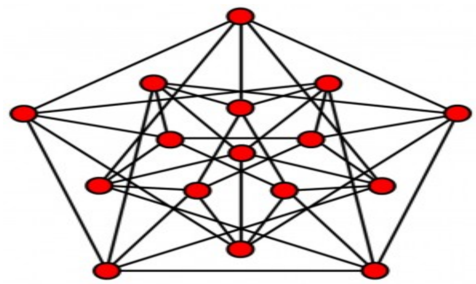
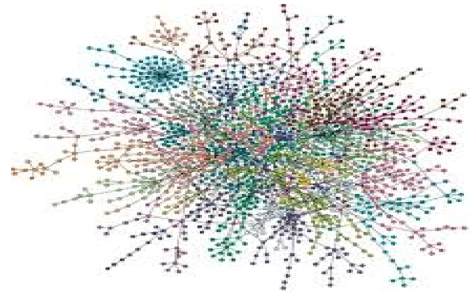
Pescara, Italy, July 2019





DIGRAPHS III
Applications: Pagerank, Contagion, Ford-Fulkerson

Based on various sources.

J. J. P. Veerman,
Math/Stat, Portland State Univ., Portland, OR 97201, USA.
email: veerman@pdx.edu

Conference Website:
www.sci.unich.it/mmcs2019

SUMMARY:

* This is a review of three important applications of graph theory presented in a way that is consistent with the earlier lectures on the theory of digraphs.

* We discuss the pagerank algorithm and give a treatment that is dual to the usual one, namely cast in terms of consensus (and not random walk).

* We discuss contagion on a graph and give some elementary results about the probability that the invading species 'takes over'.

* We discuss how to optimize transport on digraphs where each edge has a maximum capacity. This is known as the Ford Fulkerson algorithm and the max-flow is min-cut theorem.

OUTLINE:

The headings of this talk are color-coded as follows:

The Pagerank Algorithm

Teleporting and Pagerank

Contagion and Evolution

The Probability that the Invader Wins

The Ford Fulkerson Algorithm

When Ford Fulkerson Fails

.

P A G E R A N K

Recall of Definitions

We recall some definitions.

Definition: The **combinatorial adjacency matrix** Q of the graph G is defined as:
$Q_{ij} = 1$ if there is an edge ji (if **"i sees j"**) and 0 otherwise.
If vertex i has no incoming edges, set $Q_{ii} = 1$ (create a loop).

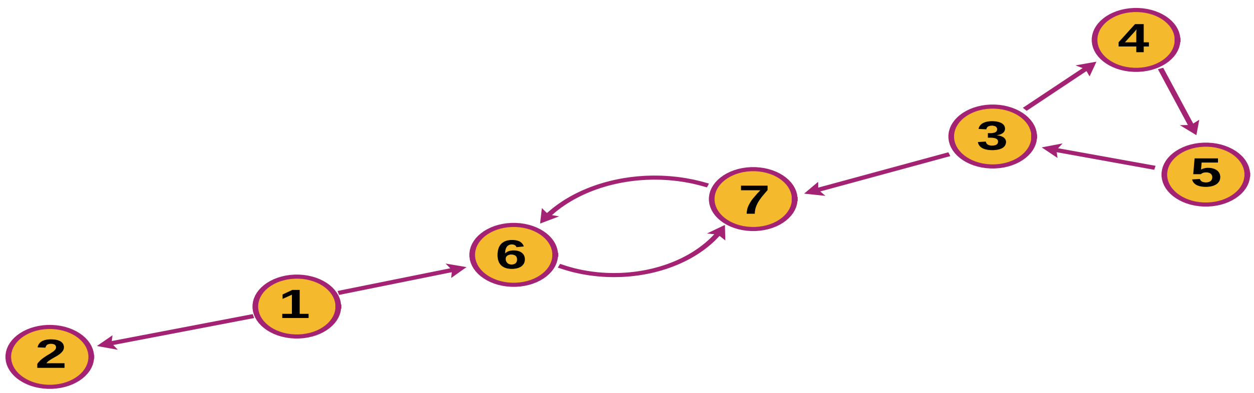
Remark: Instead of creating a loop, sometimes all elements of the ith row are given the value $1/n$. This is called Teleporting! The matrix is denoted by $\bar{Q}$.

Definition: The **in-degree matrix** D is a diagonal matrix whose i diagonal entry equals the number of (directed, incoming) edges xi, $x \in V$.

Definition: The matrices $S \equiv D^{-1}Q$ and $\bar{S} \equiv D^{-1}\bar{Q}$ are called the **normalized adjacency matrices**. By construction, they are **row-stochastic** (non-negative, every row adds to 1).

Definition: The **pagerank adjacency matrices** are given by $S_p = \beta S + \dfrac{1-\beta}{n} J$, where S may be replaced by $\bar{S}$ ("with teleporting").

The Pagerank Algorithm



Recall: **consensus** flows ***with*** the arrows, **random walk** goes ***against*** them.

The original pagerank algorithm by Page and Brin (as discussed in [5]). Our dual treatment mostly follows [1].

Definition (Pagerank): Let J be the $n \times n$ *all ones* matrix. Define, for $\beta = 0.85$, say,

$$S_p \equiv \beta S + \frac{1-\beta}{n} J$$

Determine **unique invariant probability measure** $\wp$ for the random walk S_p. Pagerank of i equals $\wp(i)$. Thus, solve:

$$\boldsymbol{\wp = \wp S_p} \; .$$

Crash Course Pagerank

$$S_p \equiv \beta S + \frac{1-\beta}{n} J$$

S_p strictly positive (every vertex "sees" every other vertex).
Therefore: one reach!
Thus $\wp$ is unique (thms 3, 4, 5, Digraphs II).

S and J are simultaneously diagonalizable.
Denote the *all ones* vector by $\mathbf{1}$.

Leading eigenpair: eval 1 with evec $\mathbf{1}$ (for S and J).
Other evecs: eval at most $\beta \approx 0.85$ for S and 0 for J.

Very fast convergence: $0.85^{57} \approx 10^{-4}$.
Can formulate the whole thing without using matrices.

Observation: Original algorithm uses $\bar{S}$ instead of S.
[1] shows that the two rankings are trivially related.

Dual Approach to Pagerank 1

Recall Thm 8 of Digraphs II: Displacements in consensus caused by initial displacement x_0:

$$\dot{x} = -\mathcal{L}x \quad \Longrightarrow \quad \lim_{t\to\infty} x^{(t)} = \Gamma x^{(0)}$$

Left multiplying by $\frac{1}{n}\mathbf{1}^T$ has the effect of taking an average of these displacements.

Definition: The **influence** $I(i)$ of the vertex i is **average** of the displacements caused by unit displacement e_i:

$$I(i) \equiv \frac{1}{n}\mathbf{1}^T\,\Gamma\,e_i = \frac{1}{n}\mathbf{1}^T\left(\sum_{m=1}^{k}\gamma_m\otimes\bar{\gamma}_m\right)e_i$$

$\mathbf{1}$ is the *all ones* vector.

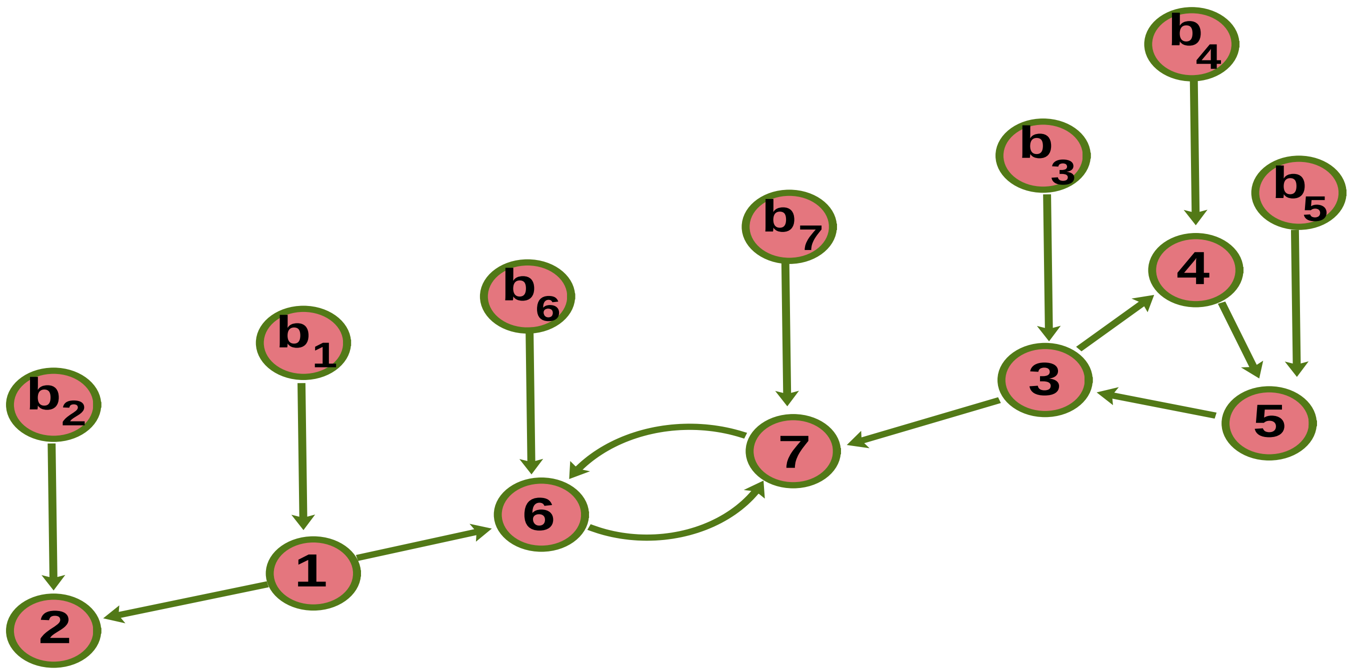
Problem:

By assoc., non-zero only if $\bar{\gamma}_m\,e_i \neq 0$ for some m.
Thus $I(i) > 0$ only if i is in a cabal (by defn $\bar{\gamma}_m$). Not interesting!

Definition: The **extended graph** G_α. for every vertex v in V, attach a new vertex b_v and an edge $b_v v$ with strength α.

Think of b_v as the *boss/owner/administrator* of the page v.

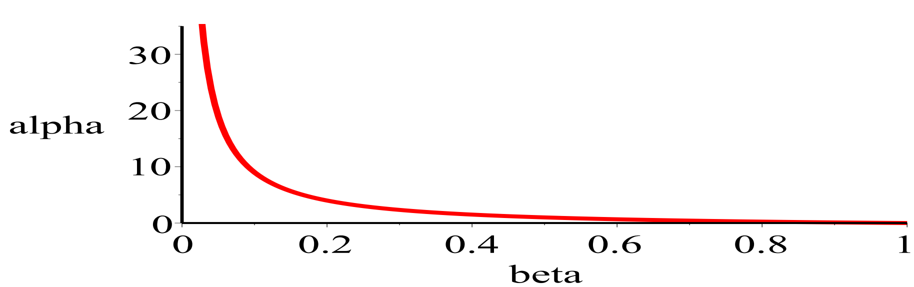
Dual Approach to Pagerank 2



G_α has n leaders b_i. Each of these has a non-zero influence $\tilde{I}(b_i)$. The tilde ($\tilde{.}$) indicates extended graph.

Theorem 1 (Pagerank Theorem) [1]: If we choose $\alpha = \frac{1-\beta}{\beta}$, then the pagerank $\wp(i)$ of i equals $2\tilde{I}(b_i) - \frac{1}{n}$.

The factor 2 is because the pagerank in G_α is averaged over $2n$ vertices. We have to subtract $\frac{1}{n}$ because we do not want to count the displacement of the "virtual" page b_i.



Sketch of Proof Pagerank Theorem

The **extended Laplacians** are:

$$\tilde{L} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\alpha I & \alpha I + \mathcal{L} \end{pmatrix} \quad \text{and} \quad \tilde{\mathcal{L}} = \frac{1}{1+\alpha} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\alpha I & \alpha I + \mathcal{L} \end{pmatrix}$$

Theorem 4 (in D II) says that the kernel of $\tilde{\mathcal{L}}$ has basis $\begin{pmatrix} e_m \\ \eta_m \end{pmatrix}$ where $m \in \{1, \cdots n\}$. Substituting gives:

$$\eta_m = (I + \alpha^{-1}\mathcal{L})^{-1} e_m$$

Thus the **influence of b_m on the "rest"** (non-leaders) is

$$I(m) = \frac{1}{n} \mathbf{1}^T (I + \alpha^{-1}\mathcal{L})^{-1} e_m$$

Theorem 10 (D II) **implies*** that $\sum_m I(m) = 1$ and so

$$\boldsymbol{p = \frac{1}{n} 1^T (I + \alpha^{-1}\mathcal{L})^{-1}}$$

is a row-vector of influences and a **probability measure**.

***Alternatively:** If all leaders move 1 unit, all others eventually do the same.

Sketch of Proof Continued

Exercise 1: J is the all ones matrix. Show that

$$\beta S + \frac{1-\beta}{n} J = I + \frac{\alpha}{1+\alpha}\left(\frac{1}{n} J - (I + \alpha^{-1}\mathcal{L})\right)$$

Hint: $\alpha = \frac{1-\beta}{\beta}$ *or* $\beta = \frac{1}{1+\alpha}$.

Exercise 2: Show that

$$\left(\frac{1}{n}\mathbf{1}^T (I + \alpha^{-1}\mathcal{L})^{-1}\right)\left(\frac{1}{n} J - (I + \alpha^{-1}\mathcal{L})\right) = 0$$

Hint: For a probability measure p*, we have* $pJ = \mathbf{1}^T$.

The exercises show that the probability measure p satisfies

$$\mathbf{p} = \mathbf{p}\left(\boldsymbol{\beta}\mathbf{S} + \frac{\mathbf{1}-\boldsymbol{\beta}}{\mathbf{n}}\mathbf{J}\right)$$

And thus p equals the pagerank $\wp$.

Exercise 3: Relate this to the influence of b_m in the extended graph.
Hint: the extended graph has 2n vertices and the initial condition $x_{b_n} = 1$ *moves none of the leaders except* b_n *itself.*

.

P A G E R A N K

W I T H   T E L E P O R T I N G

O R   W I T H O U T ?

The Two Cases

Lemma: J is the *all ones* matrix. For any probability vector p, we have

$$pJ = \mathbf{1}^T$$

So, to find the pagerank, we find the unique solution of:

$$\wp = \wp\left(\beta S + \frac{1-\beta}{n} J\right) \quad \Longrightarrow \quad \wp(I - \beta S) = \frac{1-\beta}{n}\mathbf{1}$$

There are two cases:
Case I: no teleporting.
Case II: with teleporting, marked by an overbar ($\bar{\boldsymbol{S}}$).

Partition vert's in B, set of <u>leaders</u>, and complement R. The ith rows of the S's differ only if $i \in L$.

$$(\wp_B, \wp_T)\left[\begin{pmatrix} I_B & 0 \\ 0 & I_R \end{pmatrix} - \beta \begin{pmatrix} S_{BB} & S_{BR} \\ S_{RB} & S_{RR} \end{pmatrix}\right] = \frac{1-\beta}{n}(\mathbf{1}_B, \mathbf{1}_T)$$

Case I:

$$\begin{pmatrix} S_{BB} & S_{BR} \\ S_{RB} & S_{RR} \end{pmatrix} = \begin{pmatrix} I_{BB} & \mathbf{0} \\ S_{RB} & S_{RR} \end{pmatrix}$$

Case II:

$$\begin{pmatrix} \bar{S}_{BB} & \bar{S}_{BR} \\ \bar{S}_{RB} & \bar{S}_{RR} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} J_{BB} & \frac{1}{n} J_{BR} \\ S_{RB} & S_{RR} \end{pmatrix}$$

The Two Cases

Exercise 4: Write out the **orange equation** for the two cases. Show that $\wp_B$, $\bar{\wp}_R$, and $\bar{\wp}_B$ all can be expressed in terms of $\wp_R$.
Hint: you need to use the lemma.

Definition: Use π for probability that walker is in L:

$$\pi := \wp_B \, \mathbf{1}_B \qquad \text{and} \qquad \bar{\pi} := \bar{\wp}_B \, \mathbf{1}_B$$

Exercise 5: Exercise 4 and the definition imply the following.

Theorem 2 [1]: We have

$$\begin{aligned} \bar{\wp}_B &= \wp_B - \beta(1-\bar{\pi})\,\wp_B \\ \bar{\wp}_R &= \wp_R + \tfrac{\beta}{1-\beta}\,\bar{\pi}\,\wp_R \end{aligned}$$

Upon "teleporting", leaders go down a bit, "rest" goes up. Like a card shuffle. The two subsets maintain relative rankings within them.

One Loose Thread

To complete the picture, need to express $\bar{\pi}$ in terms of "un-teleported" quantities.

Exercise 5: Sum the components of the first equation of Theorem 2 to show:

Corollary: $\bar{\pi} = \dfrac{(1-\beta)\pi}{(1-\beta\pi)}$.

Exercise 6: Substitute this into Theorem 2 to show:

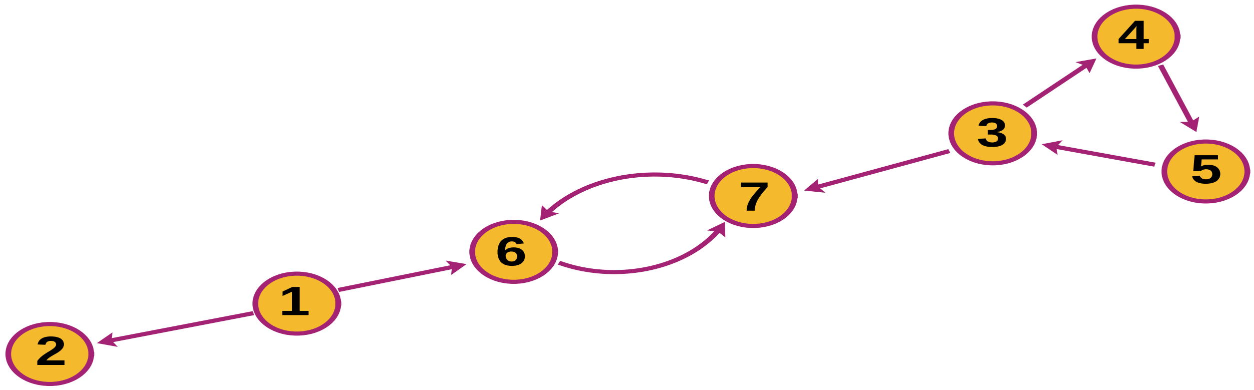
Corollary:

$$\bar{\wp}_B = \left(\frac{1-\beta}{1-\beta\pi}\right) \wp_B$$

$$\bar{\wp}_R = \left(\frac{1}{1-\beta\pi}\right) \wp_R$$

Thus pagerank with teleporting can be trivially expressed in terms of pagerank without teleporting.

Example



$$\mathcal{L} = \left(\begin{array}{c|c|ccc|cc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ \hline -1/2 & 0 & 0 & 0 & 0 & 1 & -1/2 \\ 0 & 0 & -1/2 & 0 & 0 & -1/2 & 1 \end{array}\right)$$

Pagerank as function of $\underline{\beta}$:

$$\wp = 7^{-1}\mathbf{1}^T (I + \alpha^{-1}\mathcal{L})^{-1} = 7^{-1}\mathbf{1}^T \left(I + \frac{\beta}{1-\beta}\mathcal{L}\right)^{-1}$$

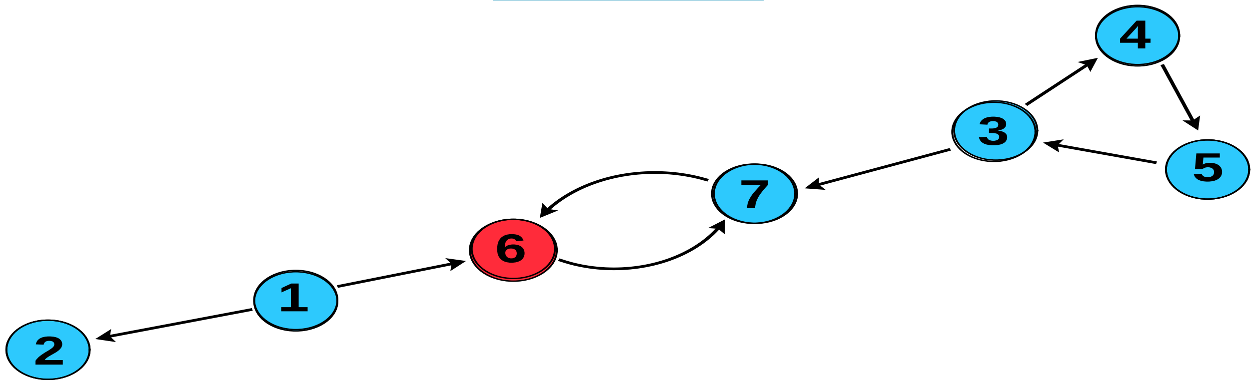
$$\begin{aligned} \wp(0.10) &= (0.165,\ 0.129,\ 0.150,\ 0.143,\ 0.144,\ 0.135,\ 0.135) \\ \wp(0.40) &= (0.236,\ 0.086,\ 0.166,\ 0.147,\ 0.152,\ 0.107,\ 0.107) \\ \wp(0.60) &= (0.290,\ 0.057,\ 0.174,\ 0.154,\ 0.162,\ 0.082,\ 0.082) \\ \wp(0.90) &= (0.388,\ 0.014,\ 0.186,\ 0.178,\ 0.182,\ 0.026,\ 0.026) \end{aligned}$$

$$\begin{aligned} \bar{\wp}(0.10) &= (0.151,\ 0.131,\ 0.152,\ 0.145,\ 0.146,\ 0.138,\ 0.138) \\ \bar{\wp}(0.40) &= (0.156,\ 0.095,\ 0.183,\ 0.162,\ 0.168,\ 0.118,\ 0.118) \\ \bar{\wp}(0.60) &= (0.140,\ 0.069,\ 0.211,\ 0.186,\ 0.196,\ 0.099,\ 0.099) \\ \bar{\wp}(0.90) &= (0.060,\ 0.022,\ 0.286,\ 0.273,\ 0.279,\ 0.040,\ 0.040) \end{aligned}$$

.

CONTAGION OR EVOLUTION IN DIGRAPHS

Fitness



G initially has **blue** vertices. Color 1 vertex **red** (the 'seed').

Definition: Fitness is the probability (a priori likelihood) of procreating. How many kids are you likely to have? More precisely: anyone of "your" population group.

Definition: Assume from now on that

$$\textbf{fitness(red)} = r \cdot \textbf{fitness(blue)}$$

Contagion/procreation occurs along a directed graph. Gene flow is information flow, so it **follows the arrows**.

First Results

Definition: Fixation probability P is the probability that 1 red takes over the entire graph by contagion.

Gene flow **follows the arrows**. So in essence we look for influence vectors (see DII).

Corollary: Given a digraph G.
a) Red cannot take all ($P = 0$) if G has more than 1 reach.
b) Red dies out ($P = 0$) if the seed is not in a cabal.

Proposition: Given a digraph G with n vertices. If red conquers cabal m, then red will average a proportion $\frac{1}{n}\mathbf{1}\gamma_m$ of the population.

Idea of Proof: By DII, Thm 6: $\gamma_m(j)$ is the probability that j's information comes from cabal m.

Thus the relevant question becomes:
Investigate P for **Strongly Connected Components** (SCC's).

Contagion on SCC's

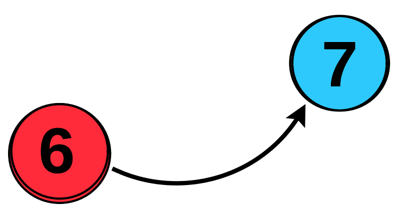
Definition: Probability measure μ on **outgoing edges**:
-Assign blue vertices a probability b (normalization).
-Assign red vertices a probability $r \cdot b$.
-Assign each of the outgoing edges at a vertex equal probability whose sum is the probability of that vertex.

From now on $x^{(n)}(i)$ is the **color** of vertex i at time n.

$$x^{(n)}(i) = 0 \textbf{ if uninfected ; } \; x^{(n)}(i) = 1 \textbf{ if infected}$$

An **"evolutionary" dynamical system $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$**:
At time step n, choose a μ-random (outgoing) edge $v \to w$.
Then v 'spreads' to w, or w assumes the color of v:

$$x^{(n+1)}(7) := x^{(n)}(6)$$



Now denote by **m** the numbers of **infected**, and by $\mathbf{n - m}$ the number of **uninfected**.

The Dynamical System F

As in DII, set $Q_{ij} = 1$ if there is edge ji and 0 otherwise. But this time the average is over **outbound edges**. There are no loops ($Q_{ii} = 0$).

Definition: Normalized out-degree adjacency matrix $W \equiv QD^{-1}$ where D is the diagonal matrix of **column** sums.

Thus the time-dependent prob. to select the edge ji equals

$$\mathrm{Pr}(\boldsymbol{ji}) = \frac{\boldsymbol{W_{ij}}}{\boldsymbol{n - m + rm}}$$

if j is **uninfected**, and $\boldsymbol{r}$ times that if j is **infected**.

$\pi_{m,m+1}$ (resp. $\pi_{m,m-1}$) is the probability that in next time step the system goes from m to $m+1$ (resp. $m-1$) infected.

Lemma: For $m \in \{1, \cdots n-1\}$ we have

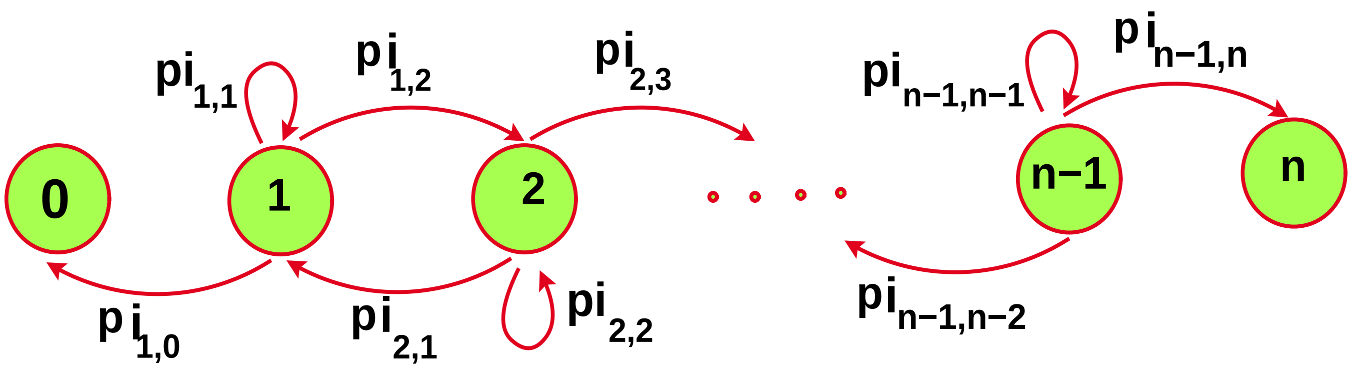
$$\begin{aligned}
\boldsymbol{\pi_{m,m+1}} &= \frac{\boldsymbol{r \sum_{ij} W_{ij}\,(1 - x(i))x(j)}}{\boldsymbol{n - m + rm}} \\
\boldsymbol{\pi_{m,m-1}} &= \frac{\boldsymbol{\sum_{ij} W_{ij}\,x(i)(1 - x(j))}}{\boldsymbol{n - m + rm}}
\end{aligned}$$

Exercise 7: Compute $\pi_{m,m}$.

Exercise 8: Use that W is **column stochastic** to verify that $\quad \pi_{m,m+1} + \pi_{m,m-1} + \pi_{m,m} = 1$.

The Associated Graph

Definition: The associated graph A is a graph on $n+1$ vertices. The vertex i stands for the total number of infected in G. The dynamical system F induces a **random walk** R on A with transition probabilities $\pi_{i,i\pm 1}$ (see figure).



Definition: Let S be the rw adjacency matrix on A. Thus

$$S_{ij} = \pi_{i,j} \quad \text{with row-sum 1}$$

Important: S flips the arrows in the graph. Random walk becomes

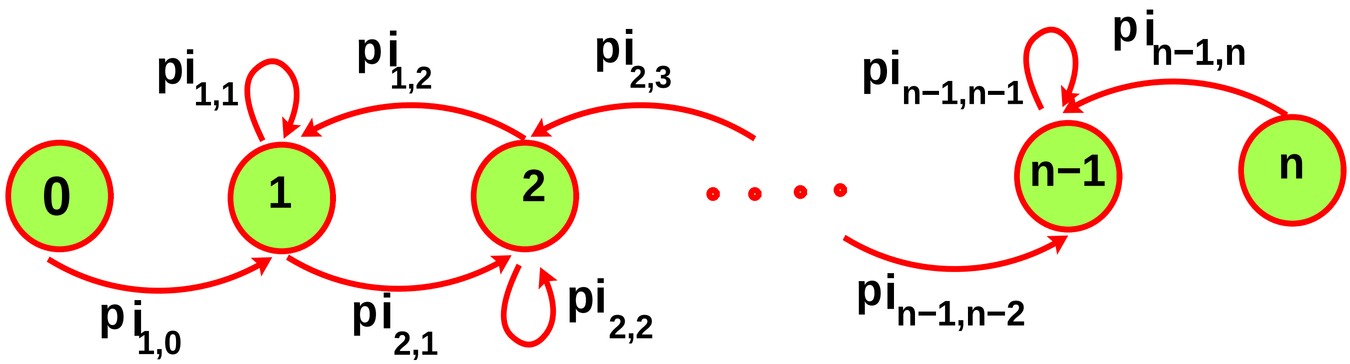
$$p^{(n+1)} = p^{(n)} S$$

The **problem** is that the transition probabilities $\pi_{i,i\pm 1}$ depend on **which** i **vertices** are infected.

Reversing the arrows, we see.....

Standard Format of Associated Graph

Reversing the arrows, we see.....



Now the rw moves against the arrows, as per the conventions in DII.

.

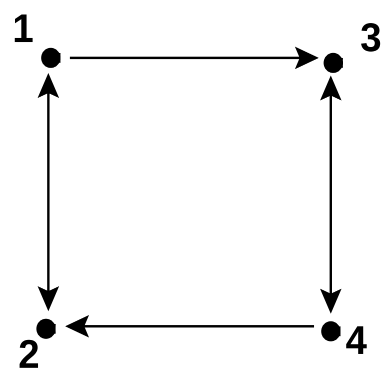
THE FIXATION PROBABILITY

Doubly Stochastic SCC's

Doubly stochastic: row sum is 1 and column sum is 1. All elemts ≥ 0.

Theorem 3: We have the following:
a) G is SCC $\Longleftrightarrow$ A has reaches $\{0, \cdots n-1\}$ with 0 as leader and $\{1, \cdots n\}$ with n as leader.
b) [4] W is doubly stochastic $\Longleftrightarrow$ $\boldsymbol{\pi_{m,m+1} = r\, \pi_{m,m-1}}$.

(b) holds if W symm. But there are interesting other examples.



Example: This graph has norm. outdegr. adj. matrix W

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \Longrightarrow \quad W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

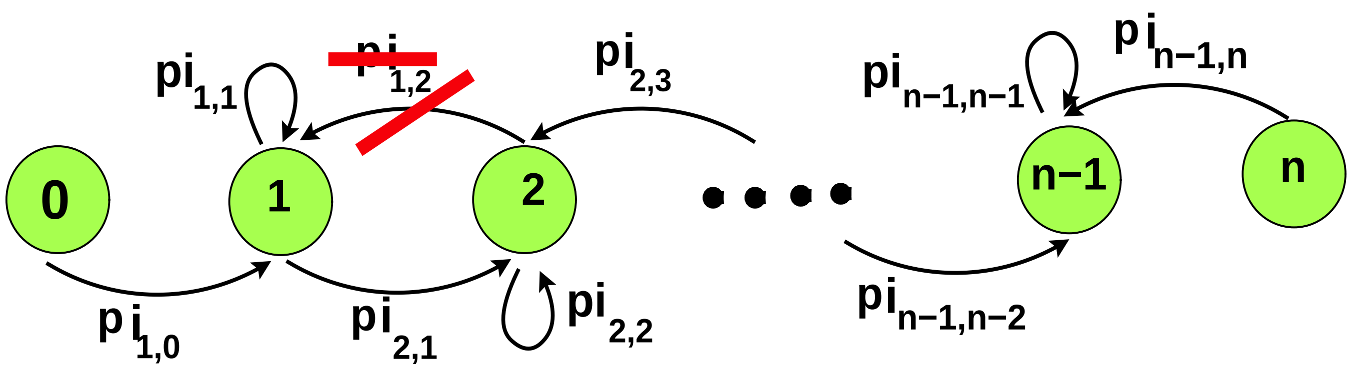
Spectrum $\{-1, 0^{(2)}, 1\}$ with one 2-dimensional Jordan block.

Sketch of Proof of Theorem 3

Proof of (a).

0 and n are leaders. If there are 0 infected, no infections can occur. So $S_{0i} = 0$ for all i. Same for n.

Recall that the $\pi_{i,i+1}$ depend on which vertices are infected. Suppose that **at any point in the process G is not SCC**. This can happen if and only if there is a non-trivial set V of i red or blue vertices that cannot infect V^c. In this case one of $\pi_{i,i\pm1}$ is zero. And that means **A has reaches different from the theorem** (see figure).



Sketch of Proof of Theorem 3, Cont'd

Proof of (b).

Suppose W doubly stochastic. Recall

$$(n - m + rm)\pi_{m,m+1} = r\,(\mathbf{1} - x)Wx = r\,(\mathbf{1}Wx - xWx)$$

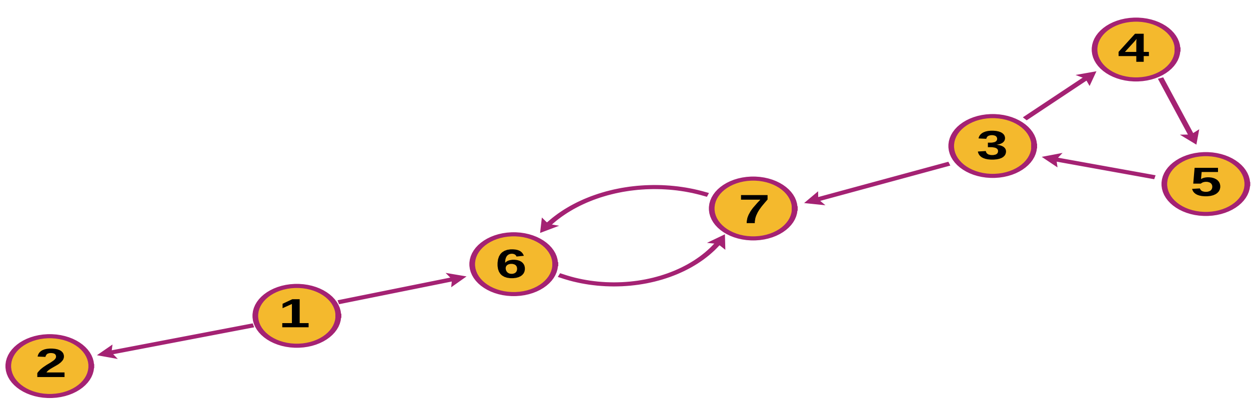
$$(n - m + rm)\pi_{m,m-1} = xW(\mathbf{1} - x) = xW\mathbf{1} - xWx$$

Use double stochasticity of W to see that $\mathbf{1}Wx = xW\mathbf{1}$.
Then $\pi_{m,m+1}$ equals $r\,\pi_{m,m-1}$.

If $\pi_{m,m+1}$ equals $r\,\pi_{m,m+1}$, set $x = e_\ell$. The same computation now shows that then $\mathbf{1}We_\ell = e_\ell W\mathbf{1}$. **Then W is doubly stochastic.**

Remark. It is possible that $\pi_{m,m\pm1} = 0$. This can happen, for example, if G is not an SCC.

Exercise 9: Analyze the associated graph (and its reaches) for the graph in the figure.



Fixation Probability for Doubly Stoch.

Recall that **infected** vertices have **relative fitness** **r**. The fixation probability, is the probability that 1 **red** vertex takes over the entire graph.

Theorem 4: If G is an SCC whose norm. out-degree adj. matrix is doubly stochastic, then G has fixation probability equal to $\frac{1-r^{-1}}{1-r^{-n}}$.

When $r = 1$, use L'Hôpital.

The fixation probability as function of r and n.

n \ r	4	2	1	0.5	0.25
4	0.753	0.53	1/4	0.15	1.18E−2
8	0.75	0.502	1/8	3.9E−3	4.5E−5
32	0.75	0.50	1/16	1.5E−5	6.9E−10
64	0.75	0.50	1/32	2^{-10}	1.6E−19

Sketch of Proof of Theorem 4

Thm 3a): The associated graph A has reach $\{1, \cdots m\}$.

DII, Thm 5): Ker$\mathcal{L}$ contains γ st $\gamma(n) = 1$ and $\gamma(0) = 0$.

DII, Thm 6): $\gamma(1)$ is the fixation probability.

Thm 3b): The rw adjacency of the assoc. graph A is

$$S = \begin{pmatrix} 1 & 0 & \cdots & & \\ \pi_{1,0} & \pi_{1,1} & r\pi_{1,0} & \cdots & \\ & \cdots & & \cdots & \\ & \cdots & \pi_{n-1,n-2} & \pi_{n-1,n-1} & r\pi_{n-1,n-2} \\ & & \cdots & 0 & 1 \end{pmatrix}$$

with row-sum 1.

Exercise 10: From $(I - S)\gamma = 0$, derive

$$(\gamma(i+1) - \gamma(i)) = r^{-1}(\gamma(i) - \gamma(i-1))$$

Furthermore, by telescoping, and the fact that $\gamma(n) = 1$:

$$\sum_{i=0}^{n-1} (\gamma(i+1) - \gamma(i)) = 1$$

Exercise 11: Show that exercise 9 implies that

$$1 = \sum_{i=0}^{n-1} (\gamma(i+1) - \gamma(i)) = \sum_{i=0}^{n-1} r^{-i}\gamma(1)$$

from which the fixation probability follows.

THE FORD FULKERSON ALGORITHM

Definitions

Our treatment is mostly based on [2] and [6].

Here: edges correspond to physical conduits. Oil or water pipes (of differing diameters), transportation networks, nutrient networks in ecology, etc. So for now: **arrows indicate direction of physical flow**.

Definition: An **FF network** N is a digraph with 1 leader (called **source** s) and 1 goose (called **sink** t) together with a flow satisfying **feasibility conditions**.

Definition: Every edge e has a **capacity** $c(e) \geq 0$ and a **flow** $f(e)$. The **value** $\text{val}(f)$ of the flow is the output at t.

Feasibility Conditions:

1. $f(e) = -f(-e) \quad ; \quad c(e) = -c(-e)$
2. $0 \leq f(e) \leq c(e) \quad (\text{or} \quad c(e) \leq f(e) \leq 0)$
3. At every vertex, except s and t: flow in = flow out.
4. flow into s = flow out of t. The VALUE of the flow.

This amounts to conservation of mass in the graph.

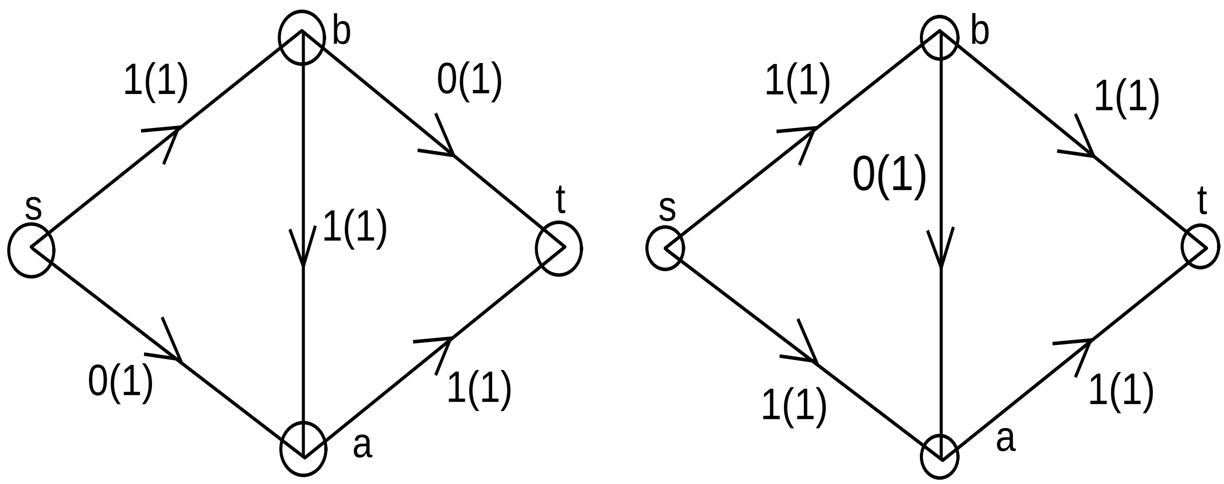
Remark: Could be any digraph with input in cabals and output in gaggles. For example, can create one **super source**.

Maximize Flow

Want to find the **maximum flow**.

Notation: $a(b)$ means flow(capacity) along arrow (see figure).

A **maximal flow**: cannot increase flow on any edge (left).
A **maximum flow**: exists no flow with greater value (right).



Definition: An **augmenting path** is a continuous path from s to t with **spare capacity**.

Example: In left figure, $\gamma = sabt$ has **spare capacity** of 1.
Feasibility onditions require:

$$f(sa) \in [0,1]\ ,\ f(ab) \in [-1,0]\ ,\ f(bt) \in [0,1]$$

Let f_γ be flow along γ with value 1. Flow of right figure:

$$f' := f + f_\gamma$$

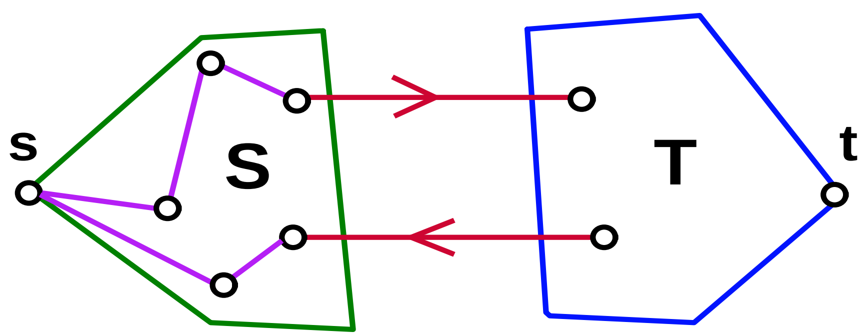
is a **feasible** flow with a higher value than f.

Max-Flow Min-Cut

Definition: An ***st*** **cut,** $[S,T]$**,** is a partition of the vertices into S containing s and T containing t. Its **capacity,** **cap**$[S,T]$, is the sum of the capacities of edges **from** S **to** T.

Theorem 5: flow is maximum $\iff$ no augm. path.

Sketch of Proof: To prove: no augm. path $\Rightarrow$ flow max.
S the set of vertices in augmenting **semi-paths** out of s (pink in the figure). T is its complement.
By defn of S, in the absence of an augm. path, we have:
1. $s \in S$ and $t \in T$.
2. $e \in [S,T] \Longrightarrow f(e) = c(e)$ and $e \in [T,S] \Longrightarrow f(e) = 0$.
Thus val(f)=cap$[S,T]$, and f must be **maximum**.



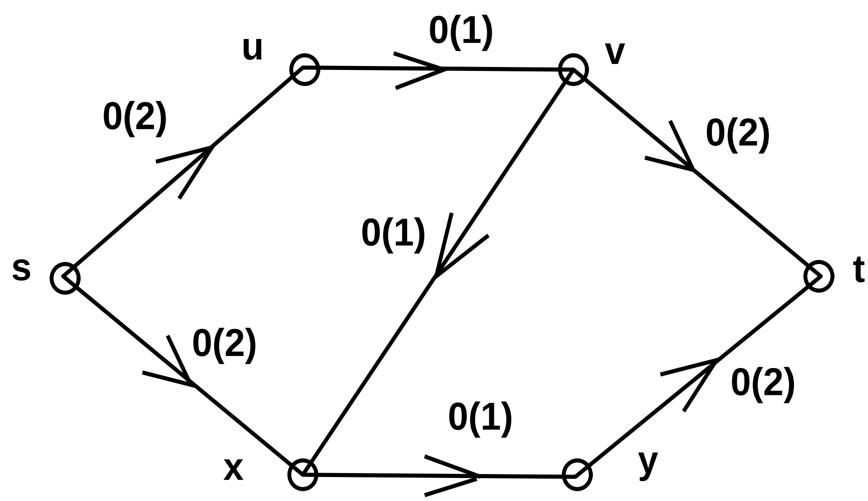
By mass conserv., no flow is greater than minimum of cap$[S,T]$.
A flow with val(f)=cap$[S,T]$ can be constructed. Thus:

Theorem 6 (Max-flow min-cut theorem, FF 1956):

$$\max_{\text{feas. flows}} \text{val}(f) = \min_{st \text{ cuts}} \text{cap}[S,T]$$

The Ford Fulkerson Algorithm

Definition: At every step of the algorithm, the set of vertices is partitioned into the following sets. S stands for searched, R stands for reached, and C, the complement of $S \cup R$.



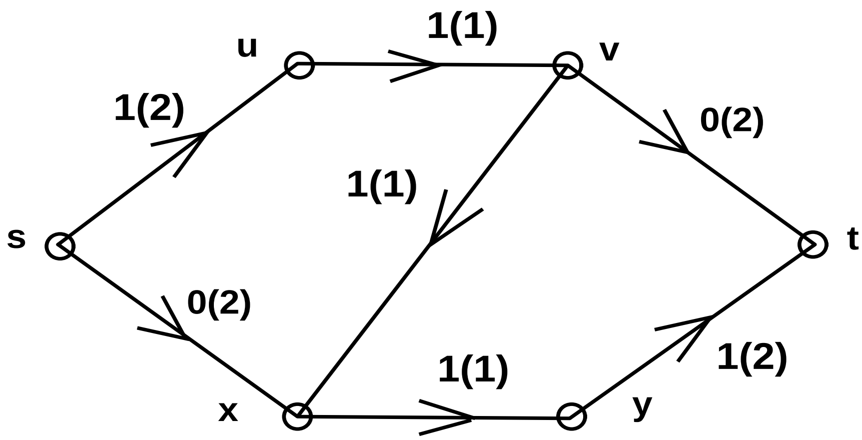
Steps of the algorithm:

step	S	R	*comment*
1.	$\emptyset$	s	start
2.	s	u, x	find spare cap. su
3.	s, u	v	find spare cap. suv
4.	s, u, v	x, t	find augm. path $suvt$
⋮	⋮	⋮	start again at s until no spare cap.

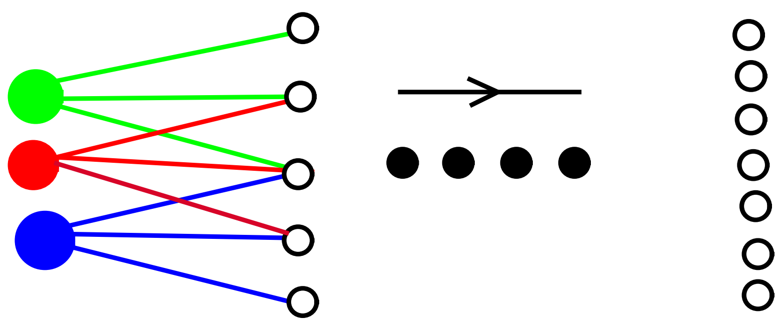
Remark. Note that when searching a vertex v, there is no strategy specified how to order edges incident to v. Improved formulations specify search strategy.

Ford Fulkerson Exercises

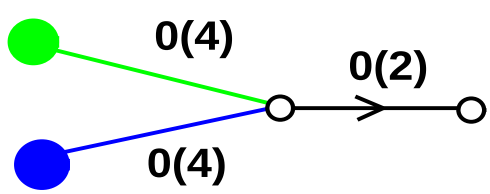
Exercise 12: Use the algorithm to find an augm. path.



Exercise 13 (FF for Artists): Given FF network with many sources and many sinks. Each sources inputs a specific color of paint. What is color mix of each output? (See below.)

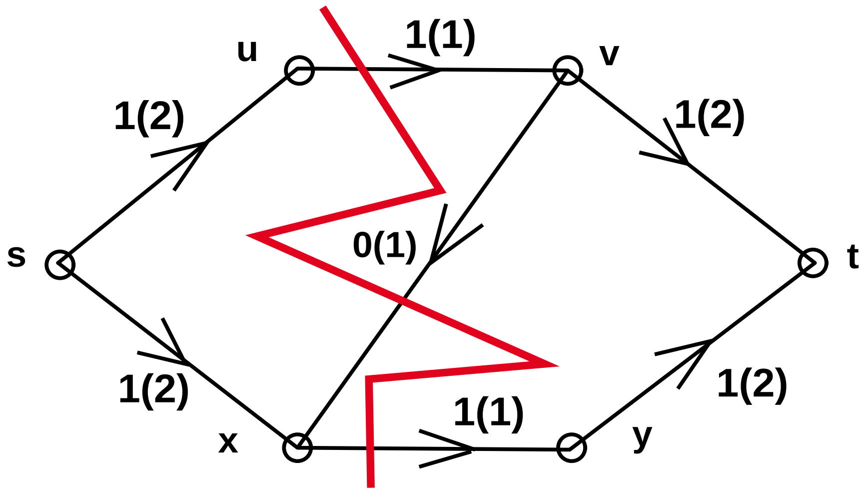


Comment. Answer is not unique as figure below shows. Use flow adjacency matrix of the flow computed by FF.



Remarks

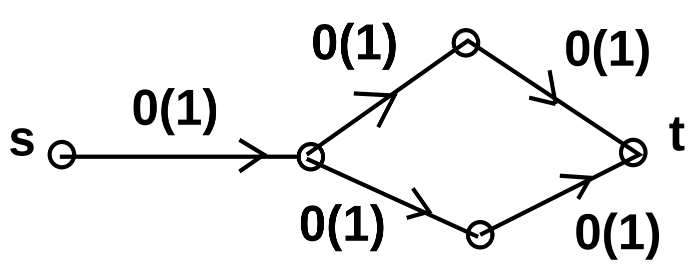
Example: In ex. 12, you should have found flow of value 2. There is cut of capacity 2. By Thm 5, max-flow=min-cut=2.



Corollary: If the capacities are **rational**, then FF converges to the max flow solution in **finitely many** steps.

Proof: Sufficient to do this for integers. Every augm. path has spare cap. at least 1. So FF terminates after finite steps.

Remark: The result of the FF algorithm depends of the search strategy. The outcome is <u>not</u> unique (see below).



.

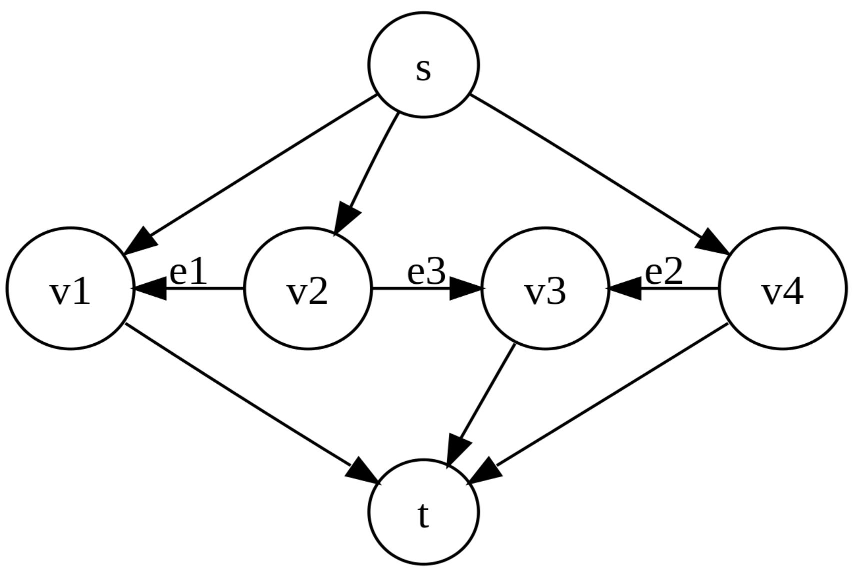
FORD FULKERSON COUNTER-EXAMPLE

The Smallest Counter-Example 1

In presence of irrational capacity, convergence can be beaten, but one has to be really clever to carefully craft a search strategy so that FF fails to converge to max flow [7].

All unmarked edges have capacity $m \geq 2$. Furthermore:

$$c(e_1) = c(e_3) = 1$$
$$c(e_2) = r := \frac{\sqrt{5}-1}{2} \approx 0.618$$
$$p_0 = (s, v_2, v_3, t)$$
$$p_1 = (s, v_4, v_3, v_2, v_1, t)$$
$$p_2 = (s, v_2, v_3, v_4, t)$$
$$p_3 = (s, v_1, v_2, v_3, t)$$



Exercise 14: Start with $f = 0$. Execute FF in such a way that the sequence of augm. paths is $(p_0, p_1, p_2, p_1, p_3, p_1, p_2, \cdots)$.

The Smallest Counter-Example 2

Exercise 15: Check the listed flow and spare capacities in the following table. (*Hint: use that* $r^2 = 1 - r$.)

step	augm. path	val(path)	sp. cap. e_1	sp. cap. e_2	sp. cap. e_3
0.	$\emptyset$	0	r^0	r^1	1
1.	p_0	r^0	r^0	r^1	0
2.	p_1	r^1	r^2	0	r^1
3.	p_2	r^1	r^2	r^1	0
4.	p_1	r^2	0	r^3	r^2
5.	p_3	r^2	r^2	r^3	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Exercise 15: Conclude that FF does not terminate, and that the value of the (total) flow converges to $1 + 2r \sum_{i \geq 0} r^i = r^{-3} \approx 4.24$.

Exercise 16: Exhibit a cut and a flow of value $2m + 1 \geq 5$.

References

[1] J. J. P. Veerman, E. Kummel, *Diffusion and Consensus on Weakly Connected Directed Graphs*, **Linear Algebra and Its Applications**, accepted, 2019.

[2] J. A. Bondy, U. S. R. Murty, *Graph Theory with Applications*, **North-Holland**, New York, 1976.

[3] M. Newman, *Networks*, 2nd edition, **Oxford University Press**, Oxford, 2018.

[4] M. A. Nowak, *Evolutionary Dynamics, Exploring the Equations of Life*, **The Belknap Press of Harvard Press**, Cambridge (MA), 2006.

[5] S. Sternberg, *Dynamical Systems*, **Dover Publications**, Mineola, NY, 2010, revised edition 2013.

[6] D. B. West, *Introduction to Graph Theory*, 2-nd Edition. Prentice-Hall, Upper SAddle River, NJ, 2001.

[7] U. Zwick, *The smallest networks on which the Ford-Fulkerson maximum flow procedure may fail to terminate*, **Theor. Comp. Sc** 148(1), 165–170, 1995.