CS 6170: Computational Topology, Spring 2019 Lecture 02

Topological Data Analysis for Data Scientists

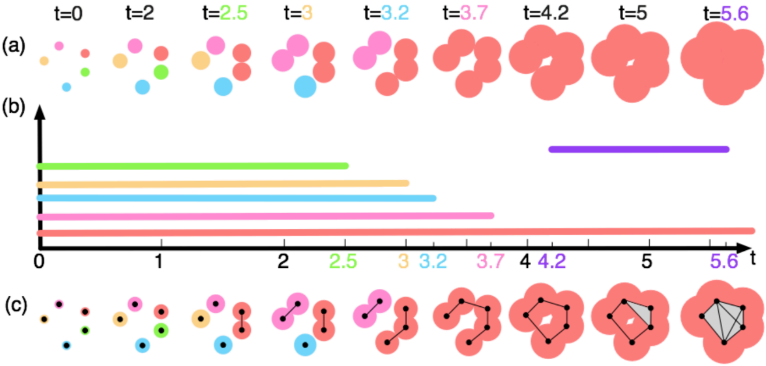
Dr. Bei Wang

School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah
www.sci.utah.edu/~beiwang
beiwang@sci.utah.edu

Jan 10, 2019

Persistent Homology in a Nutshell

Persistent homology at a glance



Wong et al. (2016)

Previous Project Ideas

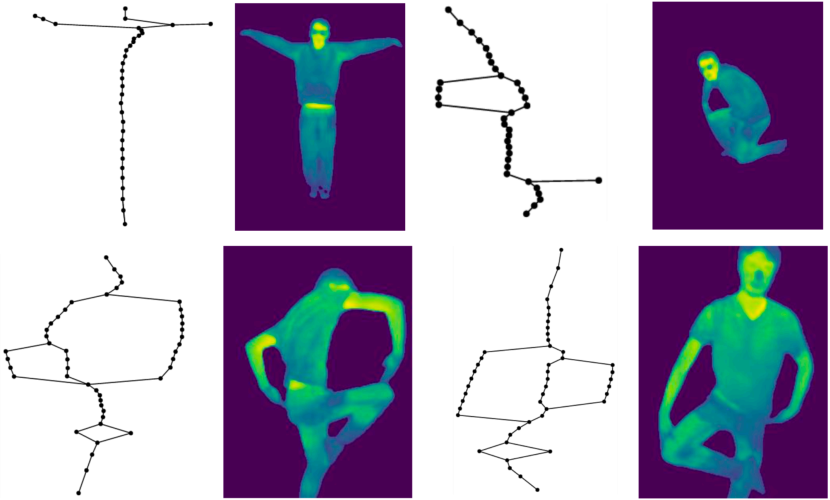
Previous Course Projects

- Topology of Elementary Thermodynamic Models
- Pose Analysis of Thermal Images
- Using Mapper on Netflix data
- Visualizing Periodicity in Time Series Data
- Stratification Learning
- Persistence applied to decomposed spaces
- Prediction of Grasp Stability using Topological Data Analysis
- Distributing the Mapper Algorithm
- Topology in Uncertain Visualization
- Topological data analysis of mice pregnancy data
- Jacobi Set in Discrete Morse Theory
- Using Mapper Explore High Dimensional Data
- Visualization of Sensor Network Coverage with Sensor Location Uncertainty

Pose Analysis of Thermal Images

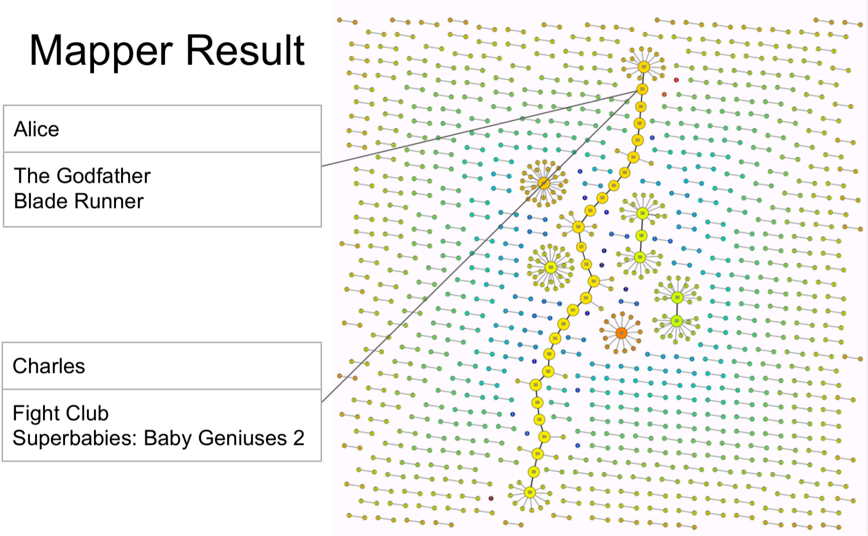
- Prediction of Grasp Stability using TDA
- Moving Sensor Coverage
- A TDA of IED Explosions in Afghanistan 2004-2009
- Study and Verification of Topological Strata of Weighted Complex Networks
- Analysis of High Dimensional Autism Brainwave Data Using Mapper
- Openspace and Disperse
- Topology guides volume exploration
- TDA for Bird Migration
- Semantic Segmentation with Topological Methods

Pose Analysis of Thermal Images



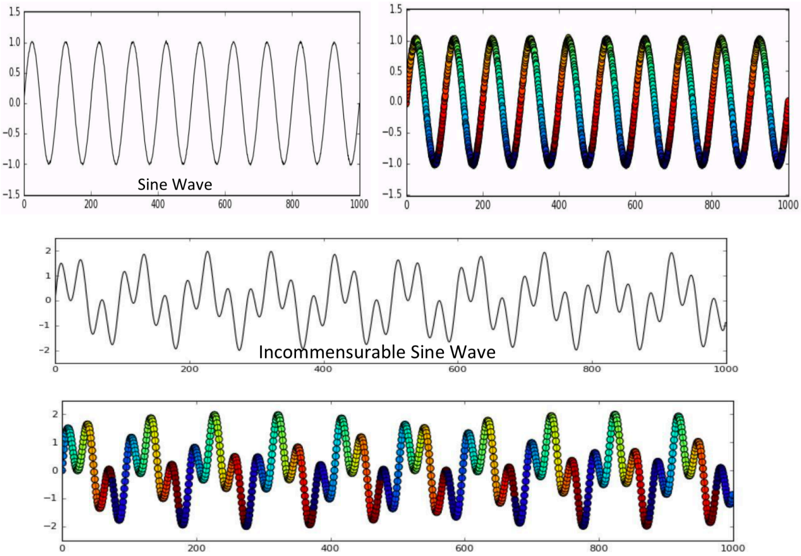
S. Anderson and D. Geisler, 2017

Analysis of Netflix Data



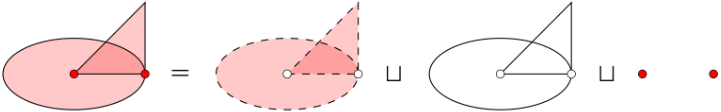
M. Annasamy and S. Brooks, 2017

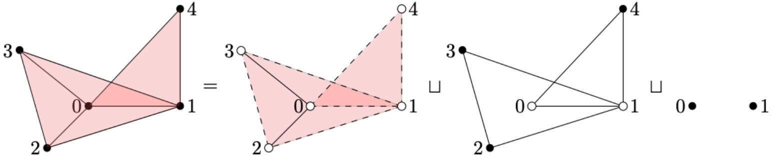
Visualizing Periodicity in Time Series Data



J. Boyer and S. Ram, 2017

Theory: Stratification Learning Using Sheaves*

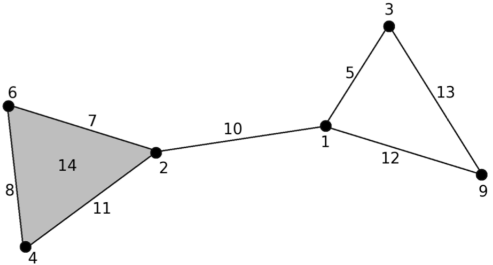




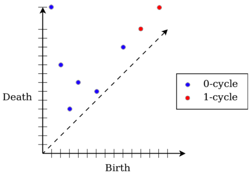
A. Brown, 2017, Brown and Wang (2018)

Theory: Persistence applied to decomposed spaces

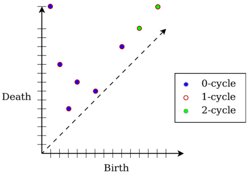
Consider the space Y:



Persistence diagram for Y

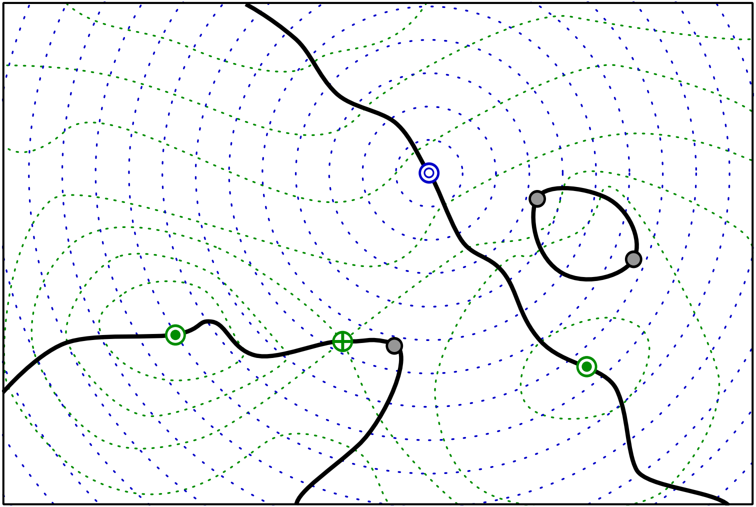


Persistence diagram for $S^1 \times Y$



K. Childers, 2017

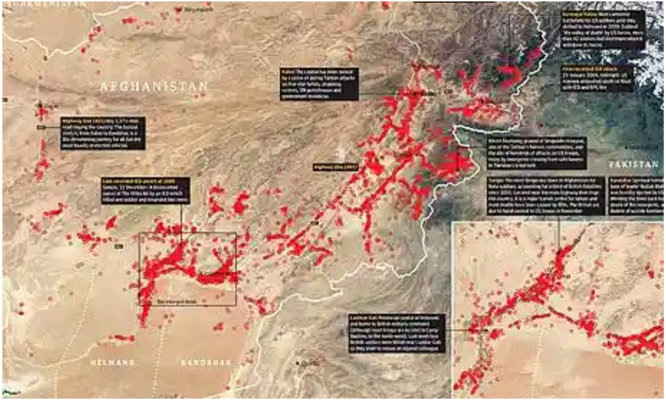
Theory: Jacobi Set in Discrete Morse Theory



P. Klacansky, 2017.

Edelsbrunner and Harer (2004)

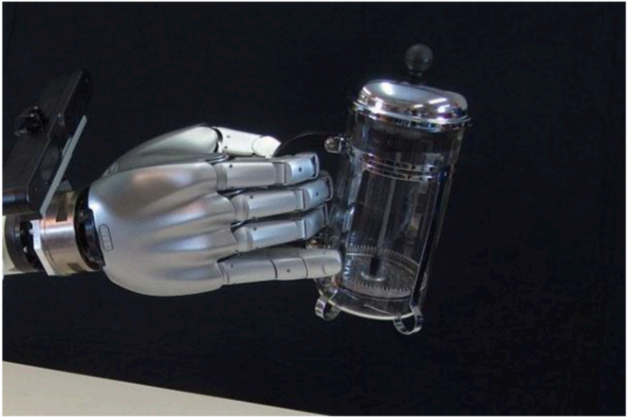
A TDA of IED Explosions in Afghanistan 2004-2009



P. Narayanan, 2017

https://www.theguardian.com/world/datablog/2010/jul/26/wikileaks-afghanistan-ied-attacks

Prediction of Grasp Stability using TDA



A. Conkey, 2017

Image Courtesy: `http://www.pacman-project.eu/news/`

Distributed Computation of the Mapper Algorithm

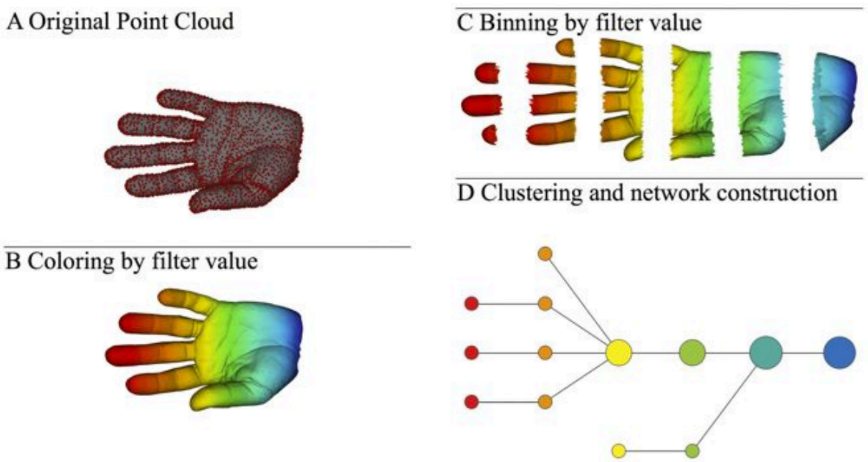
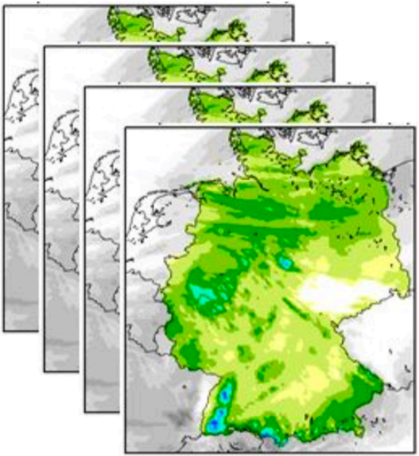


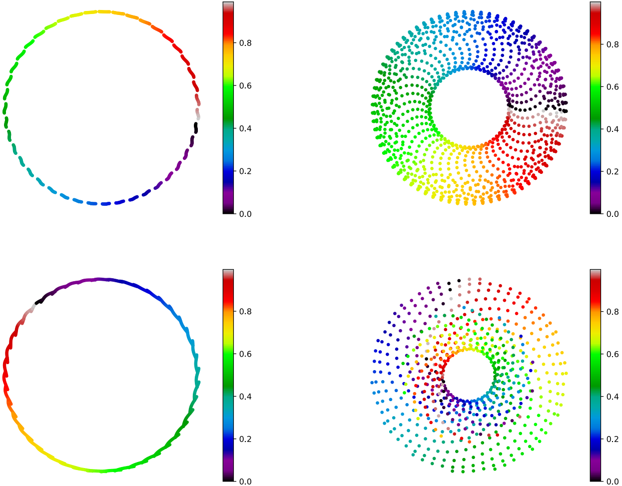
Image courtesy Lum et al. (2013)
P. English, 2017.

Topology for Uncertainty Visualization



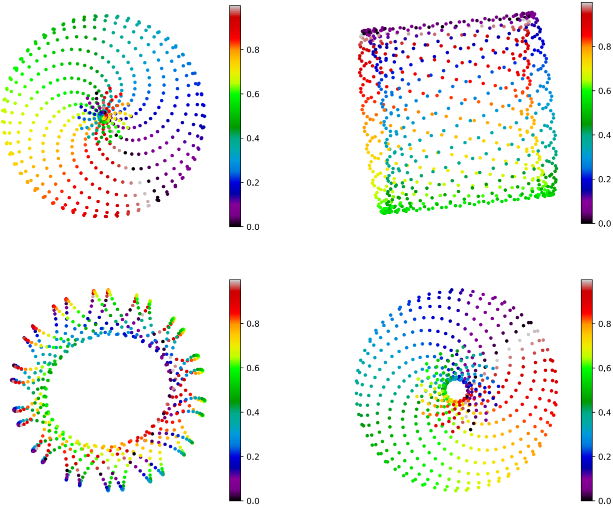
S. Mehrpour and N. Farhoudi, 2017.

TDA of Pregnancy Mice*



Which mice are pregnant? Data from Smarr et al. (2016).
V. Jose and A. Sharma, 2017. Sharma (2018)

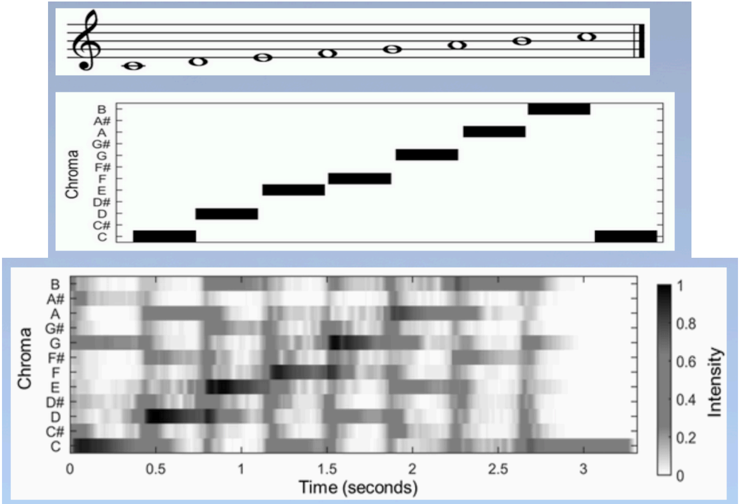
TDA of Pregnancy Mice*



Which jet lagged mice are pregnant?

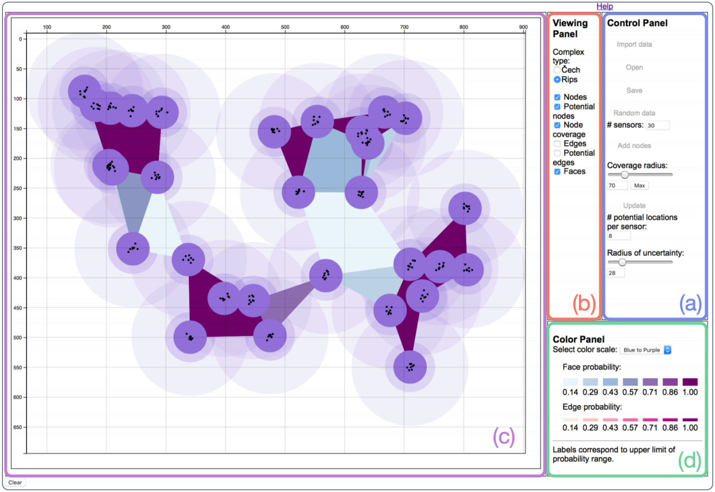
V. Jose and A. Sharma, 2017. Sharma (2018)

Song Similarities via Homology of Chroma Features



S. Leventhal and Z. Fahimfar, 2017.

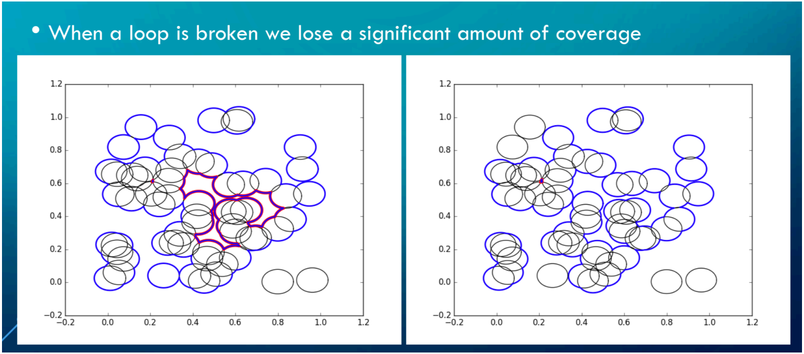
Sensor Network Coverage with Location Uncertainty*



T. Sodergren and J. Lohse, 2017. Sodergren et al. (2017)

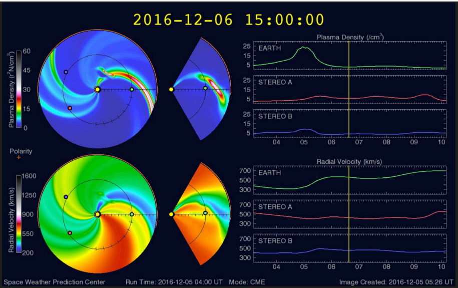
http://www.sci.utah.edu/~tsodergren/prob_net_vis_working/

Moving Sensor Coverage

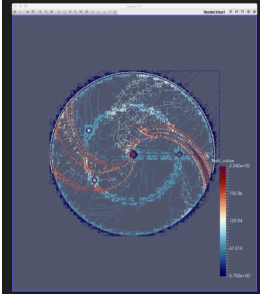


D. Mcclelland and M. Matheny, 2017.

Astronomy: Open Space and DisPerse



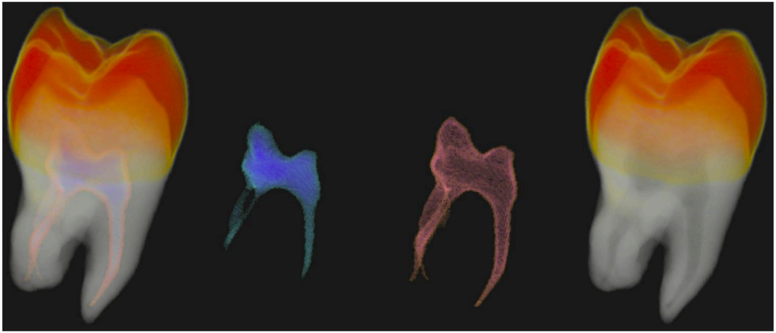
www.swpc.noaa.gov/products/wsa-enlil-solar-wind-prediction



Cut 1

M. Territo, 2017. `www.openspaceproject.com`

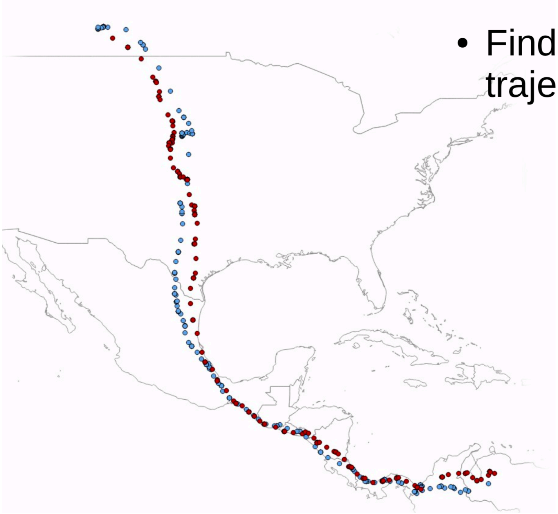
Topology Guided Volume Exploration



Tooth after per-segment classification and removal of background. The Inner pocket and boundary are clear to see and not affected by removal of noise

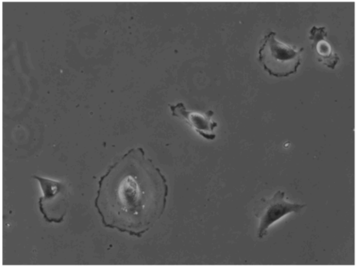
W. Usher, Q. Wu , 2017.

TDA for Birth Migration



- Find two migration trajectories

J. Wagstaff, 2017.



(a) Input from PhC-U373 Dataset

(b) Segmented Result from the Trained Network

D. Wang, 2017.

Data has shape and shape matters

Possible quote by G. Carlsson.

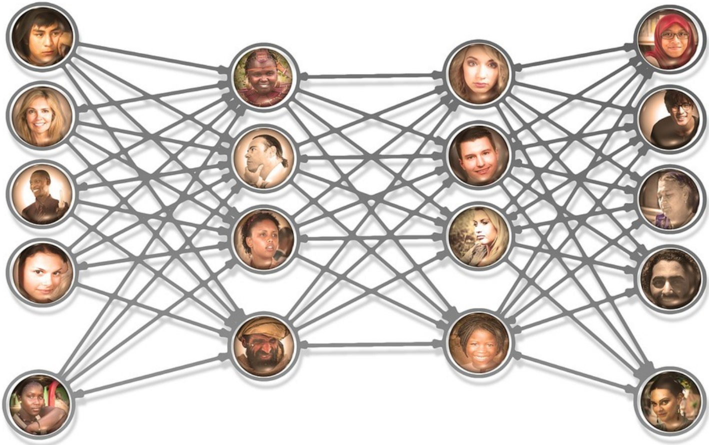
Graphs and Connected Components

Book Chapter A.I.

Graphs in DS

https://www.flickr.com/photos/frauhoelle/8464661409

Social Network



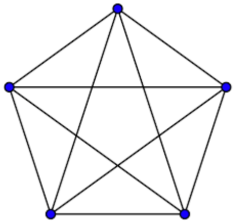
pixabay.com

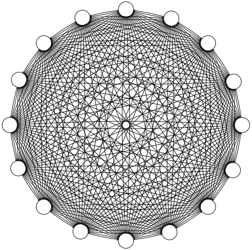
Simple Graphs

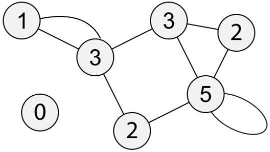
Definition (Graph)

An abstract *graph* is a pair $G = (V, E)$ consisting of a set of vertices V, a set of edges E for each pair of vertices.

- A graph is *simple* if no two edges connect the same two vertices and there is no self-loop.
- A *complete graph* contains an edge for every pair of vertices.
- A *regular graph* contains vertices with the same degree.

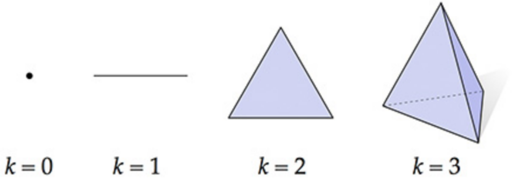


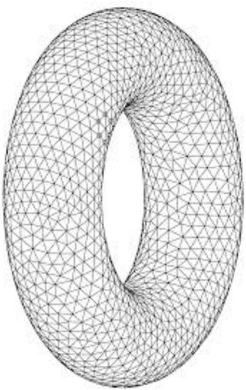




Simplex

- K_n: a complete graph with n vertices.
- K_n represents edges of a $(n-1)$-simplex.





Connectivity of a graph

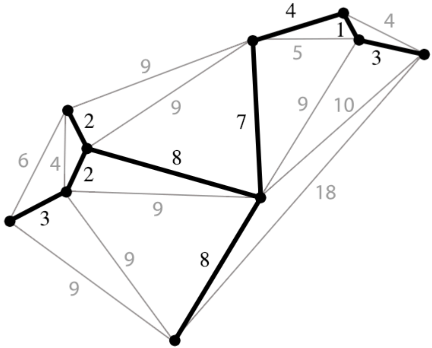
Definition (Connected Graph)

A simple graph is *connected* if there is a path between every pair of vertices. A *connected component* (CC) of a graph is a maximal subgraph that is connected.

- A *path* between vertices u and v is a sequence of vertices $u = u_0, u_1, \cdots, u_k = v$, with an edge between u_i and u_{i+1} for each $0 \leq I \leq k-1$.
- *Simple path*: all vertices are distinct (no loops)
- *Path length*: number of edges traversed

Tree, spanning tree

- A *tree* is an undirected graph in which any two vertices are connected by exactly one path.
- The smallest connected graph is a tree (n vertices, $n-1$ edges).
- Deleting any edge disconnects the tree.
- A *spanning tree* of a graph $G=(V,E)$ is a tree $T=(V,E')$ with $E' \subseteq E$.



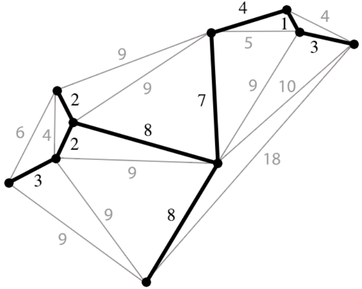
https://commons.wikimedia.org/wiki/File:Minimum_spanning_tree.svg

Separation

Definition (Separation)

A *separation* is a non-trivial partition of the vertices; that is, $V = U \cup W$ where U and W are nonempty, such that no edge connects a vertex in U with a vertex in V.

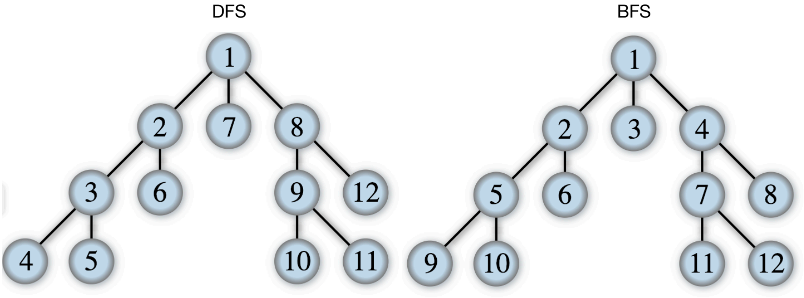
- A simple graph is *connected* iff it has no separation.
- A simple graph is *connected* iff it has a spanning tree.



https://commons.wikimedia.org/wiki/File:Minimum_spanning_tree.svg

Algorithms to test connectivity of a graph

- DFS: depth-first search
- BFS: breadth-first search
- Union-find



https://upload.wikimedia.org/wikipedia/commons/3/33/Breadth-first-tree.svg

https://commons.wikimedia.org/wiki/File:Depth-first-tree.svg

References I

Brown, A. and Wang, B. (2018). Sheaf-Theoretic Stratification Learning. *International Symposium on Computational Geometry (SOCG)*.

Edelsbrunner, H. and Harer, J. (2004). Jacobi sets of multiple Morse functions. In F. Cucker, R. DeVore, P. O. and Sueli, E., editors, *Foundations of Computational Mathematics, Minneapolis 2002*, pages 37–57. Cambridge University Press, England.

Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., Carlsson, J., and Carlsson, G. (2013). Extracting insights from the shape of complex data using topology. *Scientific Reports*, 3.

Sharma, A. (2018). Topological analysis and visualization of mice temperature data for exploring biological events. University of Utah, Master Thesis.

Smarr, B. L., Zucker, I., and Kriegsfeld, L. J. (2016). Detection of successful and unsuccessful pregnancies in mice within hours of pairing through frequency analysis of high temporal resolution core body temperature data. *PLoS One*.

References II

Sodergren, T., Hair, J., Phillips, J. M., and Wang, B. (2017). Visualizing sensor network coverage with location uncertainty. *Symposium on Visualization in Data Science (VDS) at IEEE VIS*.

Wong, E., Palande, S., Wang, B., Zielinski, B., Anderson, J., and Fletcher, P. T. (2016). Kernel partial least squares regression for relating functional brain network topology to clinical measures of behavior. *International Symposium on Biomedical Imaging (ISBI)*.