185.A83 Machine Learning for Health Informatics
2021S, VU, 2.0 h, 3.0 ECTS
Andreas Holzinger, Rudolf Freund
Marcus Bloice, Florian Endel, Anna Saranti

From Probabilistic Graphical Models to Graph Machine Learning

Contact: andreas.holzinger AT tuwien.ac.at

https://human-centered.ai/lv-185-a83-machine-learning-for-health-informatics-class-of-2021

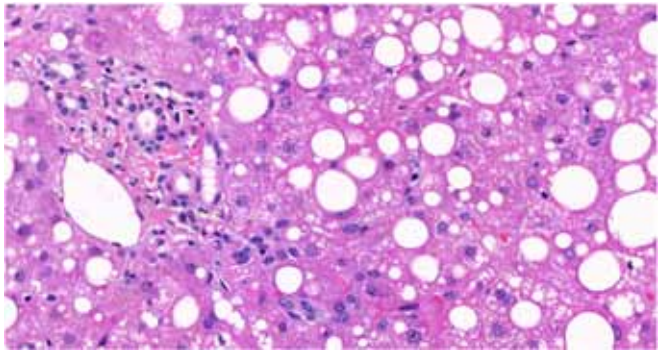
What will we learn today ?

- 00 Reflection
- 01 Machine Learning on Graphs
- 02 Graph Neural Networks
- 03 Knowledge Graph Embeddings
- 04 Applications
- 05 Graph metrics
- 06 Graph building

00 Reflection

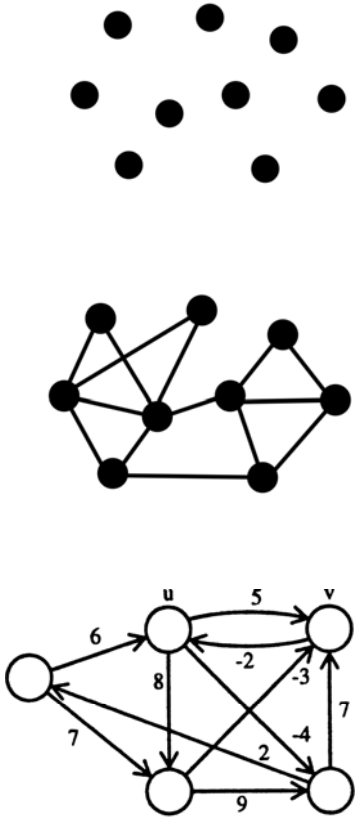
Warm-up Quiz

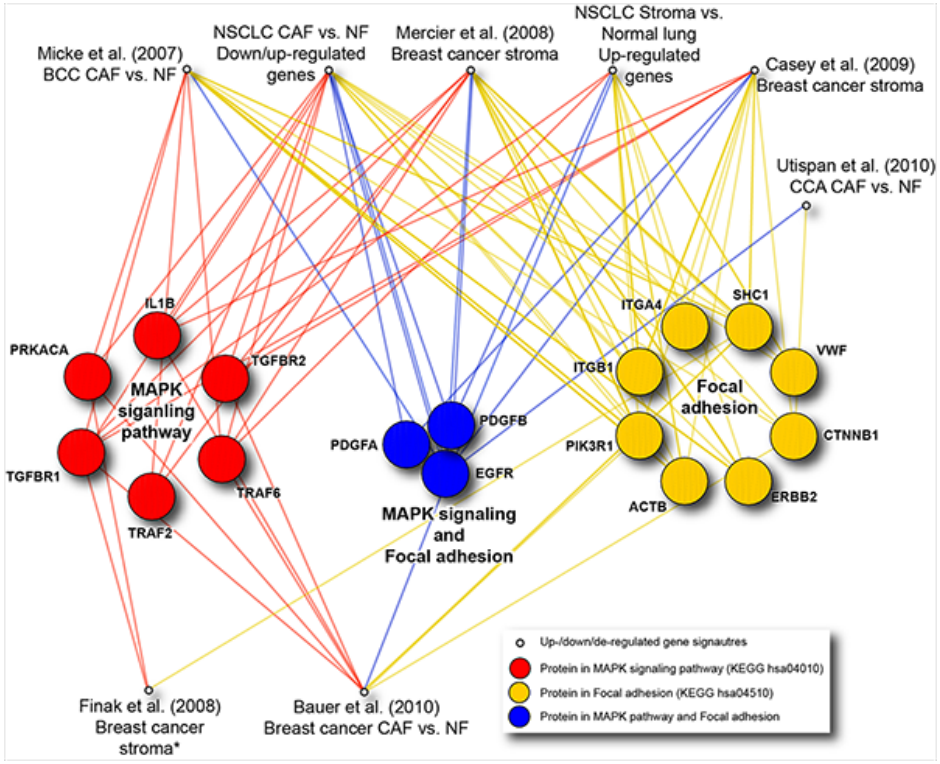
- 1) What does it mean for a computer to understand a question?
- 2) Why do humans outperform computers at in many tasks?
- 3) What is a concept ?
- 4) Why is correlation not causation ?
- 5) What is inductive reasoning ?
- 6) What is empirical inference ?



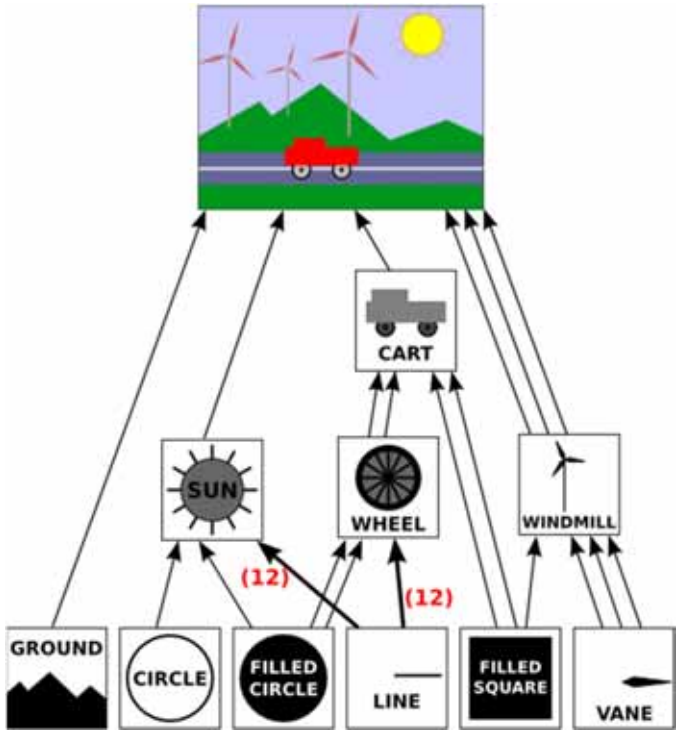
- When I tell students this, they don't believe me, so it's not me saying it, it's Jure Leskovec from Stanford:

What do you see here ?

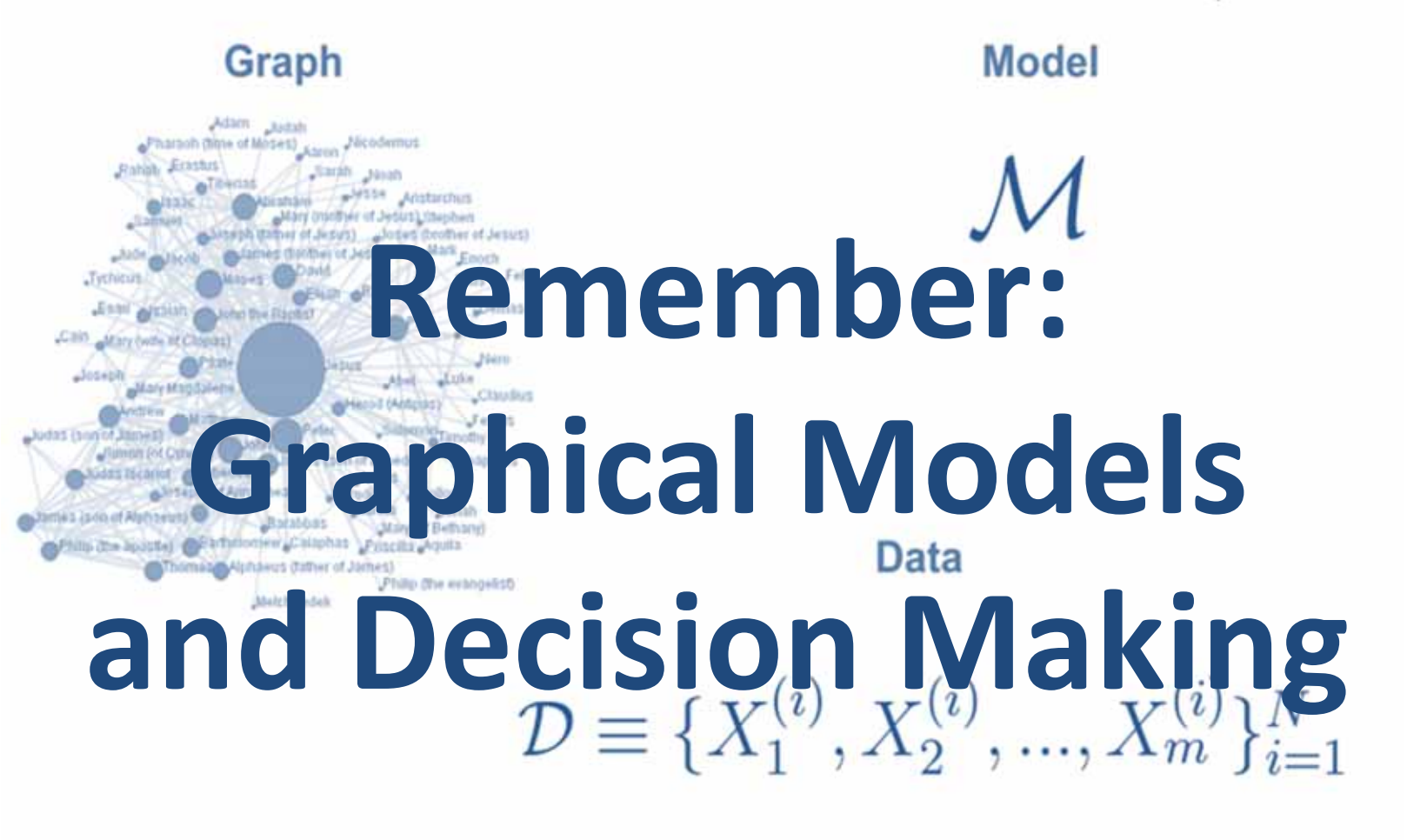




https://www.cs.toronto.edu/~juris

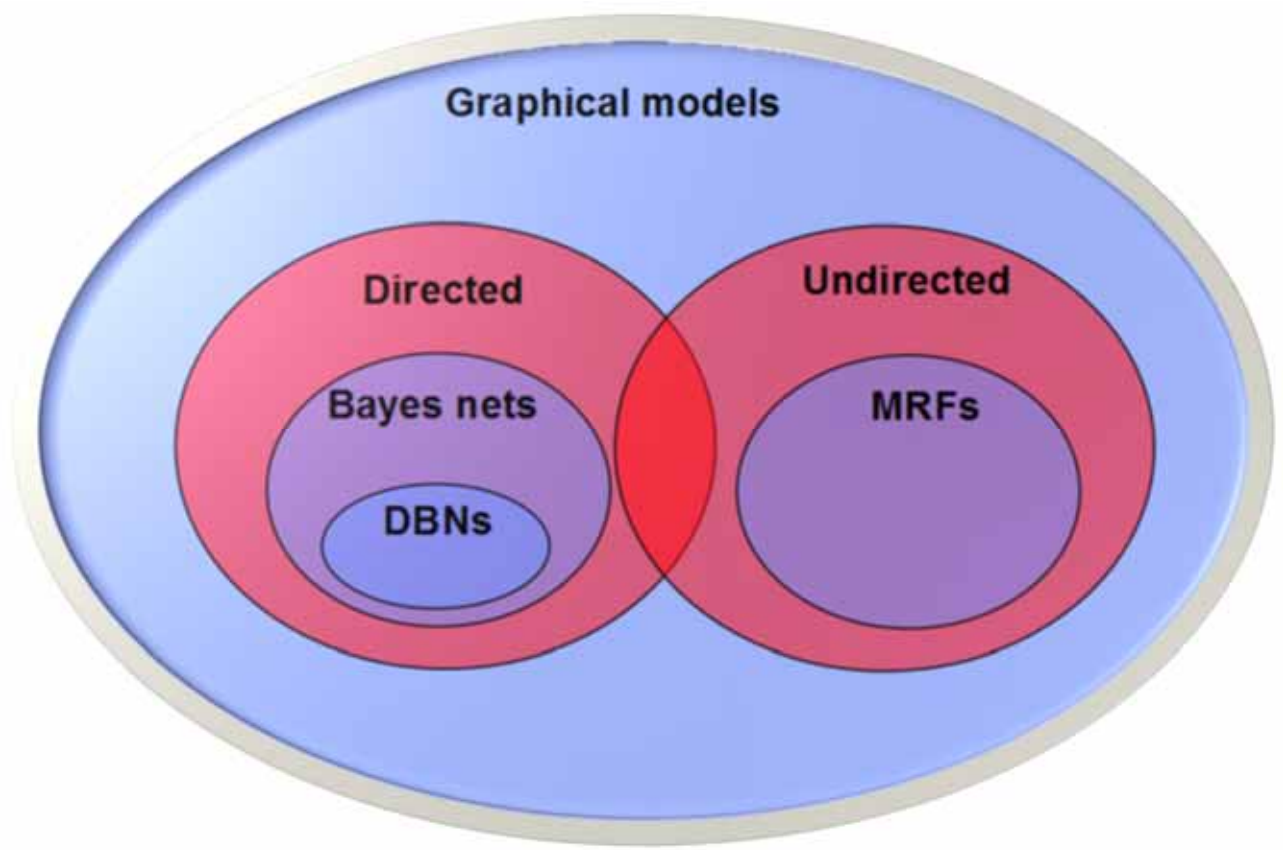


David J Eck (2016). *Introduction to Computer Graphics*, Online Book.
http://math.hws.edu/graphicsbook/c2/s4.html



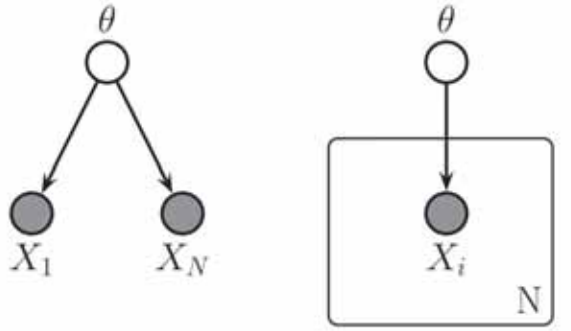
Remember: Graphical Models and Decision Making

What Classes of Graphical Models do we know ?

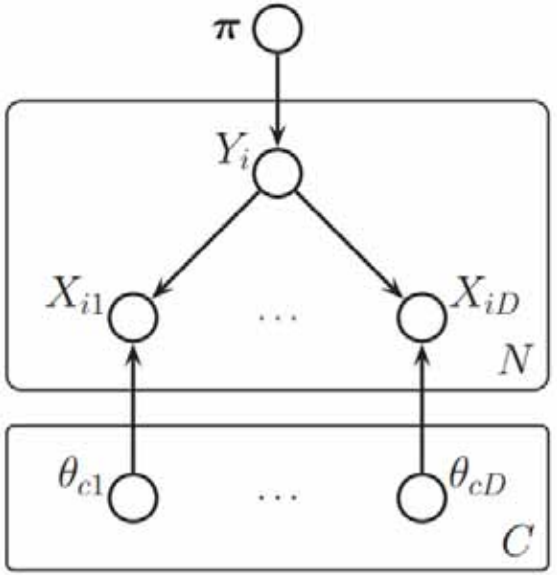


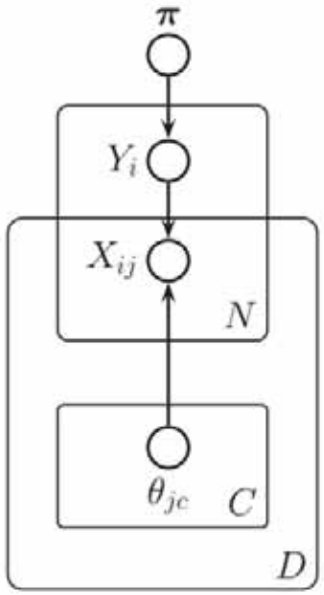
Murphy, K. P. 2012. Machine learning: a probabilistic perspective, Cambridge (MA), MIT press.

Naïve Bayes classifier as DGM (single/nested plates)



$$p(\boldsymbol{\theta}, \mathcal{D}) = p(\boldsymbol{\theta}) \left[\prod_{i=1}^{N} p(\mathbf{x}_i | \boldsymbol{\theta}) \right]$$





Murphy, K. P. 2012. Machine learning: a probabilistic perspective, Cambridge (MA), MIT press.

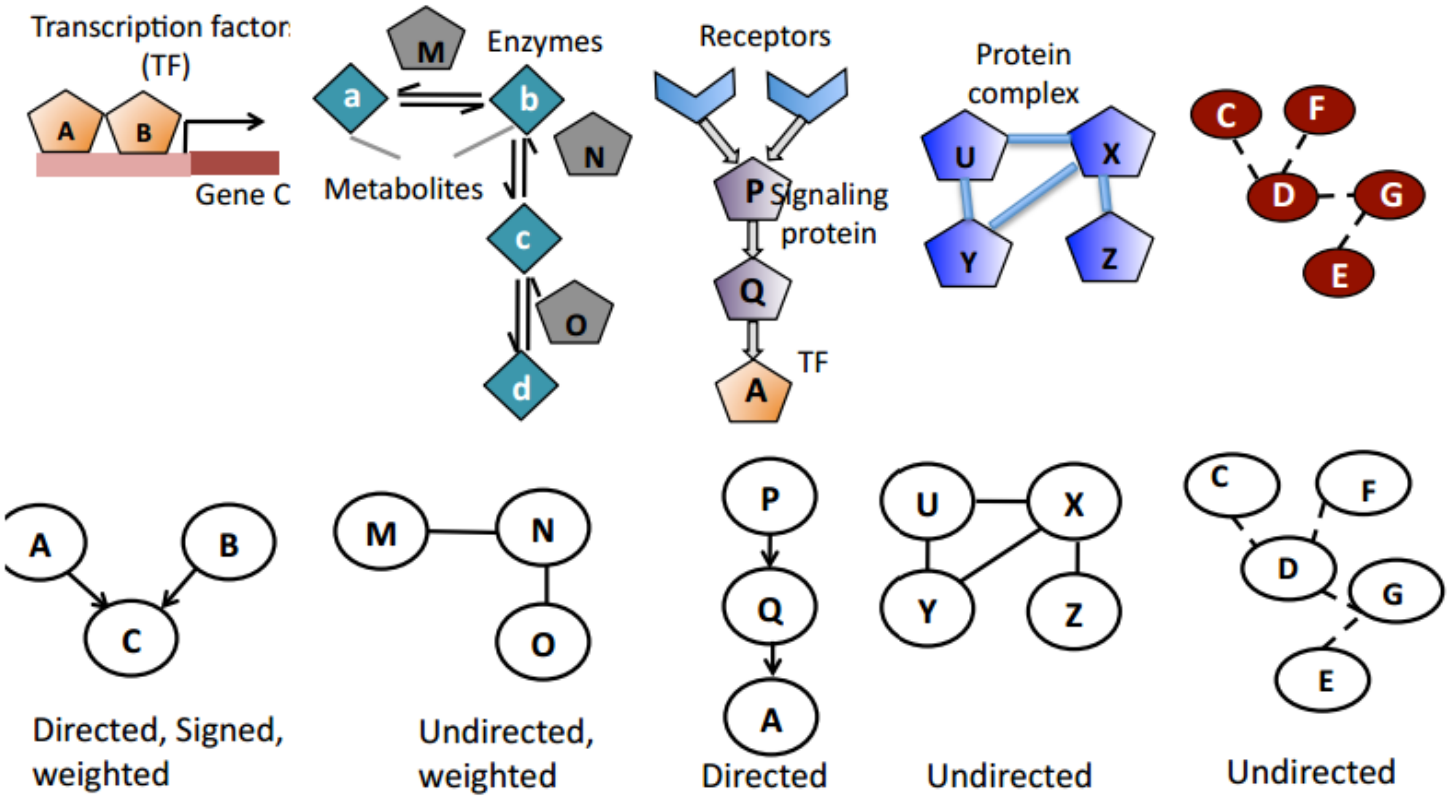
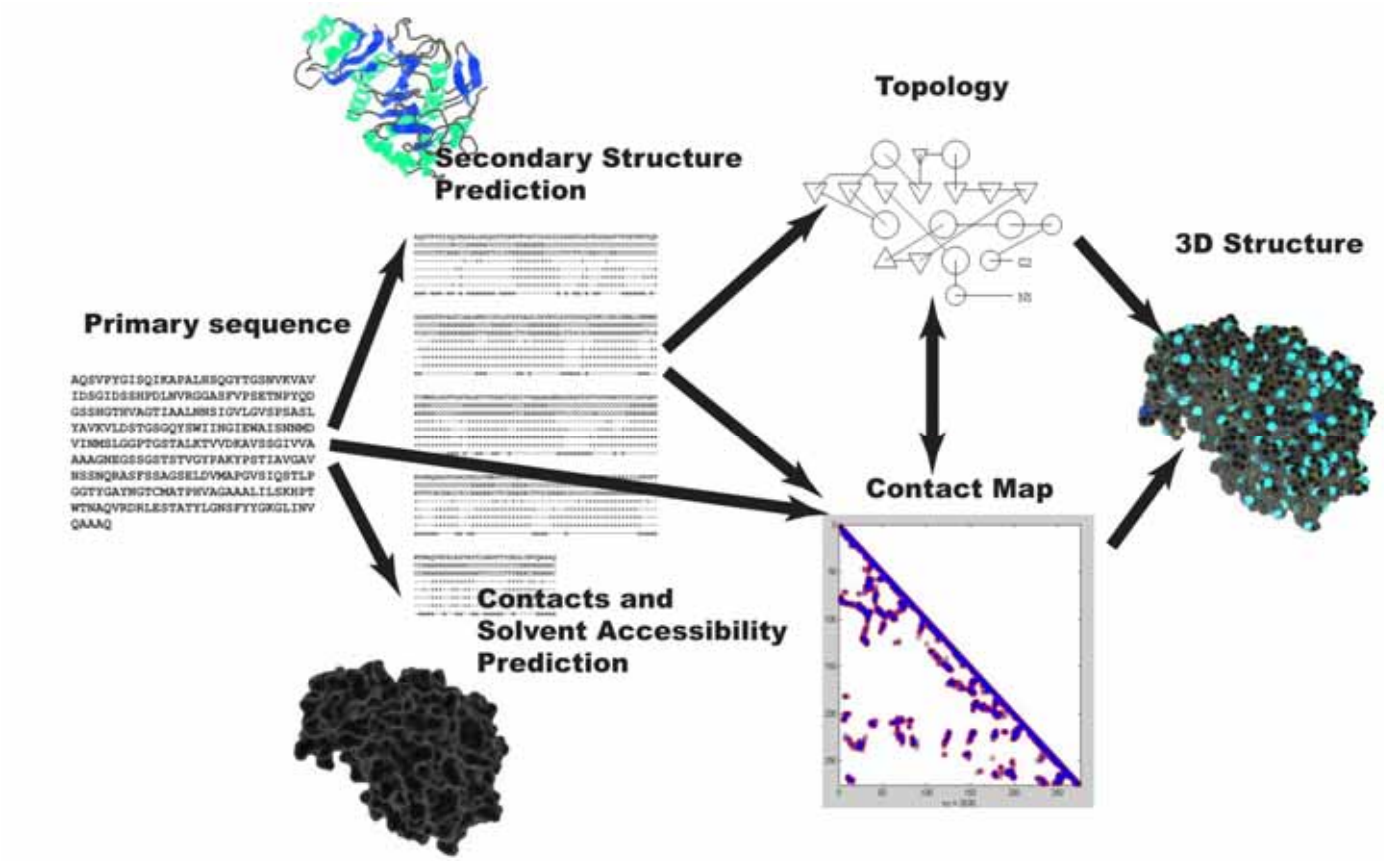


Image credit to Anna Goldenberg, Toronto

- Medicine is an extremely complex application domain – dealing most of the time with uncertainties -> **probable information!**
- When we have big data but little knowledge automatic ML can help to gain insight:
- **Structure learning and prediction in large-scale biomedical networks with probabilistic graphical models**
- If we have little data and deal with NP-hard problems we still need the human-in-the-loop

How to we get from structure to function prediction ?



Baldi, P. & Pollastri, G. 2003. The principled design of large-scale recursive neural network architectures--dag-rnns and the protein structure prediction problem. The Journal of Machine Learning Research, 4, 575-602.

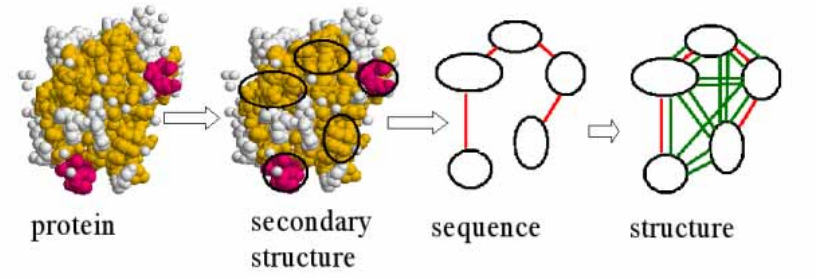
Protein Network Inference

- Hypothesis: most biological functions involve the interactions between many proteins, and the complexity of living systems arises as a result of such interactions.
- In this context, the problem of inferring a global protein network for a given organism,
- - using all (genomic) data of the organism,
- is one of the main challenges in computational biology

Yamanishi, Y., Vert, J.-P. & Kanehisa, M. 2004. Protein network inference from multiple genomic data: a supervised approach. Bioinformatics, 20, (suppl 1), i363-i370.

Problem: Is Graph Isomorphism NP-complete ?

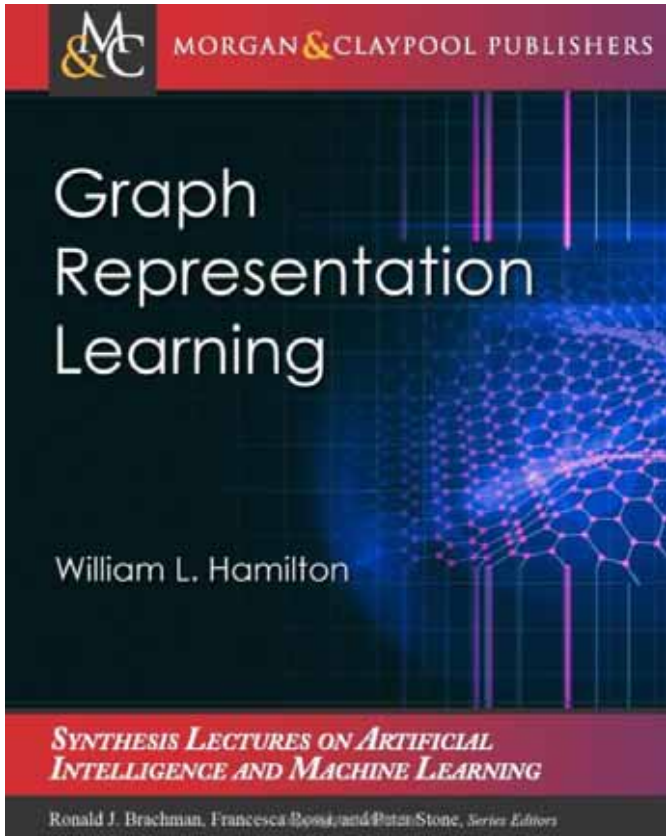
Borgwardt, K. M., Ong, C. S., Schönauer, S., Vishwanathan, S., Smola, A. J. & Kriegel, H.-P. 2005. Protein function prediction via graph kernels. Bioinformatics, 21, (suppl 1), i47-i56.

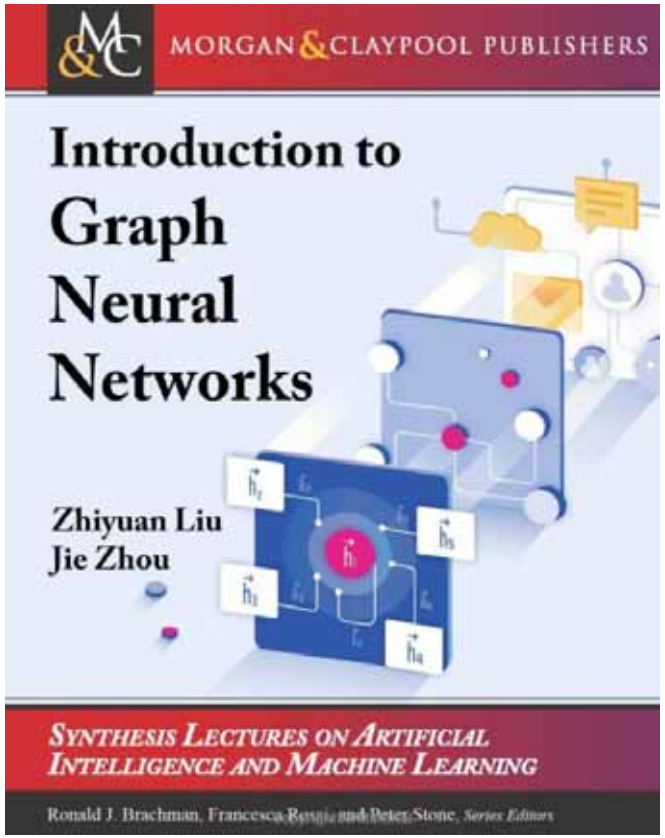


- Important for health informatics: Discovering relationships between biological components
- Unsolved problem in computer science:
- Can the graph isomorphism problem be solved in polynomial time?
  - So far, no polynomial time algorithm is known.
  - It is also not known if it is NP-complete
  - We know that subgraph-isomorphism is NP-complete

01 Machine Learning on Graphs

Which books are recommendable ?





https://www.cs.mcgill.ca/~wlh/grl_book/

What is the naive idea of machine learning with graphs ?

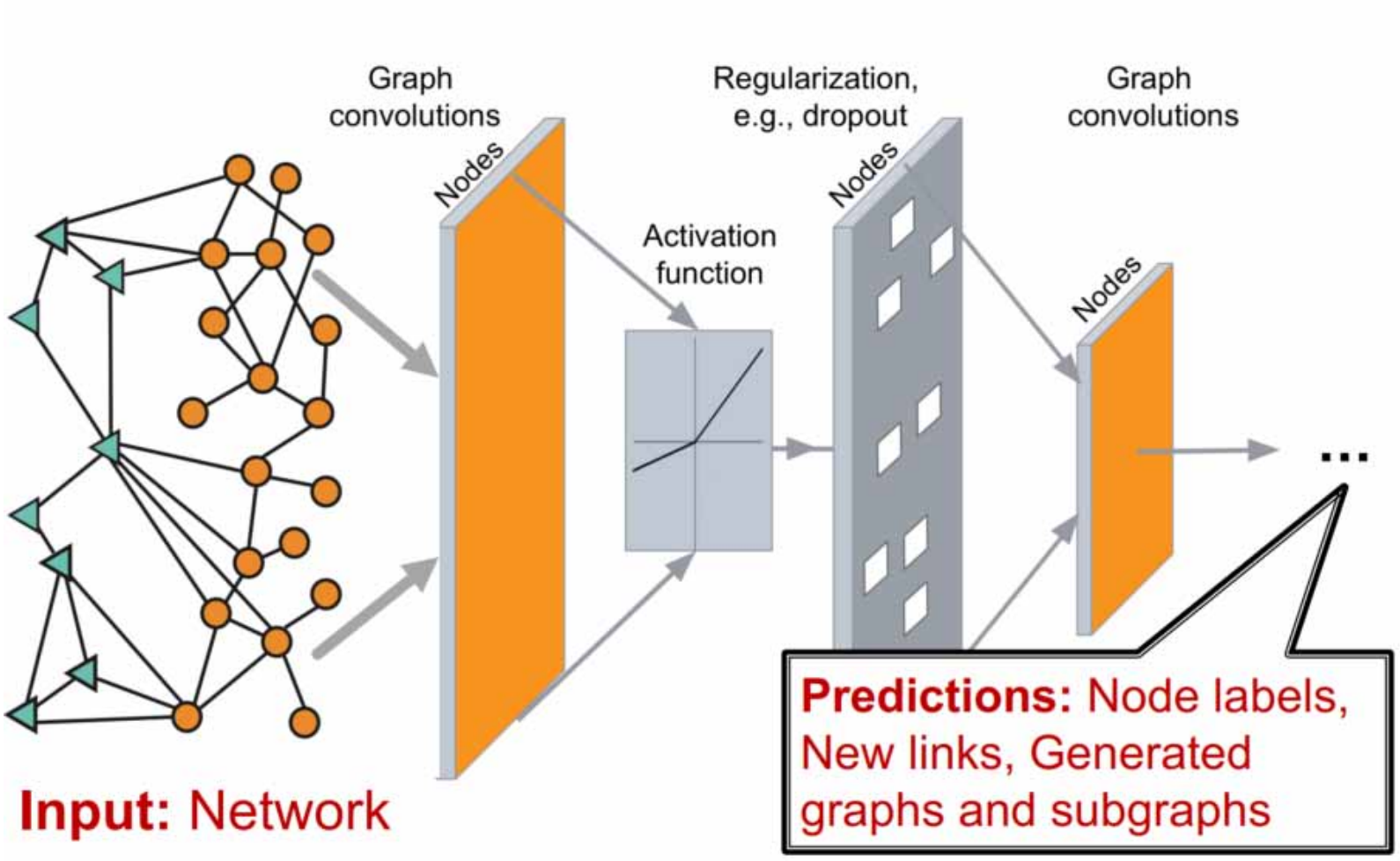


Image Source: Jure Leskovec, Deep Learning in Graphs, CS 224, Stanford Machine Learning with Graphs

How can we learn a mapping function f of a graph ?

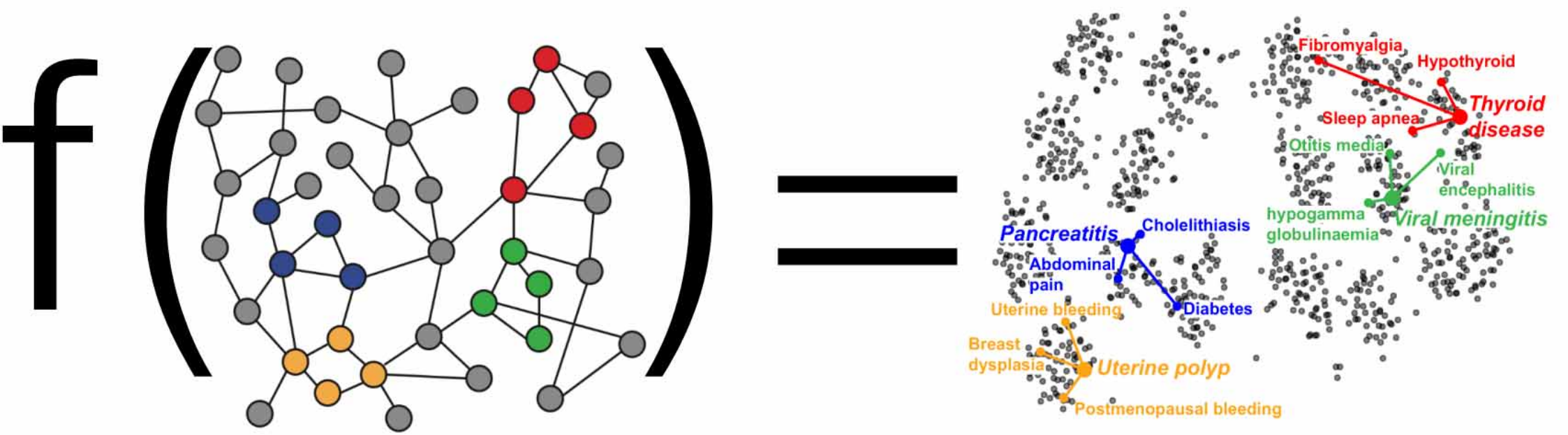


Image Source: Jure Leskovec, Deep Learning in Graphs, CS 224, Stanford Machine Learning with Graphs

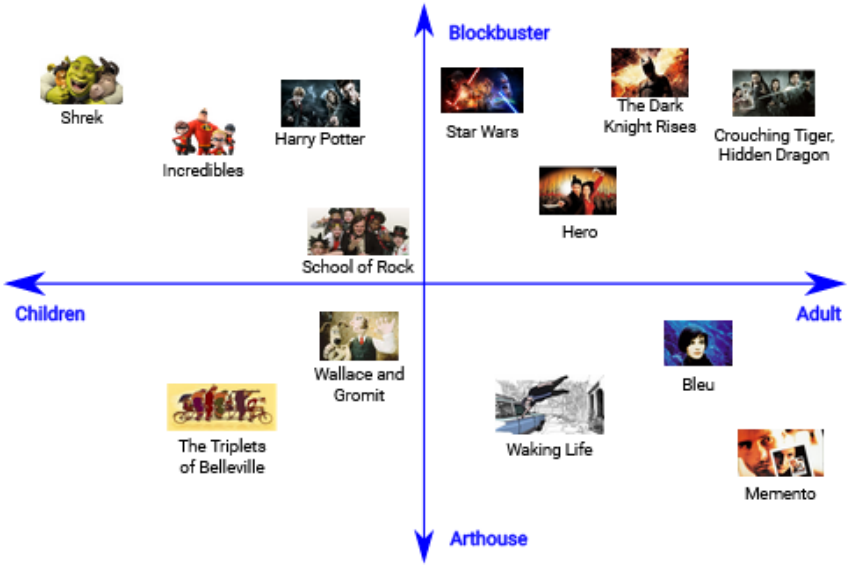
What is an embedding ?

Goal: representing each word (or common phrase) w as a d-dimensional word vector $\vec{w} \in \mathbb{R}^d$

1d

$$\overrightarrow{\text{man}} - \overrightarrow{\text{woman}} \approx \overrightarrow{\text{king}} - \overrightarrow{\text{queen}}$$

2d



Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings

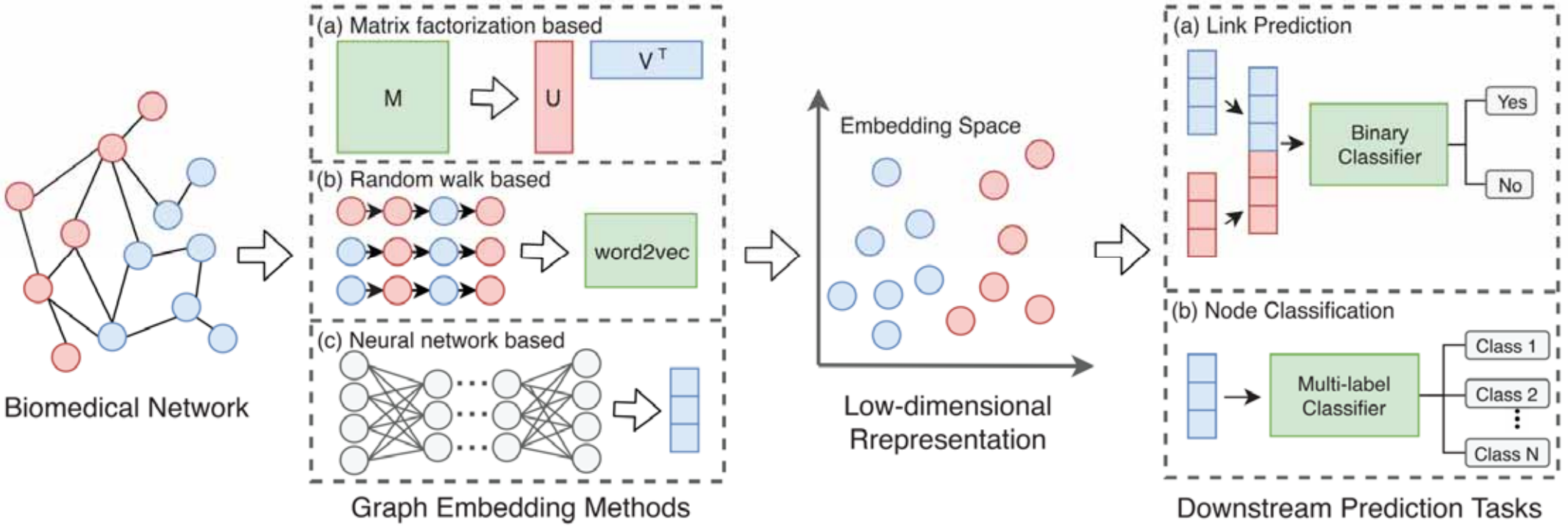
Tolga Bolukbasi[1], Kai-Wei Chang[2], James Zou[2], Venkatesh Saligrama[1,2], Adam Kalai[2]
[1]Boston University, 8 Saint Mary's Street, Boston, MA
[2]Microsoft Research New England, 1 Memorial Drive, Cambridge, MA
tolgab@bu.edu, kw@kwchang.net, jamesyzou@gmail.com, srv@bu.edu, adam.kalai@microsoft.com

Abstract

The blind application of machine learning runs the risk of amplifying biases present in data. Such a danger is facing us with *word embedding*, a popular framework to represent text data as vectors which has been used in many machine learning and natural language processing tasks. We show that even word embeddings trained on Google News articles exhibit female/male gender stereotypes to a disturbing extent. This raises concerns because their widespread use, as we describe, often tends to amplify these biases. Geometrically, gender bias is first shown to be captured by a direction in the word embedding. Second, gender neutral words are shown to be linearly separable from gender definition words in the word embedding. Using these properties, we provide a methodology for modifying an embedding to remove gender stereotypes, such as the association between between the words *receptionist* and *female*, while maintaining desired associations such as between the words *queen* and *female*. We define metrics to quantify both direct and indirect gender biases in embeddings, and develop algorithms to "debias" the embedding. Using crowd-worker evaluation as well as standard benchmarks, we empirically demonstrate that our algorithms significantly reduce gender bias in embeddings while preserving the its useful properties such as the ability to cluster related concepts and to solve analogy tasks. The resulting embeddings can be used in applications without amplifying gender bias.

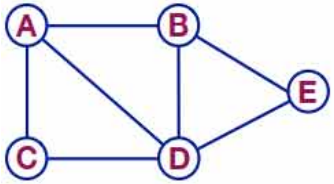
https://developers.google.com/machine-learning/crash-course/embeddings/video-lecture

How does a graph embedding pipeline for health informatics look like ?



Xiang Yue, Zhen Wang, Jingong Huang, Srinivasan Parthasarathy, Soheil Moosavinasab, Yungui Huang, Simon M Lin, Wen Zhang, Ping Zhang & Huan Sun (2020). Graph embedding on biomedical networks: methods, applications and evaluations. Bioinformatics, 36, (4), 1241-1251.

- Join adjacency matrix and features and take it as input into DL
- A naive idea, because this is problematic due to O(|V|) parameters, not applicable to graphs of different sizes and sensitive to node ordering



	A	B	C	D	E	Feat	
A	0	1	1	1	0	1	0
B	1	0	0	1	1	0	0
C	1	0	0	1	0	0	1
D	1	1	1	0	1	1	1
E	0	1	0	1	0	1	0

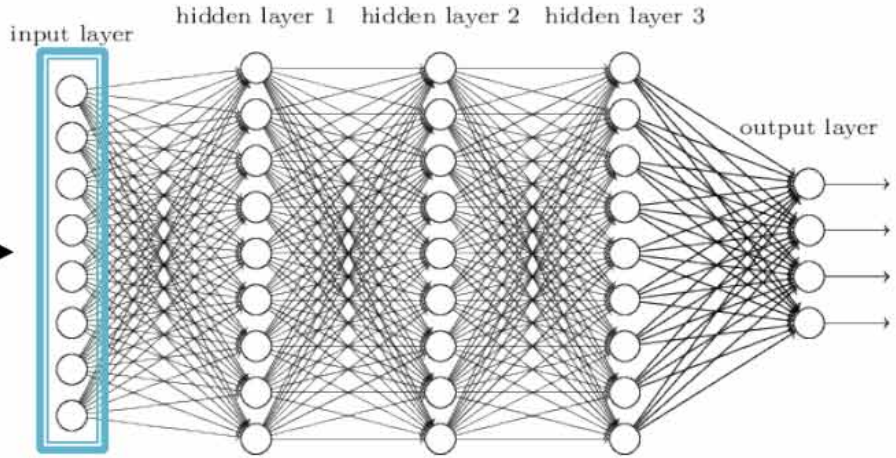
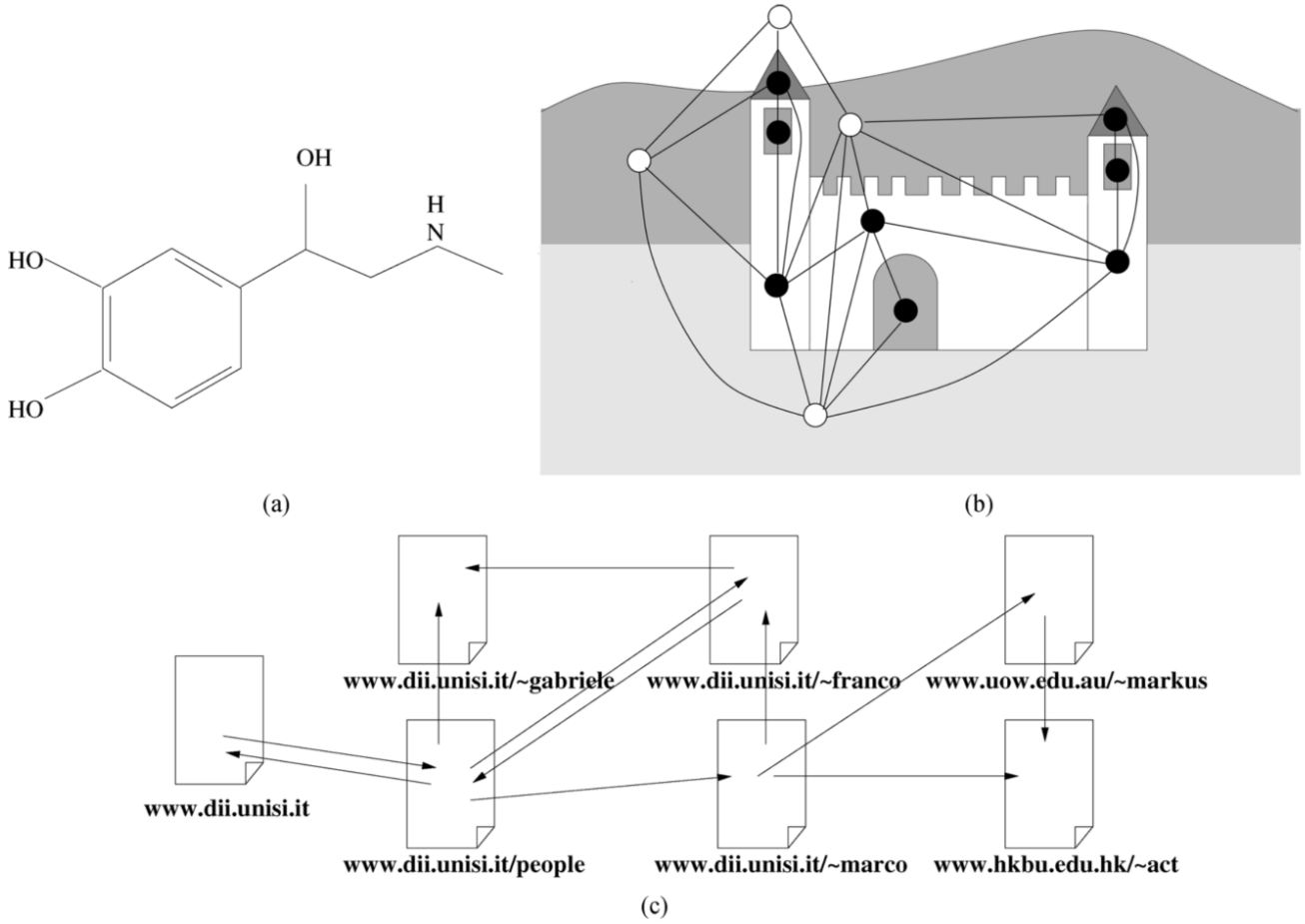


Image Source: Jure Leskovec, Deep Learning in Graphs, CS 224, Stanford Machine Learning with Graphs

02 Graph Neural Networks

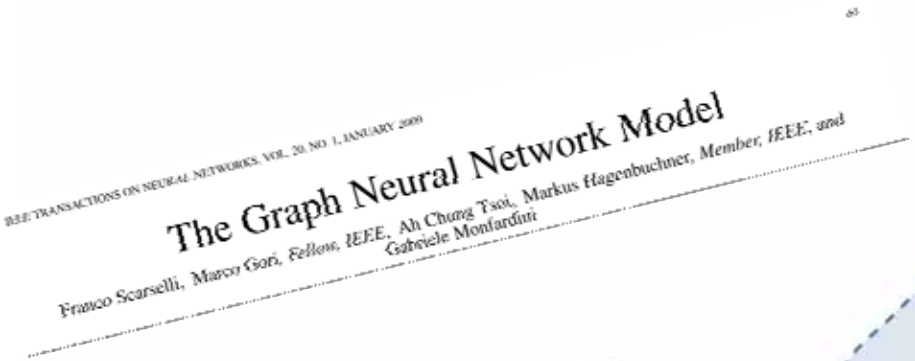
Remember: How can we represent a graph ?



Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner & Gabriele Monfardini (2008). The graph neural network model. IEEE Transactions on Neural Networks, 20, (1), 61-80, doi:10.1109/TNN.2008.2005605.

What is the intuitive idea of graph neural networks ?

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner & Gabriele Monfardini (2008). The graph neural network model. IEEE Transactions on Neural Networks, 20, (1), 61-80, doi:10.1109/TNN.2008.2005605.

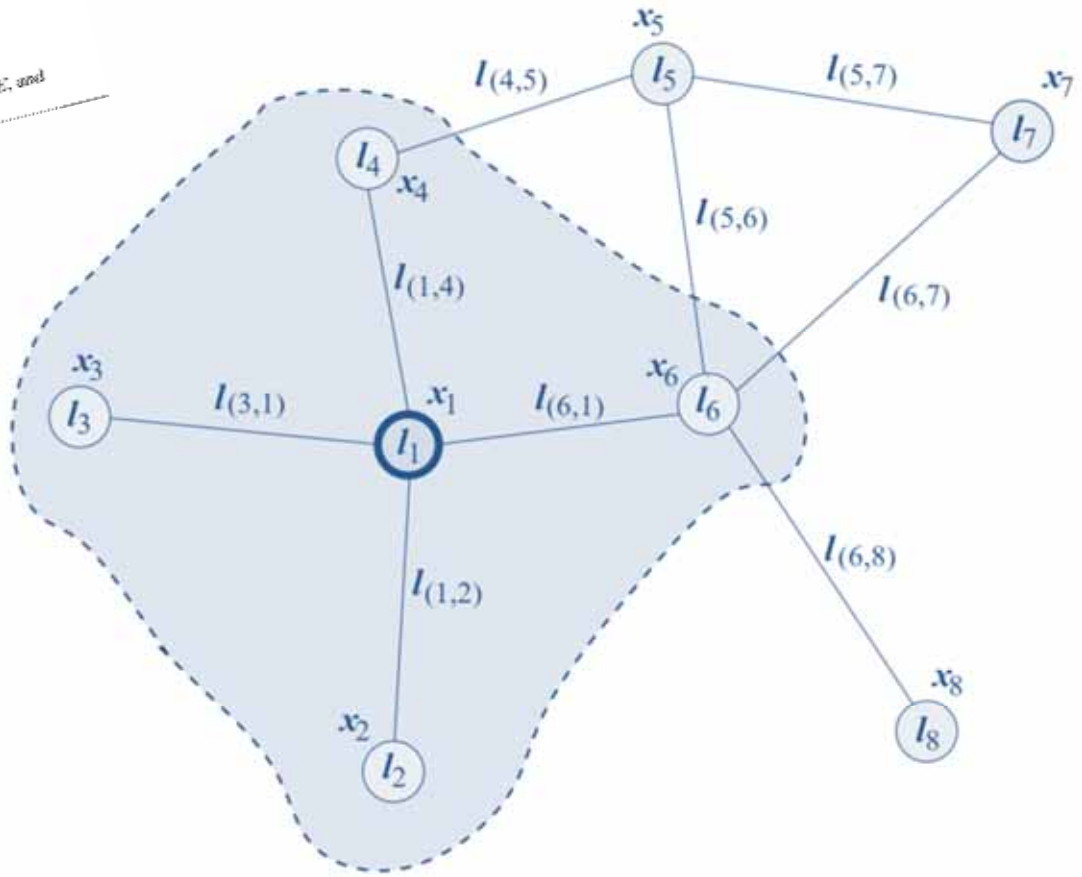


$$\boldsymbol{x}_n \in \mathbb{R}^s$$

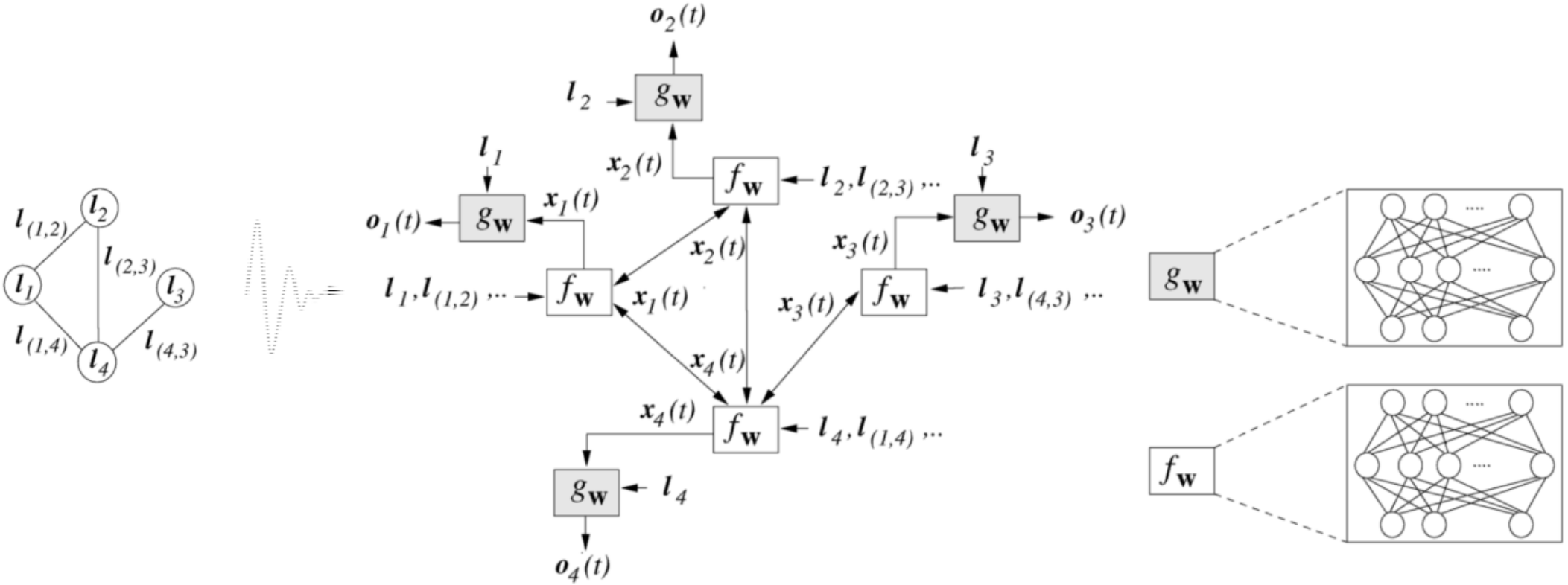
$$\begin{aligned} \boldsymbol{x}_n &= f_{\boldsymbol{w}}(\boldsymbol{l}_n, \boldsymbol{l}_{\mathrm{co}[n]}, \boldsymbol{x}_{\mathrm{ne}[n]}, \boldsymbol{l}_{\mathrm{ne}[n]}) \\ \boldsymbol{o}_n &= g_{\boldsymbol{w}}(\boldsymbol{x}_n, \boldsymbol{l}_n), \end{aligned}$$

$$\begin{aligned} \boldsymbol{x} &= F_{\boldsymbol{w}}(\boldsymbol{x}, \boldsymbol{l}) \\ \boldsymbol{o} &= G_{\boldsymbol{w}}(\boldsymbol{x}, \boldsymbol{l}_N) \end{aligned}$$

$$\varphi_{\boldsymbol{w}} : \mathcal{D} \rightarrow \mathbb{R}^m$$

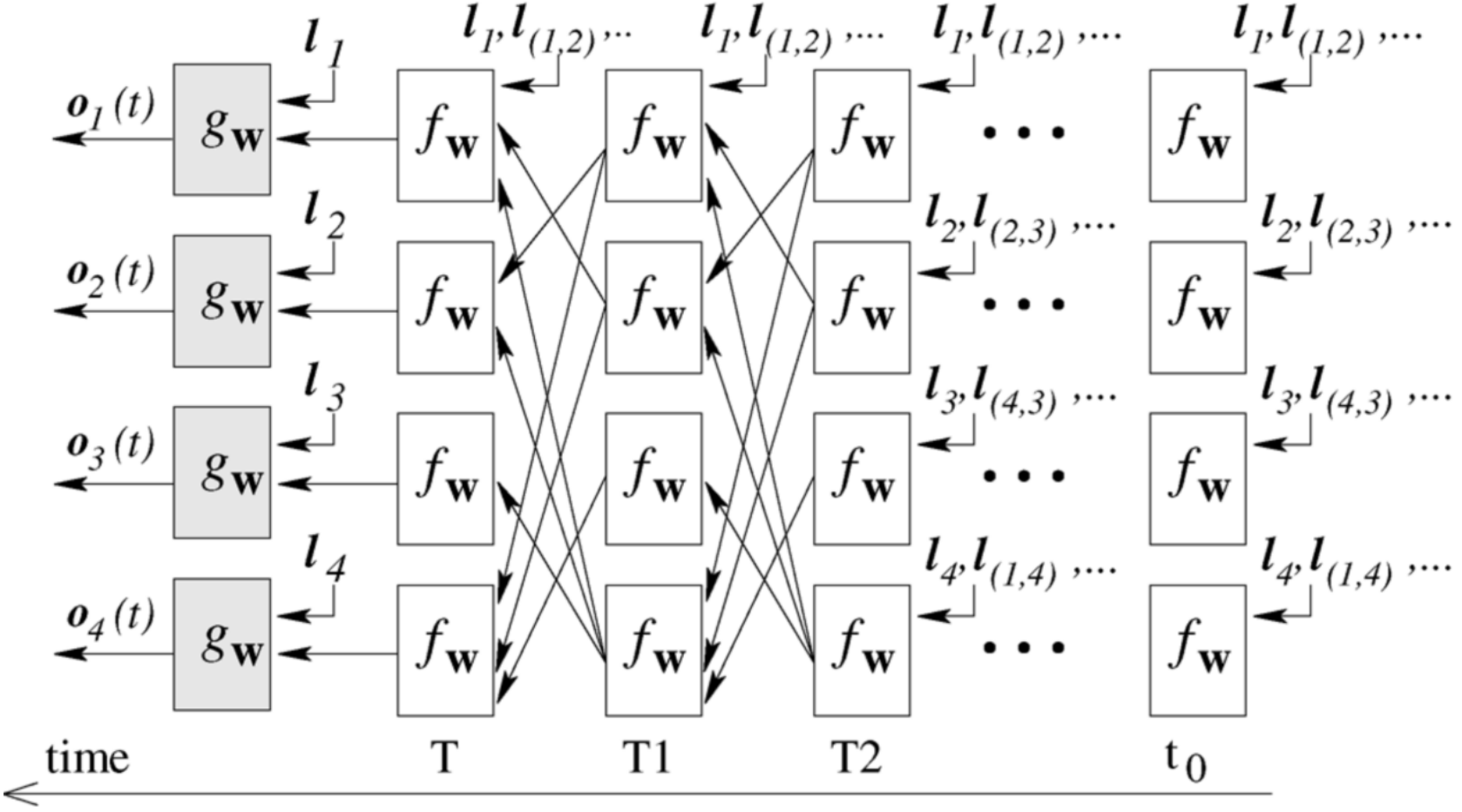


How does the corresponding encoding network look like ?



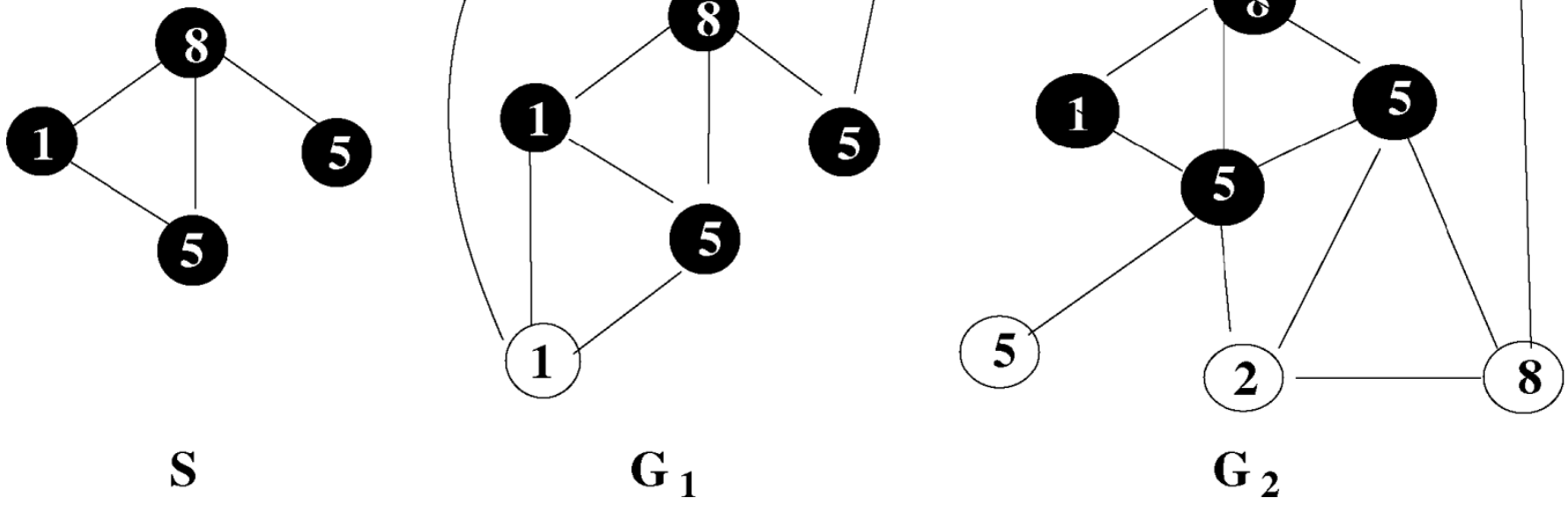
Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner & Gabriele Monfardini (2008). The graph neural network model. IEEE Transactions on Neural Networks, 20, (1), 61-80, doi:10.1109/TNN.2008.2005605

How does the network look after unfolding the encoding network ?



Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner & Gabriele Monfardini (2008). The graph neural network model. IEEE Transactions on Neural Networks, 20, (1), 61-80, doi:10.1109/TNN.2008.2005605.

Why is the subgraph isomorphism problem so difficult to solve ?



Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner & Gabriele Monfardini (2008). The graph neural network model. IEEE Transactions on Neural Networks, 20, (1), 61-80, doi:10.1109/TNN.2008.2005605.

What are the limitations and pitfalls of GNN ?

- computationally very intensive !
- the same parameters are used in every epoch (shallow approach)
- with each iteration each node increase the relations and number of weights (dynamic graphs)
- When used for non-standardized (e.g. text) data it is difficult to generate graphs from the raw data -> open for research
- unclear what happens with bias -> open for research
- Scalability unclear -> open for research

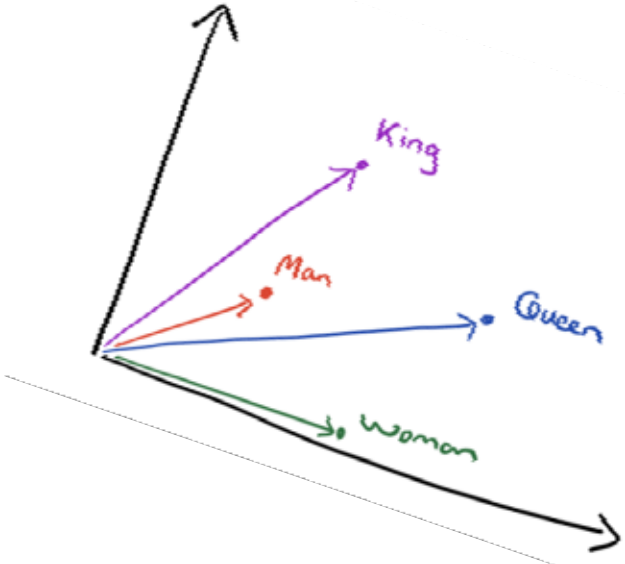
Oleksandr Shchur, Maximilian Mumme, Aleksandar Bojchevski & Stephan Günnemann (2018). Pitfalls of graph neural network evaluation. arXiv preprint arXiv:1811.05868.

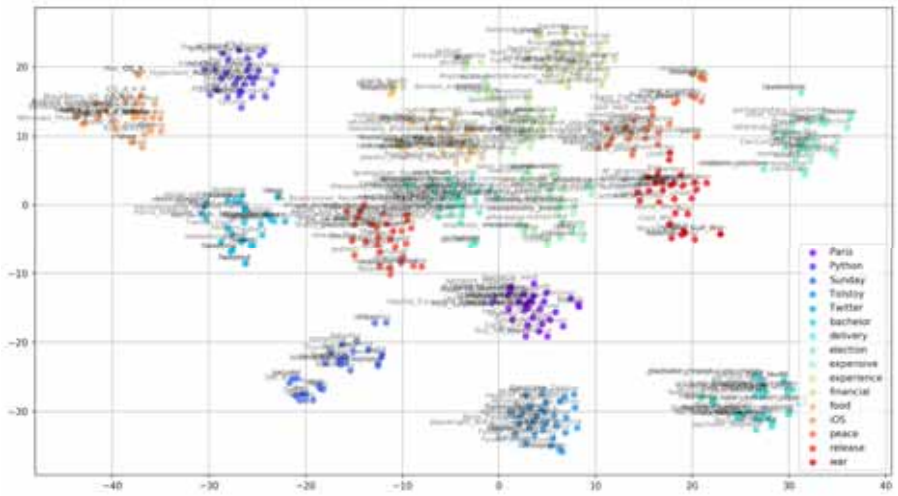
Keyulu Xu, Weihua Hu, Jure Leskovec & Stefanie Jegelka (2018). How powerful are graph neural networks? arXiv:1810.00826.

03 Knowledge Graph Embeddings

1. Visualization & Analysis of Word Vector Embeddings

- of different models
- via different techniques
- describing (explaining?) how they discriminate / cluster

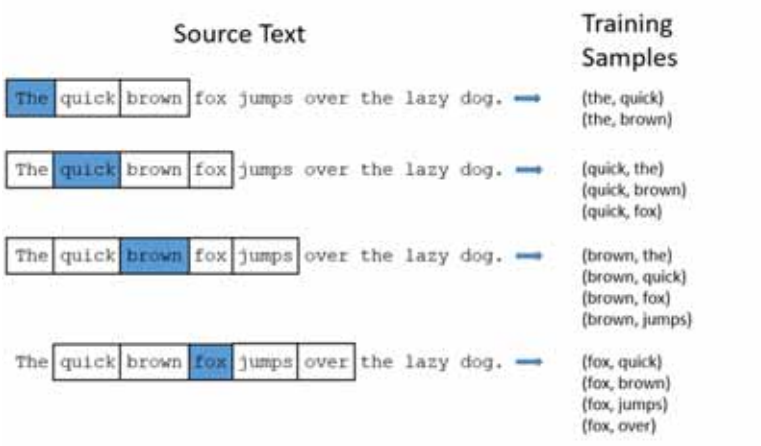


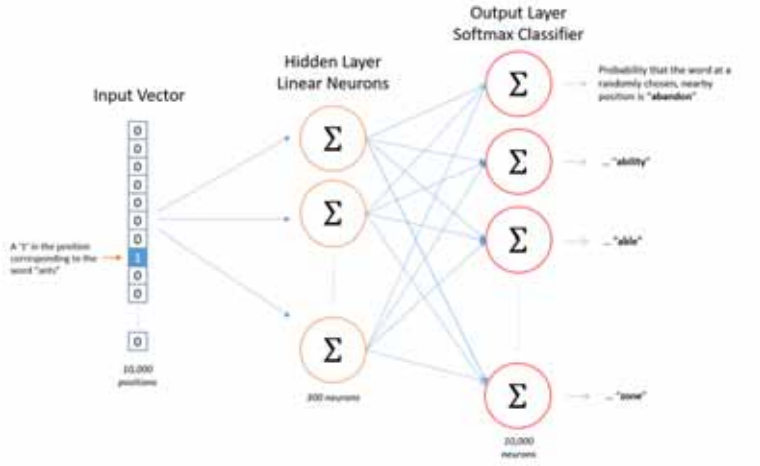


https://towardsdatascience.com/google-news-and-leo-tolstoy-visualizing-word2vec-word-embeddings-with-t-sne-11558d8bd4d

https://www.depends-on-the-definition.com/guide-to-word-vectors-with-gensim-and-keras/

*"**Word embeddings** are mathematical models that encode **word** rel ations within a **vector** space. They are created by an unsupervised training process based on **cooccurrence** information between wordsi n a large corpus"*





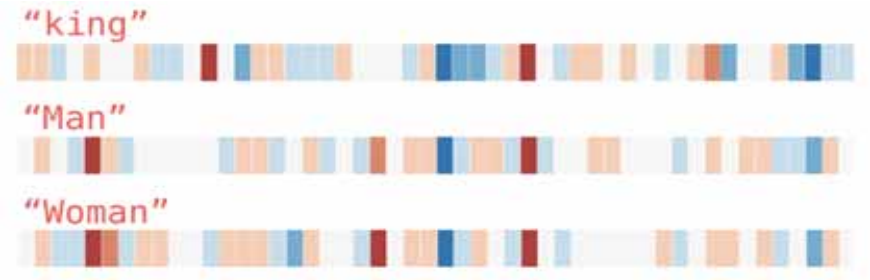
"define a fake task for the NN"
- context prediction in the case of skip-gram

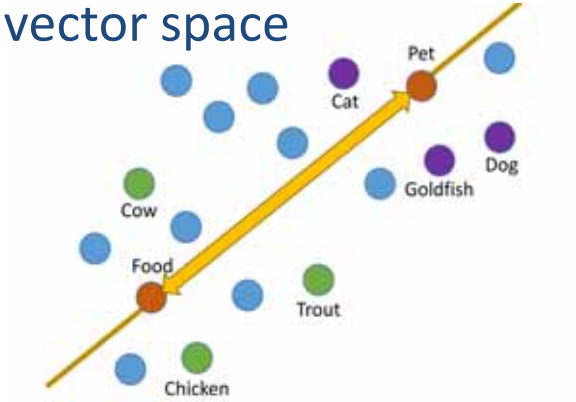
"build a shallow architecture, train & throw away the outputs"
- we only need the embeddings

http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/

Tasks:

- learn how to interpret similarity in the vector space



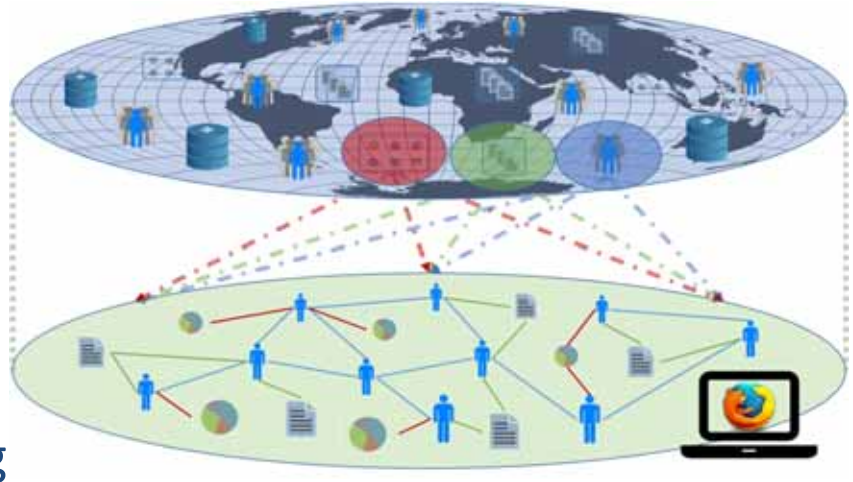


- understand the influence of source corpora on the embeddings
- analyze neighborhoods – what is (un)expected?
- predict & test consequences of changes in parameter settings / pre-processing of input data on the resulting model
- *=> "develop an intuition for embeddings as a basis for future explanations"*

Heimerl, F., & Gleicher, M. (2018). Interactive Analysis of Word Vector Embeddings. *Computer Graphics Forum*, *37*(3). doi:10.1111/cgf.13417

http://jalammar.github.io/illustrated-word2vec/

Visualization of personal recommender graphs (*"Local Sphere"*) & their change over time

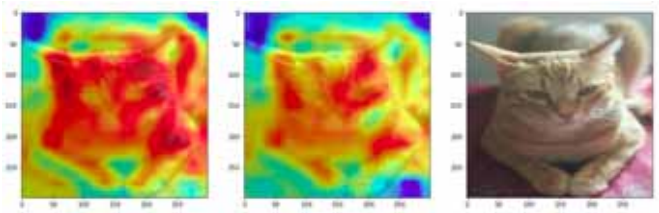


Drawing from globally available resources (e.g. a webshop database en-riched similarities) each user derives her own local sub-graph representing her context / potential interests. As she interacts with the system (explores & follows / ignores visual clues), this context is refined & the graph should respond accordingly.

Bernd Malle, Nicola Giuliani, Peter Kieseberg, and Andreas Holzinger. The More the Merrier - Federated Learning from Local Sphere Recommendations. In Machine Learning and Knowledge Extraction, IFIP CD-MAKE, Lecture Notes in Computer Science LNCS 10410, pages 367–374. Springer, Cham, 2017. doi: 10.1007/978-3-319-66808-6 24.

Project 2 – Background

- Explainability of deep learning systems is a must, but still in the early stages (except for cat pics ;-)



- non-visual and / or higher dimensional data are not intuitive to the human brain – decisions made in those spaces aren't either

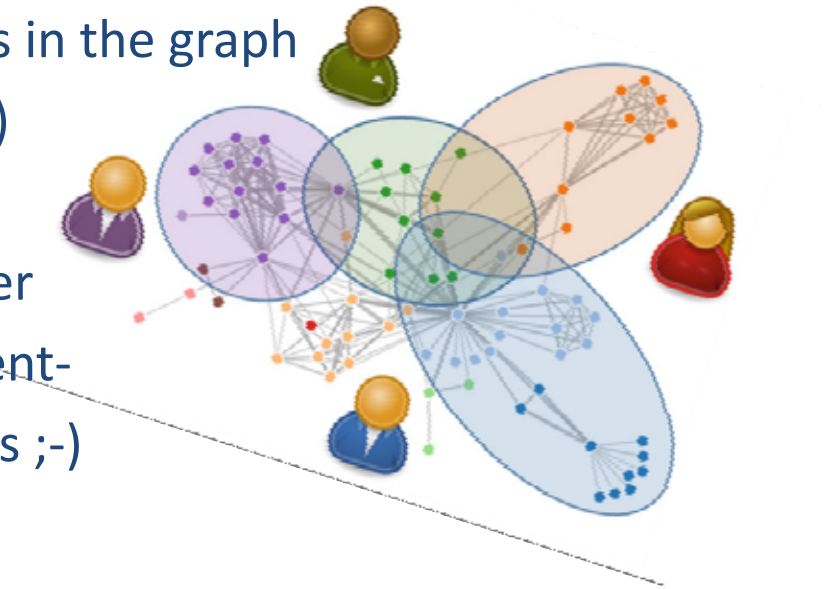
- graphs are a convenient way to break-down high-dimensional information by reducing their complexity to concepts like similarity, connection, and influence.

- understanding which graph metrics will change with user interactions will help develop an intuition about what factors in the original high-dimensional space are relevant for decisions!

https://medium.com/google-developer-experts/interpreting-deep-learning-models-for-computer-vision-f95683e23c1d

Tasks:

- Research graph visualization algorithms pertinent to recommenders (node & edge types, cluster)
- Either extend our existing (Graphinius) VIS library or decide on a different one (but make sure it's properly extensible)
- Highlight recommendations and influence factors (if available)
- Visualize continuous changes in the graph due to user interaction (vids)
- If time permits, visualize several local spheres together
- Graphs, recommender & event-stream will be provided by us ;-)

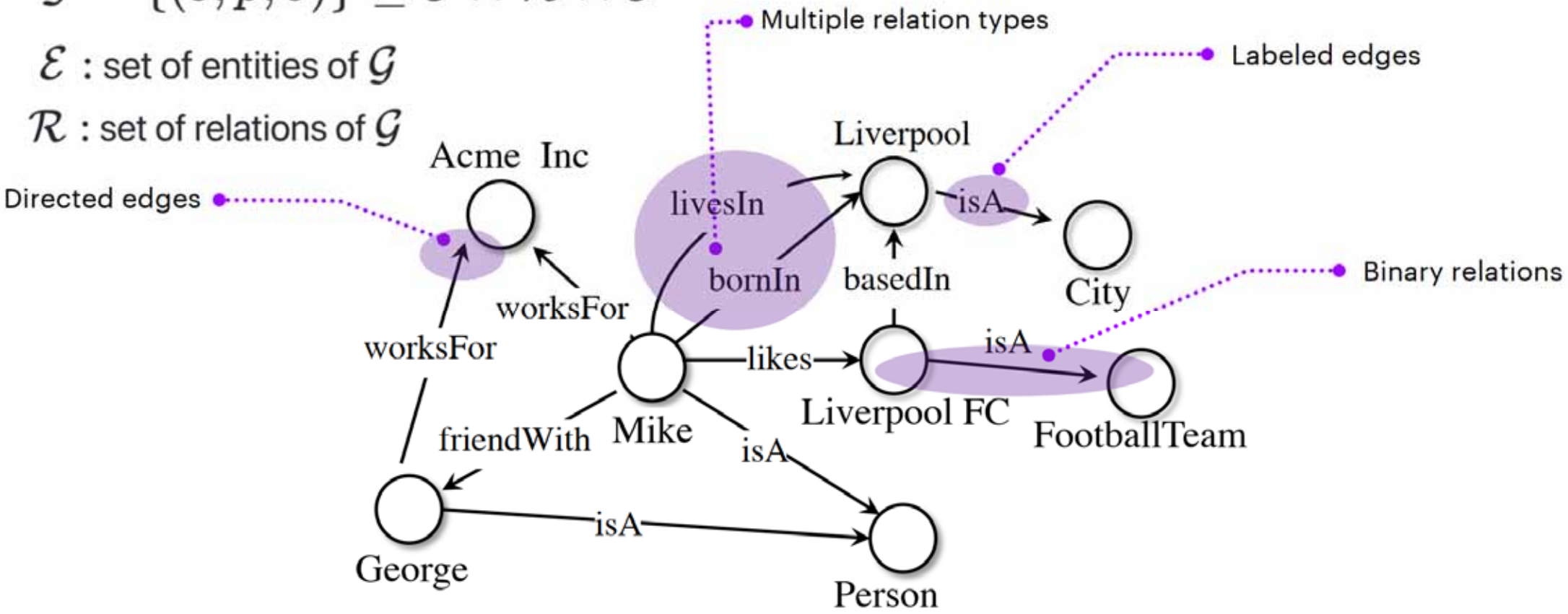


What is a Knowledge Graph ?

$$\mathcal{G} = \{(s, p, o)\} \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}$$

$\mathcal{E}$: set of entities of $\mathcal{G}$

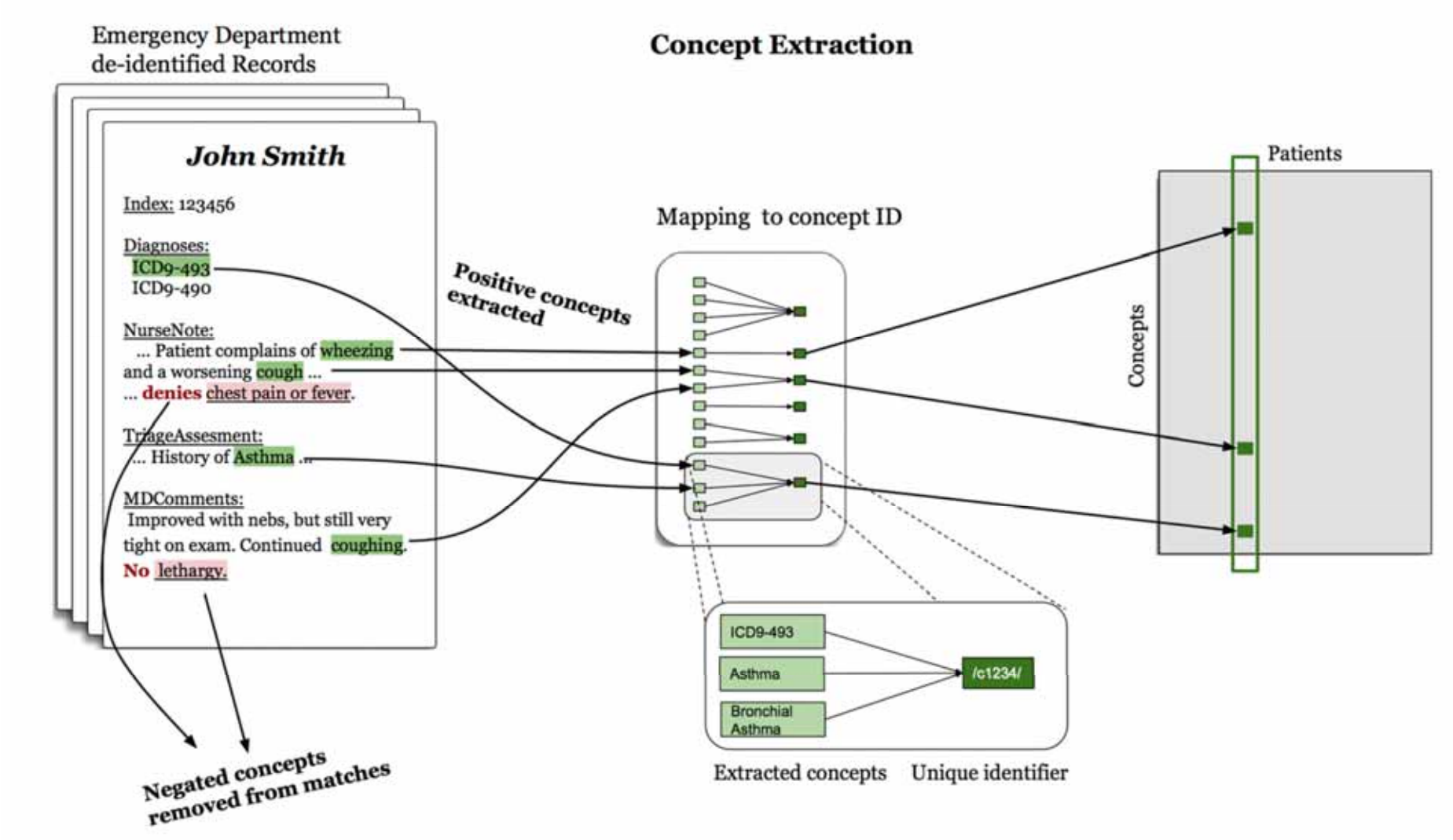
$\mathcal{R}$: set of relations of $\mathcal{G}$



Luca Costabello, Sumit Pai, Nicholas Mccarthy & Adrianna Janik (2020). Knowledge Graph Embeddings Tutorial: From Theory to Practice. doi:10.5281/zenodo.4268208.

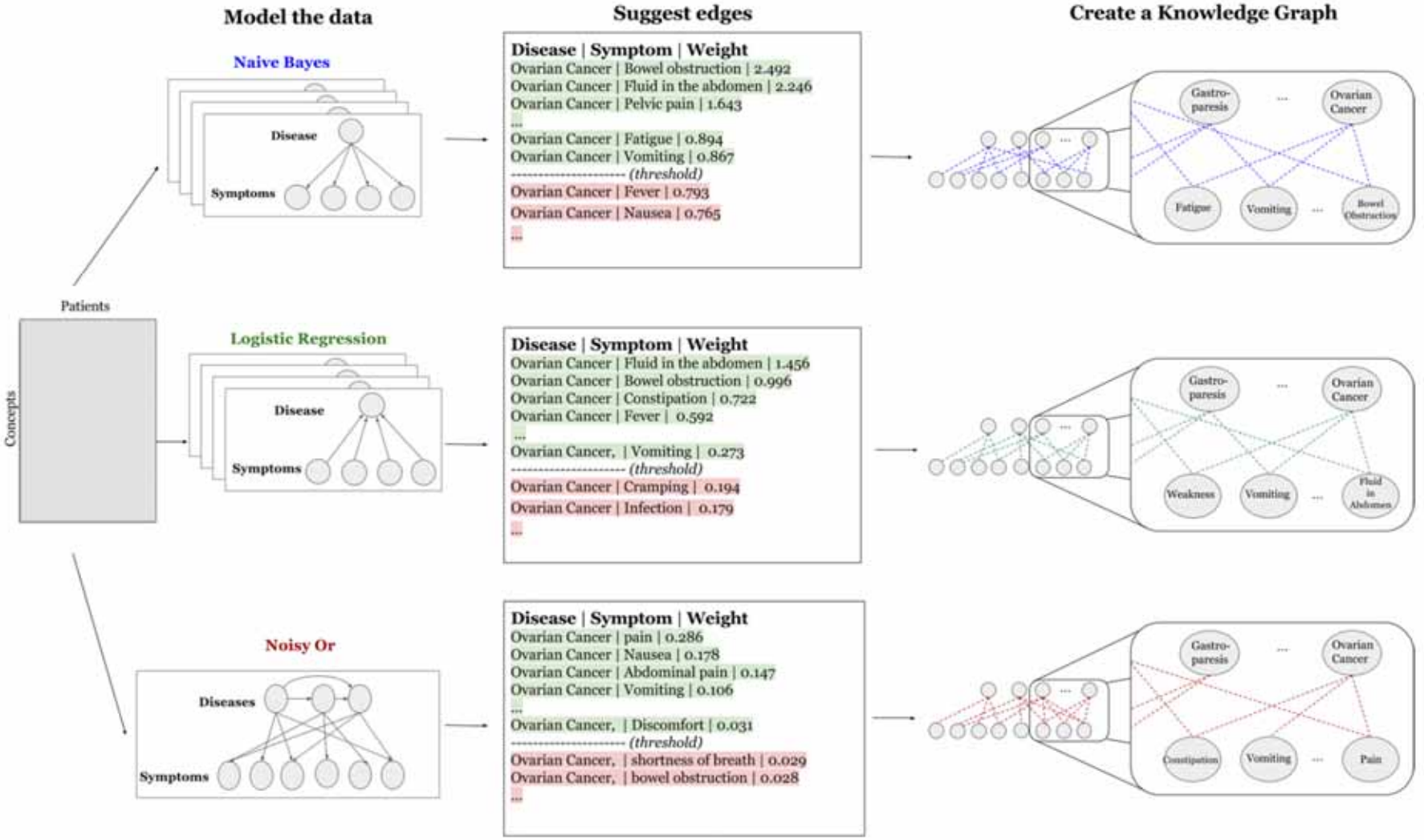
kge-tutorial-ecai2020.github.io

How does a concept extraction pipeline look like ?



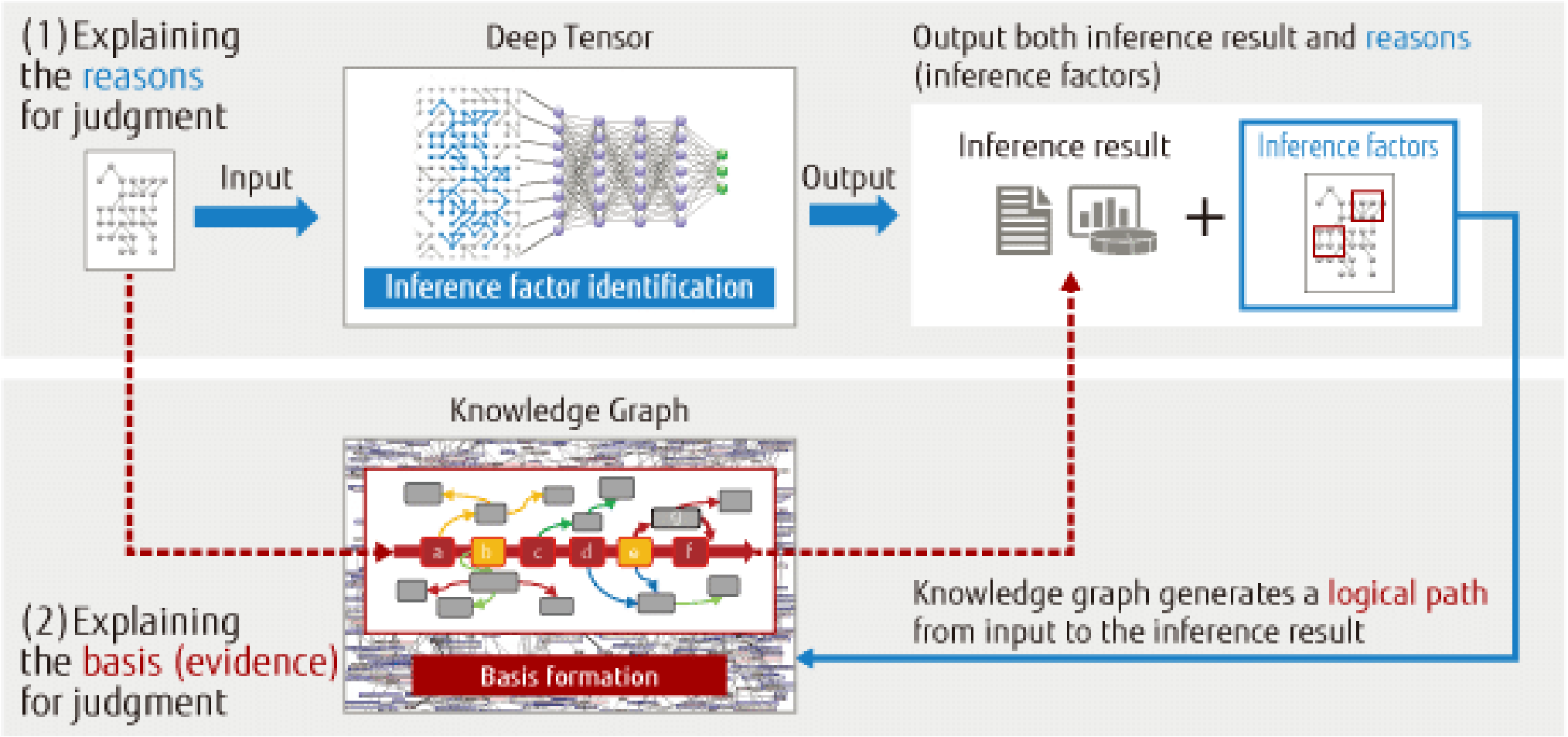
Maya Rotmensch, Yoni Halpern, Abdulhakim Tlimat, Steven Horng & David Sontag (2017). Learning a Health Knowledge Graph from Electronic Medical Records. Scientific Reports, 7, 5994, doi:10.1038/s41598-017-05778-z.

How does a concept extraction pipeline look like ?



Maya Rotmensch, Yoni Halpern, Abdulhakim Tlimat, Steven Horng & David Sontag (2017). Learning a Health Knowledge Graph from Electronic Medical Records. Scientific Reports, 7, 5994, doi:10.1038/s41598-017-05778-z.

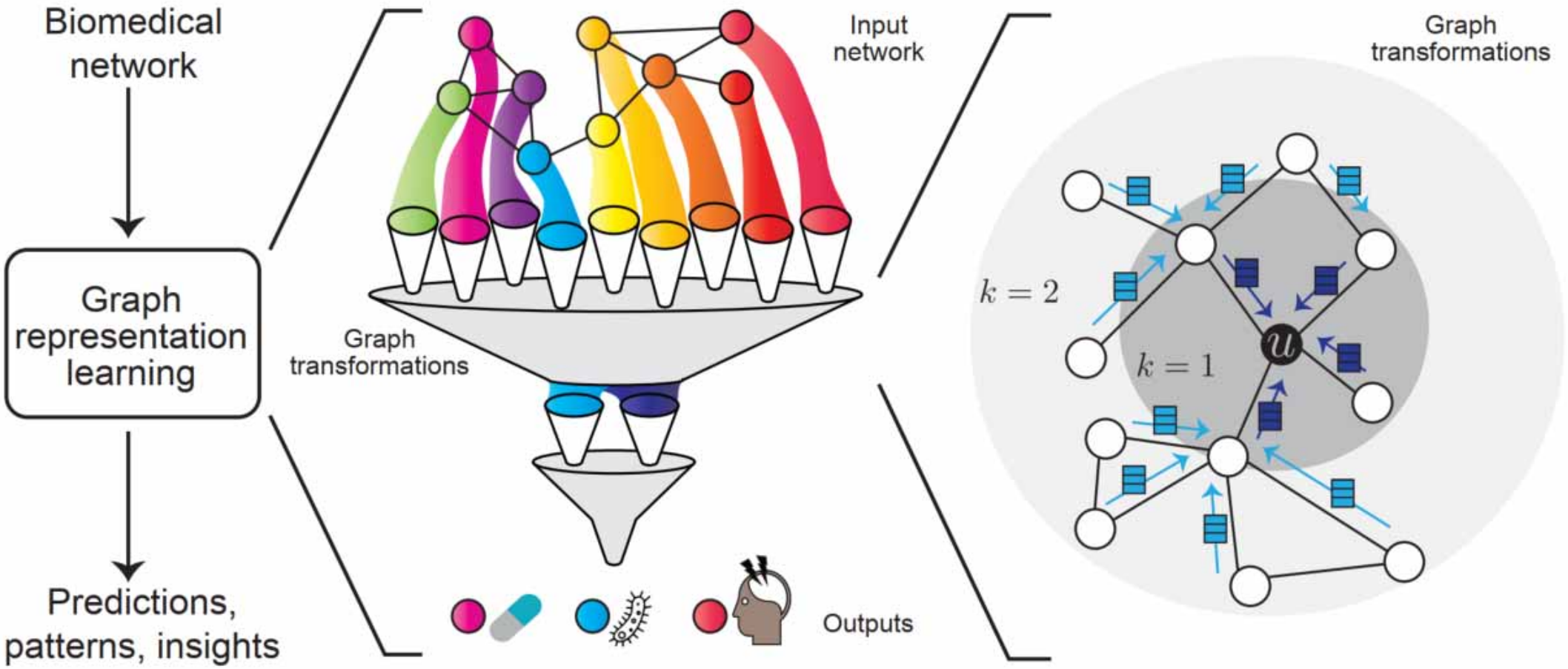
How could a hybrid solution look like ?



Randy Goebel, Ajay Chander, Katharina Holzinger, Freddy Lecue, Zeynep Akata, Simone Stumpf, Peter Kieseberg & Andreas Holzinger (2018). Explainable AI: the new 42? *LNCS 11015.* Cham: Springer, pp. 295-303, doi:10.1007/978-3-319-99740-7-21.

04 Applications

Example: Graph Machine Learning for Computational Biology



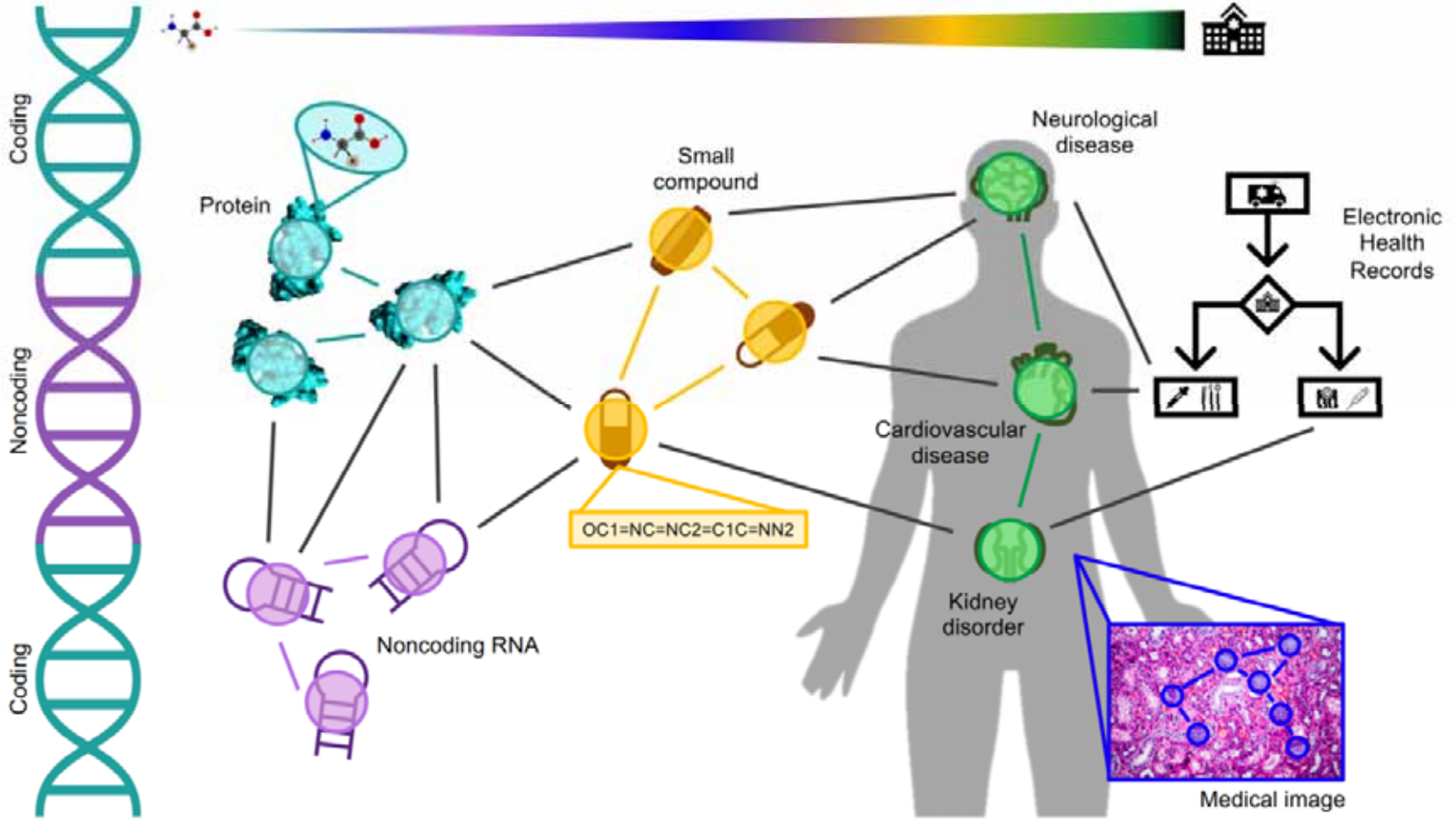
Michelle M. Li, Kexin Huang & Marinka Zitnik (2021). Representation Learning for Networks in Biology and Medicine: Advancements, Challenges, and Opportunities. arXiv:2104.04883.

What are the most prominent graph machine learning tasks?

- Canonical graph prediction. Graph prediction aims to predict a label in the graph. The label can be associated with any unit of the graph.
- Important for Biomedical graphs:
  - (1) Module detection aims to detect a subgraph module in the graph that contributes to a variable;
  - (2) Clustering or community detection aims to partition the graphs into a set of subgraphs such that each subgraph contains similar nodes;
  - (3) Subgraph classification/regression aims to predict a label for the subgraph or module;
  - (4) Dynamic graph prediction aims to perform the above prediction tasks in a sequence of dynamic graphs.
- Latent graph learning.
- Graph generation. The objective of graph generation is to generate a never-before-seen graph G with some properties of interest.

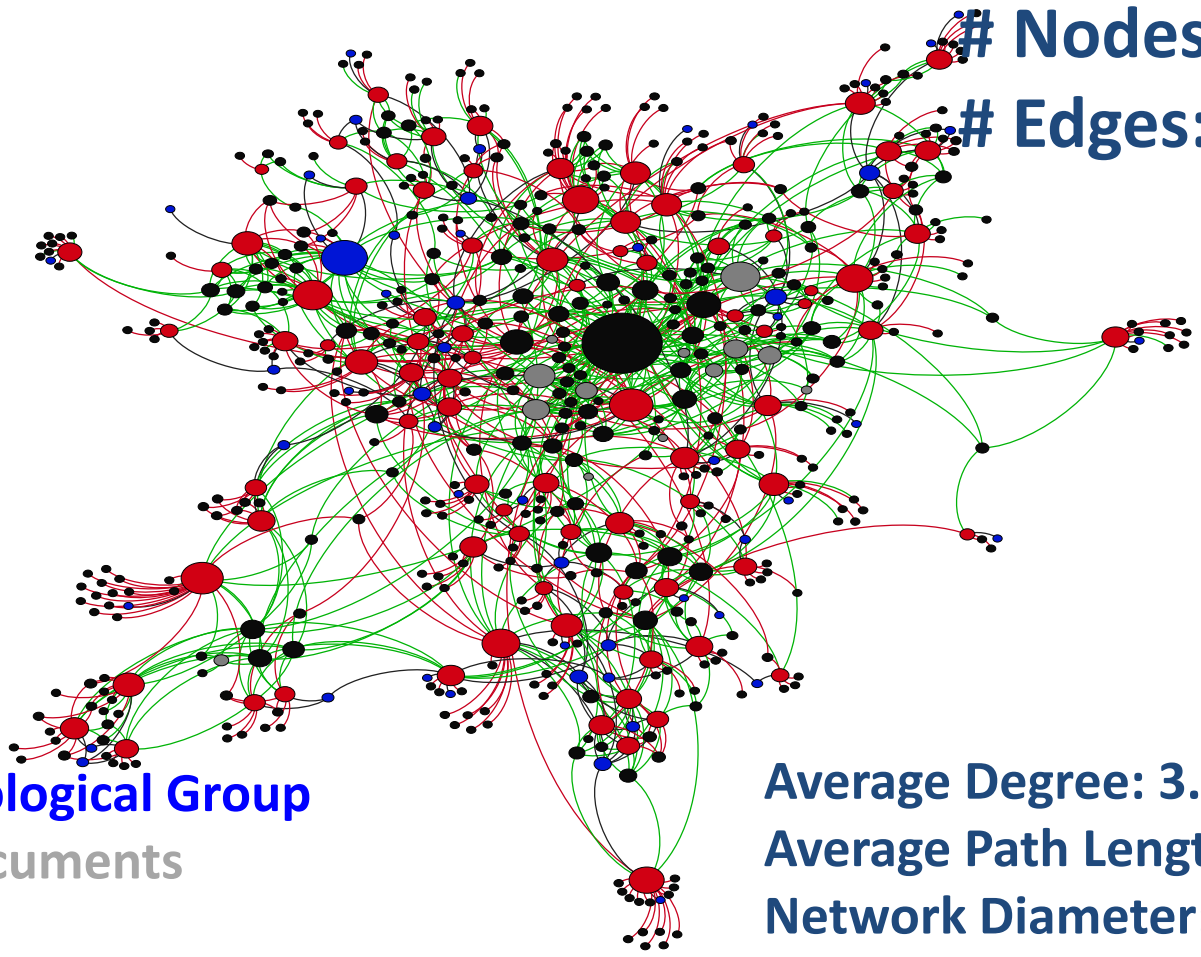
Michelle M. Li, Kexin Huang & Marinka Zitnik (2021). Representation Learning for Networks in Biology and Medicine: Advancements, Challenges, and Opportunities. arXiv:2104.04883.

What are the most important relevant biomedical application areas ?



Michelle M. Li, Kexin Huang & Marinka Zitnik (2021). Representation Learning for Networks in Biology and Medicine: Advancements, Challenges, and Opportunities. arXiv:2104.04883.

Example for a Medical Knowledge Space



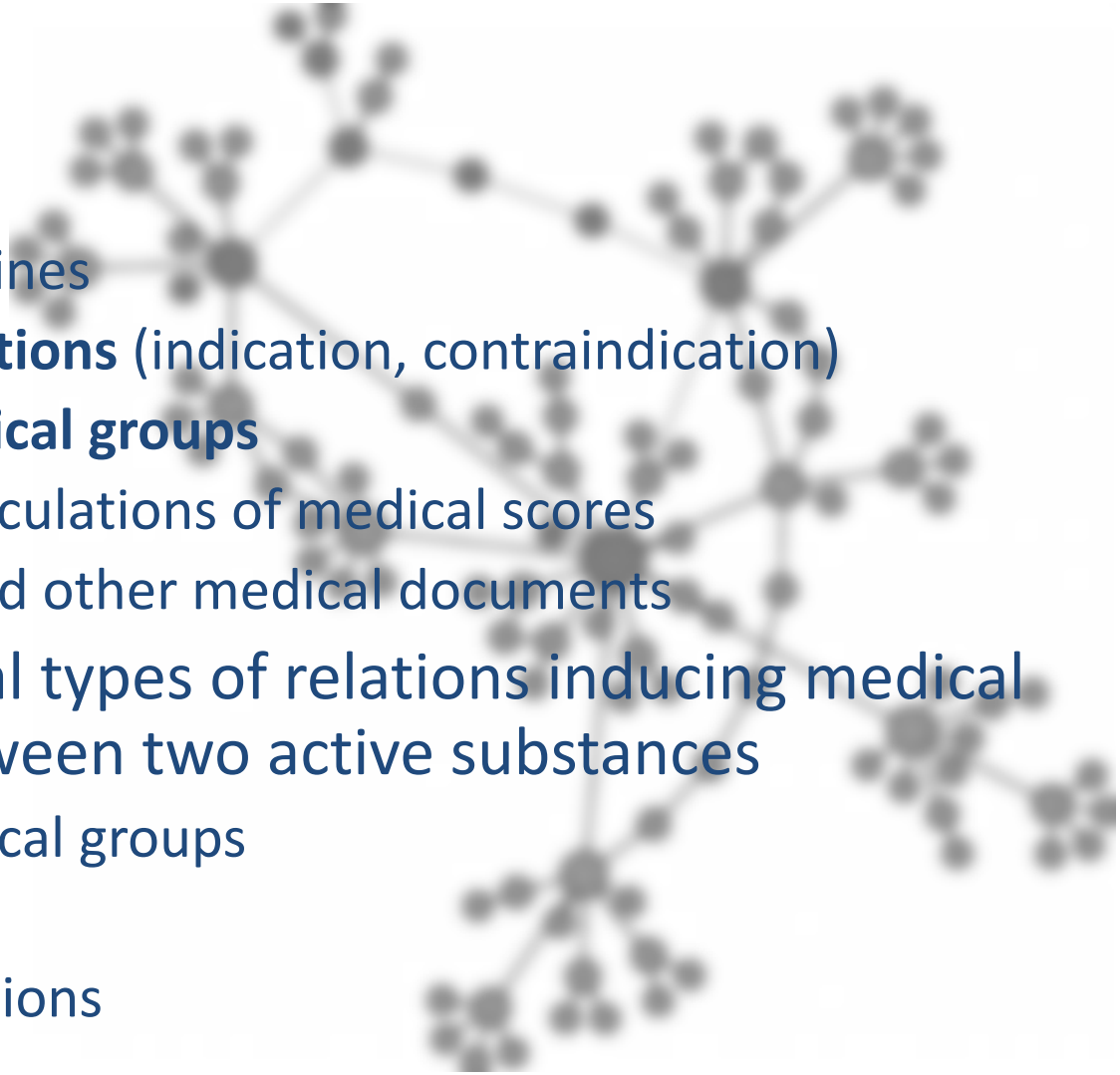
 # Nodes: 641
 # Edges: 1250

Agent
Condition
Pharmacological Group
Other Documents

Average Degree: 3.888
Average Path Length: 4.683
Network Diameter: 9

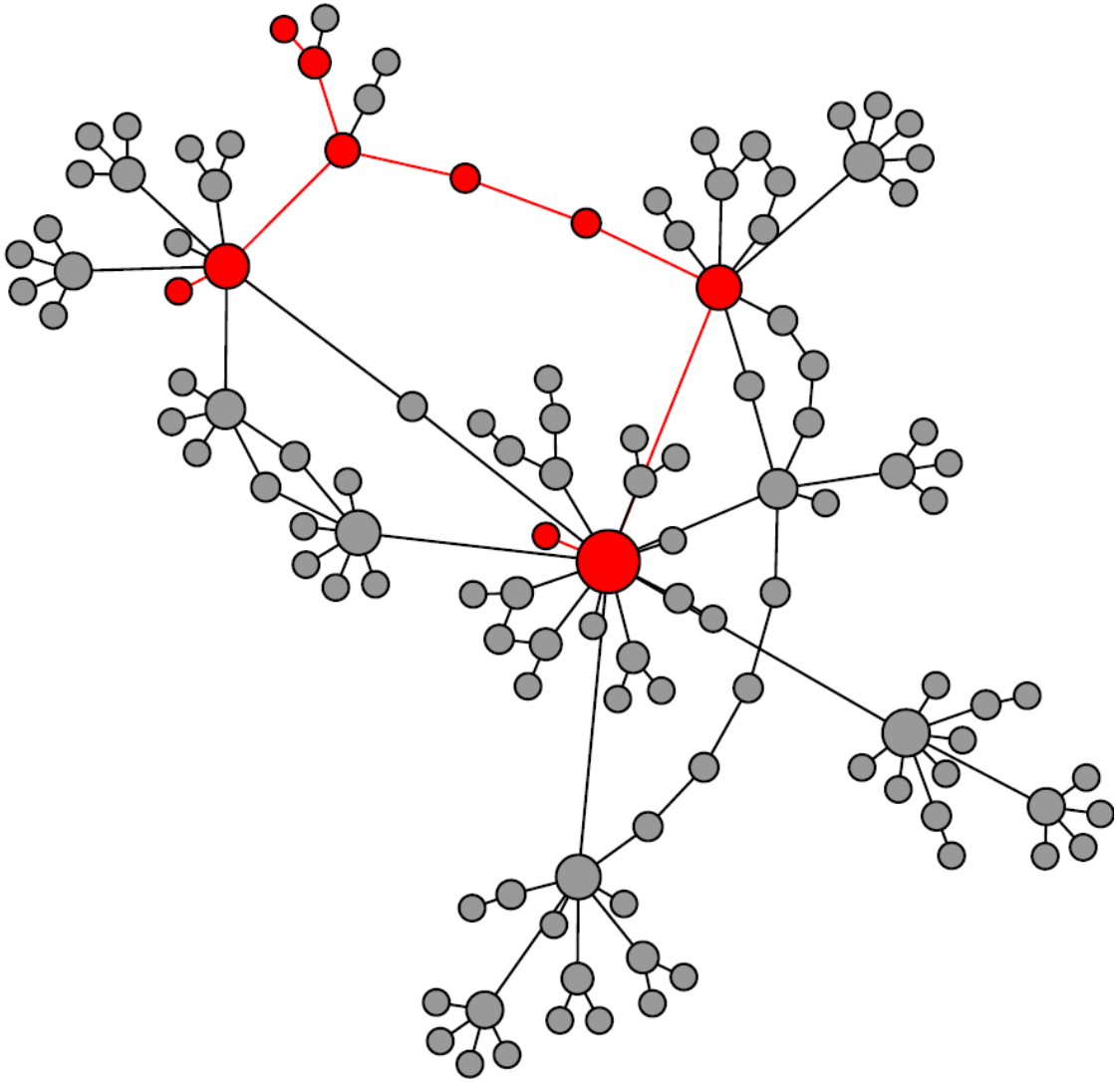
Holzinger, A., Ofner, B., Dehmer, M.: Multi-touch Graph-Based Interaction for Knowledge Discovery on Mobile Devices: State-of-the-Art and Future Challenges. In: LNCS 8401, pp. 241–254, (2014)

Medical Details of the Graph

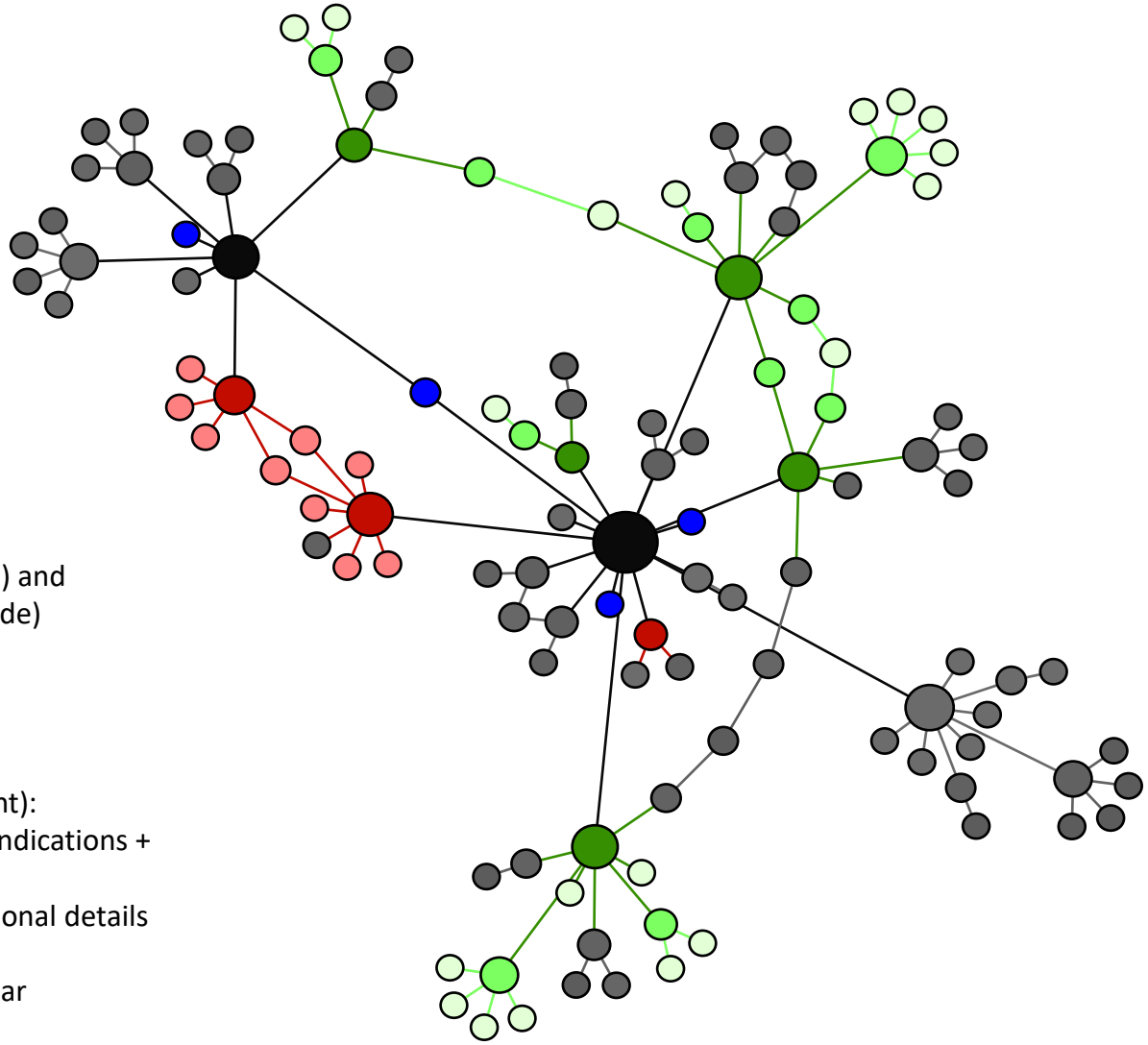


- **Nodes**
  - **drugs**
  - clinical guidelines
  - **patient conditions** (indication, contraindication)
  - **pharmacological groups**
  - tables and calculations of medical scores
  - algorithms and other medical documents
- **Edges**: 3 crucial types of relations inducing medical relevance between two active substances
  - pharmacological groups
  - indications
  - contra-indications

Example for the shortest path



Example for finding related structures

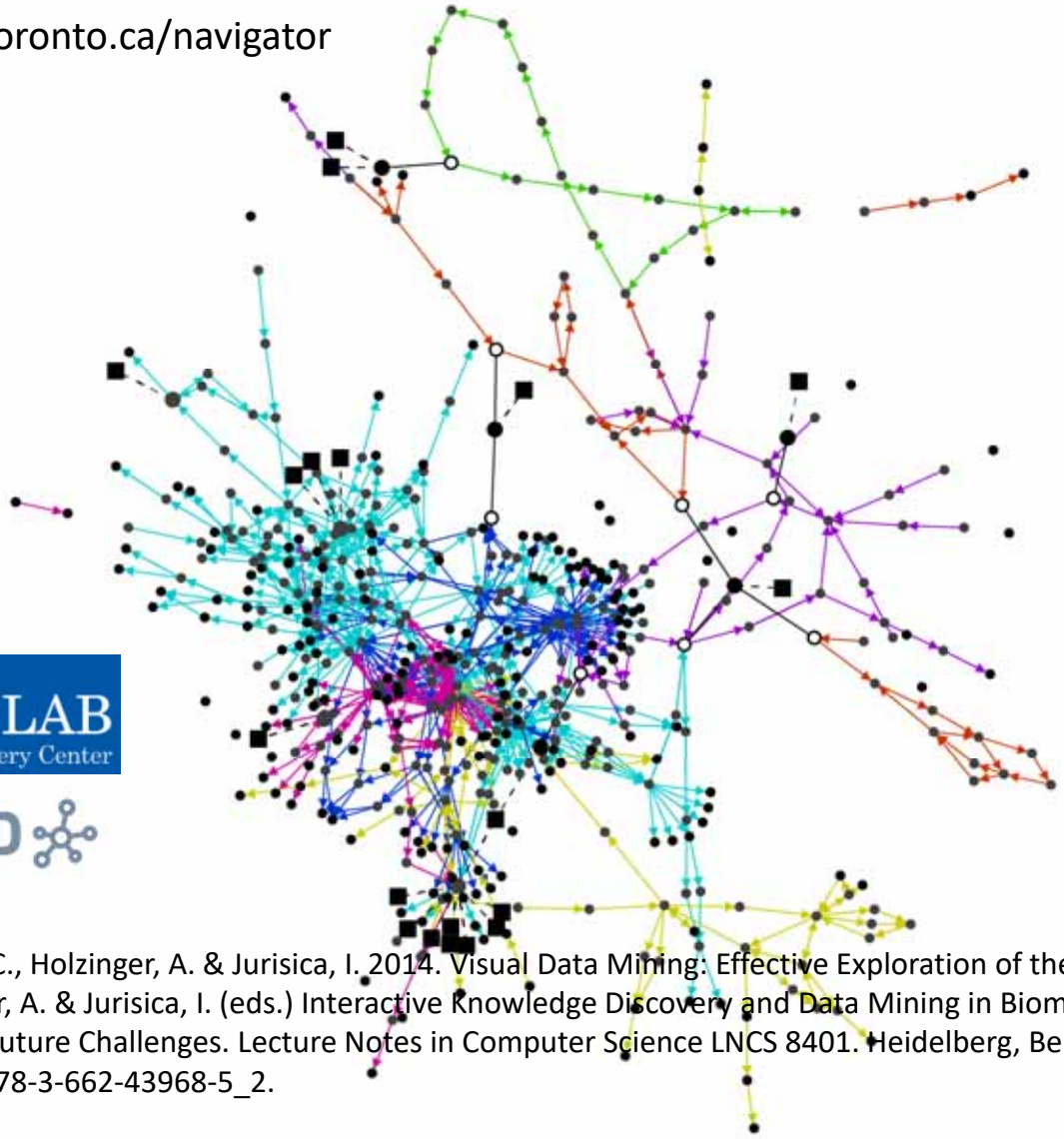


Relationship between
Adrenaline (center black node) and
Dobutamine (top left black node)
Blue: Pharmacological Group
Dark red: Contraindication;
Light red: Condition

Green nodes (from dark to light):
1. Application (one ore more indications + corresponding dosages)
2. Single indication with additional details (e. g. “VF after 3rd Shock”)
3. Condition (e.g. VF, Ventricular Fibrillation)

Interactive Visual Data Mining

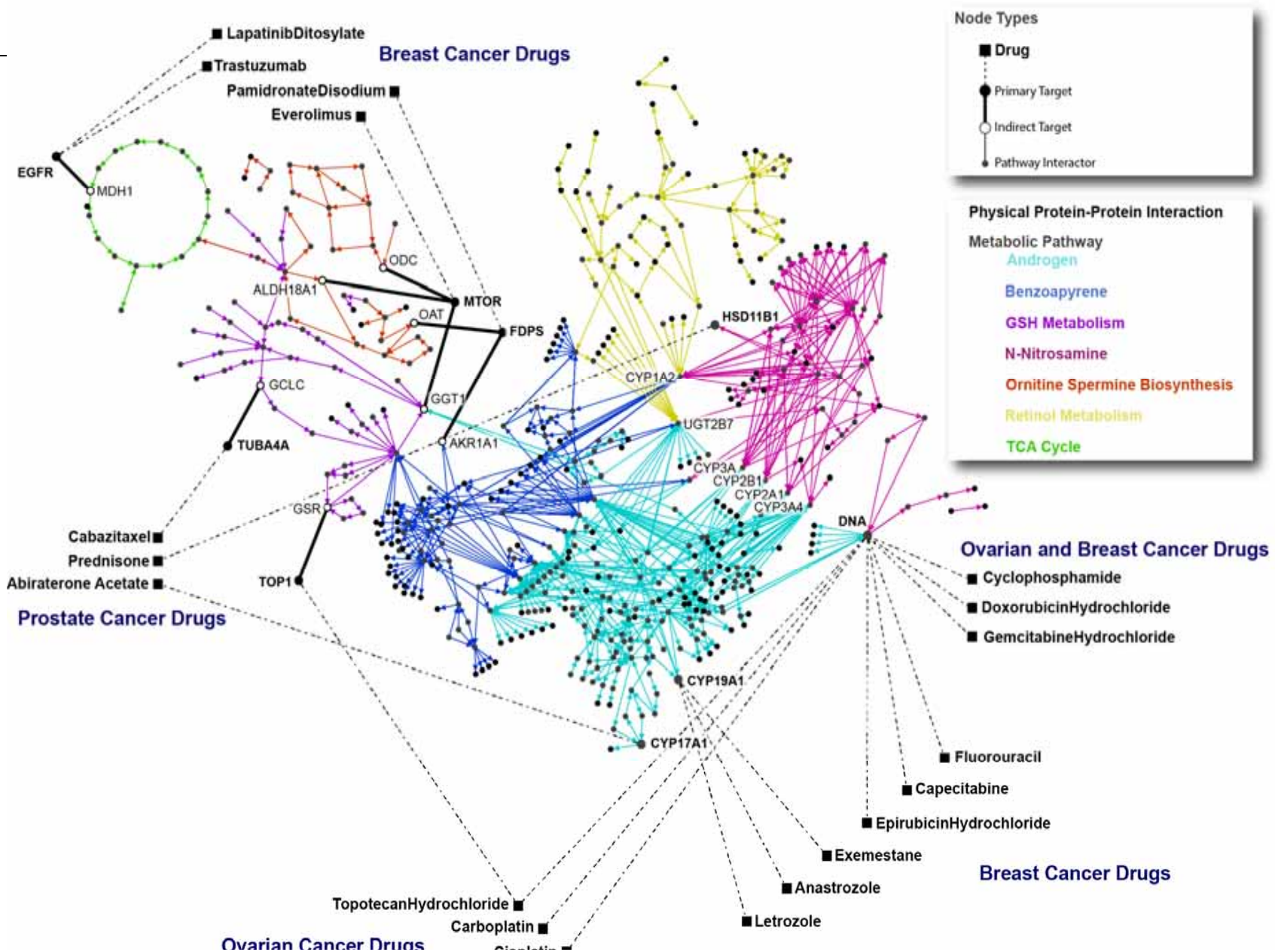
http://ophid.utoronto.ca/navigator



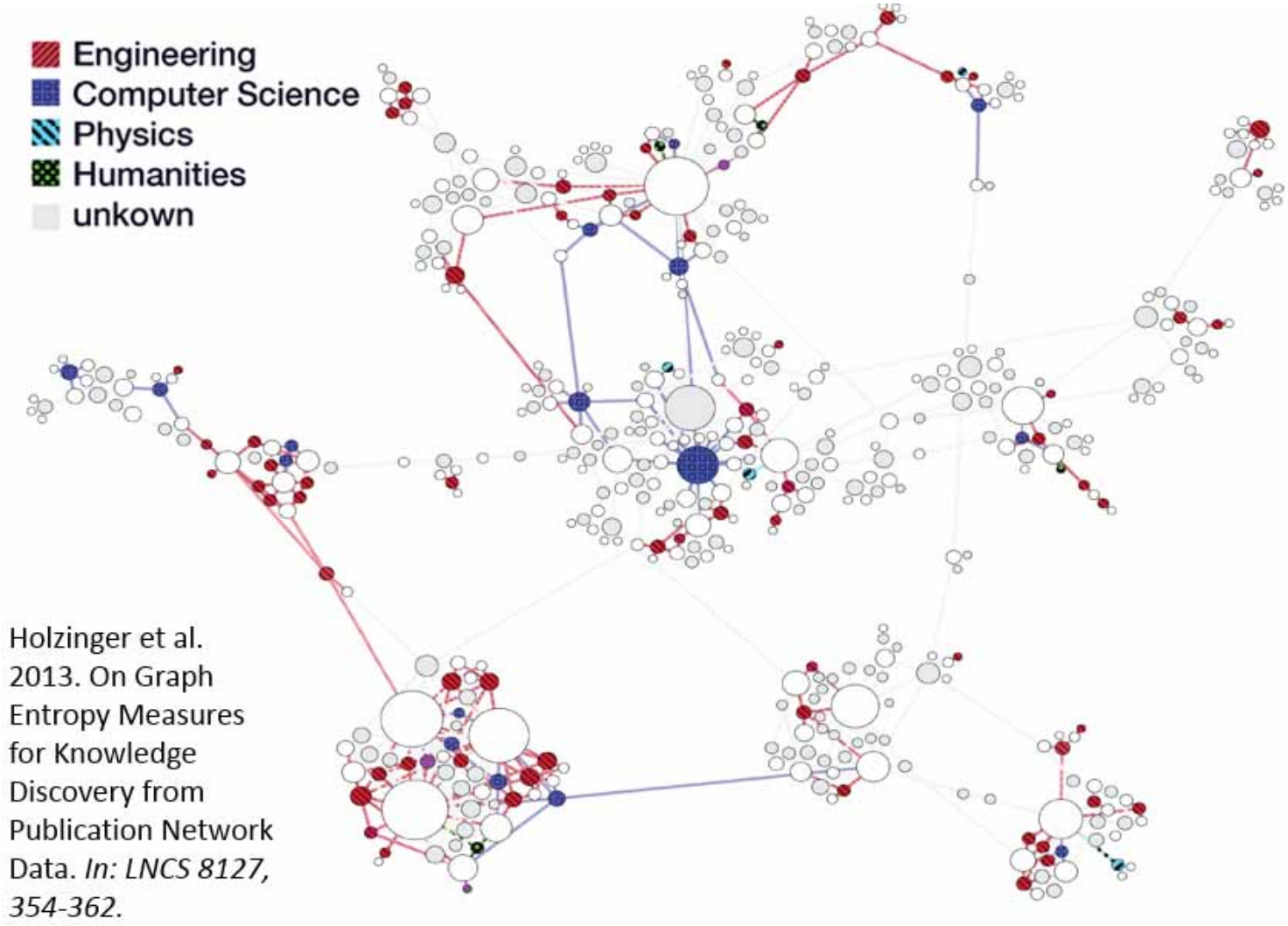
JURISICA LAB
IBM Life Sciences Discovery Center

HCI-KDD

Otasek, D., Pastrello, C., Holzinger, A. & Jurisica, I. 2014. Visual Data Mining: Effective Exploration of the Biological Universe. In: Holzinger, A. & Jurisica, I. (eds.) Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges. Lecture Notes in Computer Science LNCS 8401. Heidelberg, Berlin: Springer, pp. 19–34, doi:10.1007/978-3-662-43968-5_2.


Node Types
Drug
Primary Target
Indirect Target
Pathway Interactor
Physical Protein-Protein Interaction
Metabolic Pathway
Androgen
Benzoapyrene
GSH Metabolism
N-Nitrosamine
Ornitine Spermine Biosynthesis
Retinol Metabolism
TCA Cycle
Breast Cancer Drugs
LapatinibDitosylate
Trastuzumab
PamidronateDisodium
Everolimus
EGFR
MDH1
ODC
ALDH18A1
MTOR
OAT
FDPS
GCLC
GGT1
TUBA4A
AKR1A1
HSD11B1
CYP1A2
UGT2B7
CYP3A
CYP2B1
CYP2A1
CYP3A4
GSR
TOP1
DNA
Cabazitaxel
Prednisone
Abiraterone Acetate
Prostate Cancer Drugs
Ovarian and Breast Cancer Drugs
Cyclophosphamide
DoxorubicinHydrochloride
GemcitabineHydrochloride
CYP19A1
CYP17A1
Fluorouracil
Capecitabine
EpirubicinHydrochloride
Exemestane
Anastrozole
Letrozole
Breast Cancer Drugs
TopotecanHydrochloride
Carboplatin
Ovarian Cancer Drugs

Example: Graph Entropy Measures



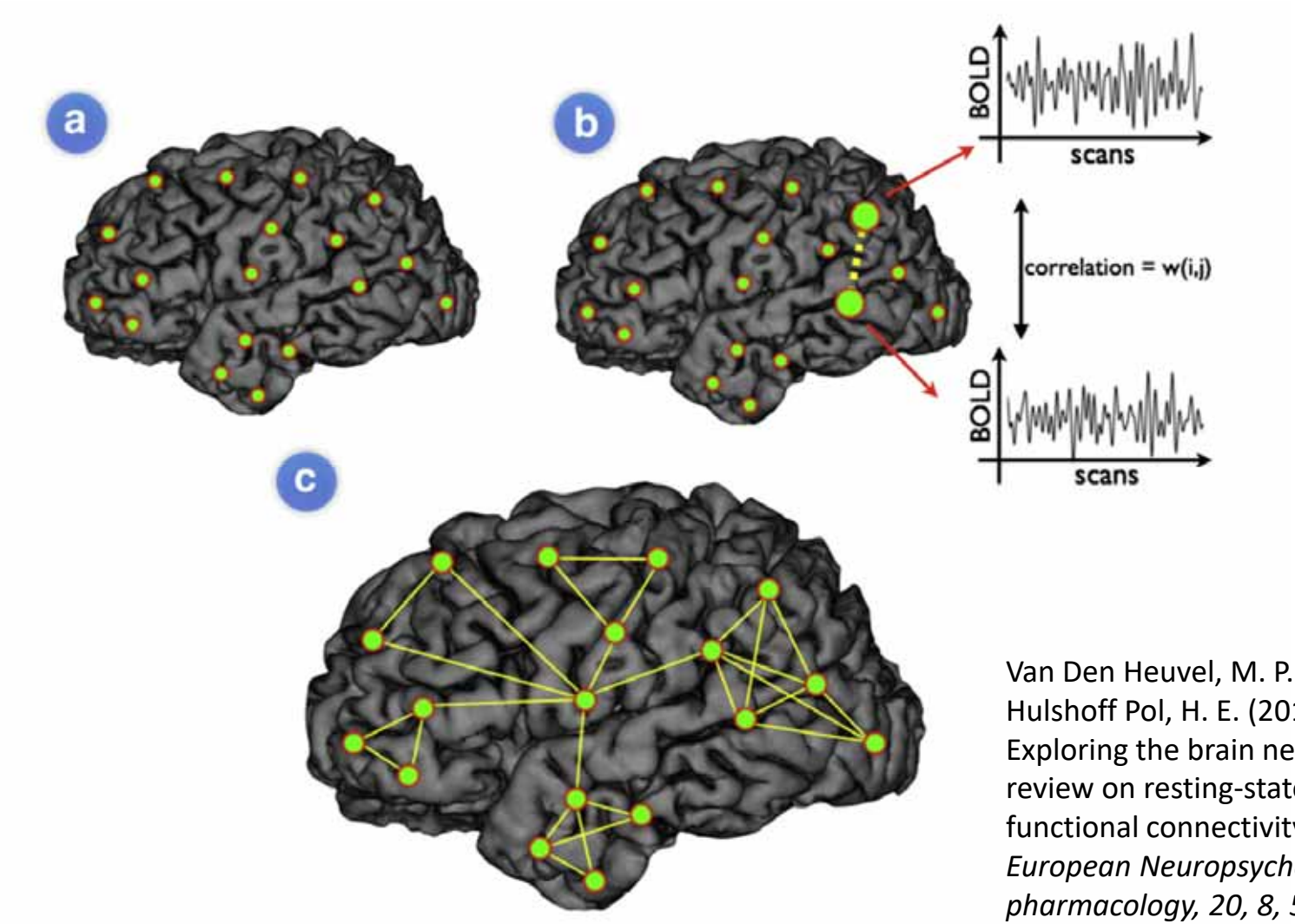
Holzinger et al. 2013. On Graph Entropy Measures for Knowledge Discovery from Publication Network Data. *In: LNCS 8127, 354-362.*

Some selected open problems

- **Problem:** What is the max. number of edges of an Relative Neighborhood Graph in R3 ? No supra-linear lower bound is known.
- **Problem:** What is the structural interpretation of graph measures ? They are mappings which maps graphs to the reals. Thus, they can be understood as graph complexity measures and investigating their structural interpretation relates to understand what kind of structural complexity they detect.
- **Problem:** It is important to visualize large networks meaningfully. So far, there has been a lack of interest to develop efficient software beyond the available commercial software.
- **Problem:** Are multi-touch interaction graphs structurally similar to other graphs (from known graph classes)? This calls for a comparison of graph classes and their structural characteristics.
- **Problem:** Which graph measures are suitable to determine the complexity of multi-touch interaction graphs? Does this lead to any meaningful classification based on their topology?
- **Problem:** What is interesting? Where to start the interaction?

Holzinger, A., Ofner, B., & Dehmer, M. (2014). Multi-touch Graph-Based Interaction for Knowledge Discovery on Mobile Devices: State-of-the-Art and Future Challenges. LNCS 8401 (pp. 241–254). Berlin, Heidelberg: Springer.

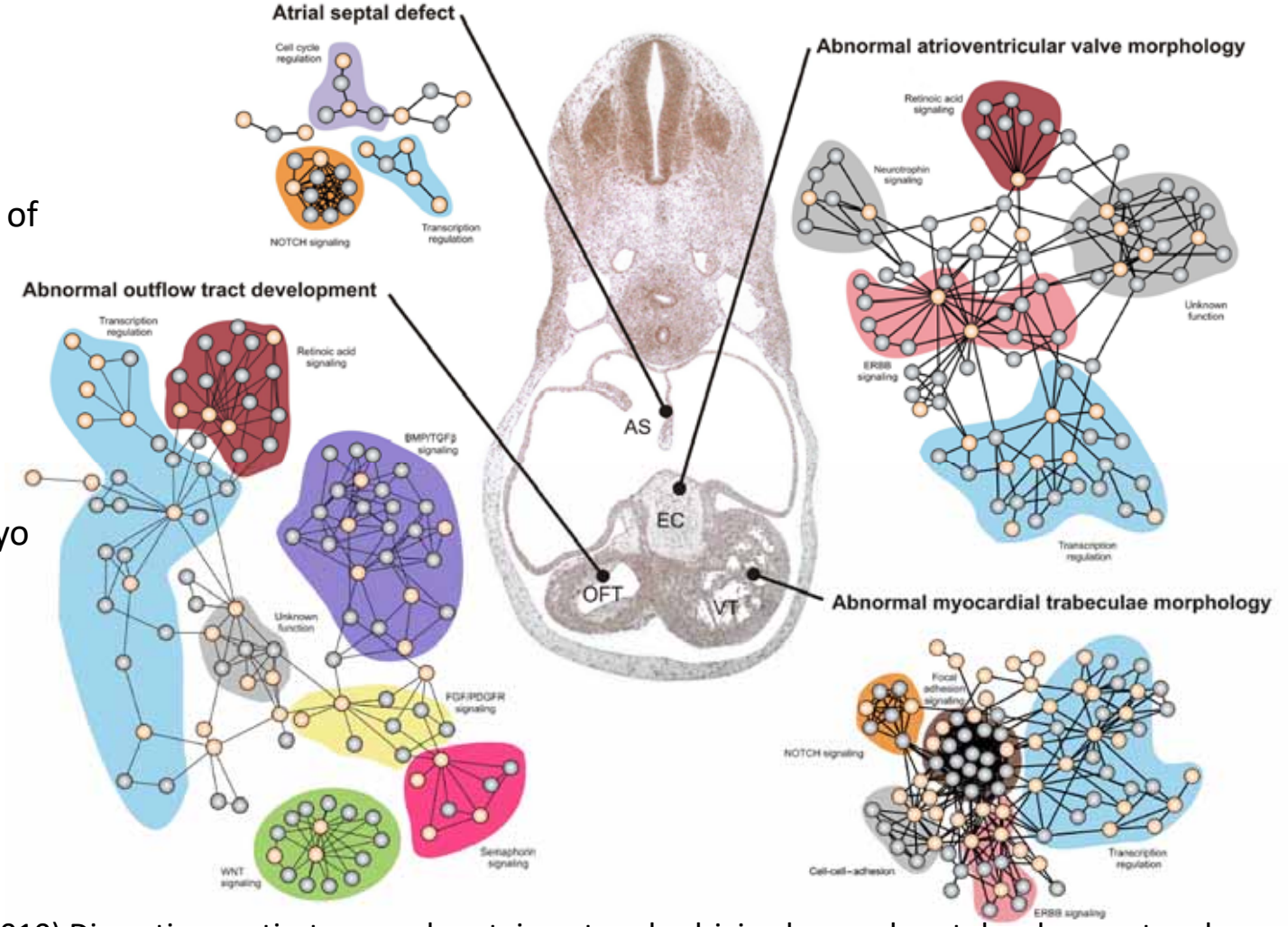
Example: The brain is a complex network



Van Den Heuvel, M. P. & Hulshoff Pol, H. E. (2010) Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsycho-pharmacology, 20, 8, 519-534.*

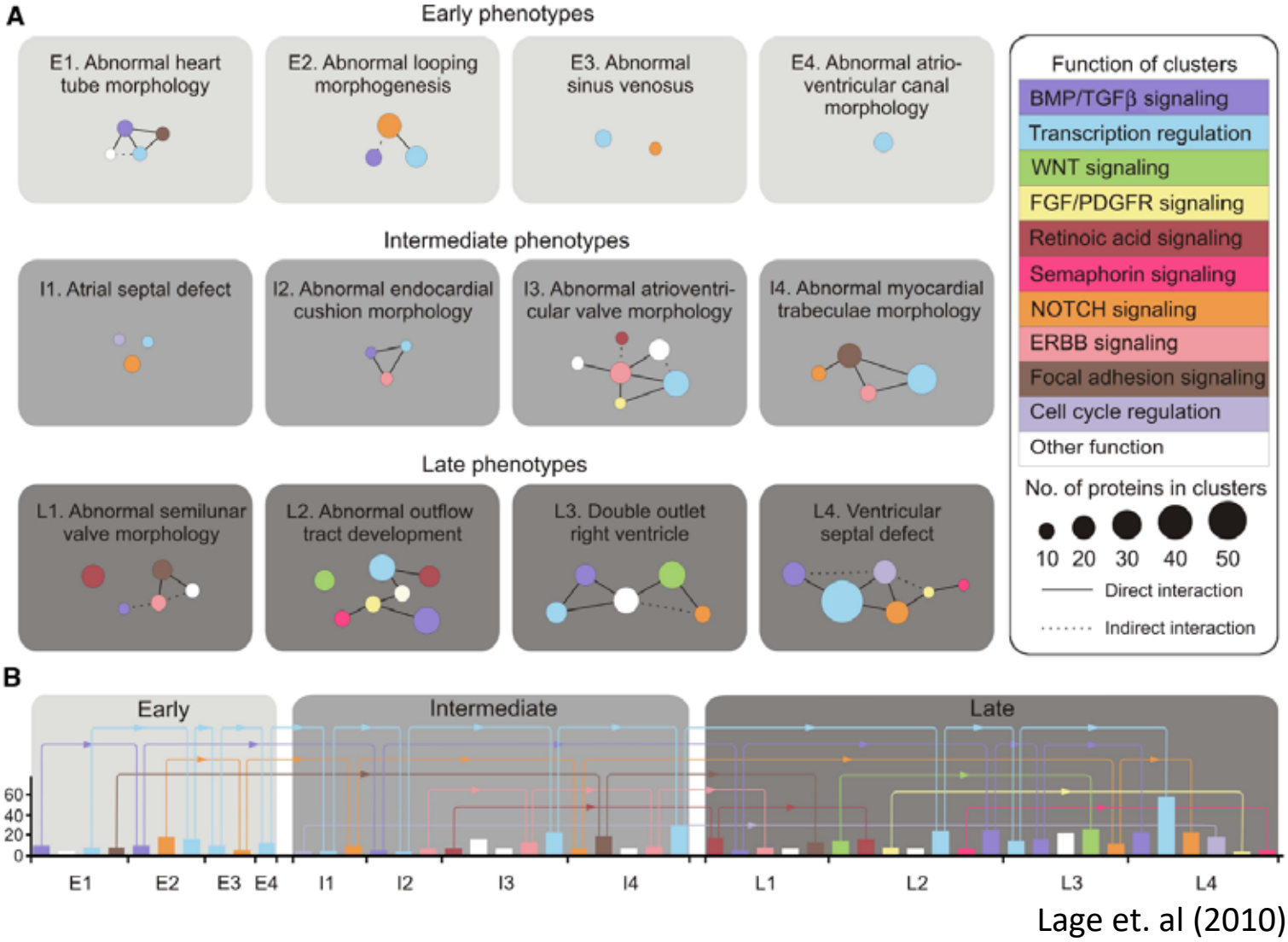
Representative Examples of disease complexes

Examples of 4 functional networks driving the development of different anatomical structures in the human heart of a 37-day old human embryo



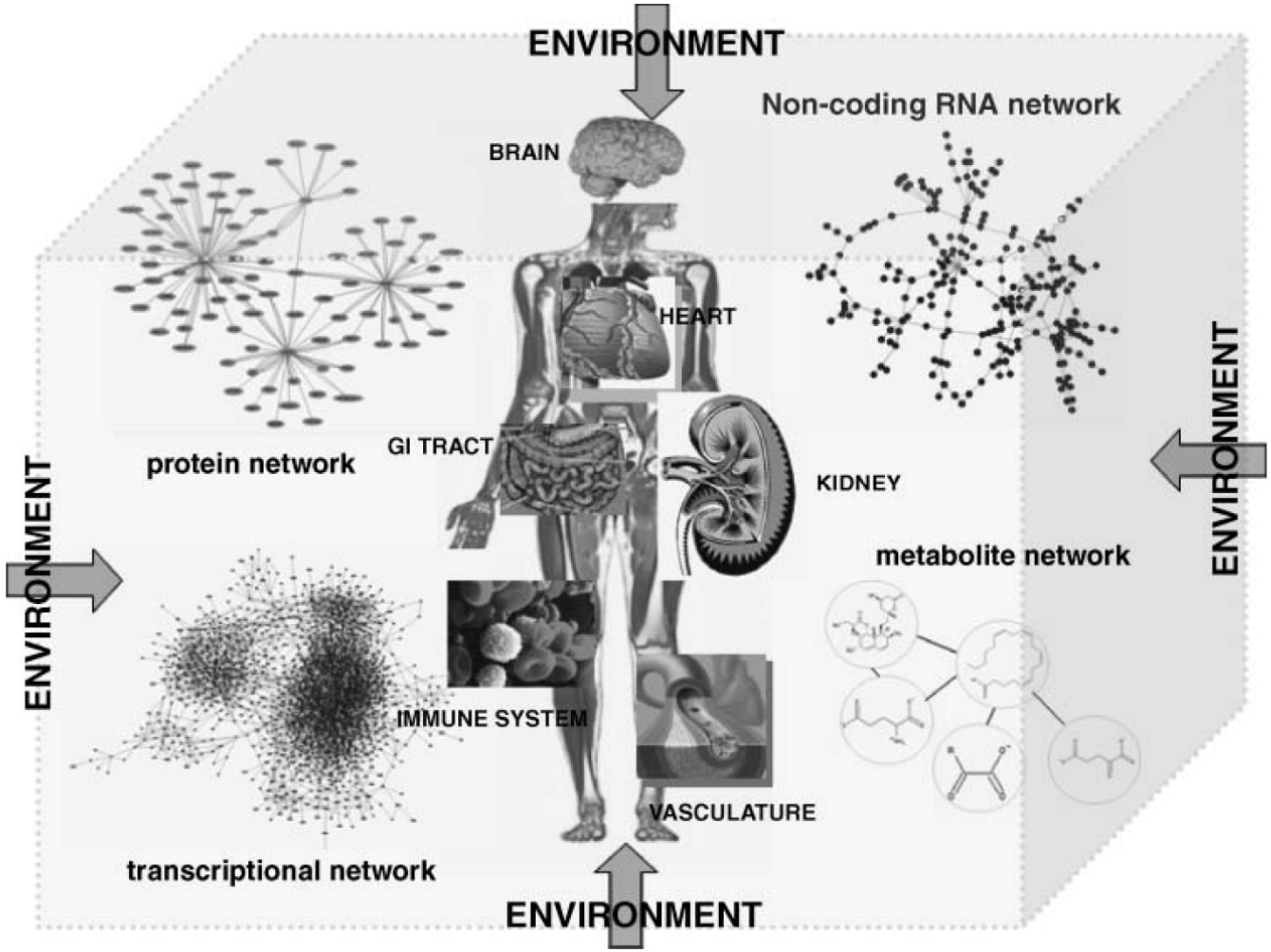
Lage, K. et. al (2010) Dissecting spatio-temporal protein networks driving human heart development and related disorders. *Molecular systems biology, 6, 1, 1-9*.

Example: Cell-based therapy



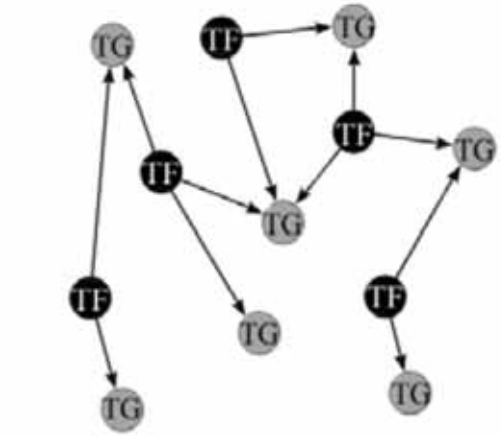
Lage et. al (2010)

Identifying Networks in Disease Research

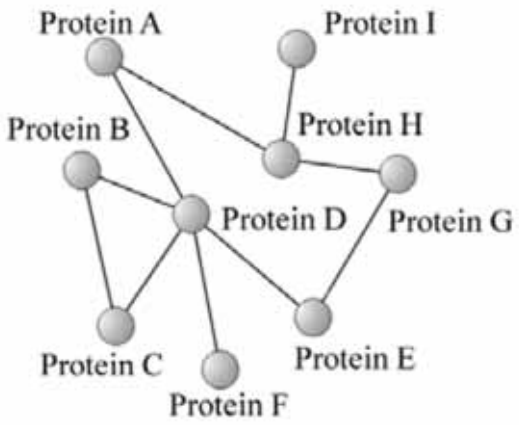


Schadt, E. E. & Lum, P. Y. (2006) Reverse engineering gene networks to identify key drivers of complex disease phenotypes. *Journal of lipid research, 47, 12, 2601-2613.*

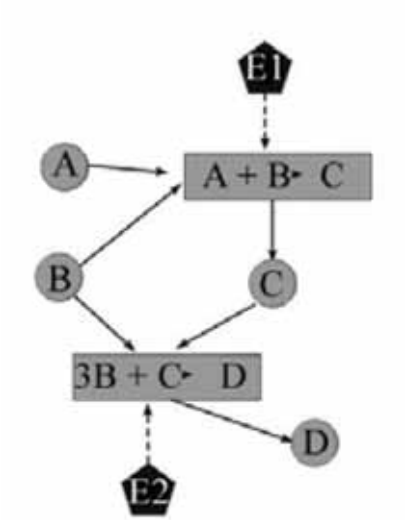
Three main types of biomedical networks



Transcriptional regulatory network with two components:
TF = transcription factor
TG = target genes
(TF regulates the transcription of TG)



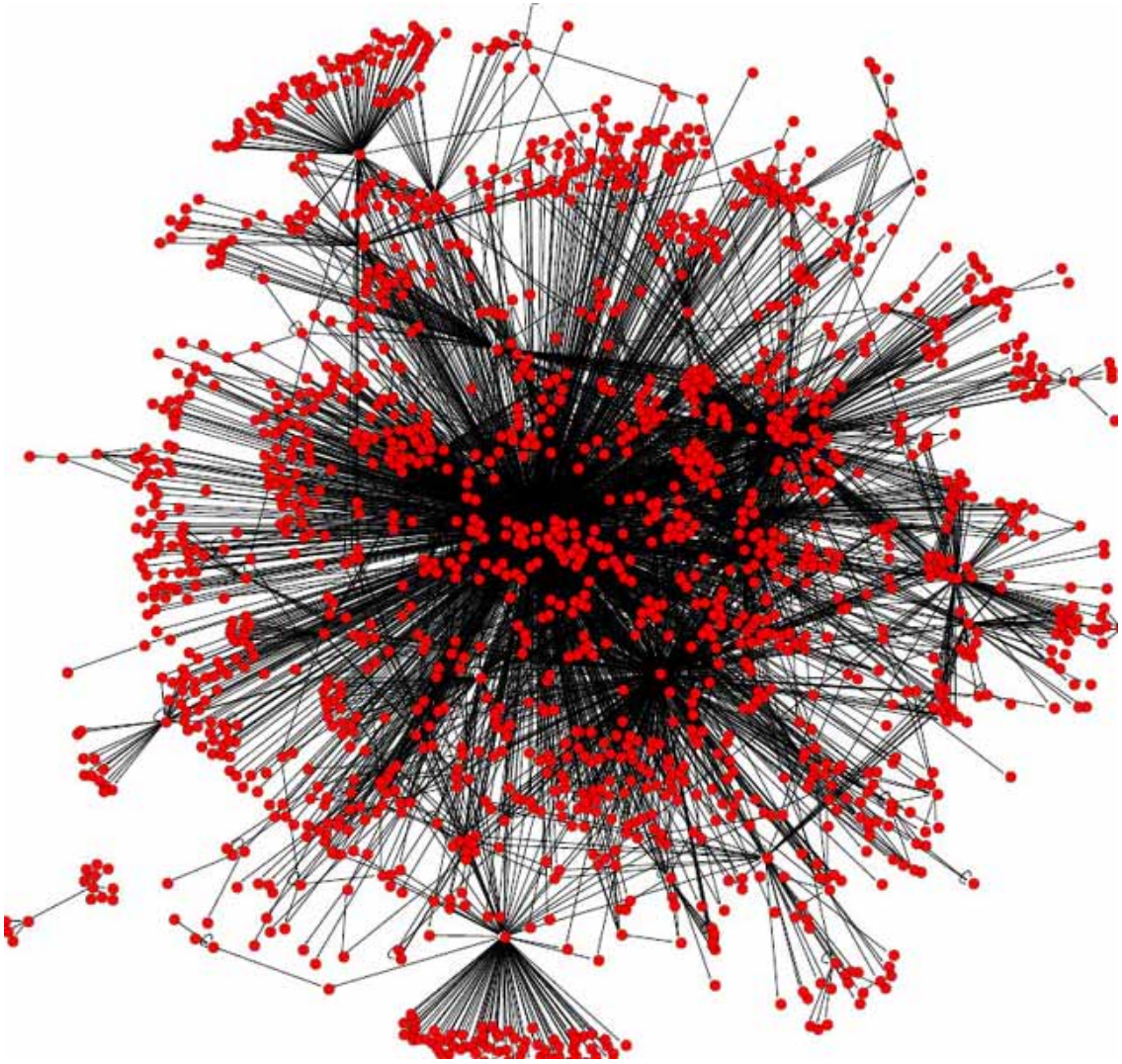
Protein-Protein interaction network



Metabolic network
(constructed considering the reactants, chemical reactions and enzymes)

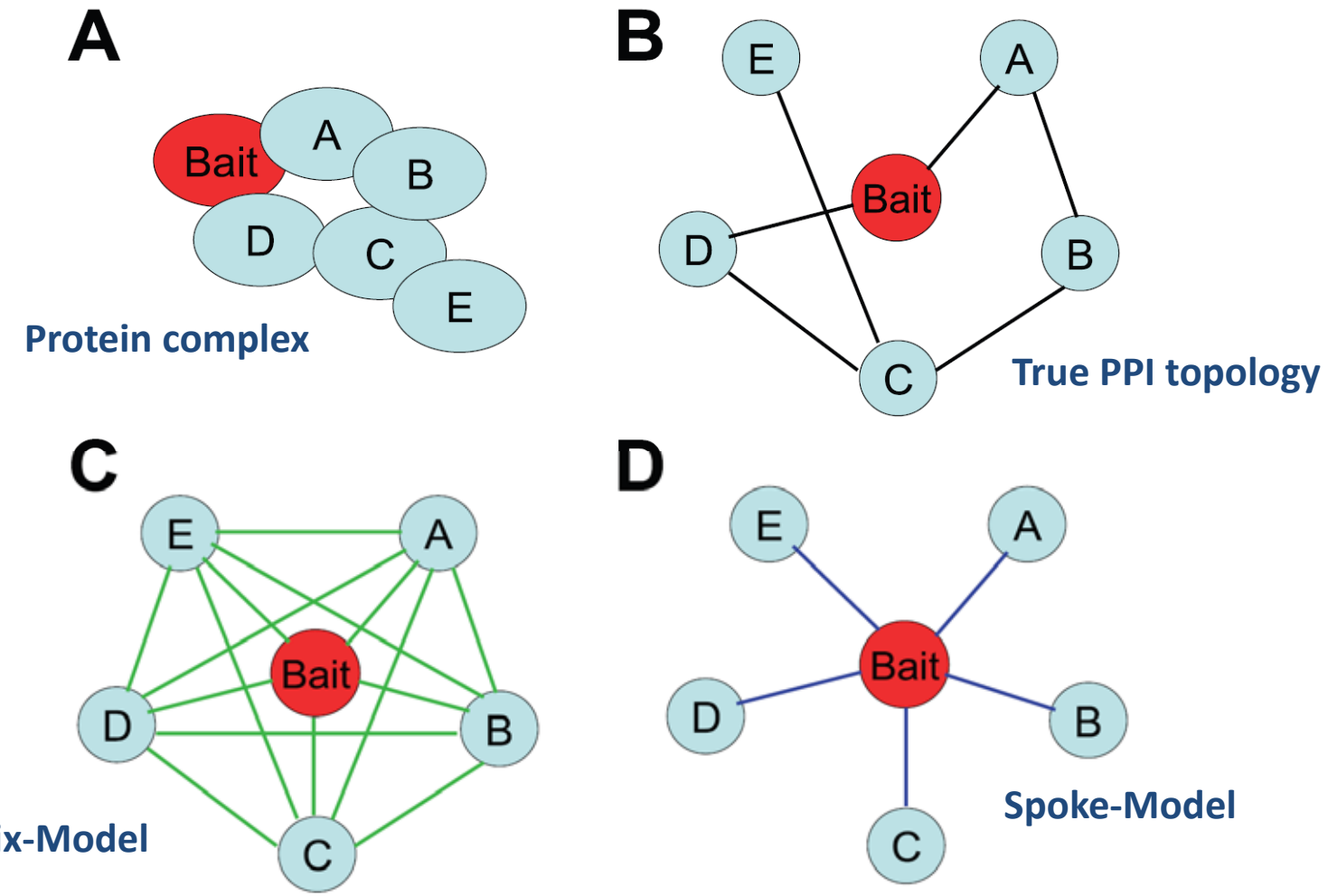
Costa, L. F., Rodrigues, F. A. & Cristino, A. S. (2008) Complex networks: the key to systems biology. *Genetics and Molecular Biology, 31, 3, 591–601.*

Example Transcriptional Regulatory Network



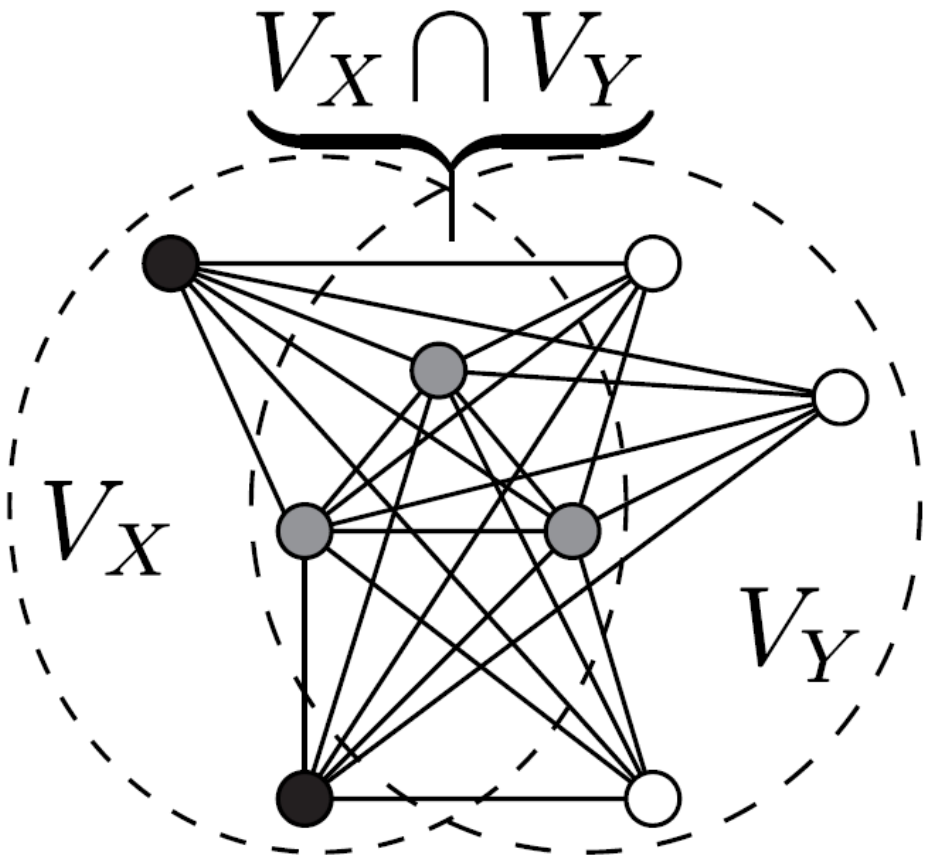
Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Peralta-Gil, M., Peñaloza-Spínola, M. I., Martínez-Antonio, A., Karp, P. D. & Collado-Vides, J. 2006. The comprehensive updated regulatory network of Escherichia coli K-12. *BMC bioinformatics, 7, (1), 5.*

Network Representations of Protein Complexes



Wang, Z. & Zhang, J. Z. (2007) In search of the biological significance of modular structures in protein networks. *PLoS Computational Biology, 3, 6, 1011-1021.*

Correlated Motif Mining (CMM)



Boyen, P., Van Dyck, D., Neven, F., van Ham, R. C. H. J. & van Dijk, A. (2011) SLIDER: A Generic Metaheuristic for the Discovery of Correlated Motifs in Protein-Protein Interaction Networks. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on, 8, 5, 1344-1357.*

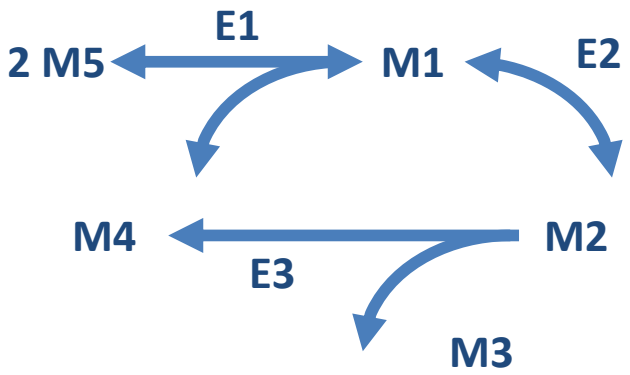
Steepest Ascent Algorithm applied to CMM

Input: PPI-network $G = (V, E, \lambda)$, $\ell, d \in \mathbb{N}$, $d < \ell$
Output: $\{X^*, Y^*\}$ best correlated motif pair found in G

1: $\{X^*, Y^*\} \leftarrow$ randomMotifPair()
2: $maxsup \leftarrow f(\{X^*, Y^*\}, G)$
3: $sup \leftarrow -\infty$
4: **while** $maxsup > sup$ **do**
5: $\quad \{X, Y\} \leftarrow \{X^*, Y^*\}$
6: $\quad sup \leftarrow maxsup$
7: $\quad$ **for all** $\{X', Y'\} \in N(\{X, Y\})$ **do**
8: $\quad\quad$ **if** $f(\{X', Y'\}, G) > maxsup$ **then**
9: $\quad\quad\quad \{X^*, Y^*\} \leftarrow \{X', Y'\}$
10: $\quad\quad\quad maxsup \leftarrow f(\{X', Y'\}, G)$

Boyen et al. (2011)

Metabolic Network



	M1	M2	M3	M4	M5
M1	0	1	0	1	1
M2	1	0	1	1	0
M3	0	0	0	0	0
M4	1	0	0	0	0
M5	1	0	0	0	0

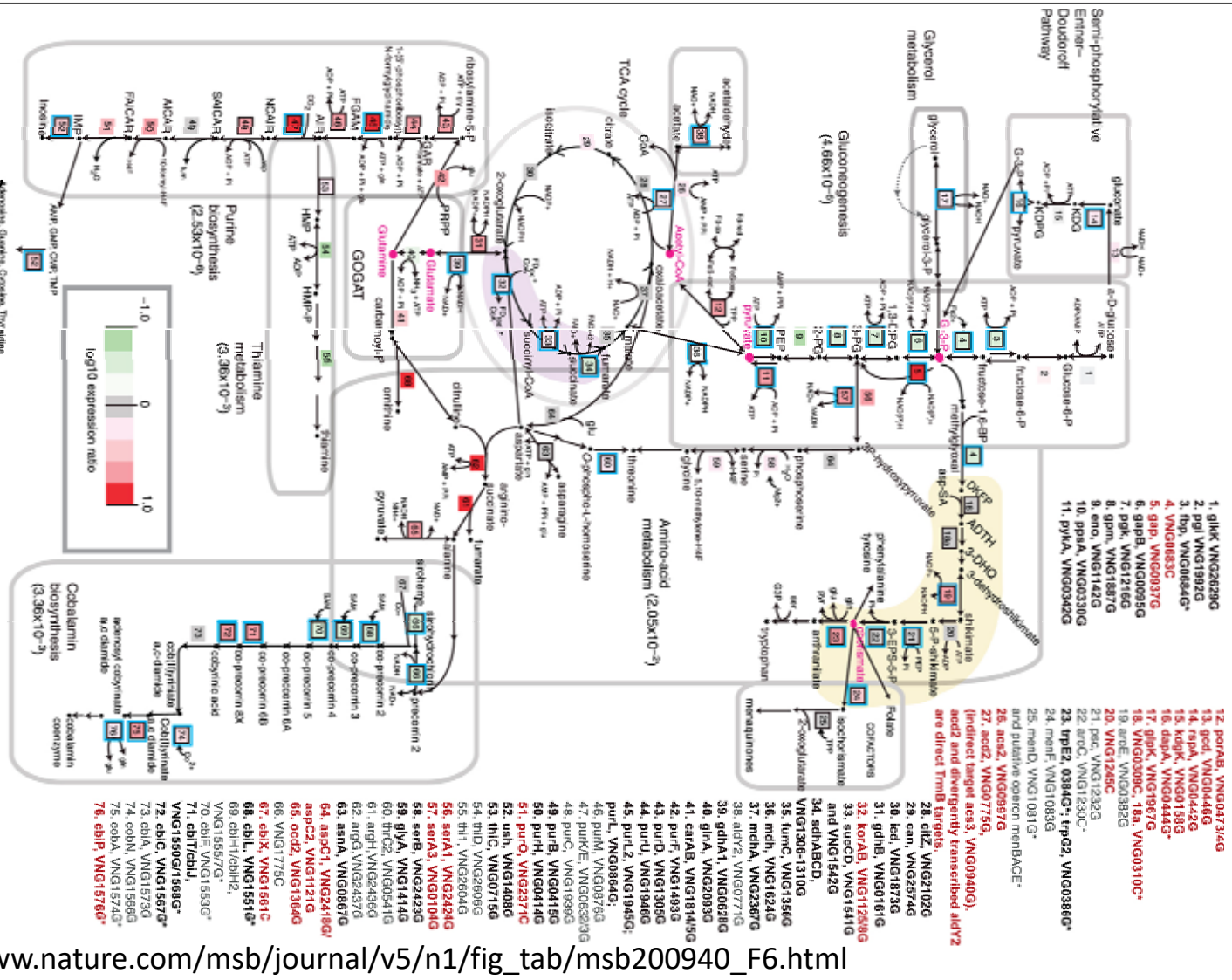
M1	M2
M1	M4
M1	M5
M2	M1
M2	M3
M2	M4
M4	M1
M5	M1

Matrix contains many sparse elements - In this case it is computationally more efficient to represent the graph as an adjacency list

Hodgman, C. T., French, A. & Westhead, D. R. (2010) *Bioinformatics. Second Edition. New York, Taylor & Francis.*

Metabolic networks are usually big … big data ☺

Schmid, A. K., Reiss, D. J., Pan, M., Koide, T. & Baliga, N. S. (2009) A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. *Molecular Systems Biology, 5, 1-9.*

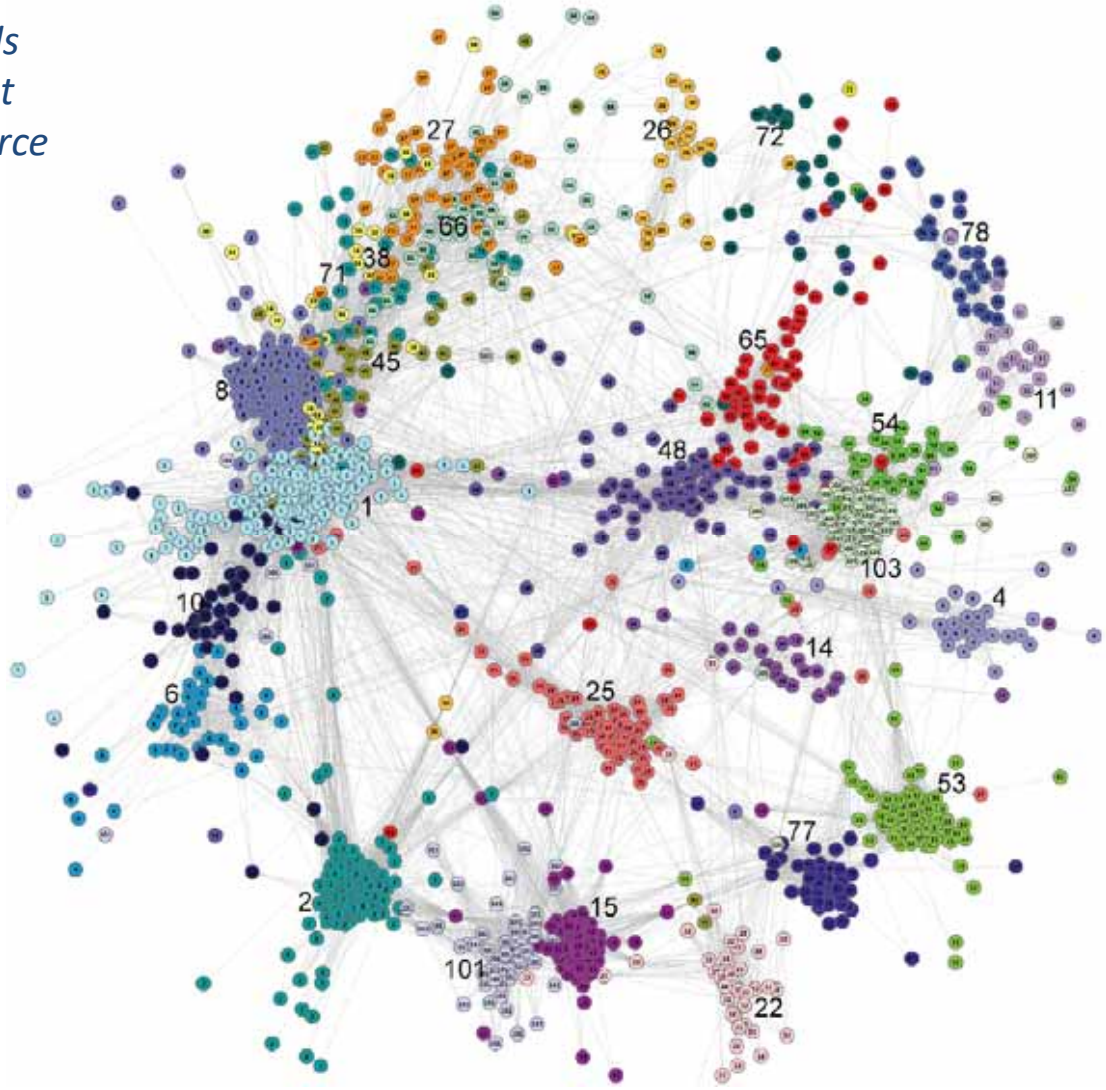


http://www.nature.com/msb/journal/v5/n1/fig_tab/msb200940_F6.html

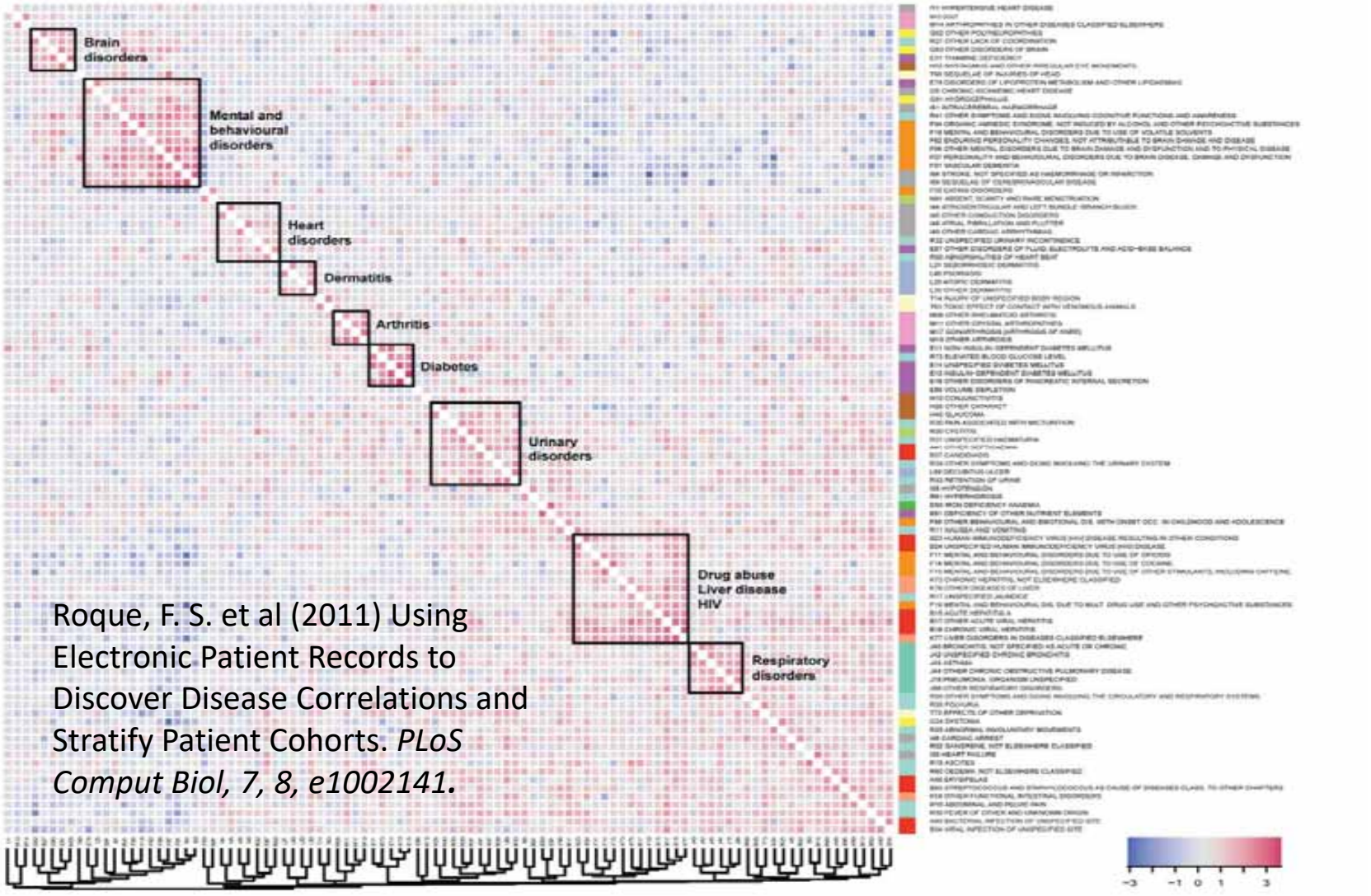
Using EPRs to Discover Disease Correlations

Electronic patient records remain a unexplored, but potentially rich data source for example to discover correlations between diseases.

Roque, F. S., Jensen, P. B., Schmock, H., Dalgaard, M., Andreatta, M., Hansen, T., Søeby, K., Bredkjær, S., Juul, A., Werge, T., Jensen, L. J. & Brunak, S. (2011) Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts. *PLoS Computational Biology, 7, 8, e1002141.*



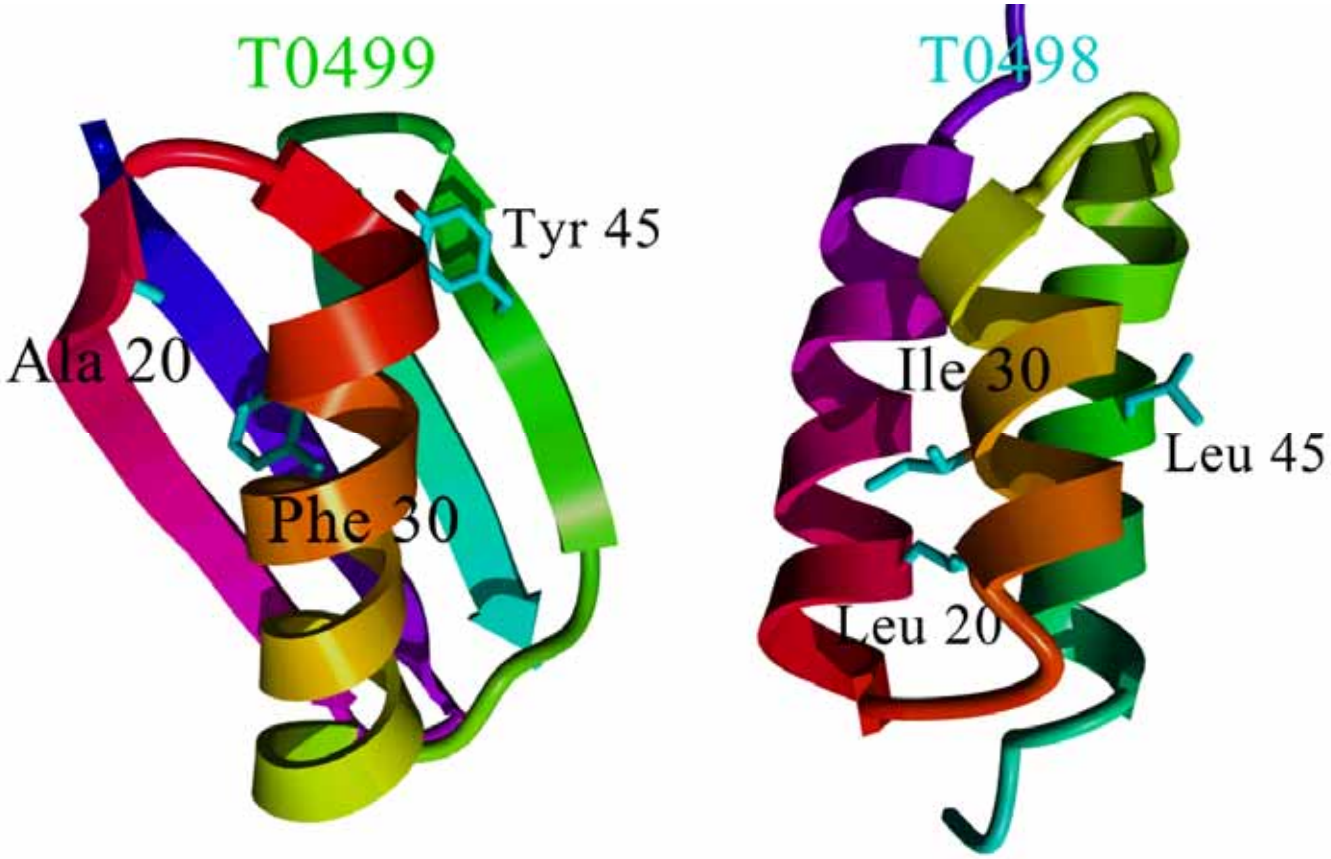
Heatmap of disease-disease correlations (ICD)



Roque, F. S. et al (2011) Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts. *PLoS Comput Biol, 7, 8, e1002141.*

Example: ὁμολογέω (homologeo)

He, Y., Chen, Y., Alexander, P., Bryan, P. N. & Orban, J. (2008) NMR structures of two designed proteins with high sequence identity but different fold and function. Proceedings of the National Academy of Sciences, 105, 38, 14412.

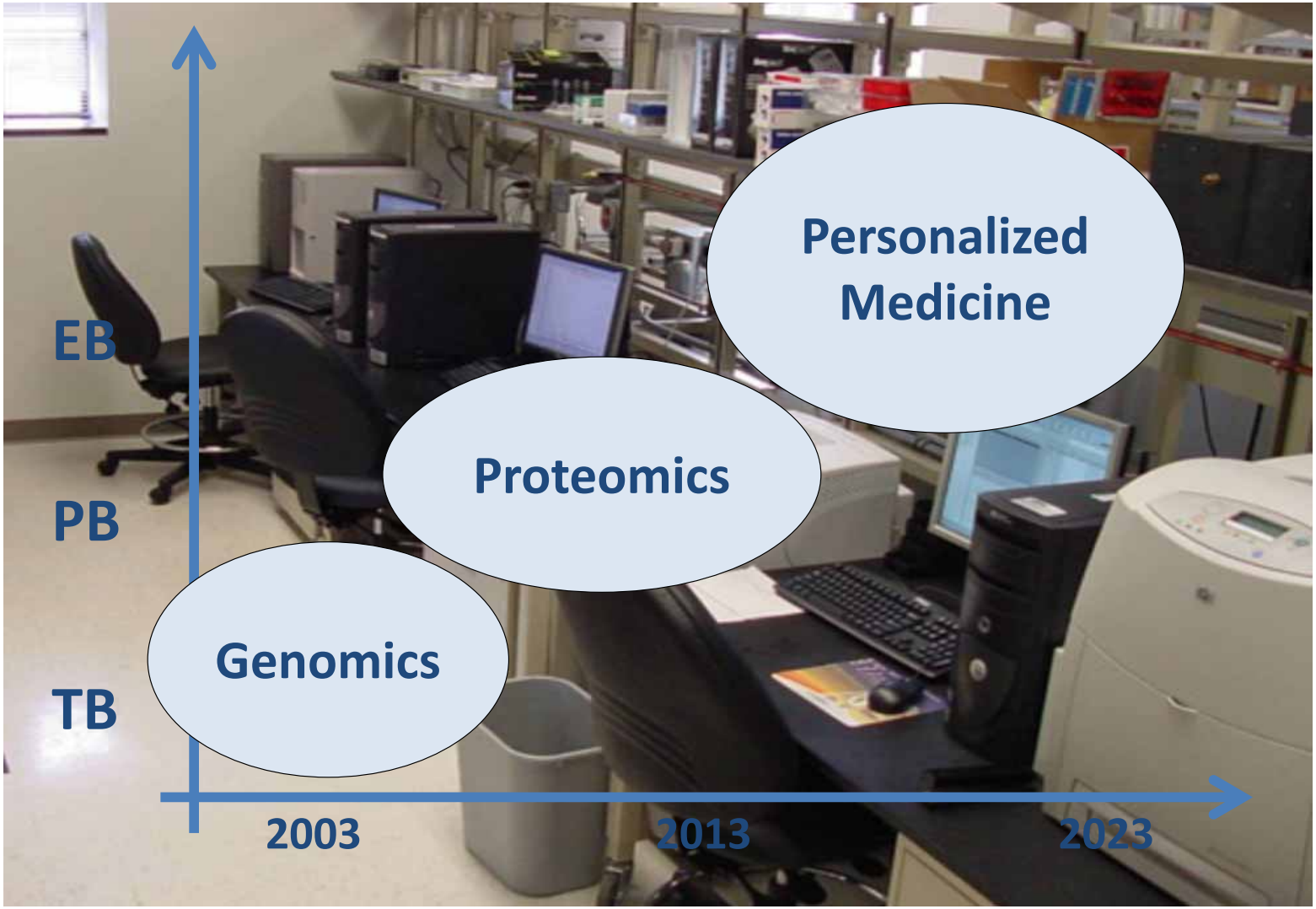



```
T0499  TTYKLILNLKQAKEEAIKEAVDAGTAEKYFKLIANAKTVEGVWTYKDEIKTFTVTE
       ||||||||||||||||||||X|||||||||X|||||||||||||X||||||||||||
T0498  TTYKLILNLKQAKEEAIKELVDAGTAEKYIKLIANAKTVEGVWTLKDEIKTFTVTE
```

Conclusion

- Homology modeling is a knowledge-based prediction of protein structures.
- In homology modeling a protein sequence with an unknown structure (the target) is aligned with one or more protein sequences with known structures (the templates).
- The method is based on the principle that homologue proteins have similar structures.
- **Homology modeling will be extremely important to personalized and molecular medicine in the future.**

Future Outlook

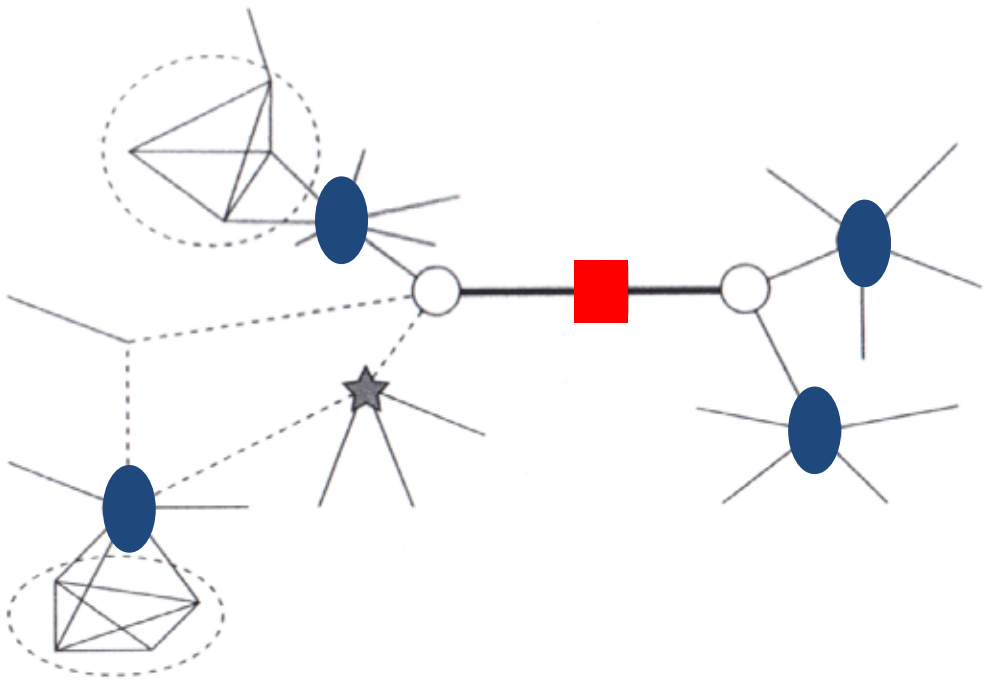


05 Graph metrics and Graph measures

Complex Biological Systems key concepts

- In order to understand complex biological systems, the three following key concepts need to be considered:
- (i) **emergence,** the discovery of links between elements of a system because the study of individual elements such as genes, proteins and metabolites is insufficient to explain the behavior of whole systems;
- (ii) **robustness,** biological systems maintain their main functions even under perturbations imposed by the environment; and
- (iii) **modularity,** vertices sharing similar functions are highly connected.
- Network theory can largely be applied for biomedical informatics, because many tools are already available

$$G(V,E)\ Graph$$
$$V \ldots vertex$$
$$E \ \ldots edge\ \{a,b\}$$
$$\mathbf{a},\mathbf{b} \in V; a \neq b$$

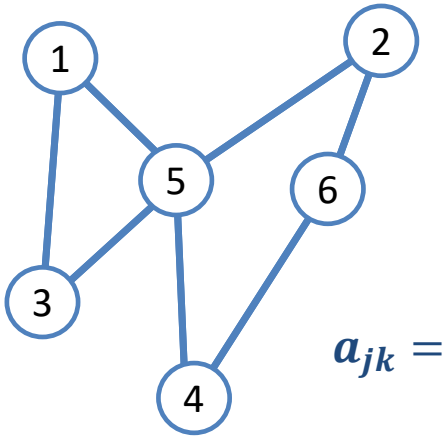


Hodgman, C. T., French, A. & Westhead, D. R. (2010) *Bioinformatics. Second Edition. New York, Taylor & Francis.*

Baby Stuff: Computational Graph Representation

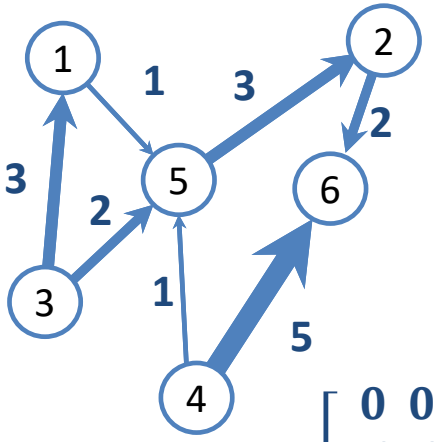
Adjacency (ə-ˈjā-sᵊn(t)-sē) **Matrix** $A = (a_{jk})$

$$a_{jk} = \begin{cases} 1, & if\ \{j,k\} \in E \\ 0, & otherwise \end{cases}$$



$$a_{jk} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Simple graph, symmetric, binary

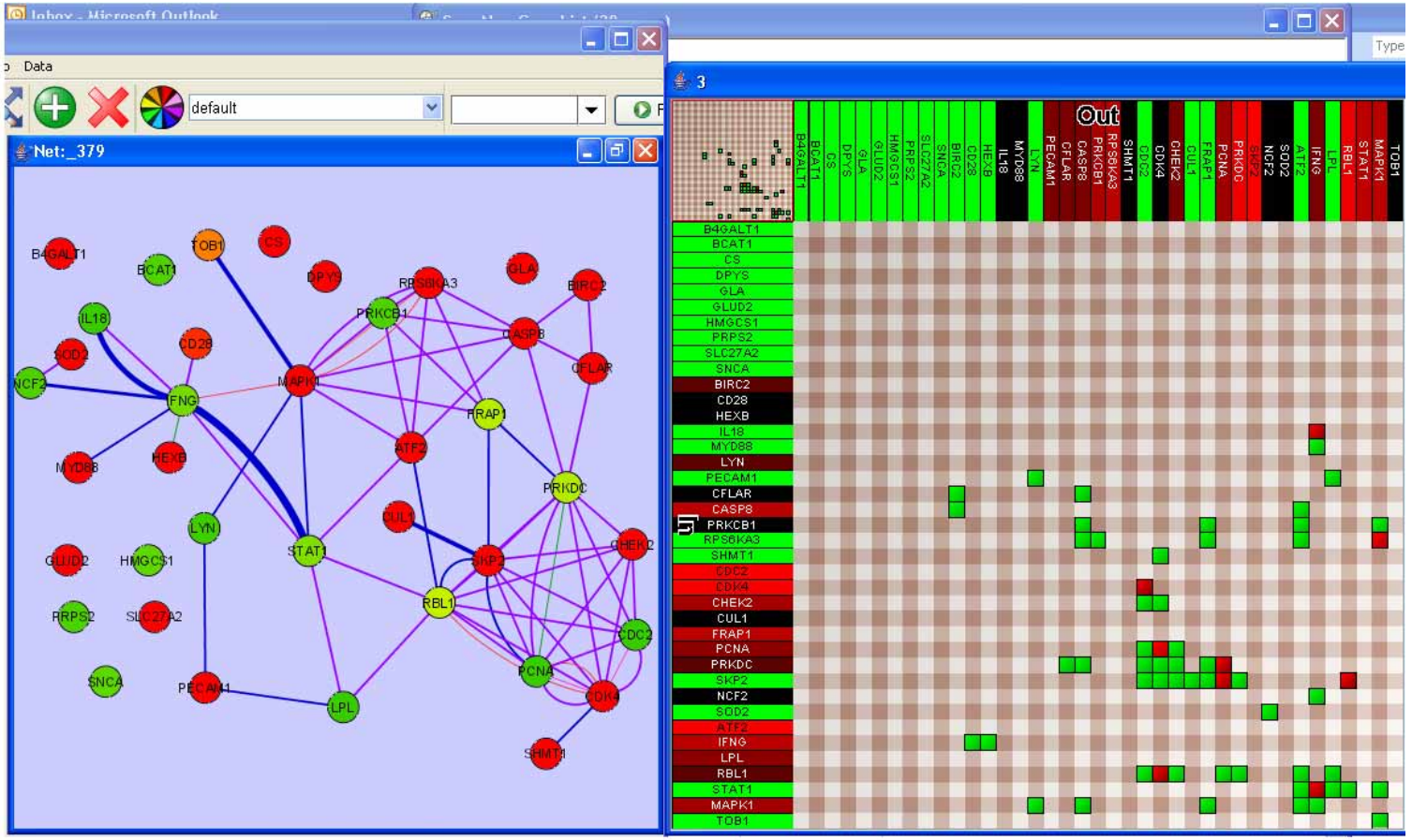


$$w_{jk} = \begin{bmatrix} 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 3 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -5 \\ 1 & -2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \end{bmatrix}$$

Directed and weighted

For more information: Diestel, R. (2010) *Graph Theory, 4th Edition. Berlin, Heidelberg, Springer.*

Example: Tool for Node-Link Visualization



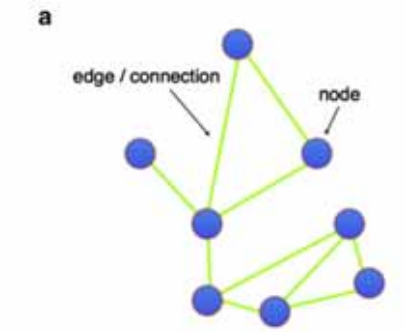
Jean-Daniel Fekete http://wiki.cytoscape.org/InfoVis_Toolkit

Fekete, J.-D. The infovis toolkit. Information Visualization, INFOVIS 2004, 2004. IEEE, 167-174.

Some Network Metrics (1/2)

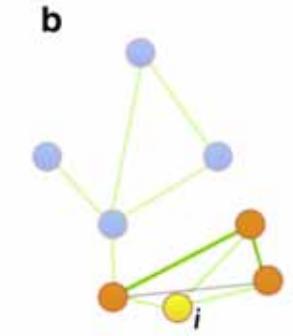
Order = total number of nodes n; **Size** = total number of links (a):

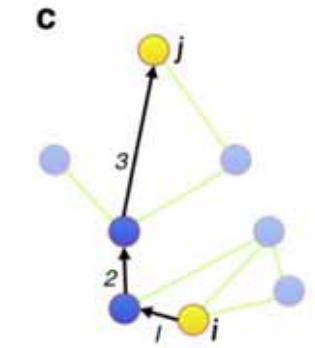
$$\sum_i \sum_j a_{ij}$$



Clustering Coefficient (b) = the degree of concentration of the connections of the node's neighbors in a graph and gives a measure of local inhomogeneity of the link density:

$$C_i = \frac{2t_i}{k(k_i - 1)} \qquad C = \frac{1}{n}\sum_i C_i$$





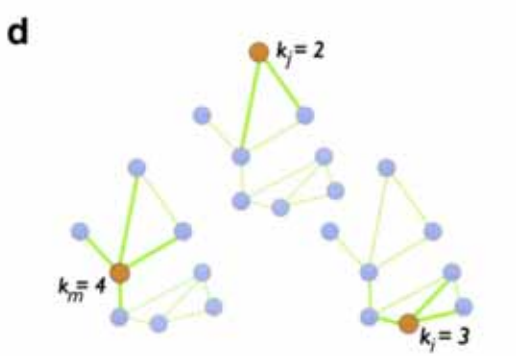
Path length (c) = is the arithmetical mean of all the distances:

$$l = \frac{1}{n(n-1)}\sum_{i \neq j} d_{ij}$$

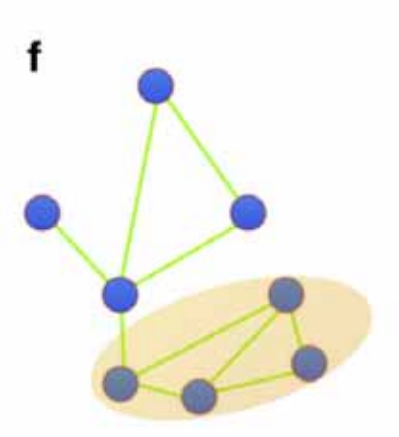
Costa, L. F., Rodrigues, F. A., Travieso, G. & Boas, P. R. V. (2007) Characterization of complex networks: A survey of measurements. *Advances in Physics, 56, 1, 167-242.*

Some Network Metrics (2/2)

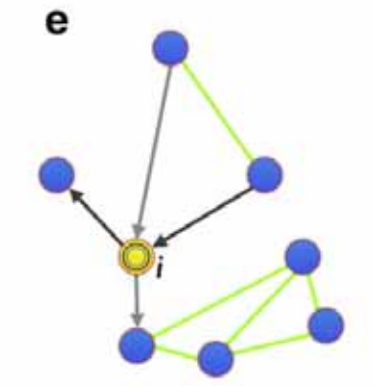
- **Centrality** (d) = the level of "betweenness- centrality" of a node I ("hub-node in Slide 28);



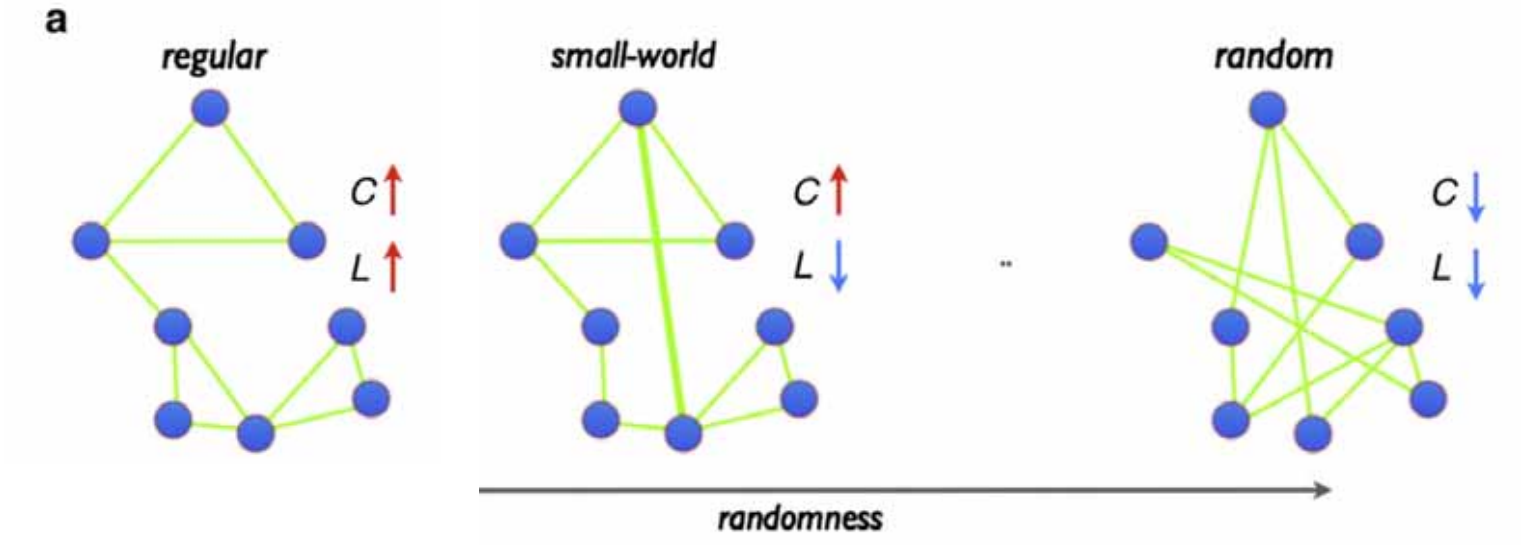
- **Nodal degree** (e) = number of links connecting *i* to its neighbors: $k_i = \sum_i a_{ij}$

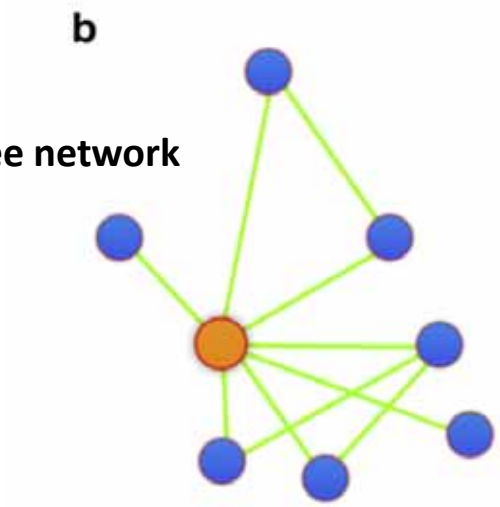


Modularity (f) = describes the possible formation of communities in the network, indicating how strong groups of nodes form relative isolated sub-networks within the full network (refer also to Slide 5-8).



Network Topologies

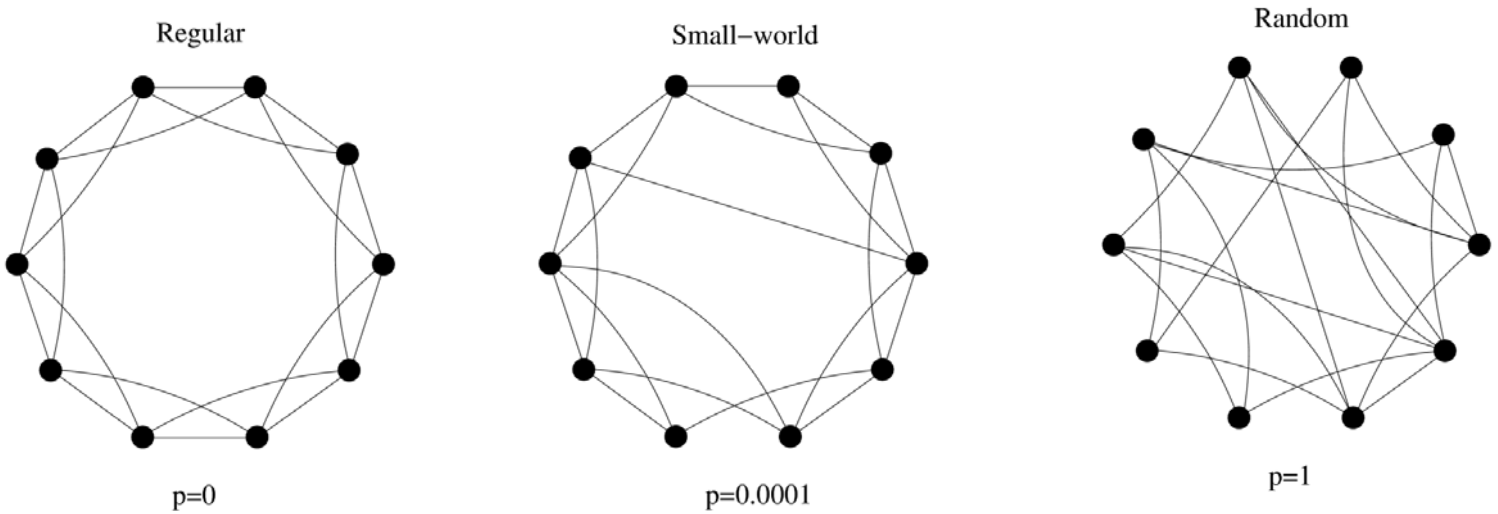




Scale-free network

Van Heuvel & Hulshoff (2010)

Small-World Networks

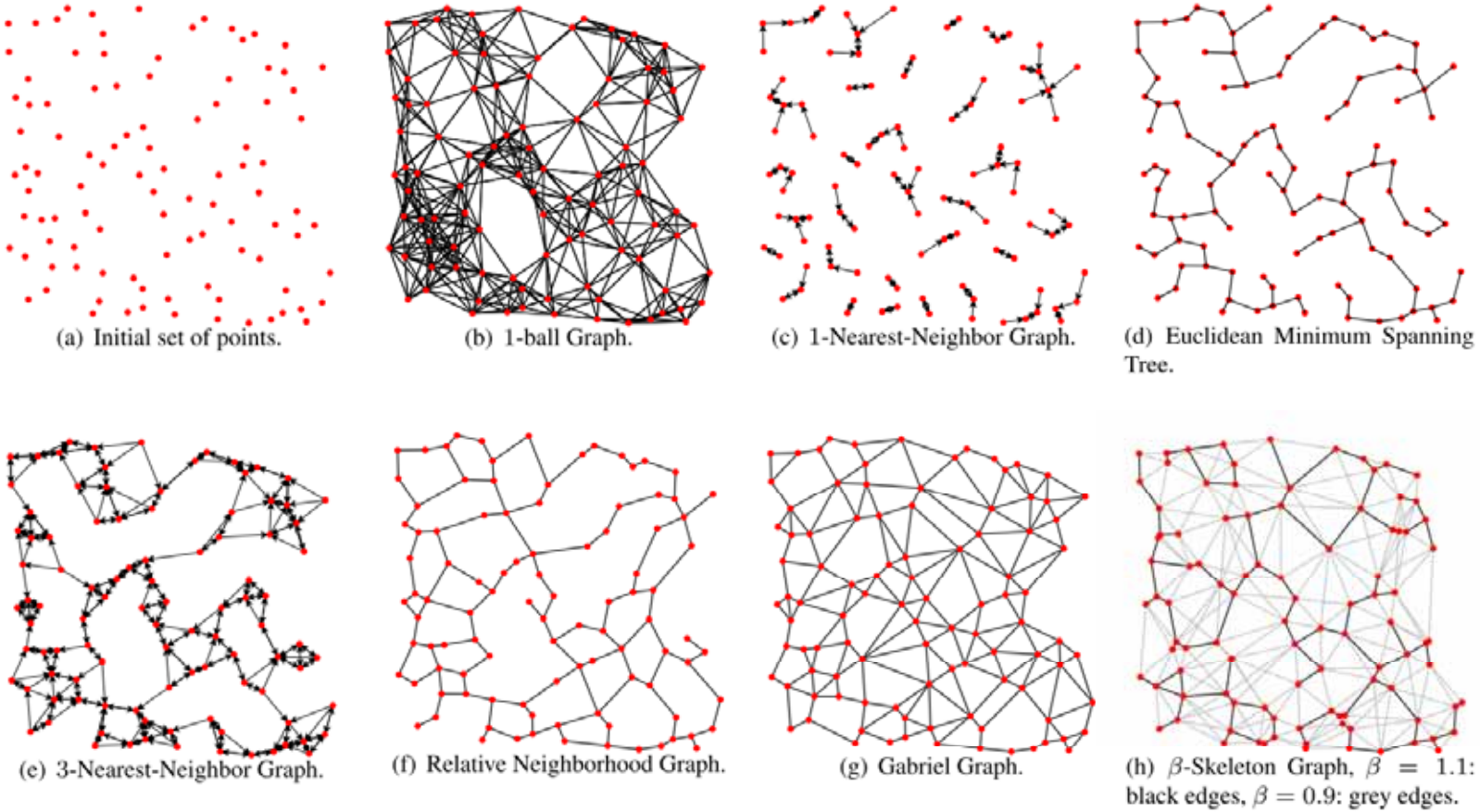


Increasing randomness

29.000 citations …

Watts, D. J. & Strogatz, S. (1998) Collective dynamics of small-world networks. *Nature, 393, 6684, 440-442.*

Milgram, S. 1967. The small world problem. *Psychology today, 2, (1), 60-67.*



(a) Initial set of points. (b) 1-ball Graph. (c) 1-Nearest-Neighbor Graph. (d) Euclidean Minimum Spanning Tree.

(e) 3-Nearest-Neighbor Graph. (f) Relative Neighborhood Graph. (g) Gabriel Graph. (h) β-Skeleton Graph, $\beta = 1.1$: black edges, $\beta = 0.9$: grey edges.

Lézoray, O. & Grady, L. 2012. Graph theory concepts and definitions used in image processing and analysis. *In: Lézoray, O. & Grady, L. (eds.) Image Processing and Analysing With Graphs: Theory and Practice. Boca Raton (FL): CRC Press, pp. 1-24.*

What are the most important graph topological metrics at a glance ?

Category	Class	Metric	Symbol
Topological	**Distance**	Hopcount	$H_{A\to B}$
		Closeness	C_i
		Eccentricity	ε_i
		Diameter	D
		Radius	R
		Girth	γ
		Expansion	e_h
		Betweenness	B_i
		Ce. Pt. of Domimance	CPD
		Distortion	t
	Connection	Degree	d_i
		Entropy	H
		Joint Degree	$\Pr[d_i, d_j]$
		Assortativity	r
		Coreness	k_i
		Clique	$n-clique$
		Clustering C.	C
		Rich Club coefficient	Φ_k
		Giant component	G_C
		Reliability	$\kappa(G), \lambda(G), ...$
		Chromatic number	$\mathcal{X}$
	Spectra	Algebraic connectivity	μ_{N-1}
		Spectral radius	ρ
		Spectral partitioning	s_i
		Principal eigenvector	x_1

The anatomy of a large-scale hypertextual web search engine

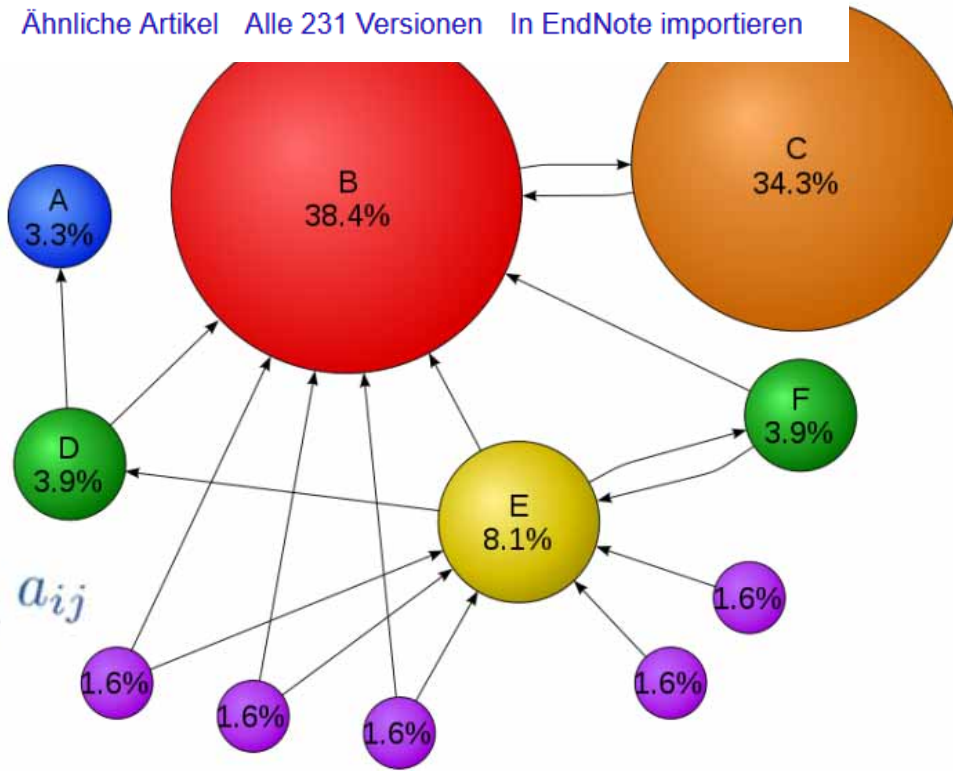
S Brin, L Page - Computer networks and ISDN systems, 1998 - Elsevier

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The ...

☆ 99 Zitiert von: 20933 Ähnliche Artikel Alle 231 Versionen In EndNote importieren

$$Ax_1 = \rho x_1$$

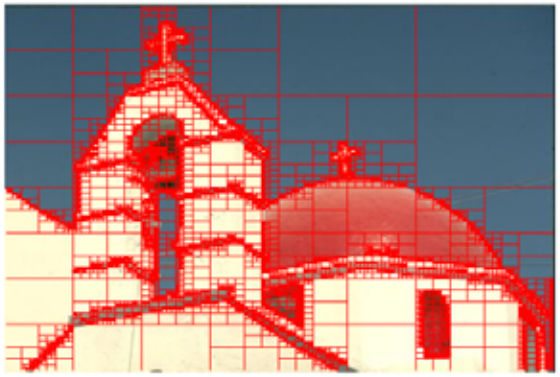
$$p = a_{ij} / \sum_{j=1}^{N} a_{ij}$$



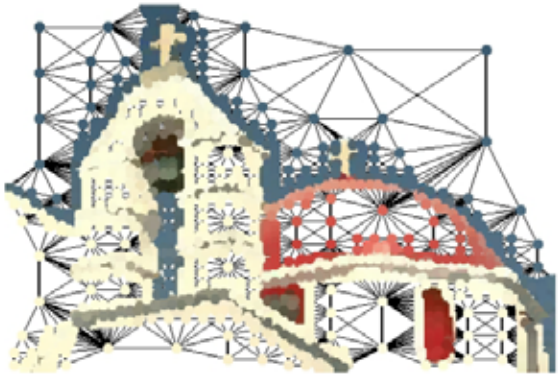
Javier Martın Hernández & Piet Van Mieghem (2011). Classification of graph metrics. Technical Report: Delft University of Technology, 1-20.

06 How do you get point cloud data from natural images – and then a graph?

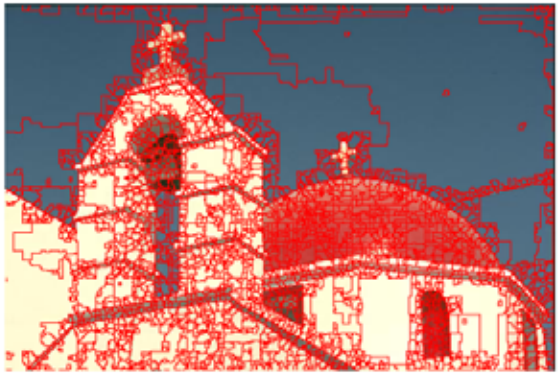
Graphs from Images



a) quadtree tessellation



b) RAG assoc. to the quadtree



c) Watershed Algorithm

d) SLIC superpixels

Lézoray, O. & Grady, L. 2012. Graph theory concepts and definitions used in image processing and analysis. *In: Lézoray, O. & Grady, L. (eds.) Image Processing and Analysing With Graphs: Theory and Practice. Boca Raton (FL): CRC Press, pp. 1-24.*

Example Watershed Algorithm

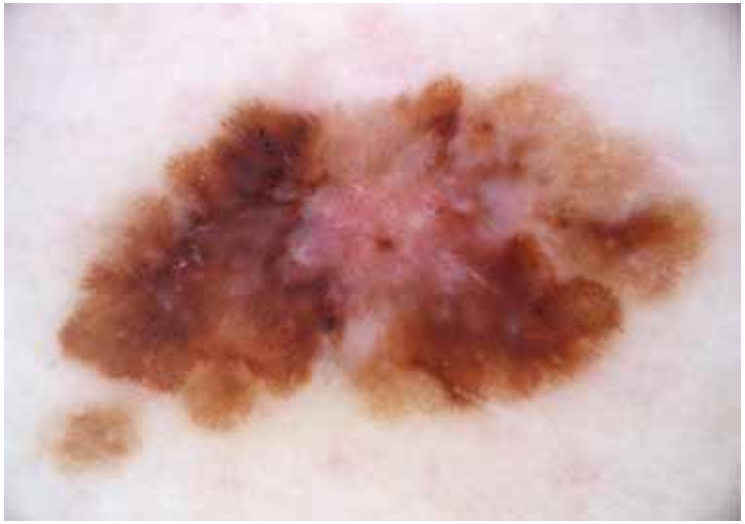
Algorithm 4.2 Watershed transform w.r.t. topographical distance based on image integration via the Dijkstra-Moore shortest paths algorithm.

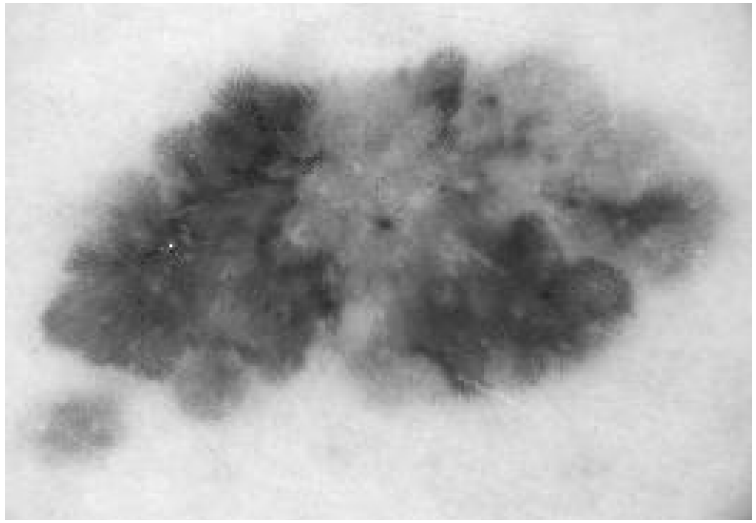
```
1: procedure ShortestPathWatershed;
2: INPUT: lower complete digital grey scale image G = (V, E, im) with cost function cost.
3: OUTPUT: labelled image lab on V.
4: #define WSHED 0                    (* label of the watershed pixels *)
5: (* Uses distance image dist. On output, dist[v] = im[v], for all v ∈ V. *)
6:
7: for all v ∈ V do                   (* Initialize *)
8:    lab[v] ← 0 ; dist[v] ← ∞
9: end for
10: for all local minima m_i do
11:    for all v ∈ m_i do
12:       lab[v] ← i ; dist[v] ← im[v]      (* initialize distance with values of minima *)
13:    end for
14: end for
15: while V ≠ ∅ do
16:    u ← GetMinDist(V)              (* find u ∈ V with smallest distance value dist[u] *)
17:    V ← V\{u}
18:    for all v ∈ V with (u, v) ∈ E do
19:       if dist[u] + cost[u, v] < dist[v] then
20:          dist[v] ← dist[u] + cost(u, v)
21:          lab[v] ← lab[u]
22:       else if lab[v] ≠ WSHED and dist[u] + cost[u, v] = dist[v] and lab[v] ≠ lab[u] then
23:          lab[v] = WSHED
24:       end if
25:    end for
26: end while
```

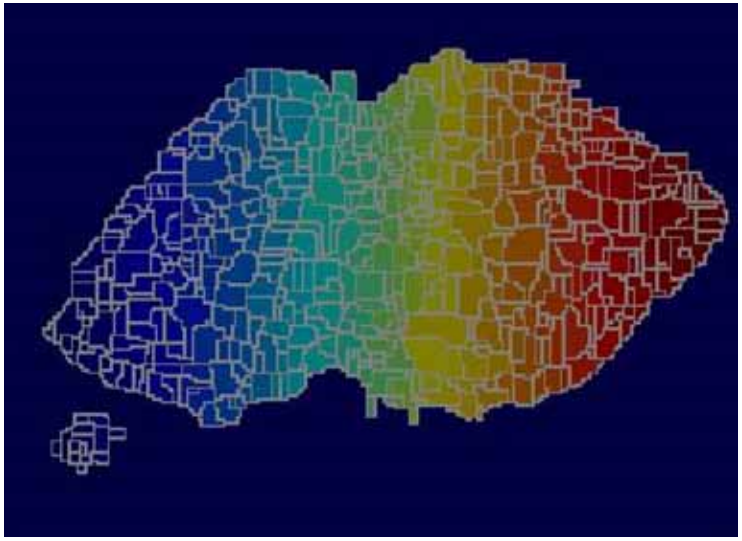
Meijster, A. & Roerdink, J. B. A proposal for the implementation of a parallel watershed algorithm. Computer Analysis of Images and Patterns, 1995. Springer, 790-795.

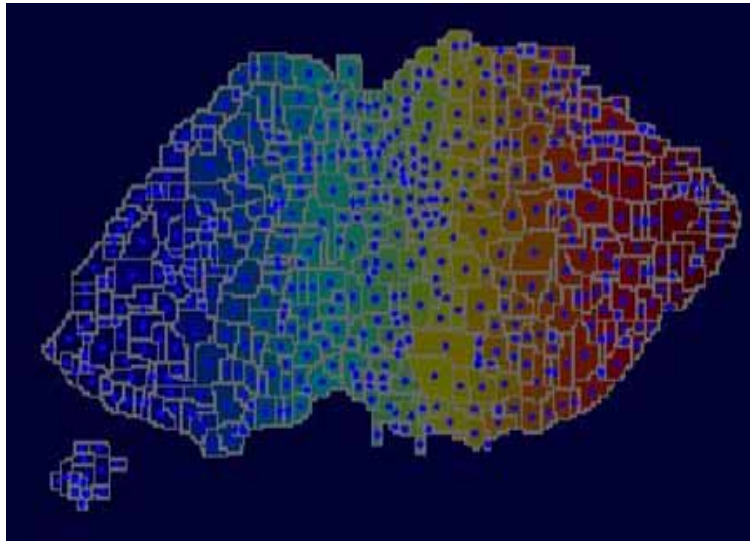
Example Image from Dermoscopy

Graphs from Images: Watershed + Centroid

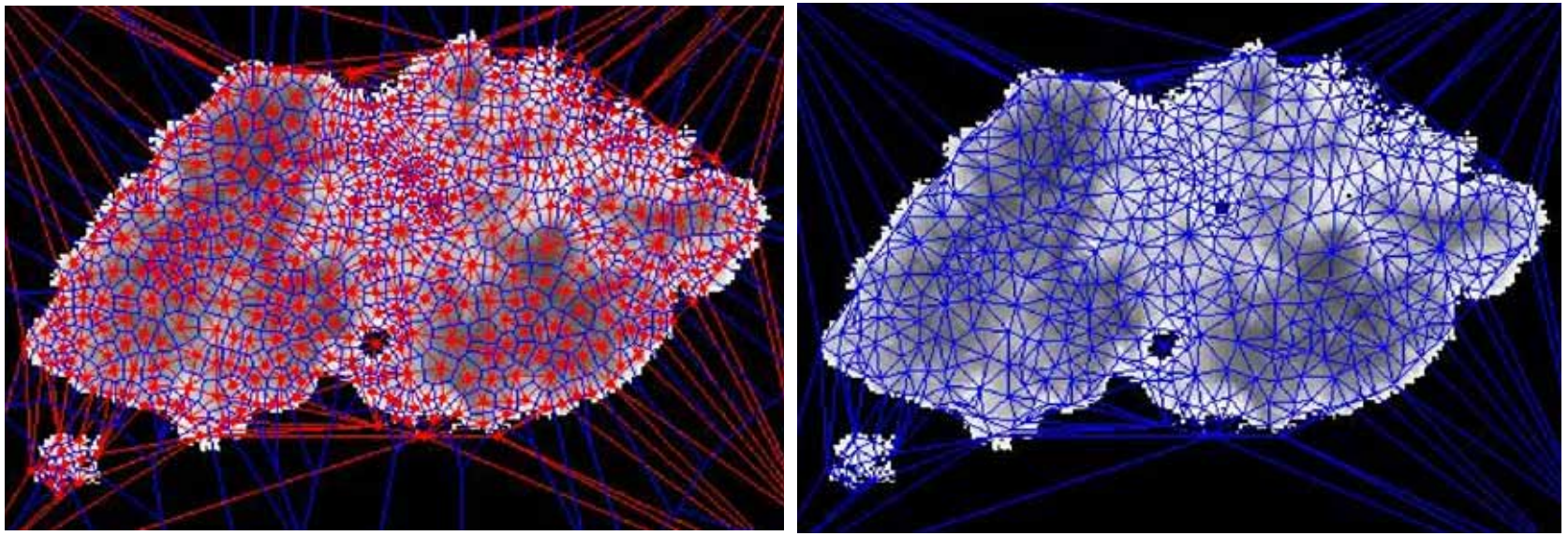








Graphs from Images: Voronoi <> Delaunay



Holzinger, A., Malle, B. & Giuliani, N. 2014. On Graph Extraction from Image Data. In: Slezak, D., Peters, J. F., Tan, A.-H. & Schwabe, L. (eds.) Brain Informatics and Health, BIH 2014, Lecture Notes in Artificial Intelligence, LNAI 8609. Heidelberg, Berlin: Springer, pp. 552-563.

For Voronoi please refer to: Aurenhammer, F. 1991. Voronoi Diagrams - A Survey of a fundamental geometric data structure. *Computing Surveys,* 23, (3), 345-405.

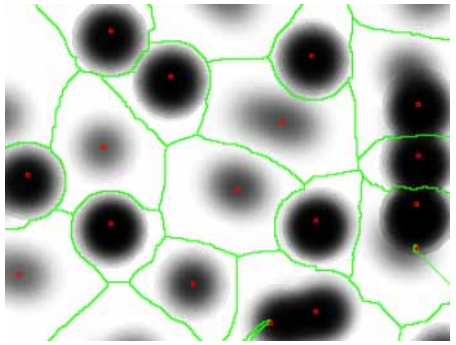
For Delaunay please refer to: Lee, D.-T. & Schachter, B. J. 1980. Two algorithms for constructing a Delaunay triangulation. Intl. Journal of Computer & Information Sciences, 9, (3), 219-242.

Are graphs better than feature vectors ?

- More expressive data structures
- Find novel connections between data objects
- Fit for applying graph based machine learning techniques
- New approaches (Belief Propagation, global understanding from local properties)

Bunke, H.: Graph-based tools for data mining and machine learning. In Perner, P., Rosenfeld, A., eds.: Machine Learning and Data Mining in Pattern Recognition, Proceedings. Volume 2734 of Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin, (Berlin) 7–19
Holzinger, A., Blanchard, D., Bloice, M., Holzinger, K., Palade, V., Rabadan, R.: Darwin, lamarck, or baldwin: Applying evolutionary algorithms to machine learning techniques. In: The 2014 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2014), IEEE (2014) in print

Watershed methods



- Topographic maps => landscapes with height structures
- Segmentation into regions of pixels
- Assuming drops of water raining on the map
- Following paths of descent
- Lakes called catchment basins
- Also possible: Flooding based
- Needs Topographical distance measures (MST)

Vincent, L. & Soille, P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE transactions on pattern analysis and machine intelligence, 13, (6), 583-598.

Watershed 4 Steps

- 1) Transformation into a topographic map
  - Convert gray values into height information
- 2) Finding local minima
  - Inspecting small regions in sequence
- 3) Finding catchment basins
  - Algorithm simulating flooding
  - Graph algorithms such as Minimum Spanning Trees
- 4) Erecting watersheds
  - Artificial divide between catchment basins
  - Final segmentation lines

Watershed Algorithm based on connected components

7	4	8	12	11	3
7	7	8	12	11	7
13	13	15	16	16	13
19	19	18	17	15	7
20	18	17	16	15	5

(a) The original image

→	m	←	←	→	m
↗	↑	↖	←	↗	↑
↑	↑	↖	↖	↗	↑
↑	↑	↑	→	↘	↓
→	→	→	→	→	m

(b) Each pixel connect to lowest minimum

0	0	0	0	1	1
0	0	0	0	1	1
0	0	0	0	1	1
0	0	0	2	2	2
2	2	2	2	2	2

(c) The Image with labels

Connects each pixel to the lowest neighbor pixel, all pixel connected to same lowest neighbor pixel form a segment

Challenges

- We want to find “interesting” novel patterns (rules, anomalies, outliers, similarities, …)
- Problem #1: How to get a graph?
- Problem #2: How do graphs evolve?
- Problem #3: What tools to apply?
- Problem #4: Scalability to TB, PB, EB …
- **Success is in repeatability and scalability**

State-of-the-Art Facts

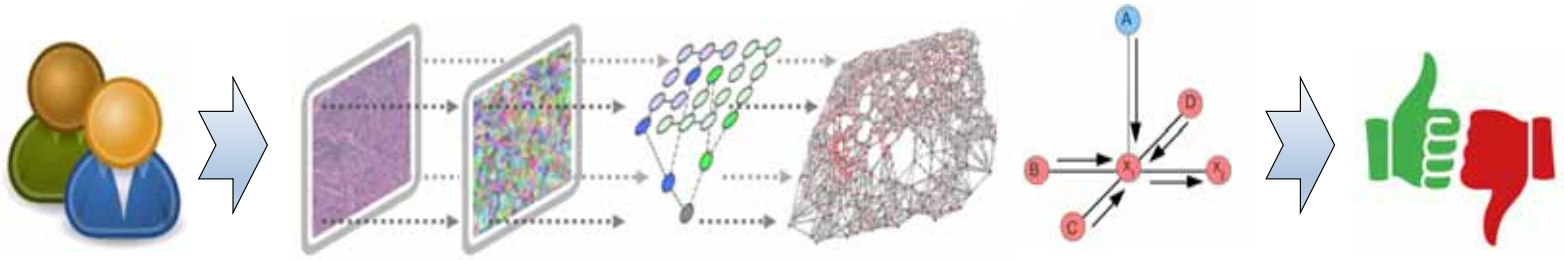
- Study of complex networks started in the 1990s with the insight that real networks contain properties not present in random (Erdös-Renyi) networks.
- Meanwhile networks and network-based approaches form an integral part of many studies throughout the sciences.
- Graph-Theory provides powerful tools to organize data structurally and in combination with statistical and machine learning methods allows a meaningful analysis of underlying processes.
- For instance, a mapping of causal disease genes and disorders as made available by the OMIM database provided novel insights into disease patterns, as recently demonstrated by investigating the diseasome (http://diseasome.eu).

07 Browser based Graph extraction from pixel images

The processing pipeline…

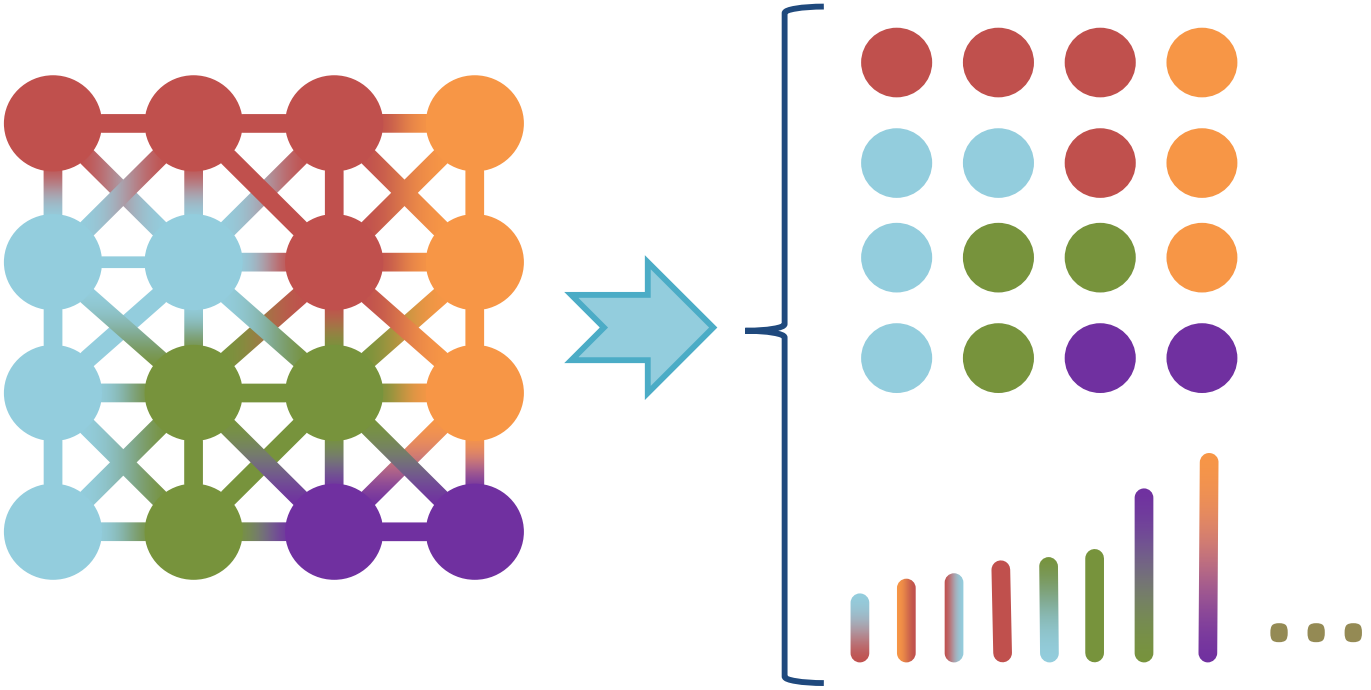
... Consists of 4 stages:

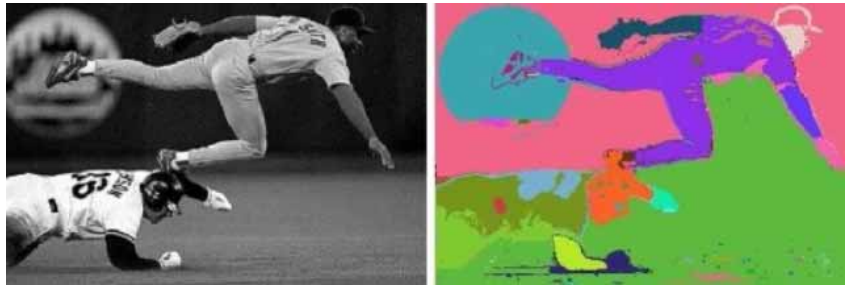
1. Image preprocessing
2. Algorithmic Preprocessing
3. Image Segmentation (Region Merging)
4. Graph extraction
5. Graph based (tumor) classification



Holzinger, Andreas, Bernd Malle, and Nicola Giuliani. "On graph extraction from image data." *International Conference on Brain Informatics and Health*. Springer International Publishing, 2014.

Graph => Node List + sorted Edge List

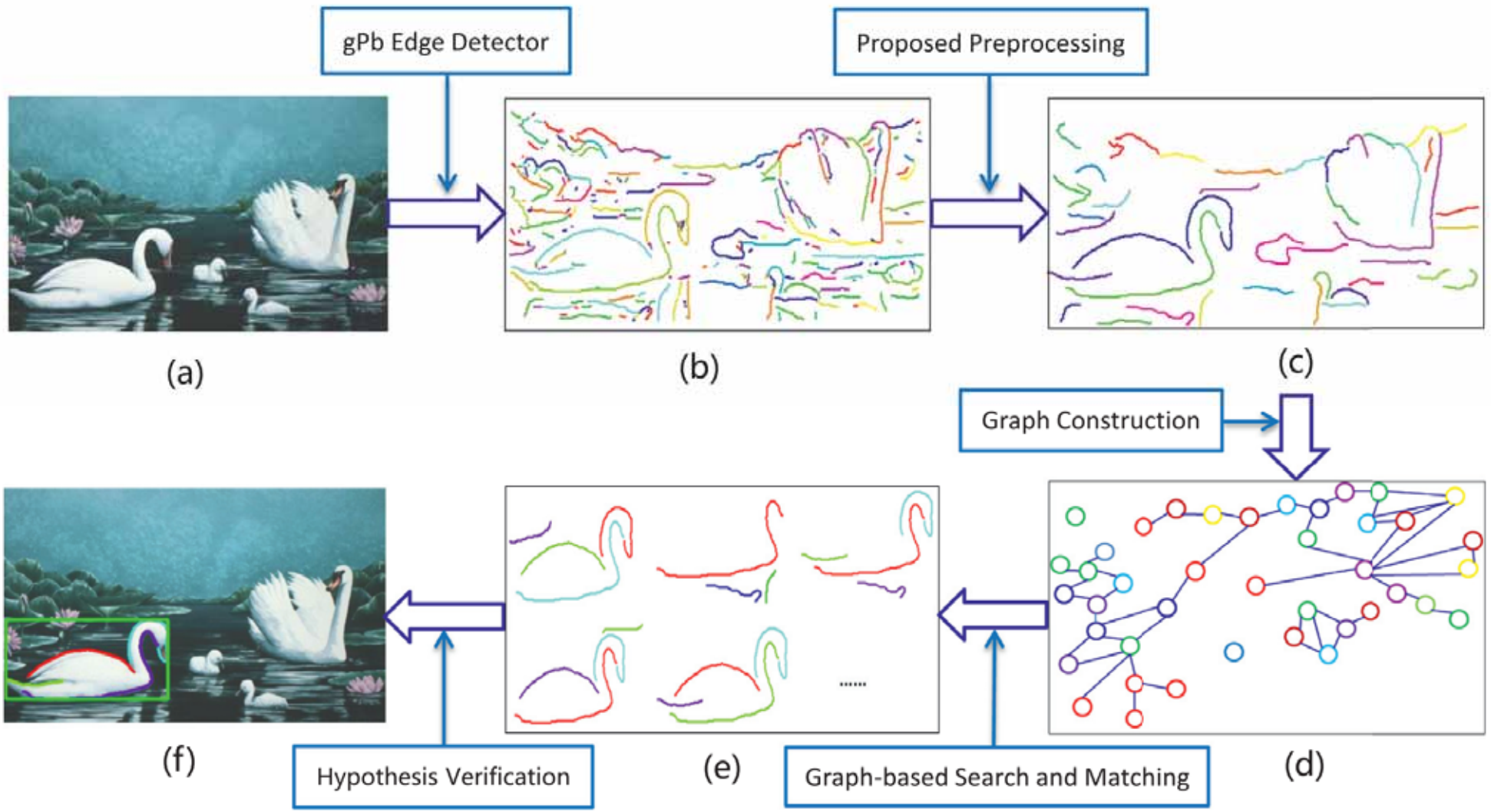




- Based on Kruskals MST algorithm
- Takes input image as natural graph with vertices := pixels and edges := pixel neighborhoods
- Visits edges in ascending order of weight and merges regions if they satisfy a certain criterion
- Flexible as merging criterion can be adapted as desired (for amount, size, or shape of regions)

Pedro F. Felzenszwalb & Daniel P. Huttenlocher (2004). Efficient graph-based image segmentation. International Journal of Computer Vision, 59, (2), 167-181, doi:10.1023/B:VISI.0000022288.19776.77.

How does a graph construction pipeline typically look like ?



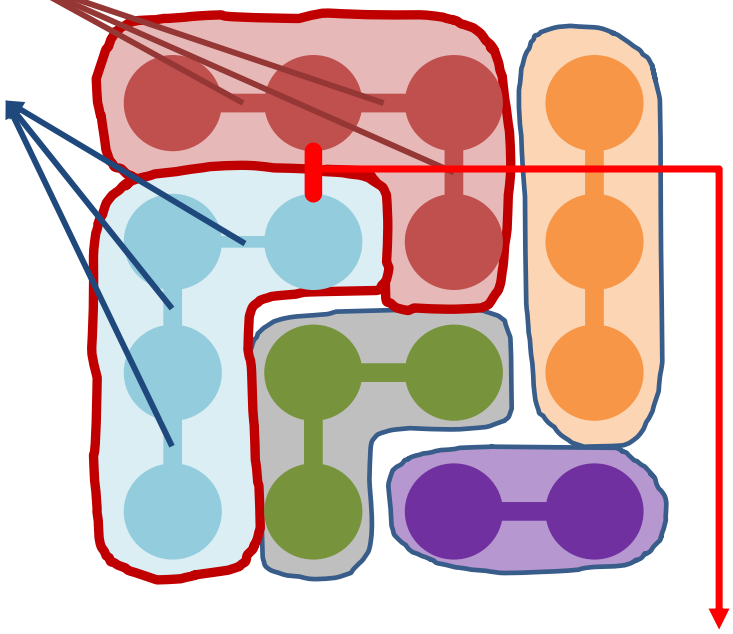
Hui Wei, Chengzhuan Yang & Qian Yu (2017). Efficient graph-based search for object detection. Information Sciences, 385, 395-414, doi:http://dx.doi.org/10.1016/j.ins.2016.12.039.

Kruskal based Region Merging

$$\mathrm{D}(C1, C2) = \begin{cases} true & \text{if } \mathrm{Dif}(C1, C2) > \mathrm{MInt}(C1, C2) \\ false & \text{otherwise} \end{cases}$$

$$\mathrm{Int}(C) = \max_{e \in \mathrm{MST}(C,E)} \omega(e)$$

$$\mathrm{Int}(C) = \max_{e \in \mathrm{MST}(C,E)} \omega(e)$$



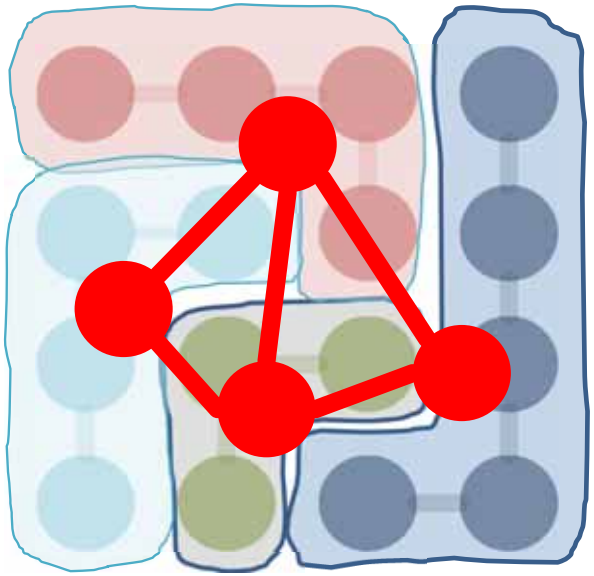
Merge!!!

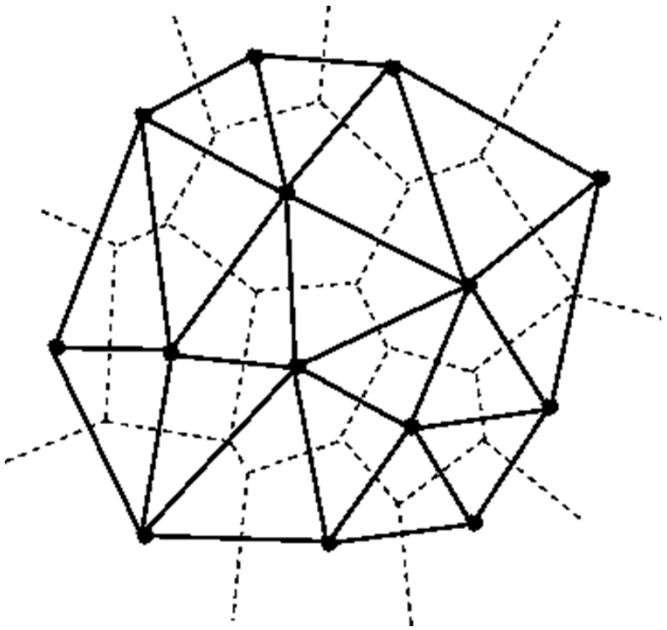
In this case:

- D(C1, C2) = true
- No merge !!!

$$\mathrm{Dif}(C1, C2) = \min_{v_i \in C1, v_j \in C2, (v_i, v_j) \in E} \omega(v_i, v_j)$$

Label map => Delaunay triangulation

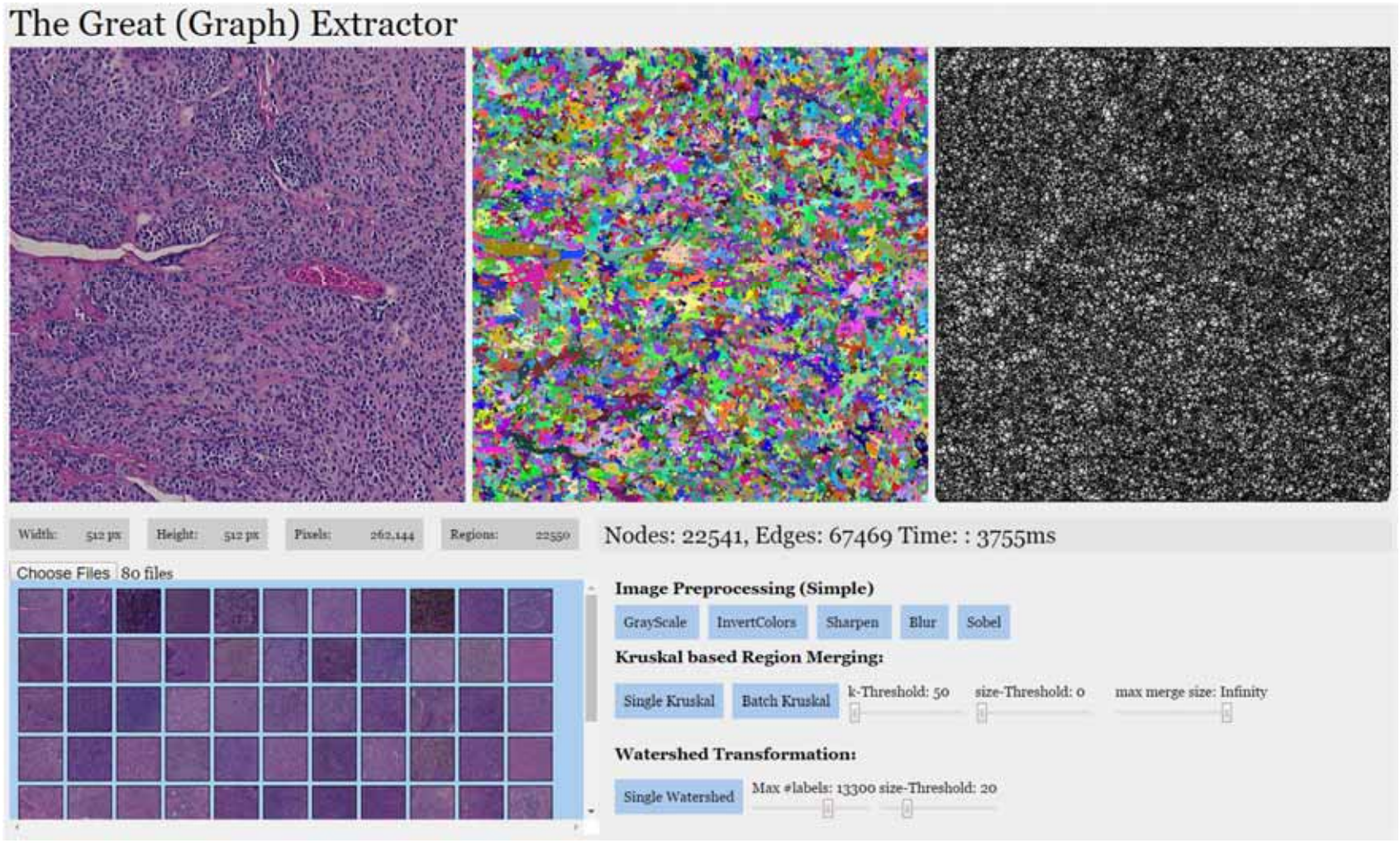




Voronoi / ***Delaunay*** triangulation (tessellation)

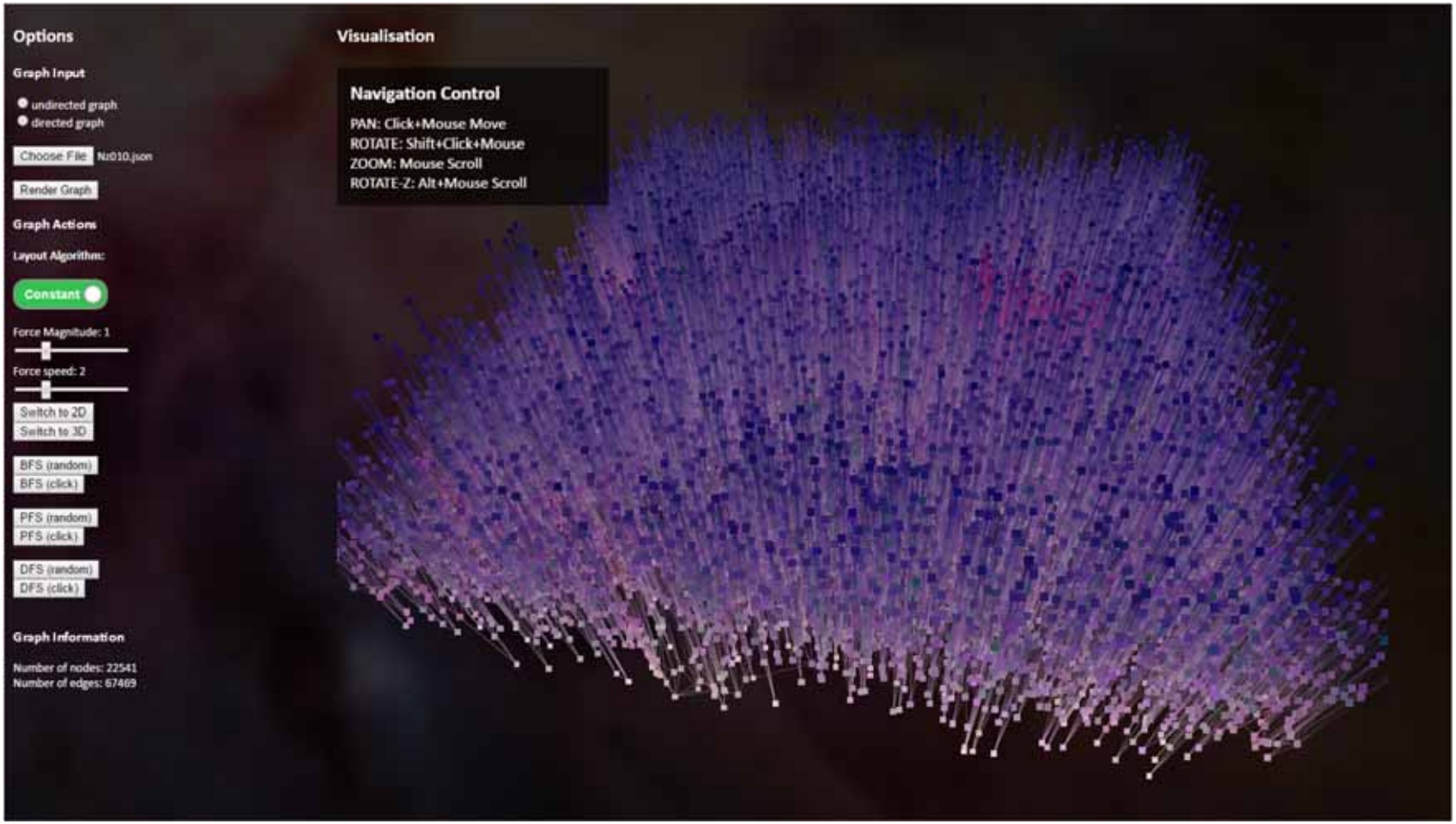
Online Graph extraction tool

http://berndmalle.com/graphext



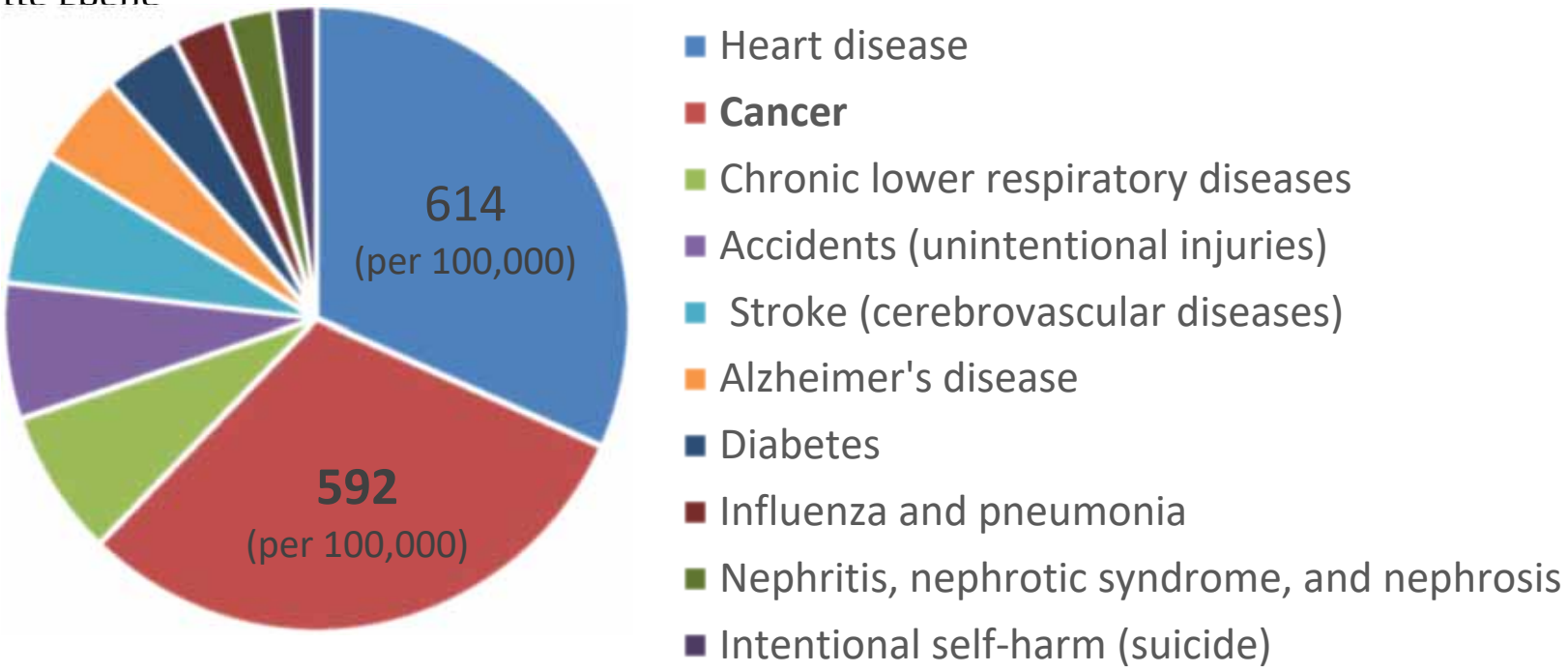
Online 3D Graph visualization tool

http://berndmalle.com/GraphiniusVis



08 Application (Web based) Tumor Growth Visualization

- Textmasterformat bearbeiten
  - Zweite Ebene
    - Dritte Ebene



Health US, 2015 – Leading Causes of Death – Centers of Disease Control and Prevention (CDC), 2015, URL: *http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm*

Key Problems

- Inter- and intracellular **dynamics**
- avoiding **hard-to-measure** variables
- **Inflexible** models
- *in silico* ***complements*** *in vivo*
- *executable (cell) biology*
- **reduce** animal experiments (resources)
- **boost** *in silico* for awareness & breakthrough
- patient-**personalized** prediction

Edelman, L. B., Eddy, J. A. & Price, N. D. 2010. In silico models of cancer. Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 2, (4), 438-459, doi:10.1002/wsbm.75.

Fisher, J. & Henzinger, T. A. 2007. Executable cell biology. Nature biotechnology, 25, (11), 1239-1249, doi:10.1038/nbt1356.

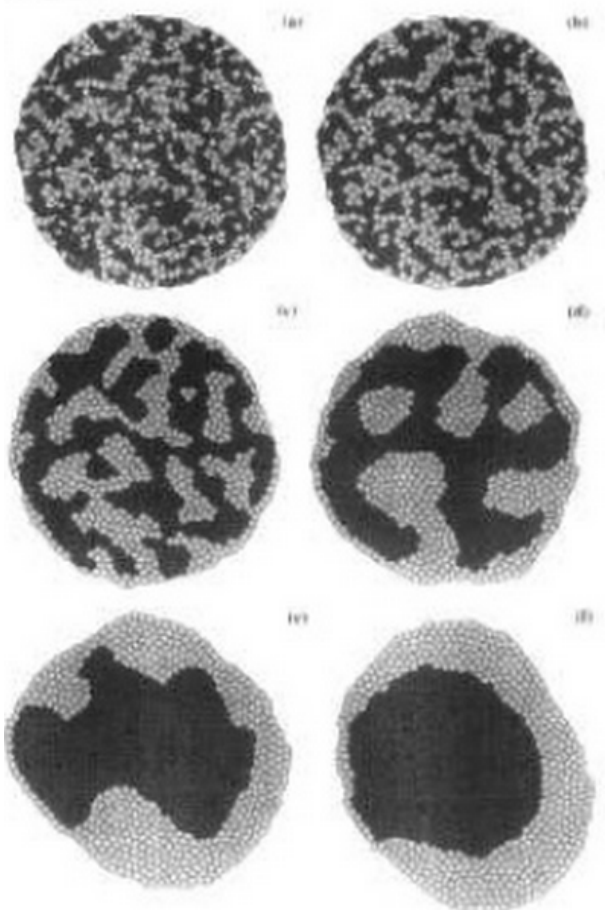


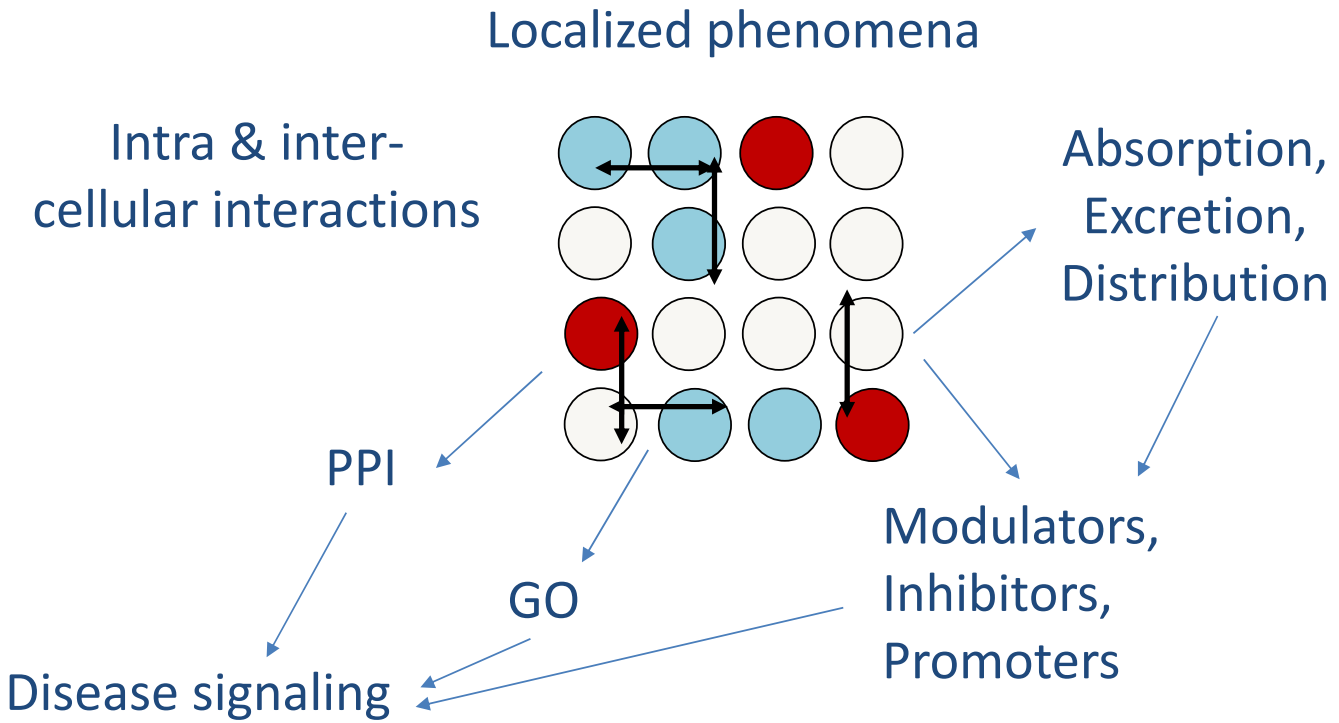
Image of a cell sorting time series

- initial: random assigned cell types
- each step represents a growing number of Monte Carlo Step (MCS)
- figure shows pattern

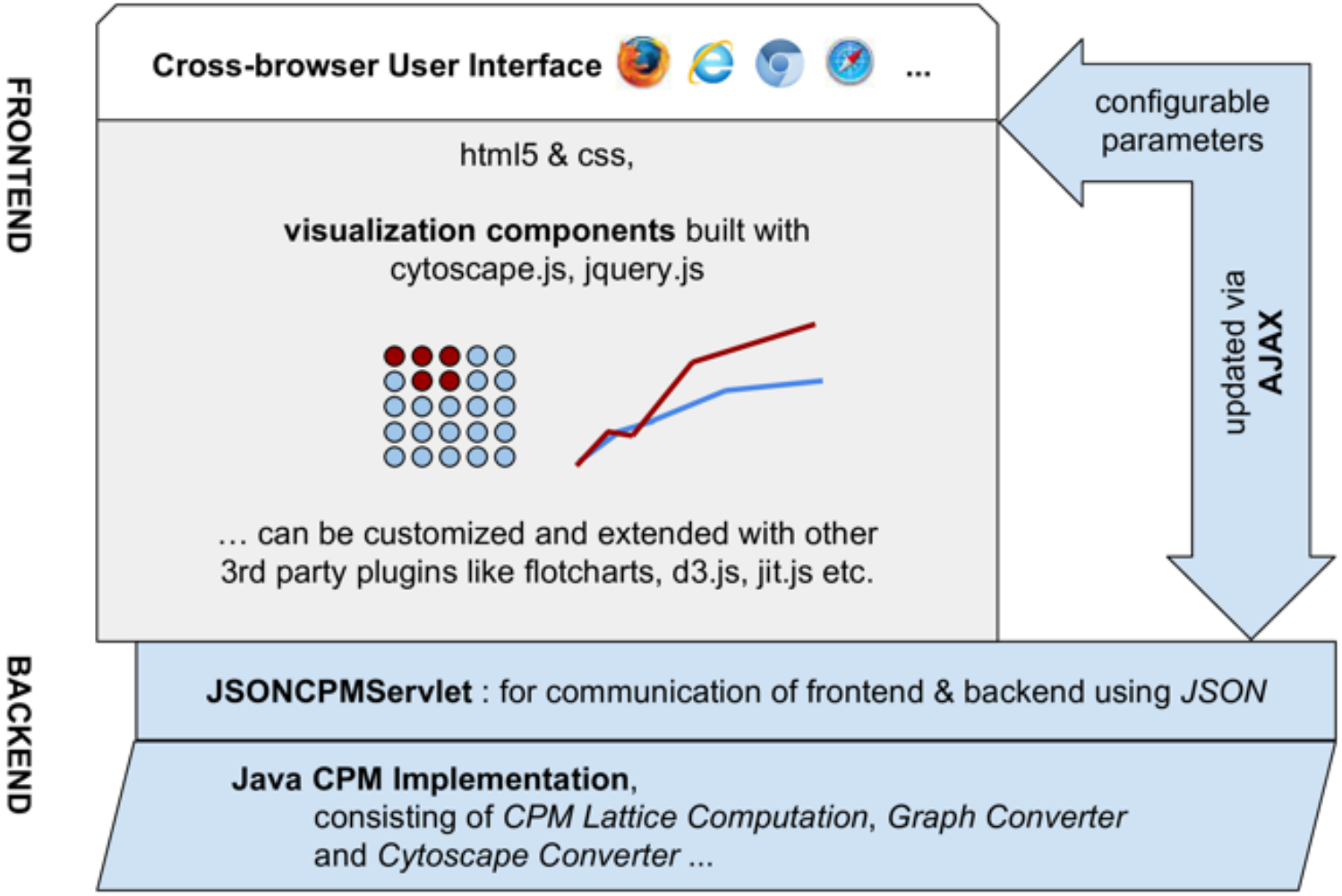
Graner, F., & Glazier, J. A. (1992). Simulation of biological cell sorting using a two-dimensional extended Potts model. Physical review letters, 69(13), 2013.

Nodes *as* ***Cellular bricks***

representing compartmental states

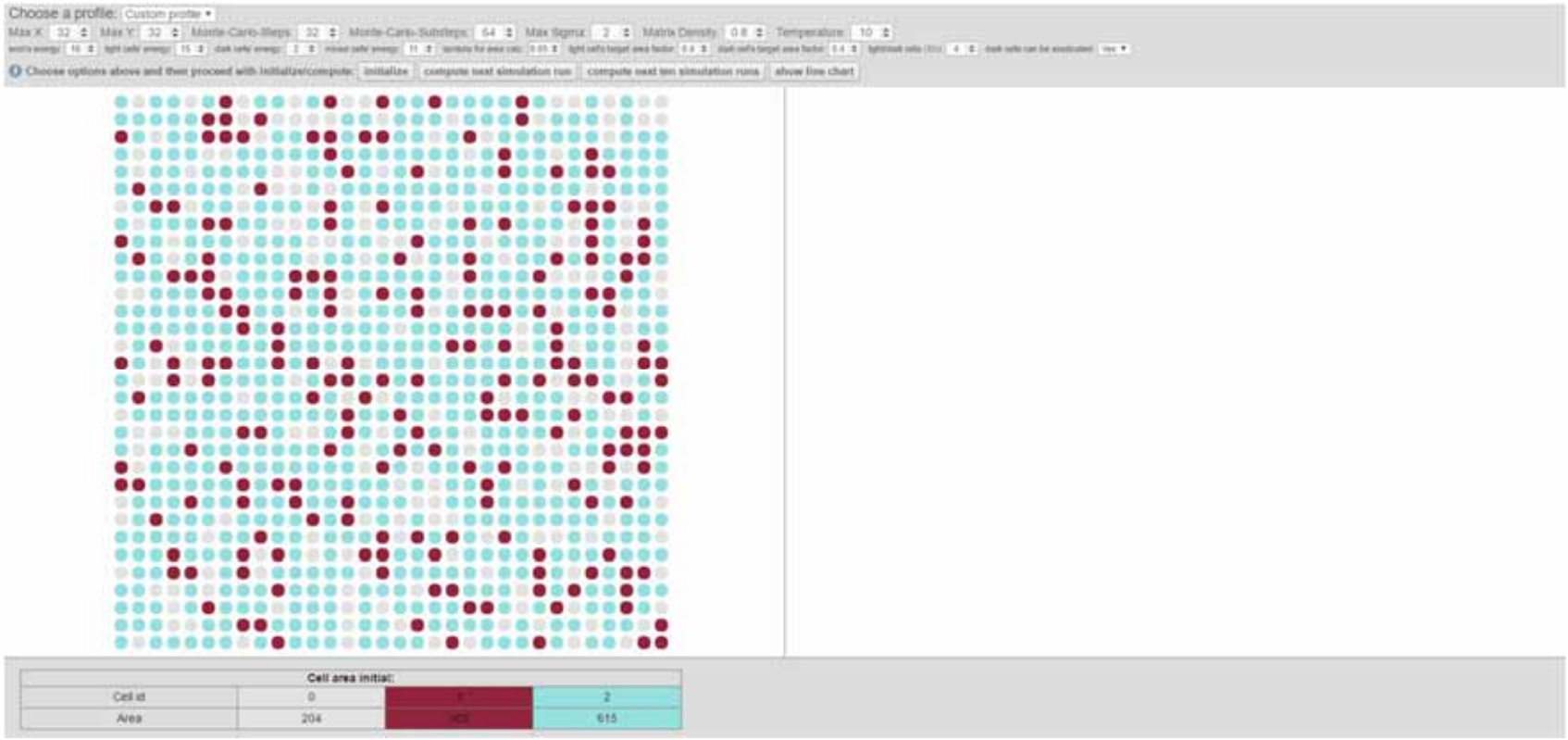


Implementation Example: Tumor Growth VIS



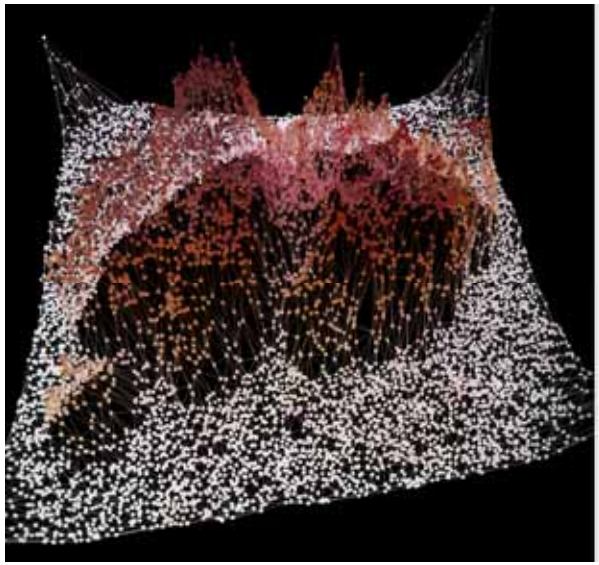
Online Tumor growth VIS tool

http://styx.cgv.tugraz.at:8080/cpm-cytoscape/



In silico modeling for tumor growth visualization – F. Jeanquartier, C. Jean-Quartier, D. Cemernek and A. Holzinger, BMC Systems Biology.2016, 10:59, DOI: 10.1186/s12918-016-0318-8
URL: http://www.biomedcentral.com/1752-0509/10/59

Project GRAPHINIUS

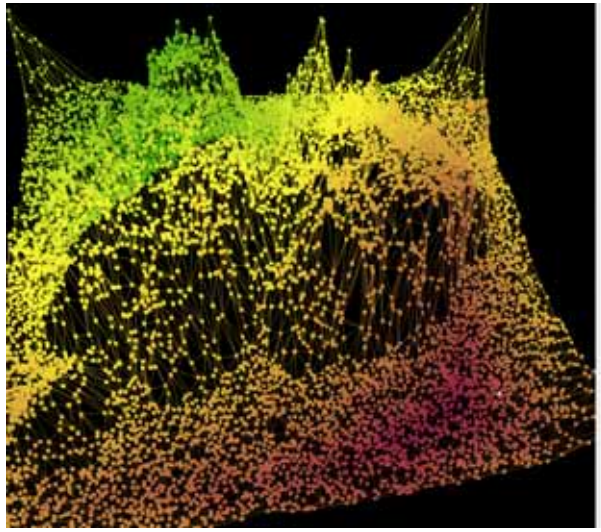


```
THREE.WebGLRenderer 74                                three.min.js:639
> var json_graph = "http://berndmalle.com/graphinius-
  demo/test_data/json/DERMATEST_1_1456756265259.json"
< undefined
> var jsonReader = new $G.input.JsonInput();
< undefined
> jsonReader.readFromJSONURL(json_graph, function(graph,
  err) {window.graph = graph})
< undefined
> graph
< e {_label: "DERMATEST_1.jpg", _nr_nodes: 10539,
  _nr_dir_edges: 0, _nr_und_edges: 31516, _mode: 2…}
    ▶ _dir_edges: Object
      _label: "DERMATEST
      _mode: 2
    ▶ _nodes: Object
      _nr_dir_edges: 0
      _nr_nodes: 10539
      _nr_und_edges: 31516
    ▶ _und_edges: Object
    ▶ __proto__: Object
> graph.degreeDistribution().all
< [0, 0, 0, 43, 850, 2855, 3644, 2119, 771, 212, 40, 5]
> $GV.core.render.renderGraph()
  rendering graph...                  graphinius.vis.js:238
```

Object

Online Graph exploration and analysis platform

- JS graph library
- Interactive console
- Real-time 2D/3D visualization
- Jupyter-style notebooks
- Platform for publishing / exchanging experiments



```
> var startNode = graph.getRandomNode()
< undefined
> var bfs = $G.search.BFS(graph, startNode)
< undefined
> Object.keys(bfs).length
< 10539
> bfs[0]
< ▶Object {distance: 23, parent: e, counter: 2209}
> bfs[0].parent
< ▶e {_id: 120, _in_degree: 0, _out_degree: 0, _und_degree:
  4, _in_edges: Object…}
> bfs[120]
< ▶Object {distance: 22, parent: e, counter: 2037}
> bfs[120].parent
< ▶e {_id: 511, _in_degree: 0, _out_degree: 0, _und_degree:
  4, _in_edges: Object…}
> bfs[511]
< ▶Object {distance: 21, parent: e, counter: 1865}
> $GV.core.mutate.colorBFS()
< undefined
```

Thank you!