OVP

6 November, 2020 09:06

Orthogonal Vector Problem (OVP)

Given $A, B \subseteq \{0,1\}^d$, $|A| = |B| = n$,

$\exists\, a \in A,\ b \in B$ s.t. $\langle a, b \rangle = 0$?

- Alg: $O(n^2 d)$ ... trivo
  $O(n (\lg n) \cdot 2^{2d})$ ... sort and remove duplicates
  $\Rightarrow$ small d is easy $\Rightarrow$ interesting case $d = \Omega(\lg n)$

- **Reduction k-SAT $\leq$ OVP** (see also the first lecture)

k-CNF $\psi = C_1 \,\&\, C_2 \,\&\, C_3 \ldots \&\, C_m$

$d = m + 2$

$\psi(\underbrace{x_1, \ldots}_{x_{1 \ldots \frac{n}{2}}} \underbrace{\ldots x_n}_{x_{\frac{n}{2}+1 \ldots n}})$

$\overline{x_1}$ $\overline{x_2}$

$\forall a \in \{0,1\}^{\frac{n}{2}}$ $\forall b \in \{0,1\}^{n/2}$

A B

$v_a = 10 \ldots {}^{0}\!/\!_{1} \ldots$ — i-bit $= 1$ iff C_i is not satisfied by $\overline{x_1} = a$

$u_b = 01 \ldots {}^{0}\!/\!_{1} \ldots$ — j-bit $= 1$ iff C_j is not satisfied by $\overline{x_2} = b$

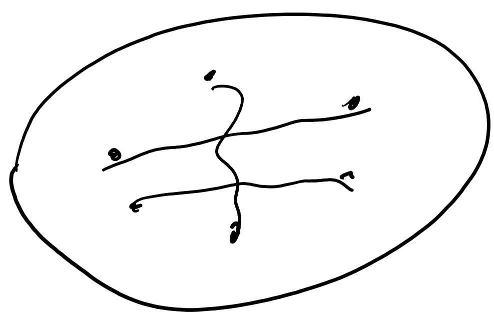
$\langle v_a, v_b \rangle = 0$ iff $\psi(a, b)$ is true by $\bar{x}_1 = a$

$|A| = 2^{n/2} = N \qquad d = m+2$ by Sparsification w.l.o.g. $d = O(\log n)$

$O(N^{2-\varepsilon})$ alg for OVP $\Rightarrow$ $2^{n(1-\frac{\varepsilon}{2})}$ alg for k-SAT

• Graph Diameter

G



$D(G) = \max_{u,v \in V(G)} d(u,v)$

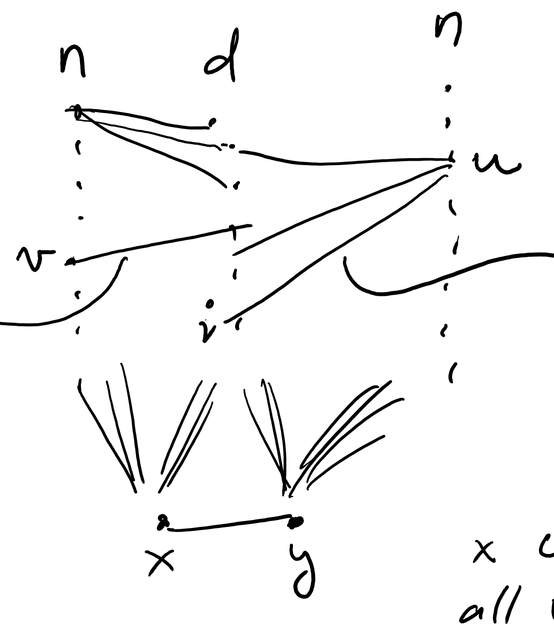
... diameter of G

Alg: Diameter can be computed in $O(n \cdot m)$ time
(compute BFS from each vertex :))

Faster alg?

• OVP $\leq$ DIAM

A, B $\rightarrow$ G



edge from v to j iff $v_j = 1$

edge from u to i iff $u_i = 1$

x connects to all v's & i's

$|V(G)| = 2n + d + 2 = O(n+d)$ $\quad$ j connects to all i's & u's.

$|E(G)| \leq 2n \cdot d + 2n + 2d + 1$

$= O(nd)$

$\langle v, u \rangle \neq 0 \iff d(u,v) = 2$

$\langle v, u \rangle = 0 \Rightarrow d(u,v) = 3$

$$D(G) = \begin{cases} 3 & \exists\, v \in A, b \in B \ \ \langle u, v \rangle = 0 \\ 2 & \text{otherwise} \end{cases}$$

$\Rightarrow$

- alg for DIAM in time $O(n^{2-\varepsilon}) \Rightarrow$ alg for OVP in time $O(n^{2-\varepsilon'})$
- alg for DIAM in time $O\left(\frac{n \cdot m}{n^{\varepsilon}}\right) \Rightarrow$ alg for OVP in time $O(n^{2-\varepsilon'})$
- $f : \Sigma^{*} \to \mathbb{N}$ $\qquad$ Alg A <u>c-approximates</u> f if $\forall x \quad A(x) \leq f(x) \leq c \cdot A(x)$.
- alg. for $\left(\frac{3}{2} - \varepsilon\right)$-approximation of DIAM in time $O\left(\frac{n \cdot m}{n^{\varepsilon}}\right) \Rightarrow$ alg for OVP in time $O(n^{2-\varepsilon'})$

<u>Known</u>: alg for $\frac{3}{2}$-approx of DIAM in time $O(m^{3/2})$
(For sparse graphs faster than $O(m \cdot n)$)

Dynamic Programming Problems

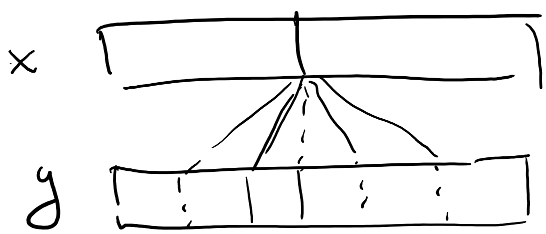
- **Longest Common Subsequence** (LCS)

x 0 1 2 0 1 1 2 3 1 0 $\quad LCS(x,y) = 7$

y 1 0 2 1 2 3 1 0 0 3

→ delete letters from x & y to get the same string → common substring, maximize its length

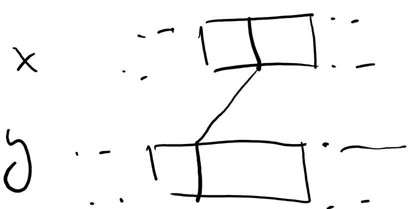
- can be solved by dynamic programming



n^2 subintervals

n^2 subintervals

$O(n^4)$ combinations



$O(n^6)$ splits

go from the smallest subintervals to largest

→ $O(n^6)$ alg.

Can be done in time $O(n^2)$ ... look on matches of subintervals which are prefixes of x & y.

- Fastest known alg. $O\left(\frac{n^2}{\log^2 n}\right)$ [Masek & Paterson]

uses word-level parallelism
(RAM can process $O(\log n)$ bits
by each operation)

• <u>Edit distance</u>

$$\begin{array}{l} X = \overset{\times\ \text{del}}{\text{TRUMP}} \\ \quad\ \downarrow \text{sub.} \\ y = \text{LUMP} \end{array}$$

edit op's
- insert symbol
- delete symbol
- substitute symbol

$$ED(x,y) = 2$$

edit distance (x,y) ... the minimum # of op's to turn x into y.

variants use different op's of various costs,

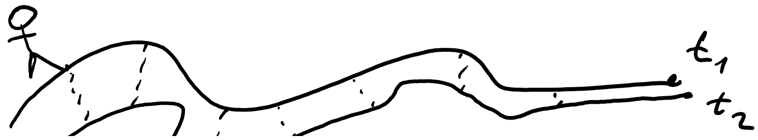
e.g. cost of substitution $= 2$
$\rightarrow ED_{ins,del}$

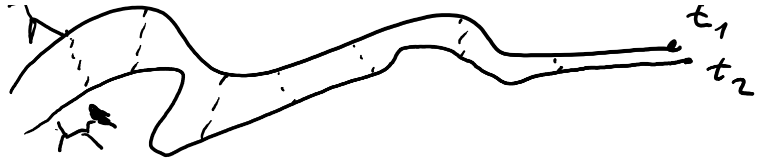
$$LCS(x,y) = n - \frac{1}{2} ED_{ins,del}(x,y)$$

• alg ... dynamic pgming $O(n^2)$ time
best $O\left(\frac{n^2}{\log^2 n}\right)$

• <u>Fréchet distance</u> (Dog Walk Distance)

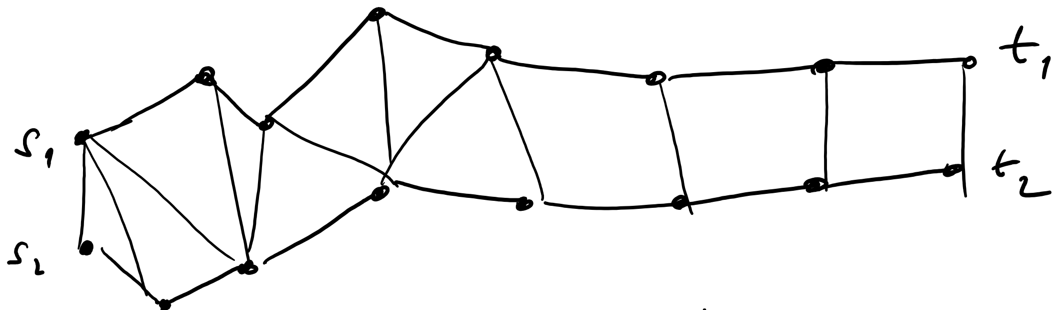
- measures similarity of curves





dog on a leash, both person & dog move only forward but can vary speed. The necessary length of leash?

- Discrete variant



given two sequences of n pt's in plane the dog and the walker can either stay at the given pt ~~st~~ or make a step to the next one. Length of the leash?

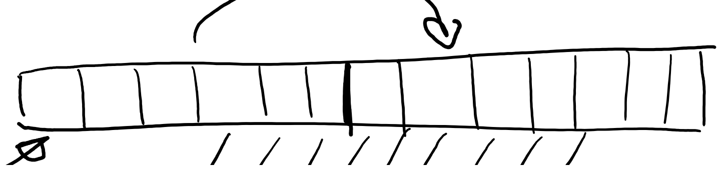
$O(n^2)$ dynamic pgming alg.

- other measures . . .

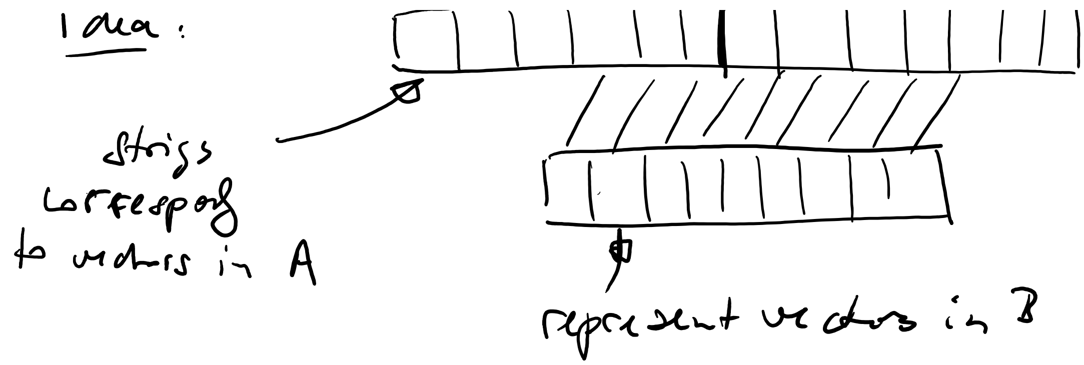
- $OVP \leq LCS$

$A, B \subseteq \{0,1\}^d$

idea:



Idea:



orthogonal vectors will give larger LCS than non orthogonal

1) encoding a bit of vectors from A & B

$$0^A = 001 \qquad\qquad 1^A = 111$$

$$2 \qquad 2 \qquad 0 \quad \leftarrow \text{LCS}$$

$$0^B = 011 \qquad\qquad 1^B = 000$$

$$V^A(a) = a_1^A\, 22\cdots 2\, a_2^A\, 22\cdots 2 \;\cdots\; 2\, a_d^A$$

$$V^B(b) = b_1^A\, \underbrace{22\cdots 2}_{4d}\, b_1^B\, 22\cdots 2\, b_3^B \cdots 2\, b_d^B$$

$$\begin{aligned} LCS(V^A(a), V^B(b)) &= 4d\cdot(d-1) + 2d && a \perp b \\ &\leq 4d\cdot(d-1) + 2d - 2 && a \not\perp b \end{aligned}$$

- you will lose by matching a_i to b_j for $i \neq j$ b/c 2's.

$$— \quad a_i\, 22222 \quad \cdots$$

$$--- \quad 2222 \quad \cdots$$

e will be lost $\geq 4d$

→ modify a's & b's by adding extra coordinate $a_{d+1} = 0$ & $b_{d+1} = 1$

$V'(a)$ $\quad V^A(a)\underbrace{44\cdots 4}^{10d^2} V(s)$ $\quad s = (\underbrace{0,0,0,\cdots,0}_{d},1)$

$V'(b)$ $\quad \underbrace{44\cdots 4}_{10d^2} V^B(b) \underbrace{44\cdots 4}_{10d^2}$

$$LCS(V'(a), V'(b)) = 10d^2 + 4d^2 + 2(d+1) \qquad a \perp b$$
$$= 10d^2 + 4d^2 + 2d \qquad a \not\perp b$$

• want to match all 4's from the top + something ⇒

$V(a)\,44\cdots 4\,V(s)$ / $444\,V(b)\,444$ $\quad a \perp b$

or $V(a)\,444\cdots 4\,V(s)$ / $4444\,V(b)\,44444$ $\quad a \not\perp b$

→

$$x = V'(a_1)(35)^{100d^2} V'(a_2)(35)^{100d^2} V'(a_3) \cdots V'(a_n)(35)^{100d^2} V'(a_1) \cdots V'(a_n)$$

$$\underbrace{33\cdots 3}_{2n\cdot 100d^2} V'(b_1)(35)^{100d^2} V'(b_2)(35)^{100d^2} \cdots V'(b_n) \underbrace{33\cdots 3}_{2n\cdot 100d^2}$$

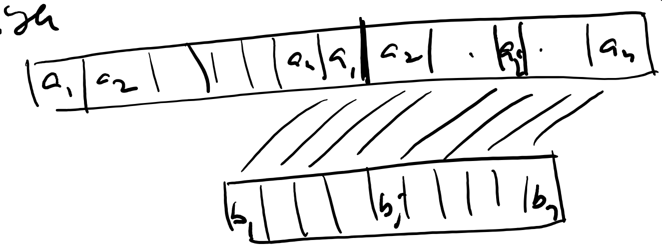
want 1) match all 3's from the top $(2n-1)100d^2$

2) match all 5's from the bottom $(n-1)100d$

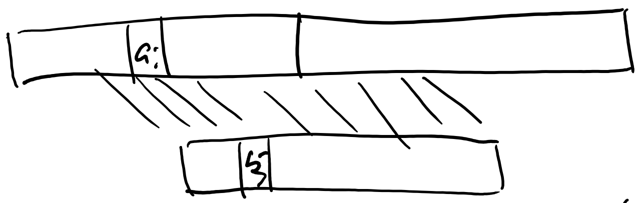
1) $\Rightarrow$ cannot match $V'(b_i)$ to two distinct $V'(a_j)$ $V'(a_{j'})$

2) $\Rightarrow$ cannot match $V'(a_j)$ to two distinct $V'(b_i)$.

best possible $\Rightarrow$ $a_i \perp b_j$



$i < j$



$j \leq i$

each $V'(b_j)$ is match to ~~some~~ $V'(a_i)$ (if possible to some orthogonal)

they have to match as block

$$LCS(x,y) \geq \overbrace{(2n-1)100d^2}^{3's} + \overbrace{(n-1)100d^2}^{5's} + \overbrace{n(14d^2+2d)}^{LCS(V'(a),V'(b))} + 2$$

if $\exists a_i \perp b_j$

$$= (2n-1)100d^2 + (n-1)100d^2 + n(14d^2+2d)$$

if $\forall a_i \not\perp b_j$