DYNAMIC MATCHING WITH DEEP REINFORCEMENT LEARNING

Samsara Counts, Linyi Chen, and Cameron Moy

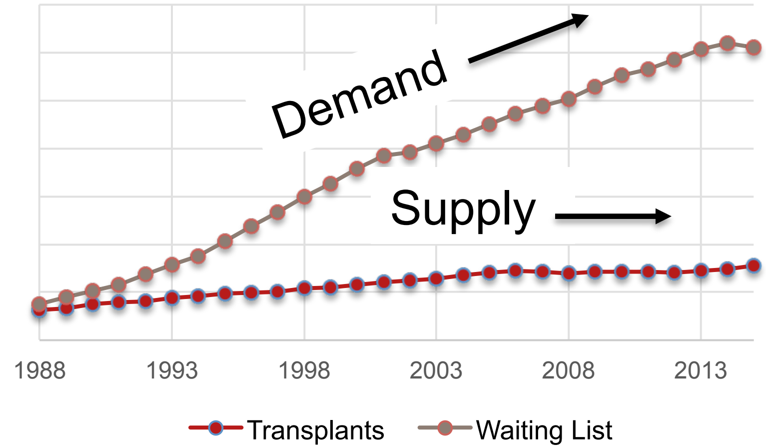
Advised by Dr. John P. Dickerson

COMPUTER SCIENCE
UNIVERSITY OF MARYLAND

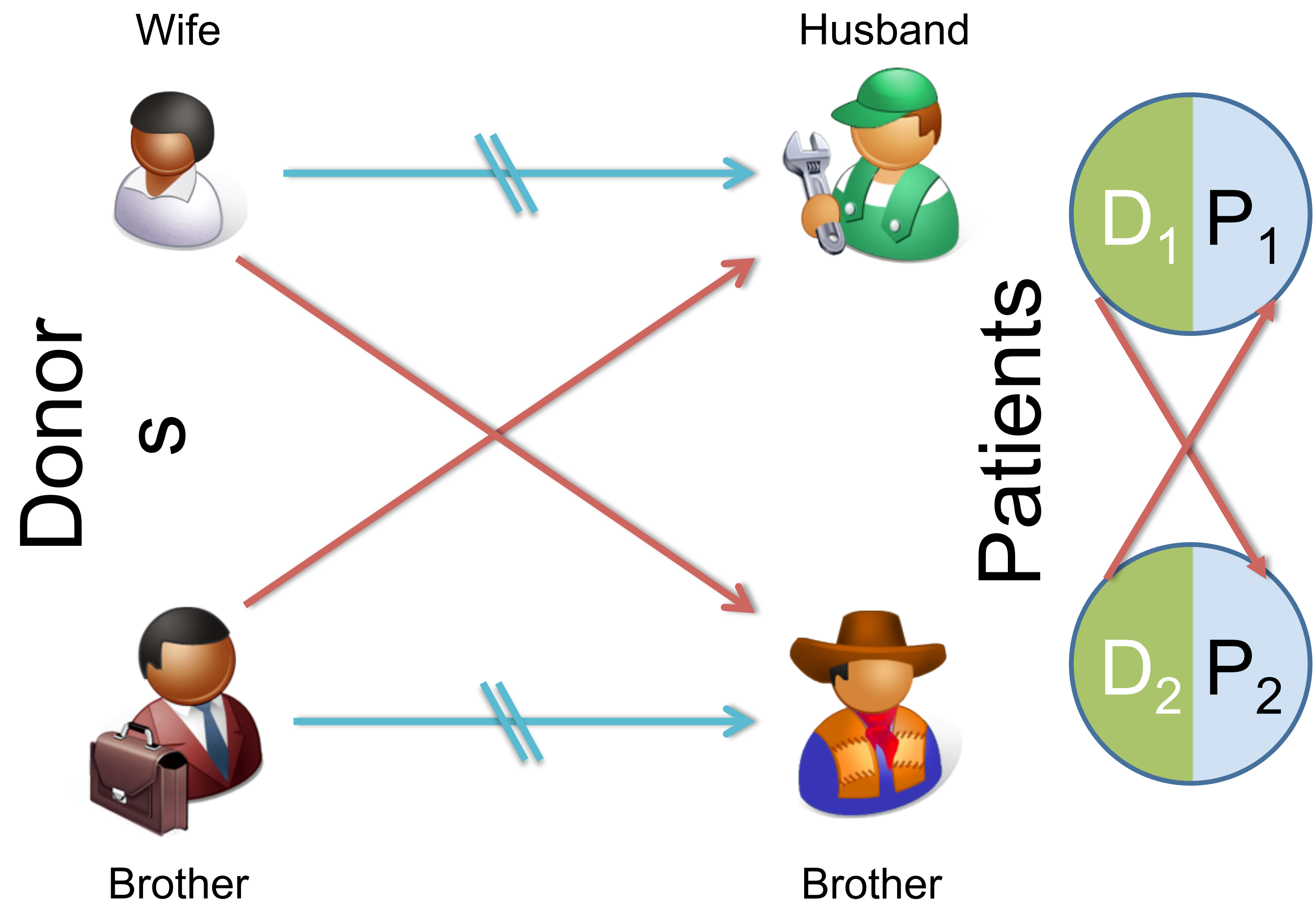
University of Maryland
Summer REU Talk
August 9, 2017

KIDNEY TRANSPLANTATION

- US waitlist: over 100,000 people
  - 36,157 added in 2014
- 4,537 people died while waiting
- 11,559 people received a kidney
- from the deceased donor waitlist
- 5,283 people received a kidney from a living donor
  - Some through **kidney exchanges**!
  - This talk: experience with **United Network for Organ Sharing** (UNOS) national kidney exchange



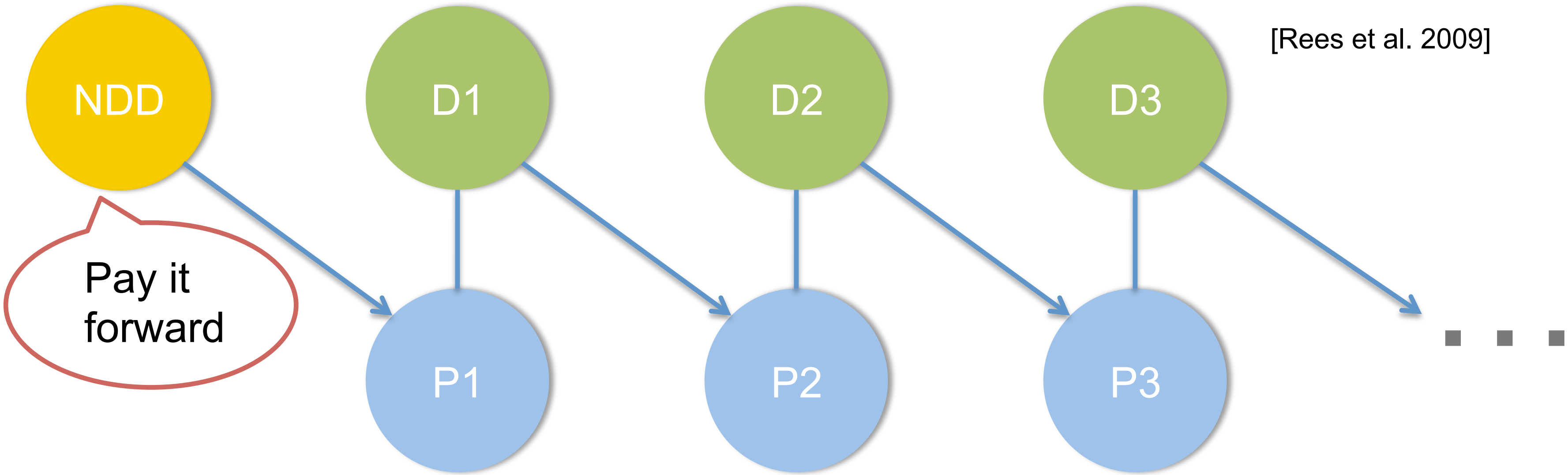
KIDNEY EXCHANGE: CYCLES



*(2- and 3-cycles, **all surgeries performed simultaneously**)*

KIDNEY EXCHANGE: CHAINS AND ALTRUISTIC DONORS

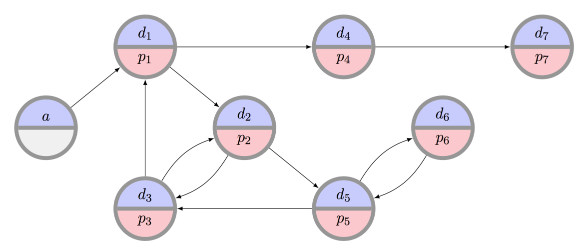
[Rees et al. 2009]



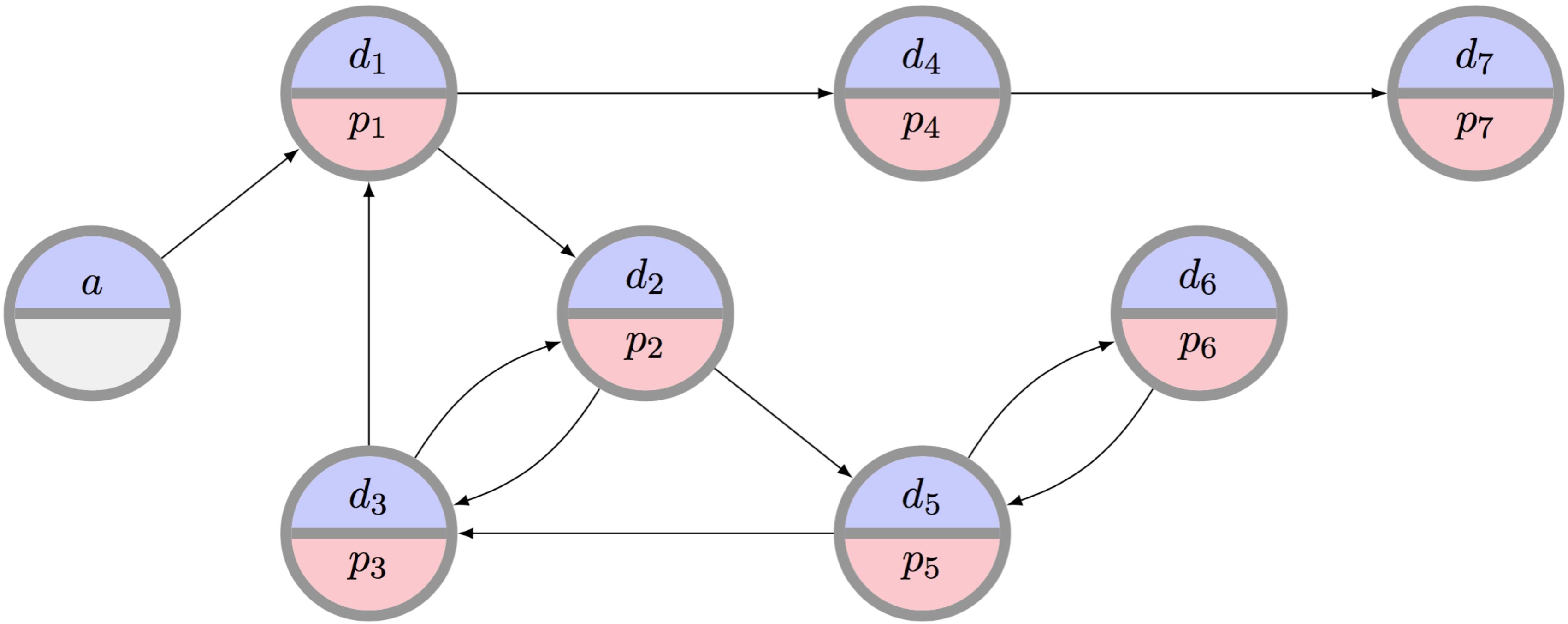
*Not executed simultaneously, so no length cap required based on logistic concerns …***but in practice edges fail, so *some* finite cap is used!**

We model these incompatible **donor/patient pairs** and **altruist donors** in a:

COMPATIBILITY GRAPH



MAX MATCHING (THE CLEARING PROBLEM)



The clearing problem is to find the "best" disjoint set of cycles of length at most *L*, and chains

- Typically, $2 \leq L \leq 5$ for kidneys (e.g., L=3 at UNOS)
- NP-hard (for L>2) in theory, really hard in practice

[Abraham et al. 07, Biro et al. 09]

[Glorie et al. 2014, Anderson et al. 2015, Plaut et al. 2016, Dickerson et al. 2016 ...]

DYNAMIC KIDNEY EXCHANGE

Kidney exchange is an inherently dynamic event that can be described by the evolution of its graph (additions and removals of edges and vertices).

THE PROBLEM:

What is "best":

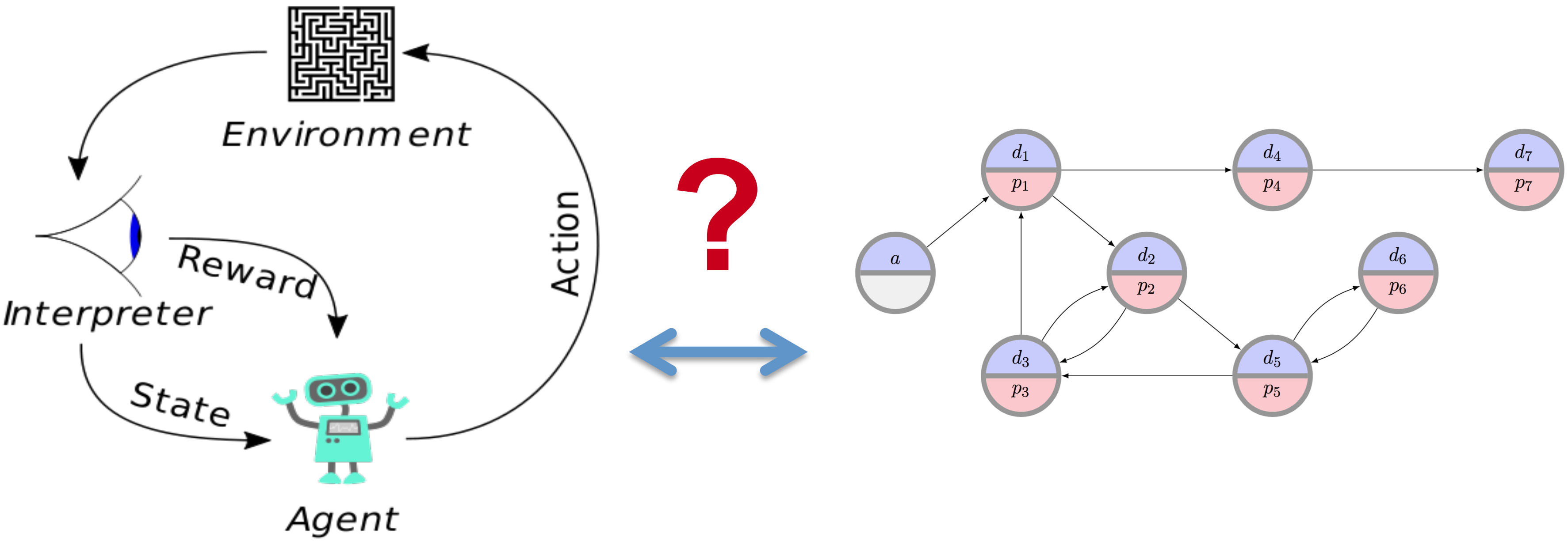
- Maximize matches right now or over time?
- Maximize transplants or matches?
- Fairness? Incentives? Ethics? Legality?

Key Questions:

How often should we match kidney-donor pairs? Should we clear the market **now** by finding the max cardinality matching, or **wait** for more vertices and edges to come in?

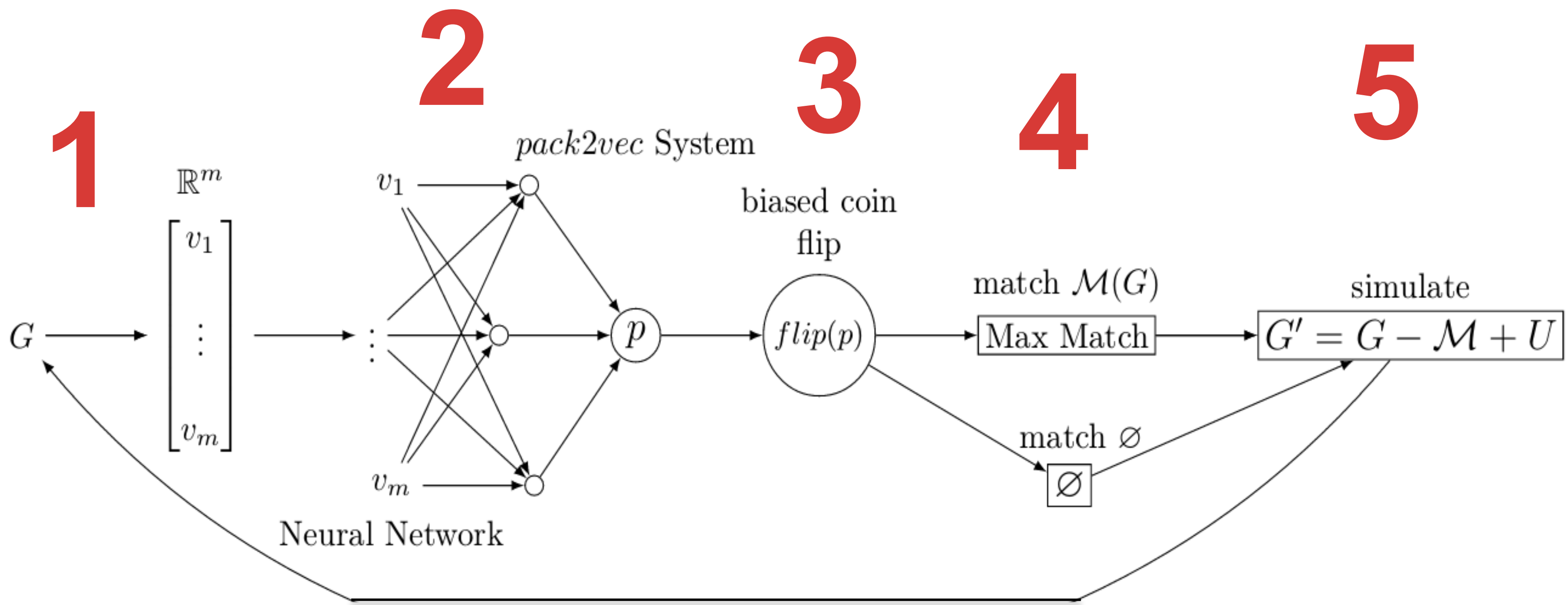
OUR SOLUTION: *LEARN* HOW, USING ARTIFICIAL INTELLIGENCE (DEEP REINFORCEMENT LEARNING)

MODELING KIDNEY EXCHANGE WITH REINFORCEMENT LEARNING



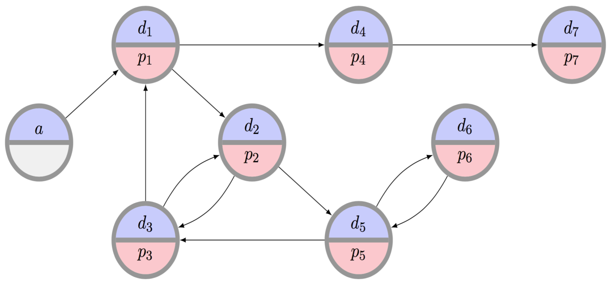
How *would* you model kidney exchange with reinforcement learning?
What do we want to reward?
What actions can we take?

LEARNING TO MATCH IN DYNAMIC ENVIRONMENTS: OUR SYSTEM



1. **Embedding** kidney exchange graphs as *fixed dimensional vectors*
2. **Neural network** uses those vectors to **learn** appropriate policy—a outputs probability
3. Flip a **biased coin**
4. Find and **match** maximum cardinality matching
5. **Simulate** kidney exchange environment and grow the graph

1. EMBEDDING



$$\Phi(G)$$

$$\mathbb{R}^m \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

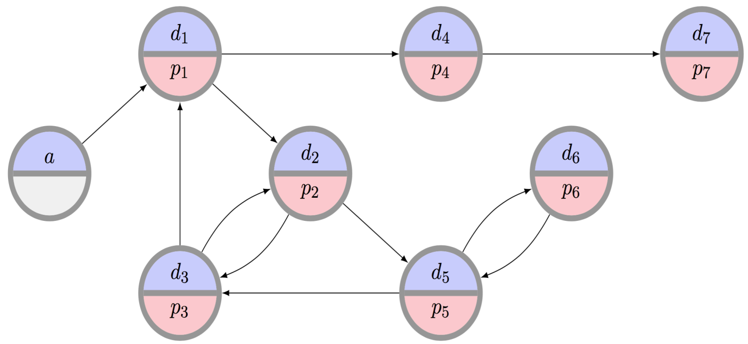
- **By convention, neural networks have fixed-sized input. However, our graphs are dynamic and change size over time!**
- **Thus, we need to embed the graph as a** *vector* **and** *maintain certain properties* **like** node neighborhood structure**.**

We use random walks to do so [Li, Campbell, Caceres 2017].

- Use random walk on a carefully selected initial distribution to capture temporal changes in probability distribution.

How does this work?

1. EMBEDDING DETAILS



- Consider an initial distribution $\vec{p_0}$ over vertices. Each temporal probability distribution is defined by a recurrence.
- $$\vec{p_t} = \frac{\alpha}{n}\vec{1} + (1-\alpha)\boldsymbol{W}\vec{p}_{t-1}$$
- Compute the pair-wise distance between distributions with respect to the steady-state distribution.

$$S_{st} = \frac{\vec{p_s}\boldsymbol{E}\vec{p_t}}{||\boldsymbol{E}^{-\frac{1}{2}}\vec{p_s}||_2||\boldsymbol{E}^{-\frac{1}{2}}\vec{p_t}||_2}$$

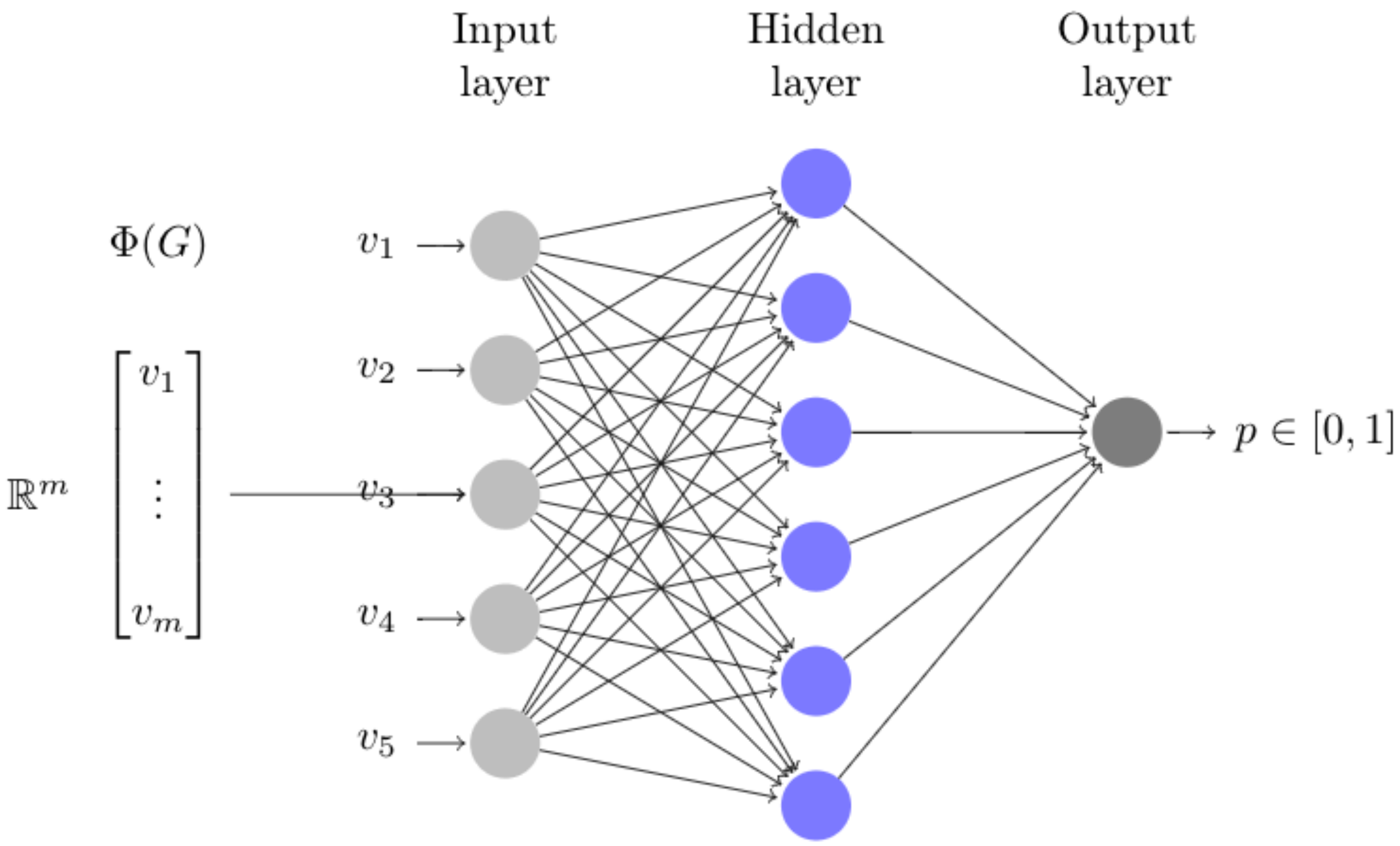
- Now stack feature vectors, embedding has nice properties.

SANITY CHECK FOR EMBEDDINGS: DISTANCE FUNCTIONS

Distance function to evaluate the degree of similarity/difference of two graphs.

- Goal: when two graphs are similar(largely different), their embedding vectors are close(far away) in terms of Euclidean distance;
- Optimal Distance Metric [Xu. et al. 2013]
  - Recognizes isomorphic graphs
  - NP-hard
- **Symmetric Kullback-Leibler Divergence**
  - Measures divergence rate between two probability distributions
  - Uses in-degree of vertices

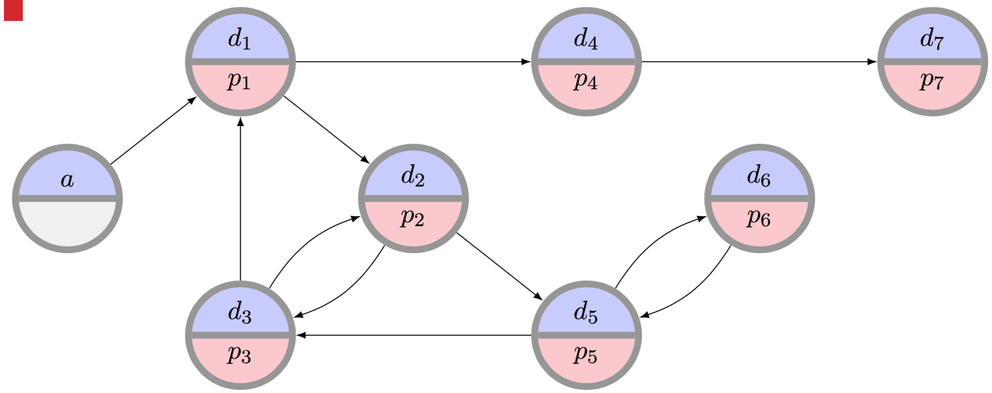
2. FEEDING INTO NEURAL NET



We feed our embedded graph into a neural network to output a learned probability for our biased coin flip

3. BIASED COIN FLIP W/LEARNED PROBABILITY

1



MAX MATCH

0

MATCH NOTHING (wait)

4. MAX MATCHING (THE CLEARING PROBLEM)—OR NOT

5. KIDNEY EXCHANGE SIMULATION – CHANGING THE INPUT GRAPH

To train the neural network, we must be able to simulate kidney exchange (graphs). We use actual exchange data from the United Network for Organ Sharing (UNOS) **to evolve our graphs.**

Previous work [Akbarpour 2017] develops a continuous model of kidney exchange that evolves based on a dynamic variant of an Erdős–Rényi graph model drawing from a **Poisson process**.

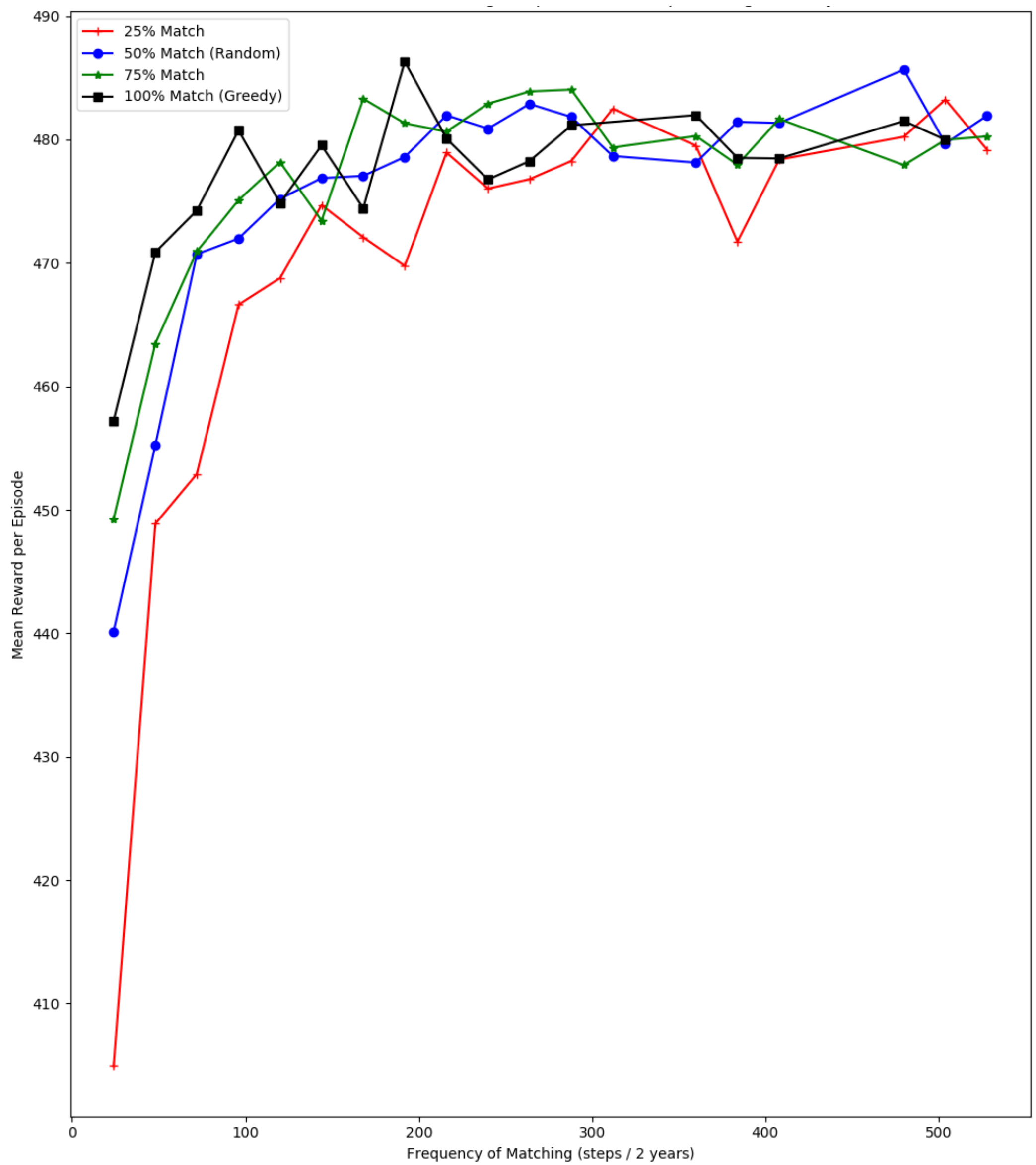
We adapt this model, making it discrete to use for reinforcement learning.

RESULTS

Our utility function—how we define our rewards—is based on **maximizing the cardinality of people matched.**

Q: What does this graph imply?

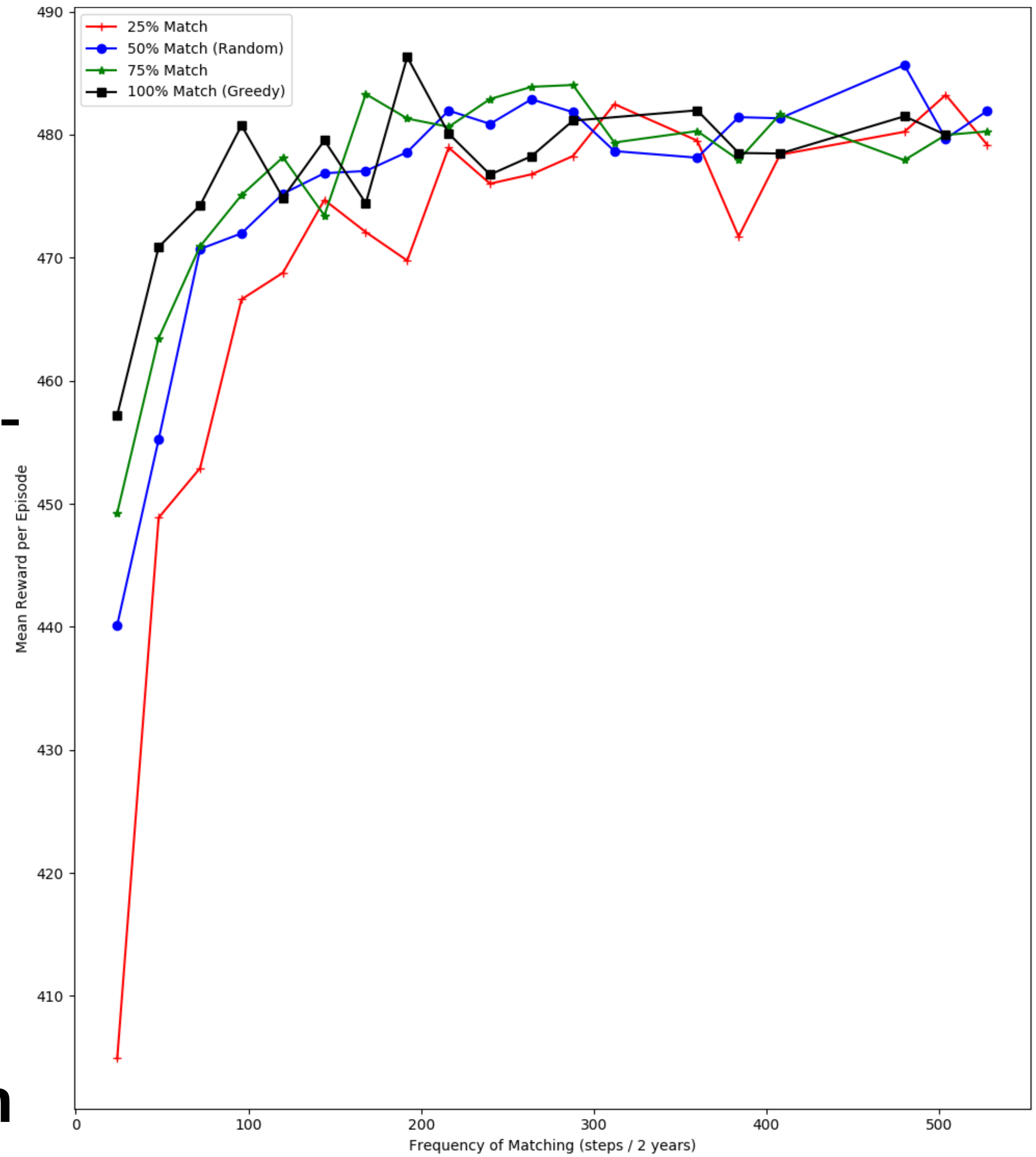
Mean Reward of Different Policies over Varying Matching Frequencies



RESULTS

- Batch frequency is **mostly irrelevant** under our cardinality-of-people-matched **utility function**.
- Hence, neural networks *can't improve much* on this by learning some **optimal batch frequency policy.**

Mean Reward of Different Policies over Varying Matching Frequencies



NEXT STEPS

- The **action space** we have is rather coarse: *max match* or *don't match*. We can make the space **continuous** by learning weights for graph features, which then inform the matching algorithm, giving us finer control over the matching.
- We may consider other methods of **embedding** that encode more information about *graph structures* in the exchange.
- We may consider other **utility functions** than the naive maximum cardinality formulation. For example, we might focus on the amount of *highly sensitized patient*s we match.

ACKNOWLEDGEMENTS

- Thank you to Prof. Gasarch and Prof. Khuller for organizing the CAAR REU program!
- Thanks to Prof. Dickerson for advising us, and Pengcheng (UMD grad student) for working with us on the project.
- Also, thank you to the National Science Foundation for supporting CAAR!