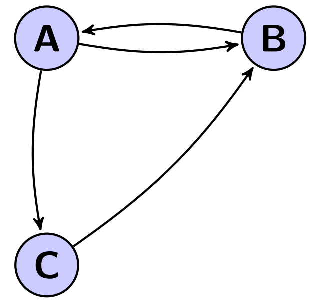
**Review.** Fibonacci numbers and Binet's formula

**Example 105.** Suppose the internet consists of only the three webpages $A, B, C$.

We wish to rank these webpages in order of "importance".




**The idea.** Instead of analyzing each webpage (which would be a lot of work!) we will try to only use the information how the pages are linked to each other. The idea being that an "important" page should be linked to from many other pages.

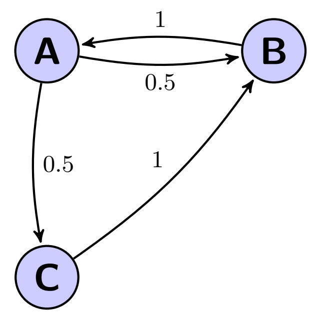
$A$ and $B$ have a link to each other. Also, $A$ links to $C$ and $C$ links to $B$. If you keep randomly clicking from one webpage to the next, what proportion of the time will you be at each page?

**The idea.** We will assign ranking to those pages according to how frequently such a random surfer would visit these pages.

**Comment.** Before we start computing, stop for a moment, and think about how you would rank the webpages.

**Solution.** Let $a_t$ be the probability that we will be on page $A$ at time $t$. Likewise, $b_t$, $c_t$ are the probabilities that we will be on page $B$ or $C$.

The transition from one state to the next now works exactly as in the previous example. We get the following transition matrix:




$$\begin{bmatrix} a_{t+1} \\ b_{t+1} \\ c_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \cdot a_t + 1 \cdot b_t + 0 \cdot c_t \\ \frac{1}{2} \cdot a_t + 0 \cdot b_t + 1 \cdot c_t \\ \frac{1}{2} \cdot a_t + 0 \cdot b_t + 0 \cdot c_t \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} a_t \\ b_t \\ c_t \end{bmatrix}$$

To find the equilibrium state, we again determine an appropriate $1$-eigenvector.

The $1$-eigenspace is $\operatorname{null}\left( \begin{bmatrix} -1 & 1 & 0 \\ \frac{1}{2} & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{bmatrix} \right)$ which has basis $\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$.

The corresponding equilibrium state is $\frac{1}{5}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$. In this context, this is also known as the **PageRank vector**.

In other words, after browsing randomly for a long time, there is (about) a $\frac{2}{5} = 40\%$ chance to be at page $A$, a $\frac{2}{5} = 40\%$ chance to be at page $B$, and a $\frac{1}{5} = 20\%$ chance to be at page $C$.

We therefore rank $A$ and $B$ highest (tied), and $C$ lowest.

**Just checking.** Maybe we were expecting $B$ to be ranked above $A$, because $B$ is the only page that has two incoming links. However, if we are at page $B$, then our next click will be to page $A$, which is why $A$ and $B$ receive equal ranking.

This method of ranking is the famous **PageRank** algorithm (underlying Google's search algorithm).

By the way, the algorithm is named, not after ranking web"pages", but after Larry Page (who founded Google in 1998 together with Sergey Brin).

Armin Straub
straub@southalabama.edu