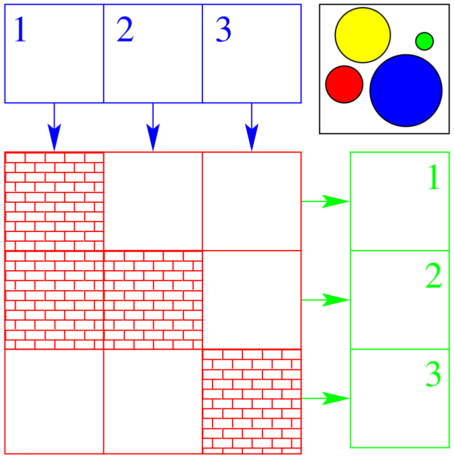


$$\boxed{\pi = (\alpha S + (1-\alpha)T)\pi}$$

Goozzles, en de Wiskunde die Google groot maakte

Jan Brandts

Korteweg-de Vries Instituut voor Wiskunde
Universiteit van Amsterdam

Nationale Wiskunde Dagen, zaterdag 2 februari 2008.

Even voorstellen ...?

- $a^2 + b^2 = c^2$
- $x = (-b \pm \sqrt{b^2 - 4ac})/2a$
- $\sin(2x) = 2\sin(x)\cos(x)$
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Even voorstellen ...?

- $a^2 + b^2 = c^2$
- $x = (-b \pm \sqrt{b^2 - 4ac})/2a$
- $\sin(2x) = 2\sin(x)\cos(x)$
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking

(Page & Brin, 1998)

Even voorstellen ...?

- π rijtje van 10 miljard "rapportcijfers"
- $x = (-b \pm \sqrt{b^2 - 4ac})/2a$
- $\sin(2x) = 2\sin(x)\cos(x)$
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1 - \alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Even voorstellen ...?

- π rijtje van 10 miljard "rapportcijfers"
- iedere maand (in 3 dagen) uitgerekend
- $\sin(2x) = 2\sin(x)\cos(x)$
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking

(Page & Brin, 1998)

Even voorstellen ...?

- π rijtje van 10 miljard “rapportcijfers”
- iedere maand (in 3 dagen) uitgerekend
- α is een (geheim) mix-getal
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1 - \alpha)E)}$$

Google’s PageRank vergelijking

(Page & Brin, 1998)

Even voorstellen ...?

- π rijtje van 10 miljard “rapportcijfers”
- iedere maand (in 3 dagen) uitgerekend
- α is een (geheim) mix-getal
- E is de teleportatie-matrix
- $\mathcal{O} = \pi r^2$

$$\boxed{\pi = \pi(\alpha S + (1 - \alpha)E)}$$

Google’s PageRank vergelijking
(Page & Brin, 1998)

Even voorstellen ...?

- π rijtje van 10 miljard "rapportcijfers"
- iedere maand (in 3 dagen) uitgerekend
- α is een (geheim) mix-getal
- E is de teleportatie-matrix
- S is de web-matrix

$$\boxed{\pi = \pi(\alpha S + (1 - \alpha) E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1-\alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1-\alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1 - \alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1-\alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1 - \alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Doel van deze workshop

Laten zien hoe de PageRank vergelijking

$$\boxed{\pi = (\alpha S + (1-\alpha)T)\pi}$$

wordt

- opgesteld aan de hand van heuristieken
- aangepast met behulp van wiskundige eigenschappen
- opgelost aan de hand van kabouterraadsels en Goozzles
- opgelost in de realiteit

Voor de meeste items is bijna geen wiskundige voorkennis vereist!

Inhoudsopgave

- De Google Parabel: Er was eens . . .
- Goozzles
- Dekpunt-iteraties
- Bewijzen

Met dank aan:

Michal Křížek, Akademie Věd Ceské Republiky, Praha

Inhoudsopgave

- De Google Parabel: Er was eens ...
- Goozzles
- Dekpunt-iteraties
- Bewijzen

Met dank aan:

Michal Křížek, Akademie Věd Ceské Republiky, Praha

Inhoudsopgave

- De Google Parabel: Er was eens ...
- Goozzles
- Dekpunt-iteraties
- Bewijzen

Met dank aan:

Ricardo da Silva, KdV Instituut, Universiteit van Amsterdam

Inhoudsopgave

- De Google Parabel: Er was eens ...
- Goozzles
- Dekpunt-iteraties
- Bewijzen

Met dank aan:

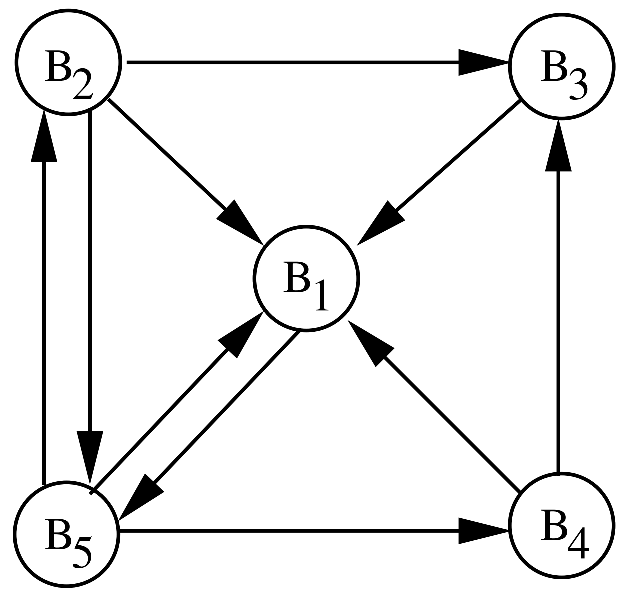
Montgomery Scott, Starfleet Academy, Starship USS Enterprise

De Google Parabel: Er was eens ...

- een goede fee (*)
- met héééł veel knikkers
- een aantal onbaatzuchtige kabouters
- ... die ieder hun vrienden uitkiezen
- ... maar dit is niet noodzakelijk wederzijds

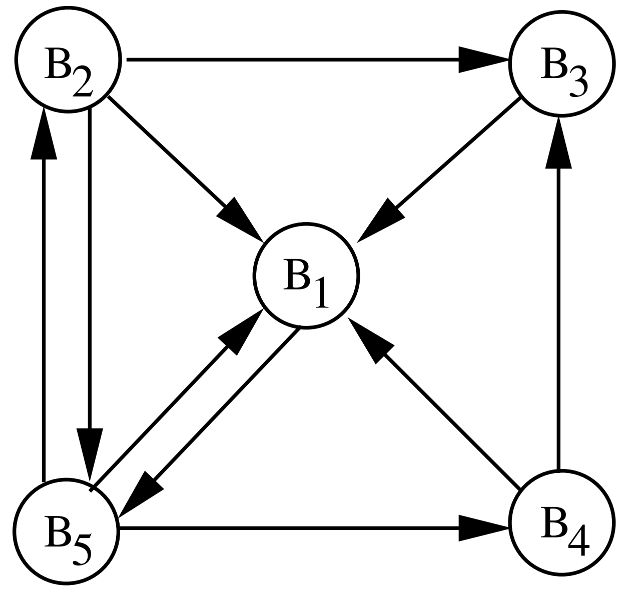
(*) Zie: Google Pictures, ”sexy fee”
(als de kinderen naar bed zijn)

Vijf kabouters die hun vrienden uitkiezen



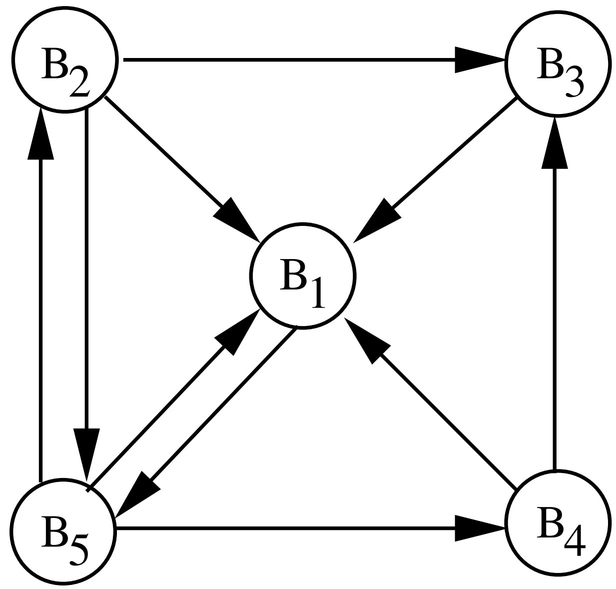
Een kabouter zal knikkers die hij van de fee krijgt, gelijkelijk verdelen onder zijn vrienden

Vijf kabouters die hun vrienden uitkiezen



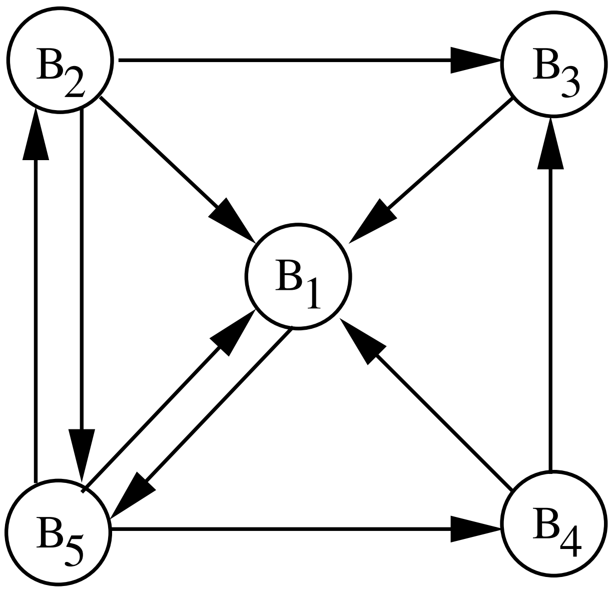
Een kabouter zal knikkers die hij van de fee krijgt, gelijkelijk verdelen onder zijn vrienden

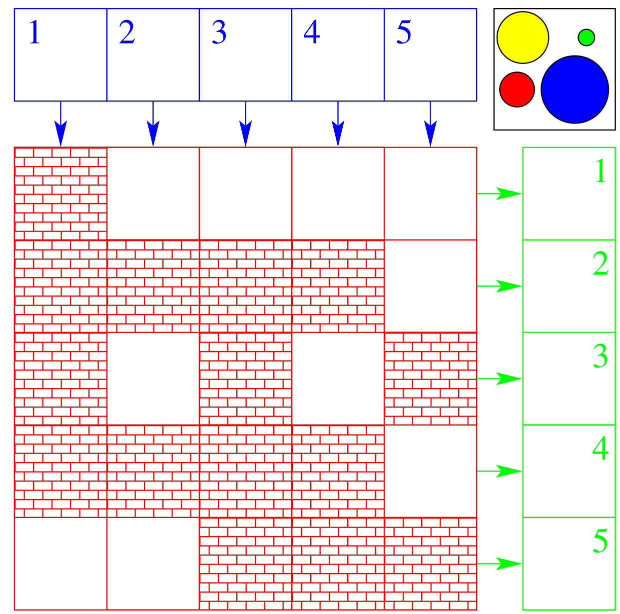
De fee kan zodanig knikkers uitdelen ...



... dat na de weggeef-actie van de kabouters, dezelfde knikker-verdeling resteert. Probeer zelf! (minimaal 51 knikkers nodig)

Goozzles

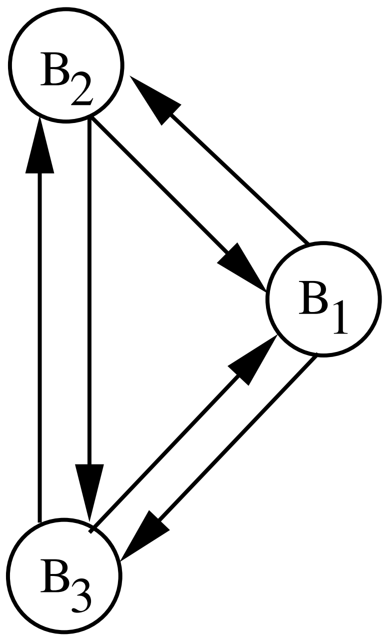


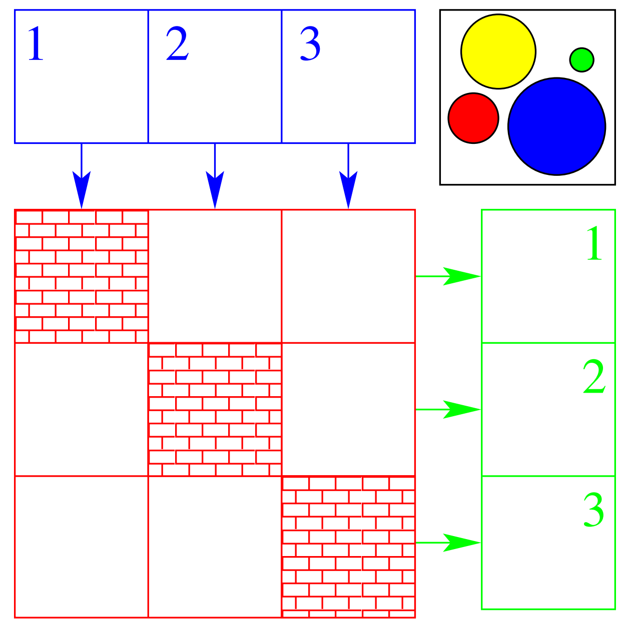


Een andere grafische voorstelling is deze 5 x 5 verzamelbak

- knikkers weggeven gaat verticaal
- knikkers incasseren gaat horizontaal

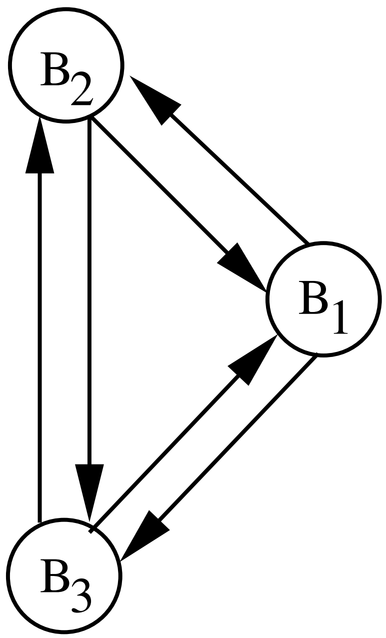
Een gemakkelijke Goozzle (1)

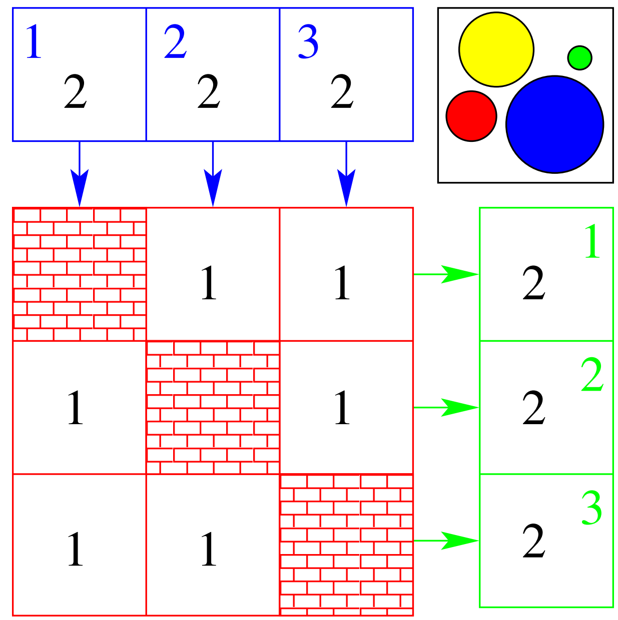




Geef de oplossing met het kleinste (positieve) aantal knikkers

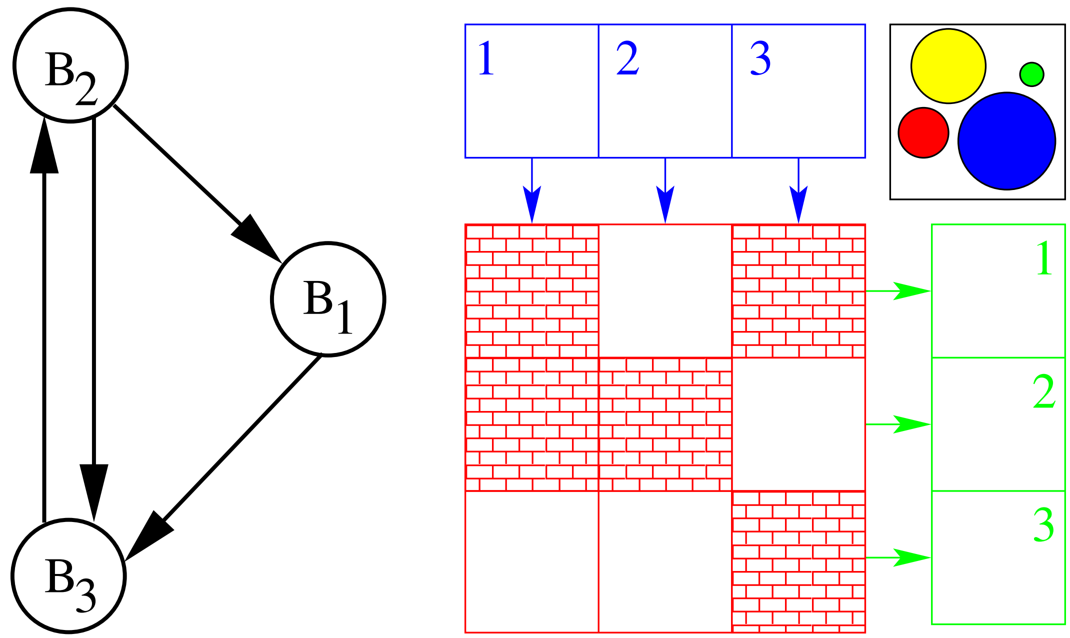
Een gemakkelijke Goozzle (1)





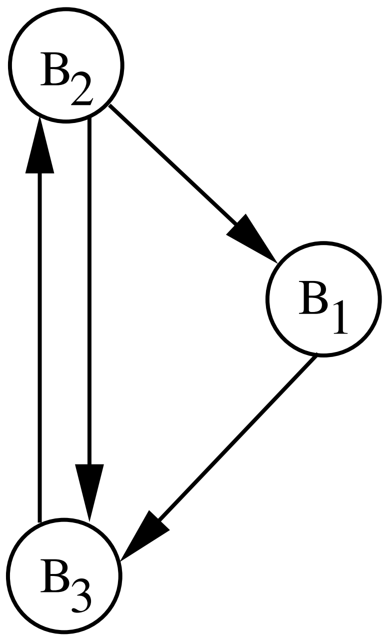
Geef de oplossing met het kleinste (positieve) aantal knikkers

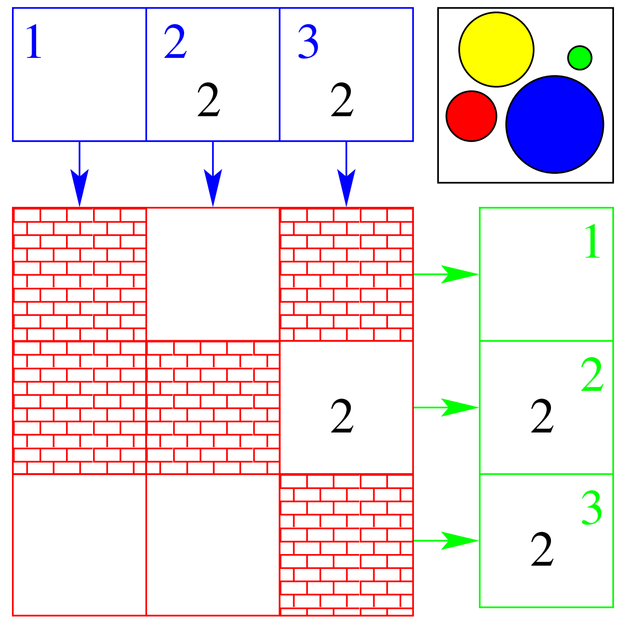
Een gemakkelijke Goozzle (2)



Geef de oplossing met het kleinste (positieve) aantal knikkers

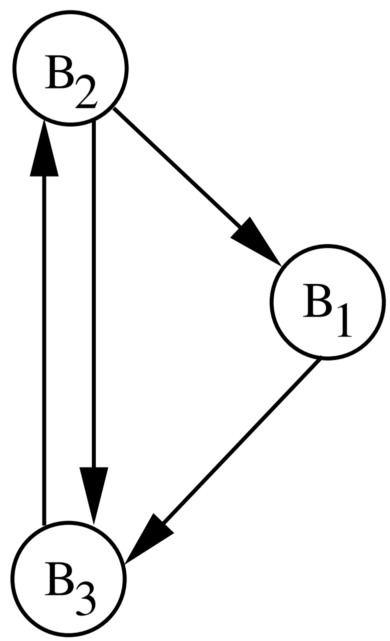
Een gemakkelijke Goozzle (2)

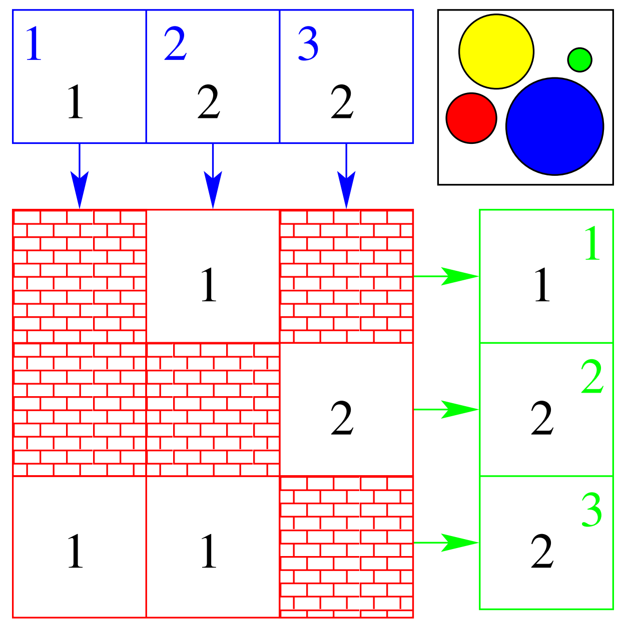




Geef de oplossing met het kleinste (positieve) aantal knikkers

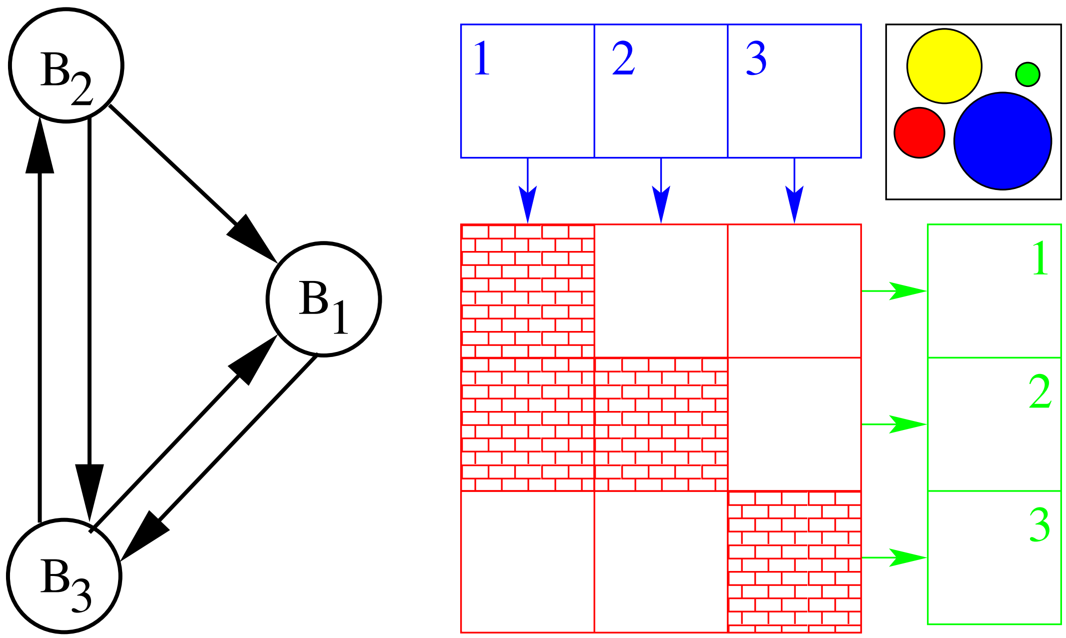
Een gemakkelijke Goozzle (2)





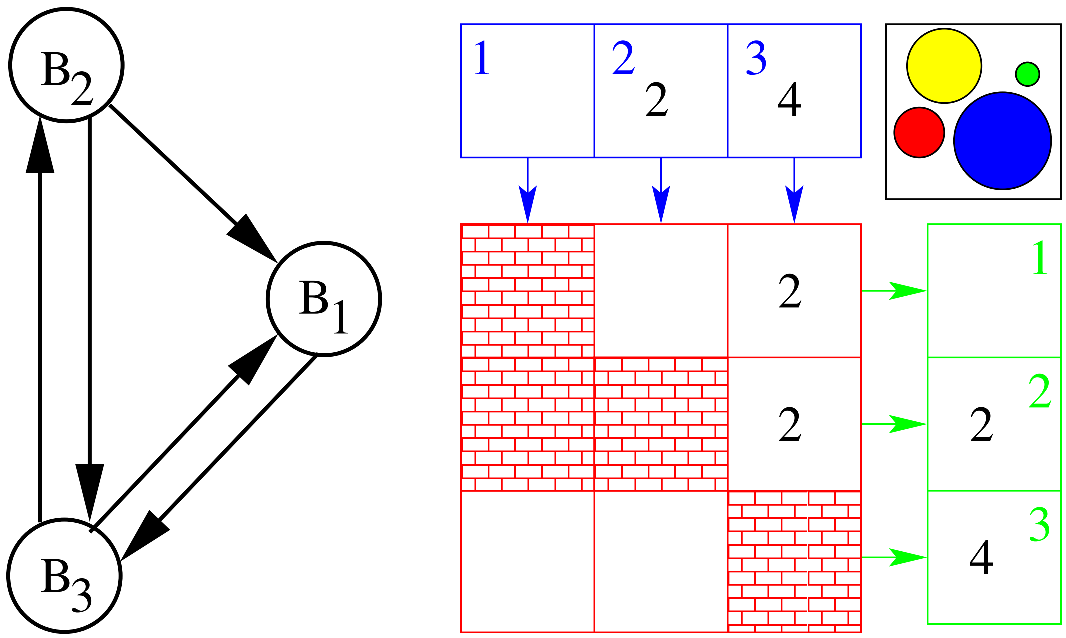
Geef de oplossing met het kleinste (positieve) aantal knikkers

Een gemakkelijke Goozzle (3)



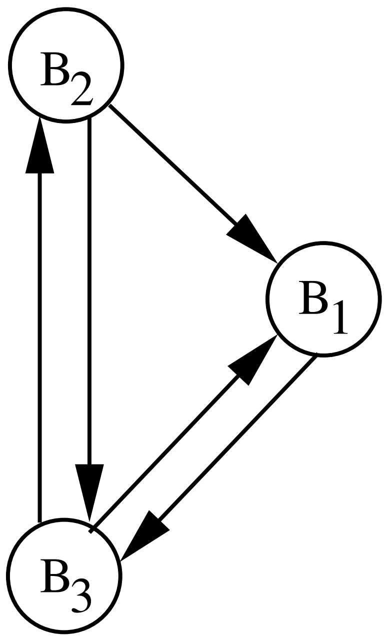
Geef de oplossing met het kleinste (positieve) aantal knikkers

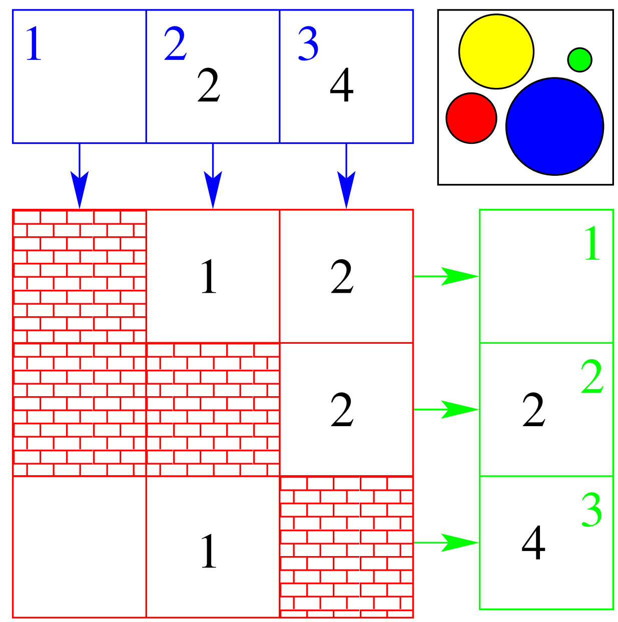
Een gemakkelijke Goozzle (3)



Geef de oplossing met het kleinste (positieve) aantal knikkers

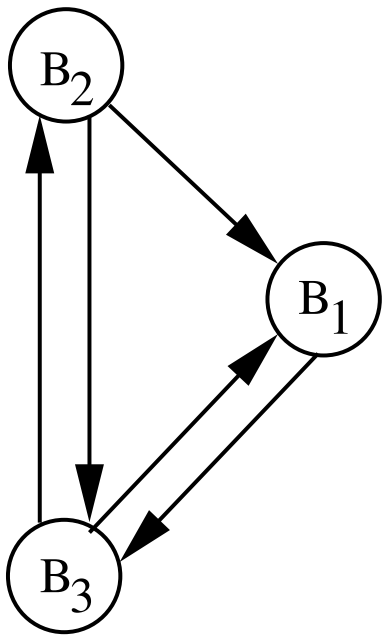
Een gemakkelijke Goozzle (3)

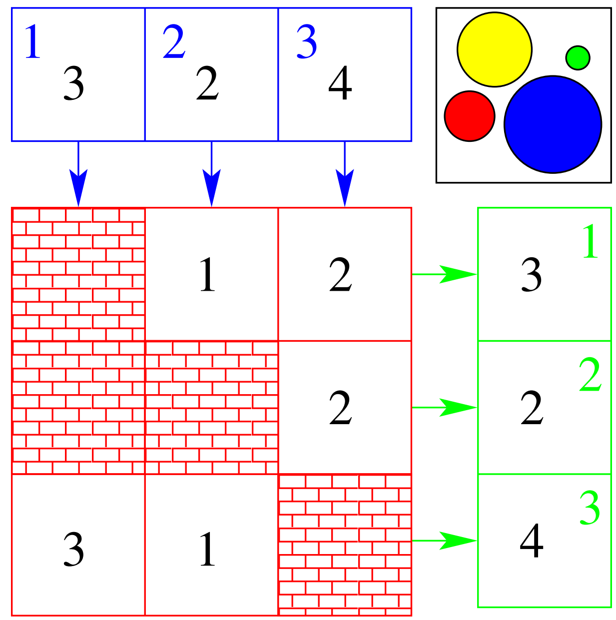




Geef de oplossing met het kleinste (positieve) aantal knikkers

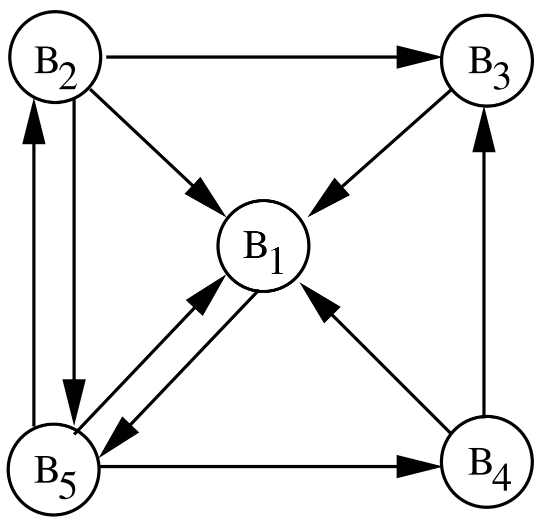
Een gemakkelijke Goozzle (3)

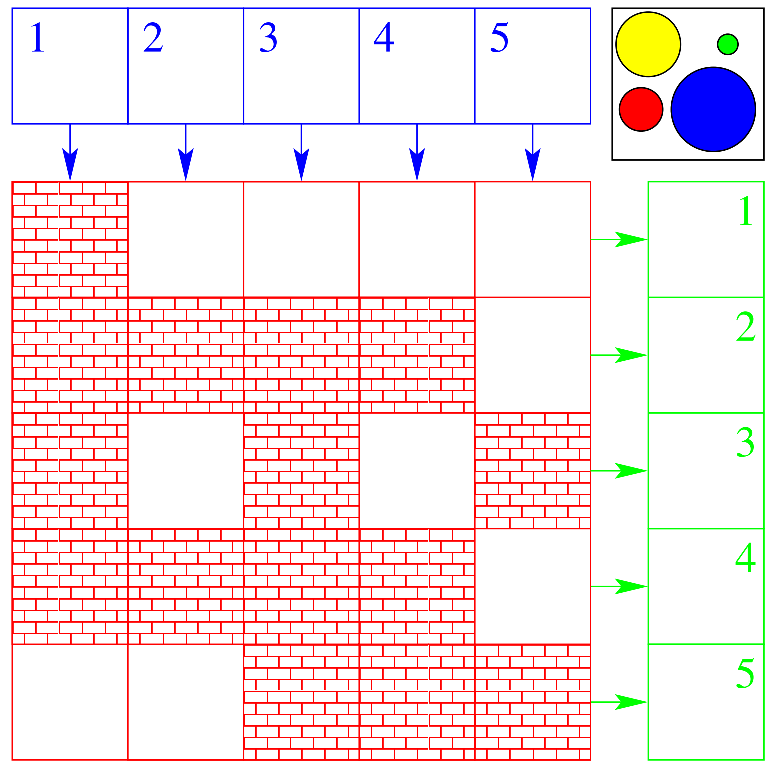




Geef de oplossing met het kleinste (positieve) aantal knikkers

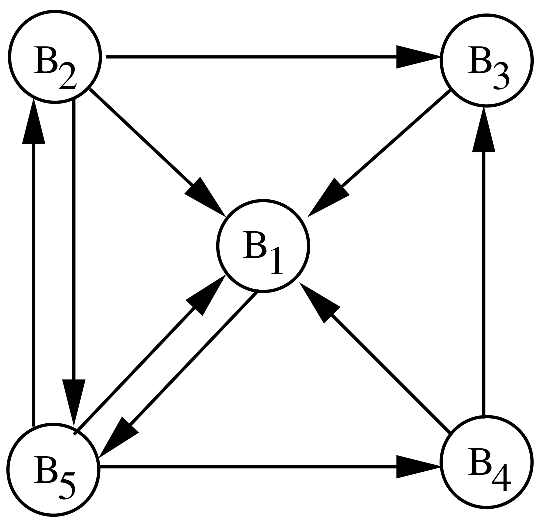
Een gemakkelijke Goozzle (4)

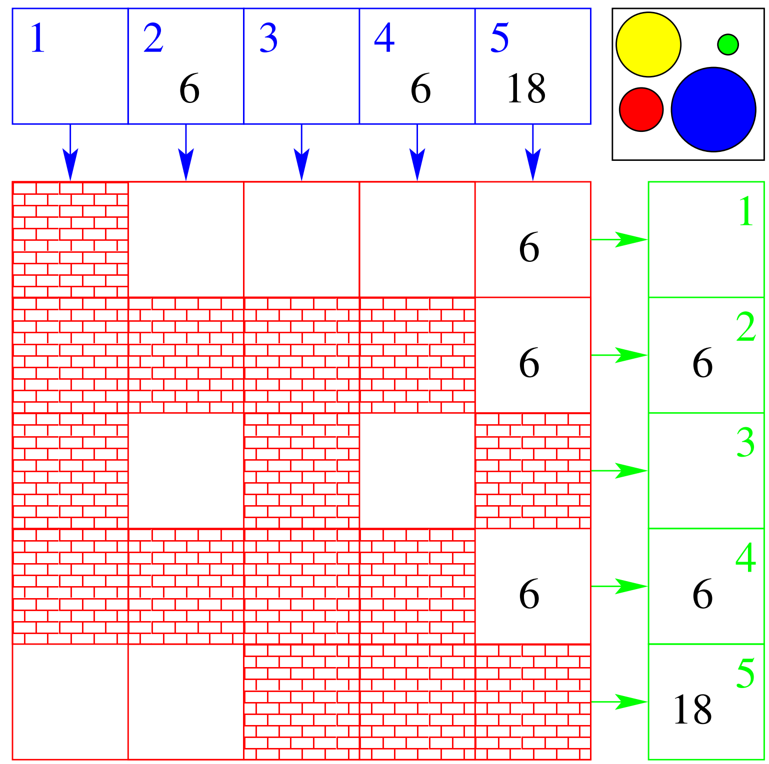




Geef de oplossing met het kleinste (positieve) aantal knikkers

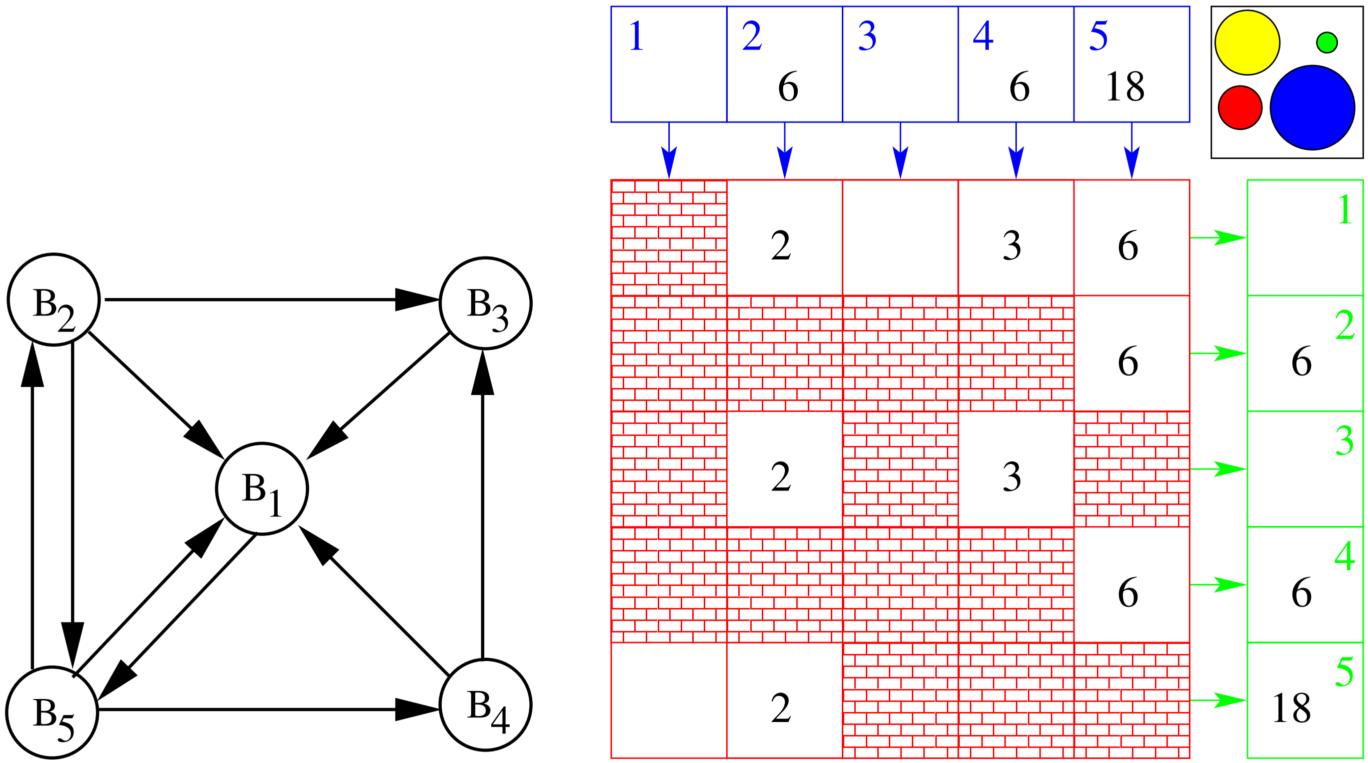
Een gemakkelijke Goozzle (4)





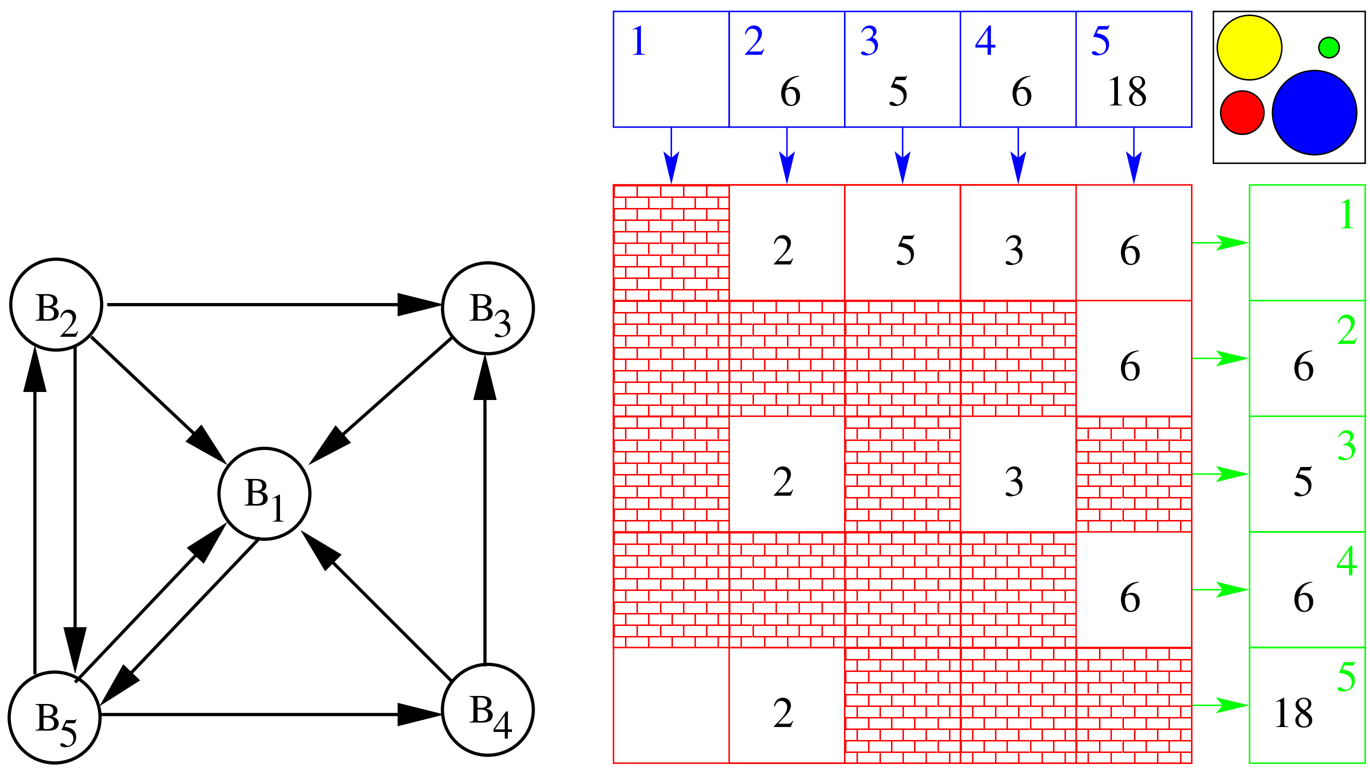
Geef de oplossing met het kleinste (positieve) aantal knikkers

Een gemakkelijke Goozzle (4)



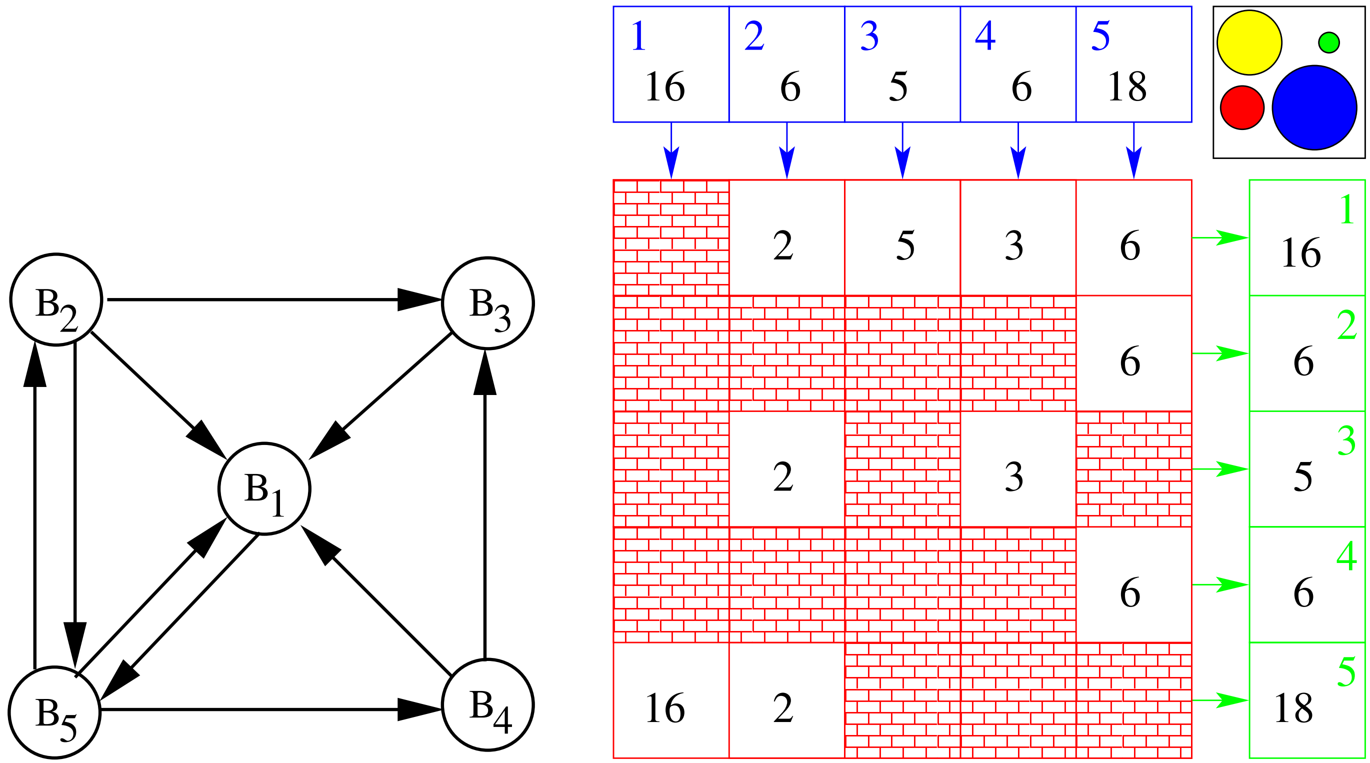
Geef de oplossing met het kleinste (positieve) aantal knikkers

Een gemakkelijke Goozzle (4)



Geef de oplossing met het kleinste (positieve) aantal knikkers

Een gemakkelijke Goozzle (4)



Geef de oplossing met het kleinste (positieve) aantal knikkers

Het eerste PageRank model van Brin en Page

Het model van Page en Brin implementeert de volgende

HEURISTIEK(EN)

- Een blz is belangrijk als belangrijke blz ernaar verwijzen ...
- ... die NIET ook naar heel veel andere bladzijden verwijzen

Het eerste PageRank model van Brin en Page

Het model van Page en Brin implementeert de volgende

HEURISTIEK(EN)

- Een blz is belangrijk als belangrijke blz ernaar verwijzen ...
- ... die NIET ook naar heel veel andere bladzijden verwijzen

Het eerste PageRank model van Brin en Page

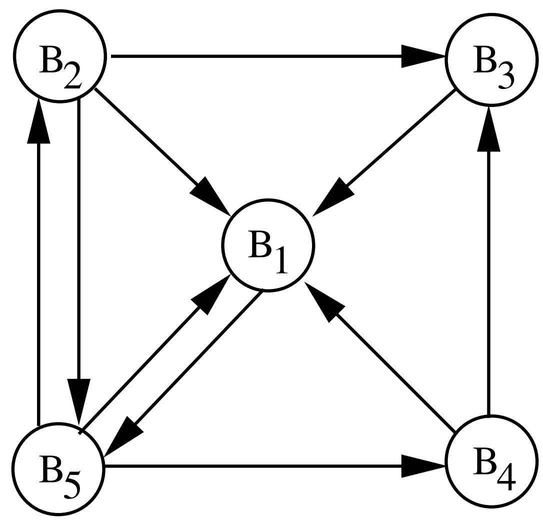
Larry Page  Sergey Brin

Het eerste model van Page en Brin implementeert de volgende

$\boxed{\text{HEURISTIEK(EN)}}$

- Een blz is belangrijk als belangrijke blz ernaar verwijzen ...
- ... die NIET ook naar heel veel andere bladzijden verwijzen

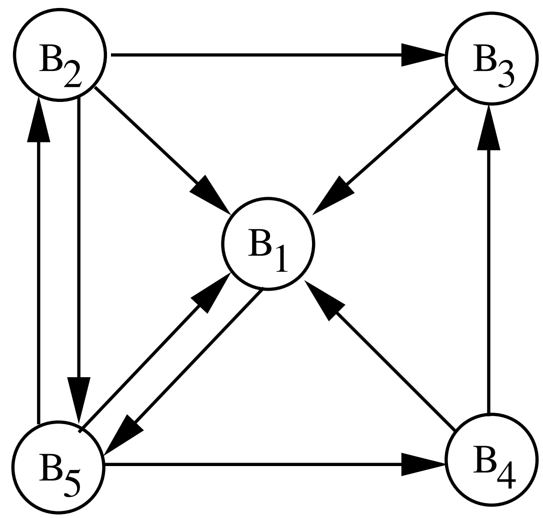
Formele definitie van PageRank (1)



Schrijf P_j voor de PageRank van bladzijde B_j.

Als bladzijde B_i naar $L_i > 0$ bladzijdes verwijst, waaronder B_j, draagt B_i een hoeveelheid P_i/L_i bij aan de PageRank van B_j

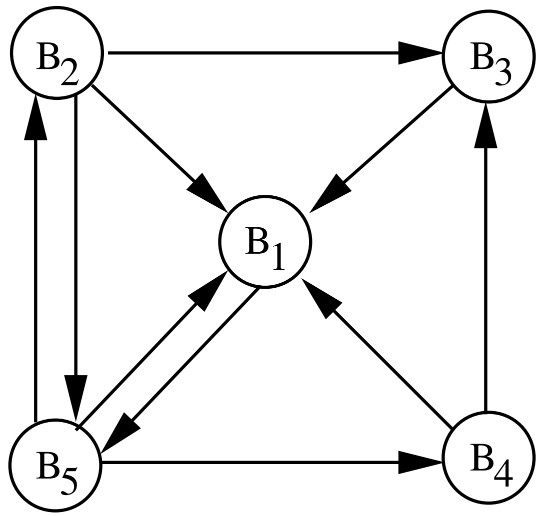
Formele definitie van PageRank (2)



$$P_1 = \frac{P_2}{3} + \frac{P_3}{1} + \frac{P_4}{2} + \frac{P_5}{3}$$

$$P_2 = \frac{P_5}{3}, \quad P_3 = \frac{P_2}{3} + \frac{P_4}{2}, \quad P_4 = \frac{P_5}{3}, \quad P_5 = \frac{P_1}{1} + \frac{P_2}{3}$$

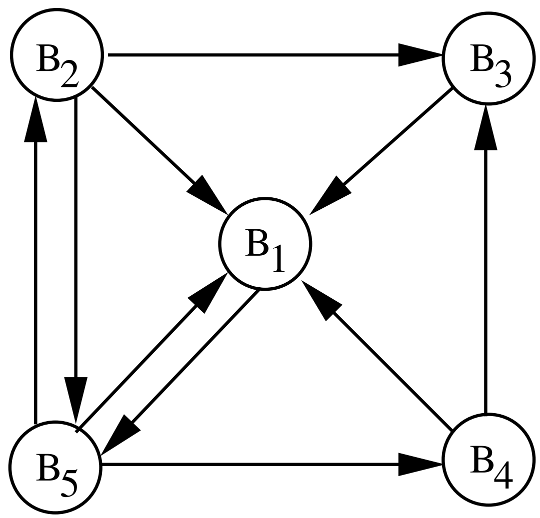
Formele definitie van PageRank (2)



$$P_1 = \frac{P_2}{3} + \frac{P_3}{1} + \frac{P_4}{2} + \frac{P_5}{3}$$

$$P_2 = \frac{P_5}{3}, \quad P_3 = \frac{P_2}{3} + \frac{P_4}{2}, \quad P_4 = \frac{P_5}{3}, \quad P_5 = \frac{P_1}{1} + \frac{P_2}{3}$$

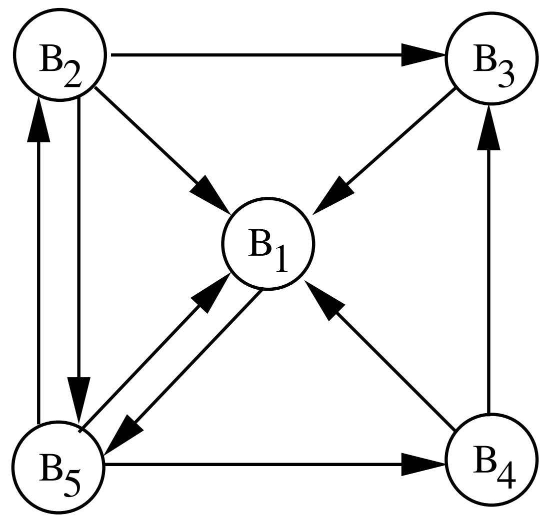
Formele definitie van PageRank (2)



$$P_1 = \frac{P_2}{3} + \frac{P_3}{1} + \frac{P_4}{2} + \frac{P_5}{3}$$

$$P_2 = \frac{P_5}{3}, \quad P_3 = \frac{P_2}{3} + \frac{P_4}{2}, \quad P_4 = \frac{P_5}{3}, \quad P_5 = \frac{P_1}{1} + \frac{P_2}{3}$$

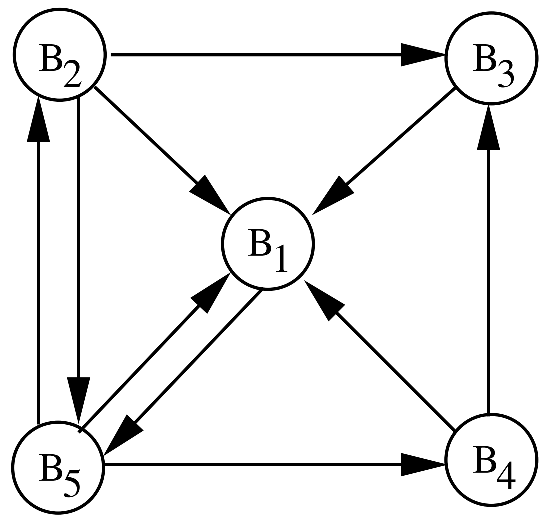
Formele definitie van PageRank (2)



$$P_1 = \frac{P_2}{3} + \frac{P_3}{1} + \frac{P_4}{2} + \frac{P_5}{3}$$

$$P_2 = \frac{P_5}{3}, \quad P_3 = \frac{P_2}{3} + \frac{P_4}{2}, \quad P_4 = \frac{P_5}{3}, \quad P_5 = \frac{P_1}{1} + \frac{P_2}{3}$$

Formele definitie van PageRank (2)



$$P_1 = \frac{P_2}{3} + \frac{P_3}{1} + \frac{P_4}{2} + \frac{P_5}{3}$$

$$P_2 = \frac{P_5}{3}, \quad P_3 = \frac{P_2}{3} + \frac{P_4}{2}, \quad P_4 = \frac{P_5}{3}, \quad P_5 = \frac{P_1}{1} + \frac{P_2}{3}$$

Formele definitie van PageRank (3)

Dit alles leidt tot de volgende matrix - vector vergelijking:

$$H\pi = \pi \;\Leftrightarrow\; \begin{bmatrix} 0 & 1/3 & 1 & 1/2 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 \\ 1 & 1/3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

In termen van Lineaire Algebra staat hier dus eigenlijk:

- De matrix H heeft een eigenwaarde $\lambda = 1$
- De bijbehorende eigenvector π is de PageRank vector

Legitieme vraag: heeft H voor ieder web een eigenwaarde één?

Formele definitie van PageRank (3)

Dit alles leidt tot de volgende matrix - vector vergelijking:

$$H\pi = \pi \;\; \Leftrightarrow \begin{bmatrix} 0 & 1/3 & 1 & 1/2 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 \\ 1 & 1/3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

In termen van Lineaire Algebra staat hier dus eigenlijk:

- De matrix H heeft een eigenwaarde $\lambda = 1$
- De bijbehorende eigenvector π is de PageRank vector

Legitieme vraag: heeft H voor ieder web een eigenwaarde één?

Formele definitie van PageRank (3)

Dit alles leidt tot de volgende matrix - vector vergelijking:

$$H\pi = \pi \ \Leftrightarrow \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

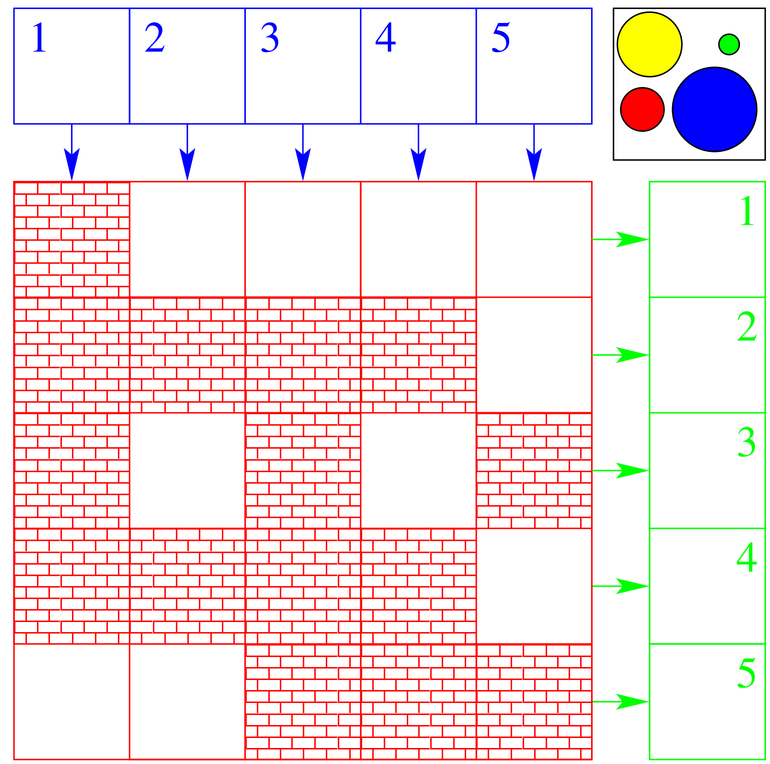
In termen van Lineaire Algebra staat hier dus eigenlijk:

- De matrix H heeft een eigenwaarde $\lambda = 1$
- De bijbehorende eigenvector π is de PageRank vector

Legitieme vraag: heeft H voor ieder web een eigenwaarde één?

Formele definitie van PageRank (3)

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



Formele definitie van PageRank (3)

Dit alles leidt tot de volgende matrix - vector vergelijking:

$$H\pi = \pi \;\; \Leftrightarrow \begin{bmatrix} 0 & 1/3 & 1 & 1/2 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 \\ 1 & 1/3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

In termen van Lineaire Algebra staat hier dus eigenlijk:

- De matrix H heeft een eigenwaarde $\lambda = 1$
- De bijbehorende eigenvector π is de PageRank vector

Legitieme vraag: heeft H voor ieder web een eigenwaarde één?

Formele definitie van PageRank (3)

Dit alles leidt tot de volgende matrix - vector vergelijking:

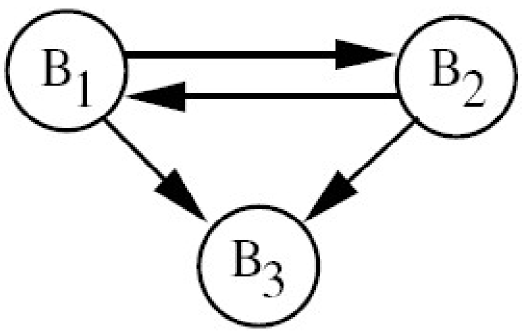
$$H\pi = \pi \;\Leftrightarrow\; \begin{bmatrix} 0 & 1/3 & 1 & 1/2 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 \\ 1 & 1/3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

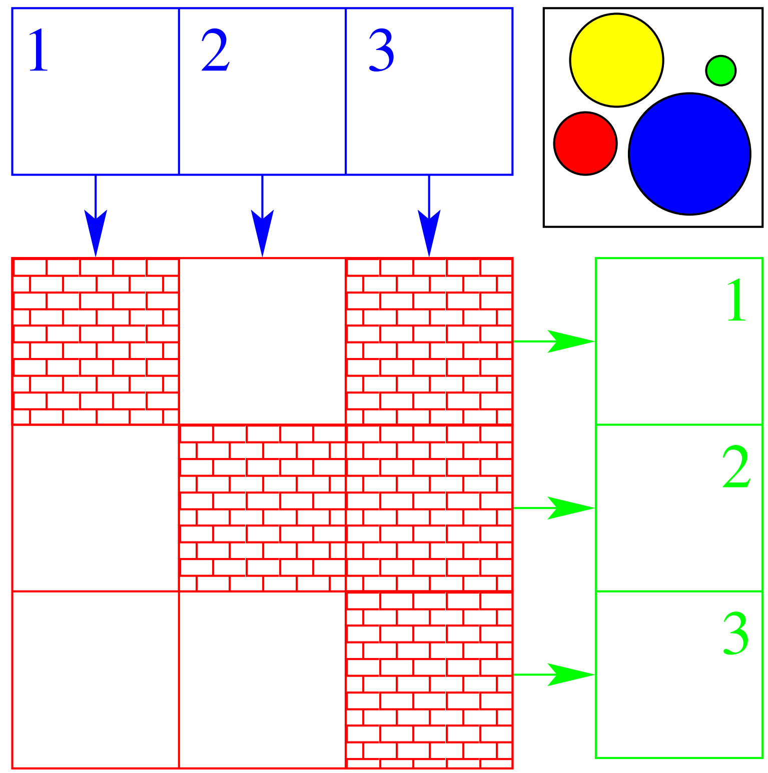
In termen van Lineaire Algebra staat hier dus eigenlijk:

- De matrix H heeft een eigenwaarde $\lambda = 1$
- De bijbehorende eigenvector π is de PageRank vector

Legitieme vraag: heeft H voor ieder web een eigenwaarde één?

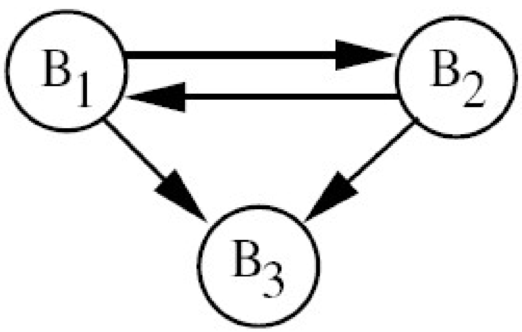
Een gemakkelijke Goozzle (5)

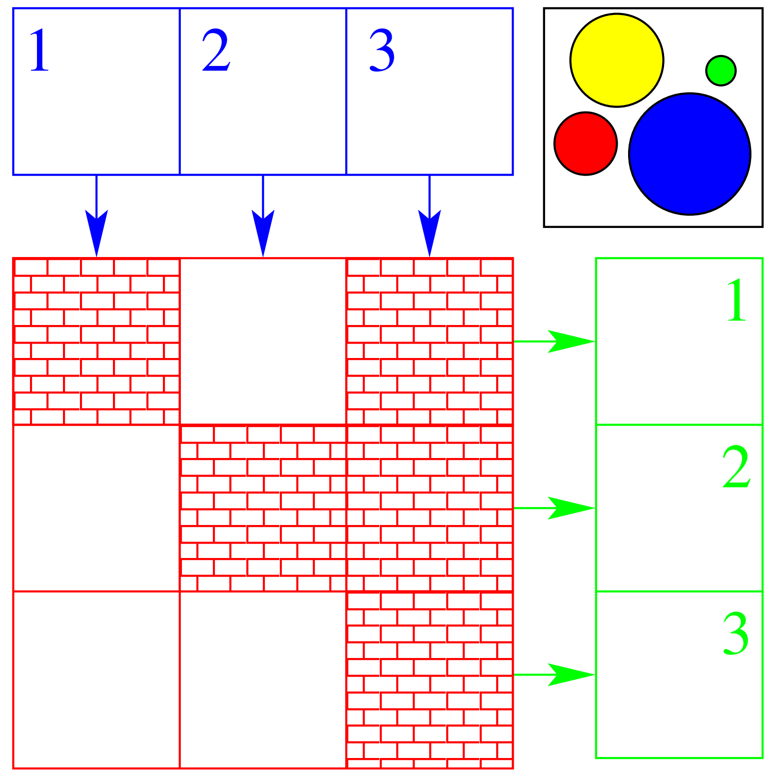




Geef de oplossing met het kleinste (positieve) aantal knikkers

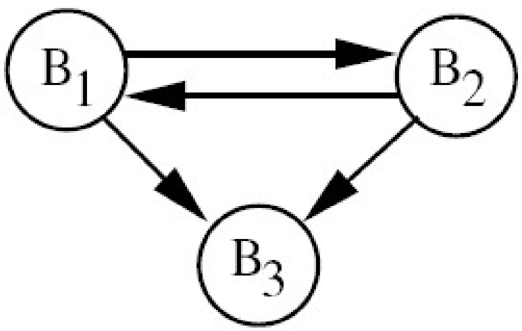
Een onoplosbare Goozzle (5)

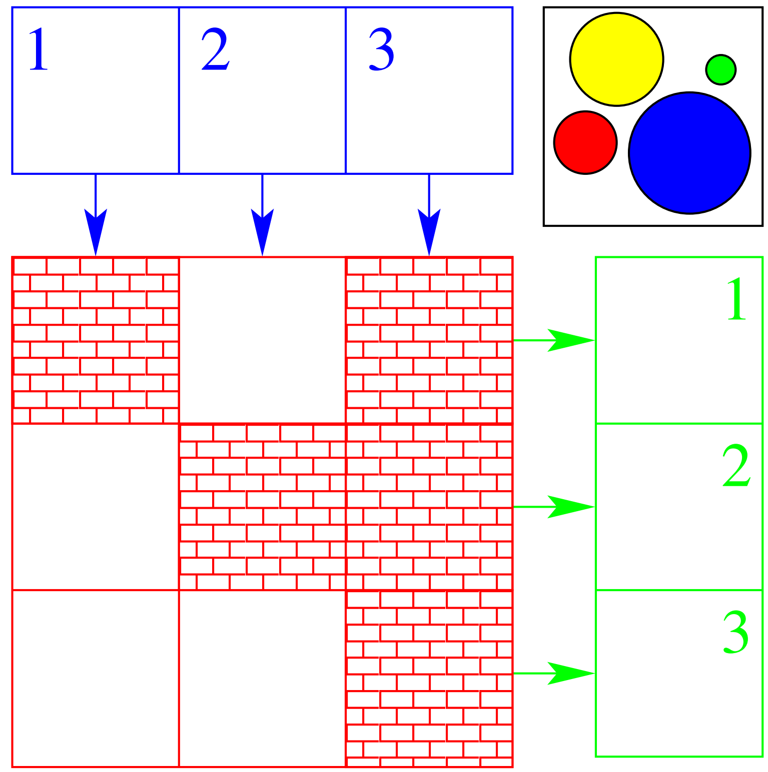




Er is alleen de nietszeggende oplossing $P_1 = P_2 = P_3 = 0$

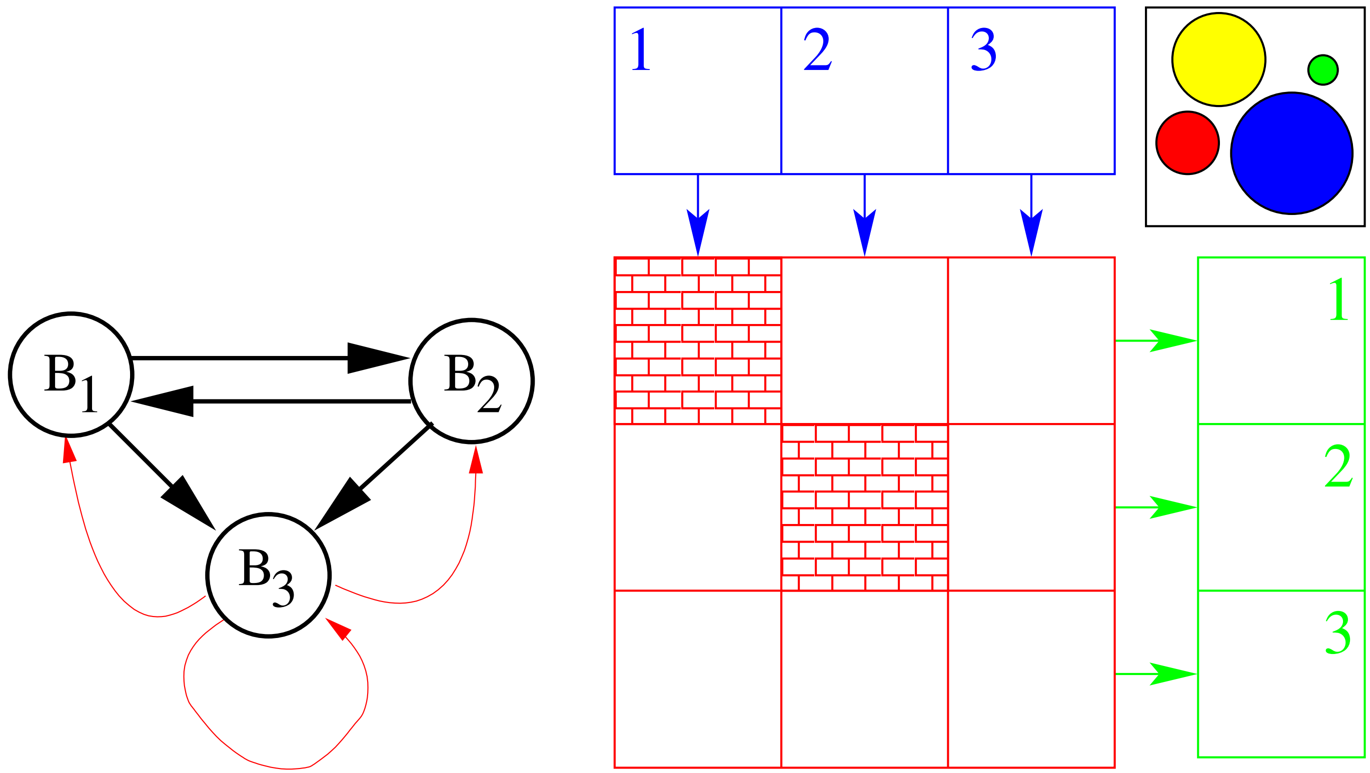
Een onoplosbare Goozzle (5)





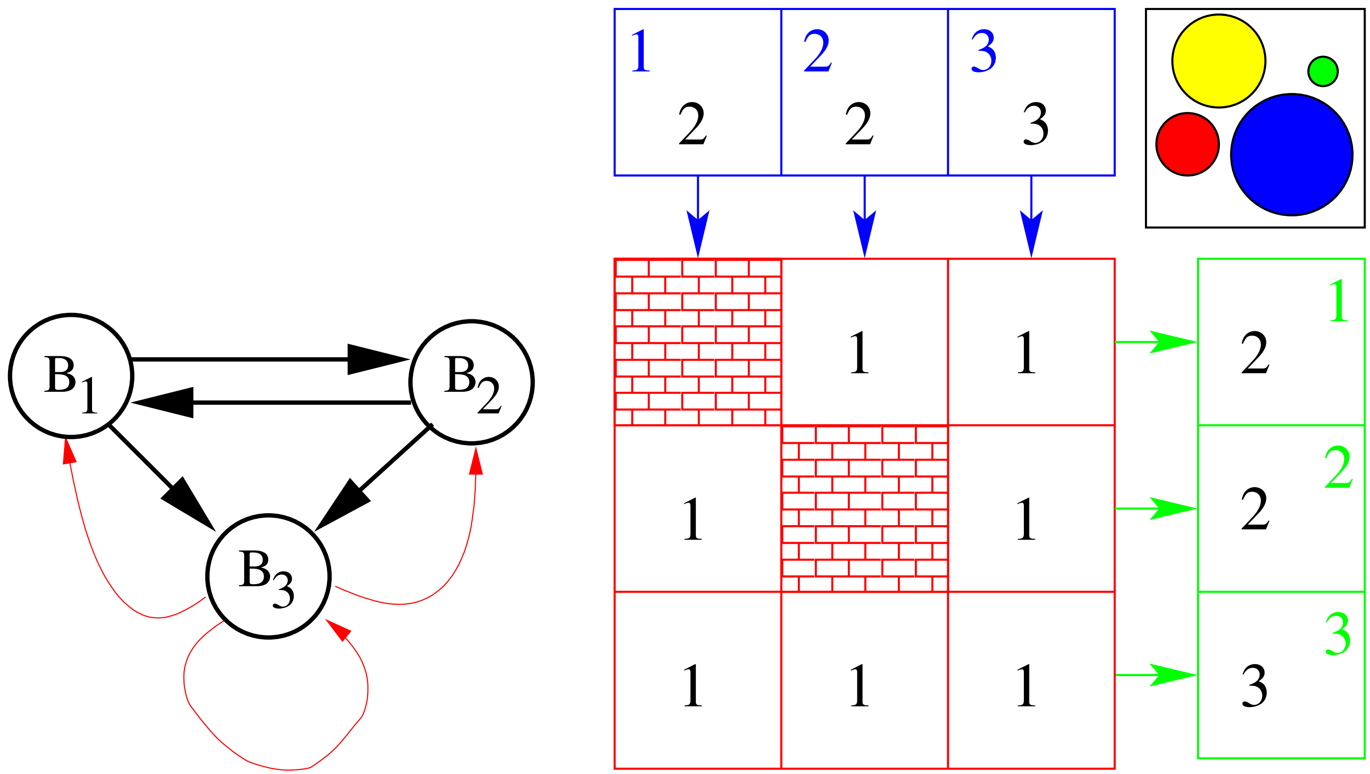
$$P_1 = \frac{1}{2}P_2 \text{ en } P_2 = \frac{1}{2}P_1 \text{ dus } P_1 = P_2 = 0 \text{ dus } P_3 = \frac{1}{2}(P_1 + P_2) = 0$$

Oplossing middels teleportatie



Extra, onpartijdige rode pijlen maken doorstroom mogelijk

Oplossing middels teleportatie



En hier is de oplossing met het kleinste aantal knikkers

Oplossing middels teleportatie

Teleportatie vervangt een verstopte Goozzle kolom door een kolom met $1/n$-nen:

$$H = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \quad \rightarrow \quad S = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$$

Merk op dat iedere kolom van S optelt tot één, dus

$$S^T e = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = e$$

Dus $\det(S^T - I) = 0 = \det(S - I)$ dus is er een $\pi \neq 0$ met $S\pi = \pi$

Oplossing middels teleportatie

Teleportatie vervangt een verstopte Goozzle kolom door een kolom met $1/n$-nen:

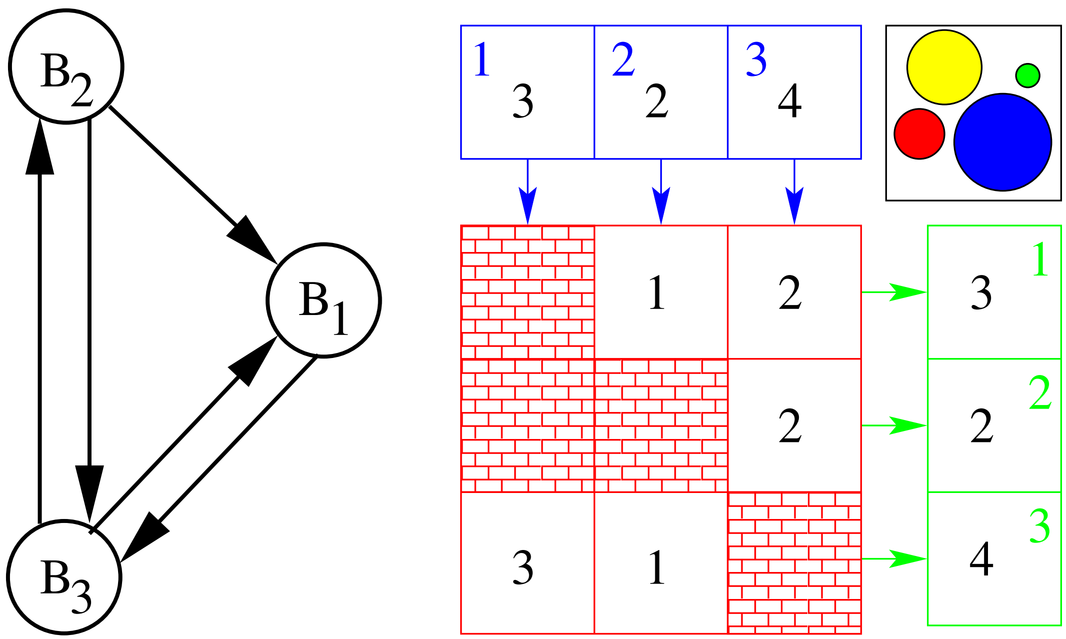
$$H = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \quad \longrightarrow \quad S = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$$

Merk op dat iedere kolom van S en rij van S^T optelt tot één, dus

$$S^T e = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = e$$

Dus $\det(S^T - I) = 0 = \det(S - I)$ dus is er een $\pi \neq 0$ met $S\pi = \pi$

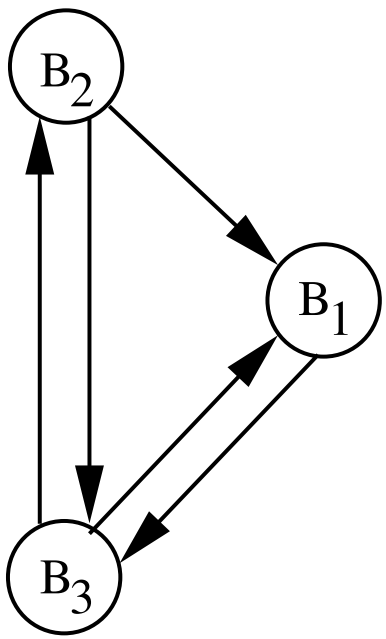
Convergentie naar de stationaire toestand

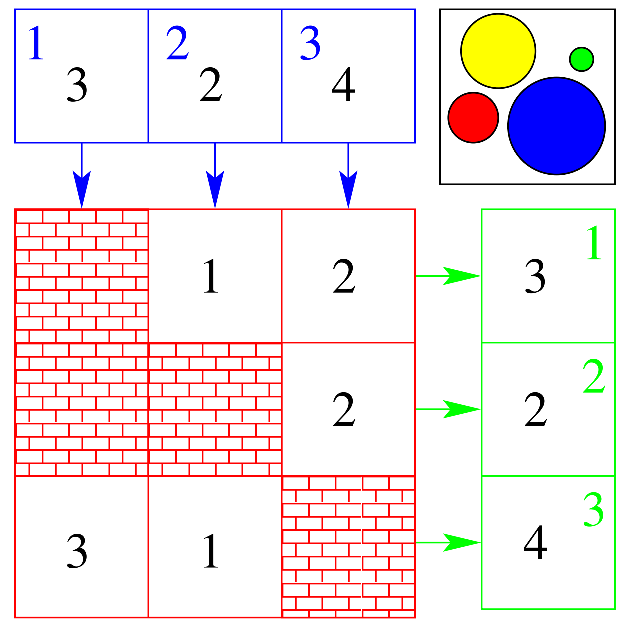


Veronderstel dat de goede fee uitdeelt:

$\boxed{\textbf{4.00} \text{ voor } B_1; \ \textbf{2.00} \text{ voor } B_2; \ \textbf{1.00} \text{ voor } B_3}$

Convergentie naar de stationaire toestand

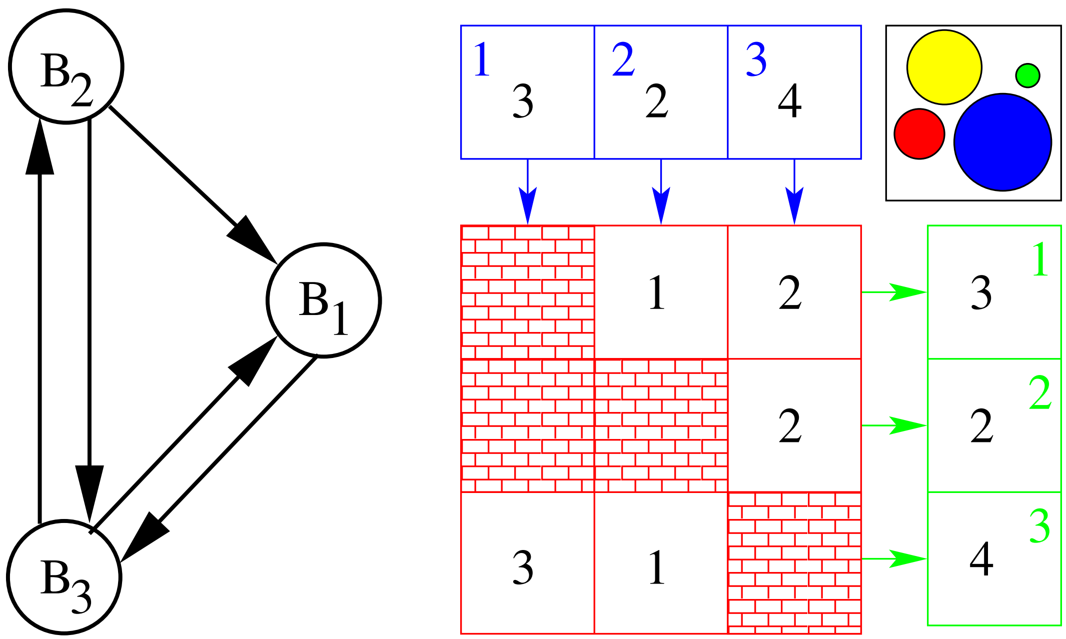




Na **01** keer knikkers weggeven is de situatie:

$\boxed{\textbf{1.50} \text{ voor } B_1;\ \ \textbf{0.50} \text{ voor } B_2;\ \ \textbf{5.00} \text{ voor } B_3}$

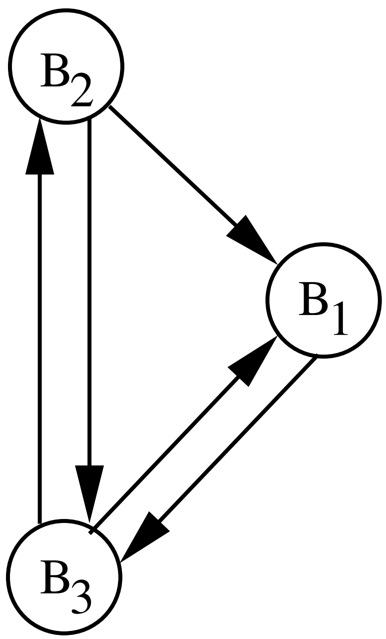
Convergentie naar de stationaire toestand

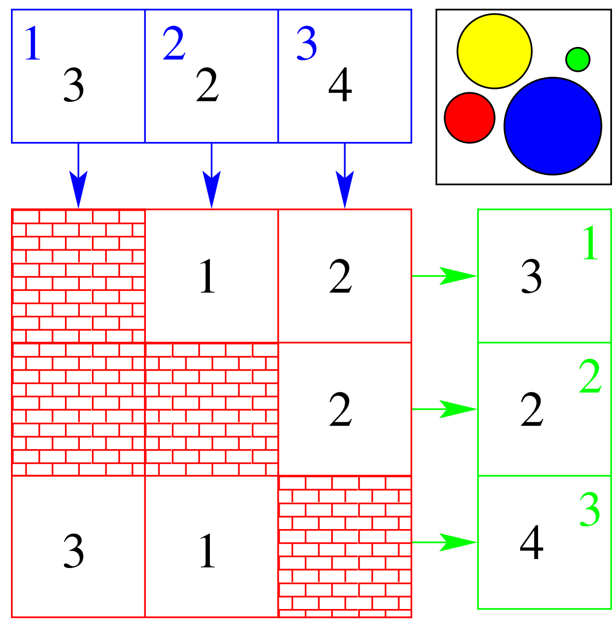


Na **02** keer knikkers weggeven is de situatie:

$\boxed{\textbf{2.75} \text{ voor } B_1; \;\; \textbf{2.50} \text{ voor } B_2; \;\; \textbf{1.75} \text{ voor } B_3}$

Convergentie naar de stationaire toestand

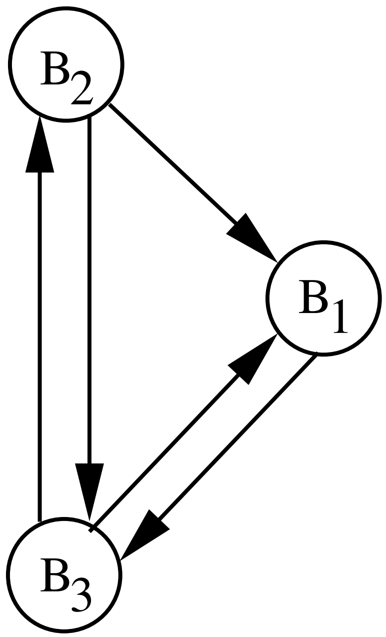


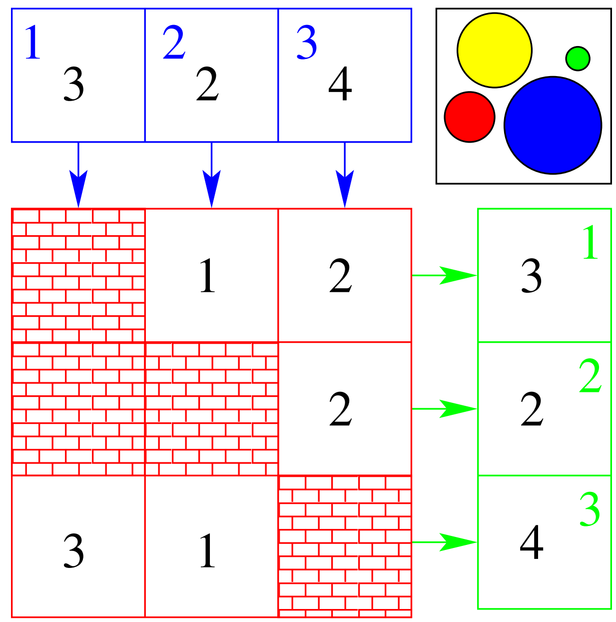


Na **03** keer knikkers weggeven is de situatie:

2.13 voor B_1; **0.88** voor B_2; **4.00** voor B_3

Convergentie naar de stationaire toestand

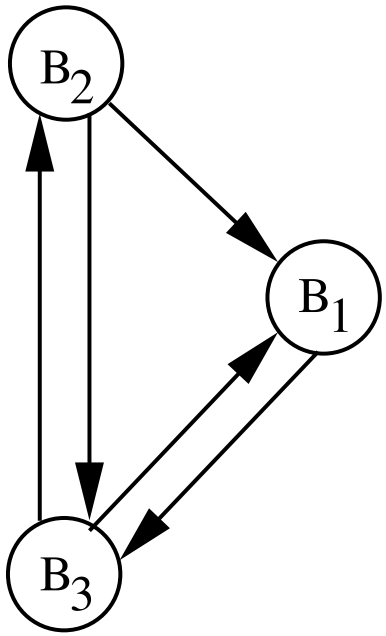


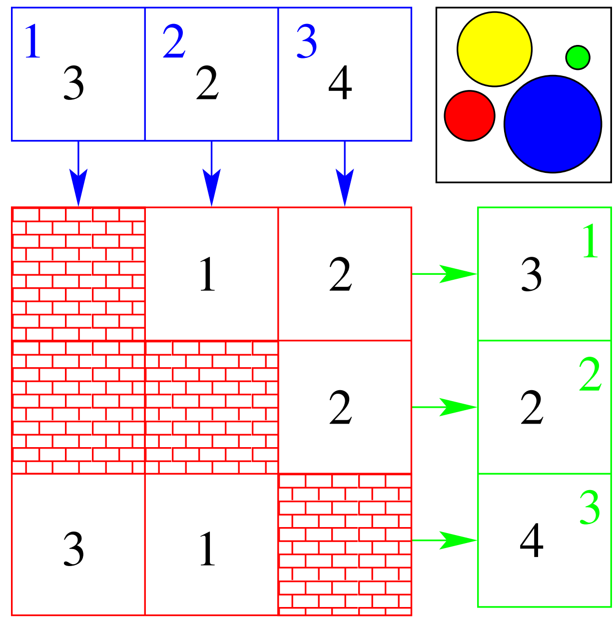


Na **05** keer knikkers weggeven is de situatie:

$\boxed{\textbf{2.28} \text{ voor } B_1; \quad \textbf{1.28} \text{ voor } B_2; \quad \textbf{3.44} \text{ voor } B_3}$

Convergentie naar de stationaire toestand

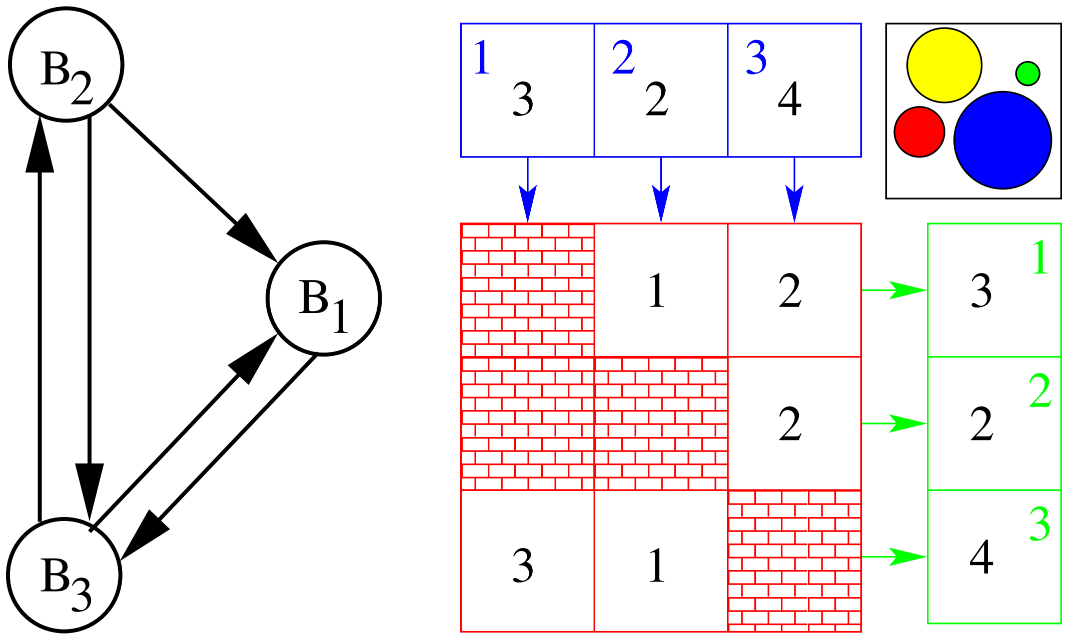




Na **10** keer knikkers weggeven is de situatie:

$$\boxed{\textbf{2.34} \text{ voor } B_1; \quad \textbf{1.57} \text{ voor } B_2; \quad \textbf{3.09} \text{ voor } B_3}$$

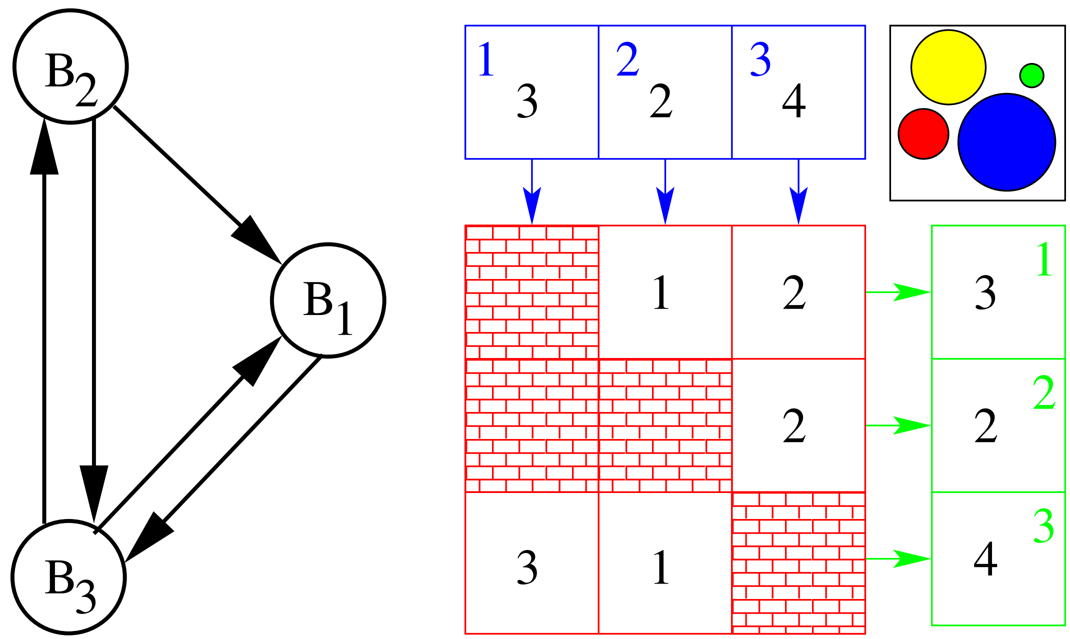
Convergentie naar de stationaire toestand



Na **13** keer knikkers weggeven is de situatie:

$\boxed{\textbf{2.33} \text{ voor } B_1; \ \ \textbf{1.55} \text{ voor } B_2; \ \ \textbf{3.11} \text{ voor } B_3}$

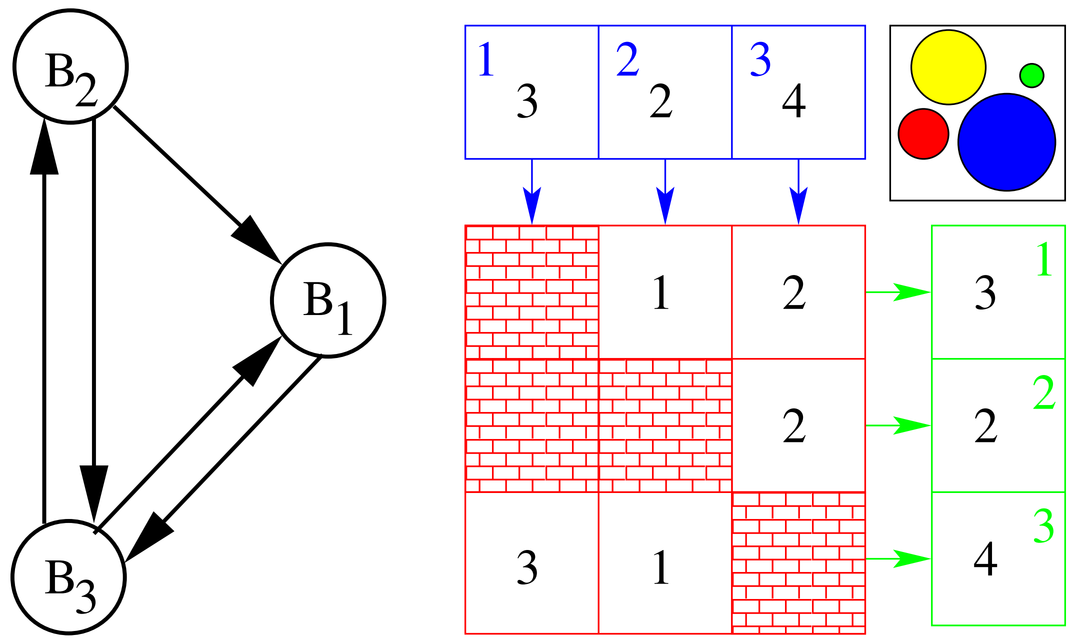
Convergentie naar de stationaire toestand



Vermenigvuldiging met **9** geeft

21.0 voor B_1; **14.0** voor B_2; **28.0** voor B_3

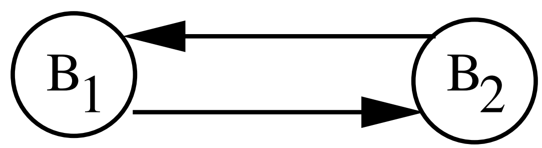
Convergentie naar de stationaire toestand



En deling door zeven uiteindelijk

$$\boxed{\textbf{3.00}\ \text{voor}\ B_1;\ \ \textbf{2.00}\ \text{voor}\ B_2;\ \ \textbf{4.00}\ \text{voor}\ B_3}$$

Geen convergentie naar de stationaire toestand

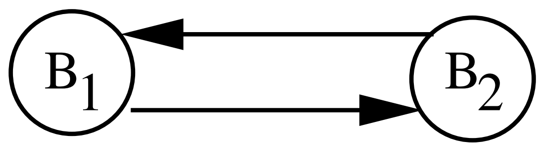


Stel, je begint met

$\boxed{\textbf{a} \text{ voor } B_1; \;\; \textbf{b} \text{ voor } B_2}$

En dus is er in dit simpele geval al geen convergentie

Geen convergentie naar de stationaire toestand

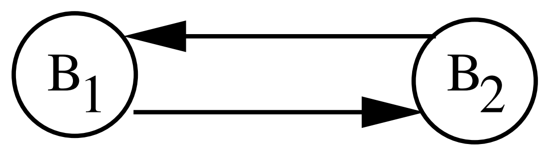


dan heb je in de volgende stap

b voor B_1; **a** voor B_2

En dus is er in dit simpele geval al geen convergentie

Geen convergentie naar de stationaire toestand

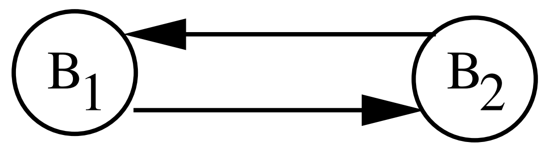


dan heb je in de volgende stap

$\boxed{\textbf{a} \text{ voor } B_1; \;\; \textbf{b} \text{ voor } B_2}$

En dus is er in dit simpele geval al geen convergentie

Geen convergentie naar de stationaire toestand



dan heb je in de volgende stap

$\boxed{\textbf{a} \text{ voor } B_1; \;\; \textbf{b} \text{ voor } B_2}$

En dus is er in dit simpele geval al geen convergentie

Nog meer teleportatie

Oplossing voor het convergentie-probleem is nog een α-snufje teleportatie

$$G = \alpha S + (1-\alpha)T \quad \text{waarbij} \quad T = \frac{1}{n}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

en α een parameter tussen de nul en één is

Enkele eigenschappen

De PageRank vector π voldoet aan $G\pi = \pi$ waarbij

$$\alpha \in [0,1], \quad G = \alpha S + (1-\alpha)T \quad \text{en} \quad T = \frac{1}{n}\begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

En

- is S de webstructuur-matrix $S = H + ea^T$
- tellen de kolommen van S op tot één: S is (kolom)-stochastisch
- tellen de kolommen van G op tot één: G is (kolom)-stochastisch
- S en G hebben daarom een eigenwaarde $\lambda = 1$
- alle entries van G zijn groter dan nul als $\alpha \neq 1$

Stelling 1

De eigenwaarden van een stochastische matrix S liggen in de complexe éénheidsschijf,

$$\lambda \in \sigma(S) \quad \Rightarrow \quad |\lambda| \leq 1$$

Bewijs. Laat $w \in \mathbb{R}^n$. Dan geldt voor de j-de entry $(S^T w)_j$ van $S^T w$

$$|(S^T w)_j| = |s_j^T w| \leq (s_j^T) w_+ \leq s_j^T e \max_k \{|w_k|\} = \max_k \{|w_k|\}$$

Laat nu $S^T w = \lambda w$, dan

$$|\lambda||w_j| \leq \max_k \{|w_k|\}$$

ook voor de j waarvoor $|w_j|$ maximaal is. Dus $|\lambda| \leq 1$. $\square$

Stelling 1

De eigenwaarden van een stochastische matrix S liggen in de complexe éénheidsschijf,

$$\lambda \in \sigma(S) \quad \Rightarrow \quad |\lambda| \leq 1$$

Bewijs. Laat $w \in \mathbb{R}^n$. Dan geldt voor de j-de entry $(S^T w)_j$ van $S^T w$

$$|(S^T w)_j| = |s_j^T w| \leq (s_j^T) w_+ \leq s_j^T e \max_k \{|w_k|\} = \max_k \{|w_k|\}$$

Laat nu $S^T w = \lambda w$, dan

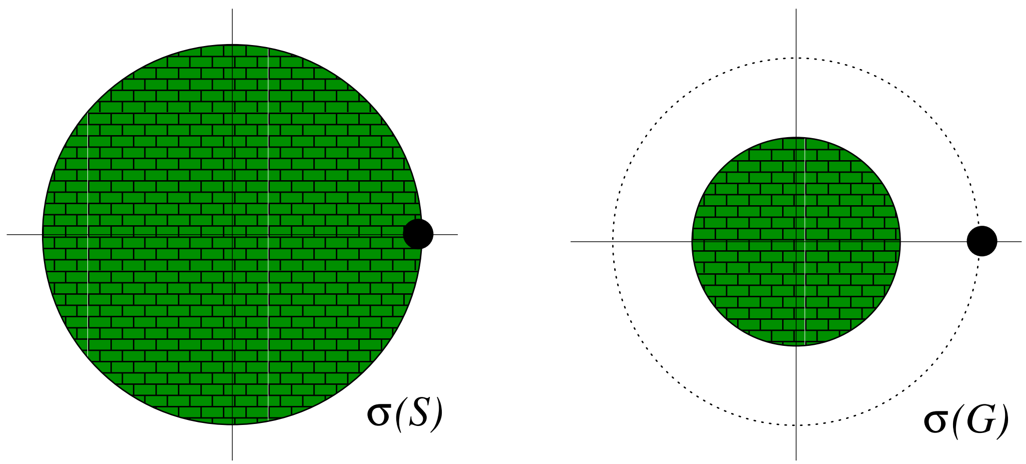
$$|\lambda||w_j| \leq \max_k \{|w_k|\}$$

ook voor de j waarvoor $|w_j|$ maximaal is. Dus $|\lambda| \leq 1$. $\square$

Stelling 2

De n eigenwaarden van $G = \alpha S + (1 - \alpha)T$ zijn:

- één
- α maal de $n-1$ eigenwaarden van S behalve één



Bewijs: Middels een orthogonale transformatie $Z = (e|X)$

Stelling 2

De n eigenwaarden van $G = \alpha S + (1-\alpha)T$ zijn:

- één
- α maal de $n-1$ eigenwaarden van S behalve één

Gevolg: $\sin \angle(\pi, \pi_k) = \mathcal{O}(\alpha^k)$

Bewijs: De convergentie-factor het quotiënt van de op één-na-grootste met de grootste absolute eigenwaarde van G, waarvoor geldt

$$\frac{|\lambda_2|}{|\lambda_1|} \leq \alpha$$

op grond van Stelling 2 $\square$

Stelling 3

Zij G stochastisch en positief. Laat $\mathcal{V}$ de eigenruimte van G zijn behorende bij $\lambda = 1$. Dan:

- $\pi \in \mathcal{V} \Rightarrow$ de entries van π zijn alle positief of alle negatief
- $\dim(\mathcal{V}) = 1$

Gevolg: De PageRank vector π als oplossing van

$$\pi = (\alpha S + (1 - \alpha)T)\pi$$

- bestaat
- is op een scalair veelvoud na uniek
- kan volledig positief gekozen worden
- en zodanig dat de entries optellen tot één

Conclusies

- PageRank is stationaire random surfer distributie
- Deze kan gevonden worden middels Goozzles
- Teleportatie zorgt voor goed-gedefinieerdheid
- Simpele dekpunt-iteraties zijn convergent
- met convergentie-factor α

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Conclusies

- PageRank is stationaire random surfer distributie
- Deze kan gevonden worden middels Goozzles
- Teleportatie zorgt voor goed-gedefinieerdheid
- Simpele dekpunt-iteraties zijn convergent
- met convergentie-factor α

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Conclusies

- PageRank is stationaire random surfer distributie
- Deze kan gevonden worden middels Goozzles
- Teleportatie zorgt voor goed-gedefinieerdheid
- Simpele dekpunt-iteraties zijn convergent
- met convergentie-factor α

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking

(Page & Brin, 1998)

Conclusies

- PageRank is stationaire random surfer distributie
- Deze kan gevonden worden middels Goozzles
- Teleportatie zorgt voor goed-gedefinieerdheid
- Simpele dekpunt-iteraties zijn convergent
- met convergentie-factor α

$$\boxed{\pi = \pi(\alpha S + (1 - \alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)

Conclusies

- PageRank is stationaire random surfer distributie
- Deze kan gevonden worden middels Goozzles
- Teleportatie zorgt voor goed-gedefinieerdheid
- Simpele dekpunt-iteraties zijn convergent
- met convergentie-factor α

$$\boxed{\pi = \pi(\alpha S + (1-\alpha)E)}$$

Google's PageRank vergelijking
(Page & Brin, 1998)


I ♥ Google™
Ready


...OG NÅ SOM OBELIX HAR FUNNET IGJEN HUN-
DEN SIN, LIVRETTEN (VILLSVIN) OG DEN GODE MAT-
LYSTEN SIN, TRØSTER ALLE VENNENE VÅRE SEG
MED EN ALLE TIDERS FEST... ALLE SÅ NÆR
SOM EN...
?!
!?
WAFF!
UDERZO & GOSCINNY