Lesson6:
Modeling the Web as a graph
Unit1:
Reviewing basic terms from graph theory

Rene Pickhardt

Introduction to Web Science Part 2

Emerging Web Properties

WeST
People and Knowledge Networks

Institute for Web Science and Technologies · University of Koblenz-Landau, Germany

Completing this unit you should

- Be familiar with a set theoretic way of denoting a graph
- Know at least 4 different types of graphs
- Have practiced your abilities in reading and writing mathematical formulas

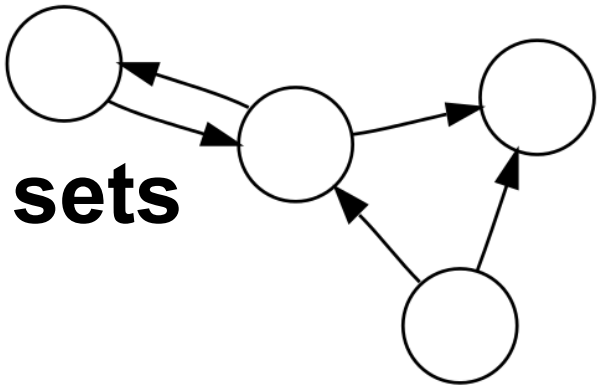
Let‘s approach graph based models like text

1. Descriptive modelming
   - Understand the topology
   - Look at distributions of occurence
   - find measures to quantify what we have seen

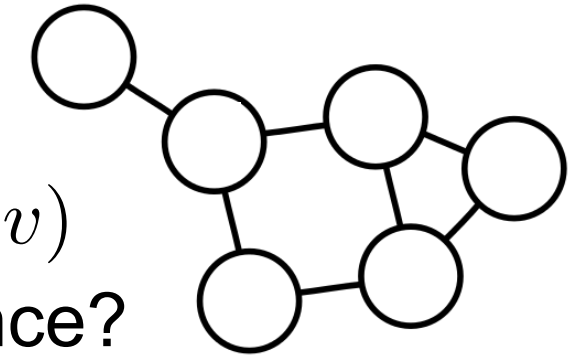
2. Linear algebra
   - Maybe useful for predictions

3. Generative modeling
   - Find a process to generate the descriptive model

Definition of a graph



- A **Graph** $G(V, E)$ consits of **two sets**
  - V the set of **vertices** or nodes
  - $E \subseteq V \times V$ the set of **edges**
  - V is usually finite, thus is E



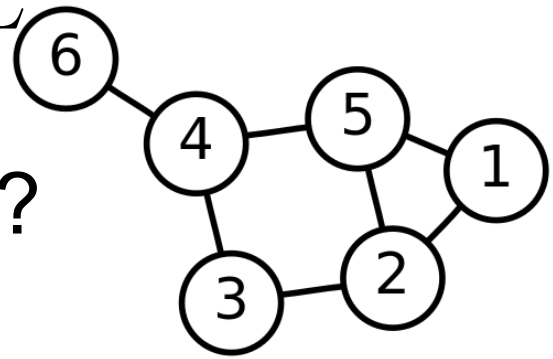
- A Graph is called **undirected** if:
  - $\forall e = (v, u) \in E \exists e' \in E : e' = (u, v)$
  - How it this expressed as a scentence?

Definition of a labeled graph

- A graph is called **labeled** if it is a **vertex-labeled** graph or an **edge-labeled** graph.

- A Graph $G(V, E)$ is called
  - **vertex-labeled** if there is a **labeling function** $l : V \longrightarrow L$ with L a set of labels.
  - **edge-labeled** if the labeling function has the edges as a domain i.e.: $l : E \longrightarrow L$

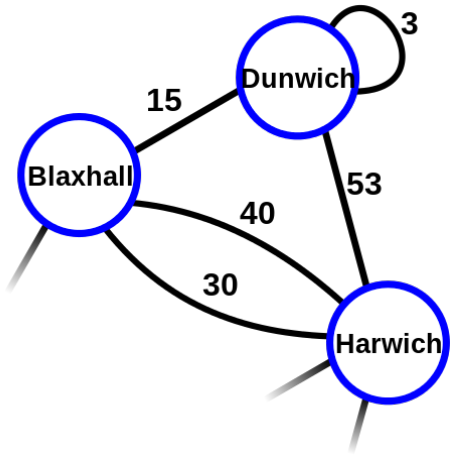
- Can you think of some examples?



Definition of weighted graphs

- A **weighted** Graph is an edge-labeled graph with a **labeling function** $\omega : E \longrightarrow \mathbb{R}$ such that every label is a real number.

- The following choices for L are pretty common
  - $L = \mathbb{N}$
  - $L = \{0, 1\}$
  - $L = \{-1, 1\}$



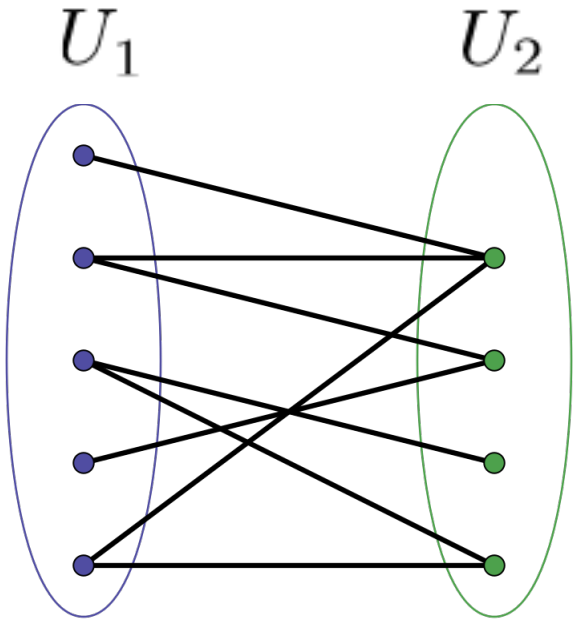
Definition of a graph

- A Graph $G(V, E)$ is called if and only if
  - Property 1
    - $\exists U_1, U_2 \subsetneq V$
    - $V = U_1 \cup U_2$
    - $U_1 \cap U_2 = \emptyset$
  - Property 2

$$\forall e = (u, v) \in E : u \in U_1 \wedge v \in U_2 \vee v \in U_1 \wedge u \in U_2$$

Definition of a bipartite graph

- A Graph $G(V, E)$ is called bipartite if and only if
  - Vertices have a **disjoint split**
    - $\exists U_1, U_2 \subsetneq V$
    - $V = U_1 \cup U_2$
    - $U_1 \cap U_2 = \emptyset$
  - Such that **all edges cross** the disjoint sets



$$\forall e = (u, v) \in E : u \in U_1 \wedge v \in U_2 \vee v \in U_1 \wedge u \in U_2$$

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