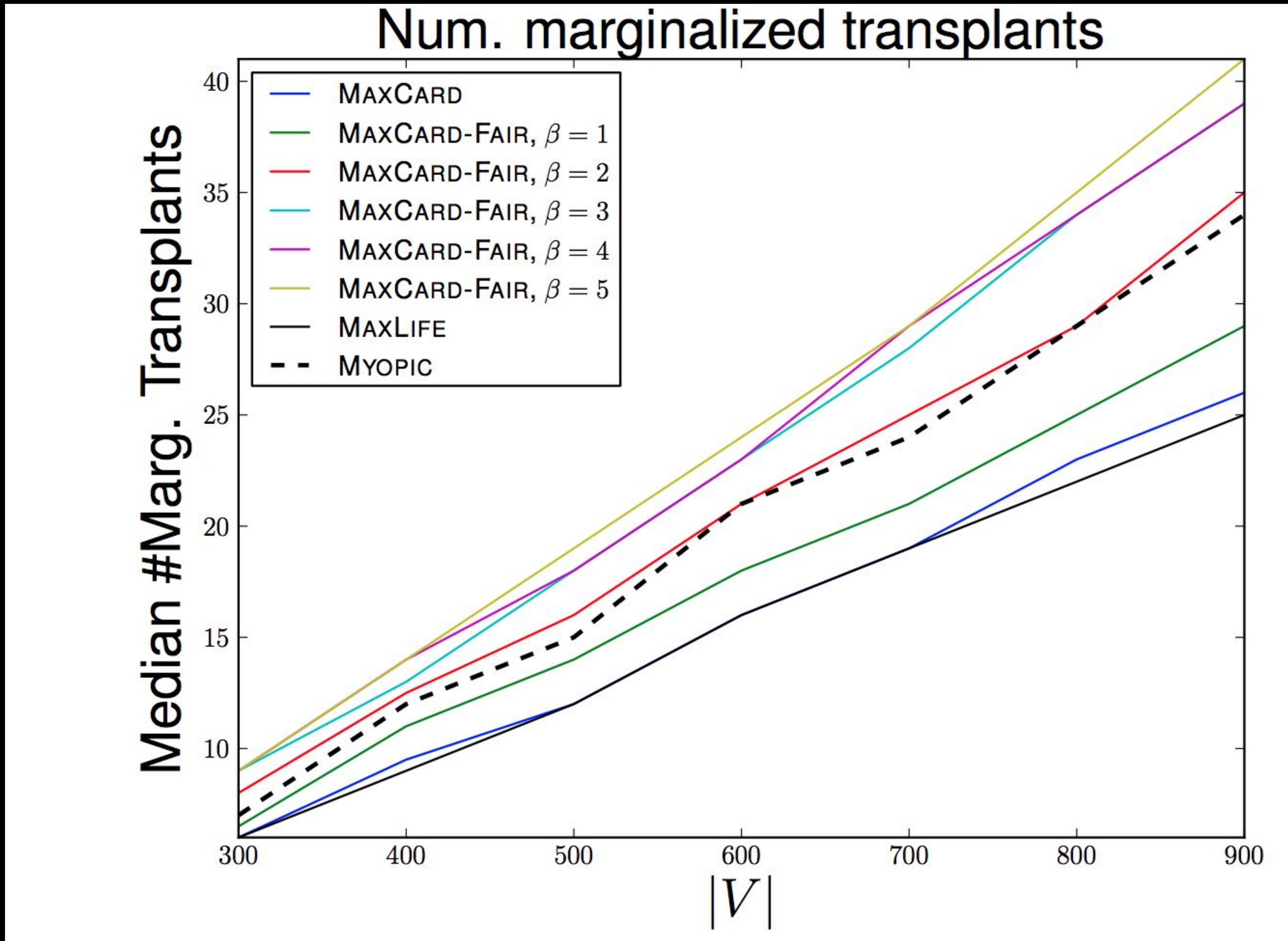
Total	$\lvert V\rvert = 300$ Gain	p	$\lvert V\rvert = 400$ Gain	p	$\lvert V\rvert = 500$ Gain	p	$\lvert V\rvert = 600$ Gain	p	$\lvert V\rvert = 700$ Gain	p	$\lvert V\rvert = 800$ Gain	p	$\lvert V\rvert = 900$ Gain	p
MAXCARD	+2	✓	+4	✓	+5	✓	+6	✓	+10	✓	+11	✓	+13	✓
MAXCARD-FAIR, $\beta = 1$	+1	✓	+4	✓	+6	✓	+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	✗

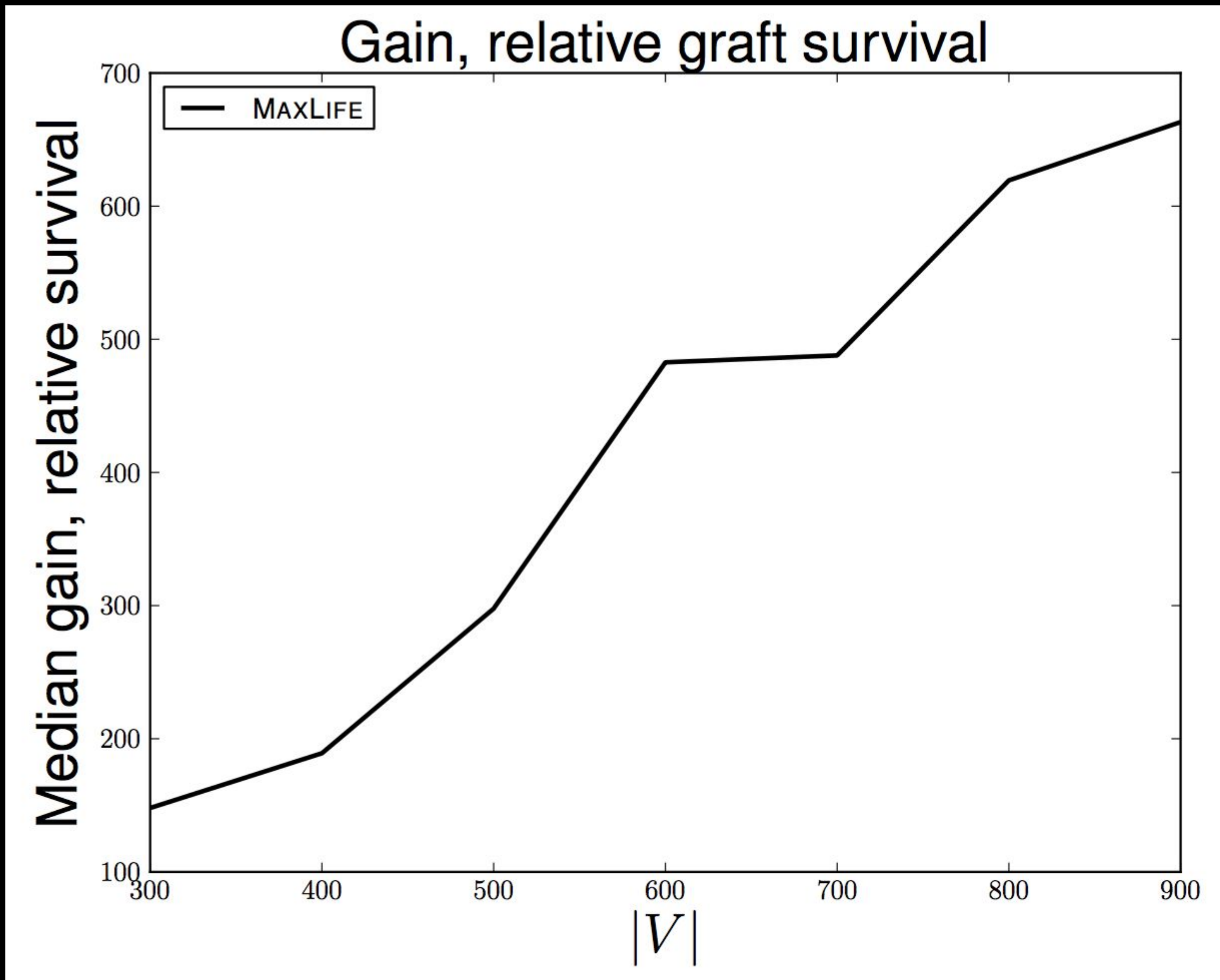
Table 2: Median gains in expected total number of transplants (top table) and total number of marginalized transplants (bottom table) under FUTUREMATCH. A ✓ represents statistical significance (Wilcoxon signed-rank test, $p \ll 0.01$).

Example (highlighted): In graphs of size 500, the median difference in number of matches for FutureMatch relative to the myopic, deterministic baseline is +3, and this is statistically significant.

Results: Marginalized Impact



Results: MaxLife Works



Conclusions

- Potentials and failure-aware modeling can be used to improve overall matches
- Special weighting is required to avoid bias towards non-marginalized patients
- Results are based on real data and attempt to mimic dynamic nature of matching

References

Unless otherwise indicated, all figures were taken from the paper or original work.

1. Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2012). Dynamic Matching via Weighted Myopia with Application to Kidney Exchange. *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 1340–1346. Retrieved from https://www.cs.cmu.edu/~sandholm/dynamicMatchingViaWeightedMyopia.aaai12.pdf%5Cnhttp://dl.acm.org/citation.cfm?id=2900728.2900918
2. Ashlagi, Itai, Duncan S. Gilchrist, Alvin E. Roth, and Michael A. Rees. "Nonsimultaneous Chains and Dominos in Kidney Paired Donation—Revisited (pdf)." *American Journal of Transplantation* 11, no. 5 (May 2011): 984–994.
3. Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2013). Failure-aware kidney exchange. Proceedings of the 14th ACM Conference on Electronic Commerce, 1(212), 323–340. https://doi.org/10.1145/2482540.2482596
4. Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2011). Sequential Model - Based Optimization for General Algorithm Configuration. *Lecture Notes in Computer Science*, *5*, 507–223. https://doi.org/10.1007/978-3-642-25566-3_40

Questions?