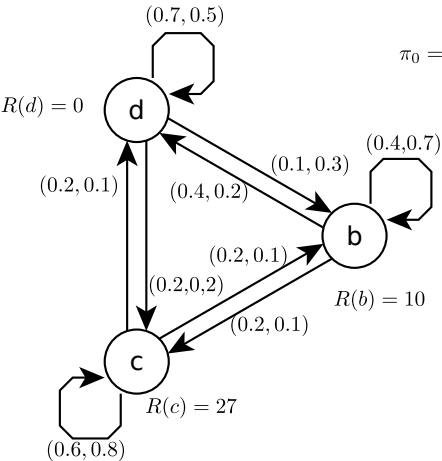


Probabilidades de transición parametrizadas por una política



$$\pi_0 = \begin{bmatrix} \neg \text{Anuncio} \\ \neg \text{Anuncio} \\ \neg \text{Anuncio} \end{bmatrix} \quad T(\pi_0) = \begin{bmatrix} 0.7 & 0.2 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.2 & 0.1 & 0.6 \end{bmatrix} \quad V_0(s) = \begin{bmatrix} 91.07 \\ 104.20 \\ 135.10 \end{bmatrix}$$

Encontramos la política  $\pi_1$  para la iteración siguiente

$$\pi_1 = \begin{bmatrix} \text{Anuncio} \\ \neg \text{Anuncio} \\ \text{Anuncio} \end{bmatrix}$$

$$\pi_0 \neq \pi_1$$

$$T(\pi_1) = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.1 \\ 0.2 & 0.2 & 0.8 \end{bmatrix} \quad V_1(s) = \begin{bmatrix} 127.58 \\ 138.57 \\ 181.98 \end{bmatrix}$$