LECTURE # 9 **Nov 13 2014**

AGONY.

Hierarchy $h(G)$ of a directed graph.

DATA: $G = (V, E)$ = Directed Graph. $n = |V|$, $m = |E|$.

r = Scoring function on vertices V

$$r : V \longrightarrow \mathbb{N}$$

Hierarchy relative to r

$$h(G,r) = 1 - \frac{1}{m} \sum_{(i,j) \in E} \max\left(r(i) - r(j) + 1, 0 \right)$$

Hierarchy $h(G)$ in a directed graph.

$$h(G) = \max_{r \in \text{ranking}} h(G, r)$$

$$= 1 - \frac{1}{m} \min_{r \in \text{ranking}} \sum_{(i,j) \in E} \left(r(i) - r(j) + 1 \right)^+$$

AGONY

$$A(G) = \min_{r \in \text{ranking}} \sum_{(i,j) \in E} \left(r(i) - r(j) + 1 \right)^+$$

Agony with respect to a ranking

$$A(G,r) = \sum_{(i,j) \in E} \left(r(i) - r(j) + 1 \right)^+$$

$$A(G) = \min_{r} A(G,r).$$

Integer Programming:

$$\min \sum_{(i,j)\in E} x(i,j)$$

subject to

$$x(i,j) \geq r(i) - r(j) + 1 \qquad \forall (i,j) \in E$$

$$x(i,j) \geq 0 \qquad \forall (i,j) \in E$$

$$r(i) \geq 0 \qquad \forall i \in V$$

$$x(i,j),\ r(i) \in \mathbb{Z}.$$

Upper Bound. $= m = |E|$

$$r(i) = 0 \quad \forall i \in V$$
$$x(i,j) = 1 \quad \forall (i,j) \in E.$$

Lower Bound: LP-relaxation. (Dual Form)

$$\max \sum_{(i,j)\in E} z(i,j)$$

$$z(i,j) \leq 1 \qquad \forall (i,j) \in E$$

$$\sum_{j\in V} z(k,j) \leq \sum_{i\in V} z(i,k) \qquad \forall k \in V$$

$\longrightarrow$ $\boxed{\text{Show This!}}$

$$z(i,j) \geq 0.$$

Recall

$$\min c^T x \quad \text{s.t } Ax \leq b,\ x \geq 0$$

$$\longleftrightarrow \max b^T y \quad \text{s.t. } A^T y \geq c,\ y \geq 0.$$

OBSERVATIONS.

$$\sum_{k \in V} \sum_{j \in V} z(k,j) \leq \sum_{k \in V} \sum_{i \in V} z(i,k) \quad \forall k \in V.$$

$$\Rightarrow \quad \sum_{j \in V} z(k,j) = \sum_{i \in V} z(i,k) \quad \forall k \in V.$$

Force $z(i,j) \in \{0,1\} \Rightarrow$ Induces a subgraph $E_z \subseteq E$.

$$G_z = (V, E_z) = \text{Eulerian.}$$

Step 1. Graph $G = (V,E)$ is decomposed into H = Eulerian Subgraph and a DAG s.t. $G = (V, H \cup DAG)$

Edges in the Eulerian graph are labelled +1
Edges in DAG are labeled −1.

$$G = (V, E, \omega) \qquad \omega : E \rightarrow \{+1, -1\}$$

Step 2. For all $v \in V$ $\ell(v) := 0$;

Step 3. <u>While</u> $\exists_{(u,v) \in E}$ s.t. $\ell(v) < \ell(u) - \omega(u,v)$ <u>do</u>

$$\ell(v) := \ell(u) - \omega(u,v)$$

Step 4.
$$x(u,v) := \begin{cases} 0 & \text{if } (u,v) \in DAG \\ \ell(u) - \ell(v) + 1 & \text{if } (u,v) \in H \end{cases}$$

Eulerian Subgraph.

Minimizing Agony in a social network (64)
leads to a ranking $\rightarrow$
Pecking order / Social Hierarchy
Org Chart
Stratification
Influencing Graph.
~~So~~.

Influence.

$$A \xrightarrow{inf} B. \quad \text{with respect to item } i.$$

$x(A, i, t_A) \Rightarrow$ A buys item i at time t_A

$x(B, i, t_B)$

Temporal Priority $\quad t_A < t_B$

Probability Raising $\quad Pr[x(B,i,\cdot) \mid x(A,i,\cdot)]$
$\gg Pr[x(B,i,\cdot) \mid \neg x(A,i,\cdot)]$.

Probabilistic Causality.
$\Rightarrow$ DAG containing individuals
$\Rightarrow$ One DAG G_i for each item
$\Rightarrow$ Social Ranking
$r(A) < r(B)$

$A \rightsquigarrow B$
Causal Path.

Q. Separating Homophily from Causal Influence.

Comparing two items i and j.

$\Rightarrow$ Causal Graphs G_i and G_j

Note $G_i = (V_i, E_i)$ & $G_j = (V_j, E_j)$

$V_i = V_j$ but $E_i \neq E_j$ (most likely)

$\Rightarrow$ Define

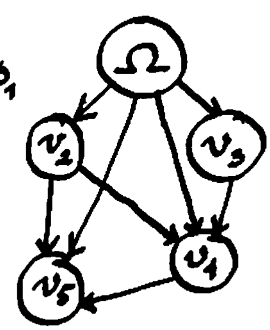
$$dist(i,j) = dist(G_i, G_j) = A(G_i \cup G_j)$$

= Agony with respect to a ranking in $G_i \cup G_j$ (which may have cycles even though G_i & G_j are DAGs).

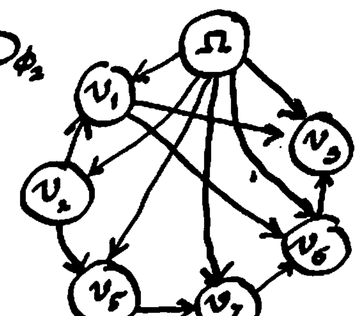
Note $\begin{cases} A(G_i \cup G_i) = 0 \\ A(G_i \cup G_j) = A(G_j \cup G_i) \end{cases}$

v	ϕ	t
Ω	ϕ_1	0
v_2	ϕ_1	2
v_3	ϕ_1	4
v_4	ϕ_1	5
v_5	ϕ_1	7
Ω	ϕ_2	0
v_2	ϕ_2	1
v_1	ϕ_2	3
v_5	ϕ_2	6
v_7	ϕ_2	7
v_6	ϕ_2	8
v_3	ϕ_2	9

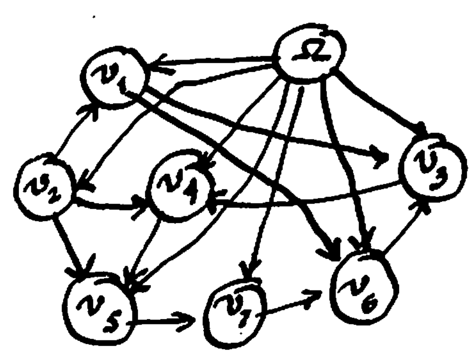
D_{ϕ_1}



D_{ϕ_2}



$D_{\phi_1} \cup D_{\phi_2}$



$v_3 \to v_4 \to v_5 \to v_7 \to v_6 \to v_3$

$A(D_{\phi_1} \cup D_{\phi_2}) = 5$

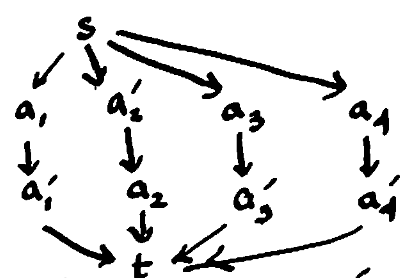
$dist(\phi_1, \phi_2) = 5$

$\vec{v} = \{0,1\}^n$ vector

Create a graph $G_{\vec{v}} = \Big(V = \{s, t, a_1, a_1', \ldots, a_n, a_n'\},$

$E = \{(s, a_i) \mid v_i = 1\}$
$\cup \{(a_i, a_i') \mid v_i = 1\}$
$\cup \{(a_i', t) \mid v_i = 1\}$
$\cup \{(s, a_i') \mid v_i = 0\}$
$\cup \{(a_i', a_i) \mid v_i = 0\}$
$\cup \{(a_i, t) \mid v_i = 0\} \Big).$

10 11 $\longrightarrow$



$$\text{Hamming Dist}(\vec{v}, \vec{w}) = \text{dist}(G_{\vec{v}}, G_{\vec{w}}) = A(G_{\vec{v}} \cup G_{\vec{w}}).$$

HYPERCUBE CLUSTERING PROBLEM. (HCP)

Data: m Binary Vectors $v_1, \ldots, v_m$ of length n and an integer k.

Desiderata: k binary vectors $c_1, \ldots, c_k$ of length n (the centroids) and
a function $f: \{1, \ldots, m\} \to \{1, \ldots, k\}: v_i \mapsto c_j$

which minimizes distortion

$$E = \sum_{t=1}^{m} \Delta(v_t, f(v_t))$$

where Δ: Hamming Distance

AGONY BOUNDED CLUSTERING:

Data: A set of DAGs $\mathbb{D}$. A threshold $\eta \in \mathbb{N}$.
A cardinality $k \in \mathbb{N}$.

Desiderata: A set of groups

$$\mathcal{P} = \{S_1, \ldots, S_k\}; \quad \bigcup_{i=1}^{k} S_i = \mathbb{D}.$$

Let $G(S_i) \equiv \bigcup [\mathbb{D} \cap S_i]$

$\forall_{i \in \{1,..k\}}$ $G(S_i)$ = Weakly-connected

and $A(G(S_i)) \leq \eta$.

Goal: Minimize k = Number of groups.