19th Jan, 2025

Djokoviks embedding characterisation

Om Joglekar

In this presentation we characterise the graphs which can be isometrically embedded into the hypercube graph H_n (the n-dim 2^n vertices graph)
All graphs considered shall be simple graphs

Definition: We say that a set of vertices $S \subseteq V(G)$ is a closed set in the connected graph G if the following holds $\forall\, a,b \in S$ and $x \in V(G)$: $d(a,x) + d(x,b) = d(a,b) \Rightarrow x \in S$

One observes from the above definition that $\forall\, a,b \in S$, all vertices along all shortest paths from a to b are in S (and that this is a sufficient condition)

Proposition 1

Let $S, T \subseteq V(G)$ be closed sets in a connected graph G.
Then so is $S \cap T$

proof

Let $a,b \in S \cap T$, $x \in G$ s.t. $d(a,x) + d(x,b) = d(a,b)$
Since $a,b \in S$ & $x \in G$, $x \in S$
Similarly $x \in T \Rightarrow x \in S \cap T$ ∎

Corollary:

Given any $S \subseteq V(G)$, we can uniquely obtain $\overline{S}$, the smallest closed set in G containing S (G connected)

Proposition 2

If G has an isometric embedding (distance preserving homomorphism from G to another graph) into H_n, then
(i) G is connected & bipartite
(ii) If a,b are adjacent in G, then $W(a,b)$ is closed in G

(Here $\mathcal{U}(a,b) = \{ x \in V(G) \mid d(a,x) < d(b,x) \}$)

proof

Firstly, let us treat G as a subgraph of H_n (with the obvious fact that $G \hookrightarrow H_n$ is a distance preserving homo)

(i) Let $a, b \in V(G)$

By definition of dist preserving homomorphism,

$d_G(a,b) = d_{H_n}(a,b)$ which is non-zero & well defined

Thus a, b are path connected

Thus G is connected

Since G is a subgraph, $\chi(G) \leq \chi(H_n) = 2$

Being connected, $\chi(G) = 2$ (life is boring if $|V(G)| = 1$)

(ii) Let $p, q \in \mathcal{U}(a,b)$. Let x lie along some shortest path from p to q.

$d(p,x) + d(x,q) = d(p,q)$

$\therefore d_H(p,x) + d_H(x,q) = d_H(p,q)$ (d_H is hamming distance)

Since $d(p,a) < d(p,b)$, $d(q,a) < d(q,b)$ and hence $d_H(p,a) < d_H(p,b)$ & $d_H(q,a) < d_H(q,b)$. WLOG, let a and b differ in the i^{th} coordinate – a_i being 0 & b_i being 1

Then, p & q both have i^{th} coordinate 0.

Now consider $d(p,q) = d_H(p,q)$. The shortest path between p & q will necessarily have all i^{th} coordinates 0.

Thus if $x \in \mathcal{U}(b,a)$, $d_H(p,x) + d_H(x,q) > d_H(p,q)$ & thus $d(p,x) + d(x,q) > d(p,q)$ giving a contradiction.

∎

Definition : Define a relation on the edge set of a connected bipartite graph as $e_1 \Theta e_2$ (where e_1 joins a & b) if e_2 joins vertices u and v s.t. $u \in \mathcal{U}(a,b)$ & $v \in \mathcal{U}(b,a)$. We define one more relation $e_1 \tilde{\Theta} e_2$ if $d(a,u) + d(b,v) \neq d(a,v) + d(b,u)$ where e_1 joins a,b & e_2 joins u,v

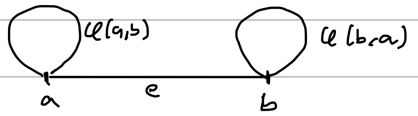
Proposition 3

$e \Theta f$ iff $e \tilde{\Theta} f$

proof

Suppose $e \Theta f$. G is bipartite and hence $\nexists\, x \in V(G)$ s.t. $d(a,x) = d(b,x)$ else we would get an odd cycle.

Consider the following diagrammatic representation for further clarity on the proof that follows.



We suppose that e joins a, b and f joins u, v. WLOG let $u \in \mathcal{C}(a,b)$. By definition of Θ, $v \in \mathcal{C}(b,a)$.

Now since ab is an edge,

$d(b,u) \leq d(b,a) + d(a,u) = 1 + d(a,u)$

$\therefore \quad 0 \leq d(b,u) - d(a,u) \leq 1$

$\therefore \quad d(b,u) = 1 + d(a,u)$

Similarly $d(a,v) = 1 + d(b,v)$

$\therefore \quad d(b,u) + d(a,v) = 2 + d(a,u) + d(b,v) \neq d(a,u) + d(b,v)$

$\therefore \quad e \,\tilde{\Theta}\, f$

Conversely, let $e \,\tilde{\Theta}\, f$ with e joining a,b & f joining u,v

WLOG, let $u \in \mathcal{C}(a,b)$. We need to show that $v \in \mathcal{C}(b,a)$

Suppose not. Then $v \in \mathcal{C}(a,b)$

$\therefore \quad d(b,u) + d(a,v) = 1 + d(a,u) + d(a,v)$

$\quad\quad d(b,v) + d(a,u) = 1 + d(a,v) + d(a,u) \quad (\because v \in \mathcal{C}(a,b))$

This contradicts $e \,\tilde{\Theta}\, f$

Thus $v \in \mathcal{C}(b,a)$ implying $e \,\Theta\, f$

∎

Proposition 4

$\Theta = \tilde{\Theta}$ is symmetric & reflexive

Proof

Follows directly from the definition of $\tilde{\Theta}$.

∎

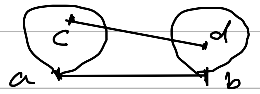
Now we move on to transitivity. This is guaranteed only when G is connected, bipartite & $\forall$ edges $\{a,b\} \in E(G)$, $\mathcal{C}(a,b)$ is closed in G ——— $(*)$

Proposition 5

Let $e \,\Theta\, f$ with e, f joining a,b & c,d respectively. Let $c \in \mathcal{C}(a,b)$. Then $\mathcal{C}(c,d) = \mathcal{C}(a,b)$ (G as described in $*$)

proof

The pictorial representation of the situation is as follows



It clearly suffices to show that $\mathcal{C}\ell(a,b) \subseteq \mathcal{C}\ell(c,d)$. Since $a \in \mathcal{C}\ell(c,d)$ & $b \in \mathcal{C}\ell(d,c)$, $\mathcal{C}\ell(c,d) \subseteq \mathcal{C}\ell(a,b)$ follows.

Let $x \in \mathcal{C}\ell(a,b) \Rightarrow d(x,a) \leq d(x,b)$ ie. $d(x,b) = 1 + d(x,a)$

Suppose $d(x,c) > d(x,d)$ ie. $1 + d(x,d) = d(x,c)$

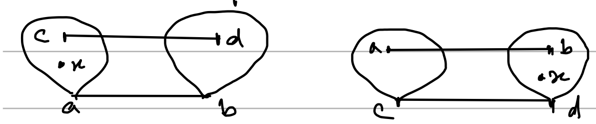
$d(x,b) - d(x,a) = d(x,c) - d(x,d)$

$d(x,b) + d(x,d) = d(x,a) + d(x,c)$

Now $c \in \mathcal{C}\ell(a,b)$, $d \in \mathcal{C}\ell(b,a) \Rightarrow a \in \mathcal{C}\ell(c,d)$, $b \in \mathcal{C}\ell(d,c)$

(Check this fact. This is also used when we said it suffices to show the inclusion in a single direction)

Relevant pictures are



Clearly, there is a contradiction since $x \in \mathcal{C}\ell(b,a)$ & $x \in \mathcal{C}\ell(a,b)$

■

Proposition 6

If G satisfies $(*)$, then Θ is an equivalence relation

proof

Symmetric & transitive has already been established by prop 4

Let an edge e joining a,b be denoted ab (same as ba)

Let $ab \,\Theta\, cd$, $cd \,\Theta\, pq$. WLOG let $c \in \mathcal{C}\ell(a,b)$ & $p \in \mathcal{C}\ell(c,d)$

using prop 5, $p \in \mathcal{C}\ell(c,d) = \mathcal{C}\ell(a,b)$ & hence $q \in \mathcal{C}\ell(d,c)$ $= \mathcal{C}\ell(b,a)$. Thus $ab \,\Theta\, pq$

■

Notation: Let $\overline{\mathfrak{A}}$ denote the collection of equivalence classes defined as $\overline{\mathfrak{A}} = E(G)/\Theta$. Let $\bar{e}$ denote the equivalence class of the edge $e \in E(G)$ (G is assumed to satisfy $(*)$)

For a fixed vertex c of G, at every vertex $x \in G$, associate a collection $F(x) \subseteq \overline{\mathfrak{A}}$ as follows: $\bar{e} \in F(x)$ iff x and c do not lie in the same set $G(a,b)$ or $G(b,a)$ where $e = ab \in E(G)$

Proposition 7

The above construction of $F(x)$ is well-defined

proof

Consider $x_0 \in G$

Let $\overline{ab} = \overline{cd}$ & $\overline{ab} \in F(x_0)$

Then x_0, c are in different parts $\mathcal{C}(a,b)$, $\mathcal{C}(b,a)$

Since $\overline{ab} = \overline{cd}$, WLOG, let $c \in \mathcal{C}(a,b)$, $d \in \mathcal{C}(b,a)$

Then $\mathcal{C}(a,b) = \mathcal{C}(c,d)$, $\mathcal{C}(b,a) = \mathcal{C}(d,c)$ (prop 5) and x_0, c are again in different parts $\mathcal{C}(c,d)$, $\mathcal{C}(d,c)$

$\blacksquare$

Proposition 8

If G satisfies $(*)$, then we can get an isometric embedding of G in H_n (this is the converse of prop 2)

proof

Our first claim is that $F(x)$ is finite subset of ख.

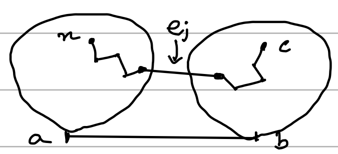
Suppose $d(x,c) = m$. Let $e_1, e_2, \ldots, e_m$ be the edges of some shortest path from x to c

Then, each $\overline{e_i} \in F(x)$ since x is closer to one end of e_i while c is closer to the other.

Thus $\{\overline{e_1}, \ldots, \overline{e_m}\} \subseteq F(x)$. The converse containment follows since if $\overline{ab} \in F(x)$, then $x \in \mathcal{C}(a,b)$, $c \in \mathcal{C}(b,a)$ (WLOG) and when the path $e_1 \cdots e_m$ is considered, at least one edge e_j runs between $\mathcal{C}(a,b)$ & $\mathcal{C}(b,a)$ giving $\overline{ab} = \overline{e_j}$.

Thus $F(x) = \{\overline{e_1}, \ldots, \overline{e_m}\}$.

The picture is drawn below

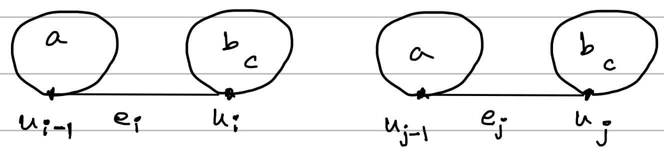


For any set S, we define the hypercube on S to be $H(S)$ whose vertices are subsets of S & $A \subseteq S$, $B \subseteq S$ are connected by an edge iff $|A \triangle B| = |(A \cup B) \setminus (A \cap B)| = 1$

The usual H_n is just $H(\{1, 2, \ldots, n\})$. Essentially, we have just reworded the definition of H_n.

Now we give our isometric embedding explicitly as $f : G \longrightarrow H(ख)$ defined as $f(x) = F(x)$

We first show that this is indeed a homomorphism.
Let $e = ab \in E(G)$
We want to show that $|F(a) \triangle F(b)| = 1$.
It is clear that $\bar{e} \in F(a) \triangle F(b)$ ($\because$ c is either closer to a or b but not both)
Suppose $\overline{e_1} \in F(a) \triangle F(b)$ with e_1 joining p, q in G
WLOG let $\overline{e_1} \in F(a) \setminus F(b)$. Then, a & c are in different parts & b and c are in the same (parts being $\mathcal{C}(p,q)$, $\mathcal{C}(q,p)$)
Thus a & b are in different & since ab is an edge, we have $ab \,\Theta\, pq \Rightarrow \overline{e_1} = \bar{e}$
Thus $F(a) \triangle F(b) = \{\bar{e}\}$
Hence f is a homomorphism
We are left to show that f is an isometry
Suppose $d(a,b) = m$ in G for some $a, b \in V(G)$
Let the path be along $e_1, e_2, \dots, e_m$
Claim: $F(a) \triangle F(b) = \{\overline{e_1}, \dots, \overline{e_m}\}$
It is clear that RHS $\subseteq$ LHS. Suppose $\overline{\text{कग}} \in$ LHS. WLOG, $\overline{\text{कग}} \in F(a) \setminus F(b)$ i.e. a & c are in different parts but b and c are in the same part (parts given by edge कग). Then similar to what we did previously, $\overline{\text{कग}} = \overline{e_j}$ for some e_j and LHS $\subseteq$ RHS.
This proves the claim
Claim: In $F(a) \triangle F(b) = \{\overline{e_1} \dots \overline{e_m}\}$, $\overline{e_i} \neq \overline{e_j}$ $\forall\, i \neq j$
Suppose $e_i \,\Theta\, e_j$ for some $i < j$.
Let the path be $a, e_1, u_1, e_2, u_2, \dots, e_{m-1}, u_{m-1}, e_m, b$
WLOG assume $\overline{e_i} = \overline{e_j}$ $(i < j) \in F(a) \setminus F(b)$ (Thus a is in one 'cell' while b & c are in another wrt both e_i & e_j)
The figure is drawn below



(just define $u_0 = a$)

(Clearly $a \in \mathcal{C}(u_{i-1}, u_i)$ & $b \in \mathcal{C}(u_j, u_{j-1})$)
$d(c, u_{i-1}) = d(c, u_i) + 1$, $d(c, u_{j-1}) = d(c, u_j) + 1$
One can now show that $d(u_j, u_{i-1}) < d(u_i, u_{j-1})$ which contradicts the shortestness of the path.