# Lesson6: Modeling the Web as a graph
# Unit5: Linear Algebra for graphs

Rene Pickhardt

Introduction to Web Science Part 2

Emerging Web Properties

WeST People and Knowledge Networks

## Completing this unit you should

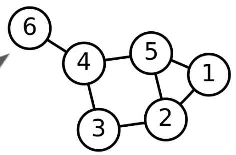
- Be able to read and build an adjacency matrix of a graph
- Know some basic matrix vector multiplications to generate some statistics out of the adjacency matrix
- Understand what is encoded in the components of the k-th power of the Adjacency matrix of a graph

# A meta model for graph based models

Simple English Wikipedia

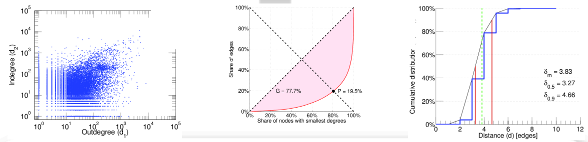
What does simple Wikipedia look like?

modelling




**Graph as a model**

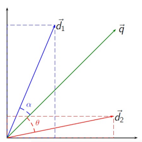
calculating



**Descriptive statistics**

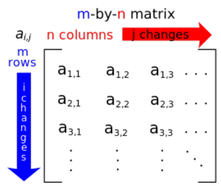
interpreting

modelling



**Vector space model**

calculating

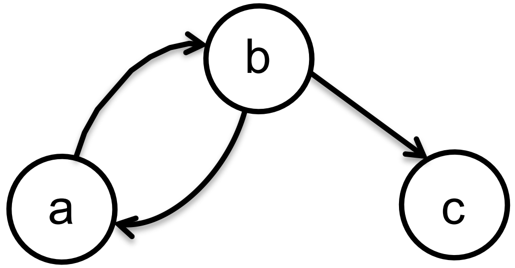



**Matrix calculations**

interpreting

# Modelling graphs with linear algebra

- Given this graph G
- Which of the following is an adjacency matrix for G?




$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

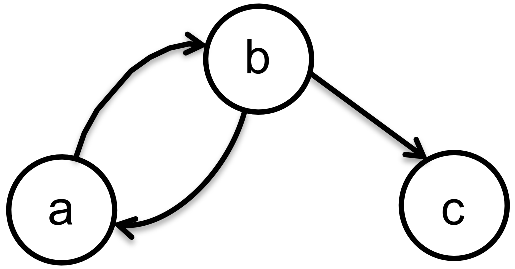
$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

# All four matrices represent the same graph!

- **Matrices are not the same**
- How can we make sure to talk about the **same thing**?



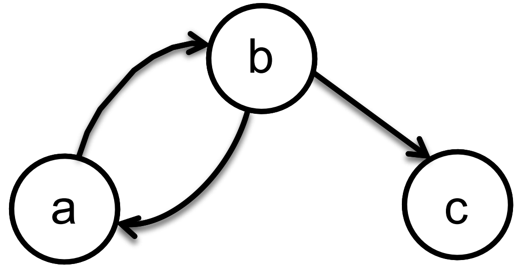

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

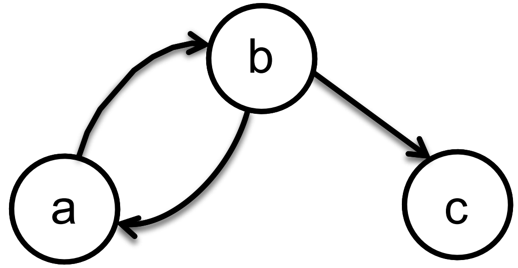
# Fix a base of our vector space




- We have out graph
  - With a set of vertices $V = \{a, b, c\}$
- We have our Vector space
  - With a set of base vectors $\mathbb{V} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$
  - Without loosing generality we can choose unit vectors

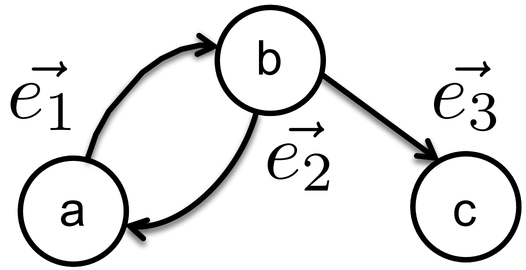
$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

# Defining the adjacency matrix

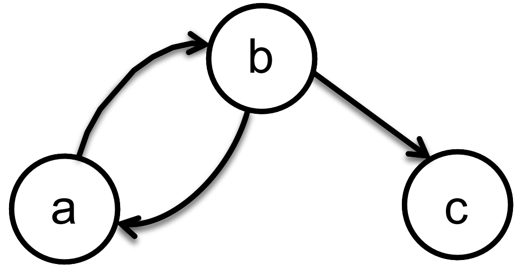



- Fix a base i.e. define:
  - choose $f_{base} : V \longrightarrow \mathbb{V}$
  - With $V = \{a, b, c\}$ and $\mathbb{V} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$
  - And $f_{base}$ must be bijective
  - $f_{base}(\{x, y\}) = f_{base}(x) + f_{base}(y)$

- For example one choice might be
  - $f_{base}(a) = \vec{e_1}$
  - $f_{base}(b) = \vec{e_2}$
  - $f_{base}(c) = \vec{e_3}$



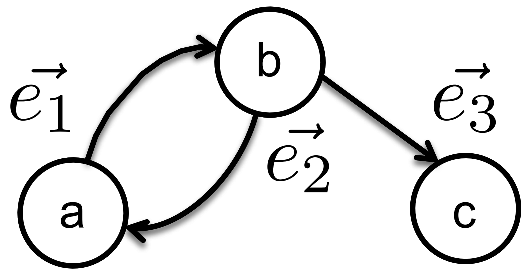
# Defining the adjacency matrix



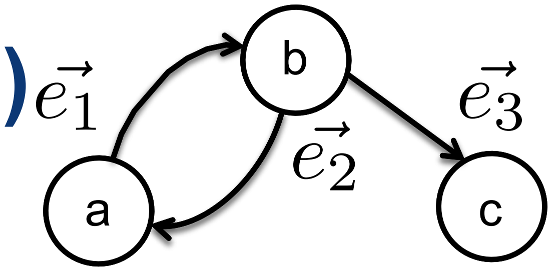

- Fix a base i.e. define:
  - choose $f_{base} : V \longrightarrow \mathbb{V}$
  - With $V = \{a, b, c\}$ and $\mathbb{V} = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$
  - And $f_{base}$ must be bijective
  - $f_{base}(\{x, y\}) = f_{base}(x) + f_{base}(y)$

Pay attention: Notation is somehow sloppy. Here f is not defined on sets.

- For example one choice might be
  - $f_{base}(a) = \vec{e_1}$
  - $f_{base}(b) = \vec{e_2}$
  - $f_{base}(c) = \vec{e_3}$

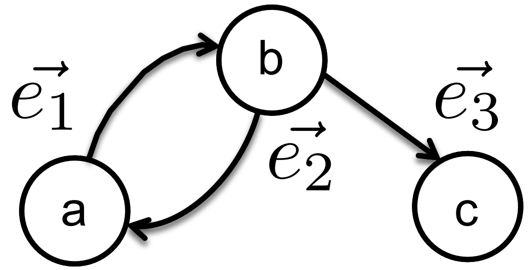


# Defining the adjacency matrix (II)




- We are looking for a matrix
  $A : \mathbb{V} \longrightarrow \mathbb{V}$ such that
  $$A(f_{base}(a)) = f_{base}(In(a))$$
- Don‘t panic!
- A just encodes the outlink of every vertex
- It is made in a way that it respects the choice of our base

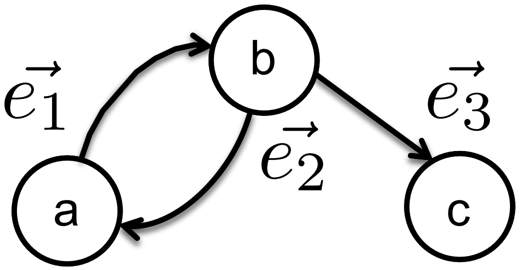
# Example $A(f_{base}(a)) = f_{base}(In(a))$




$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rightarrow A\vec{e_1} = \vec{e_2}$$

Multiplying a matrix with the i-th base vector from **right** gives the i-th column!

# Example $A(f_{base}(a)) = f_{base}(In(a))$



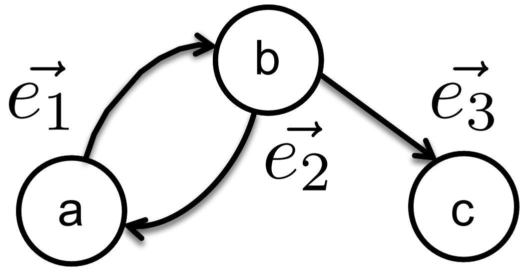

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rightarrow A\vec{e_1} = \vec{e_2}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rightarrow A\vec{e_2} = \vec{e_1}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rightarrow A\vec{e_3} = \vec{e_2}$$

For each multiplication the result is the vectors of nodes represnting in links

# Similarly for out nodes we get




- Left multiply with a base vector

$$\begin{pmatrix}0 & 1 & 0\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1\end{pmatrix} \rightarrow \vec{e_2}^{\,t} A = \vec{e_1}^{\,t} + \vec{e_3}^{\,t}$$

$f_{base}(b) = \vec{e_2}$ $\quad f_{base}(a) = \vec{e_1}$ $\quad f_{base}(c) = \vec{e_3}$

- Result is a vector representing $Out(b) = \{a, c\}$

Multiplying a matrix with the i-th base vector from **left** gives the i-th row!

# What happens if A is applied several times?

- We saw apply it once to a base vector yields all neighbours linking in or out
  - Neighbours can be seen as paths of lenght 1
  - Each component counts how many paths of length 1 exist from $\vec{e_i}^t$ to the node of the component

- So what does $\vec{e_i}^t A^k$ represent?

- What about $A^k \vec{e_i}$ ?

# What happens if A is applied several times?

- $\vec{e_i}^t A^k$ represents
  - A vector where in each compnent is the number of **paths of length k from** the node represented by $\vec{e_i}^t$ **to** the node represented by that component.

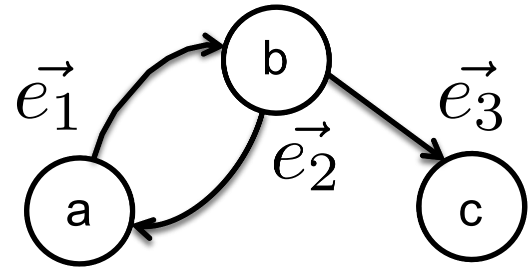
- $A^k \vec{e_i}$ represents
  - A vector where in each compnent is the number of **paths of length k to** the node represented by $\vec{e_i}$ **from** the node represented by that component

# Quiz question! Who is still with me (:

If *A* represents the adjacency matrix of the strongly connected component of Simiple English Wikipedia

**For what power of $k$ will $A^k$ have no zero entries?**

# Degree distribution multiply with a counting vector we get




$$\begin{pmatrix}1 & 1 & 1\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\1 & 0 & 1\\0 & 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 1 & 1\end{pmatrix} = 1\vec{e_1}^t + 1\vec{e_2}^t + 1\vec{e_3}^t$$

In degree distribution

Counting vector

$$\begin{pmatrix}0 & 1 & 0\\1 & 0 & 1\\0 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\2\\0\end{pmatrix} = 1\vec{e_1} + 2\vec{e_2} + 0\vec{e_3}$$

Out degree distribution

# Thank you for your attention!


Contact:
Rene Pickhardt
Institute for Web Science and Technologies
Universität Koblenz-Landau
rpickhardt@uni-koblenz.de

WeST
People and Knowledge Networks

# Pictures:

- https://commons.wikimedia.org/wiki/File:6n-graf.svg By User:AzaToth (Image:6n-graf.png simlar input data) [Public domain], via Wikimedia Commons
- https://pixabay.com/en/lightbulb-electric-light-bright-idea-31254/ Public domain
- https://commons.wikimedia.org/wiki/File:Simple-bipartite-graph.svg By MistWiz (Own work) [Public domain], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:CPT-Graphs-undirected-weighted-loop-multiedge.svg By Pluke (Own work) [CC0], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:Matrix.svg By Lakeworks (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 4.0-3.0-2.5-2.0-1.0 (http://creativecommons.org/licenses/by-sa/4.0-3.0-2.5-2.0-1.0)], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:Ego_network.png By Houl0078 (Own work) [CC0], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:Path-graph-theory.svg By MartinThoma (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY 3.0 (http://creativecommons.org/licenses/by/3.0)], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:%22Bow-tie%22_diagram_of_components_in_a_directed_network_SVG.svg By Limaner (Own work) [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:Economics_Gini_coefficient2.svg By Reidpath [Public domain], via Wikimedia Commons
- http://konect.uni-koblenz.de/statistics/volume
- https://www.flickr.com/photos/dannysullivan/369540238 via flickr CC-BY 2.0 by Danny Sullivan
- Graph statistics via http://konect.uni-koblenz.de/networks/link-dynamic-simplewiki

- https://commons.wikimedia.org/wiki/File:WWW_logo_by_Robert_Cailliau.svg public domain
- https://commons.wikimedia.org/wiki/File:WorldWideWeb.1.png public domain
- https://en.wikipedia.org/w/index.php?title=World_Wide_Web&oldid=692731715 (definition of world wide web)
- https://commons.wikimedia.org/wiki/File:Visualization_of_the_world_wide_web_common_crawl_2012.png By Sebastian Schelter [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:WorldWideWebAroundWikipedia.png Chris 73 / Wikimedia Commons [GFDL 1.3 (www.gnu.org/licenses/fdl-1.3.html) or CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons
- https://commons.wikimedia.org/wiki/File:Word_Cloud_Readers_Survey.jpg By Manipande (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons
- https://www.flickr.com/photos/valeriebb/290711738 via Flickr and Valerie Everett
- https://commons.wikimedia.org/wiki/File:Inner-product-angle.png CC-BY-SA by CSTAR & Oleg Alexandrov
- https://commons.wikimedia.org/wiki/File:Synoptic_word-for-word.png By Alecmconroy (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY 3.0 (http://creativecommons.org/licenses/by/3.0)], via Wikimedia Commons