

Stochastic Online Optimization

Jian Li

Institute of Interdisciplinary Information Sciences

Tsinghua University



清华大学
Tsinghua University

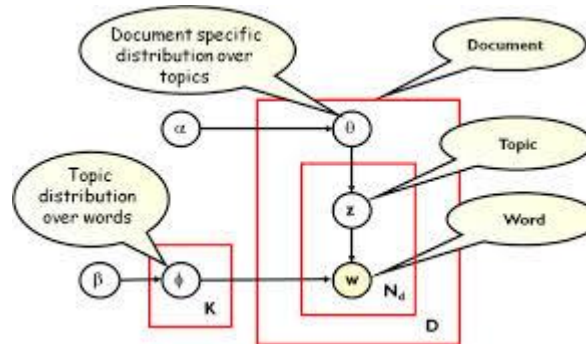
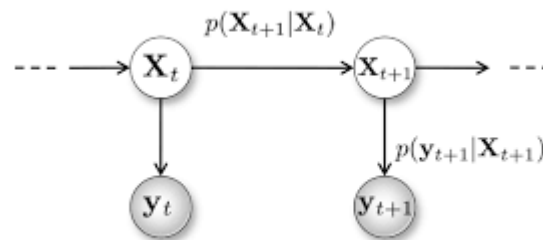
- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Uncertain Data and Stochastic Model

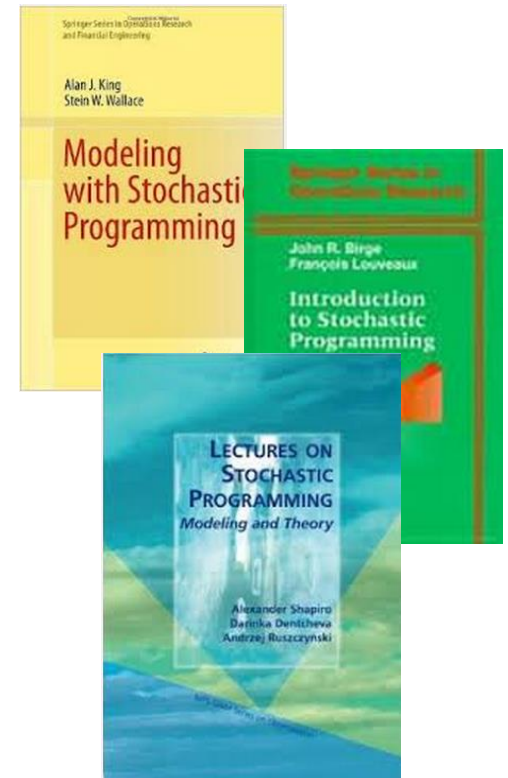
- Data Integration and Information Extraction
- Sensor Networks; Information Networks
- Probabilistic models in machine learning

Sensor ID	Temp.
1	Gauss(40,4)
2	Gauss(50,2)
3	Gauss(20,9)
...	...

Probabilistic databases



Probabilistic Models in machine learning



Stochastic models in operation research

Stochastic Optimization

- Initiated by Danzig (linear programming with stochastic coefficients)
- Instead of having a deterministic input, we have a distribution of inputs. Goal: optimize the expectation of some functional of the objective value.
- Many problems are #P-hard (even PSPACE-hard)
- Focus: polynomial time approximation algorithms
 - α -approximation (approximation factor)
 - $\frac{ALG}{OPT} \leq \alpha$ (minimization problem)

Online Algorithms

- Time = 1, 2, 3, ...
- At time t , make your decision irrevocably (only know the input up to time t)
- Competitive analysis: $\frac{ALG}{\text{Offline } OPT}$
 - The competitive ratio is typically determined by the worst case input sequence (**too pessimistic sometimes**)
 - **Stochastic Online Optimization:** Instead of considering the worst case, assume that there is a distribution of inputs (especially in the era of big data)

Simons Institute

- <https://simons.berkeley.edu/>

algorithms & uncertainty

Algorithms and Uncertainty

Aug. 17 – Dec. 16, 2016

Workshops

Algorithms and Uncertainty Boot Camp

Aug. 22 – Aug. 26, 2016

Organizers: [Avrim Blum](#) (Carnegie Mellon University), [Anupam Gupta](#) (Carnegie Mellon University), [Robert Kleinberg](#) (Cornell University), [Stefano Leonardi](#) (Sapienza University of Rome), [Eli Upfal](#) (Brown University), [Adam Wierman](#) (California Institute of Technology)

Optimization and Decision-Making Under Uncertainty

Sep. 19 – Sep. 23, 2016

Organizers: [Nikhil Bansal](#) (Technische Universiteit Eindhoven; chair), [Shipra Agrawal](#) (Columbia University), [Robert Kleinberg](#) (Cornell University), [Kamesh Munagala](#) (Duke University), [Jay Sethuraman](#) (Columbia University), [Adam Wierman](#) (California Institute of Technology)

Learning, Algorithm Design and Beyond Worst-Case Analysis

Nov. 14 – Nov. 18, 2016

Organizers: [Avrim Blum](#) (Carnegie Mellon University; chair), [Nir Ailon](#) (Technion Israel Institute of Technology), [Nina Balcan](#) (Carnegie Mellon University), [Ravi Kumar](#) (Google), [Kevin Leyton-Brown](#) (University of British Columbia), [Tim Roughgarden](#) (Stanford University)

Organizers:

[Anupam Gupta](#) (Carnegie Mellon University; chair; co-chair), [Stefano Leonardi](#) (Sapienza University of Rome; co-chair), [Avrim Blum](#) (Carnegie Mellon University), [Robert Kleinberg](#) (Cornell University), [Eli Upfal](#) (Brown University), [Adam Wierman](#) (California Institute of Technology).

Long-Term Participants (including Organizers):

[Nir Ailon](#) (Technion Israel Institute of Technology), [Susanne Albers](#) (Technische Universität München), [Aris Anagnostopoulos](#) (Sapienza University of Rome), [Peter Auer](#) (University of Leoben), [Yossi Azar](#) (Tel Aviv University), [Nikhil Bansal](#) (Technische Universiteit Eindhoven), [Peter Bartlett](#) (UC Berkeley), [Eliyan Bitar](#) (Cornell University), [Avrim Blum](#) (Carnegie Mellon University), [Nicolo Cesa-Bianchi](#) (University of Milan), [Shiri Chechik](#) (Tel Aviv University), [Edith Cohen](#) (Google Research), [Artur Czumaj](#) (University of Warwick), [Amit Daniely](#) (Google Research), [Amos Fiat](#) (Tel Aviv University), [Fabrizio Grandoni](#) (IDSIA), [Anupam Gupta](#) (Carnegie Mellon University; chair; co-chair), [MohammadTaghi Hajiaghayi](#) (University of Maryland), [Longbo Huang](#) (Tsinghua University), [Sungjin Im](#) (UC Merced), [Ravi Kannan](#) (Microsoft Research India), [Sampath Kannan](#) (University of Pennsylvania), [Anna Karlin](#) (University of Washington), [Robert Kleinberg](#) (Cornell University), [Elias Koutsoupias](#) (University of Oxford), [Ravi Kumar](#) (Google), [Stefano Leonardi](#) (Sapienza University of Rome; co-chair), [Kevin Leyton-Brown](#) (University of British Columbia), [Jian Li](#) (Tsinghua University), [Na Li](#) (Harvard University), [Katrina Ligett](#) (Hebrew University and Caltech), [Aleksander Mądry](#) (Massachusetts Institute of Technology), [Yishay Mansour](#) (Tel Aviv University), [Ruta Mehta](#) (University of Illinois, Urbana-Champaign), [Jamie Morgenstern](#) (University of Pennsylvania), [Kamesh Munagala](#) (Duke University), [Viswanath Nagarajan](#) (University of Michigan), [Seffi Naor](#) (Technion Israel Institute of Technology), [Kameshwar Poola](#) (UC Berkeley), [Kirk Pruhs](#) (University of Pittsburgh), [Ram Rajagopal](#) (Stanford University), [Satish Rao](#) (UC Berkeley), [Benjamin Recht](#) (UC Berkeley), [Rhonda Righter](#) (UC Berkeley), [Tim Roughgarden](#) (Stanford University), [Piotr Sankowski](#) (University of Warsaw), [C. Seshadhri](#) (UC Santa Cruz), [Jay Sethuraman](#) (Columbia University), [Cliff Stein](#) (Columbia University), [Chaitanya Swamy](#) (University of Waterloo), [Marc Uetz](#) (University of Twente), [Eli Upfal](#) (Brown University), [Marilena Vendittelli](#) (Sapienza University of Rome), [Maria Viasiou](#) (Eindhoven University of Technology), [Jan Vondrák](#) (Stanford University), [Jean Walrand](#) (UC Berkeley), [Gideon Weiss](#) (University of Haifa), [Adam Wierman](#) (California Institute of Technology), [Bert Zwart](#) (CWI Amsterdam).

Research Fellows:

[Ilan Cohen](#) (Tel Aviv University), [Varun Gupta](#) (University of Chicago), [Thomas Kesselheim](#) (Max-Planck-Institute for Informatics and Saarland University), [Marco Molinaro](#) (PUC-Rio de Janeiro; Microsoft Research Fellow), [Benjamin Moseley](#) (Washington University in St. Louis), [Debmalya Panigrahi](#) (Duke University), [Xiaorui Sun](#) (Columbia University; Google Research Fellow), [Matt Weinberg](#) (Princeton University), [Qiaomin Xie](#) (University of Illinois at Urbana-

- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Problem Definition

Stochastic Matching [Chen, et al. ICALP'09]

- Given:
 - A probabilistic graph $G(V,E)$.
 - Existential prob. p_e for each edge e .
 - Patience level t_v for each vertex v .
- **Probing** $e=(u,v)$: The only way to know the existence of e .
 - We can probe (u,v) only if $t_u > 0, t_v > 0$.
 - If e indeed exists, we should add it to our matching.
 - If not, $t_u = t_u - 1, t_v = t_v - 1$.

Problem Definition

- Output: A strategy to probe the edges
 - Edge-probing: an (**adaptive** or **non-adaptive**) ordering of edges.
 - Matching-probing: k rounds; In each round, probe a set of disjoint edges
- Objectives:
 - Unweighted: *Max. $E[\text{cardinality of the matching}]$.*
 - Weighted: *Max. $E[\text{weight of the matching}]$.*

Motivations

- **Online dating**
 - **Existential prob. p_e** : estimation of the success prob. based on users' profiles.

eHarmony Relationship Questionnaire

Section 12: Communication Style

Please use the scale below to rate how well you believe each of the following words generally describes you.

	not at all		somewhat		very well	
1. I try to accommodate the other person's position	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I try to understand the other person	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
3. I try to be respectful of all opinions different from my own	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I try to resolve the conflict quickly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
5. I try to avoid disagreement	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Motivations

- **Online dating**

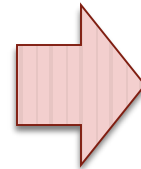
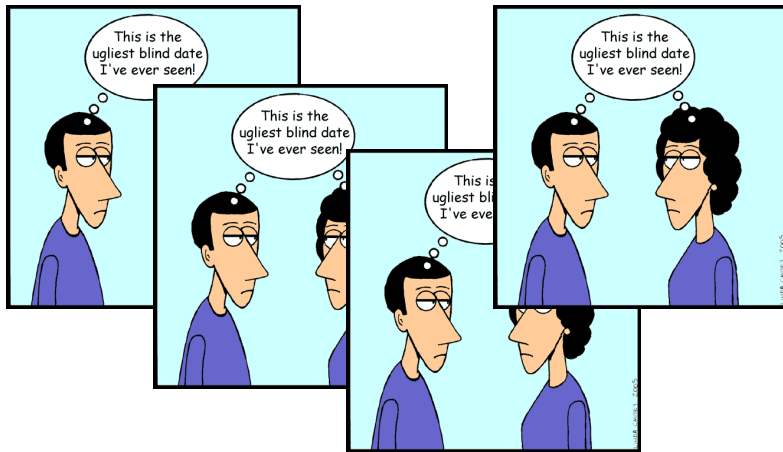
- **Existential prob. p_e** : estimation of the success prob. based on users' profiles.
- **Probing edge $e=(u,v)$** : u and v are sent to a date.



Motivations

- **Online dating**

- **Existential prob. p_e** : estimation of the success prob. based on users' profiles.
- **Probing edge $e=(u,v)$** : u and v are sent to a date.
- **Patience level**: obvious.



Motivations: Kidney Exchange

Lloyd Shapley

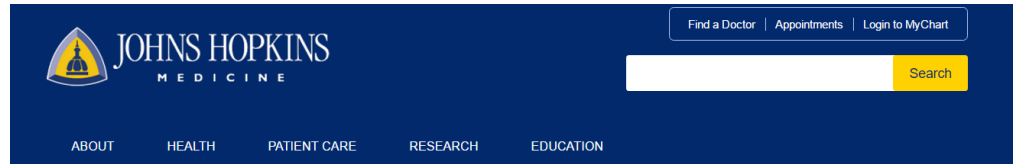


Shapley in 1980

Alvin E. Roth



Alvin E. Roth in Stockholm 2012



Comprehensive Transplant Center

Home About Us Programs Patient Information Living Donors Transplant Team Referring Physicians News and Events

- Overview
- Auto Islet
- Bone Marrow
- Cornea
- Heart
- Kidney / Pancreas
 - Overview
 - Incompatible Kidney Transplant Program
 - Overview
 - Paired Kidney Exchange

FONT SIZE PRINT THIS PAGE Share this page: [social media icons]

Home > Comprehensive Transplant Center > Programs > Kidney > Incompatible Kidney Transplant Program

Paired Kidney Exchange

- What is a paired kidney exchange?
- How are donor/recipient pairs matched?
- How long will it take to find a matching donor/recipient pair?
- How does the operation work?
- Can pairs meet?

What is a Paired Kidney Exchange?

Since 2001, Johns Hopkins Comprehensive Transplant Center has participated in paired kidney exchanges. A paired kidney exchange, also known as a "kidney swap" occurs when a [living kidney donor](#) is incompatible with the recipient, and so exchanges kidneys with another donor/recipient pair.

Two live donor transplants would occur. Suppose there were two donor/recipient pairs, Donor and Recipient 1 and Donor and Recipient 2.

- Appointment Request
- Refer a Patient
- Meet Our Team
- Contact Us
- Make a Gift

Follow Us on Facebook

Congratulations to the Kidney



Kidney Exchange

Alvin E. Roth, Tayfun Sonmez, M. Utku Unver

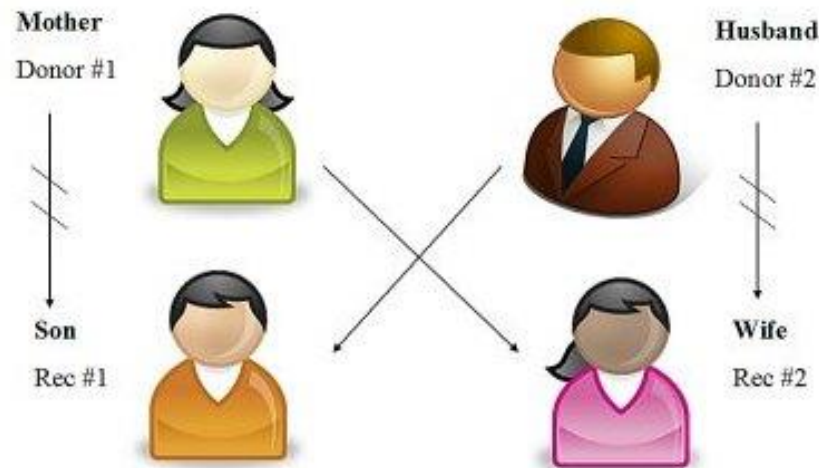
NBER Working Paper No. 10002
Issued in September 2003
NBER Program(s): HE PE

Most transplanted kidneys are from cadavers, but there are also substantial numbers of transplants from live donors. Recently, there have started to be kidney exchanges involving two donor-patient pairs such that each donor cannot give a kidney to the intended recipient because of immunological incompatibility, but each patient can receive a kidney from the other donor. Exchanges are also made in which a donor-patient pair makes a donation to someone on the queue for a cadaver kidney, in return for the patient in the pair receiving the highest priority for a compatible cadaver kidney when one becomes available. We explore how such exchanges can be arranged efficiently and incentive compatibly. The problem resembles some of the housing problems studied in the mechanism design literature for indivisible goods, with the novel feature that while live donor kidneys can be assigned simultaneously, the cadaver kidneys must be transplanted immediately upon becoming available. In addition to studying the theoretical properties of the design we propose for a kidney exchange, we present simulation results suggesting that the welfare gains would be substantial, both in increased number of feasible live donation transplants, and in improved match quality of transplanted kidneys.

Motivations: Kidney Exchange

- **Pairwise Kidney exchange**

- **Existential prob. p_e** : estimation of the success prob. based on blood type etc.
- **Probing edge $e=(u,v)$** : the crossmatch test (which is more expensive and time-consuming).



Our Results

- Previous results for unweighted version [Chen et al. '09]:
 - Edge-probing: Greedy is a 4-approx.
 - Matching-probing: $O(\log n)$ -approx.
- A simple **8-approx.** for weighted stochastic matching.
 - For edge-probing model.
 - Can be generalized to set packing.
- An improved **3-approx.** for bipartite graphs and **4-approx.** for general graphs based on **dependent rounding** [Gandhi et al. '06].
 - For both edge-probing and matching-probing models.
 - This implies the gap between the best matching-probing strategy and the best edge-probing strategy is a small const.

Our Results

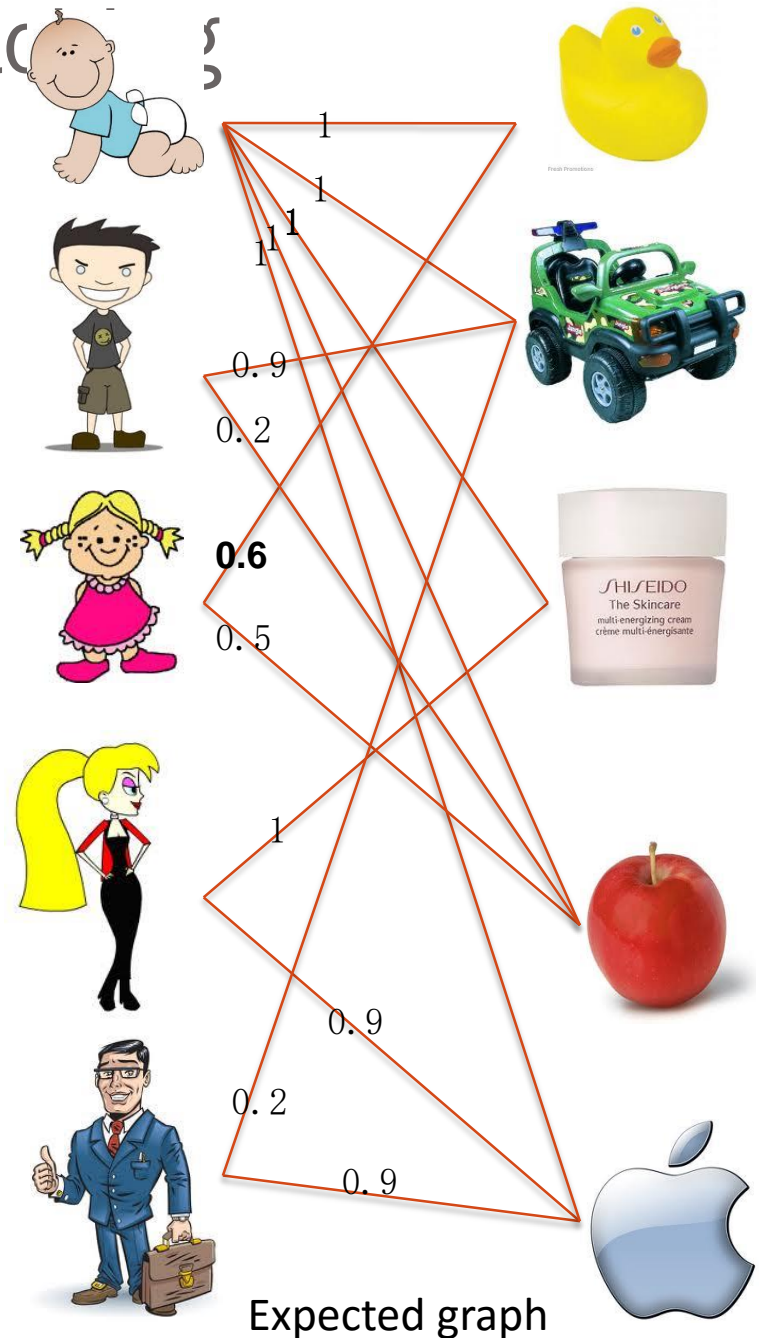
- Previous results for unweighted version [Chen et al. '09]:
 - Edge-probing: Greedy is a 4-approx.
 - Matching-probing: $O(\log n)$ -approx.
- A simple **8-approx.** for weighted stochastic matching.
 - For edge-probing model.
 - Can be generalized to set/hypergraph packing.
- An improved **3-approx.** for bipartite graphs and **4-approx.** for general graphs based on **dependent rounding** [Gandhi et al. '06].
 - For both edge-probing and matching-probing models.
 - This implies the gap between the best matching-probing strategy and the best edge-probing strategy is a small const.

Our Results

- Previous results for unweighted version [Chen et al. '09]:
 - Edge-probing: Greedy is a 4-approx.
 - Matching-probing: $O(\log n)$ -approx.
- ✓ A simple **8-approx.** for weighted stochastic matching.
 - For edge-probing model.
 - Can be generalized to set/hypergraph packing.
- An improved **3-approx.** for bipartite graphs and **4-approx.** for
- ✓ general graphs based on **dependent rounding** [Gandhi et al. '06].
 - For both edge-probing and matching-probing models.
 - This implies the gap between the best matching-probing strategy and the best edge-probing strategy is a small const.

Stochastic online match

- A set of items and a set of buyer types. A buyer of type b likes item a with probability p_{ab} .
 - $G(\text{buyer types, items})$: Expected graph
- The buyers arrive online.
 - Her type is an **i.i.d. r.v.** .
- The algorithm shows the buyer (of type b) at most t items one by one.
- The buyer buys the first item she likes or leaves without buying.
- **Goal:** Maximizing the **expected number of satisfied users.**



Stochastic online matching

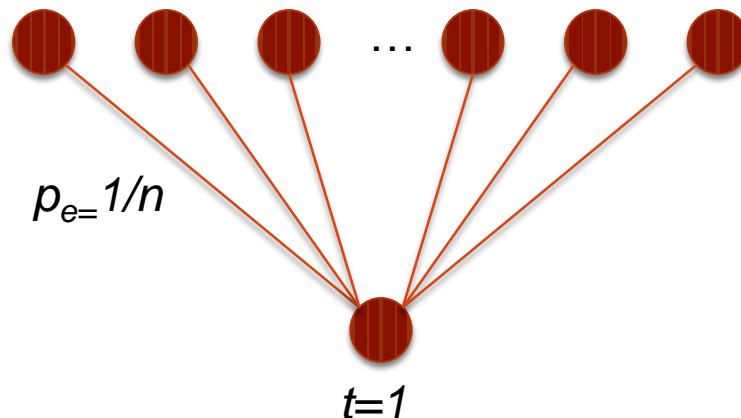
- This models the **online AdWords allocation** problem.



- This generalizes the stochastic online matching problem of [Feldman et al. '09, Bahmani et al. '10, Saberi et al '10] where $p_e = \{0, 1\}$.
- We have a **4.008**-approximation.

Approximation Ratio

- We compare our solution against the optimal (adaptive) strategy (not the offline optimal solution).
- An example:



$$E[\text{offline optimal}] = 1 - (1 - 1/n)^n \approx 1 - 1/e$$

$$E[\text{any algorithm}] = 1/n$$

A LP Upper Bound

- Variable y_e : Prob. that any algorithm probes e .

$$\text{maximize } \sum_{e \in E} w_e \cdot x_e$$

$$\text{subject to } \sum_{e \in \partial(v)} x_e \leq 1 \quad \forall v \in V \quad \text{At most 1 edge in } \partial(v) \text{ is matched}$$

$$\sum_{e \in \partial(v)} y_e \leq t_v \quad \forall v \in V \quad \text{At most } t_v \text{ edges in } \partial(v) \text{ are probed}$$

$$x_e = p_e \cdot y_e \quad \forall e \in E \quad x_e: \text{ Prob. } e \text{ is matched}$$

$$0 \leq y_e \leq 1 \quad \forall e \in E$$

A Simple 8-Approximation

An edge (u,v) is *safe* if $t_u > 0, t_v > 0$ and neither u nor v is matched

Algorithm:

- Pick a permutation π on edges uniformly at random
- For each edge e in the ordering π , do:
 - If e is not safe then do not probe it.
 - If e is safe then probe it w.p. y_e/α .

An Improved Approx. – Bipartite Graphs

Algorithm:

- $y \leftarrow$ Optimal solution of the LP.
- $y' \leftarrow$ Round y to an integral solution using *dependent rounding* [Gandhi et al. JACM06] and Let $E' = \{e \mid y'_e = 1\}$.
 - (Marginal distribution) $\Pr(y'_e = 1) = y_e$;
 - (Degree preservation) $\text{Deg}_{E'}(v) \leq t_v$; (Recall $\sum_{e \text{ in } \partial(v)} y_e \leq t_v$)
 - (Negative Correlation). $\Pr[\bigwedge_{e \in S} (Z_e = 1)] \leq \prod_{e \in S} \Pr[Z_e = 1]$. for any $S \subseteq \partial(u)$.
- Probe the edges in E' in random order.
- THM: it is a 3-approximation for bipartite graphs

- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Stochastic Probing

- A general formulation [Gupta and Nagarajan, IPCO13]
- Input:
 - Element e has weight w_e , prob of being active p_e
 - Outer packing constraints (what you can probe)
 - Downward closed (e.g., deg constraints)
 - Inner packing constraints (what your solution should be)
 - Downward closed (e.g., matchings)
 - We can adaptively probe the elements. If a probed element is active, we have to choose it irrevocably.
- Goal: Design an adaptive policy which maximizes the total weight of active probed elements

Contention Resolution Scheme

A very general and powerful rounding scheme [Chekuri et al. STOC11, SICOMP14]:

- Given a fractional point x in a polytope (the LP relaxation)
- We can do independent rounding ($X_i \leftarrow 1$ with prob x_i)
 - **But this can't guarantee feasibility**
- (b,c) -CR scheme rounds x to an feasible integer solution s.t.

$$\Pr[X_i \leftarrow 1] \geq bcx_i$$

Definition (CR Scheme) A (b,c) -balanced CR scheme π for a downwards-closed set system \mathcal{I} is a scheme such that for any $x \in P_{\mathcal{I}}$, the scheme returns a set $\pi(I) \subseteq I = R(bx)$ with the following property:

1. $\pi(I) \in \mathcal{I}$;
2. $\Pr[i \in \pi(I) \mid i \in I] \geq c$ for every element i .

Many combinatorial constraints admit good CR schemes, such as matroids, intersection of matroids (matching), knapsack etc.

Algorithm

- LP upper bound:

$$\begin{aligned} & \text{Maximize} && \sum_{e \in V} w_e x_e \\ & \text{s.t.} && x_e = p_e y_e, \quad \forall e \in V \\ & && x \in \mathcal{P}(\mathcal{I}_{\text{in}}) \\ & && y \in \mathcal{P}(\mathcal{I}_{\text{out}}) \end{aligned}$$

-
- 1 Solve the LP relaxation and obtain the optimal LP solution (x_e, y_e) ;
 - 2 Pick $I \subset 2^V$ by choosing each $e \in V$ independently with probability by_e ;
 - 3 Let $P = \pi_{\text{out}}(I)$;
 - 4 Order elements in P according to the permutation given by the ordered CR scheme π_{in} ;
 - 5 **for** $i = 1 \rightarrow |P|$ **do**
 - 6 **if** $S \cup \{e_i\} \in \mathcal{I}_{\text{in}}$ **then**
 - 7 Probe e_i ;
 - 8 **If** e_i is active, let $S \leftarrow S \cup \{e_i\}$;
-

Algorithm

Theorem *Consider an instance of the stochastic probing problem. Suppose the following hold:*

- 1. There is a (b, c_{out}) -CR scheme π_{out} for $\mathcal{P}(\mathcal{I}_{\text{out}})$;*
- 2. There is a monotone (b, c_{in}) ordered CR scheme π_{in} for $\mathcal{P}(\mathcal{I}_{\text{in}})$;*

Then, there is a polynomial time approximation algorithm which can achieve an approximation factor of $b(c_{\text{out}} + c_{\text{in}} - 1)$.

- Online content resolution scheme [Feldman et al. SODA16]
- Connection to Prophet inequalities, Bayesian Mechanism Design

- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Bayesian Online Selection

- Motivated by **Bayesian Mechanism Design**
- Input:
 - A set of elements
 - Each element is associated with a random value X_e (with known distribution)
 - We can adaptively observe the elements one by one
 - Once we see the true value of X_e , we can decide to choose it or not (main difference from stochastic probing: first see the value)
 - A combinatorial inner packing constraint as well
- Goal: Design an adaptive policy which maximizes the expected total value of chosen elements
- We can use CR scheme to solve this problem as well [Feldman et al. SODA16]

Prophet Inequality [Krengel et al. 78]

- A special case of BOS, an important problem in **optimal stopping theory**
- Input:
 - A set of elements
 - Each element is associated with a random value X_e (with known distribution)
 - We can choose one value
- Goal: Design an adaptive policy which maximizes the expected value of the chosen element

Prophet Inequality

- Prophet inequality:

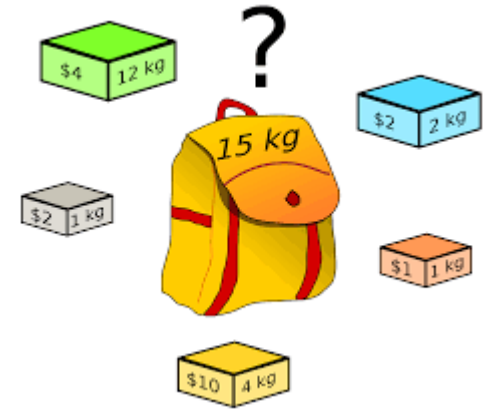
$$\mathbb{E}[X_e] \geq \frac{1}{2} \mathbb{E}[\max_{e \in V} X_e].$$

- Algorithm:
 - compute a threshold value $T = \mathbb{E}[\max_i X_i]/2$ and accept the first element whose weight exceeds this threshold
- Optimality: $1/2$ is tight

- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Stochastic Knapsack

- A knapsack of capacity C
- A set of items, each having a fixed profit
- Known: **Prior distr of size of each item.**
- Each time we choose an item and place it in the knapsack irrevocably
- The actual size of the item becomes known after the decision
- **Knapsack constraint:** **The total size of accepted items $\leq C$**
- Goal: maximize $E[\text{Profit}]$



Motivation

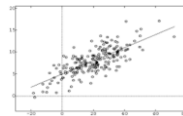
- Stochastic Scheduling
 - Jobs, each having an uncertain length, and a fixed profit
 - You have C hours
 - How to (adaptively) schedule them (maximize $E[\text{profit}]$)

Jobs:

```
# calc.py - a Python calculator
from tkinter import *

# the main class
class Calc():
    def __init__(self):
        self.total = 0
        self.current = ""
        self.new_num = True
        self.op_pending = False
        self.op = ""
        self.eq_flag = False
```

2	6	8	
5	8	9	7
3	7	4	5
6			4
	8		1
		2	3
9	8		3
	3	6	9



$C=5$ hours

Profits:

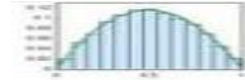
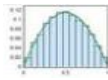
20\$

5\$

10\$

50\$

Running time:



Stochastic Knapsack

Previous work

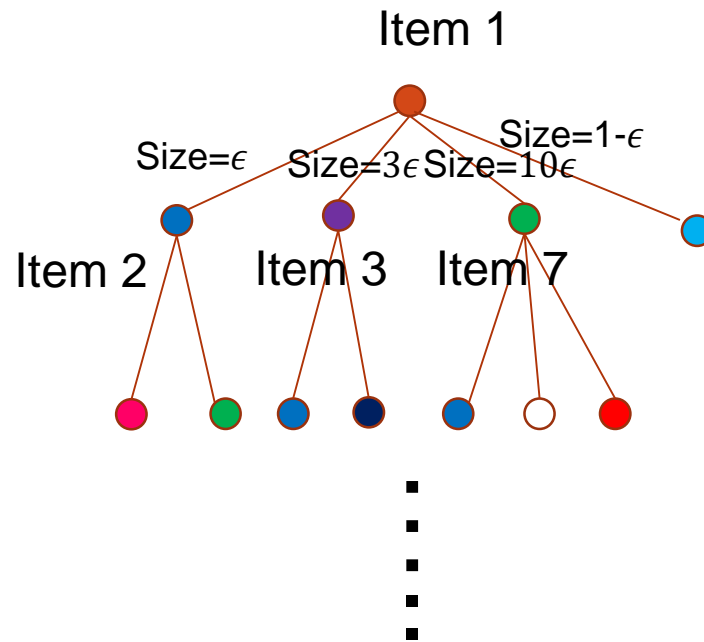
- 5-approx [Dean, Goemans, Vondrak. FOCS'04]
- 3-approx [Dean, Goemans, Vondrak. MOR'08]
- $(1+\epsilon, 1+\epsilon)$ -approx [Bhalgat, Goel, Khanna. SODA'11]
- 2-approx [Bhalgat 12]
- 8-approx (size&profit correlation, cancellation)
[Gupta, Krishnaswamy, Molinaro, Ravi. FOCS'11]

Our result:

- $(1+\epsilon, 1+\epsilon)$ -approx (size&profit correlation, cancellation)
- 2-approx (size&profit correlation, cancellation)

Stochastic Knapsack

- Decision Tree



Exponential size!!!! (depth= n)

How to represent such a tree? Compact solution?

Stochastic Knapsack

- By discretization, we make some simplifying assumptions:
 - Support of the size distribution: $(0, \epsilon, 2\epsilon, 3\epsilon, \dots, 1)$.

Still way too many possibilities, how to narrow the search space?

Block Adaptive Policies

- Block Adaptive Policies: Process items block by block



LEMMA: [Bhalgat, Goel, Khanna. SODA'11] There is a block adaptive policy that is nearly optimal (under capacity $(1 + \epsilon)C$)

Block Adaptive Policies

- Block Adaptive Policies: Process items block by block



Still exponential many possibilities, even in a single block

LEMMA: [Bhalgat, Goel, Khanna. SODA'11] There is a block adaptive policy that is nearly optimal (under capacity $(1 + \epsilon)C$)

Poisson Approximation

- Each heavy item consists of a singleton block
- Light items:
 - Using the **Poisson Approximation Technique**
 - Generate a signature for each block
 - If two blocks have the same signature, their size distributions are similar
 - **So, enumerate Signatures!** (instead of enumerating subsets)

Poisson Approximation

Le Cam's theorem (rephrased):

n r.v. X_i (with common support $(0, 1, 2, 3, 4, \dots)$) with signature

$$\mathbf{sg}_i = (\Pr[X_i = 1], \Pr[X_i = 2], \dots)$$

Let $\mathbf{sg} = \sum_i \mathbf{sg}_i$

Y_i are i.i.d. r.v. with distr $\mathbf{sg}/|\mathbf{sg}|_1$

Y follows the **compound Poisson distr (CPD)** corresponding to \mathbf{sg}

$$Y = \sum_{i=1}^N Y_i \text{ where } N \sim \text{Poisson}(|\mathbf{sg}|_1)$$

Then, $\Delta(\sum X_i, Y) \leq \sum p_i^2$ where $p_i = \Pr[X_i \neq 0]$

Variational distance:

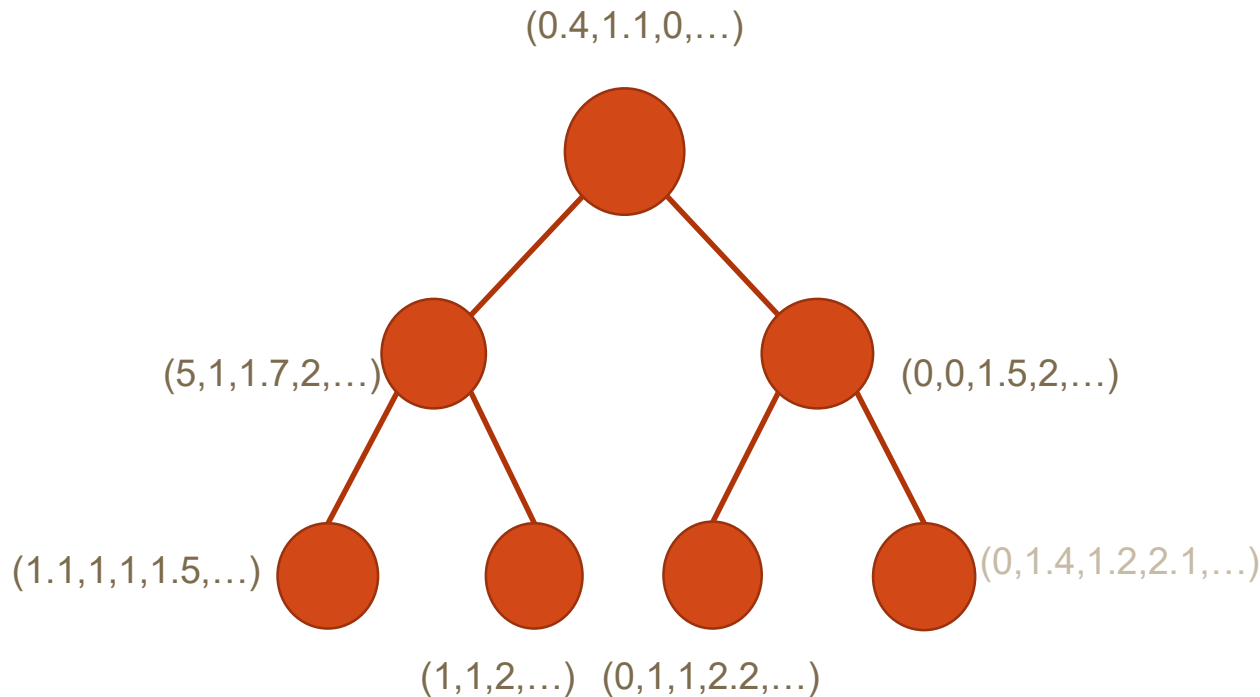
$$\Delta(X, Y) = \sum_i |\Pr[X = i] - \Pr[Y = i]|$$

Poisson Approximation

- **Le Cam's theorem:** $\Delta(\sum X_i, Y) \leq \sum p_i^2$
- Ob: If S_1 and S_2 have the same signature, then they correspond to the same CPD
- So if $\sum_{i \in S_1} p_i^2$ and $\sum_{i \in S_2} p_i^2$ are sufficiently small, the distributions of $X(S_1)$ and $X(S_2)$ are close
- Therefore, enumerating the signature of light items suffices (instead of enumerating subsets)

Algorithm

- Outline: Enumerate all block structures with a signature associated with each node



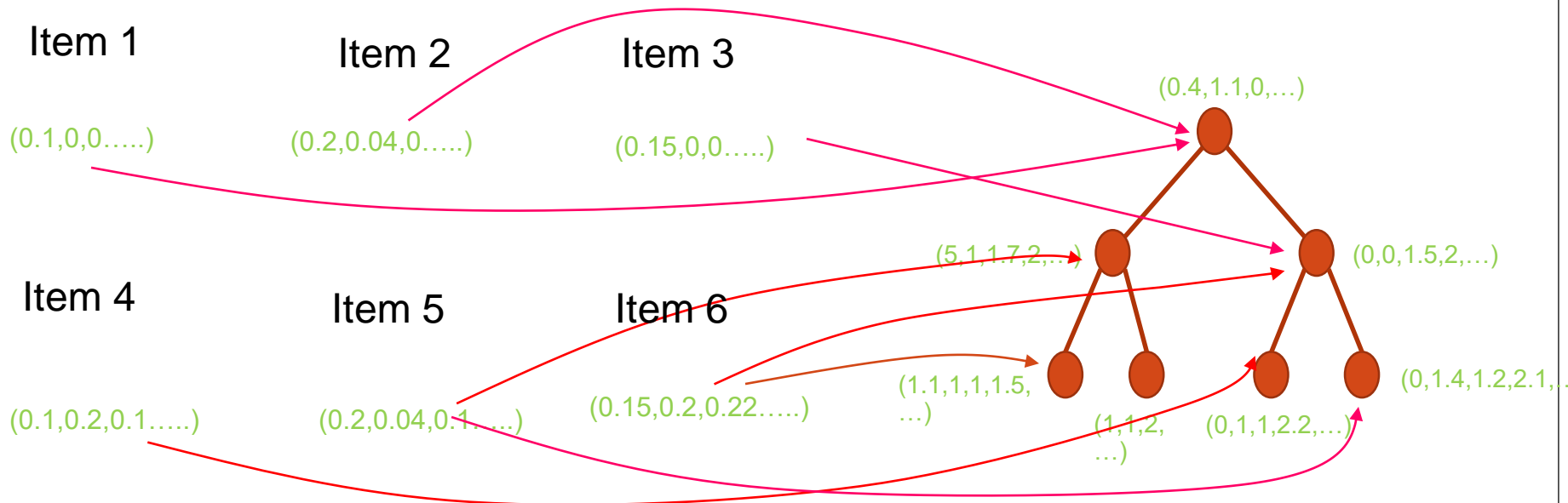
- $O(1)$ nodes
- Poly(n) possible signatures for each node
- So total #configuration = poly(n)

Algorithm

2. Find an assignment of items to blocks that matches all signatures
 - (this can be done by standard dynamic program)

Algorithm

2. Find an assignment of items to blocks that matches all signatures
 - (this can be done by standard dynamic programming)



On any root-leaf path, each item appears at most once

- Stochastic Online Optimization
- Stochastic Matching
- Stochastic Probing
- Bayesian Online Selection/Prophet inequality
- Stochastic Knapsack
- Conclusion

Conclusion

- Many interesting problems in the stochastic models
- Lots of open problems
- Deep connection to many other areas of TCS: LP primal-dual, online learning, game theory and mechanism design, counting, coresets, computational geometry
- BUT, very few researchers from China
- A survey paper:
 - Approximation Algorithms for Stochastic Combinatorial Optimization Problems. Jian Li and Yu Liu. Journal of the Operations Research Society of China. 2016

Thanks

lapordge@gmail.com

Weibo: 李建THU

Webchat: lapordge

Survey: Approximation Algorithms for Stochastic Combinatorial Optimization Problems.
Jian Li and Yu Liu. Journal of the Operations Research Society of China. 2016