# Alg. Comb: Lesson Plan

27/1/2015

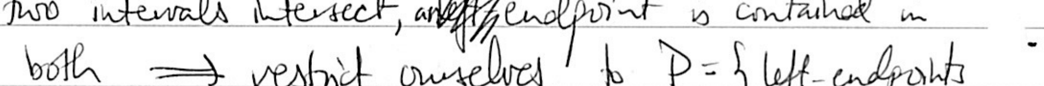
## I. Transversals of $d$-Intervals

### A. Recap ~~(pre-write)~~

1. Helly's theorem: if $C_1, \dots, C_n \subseteq \mathbb{R}^d$ are convex sets s.t. $\bigcap_{i \in I} C_i \neq \emptyset$ $\forall I \subseteq [n]$, $|I| = d+1$, then $\bigcap_{i \in [n]} C_i \neq \emptyset$.
   a. Remark: (i) e.g.: pairwise-intersecting intervals in $\mathbb{R}$.
   (ii) $(d+1)$-wise intersections necessary; e.g. triangle in $\mathbb{R}^2$.
2. $d$-intervals: Def$^n$: Union of $d$ closed intervals in $\mathbb{R}$.
3. transversal: Def$^n$: Given a family $\mathcal{F}$ of sets, $T$ is a transversal if $F \cap T \neq \emptyset$ $\forall F \in \mathcal{F}$.
4. Thm 1: (Alon) Let $\mathcal{J}$ be a finite family of $d$-intervals such that $J_1 \cap J_2 \neq \emptyset$ for every $J_1, J_2 \in \mathcal{J}$. Then $\mathcal{J}$ has a transversal of size $2d^2$; i.e., $\exists$ $2d^2$ points such that every $d$-interval of $\mathcal{J}$ contains at least one.

### B. Linear Programming formulation.

1. Endpoints:
   a. If two intervals intersect, a ~~left~~ left-endpoint is contained in both $\Rightarrow$ restrict ourselves to $P = \{$left-endpoints of intervals in $d$-intervals in $\mathcal{J}\}$.
   b. Picture:



2. Integer Programming formulation.
   a. ~~Let~~ Variables: $x_p \in \{0,1\}$ indicators of whether these points are in the transversal.
   b. Objective: $\min \sum_{p \in P} x_p$ = size of transversal
   c. Constraint: Capture a point from each $d$-interval.

$$\Rightarrow \forall J \in \mathcal{J}, \quad \sum_{p \in J} x_p \geq 1.$$

3. Linear Programming Relaxation:

(D) $\min \sum_{p \in P} x_p$

subj. to $\forall J \in \mathcal{J}, \quad \sum_{p \in J} x_p \geq 1.$

$x_p \geq 0 \quad \forall p \in P.$

a. Remark: no need for upper bound: $x_p \leq 1$ automatically.

4. $d$-approximation:

a. Let $\mathbf{x}$ be an optimal solution to LP above.

b. We will find $X \subseteq P$, $|X| \leq d \sum_p x_p$, s.t. $X$ is a transversal for $\mathcal{J}$.

c. ~~Aim: Choose $X$ s.t. the $|X|+1$ open intervals it divides $\mathbb{R}$ into have weight $< \frac{1}{d}$ under $\mathbf{x}$.~~

(i) Start counter at 0, left of min $p$.

(ii) Sweep towards the right, add $x_p$ to counter at $p$.

~~(iii) Stop when sum $\geq \frac{1}{d}$~~

(iii) If counter $\geq \frac{1}{d}$: add $p$ to $X$, reset counter to 0.

d. Observations: (i) Only add to $X$ when sum of weights is at least $\frac{1}{d}$

$\Rightarrow \frac{1}{d}|X| \leq \sum_p x_p \Rightarrow |X| \leq d \sum_p x_p.$

(ii) sum of weights in any interval avoiding $X$ is $\leq \frac{1}{d}$.

e. Transversal: If $J \in \mathcal{J}$, then $\sum_{p \in J} x_p \geq 1$

$\Rightarrow$ one of the $d$ intervals in $J$ has weight $\geq \frac{1}{d}$

$\Rightarrow$ interval meets $X \Rightarrow X \cap J \neq \emptyset$. ✓

f. Conclusion: Suffices to show value of LP is $\leq 2d$.

↳ Lemma 2: value of (D) $\leq 2d$.

C. Duality (Pf of Lemma 2)

1. ~~Pf of Lemma 2:~~ Feasibility

a. (D) is feasible: $x_p \equiv 1$.

b. Need to show it is bounded by $2d$.

c. Convenient to work w/ the dual (primal) problem instead.

2. Primal.

a. ~~Variables~~ $y_J$ for each $J \in \mathcal{J}$.

(P) $\max \sum_J y_J$, $(J \in \mathcal{J})$
subj to $\sum_{J \ni p} y_J \leq 1 \quad \forall p \in P$
$y_J \geq 0$.

b. Feasibility:

(i) (P) is feasible: $y_J \equiv 0$.

c. Strong Duality

$\Longrightarrow$ both problems bdd w/ same value

$\Longrightarrow$ suffices to show val(P) $\leq 2d$.

3. Rational solution

a. Optimal soln $y^*$ to P obtained at a b.f.s.

$\Longrightarrow$ may assume $y^*$ to be rational (determinant formula)

b. $\Rightarrow$ ~~every~~ no pt of P is contained in d-intervals of total weight $> 1$; want to show wt of <u>all</u> d-intervals $\leq 2d$.

c. Will use pairwise intersections to find a <u>popular</u> point $p \in P$ contained in intervals of large relative weight.

(i) Convenient to have "integral" set up so we can count.

d. Blow-ups.

(i) Let $D$ = common denominator of $y^*$, so $y^*_J = \frac{r_J}{D}$, $r_J \in \mathbb{Z}$.

(ii) Form a sequence (multiset) $\tilde{\mathcal{J}}$ of d-intervals, with each $J \in \mathcal{J}$ appearing $r_J$ times.

(iii) ~~Hi~~ multiset is still pairwise intersecting

(iv) wt condition $\Longrightarrow$ no pt in $> D$ d-intervals of $\tilde{\mathcal{J}}$

(v) want to deduce $|\tilde{\mathcal{J}}| \leq 2Dd$.

e. Popularity Lemma

(i) $\therefore$ suffices to prove: $\{J_1, J_2, \dots, J_n\}$

<u>Lemma 3</u>: Let $\underline{\underline{\mathcal{J}}}$ be a multiset of ~~of~~ pairwise-intersecting $d$-intervals, and let $P$ be the left endpoints of the constituent intervals. Then there is some $p \in P$ that is contained in at least $\frac{n}{2d}$ $d$-intervals (with multiplicity) of $\mathcal{J}$.

$\square$

## D. <u>Popularity Lemma</u>

1. Counting triples:

   a. We count triples $(p, i, j)$ where $p$ is a left-endpt of $J_i$ and $p \in J_j$.

2. Lower bd.

   a. Fix $i \leq j$. $J_i \cap J_j \neq \emptyset \Rightarrow$ two of the intervals of $J_i, J_j$ intersect

   b. $\Rightarrow$ left endpoint of $J_i$ or $J_j$ contained in $J_i \cap J_j$

   c. $\Rightarrow \exists p$ s.t. either $(p, i, j)$ or $(p, j, i)$ counted.

   d. $\Rightarrow$ $\#$ $(p, i, j)$ counted $\geq \binom{n}{2} + n$

3. Average.

   a. Each ~~p,~~ $i$ ~~can be counted~~ has $\leq d$ left-intervals $\Rightarrow \leq nd$ $(p, i)$ s.t. $(p, i, j)$ can be counted

   b. Averaging $\Rightarrow \exists (p, i)$ counted in $\geq \dfrac{\binom{n}{2} + n}{nd} > \dfrac{n}{2d}$. triples

$\cancel{n(\log n) \geq n^2 d}$.

   c. $\Rightarrow p$ contained in $> \frac{n}{2d}$ $d$-intervals $J_j \in \mathcal{J}$. $\boxed{\square}$

E. <u>Remarks</u>:

1. Alg. topology $\Rightarrow \exists$ transversal of size $\leq d^2$
2. Lower bounds $\Rightarrow$ cannot do better than $\frac{cd^2}{\log d}$.

## II. Randomised Algorithms for Perfect Matchings

### A. Matchings Review (spoken)

1. Bipartite graphs
   a. Hall's condition → co-NP certificate
   b. Augmenting Path Algorithm
2. General graphs
   a. Tutte's Theorem → co-NP certificate
   b. Edmond's Bloom Algorithm (complicated, not part of course)

### B. Randomised Algorithms

1. Trade-off
   a. ~~Use of~~ Clever use of randomness allows one to trade certainty for efficiency
   b. Sometimes randomised algs. can even be efficiently derandomised.
2. One-sided errors
   a. ~~Basic scen~~ Setting: Yes/No question
   b. False YES would be catastrophic, but time of the essence
      eg: Is this donor a match?
   c. One-sided tester:
      (i) If NO → No always
      (ii) If YES → No 50%
      Yes 50%.
   d. From our point-of-view: Yes ⇒ YES.
   e. Boost confidence by repeating independent trials
      (i) NO → No always
      (ii) YES → No $2^{-k}$
      Yes $1-2^{-k}$
      (iii) Error prob can be insignificantly small.