Learning to Build Solutions in Stochastic Matching Problems Using Flows (Student Abstract)

William St-Arnaud[1, 3], **Margarida Carvalho**[1], **Golnoosh Farnadi**[1, 2, 3]
[1]Université de Montréal
[2]McGill University
[3]Mila
william.st-arnaud@umontreal.ca

Abstract

Generative Flow Networks, known as GFlowNets, have been introduced in recent times, presenting an exciting possibility for neural networks to model distributions across various data structures. In this paper, we broaden their applicability to encompass scenarios where the data structures are optimal solutions of a combinatorial problem. Concretely, we propose the use of GFlowNets to learn the distribution of optimal solutions for kidney exchange problems (KEPs), a generalized form of matching problems involving cycles.

Introduction

In the literature, there are many combinatorial optimization models motivated by real-world problems. Most of these problems are usually tackled using mixed-integer programming techniques such as branch-and-bound. They are usually formulated as mathematical programs and fed to a solver that finds an optimal solution. These optimization techniques aim to be effective methods to search the solution space for a solution that is optimal, or close to optimal, and feasible, i.e., satisfying a set of constraints.

One particular example of a combinatorial optimization problem that has a real-life application is a set of problems referred to as kidney exchange problems (KEPs). These are allocation problems where the resources to be distributed are organs. Their solutions are referred to as exchange plans and their goal is typically concerned with designing a mechanism to select an exchange plan. In kidney exchanges, the evolution of the patients-donor pool is subject to uncertainty and they often have multiple optimal solutions; any policy for selecting a solution needs to account for the impact it can have on its patients. Different approaches have been tried to define fairness in the context of kidney exchanges. The two main flavours either prioritize outcomes with certain properties (Dickerson, Procaccia, and Sandholm 2014) or give guarantees to patients in the form of distributions over solutions (Farnadi et al. 2021). These definitions of fairness are related to the static case (i.e., single-round KEP), but in the dynamic case, it is not obvious what is a good definition of fairness. The two main obstacles to definitions in this setting are uncertainty and computing candidate solutions to build a distribution in the first place.

Objective: The goal of this research project is to devise efficient methods to sample exchange plans that match a high number of patients, while also guaranteeing some fairness notions as part of the allocation mechanism. We will characterize some natural fairness notions that can arise in the context of dynamic KEPs and propose methods to compute exchange plans in order to push the evolution of the KEP pool in a desired direction, i.e. whose matching dynamics fulfill high transplant counts and fair allocations.

Related Work

Before diving into the proposed research projects on the dynamic setting, we first summarize the work that has been done so far on the static case. In St-Arnaud, Carvalho, and Farnadi (2023), we combine a utilitarian objective with a fair allocation mechanism, referred to as a fairness scheme. This idea is motivated by the need for kidney exchange allocation policies to match as many patients as possible, while also ensuring that the allocation of compatible pairs is done in a fair and equitable way. We successfully demonstrate that it is possible to effectively balance utility and fairness in two multi-objective frameworks which we refer to as the *social welfare program* and the *Nash social welfare program* (Gass and Saaty 1955). We provide an efficient column generation method for both frameworks in order to compute our solutions. We refer the reader to supplementary material for a less cursory overview of our results.

In order to design fair allocation mechanisms for kidney exchange systems, we have identified the need to efficiently compute multiple optimal solutions or exchange plans. We have surveyed the literature and identified several candidate methods that can sample solutions. From the literature on Monte Carlo sampling, we have found Monte Carlo Markov chain methods (MCMC) such as variations of the Metropolis-Hastings algorithm that can converge to some desired probability distribution over time. In the reinforcement learning (RL) literature, policy gradient methods such as proximal policy optimization (PPO) (Schulman et al. 2017) and deep Q-learning (Hausknecht and Stone 2015) have both been deployed to learn policies. This exploration of available methods will allow us to identify the most promising sampling methods and also understand their limitations in the context of KEPs.

Approach

Because kidney exchange systems are run over time and since the pool is a function of previously realized exchange plans, we must extend the dual utilitarian-fair design of policies to the dynamic setting. As an initial step towards this goal, we apply GFlowNets (Bengio et al. 2023) to learn to sample exchange plans proportionally to a reward function. This enables us to explore the space of solutions efficiently by providing a sampler that can return high-reward exchange plans. We train a GFlowNet with multiple trajectories on both a single KEP graph (OptLearn) and multiple graphs (OptLearn-Multi), where trajectories are built from sequences of selected cycles in order to form exchange plans. We use a trajectory-decomposable loss in order to satisfy the trajectory balance equation, giving us a flow (in practice, this equation is approximately satisfied). We compare the number of patients included in sampled exchange plans against the optimal solution given by a MIP solver (OptMIP) and two random baselines selecting cycles greedily according to size (RandGreedy) and uniformly at random (RandUniform). The results can be found in table 1.

Method	Expected ratio of optimal solution **Single instance**	 **Multiple instances**
OptMIP	1.000	1.000
RandUniform	0.873	0.795
RandGreedy	0.877	0.821
OptLearn	**0.928**	-
OptLearn-Multi	-	**0.897**

Table 1: Performance of OptLearn against baselines

We also measure the capacity of the learning procedure to generalize to larger instances (i.e. KEP graphs) by training on trajectories from multiple KEP graphs of size 50 (i.e. 50 patient-donor pairs) and sampling solutions from graphs of size 100 and 200 (see table 2). Again, we see that our approach outperforms the baselines on larger instances. While not definitive, our initial results for the static case are encouraging by showing that it is possible to sample such solutions. Again, we encourage the reader to seek a more detailed description of our results in the supplementary materials. We wish to stress the fact that this work on static KEPs is a preliminary step towards tackling the general case of dynamic KEPs. We will eventually need to extend the learning of the flows to a dynamic setting and to also account for fairness requirements. This is what we will set out to do in our future research, as discussed below.

Method	Expected Ratio of Optimal Solution 50	 100	 200
OptMIP	1.000	1.000	1.000
RandUniform	0.829	0.795	0.631
RandGreedy	0.841	0.821	0.706
OptLearn-Multi	**0.936**	**0.872**	**0.782**

Table 2: Training on instances of size (i.e. λN) 50 and testing against instances of size 50, 100, and 200.

Future Work

The work that we have done so far in static KEPs forms a strong basis on which we can build extensions to tackle the dynamic setting. The ultimate goal of this research project is to compute fair and high-utility KEP policies that are not myopic. To help us achieve this goal in the dynamic setting, we propose to extend kidney exchange optimization, define fairness notions, combine utilitarian and fair allocation schemes and provide interpretable policies to patients and decision-makers to help guide their decisions and also better predict changes in the environment. For this purpose, we will continue to apply GFlowNets to KEPs by extending their use to the dynamic setting. With this line of inquiry, we wish to determine if this method is able to effectively learn to sample solutions proportionally to a pre-determined reward when we extend the optimization horizon. We will make use of a simulator to capture the uncertainty of the environment in the dynamic case (e.g. arrival and departure of patient-donor pairs). As part of our experiments, we will try out different flow parameterizations in order to select the best one for our needs. The modelization will follow, as a basis, the work of Pan et al. (2023). Our hypothesis is that we will be able to show that this method provides an improvement over other alternatives such as MCMC and energy-based methods. We would also like to compare against other RL baselines such as deep-Q learning.

References

Bengio, Y.; Lahlou, S.; Deleu, T.; Hu, E. J.; Tiwari, M.; and Bengio, E. 2023. GFlowNet Foundations. *Journal of Machine Learning Research*, 24(210): 1–55.

Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2014. Price of fairness in kidney exchange. In *AAMAS*, 1013–1020.

Farnadi, G.; St-Arnaud, W.; Babaki, B.; and Carvalho, M. 2021. Individual Fairness in Kidney Exchange Programs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13): 11496–11505.

Gass, S. I.; and Saaty, T. L. 1955. Parametric Objective Function (Part 2)—Generalization. *Journal of the Operations Research Society of America*, 3(4): 395–401.

Hausknecht, M.; and Stone, P. 2015. Deep recurrent q-learning for partially observable mdps. In *2015 aaai fall symposium series*.

Pan, L.; Zhang, D.; Jain, M.; Huang, L.; and Bengio, Y. 2023. Stochastic Generative Flow Networks. In Evans, R. J.; and Shpitser, I., eds., *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, 1628–1638. PMLR.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal Policy Optimization Algorithms. *ArXiv*, abs/1707.06347.

St-Arnaud, W.; Carvalho, M.; and Farnadi, G. 2023. Adaptation, Comparison and Practical Implementation of Fairness Schemes in Kidney Exchange Programs. arXiv:2207.00241.