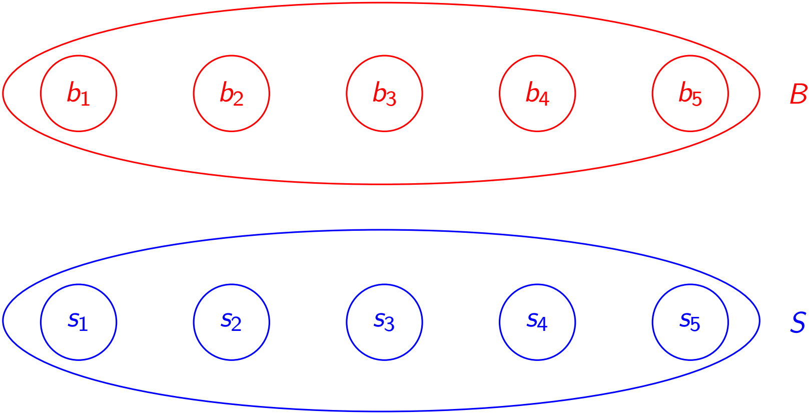
Optimal control of the N bipartite matching model

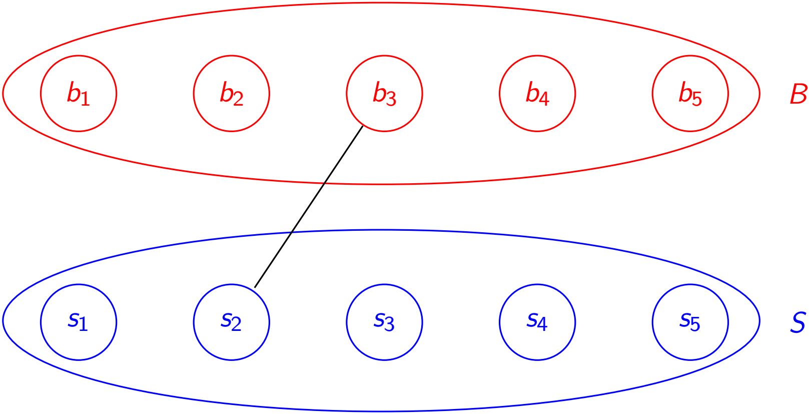
Arnaud Cadas[1 2] Josu Doncel[2 3] Ana Bušić[1 2]
[1]Département d'informatique de l'École normale supérieure, CNRS, PSL Research University, Paris, France
[2]Inria, Paris, France,
[3]Université du Pays Basque, Bilbao, Espagne

12 march 2018

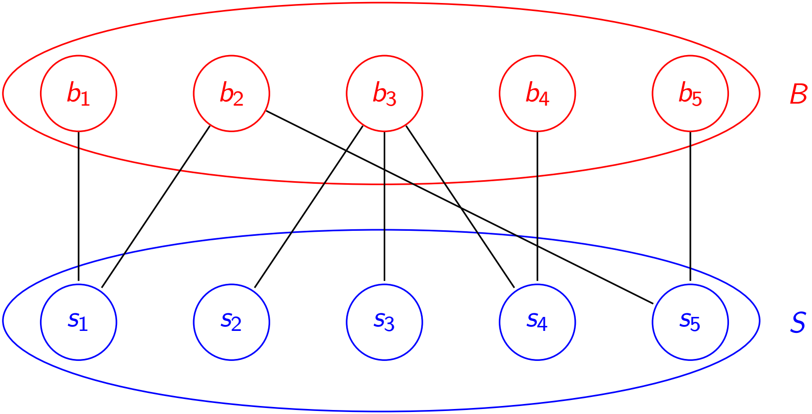
What is a matching model ?



What is a matching model ?



Hard constraints



Different types of matching models

- Stable marriage
- Online stochastic matching
- Fully dynamic matching model

Results on fully dynamic matching models

FCFS

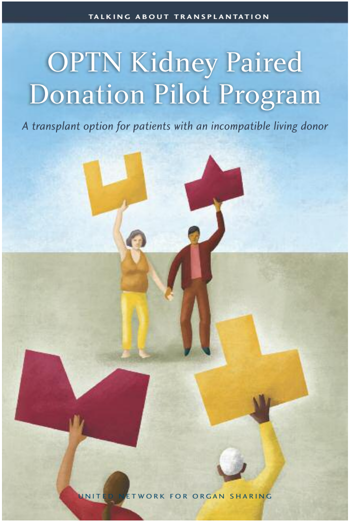
R. Caldentey, E. Kaplan, G. Weiss. FCFS infinite bipartite matching of servers and customers. *Adv. Appl. Probab.*, 2009.
I. Adan and G. Weiss. Exact FCFS matching rates for two infinite multitype sequences. *Operations Research*, 2012.
I. Adan, A. Bušić, J. Mairesse and G. Weiss. Reversibility and further properties of FCFS infinite bipartite matching. *ArXiv*, 2015.

Stability

A. Bušić, V. Gupta, and J. Mairesse. Stability of the bipartite matching model. *Adv. in Appl. Probab.*, 2013.
J. Mairesse and P. Moyal. Stability of the stochastic matching model. *J. Appl. Probab.*,2016.

Optimization

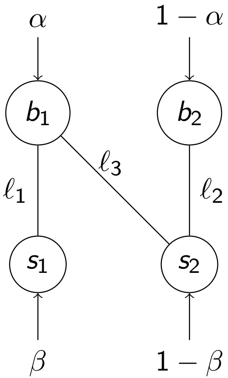
A. Bušić and S. Meyn. Approximate optimality with bounded regret in dynamic matching models. *ArXiv*, 2016.



I. Ashlagi, P. Jaillet, and V. H. Manshadi. Kidney exchange in dynamic sparse heterogenous pools. *ACM EC*, 2013.

The case N

A matching graph with two types of buyers and two types of sellers.



Stability condition: $\beta < \alpha$

The vector of the queue length of all the nodes (at time t):

$Q(t) = (q_k(t))_{k \in \{b_1, b_2, s_1, s_2\}}$.

Markov Decision Process

The matchings are carried out after the arrivals at time t:

$$X(t) = Q(t) + A(t)$$
$$Q(t+1) = X(t) - U(X(t))$$

Dynamics of the system: $X(t+1) = X(t) - U(X(t)) + A(t+1)$
Markov Decision Process:

- Transitions: $P_u(x, x') = \sum_a \mathbb{1}_{x-u+a=x'} \mathbb{P}(A = a)$
- Costs: $C_u(x, x') = \sum_{k \in \{b_1, b_2, s_1, s_2\}} c_k x'_k$
- Policy: $\pi = (u_t)_{t \geq 0}$

Goal

$$v^*(x_0) = \inf_{\pi} \lim_{N\to\infty} \sum_{t=0}^{N-1} \theta^t \mathbb{E}^{\pi}_{x_0}[c(X(t))]$$

where $\theta \in (0,1)$ is a discount factor.

$$\begin{aligned} v^*(x_0) &= \inf_{\pi} \sum_{t=0}^{\infty} \theta^t \mathbb{E}^{\pi}_{x_0}[c(X(t))] \\ &= \inf_{u_0}\{c(x_0) + \theta \sum_{x_1} \inf_{\pi}\{\sum_{t=1}^{\infty} \theta^{t-1} \mathbb{E}^{\pi}_{x_1}[c(X(t))]\} P_{u_0}(x_0, x_1)\} \\ &= \inf_{u_0}\{c(x_0) + \theta \sum_{x_1} v^*(x_1) P_{u_0}(x_0, x_1)\}\} \end{aligned}$$

$$\text{Operators: } L_u v(x) = c(x) + \theta \sum_{x'} v(x') P_u(x, x')$$

$$Lv(x) = \min_u L_u v(x)$$

Bellman equation

$$v^* = Lv^*$$

How to solve this equation ?

Value iteration

$$v^{n+1} = Lv^n \text{ with } v^0 \equiv 0$$

L is a contraction mapping ($\|Lv - Lw\| \leq \theta \|v - w\|$) and using the Banach Fixed-Point Theorem, v^n will converge towards v^* when n tends to infinity.

See chapter 6 from M.L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. Wiley, 2005.

Theorem (Puterman 2005)

Assumptions to handle the unbounded costs. Let V^σ (the set of structured value functions) and D^σ (the set of structured decisions) be such that

(a) $v \in V^\sigma$ *implies that* $Lv \in V^\sigma$;

(b) $v \in V^\sigma$ *implies that there exists a decision* $u' \in D^\sigma$ *such that* $u' \in \arg\min_u L_u v$;

(c) V^σ *is a closed subset of the set of value functions by simple convergence.*

Then, there exists an optimal stationary policy that belongs to Π^σ *(the set of structured policies).*

e_{ℓ_j} is the vector with 1 on both nodes of the edge j and 0 elsewhere.

Definition (Increasing)

We say that a function v is increasing in ℓ_1 or $v \in \mathcal{I}_{\ell_1}$ if

$$v(x + e_{\ell_1}) \geq v(x) \quad \forall x \in \mathcal{X}.$$

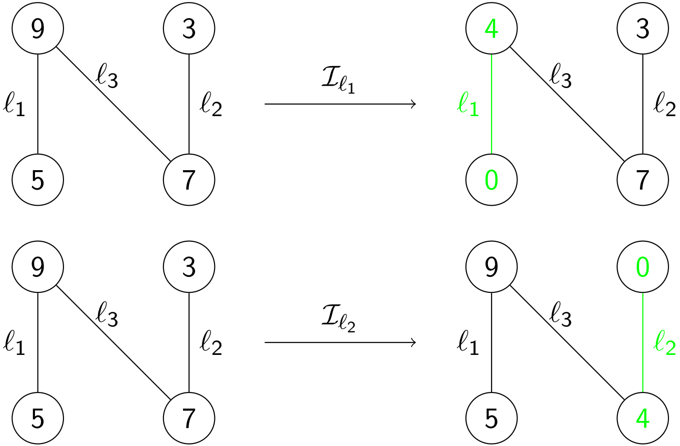
Likewise, v is increasing in ℓ_2 or or $v \in \mathcal{I}_{\ell_2}$ if

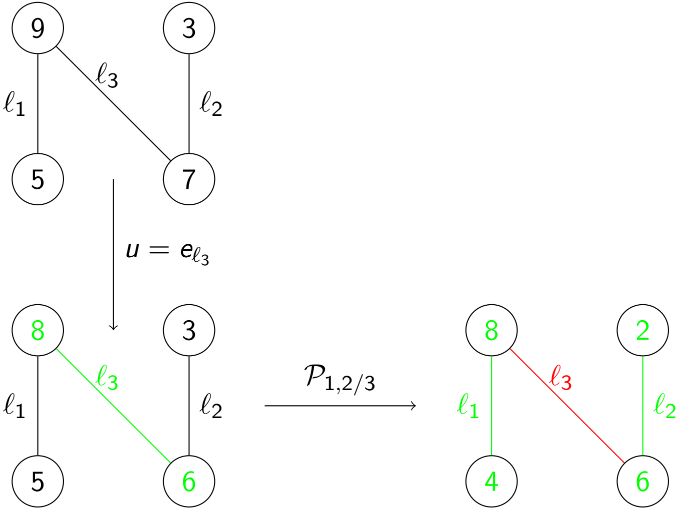
$$v(x + e_{\ell_2}) \geq v(x) \quad \forall x \in \mathcal{X}.$$

Definition (Priority of ℓ_1 and ℓ_2 over ℓ_3)

A function $v \in \mathcal{P}_{1,2/3}$ if

$$v(x + e_{\ell_1} + e_{\ell_2} - e_{\ell_3}) \geq v(x) \quad \forall x \in \mathcal{X}.$$


9
3
ℓ_3
ℓ_1
ℓ_2
5
7
$\mathcal{I}_{\ell_1}$
4
3
ℓ_3
ℓ_1
ℓ_2
0
7
9
3
ℓ_3
ℓ_1
ℓ_2
5
7
$\mathcal{I}_{\ell_2}$
9
0
ℓ_3
ℓ_1
ℓ_2
5
4


9
3
5
7
ℓ_1
ℓ_2
ℓ_3
$u = e_{\ell_3}$
8
3
5
6
$\mathcal{P}_{1,2/3}$
8
2
4
6

Proposition

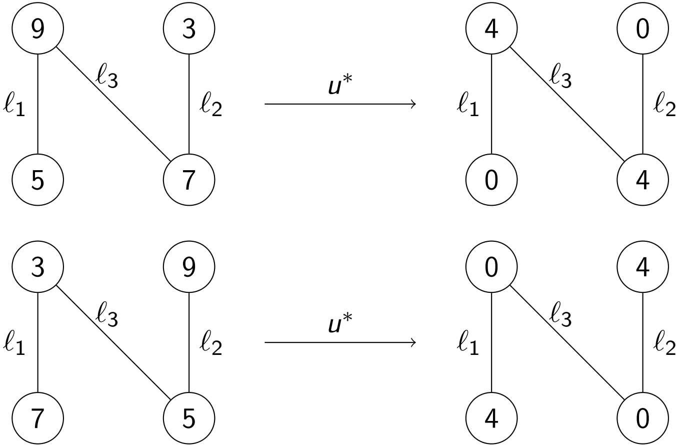
Let $v \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2} \cap \mathcal{P}_{1,2/3}$. For any $x \in \mathcal{X}$, there exists $u^* \in U_x$ such that $u^* \in \arg\min_{u \in U_x} L_u v(x)$, $(u^*)_{s_1} = min(x_{b_1}, x_{s_1})$ and $(u^*)_{b_2} = min(x_{b_2}, x_{s_2})$.

Lemma

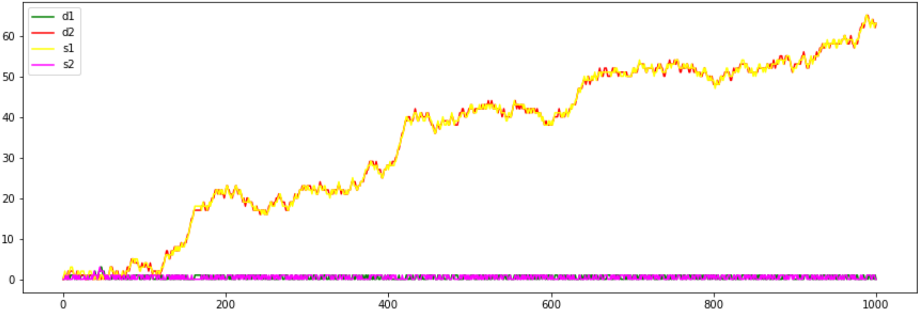
If $v \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2} \cap \mathcal{P}_{1,2/3}$, then $Lv \in \mathcal{P}_{1,2/3}$.

Lemma

If a function $v \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2} \cap \mathcal{P}_{1,2/3}$, then $Lv \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2}$.


9
3
5
7
ℓ_1
ℓ_2
ℓ_3
u^*
4
0
0
4
ℓ_1
ℓ_2
ℓ_3
3
9
7
5
ℓ_1
ℓ_2
ℓ_3
u^*
0
4
4
0
ℓ_1
ℓ_2
ℓ_3

Motivation for threshold policy



One realisation of $X(t)$ for $t = 1, \cdots, 1000$

A. Bušić and S. Meyn. Approximate optimality with bounded regret in dynamic matching models. *ArXiv*, 2016.

Threshold policy

Definition (Threshold-type matching policy)

A matching policy u_x is said to be of threshold type in ℓ_3 with priority to ℓ_1 and ℓ_2 if:

- $(u_x)_{s_1} = \min(x_{b_1}, x_{s_1})$
- $(u_x)_{b_2} = \min(x_{b_2}, x_{s_2})$
- $(u_x)_{b_1} = \min(x_{b_1}, x_{s_1}) + k_t(x)$
- $(u_x)_{s_2} = \min(x_{b_2}, x_{s_2}) + k_t(x)$

where $k_t(x) = \begin{cases} 0 & \textit{if } x_{b_1} - x_{s_1} \leq t \\ x_{b_1} - x_{s_1} - t & \textit{else} \end{cases}$

Convexity in ℓ_3

Definition (Convexity)

A function v is convex in ℓ_3 or $v \in \mathcal{C}_{\ell_\ni}$ if $v(x + e_{\ell_3}) - v(x)$ is nondecreasing in ℓ_3. i.e.,
$v(x + 2(e_{\ell_3})) - v(x + e_{\ell_3}) \geq v(x + e_{\ell_3}) - v(x)$.

Let $x \in \mathcal{X}$, $m_{\ell_1} = \min(x_{s_1}, x_{b_1})$, $m_{\ell_2} = \min(x_{s_2}, x_{b_2})$. We define:

- $\tilde{u}_x = (m_{\ell_1}, m_{\ell_2}, m_{\ell_1}, m_{\ell_2})$
- $K_x = \begin{cases} \{0\} & \textit{if } x_{b_1} \leq x_{s_1} \\ \{0, \cdots, \min(x_{b_1} - x_{s_1}, x_{s_2} - x_{b_2})\} & \textit{else} \end{cases}$
- $k^*(x) = \max\{k' \in \arg\min_{k \in K_x} L_{\tilde{u}_x + k e_{\ell_3}} v(x)\}$

Lemma

Let v be a function that is convex in ℓ_3. Let $\underline{x} \in \mathcal{X}$, $\overline{x} = \underline{x} + e_{\ell_3}$. Let $k \in K_{\underline{x}}$ such that $k \leq k^*(\underline{x})$. Then,

$$k^*(\underline{x}) \leq k^*(\overline{x}) \leq k^*(\underline{x}) + 1$$

Moreover, if we suppose that $k^*(\underline{x}) > 0$, we have: $k^*(\overline{x}) = k^*(\underline{x}) + 1$.

Proposition

Let $v \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2} \cap \mathcal{P}_{1,2/3} \cap \mathcal{C}_{\ell_3}$. There exists $u^* \in U_x$ such that u^* is a matching policy of threshold type in ℓ_3 with priority to ℓ_1 and ℓ_2 and $u^* \in \arg\min_{u \in U_x} L_u v(x)$.

Conjecture: the convexity property propagates, i.e if $v \in \mathcal{I}_{\ell_1} \cap \mathcal{I}_{\ell_2} \cap \mathcal{P}_{1,2/3} \cap \mathcal{C}_{\ell_3}$, then $Lv \in \mathcal{C}_{\ell_3}$

Average cost problem

We also considered the average cost problem:

$$g^*(x_0) = \inf_{\pi} \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathbb{E}_{x_0}^{\pi}[c(X(t))]$$

In this case, we use the structured policies method in the discounted problem and make θ tends towards 1.

Optimal threshold

Proposition

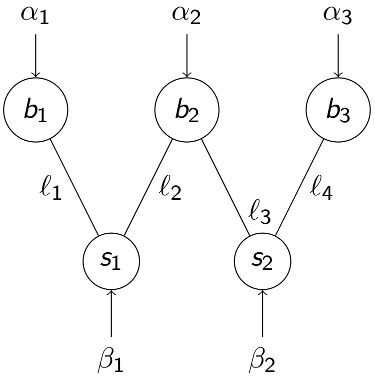
Let $\rho = \frac{\beta(1-\alpha)}{\alpha(1-\beta)} \in (0,1)$, $R = \frac{c_{s_1}+c_{b_2}}{c_{b_1}+c_{s_2}}$ *and* $\Pi^{T_{\ell_3}}$ *be the set of matching policy of threshold type in* ℓ_3 *with priority to* ℓ_1 *and* ℓ_2*. Assume that the cost function is a linear function. The optimal threshold* t^**, which minimize the average cost on* $\Pi^{T_{\ell_3}}$*, is*

$$t^* = \begin{cases} \lceil k \rceil & \text{if } f(\lceil k \rceil) \leq f(\lfloor k \rfloor) \\ \lfloor k \rfloor & \text{else} \end{cases}$$

where $k = \frac{\log \frac{\rho-1}{(R+1)\log \rho}}{\log \rho} - 1$ *and*
$f(x) = (c_{b_1} + c_{s_2})x + (c_{b_1} + c_{b_2} + c_{s_1} + c_{s_2})\frac{\rho^{x+1}}{1-\rho} - (c_{b_1} + c_{s_2})\frac{\rho}{1-\rho} +$
$((c_{b_1} + c_{s_1})\alpha\beta + (c_{b_2} + c_{s_2})(1-\alpha)(1-\beta) + (c_{b_2} + c_{s_1})(1-\alpha)\beta + (c_{b_1} + c_{s_2})\alpha(1-\beta)$

Future work

The case W:



General bipartite graph.