Elements of Complex Networks

Pedro Larrañaga
Computational Intelligence Group
Artificial Intelligence Department
Universidad Politécnica de Madrid

Arantzazu, November 21, 2024

Outline

Techno-Social Systems as Complex Systems

Complex Systems

- It is made up of several interconnected components
- Their interaction is capable of creating new emergent properties
- These emergent properties are unobservable when the system is seen from a disaggregated perspective
- The human brain

Techno-Social Systems

- Infrastructures made up of different technological layers that interoperate within society to provide global public services of a technological nature
- They are complex systems consisting of large-scale (or cyber-) physical infraestructures
- The prediction of their behavior is based on the mathematical modeling of data collected in the real world

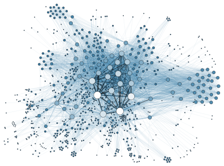
Complex Networks

Complex networks: mathematical paradigm to model techno-social systems

Graphical Representation of Complex Systems

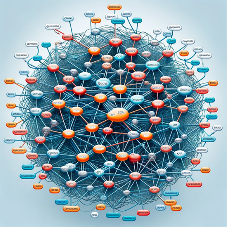
- Graph with non-trivial topological properties
- Nodes: elements of the system
- Edges: relationships between pairs of nodes

Example

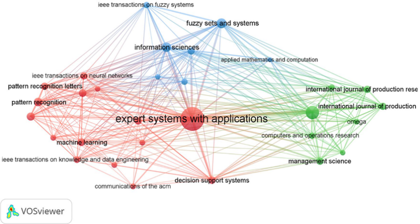


Examples of Complex Networks

Human connectome



Gene regulatory network

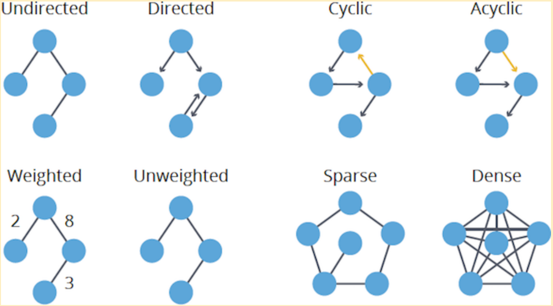


Co-citations of AI journals

Air transportation network

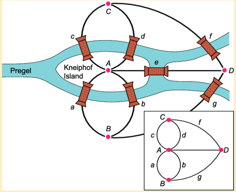
Types of Graphs

Undirected vs Directed. Cyclic vs Acyclic. Weighted vs Unweighted. Sparse vs Dense



Complex Networks. Mathematical Foundations

Graph Theory



- Euler and the seven bridge problem at Königsberg (1735)
- How can you across all seven bridges crossing only once for each one?
- The answer is that it is no possible

Probability Theory

Kolmogorov axioms of probability (1933)

- $P(A) \geq 0$ for any event A
- $P(\Omega) = 1$, for the sample space Ω
- $P(A \cup B) = P(A) + P(B)$ for two mutually exclusive events A and B

Complex Networks. Seminal Papers

On random graphs I.

Dedicated to O. Varga, at the occasion of his 50th birthday.

By P. ERDŐS and A. RÉNYI (Budapest).

Publicationes Mathematicae (1959)

RANDOM GRAPHS

BY E. N. GILBERT

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

Annals of Mathematical Statistics (1959)

Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Nature (1998)

Emergence of Scaling in Random Networks

Albert-László Barabási* and Réka Albert

Science (1999)

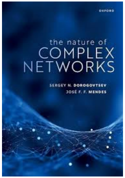
The large-scale organization of metabolic networks

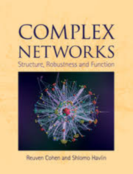
H. Jeong*, B. Tombor†, R. Albert*, Z. N. Oltvai† & A.-L. Barabási*

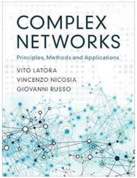
* *Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA*
† *Department of Pathology, Northwestern University Medical School, Chicago, Illinois 60611, USA*

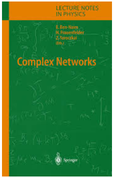
Nature (2000)

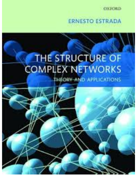
Complex Networks. Books

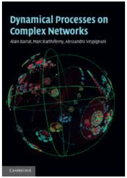


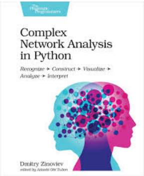


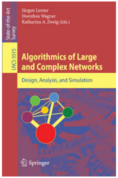


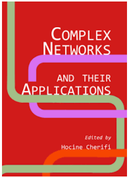


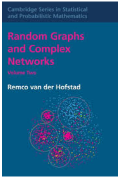


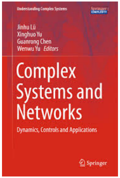


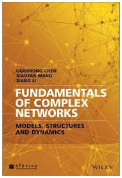


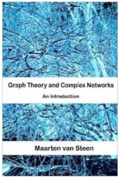


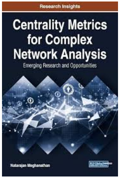








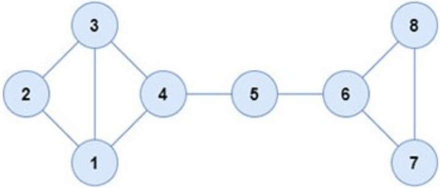




Outline

Complex Networks. Centrality measures

Node degree is the number of its neighbor nodes in the graph

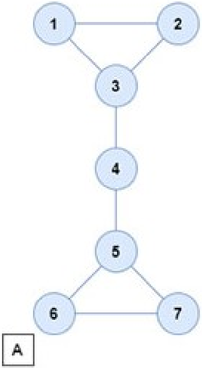


$$
\begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8 \end{array}
\begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\begin{bmatrix}
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0\\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0\\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0
\end{bmatrix} \end{array}
\times
\begin{bmatrix} 1\\1\\1\\1\\1\\1\\1\\1 \end{bmatrix}
=
\begin{bmatrix} 3\\2\\3\\3\\2\\3\\2\\2 \end{bmatrix}
\begin{array}{c} \text{Node} \\ 1\\2\\3\\4\\5\\6\\7\\8 \end{array}
$$

Node degree

Complex Networks. Centrality measures

Betweenness centrality for each node is the quotient between the number of geodesics (shortest paths) between any two nodes that pass through the given node and the number of geodesics that join any two nodes



Betweenness of node 3:

Node A	Node B	Value
4	1	1
4	2	1
5	1	1
5	2	1
6	1	1
6	2	1
7	1	1
7	2	1

Betweenness centrality of node 3: 8

Betweenness of node 5:

Node A	Node B	Value
6	1	1
6	2	1
6	3	1
6	4	1
7	1	1
7	2	1
7	3	1
7	4	1

Betweenness centrality of node 5: 8

Betweenness of node 4:

Node A	Node B	Value
5	1	1
5	2	1
5	3	1
6	1	1
6	2	1
6	3	1
7	1	1
7	2	1
7	3	1

Betweenness centrality of node 4: 9

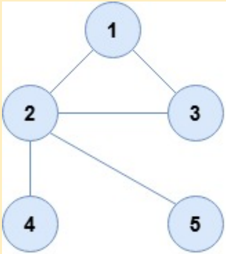
Betweenness centrality of node 1: 0

Betweenness centrality of node 2: 0

Betweenness centrality of node 6: 0

Betweenness centrality of node 7: 0

Closeness centrality of a node i measures how short the shortest paths are from node i to all nodes



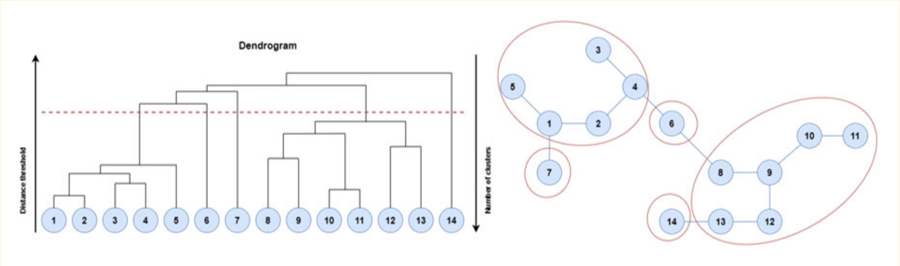
$n = 5$

$$\begin{array}{c|c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 0 & 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 \\ 2 & 1 & 2 & 0 & 2 \\ 2 & 1 & 2 & 2 & 0 \end{bmatrix} \end{array}$$

$\sum_j d(i,j)$	$CC(i)$
6	$\frac{(5-1)}{6} = 0.67$
4	$\frac{(5-1)}{4} = 1.00$
6	$\frac{(5-1)}{6} = 0.67$
7	$\frac{(5-1)}{7} = 0.57$
7	$\frac{(5-1)}{7} = 0.57$

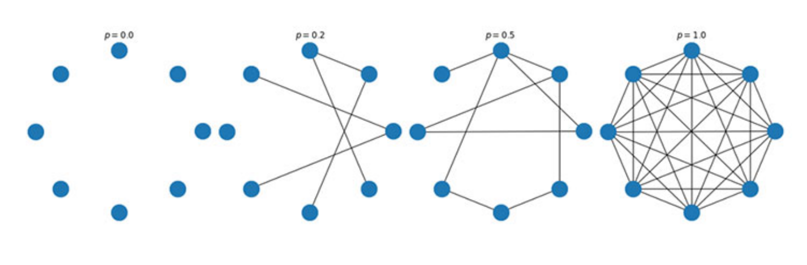
Complex Networks. Community detection

Community detection clustering algorithms are based on defining a distance between pairs of nodes to subsequently apply standard clustering methods



Complex Networks. Models

Erdös-Rényi-Gilbert model: each edge has a fixed probability of being present or absent, independently of the other edges

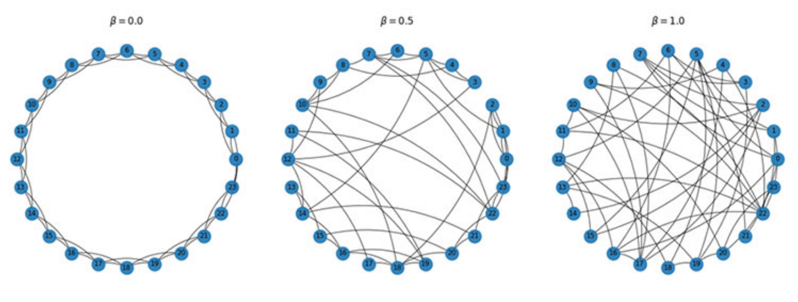


- The probability of generating a graph with n nodes and m edges is $p^m(1-p)^{\binom{n}{2}-m}$
- The average number of edges is $\binom{n}{2}p$
- The distribution of the degree of any particular node is binomial $\binom{n-1}{k}p^k(1-p)^{n-1-k}$

Complex Networks. Models

The two phases of the Watts-Strogatz generation model:

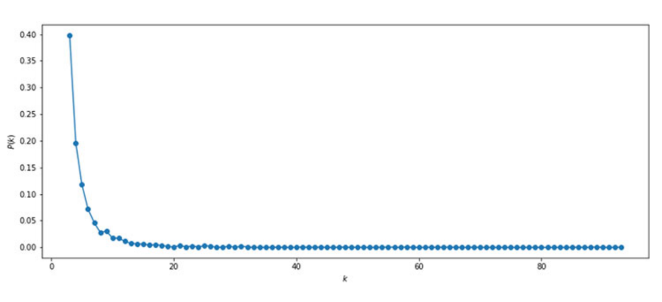
1. n ordered nodes connected each one of them with an even k number of nodes, the $\frac{k}{2}$ immediate predecessors and the $\frac{k}{2}$ following consecutive nodes
2. All edges are traversed, each one of them being 'rewired' towards another node chosen randomly, from the $n-1$ nodes, with β probability



Complex Networks. Models

The Barabási-Albert model generates random scale-free networks.

- The network starts with a set of $m_0 \geq m$ randomly connected nodes. This m is the only parameter
- New nodes are added to the network one by one. Each node is connected to m network nodes. For each of these nodes the probability p_i to be connected with node i is $p_i = \frac{k_i}{\sum_j k_j}$, where k_i is the degree of node i



The probability of a node in the network having k connections to other nodes follows a power law distribution: $P(k) \sim k^{-\gamma}$ with $2 < \gamma < 3$

Outline

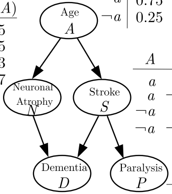
Bayesian Networks

DAG + CPTs

- Directed acyclic graph
- Conditional independence: **W** and **T** are conditionally independent given **Z** $\Leftrightarrow p(\mathbf{W}|\mathbf{T},\mathbf{Z}) = p(\mathbf{W}|\mathbf{Z})$
- $p(X_1, \ldots, X_n) = \prod_{i=1}^{n} p(X_i \mid \mathbf{Pa}(X_i))$

A	$p(A)$
a	0.75
$\neg a$	0.25

A	N	$p(N\|A)$
a	n	0.15
a	$\neg n$	0.85
$\neg a$	n	0.03
$\neg a$	$\neg n$	0.97



A	S	$p(S\|A)$
a	s	0.10
a	$\neg s$	0.90
$\neg a$	s	0.02
$\neg a$	$\neg s$	0.98

N	S	D	$p(D\|N,S)$
n	s	d	0.96
n	s	$\neg d$	0.04
n	$\neg s$	d	0.40
n	$\neg s$	$\neg d$	0.60
$\neg n$	s	d	0.45
$\neg n$	s	$\neg d$	0.55
$\neg n$	$\neg s$	d	0.10
$\neg n$	$\neg s$	$\neg d$	0.90

S	P	$p(P\|S)$
s	p	0.75
s	$\neg p$	0.25
$\neg s$	p	0.05
$\neg s$	$\neg p$	0.95

$p(A, N, S, D, P) = p(A)p(N|A)p(S|A)p(D|N,S)p(P|S)$

DAG + CPTs

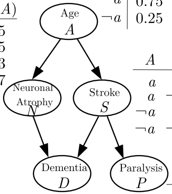
- Directed acyclic graph
- Conditional independence: **W** and **T** are conditionally independent given **Z** $\Leftrightarrow p(\mathbf{W}|\mathbf{T},\mathbf{Z}) = p(\mathbf{W}|\mathbf{Z})$
- $p(X_1,\ldots,X_n) = \prod_{i=1}^{n} p(X_i \mid \mathbf{Pa}(X_i))$

Inference

- Exact: variable elimination, message passing...
- Approximate: sequential simulation, MCMC...

A	$p(A)$
a	0.75
$\neg a$	0.25

A	N	$p(N\|A)$
a	n	0.15
a	$\neg n$	0.85
$\neg a$	n	0.03
$\neg a$	$\neg n$	0.97

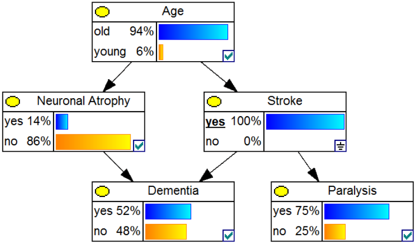


A	S	$p(S\|A)$
a	s	0.10
a	$\neg s$	0.90
$\neg a$	s	0.02
$\neg a$	$\neg s$	0.98

N	S	D	$p(D\|N,S)$
n	s	d	0.96
n	s	$\neg d$	0.04
n	$\neg s$	d	0.40
n	$\neg s$	$\neg d$	0.60
$\neg n$	s	d	0.45
$\neg n$	s	$\neg d$	0.55
$\neg n$	$\neg s$	d	0.10
$\neg n$	$\neg s$	$\neg d$	0.90

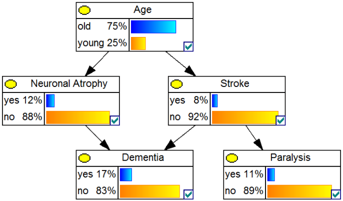
S	P	$p(P\|S)$
s	p	0.75
s	$\neg p$	0.25
$\neg s$	p	0.05
$\neg s$	$\neg p$	0.95

$p(A,N,S,D,P) = p(A)p(N|A)p(S|A)p(D|N,S)p(P|S)$

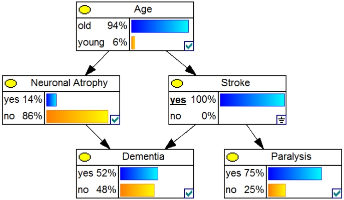


$p(X_i|\texttt{Stroke=yes})$

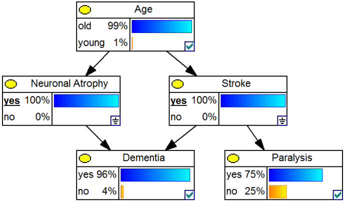
Conditional Independence



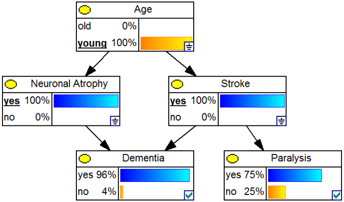
$p(X_i)$



$p(X_i|\texttt{Stroke=yes})$



$p(X_i|\texttt{Stroke=yes, Neuronal Atrophy=yes})$



$p(X_i|\texttt{Stroke=yes, Neuronal Atrophy=yes, Age=young})$

Learning Bayesian Networks from Data

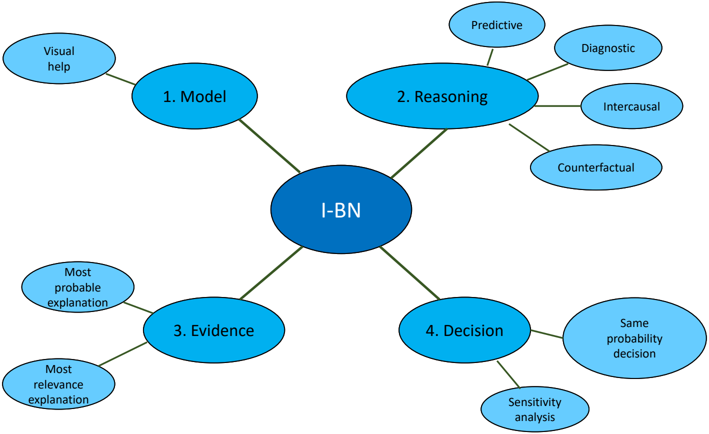
Structure learning: the DAG

- Detecting conditional independencies among triples of variables by means of statistical hypothesis testing
- Methods based on score + search

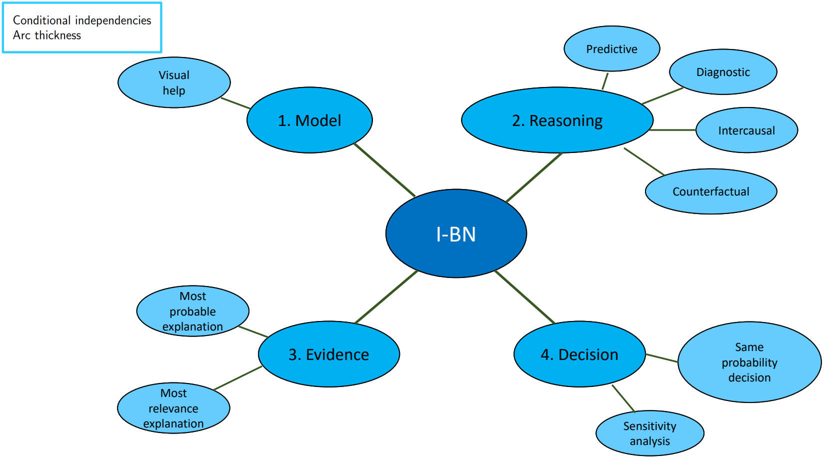
Parametric learning: $p(X_i = x_i \mid \mathbf{Pa}(X_i) = \mathbf{pa}_i^j)$

- Maximum likelihood estimation
- Bayesian estimation

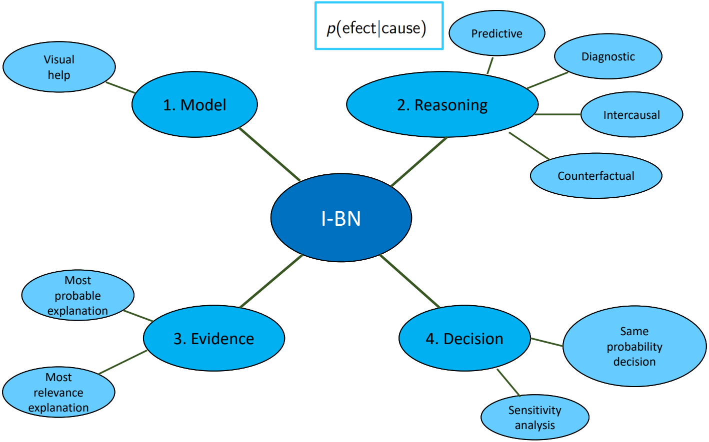
Interpretability with Bayesian Networks



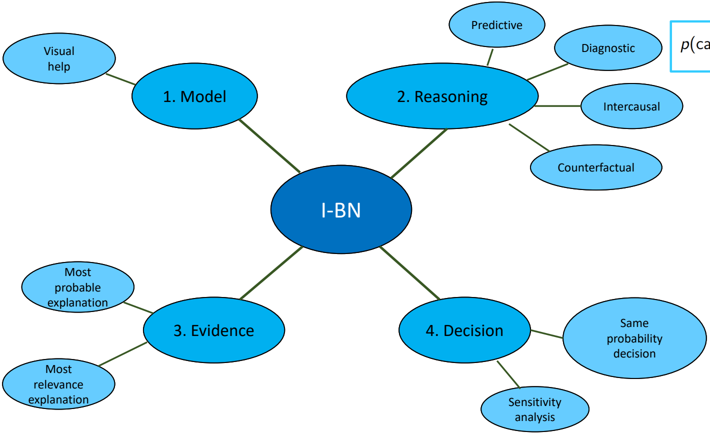
Interpretability with Bayesian Networks



Interpretability with Bayesian Networks

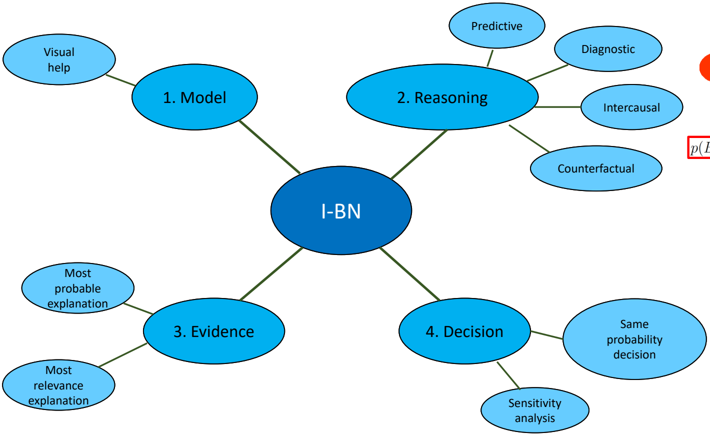


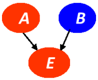
Interpretability with Bayesian Networks



$p(\text{cause}|\text{efect})$

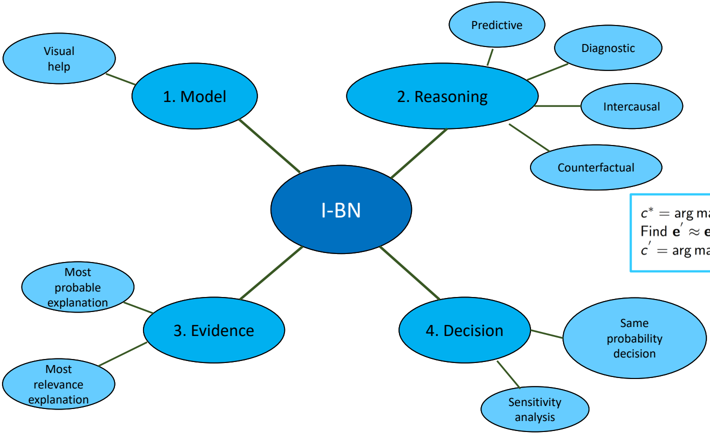
Interpretability with Bayesian Networks





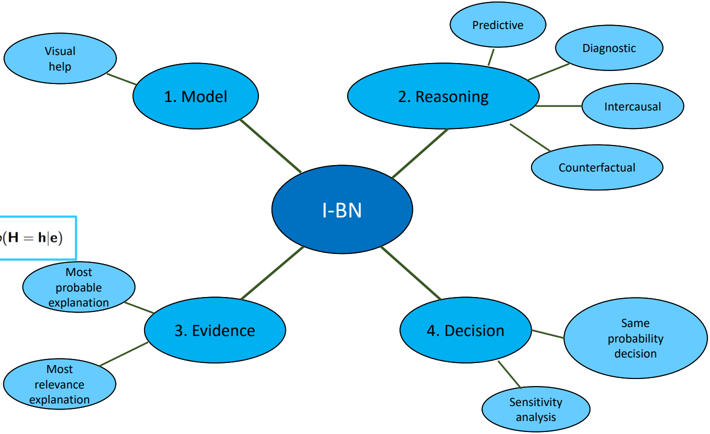
$\boxed{p(B|A=a, E=e)}$

Interpretability with Bayesian Networks



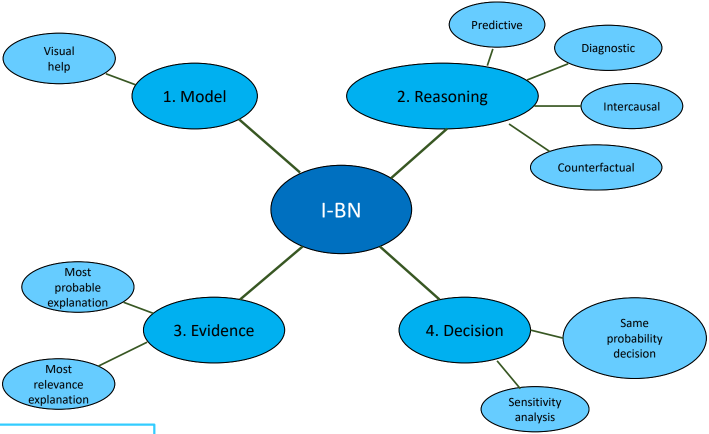
$c^* = \arg\max_c p(c|\mathbf{e})$
Find $\mathbf{e}^{'} \approx \mathbf{e}$ such that
$c^{'} = \arg\max_c p(c|\mathbf{e}^{'}) \neq c^*$

Interpretability with Bayesian Networks



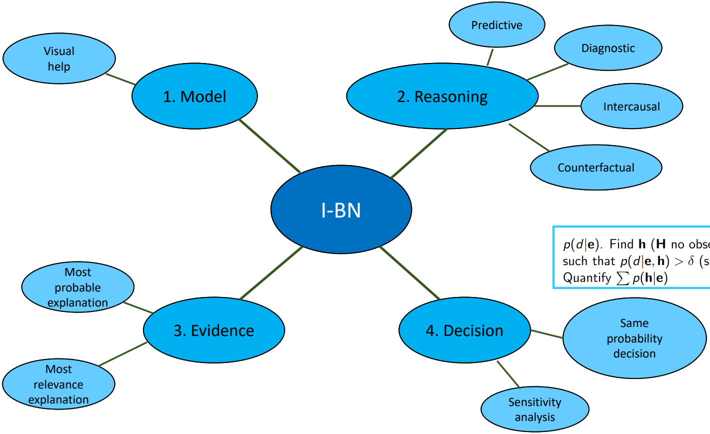
$\mathbf{h}^* = \arg\max_{\mathbf{h}} p(\mathbf{H} = \mathbf{h}|\mathbf{e})$

Interpretability with Bayesian Networks



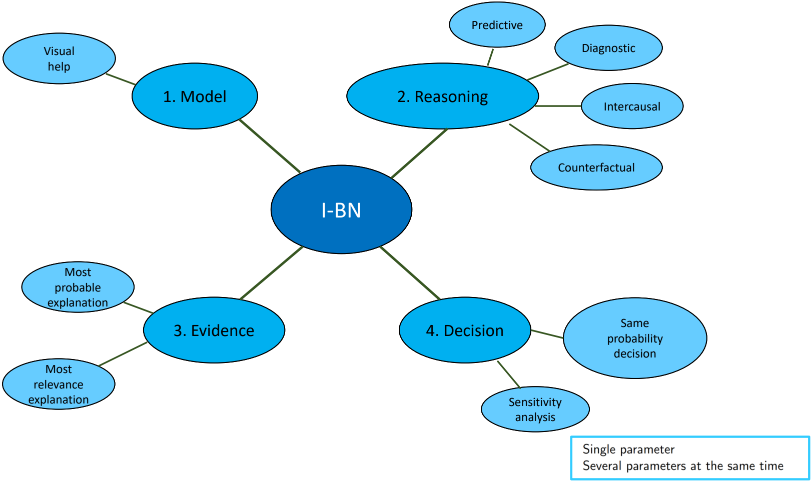
$$\boxed{\mathbf{r}^* = \arg\max_{\mathbf{r}} \frac{p(\mathbf{e}|\mathbf{r})}{p(\mathbf{e}|\neg\mathbf{r})}, \mathbf{R} \subset \mathbf{H}}$$

Interpretability with Bayesian Networks

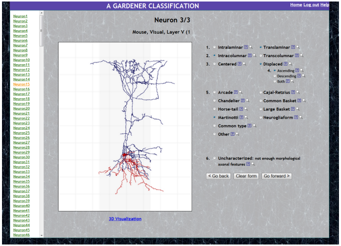


$p(d|\mathbf{e})$. Find $\mathbf{h}$ ($\mathbf{H}$ no observed)
such that $p(d|\mathbf{e}, \mathbf{h}) > \delta$ (same decision)
Quantify $\sum p(\mathbf{h}|\mathbf{e})$

Interpretability with Bayesian Networks

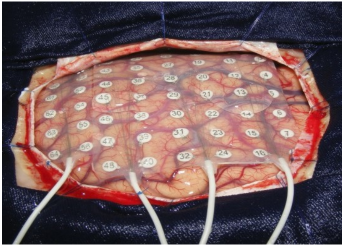


Bayesian Networks in Neuroscience



Nature Reviews Neuroscience (2014)

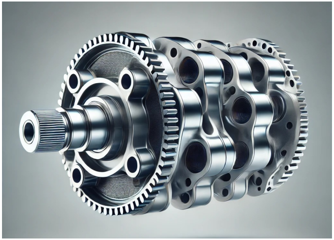
Nature Neuroscience (2020)



PLOS ONE (2014)

Scientific Reports (2021)

Bayesian Networks in Industry



Engineering Applications of Artificial Intelligence (2020)

IABMAS (2024)

Journal of Combinatorial Optimization (2022)

Submitted

Outline

Conclusions

Complex Networks

- Complex networks as mathematical paradigms to model techno-social systems
- Complex networks with undirected graphs
  - Centrality measures: node degree, betweenness centrality, closeness centrality
  - Community detection by means of clustering algorithms
  - Complex network models: Erdös-Rényi-Gilbert model, Watts-Strogatz model, Barabási-Albert model
- Complex networks with directed acyclic graphs: Bayesian networks
  - Inference: exact and approximate
  - Learning Bayesian networks from data: by testing conditional independencies or by score + search
  - Interpretability: model, reasoning, evidence, decision

Acknowledgements

Elements of Complex Networks

Pedro Larrañaga
Computational Intelligence Group
Artificial Intelligence Department
Universidad Politécnica de Madrid

Arantzazu, November 21, 2024