

A 2-edge-connected spanning subgraph problem: Robert Carr, Ojas Parekh, Sandia Labs

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

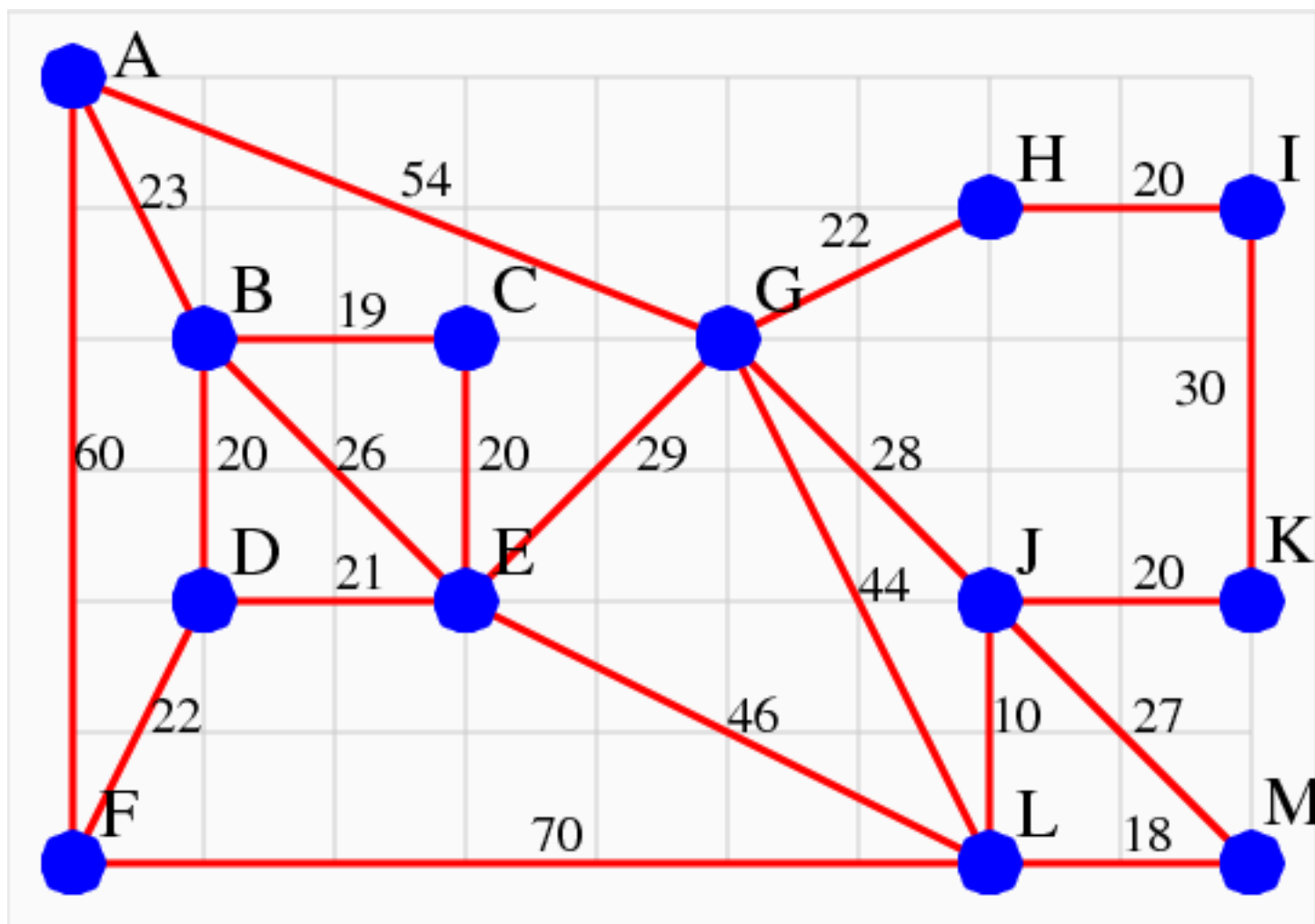


illustration from http://people.sc.fsu.edu/~jburkardt/latex/asa_2011_graphs_homework/

Integer programming solves the Labs difficult **discrete optimization problems**.

Water, Road **sensor** placement, subway, building **sensor** management, **Network** interdiction, **Scheduling** quantum EC, **Protein** structure, **Peptide** docking, **Meshing**, **Space-filling** curves, **Energy** systems, **Pantex** planning, **Vehicle routing**, **Conference** schedule.

Integer Program: Minimize a linear **cost function** subject to **linear inequality and integrality constraints**.

$$\begin{array}{ll}\text{minimize} & c \cdot x \\ \text{subject to} & \\ & A \cdot x \geq b \\ & x \in \mathbf{Z}^n.\end{array}$$

Integer Programming is NP-hard.

Linear Program: Minimize a linear **cost function** subject to **linear inequality constraints only**.

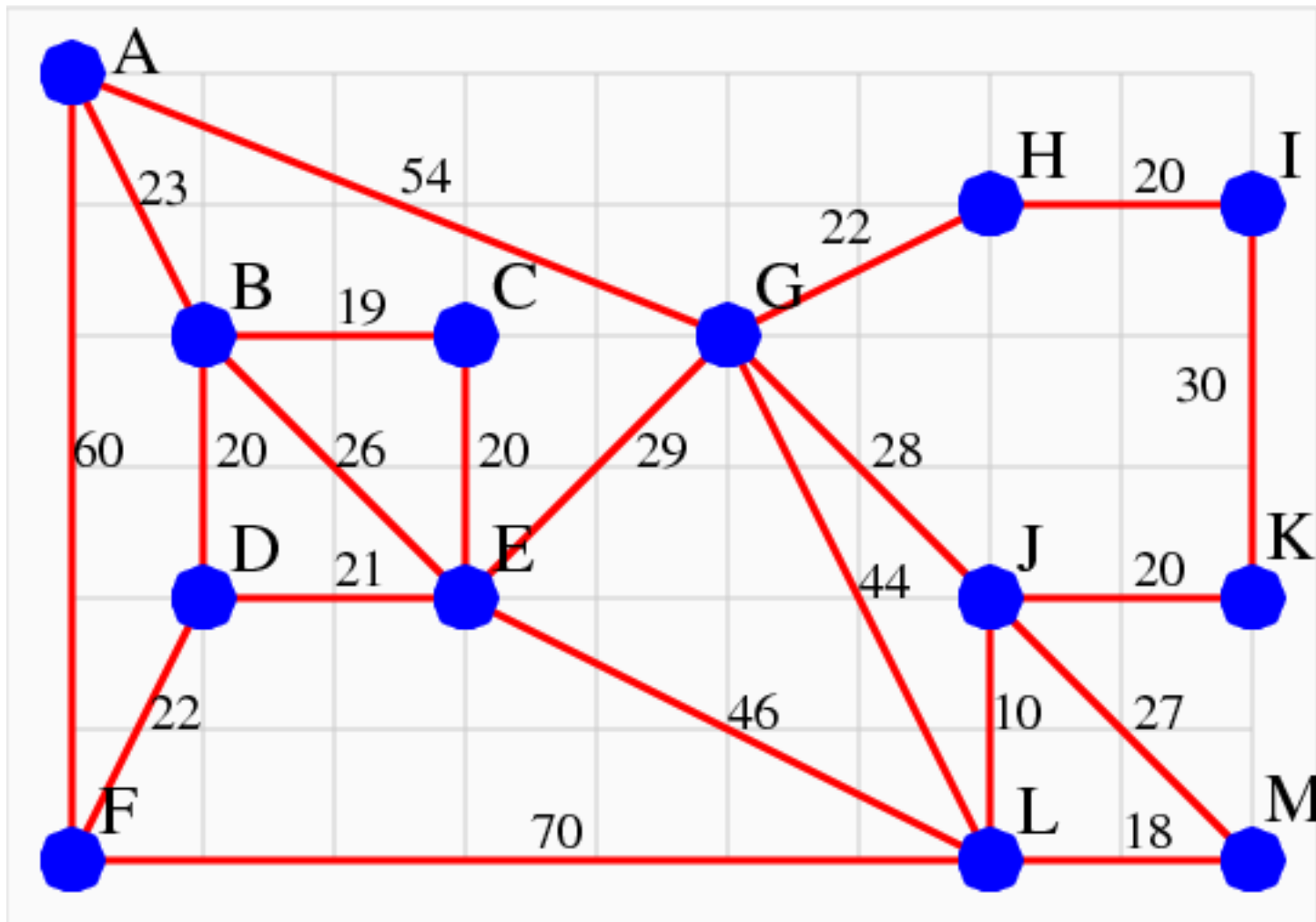
Linear Programming is Polynomially solvable.

Recent focus on **creating formulations** of identified (tractible) problem structures.

Hardness arguments of modeling difficult structures.

Predict solution efficiency of a formulation.

Lockheed Martin Tech Refresh (Watson), improved formulation changed solution times from days to minutes.



Find minimum cost **2-edge connected** spanning subgraph.

Protect shipments **against single failure**.

Doubled edges are allowed and provided at a **discount**.

$$\begin{aligned}
\delta(S) &:= \{e = \{i, j\} \in E : |S \cap e| = 1\}, \\
E(S) &:= \{e = \{i, j\} \in E : |S \cap e| = 2\} \quad \forall S \subset V, \\
x(F) &:= \sum_{e \in F} x_e \quad \forall F \subset E.
\end{aligned}$$

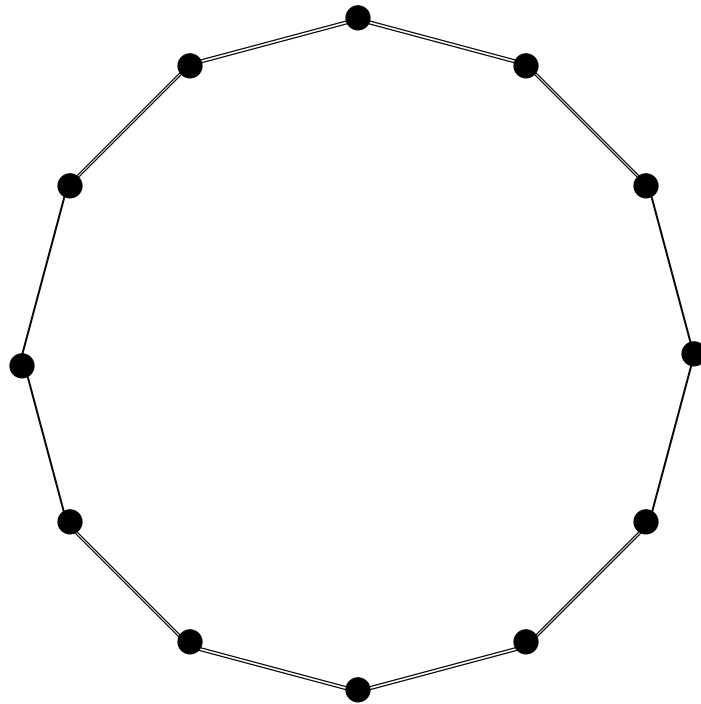
A Classic 2-edge connected spanning subgraph problem.(2ECSS)

$x_e \in \{0, 1, 2\}$ vars: Buy **edge** at **price** c_e .

$$\begin{aligned}
&\min \quad c \cdot x \\
&\text{subj to} \\
&\quad x(\delta(S)) \geq 2 \quad \forall S \subset V, \\
&\quad 0 \leq x_e \leq 2 \quad \forall e \in E, \\
&\quad x_e \in \mathbf{Z} \quad \forall e \in E.
\end{aligned}$$

Drop integrality constraints to get LP relaxation.

The Classic Traveling Salesman Problem (TSP)



A Hamilton cycle is a simple cycle through all the nodes in a graph.

TSP: Find a minimum cost **Hamilton cycle**.

The Graphical Traveling Salesman Problem

TSP: Find a minimum cost **Hamilton cycle**.

Eulerian graph: Connected graph with every node having even degree.

GTSP: Find a minimum cost **Eulerian subgraph**.

An **Eulerian subgraph** can be shortcut to a Hamilton cycle.

The Classic Traveling Salesman Problem (TSP)

TSP: Find a minimum cost Hamilton cycle.

$x_e \in \{0, 1\}$ vars: Buy **edge** at **price** c_e .

$$\begin{array}{ll}\min & c \cdot x \\ \text{subj to} & \\ & x(\delta(S)) \geq 2 \quad \forall S \subset V, \\ & x(\delta(v)) = 2 \quad \forall v \in V, \\ & 0 \leq x_e \leq 1 \quad \forall e \in E, \\ & x_e \in \mathbf{Z} \quad \forall e \in E.\end{array}$$

Double-tree and Christofides heuristics

Select a minimum cost spanning tree $T = (V, E^T)$.

The edge incidence vector:

$\chi_e^T = 1$ iff $e \in E^T$ else $\chi_e^T = 0$.

Double each $e \in E^T$: $2\chi^T$ is the multi-edge incidence vector (has 2s) of our 2-edge connected graph.

Take the set T^{odd} of odd degree nodes of T .

A T^{odd} -join is a graph $M = (V, E^M)$ such that the degree of $v \in V$ is odd iff $v \in T^{odd}$.

Select a minimum cost T^{odd} -join M .

$\chi^T + \chi^M$ is the multi-edge incidence vector of a connected, Eulerian, hence 2-edge connected graph.

Double-tree heuristic for TSP

OPT is cost of minimum cost Hamilton cycle.

Minimum cost spanning tree $T = (V, E^T)$ is at most minimum cost Hamilton cycle without its most expensive edge, a spanning tree!
 $cost(T) \leq OPT$.

Eulerian Graph: connected, even degree at each node. Can be shortcut to Hamilton cycle of no greater cost (triangle inequality costs assumption).

Double each $e \in E^T$: $2\chi^T$ is the multi-edge incidence vector (has 2s) of an Eulerian graph.

Shortcut to a Hamilton cycle. Cost of this heuristic Hamilton cycle is at most $2OPT$!

Christofides heuristic for TSP

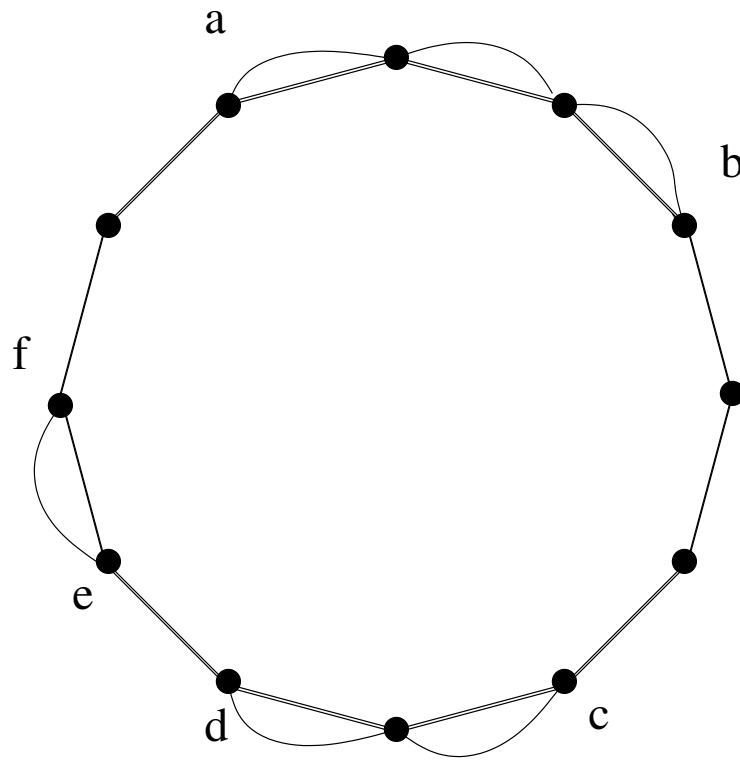
Minimum cost spanning tree $T = (V, E^T)$ has $\text{cost}(T) \leq OPT$.

Take the set T^{odd} of odd degree nodes of T .

A T^{odd} -join is a graph $M = (V, E^M)$ such that the degree of $v \in V$ is odd iff $v \in T^{odd}$.

Select a minimum cost T^{odd} -join M . Edges from a minimum cost Hamilton cycle form 2 such T^{odd} -joins. Hence $\text{cost}(M) \leq 1/2OPT$.

Christofides heuristic for TSP



$\chi^T + \chi^M$ is the multi-edge incidence vector of a connected, Eulerian graph of cost at most $3/2OPT$.

This Eulerian graph can be shortcut to a Hamilton cycle of cost at most $3/2OPT$.

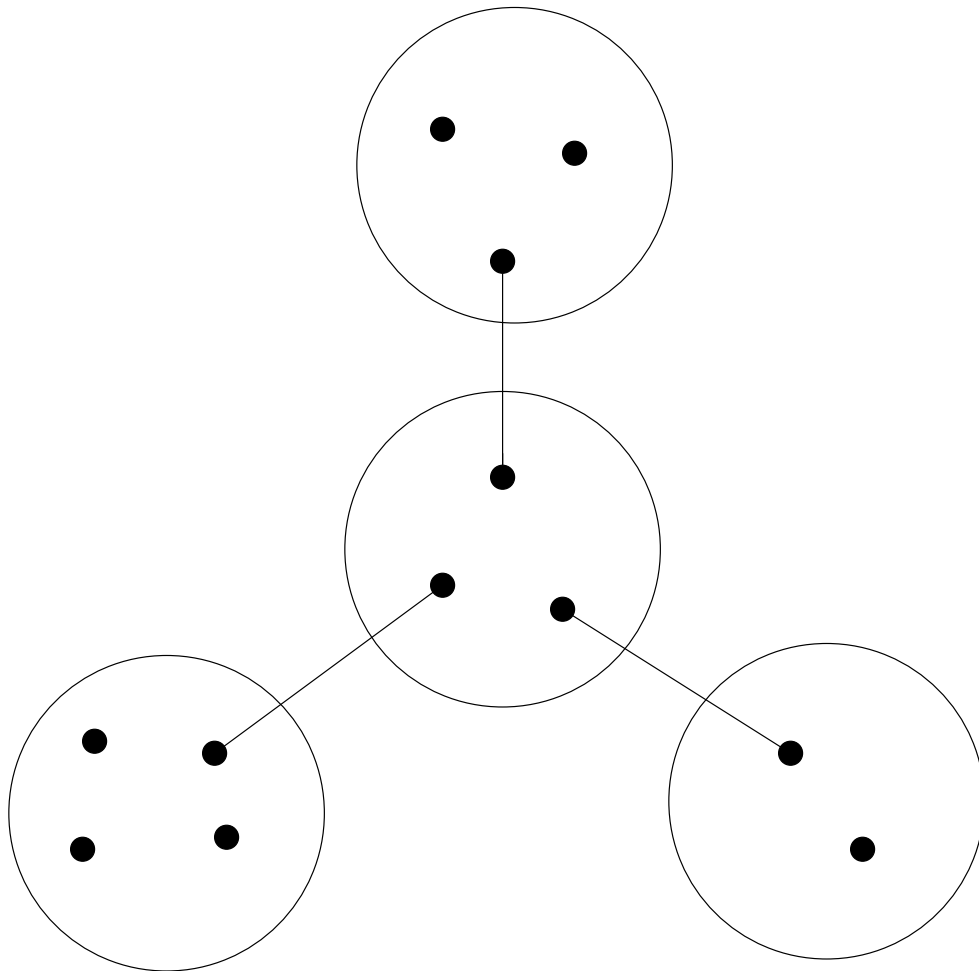
Convex combination arguments

Let \mathcal{W} be a set of vectors W_i having a certain structure of interest (cycles, trees, matchings)

$\sum_i \lambda_i W_i$ with $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$ is a **convex combination** of these vectors.

Convex combination arguments are used to bound the quality of an approximation.

Partition Inequalities for spanning trees



$$x(P) = \frac{1}{2} \sum_{i=1}^k x(\delta(S_i)).$$

Partition Inequality: $x(P) \geq k - 1$.

Double-tree approximation

Let x^* be optimal for LP relaxation.

$$x^*(\delta(S)) \geq 2 \quad \forall S \subset V$$

$$\begin{aligned} x^*(P) &= \frac{1}{2} \sum_{i=1}^k x^*(\delta(S_i)) \\ &\geq \frac{1}{2} 2k = k. \end{aligned}$$

Hence partition inequality $x(P) \geq k - 1$ is satisfied by x^* .

Double-tree approximation

Let x^* be optimal for LP relaxation.

$x^*(\delta(S)) \geq 2 \quad \forall S \subset V$ implies that x^* satisfies the partition inequalities for spanning trees.

Since x^* satisfies the partition inequalities, x^* dominates a convex combination of incidence vectors of spanning trees:

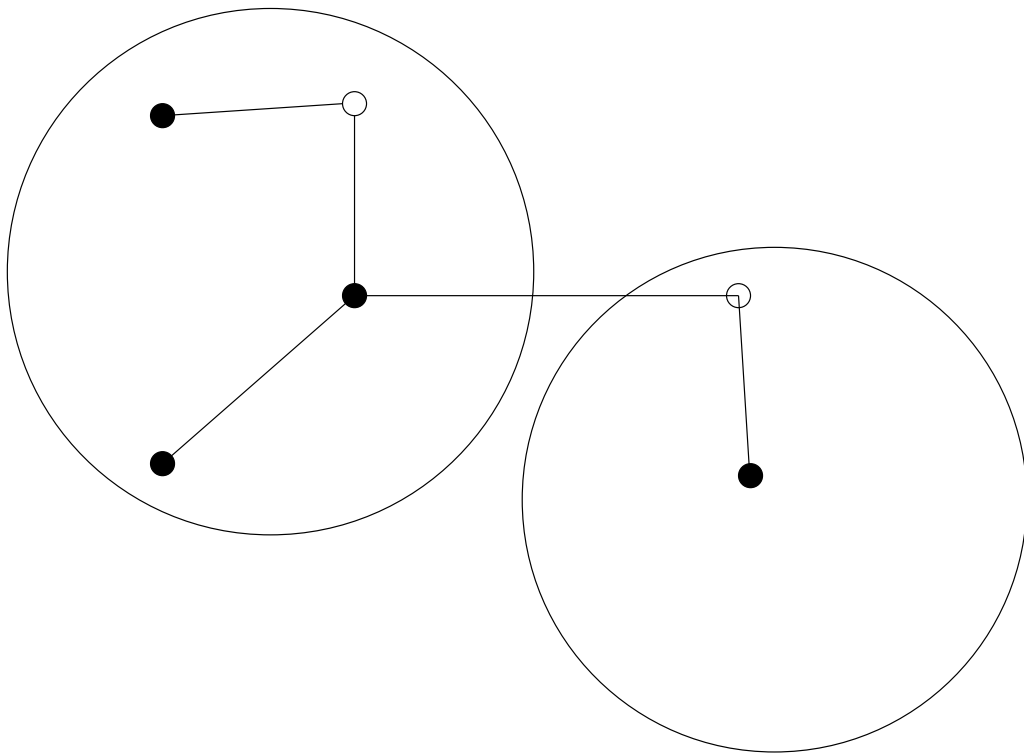
$$x^* \geq \sum_i \lambda_i \chi^{T,i}, \quad (\sum_i \lambda_i = 1).$$

Each tree can be doubled to get a 2-edge connected graph:

$$2x^* \geq \sum_i \lambda_i (2\chi^{T,i}).$$

By averaging argument, one 2-edge connected $2\chi^{T,i}$ costs at most that of $2x^*$.

T-join Inequalities



Black nodes are the T nodes.

Select set of edges subject to
 T nodes and only T nodes have odd degree.

T -join inequalities: $x(\delta(S)) \geq 1$.

Christofides approximation

x^* dominates a convex combination of tree vectors:

$$x^* \geq \sum_i \lambda_i \chi^{T,i}.$$

Let T_i be set of odd degree nodes for tree i .

$\frac{1}{2}x^*(\delta(S)) \geq 1 \quad \forall S \subset V$ implies that $\frac{1}{2}x^*$ satisfies the T_i -join inequalities for each i .

Since $\frac{1}{2}x^*$ satisfies the T_i -join inequalities, $\frac{1}{2}x^*$ dominates a convex combination of T_i -join vectors:

$$\frac{1}{2}x^* \geq \sum_j \mu_{ij} \chi^{M,ij}, \quad (\sum_j \mu_{ij} = 1).$$

For each i, j , $\chi^{T,i} + \chi^{M,ij}$ is 2-edge connected.

$$\frac{3}{2}x^* \geq \sum_i \sum_j \lambda_i \mu_{ij} (\chi^{T,i} + \chi^{M,ij}).$$

By averaging argument, one 2-edge connected $\chi^{T,i} + \chi^{M,ij}$ costs at most that of $\frac{3}{2}x^*$.

Notation for rest of talk

Let $x \in \mathbf{R}^3$. Say $x := (1, 2, 3)$.

let $y \in \mathbf{R}^3$. Say $y := (4, 5, 6)$.

Then $x \oplus y \in \mathbf{R}^3 \times \mathbf{R}^3$.

$x \oplus y = (1, 2, 3, 4, 5, 6)$.

Our new 2-edge connected problem(cap-2-ECSS)

$x_e \in \{0, 1, 2\}$ vars: Buy each **edge** at **price** c_e ,
 $y_e \in \{0, 1\}$ Buy **doubled edge** at a **discount**,
 $x \oplus y \in \mathbf{R}^E \times \mathbf{R}^E$, $c \oplus c' \in \mathbf{R}^E \times \mathbf{R}^E (c'_e \leq 2c_e)$.

$$\begin{aligned} & \min \quad (c \oplus c') \cdot (x \oplus y) \\ & \text{subj to} \\ & \quad x(\delta(S)) + 2y(\delta(S)) \geq 2 \quad \forall S \subset V \\ & \quad 0 \leq x_e \leq 2, 0 \leq y_e \leq 1 \quad \forall e \in E \\ & \quad x_e, y_e \in \mathbf{Z}. \end{aligned}$$

Drop integrality constraints to get LP relaxation.

Integrality gap of 2: $y_e^* = \frac{1}{2}$ for edges of a Hamilton (n edge) cycle. But optimal integer solution is $y_e^{opt} = 1$ for all but one edge of cycle.

A better LP relaxation

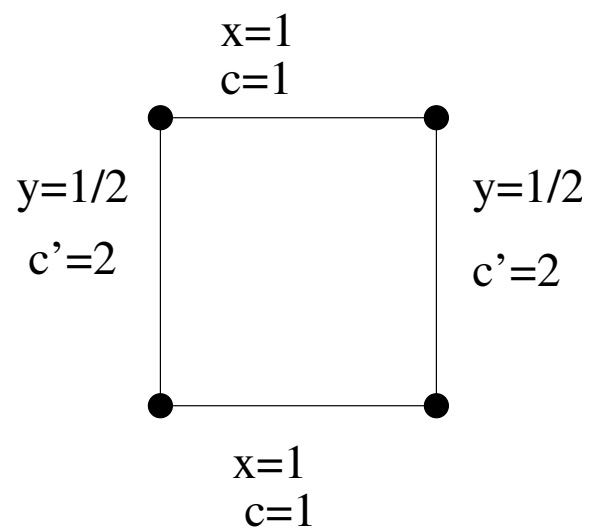
Idea: $x_e + y_e$ dominates a spanning tree vector, denoted by z_e . That is, $x + y$ has enough mass ($n - 1$ edges) to contain a spanning tree, and the tree (z) has acyclic structure.

$$\begin{aligned} \text{Add } x_e + y_e &\geq z_e & z(E(V)) &= n - 1, \\ & & \forall S \subset V \quad z(E(S)) &\leq |S| - 1. \end{aligned}$$

Now $y_e^* = \frac{1}{2}$ on a Hamilton cycle no longer feasible.

Worst **gap** seen is **now** $\frac{3}{2}$.

A better LP relaxation



Worst **gap** seen is **now** $\frac{3}{2}$ when horizontal edges of a square have $x_e^* = 1$ and a cost of 1 and vertical edges of that square have $y_e^* = \frac{1}{2}$ and a cost of 2.

The 5/3 approximation (conclusion of talk)

From approximation one, $2x^* \oplus y^*$ dominates a convex combination of 2-edge connected graphs G_i^1 .

From approximation two, $\frac{3}{2}x^* \oplus 2y^*$ dominates a convex combination of 2-edge connected graphs G_i^2 .

We can combine these as follows:

$$\begin{array}{rcl} & \frac{1}{3} & (\quad 2x^* \oplus y^* \quad \geq \quad \sum_i \lambda_i G_i^1) \\ + & \frac{2}{3} & (\quad \frac{3}{2}x^* \oplus 2y^* \quad \geq \quad \sum_i \lambda_i G_i^2) \\ \hline \end{array}$$

$$\frac{5}{3}x^* \oplus \frac{5}{3}y^* \geq \sum_i \lambda_i G_i.$$

The $\frac{5}{3}$ approximation and integrality gap follows since one of the G_i s cost at most that of $\frac{5}{3}(x^* \oplus y^*)$ by our averaging argument.

Let $x^* \oplus y^*$ be an optimal extreme point solution to our LP.

To keep things simple, assume $x_e^* = 0$ or $y_e^* = 0$ for each $e \in E$.

$x_e^* + y_e^* \geq z_e^*$ and the spanning tree constraints on z^* imply that $x^* + y^*$ dominates a convex combination of incidence vectors $\chi^{T,i}$ of spanning trees $x^* + y^* \geq \sum_i \lambda_i \chi^{T,i}$.

For each spanning tree, break up its set of edges into a set of x -edges and a set of y -edges. Then its incidence vector $\chi^{T,i}$ is broken up into incidence vectors $\chi^{T,x,i}$ and $\chi^{T,y,i}$.

So, $\chi^{T,i} = \chi^{T,x,i} + \chi^{T,y,i}$. Thus,
 $x^* + y^* \geq \sum_i \lambda_i (\chi^{T,x,i} + \chi^{T,y,i})$.

Finally, $x^* \oplus y^* \geq \sum_i \lambda_i (\chi^{T,x,i} \oplus \chi^{T,y,i})$.

Double-tree approximation

Let $x^* \oplus y^*$ be optimal for LP relaxation.

$x_e^* + y_e^* \geq z_e^*$ and the constraints on z^* imply that $x^* \oplus y^*$ dominates a convex combination of incidence vectors of spanning trees in $x \oplus y$ variable space:

$$x^* \oplus y^* \geq \sum_i \lambda_i (\chi^{T,x,i} \oplus \chi^{T,y,i}).$$

The x -part of each tree can be doubled to get a 2-edge connected graph:

$$2x^* \oplus y^* \geq \sum_i \lambda_i (2\chi^{T,x,i} \oplus \chi^{T,y,i}).$$

By averaging argument, one 2-edge connected $2\chi^{T,x,i} \oplus \chi^{T,y,i}$ costs at most that of $2x^* \oplus y^*$.

Christofides approximation

$x^* \oplus y^*$ dominates a convex combination of tree vectors: $x^* \oplus y^* \geq \sum_i \lambda_i (\chi^{T,x,i} \oplus \chi^{T,y,i})$.

Let T_i be set of odd degree nodes for tree i .

$\frac{1}{2}x^*(\delta(S)) + y^*(\delta(S)) \geq 1 \quad \forall S \subset V$ implies that $\frac{1}{2}x^* + y^*$ satisfies T_i -join inequalities for each i .

Since $\frac{1}{2}x^* + y^*$ satisfies the T_i -join inequalities, $\frac{1}{2}x^* \oplus y^*$ dominates a convex combination of T_i -join vectors:

$$\frac{1}{2}x^* \oplus y^* \geq \sum_j \mu_{ij} (\chi^{M,x,ij} \oplus \chi^{M,y,ij}).$$

Christofides approximation

For each i, j , $(\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij})$ is 2-edge connected.

$$x^* \oplus y^* + \frac{1}{2}x^* \oplus y^* = \frac{3}{2}x^* \oplus 2y^*.$$

$$\frac{3}{2}x^* \oplus 2y^* \geq \sum_i \sum_j \lambda_i \mu_{ij} (\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij}).$$

By averaging argument, one 2-edge connected $(\chi^{T,x,i} + \chi^{M,x,ij}) \oplus (\chi^{T,y,i} + \chi^{M,y,ij})$ costs at most that of $\frac{3}{2}x^* \oplus 2y^*$.

The 5/3 approximation

From the Double-tree approximation, $2x^* \oplus y^*$ dominates a convex combination of 2-edge connected graphs G_i^1 .

From the Christofides approximation, $\frac{3}{2}x^* \oplus 2y^*$ dominates a convex combination of 2-edge connected graphs G_i^2 .

We can combine these as follows:

$$\begin{array}{rcl} & \frac{1}{3} & (\quad 2x^* \oplus y^* \quad \geq \quad \sum_i \lambda_i G_i^1) \\ + & \frac{2}{3} & (\quad \frac{3}{2}x^* \oplus 2y^* \quad \geq \quad \sum_i \lambda_i G_i^2) \\ \hline \end{array}$$

$$\frac{5}{3}x^* \oplus \frac{5}{3}y^* \geq \sum_i \lambda_i G_i.$$

The $\frac{5}{3}$ approximation and integrality gap follows since one of the G_i s cost at most that of $\frac{5}{3}(x^* \oplus y^*)$ by our averaging argument.

A 5/3 approximation algorithm

Let $c_{ij}^1 := \min(2c_{ij}, c'_{ij})$.

Find minimum cost spanning tree T with c^1 costs.

Insert a doubled- x -edge or a y -edge for each edge in spanning tree, whichever is cheaper.

This results in a 2-edge connected graph T^1 .

$$\text{cost}(T^1) \leq (c \oplus c') \cdot (2x^* \oplus y^*)$$

A 5/3 approximation algorithm

Let $c_{ij}^2 := \min(c_{ij}, c'_{ij})$.

Find minimum cost spanning tree T with c^2 costs.

Insert an x -edge or a y -edge for each edge in spanning tree, whichever is cheaper.

This results in a connected graph T^2 .

$$\text{cost}(T^2) \leq (c \oplus c') \cdot (x^* \oplus y^*)$$

Let T_{odd} be the set of odd degree nodes for tree T^2 .

Find minimum cost T_{odd} -join M with c^2 costs.

Insert an x -edge or a y -edge for each edge in T_{odd} -join, whichever is cheaper.

This results in M^2 .

$$\text{cost}(M^2) \leq (c \oplus c') \cdot (\frac{1}{2}x^* \oplus y^*)$$

By Christofides approximation $T^2 + M^2$ is a 2-edge connected graph.

$$\text{cost}(T^2 + M^2) \leq (c \oplus c') \cdot (\frac{3}{2}x^* \oplus 2y^*)$$

A 5/3 approximation algorithm

$$\min(\text{cost}(T^1), \text{cost}(T^2 + M^2)) \leq 1/3 \text{cost}(T^1) + 2/3 \text{cost}(T^2 + M^2)$$

$$\leq 1/3(c \oplus c') \cdot (2x^* \oplus y^*) + 2/3(c \oplus c') \cdot (\frac{3}{2}x^* \oplus 2y^*)$$

$$= 5/3(c \oplus c') \cdot (x^* \oplus y^*).$$

The best of T^1 and $T^2 + M^2$ is thus a 5/3-approximation.