Artificial intelligence (AI)

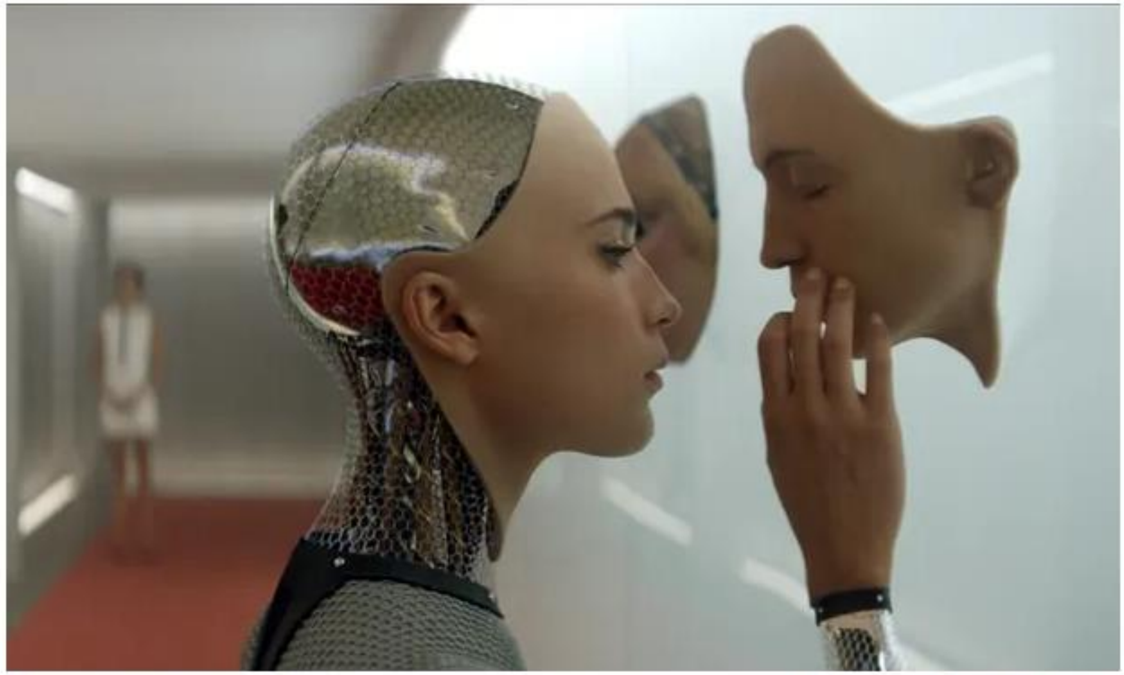
'Yeah, we're spooked': AI starting to have big real-world impact, says expert

Prof Stuart Russell says field of artificial intelligence needs to grow up quickly to ensure humans remain in control

Nicola Davis *Science correspondent*

@NicolaKSDavis

Fri 29 Oct 2021 16.00 BST



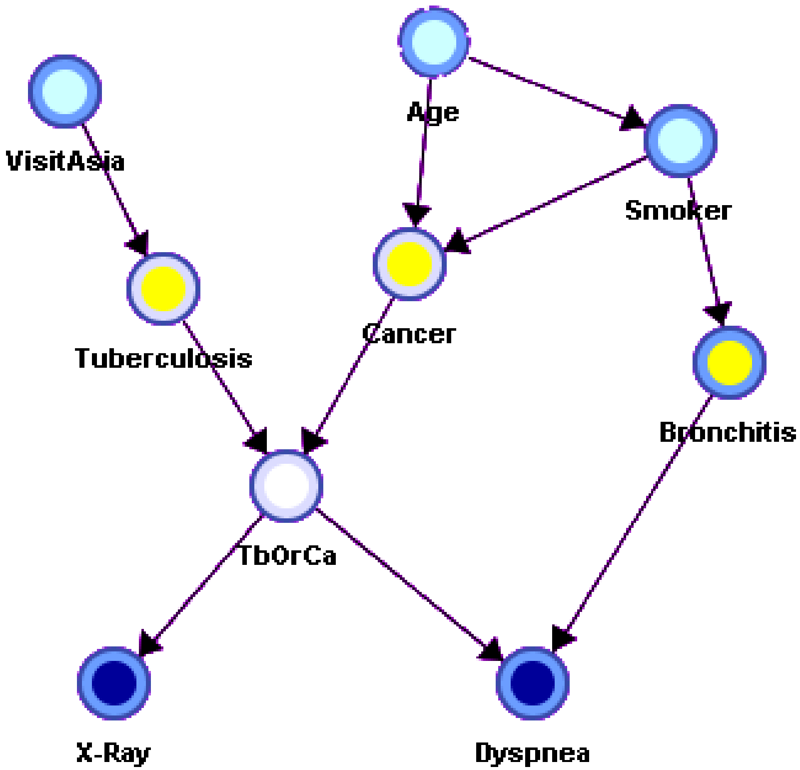
▲ There is still a big gap between the AI of today and that depicted in films such as Ex Machina, Prof Stuart Russell says. Photograph: Film4/Allstar

A scientist who wrote a leading textbook on artificial intelligence has said experts are "spooked" by their own success in the field, comparing the

https://www.theguardian.com/technology/2021/oct/29/yeah-were-spooked-ai-starting-to-have-big-real-world-impact-says-expert?CMP=Share_AndroidApp_Other

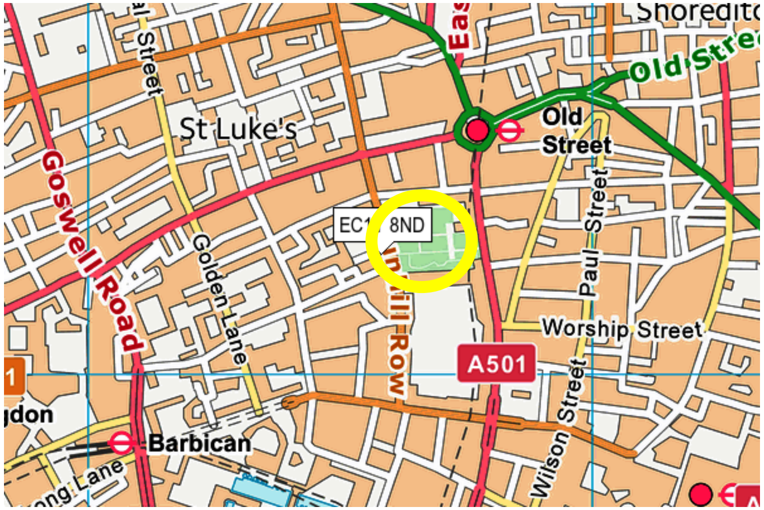
Brunel University London

Bayesian Networks – An Introduction



Brunel University London

Bayesian Networks – An Introduction

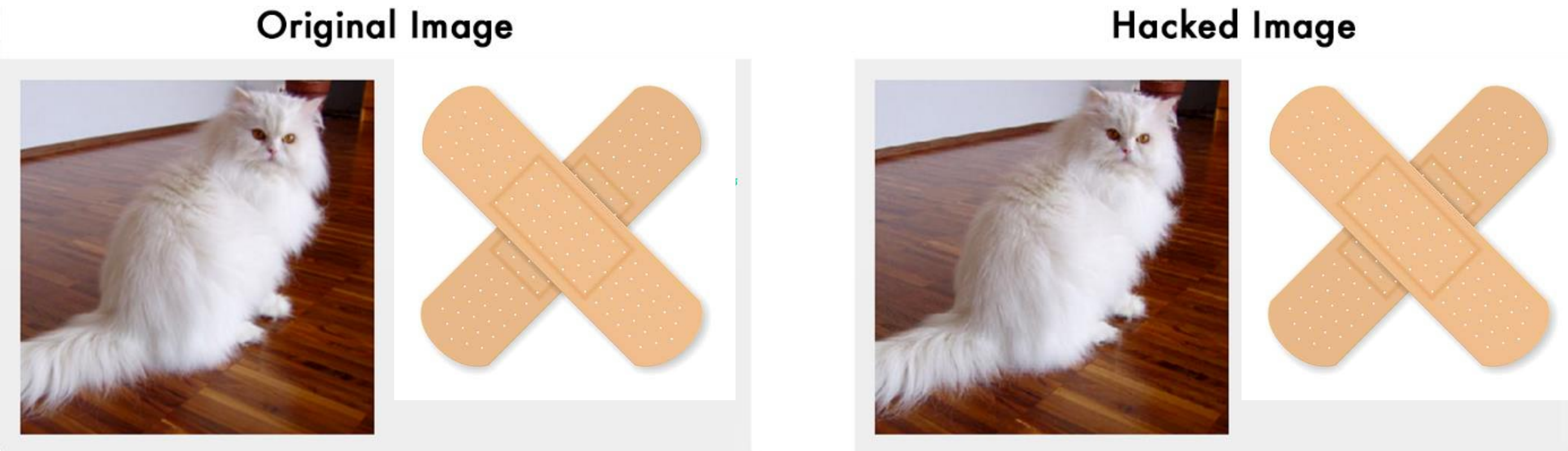


Brunel University London

Bayesian Networks

- In this lecture and lab
  - Black Box Models
  - Probability Distributions
  - Definition of Bayesian network
  - Inference and Learning Models

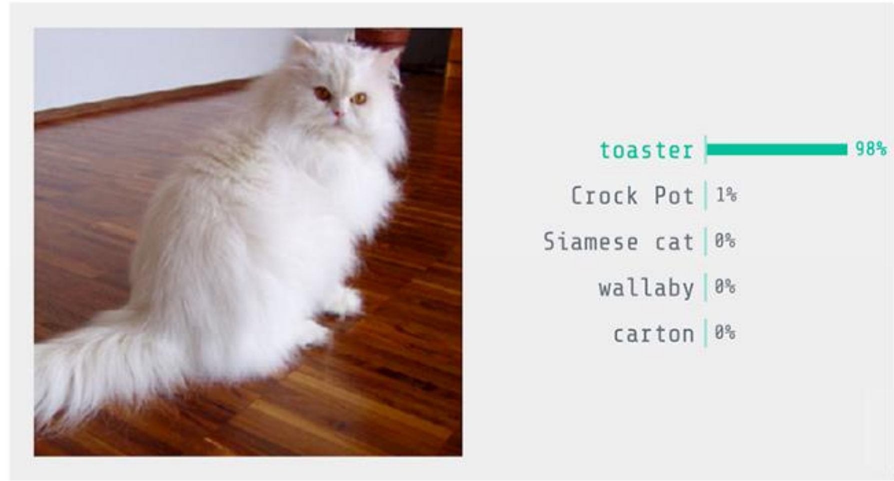
Easy to trick a Deep NN



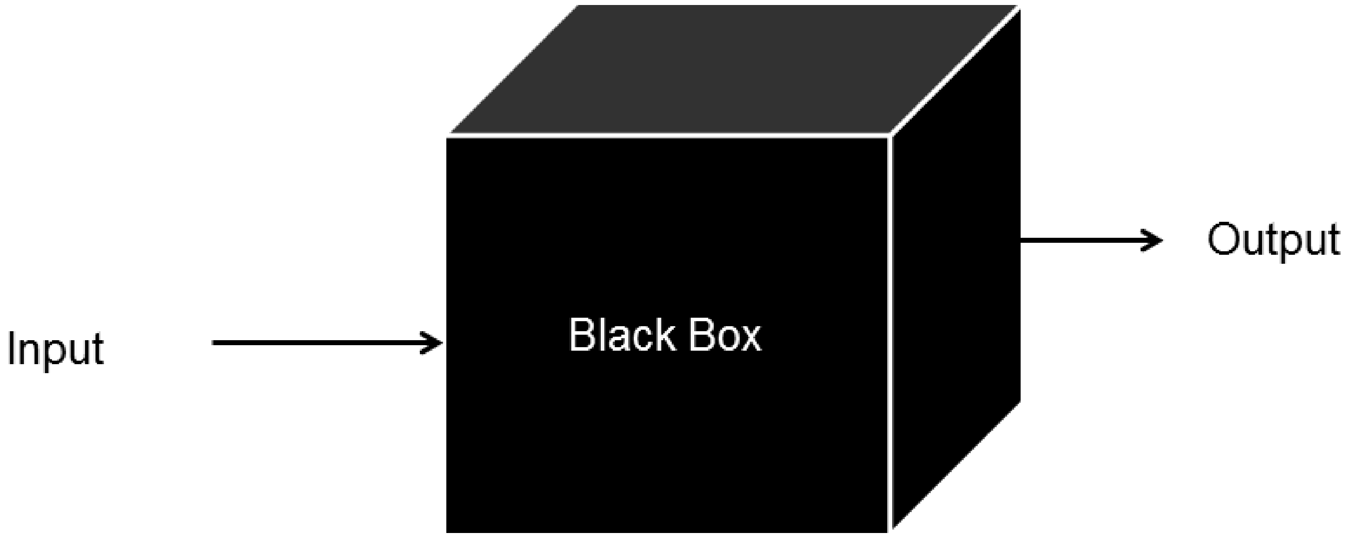
Easy to trick a Deep Learner

Original Image

Hacked Image



Opening the Black Box

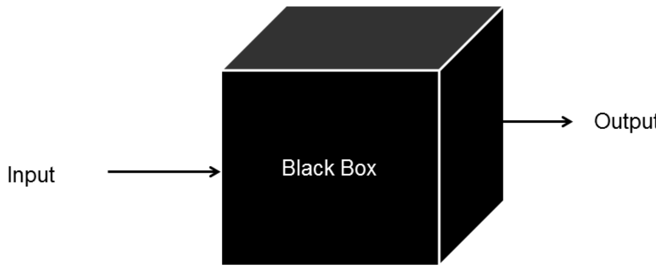


Internal behavior of the code is unknown

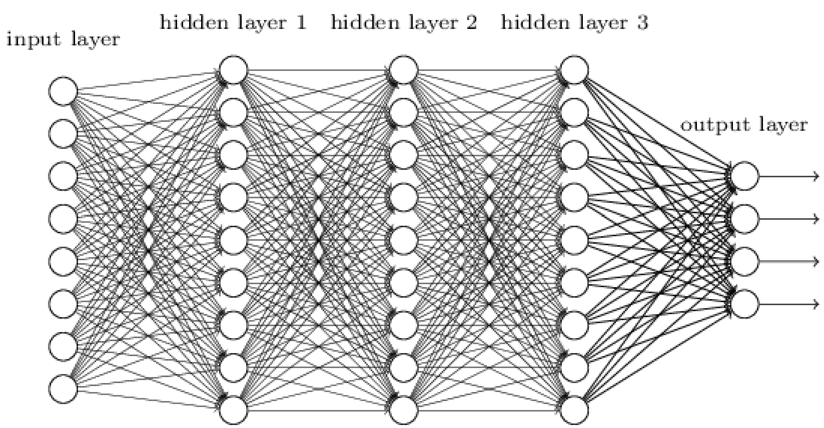
Black Box Models

Too complex for us to understand

- Massively parallel
- Huge numbers of parameters



Internal behavior of the code is unknown



Do we care?

“I don’t care if the decision cannot be explained if it is better than a human”

The Geoff Hinton “Is this a 2?” argument

Do we care?

"I don't care if the decision cannot be explained if it is better than a human"

The Geoff Hinton "Is this a 2?" argument



General Data Protection Reg. 2018

Rights related to automated decision making and profiling

In brief...

The GDPR provides safeguards for individuals against the risk that a potentially damaging decision is taken without human intervention. These rights work in a similar way to existing rights under the DPA.

Identify whether any of your processing operations constitute automated decision making and consider whether you need to update your procedures to deal with the requirements of the GDPR.

In more detail...

When does the right apply?

Individuals have the right *not to be subject to a decision* when:

- it is based on automated processing; and
- it produces a legal effect or a similarly significant effect on the individual.

You must ensure that individuals are able to:

- obtain human intervention;
- express their point of view; and
- obtain an explanation of the decision and challenge it.

Does the right apply to all automated decisions?

No. The right does not apply if the decision:

- is necessary for entering into or performance of a contract between you and the individual;
- is authorised by law (eg for the purposes of fraud or tax evasion prevention); or
- based on explicit consent. (Article 9(2)).

Furthermore, the right does not apply when a decision does not have a legal or similarly significant effect on someone.

Urgent need to open the black box

- We need to know the underlying mechanisms of the black box to
  - Gain trust of users
  - Gain new insights
  - Make better decisions / interventions

AI researchers allege that machine learning is alchemy

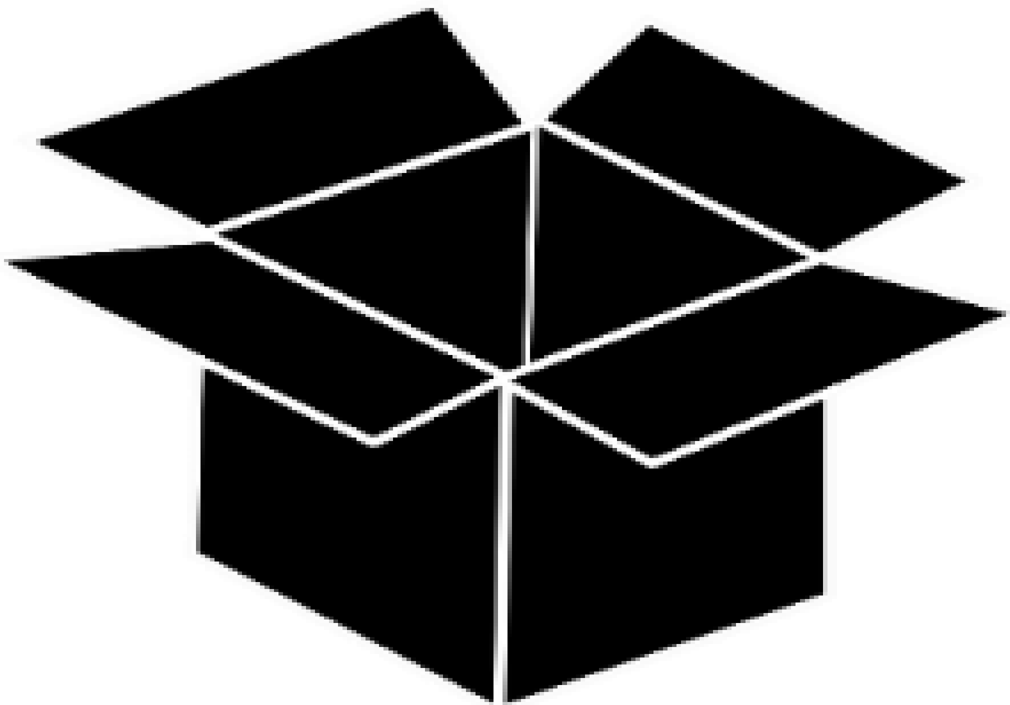
By **Matthew Hutson** | May. 3, 2018 , 11:15 AM

Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, took a swipe at his field last December—and received a 40-second ovation for it. Speaking at an AI conference, Rahimi charged that machine learning algorithms, in which computers learn through trial and error, **have become a form of "alchemy."** Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another. Now, in a paper presented on 30 April at the International Conference on Learning Representations in Vancouver, Canada, Rahimi and his collaborators **document examples** of what they see as the alchemy problem and offer prescriptions for bolstering AI's rigor.

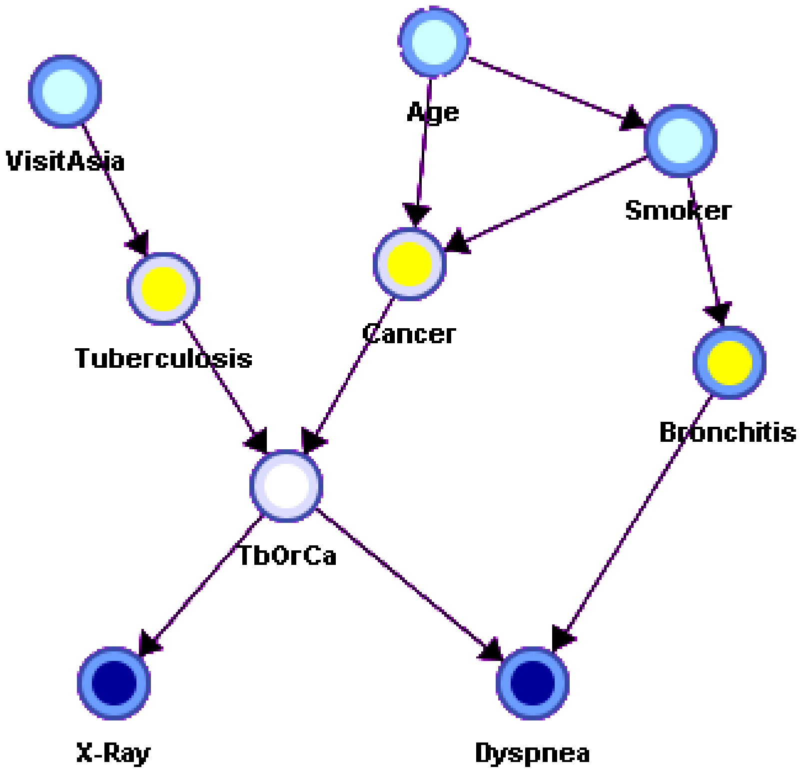
"There's an anguish in the field," Rahimi says. "Many of us feel like we're operating on an alien technology."

Brunel University London

Opening the Black Box



Bayesian Networks – An Introduction



Probability

- "Event": x (e.g. tossing coin)
- "Outcome": e.g. Heads
- Necessarily true event, x: $p(x) = 1$
- Necessarily false event, x: $p(x) = 0$

- $p(x=\text{Heads}) = 0.5$
- $p(x=\text{Winning the Lottery}) = 0.000001$
- $p(x=\text{Passing a viva}) = 0.999$

Brunel University London

Conditional Probability

- What is the probability that you have a headache p(H)?
- What is the probability that you have headache *given* that you have the flu p(H|F)?

- Flu and Headache are NOT independent

Conditional Probability

- Probability of B given A:
  - E.g. P(Headache|Flu)

$p(A \& B) = p(B) \times p(A|B)$

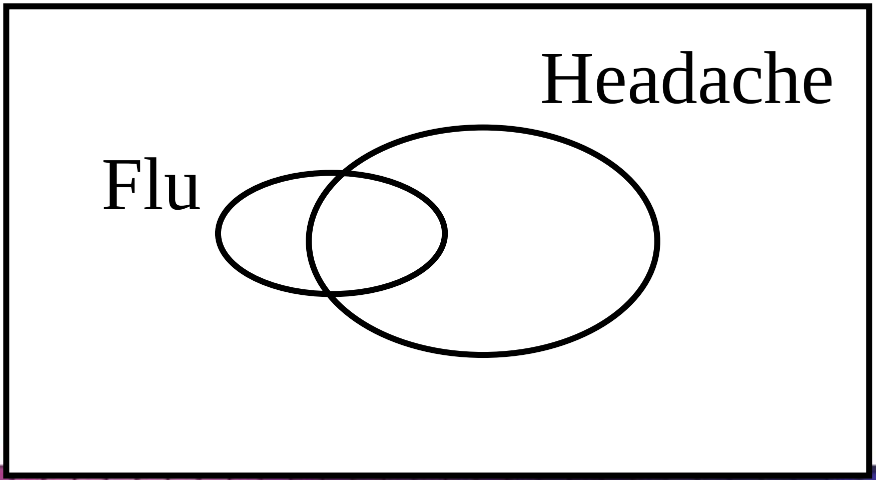
$$p(B|A) = \frac{p(A \& B)}{p(A)}$$

Conditional Probability

- Headaches are not that common e.g. p(H)=1/10
- Flu is even less common e.g. p(F)=1/40
- They are not independent
- However, if you have a headache then flu is more likely e.g. p(H|F)=1/2

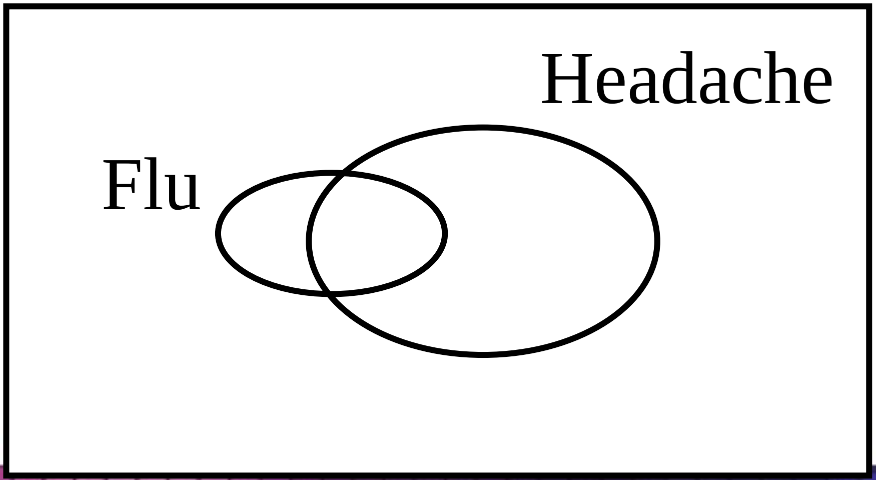
Conditional Probability

- Headaches are not that common e.g. p(H)=1/10
- Flu is even less common e.g. p(F)=1/40
- They are not independent
- However, if you have a headache then flu is more likely e.g. p(H|F)=1/2



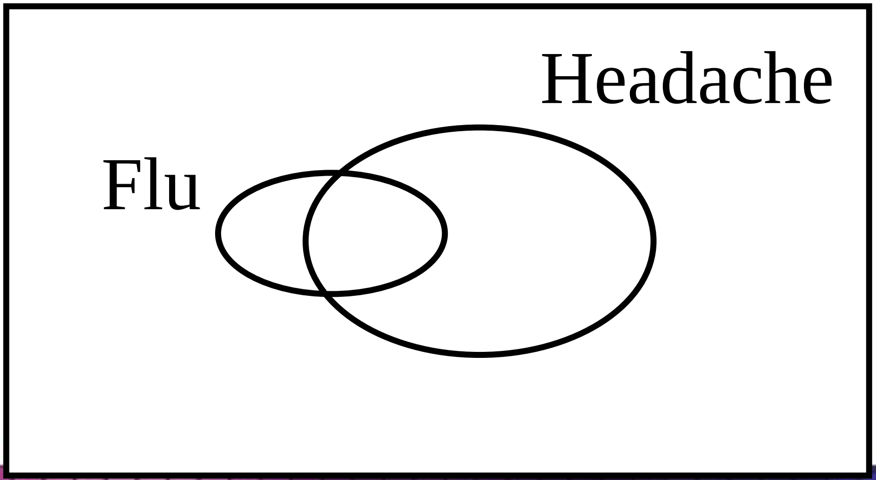
Probabilistic Inference

- Wake with a headache and say “Doh! 50% of people with flu get a headache. Therefore, I probably have the flu”
- Is this reasonable?



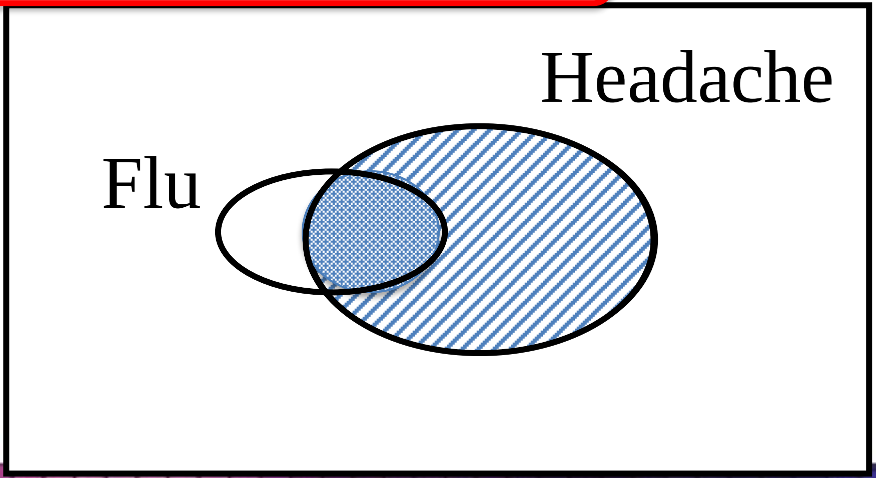
Probabilistic Inference

- Basically this is saying “probability of flu given I have a headache p(F|H) is high” but:
    - p(H) = 1/10, p(F) = 1/40, p(H|F) = 1/2
    - p(F & H) = 1/40 x 1/2 = 1/80
    - p(F|H) = (1/80) / (1/10) = 1/8



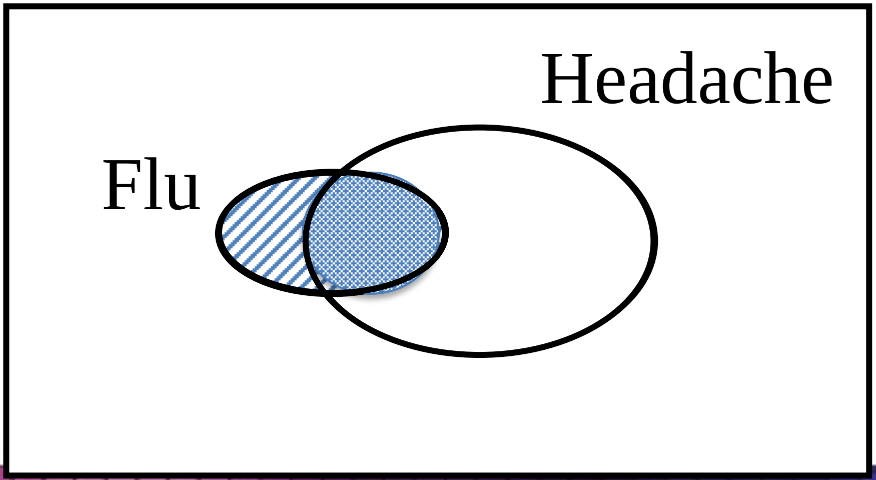
Probabilistic Inference

- Basically this is saying "probability of flu given I have a headache $p(F|H)$ is high" but:
  - $p(H) = 1/10$, $p(F) = 1/40$, $p(H|F) = 1/2$
  - $p(F \& H) = 1/40 \times 1/2 = 1/80$
  - $\boxed{p(F|H) = (1/80) / (1/10) = 1/8}$



Probabilistic Inference

- Basically this is saying "probability of flu given I have a headache p(F|H) is high" but:
  - p(H) = 1/10, p(F) = 1/40, p(H|F) = 1/2
  - p(F & H) = 1/40 x 1/2 = 1/80
  - p(F|H) = (1/80) / (1/10) = 1/8



Bayes' Theorem

- What we just did is use Bayes' rule:

$$p(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- An extremely important tool for calculating new conditional probabilities from old ones
- A major Branch of AI

Bayes' Theorem

- What we just did is use Bayes' rule:

$$\boxed{p(B|A)} = \frac{\boxed{P(A|B)}P(B)}{P(A)}$$

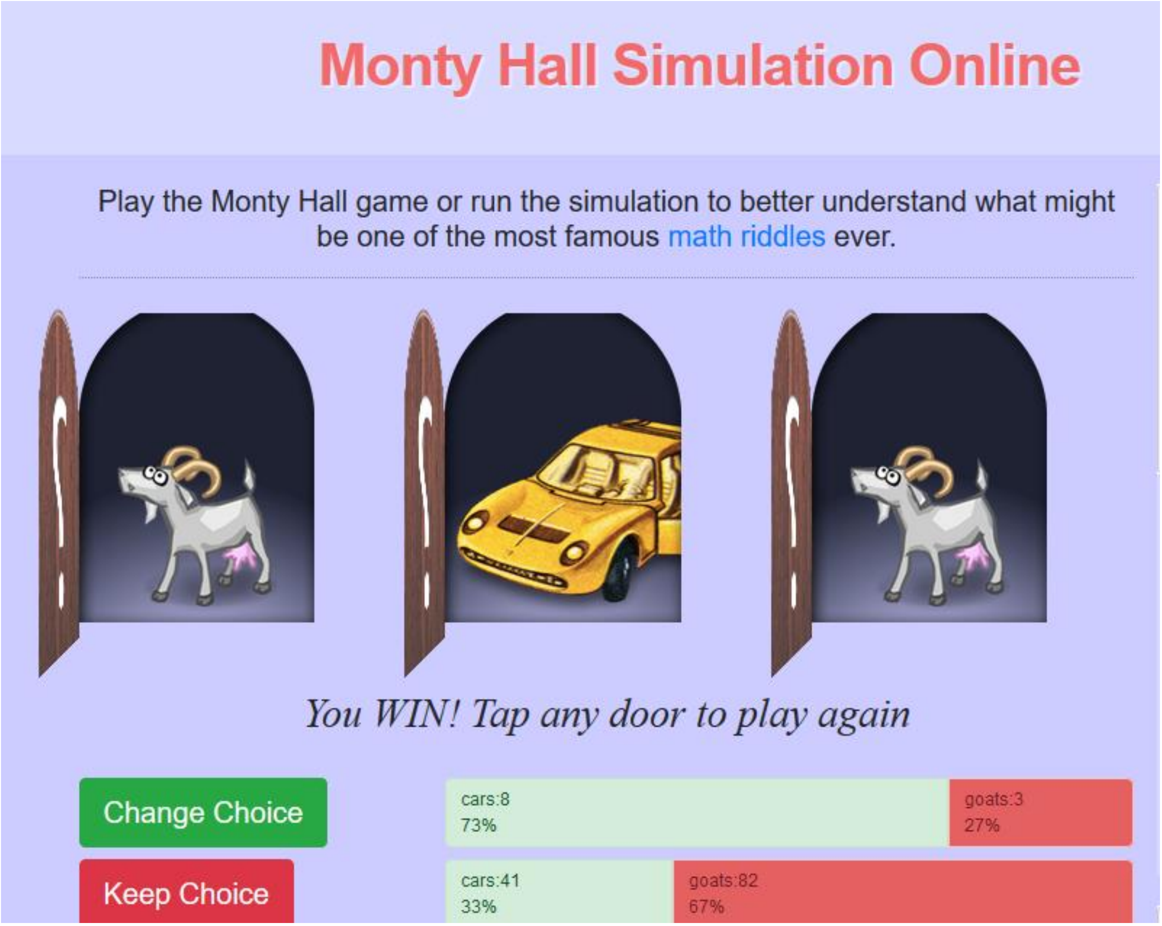
- An extremely important tool for calculating new conditional probabilities from old ones
- A major Branch of AI

Bayes’ Theorem

- Also highlights how bad we are at reasoning with small probabilities:
  - P(heads…heads….heads)
  - P(winning lottery)

Brunel University London

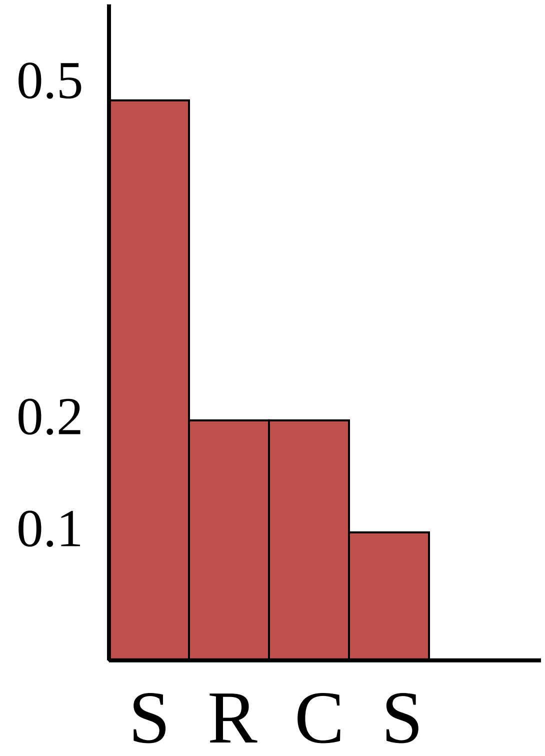
Monty Hall Problem



https://www.mathwarehouse.com/monty-hall-simulation-online/

Probability Distributions

- Probability Distribution:
  - p(Weather=Sunny) = 0.5
  - p(Weather=Rain)= 0.2
  - p(Weather=Cloud)= 0.2
  - p(Weather=Snow)= 0.1



- NB Distribution sums to 1.

Joint Probability

- Completely specifies all beliefs in a problem domain.
- Joint prob Distribution is an n-dimensional table with a probability in each cell of that state occurring.
- Written as $P(X_1, X_2, X_3 \ldots, X_n)$
- When "instantiated" as $P(x_1, x_2 \ldots, x_n)$ – where all variables are assigned a value

Joint Distribution Example

- Domain with 2 variables each of which can take on 2 states:

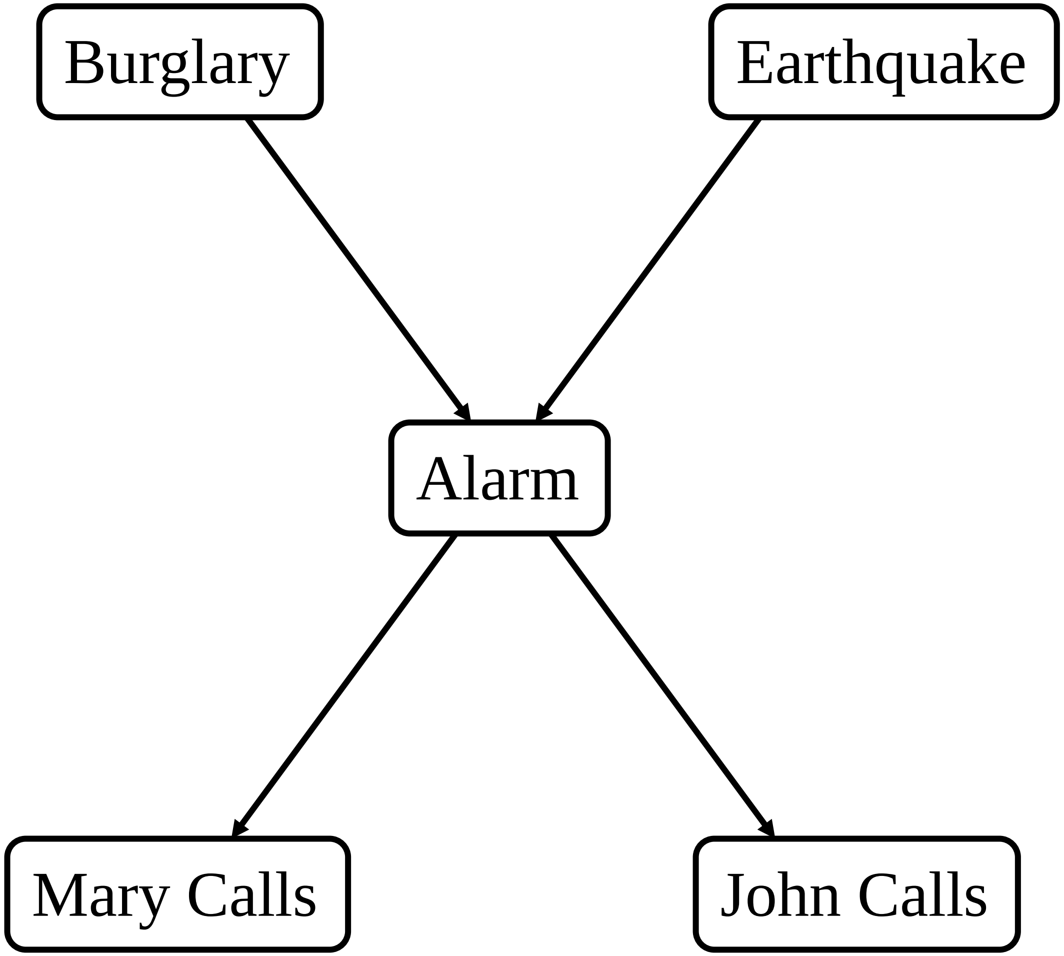
P(Toothache, Cavity)

	Toothache	¬Toothache
Cavity	0.04	0.06
¬Cavity	0.01	0.89

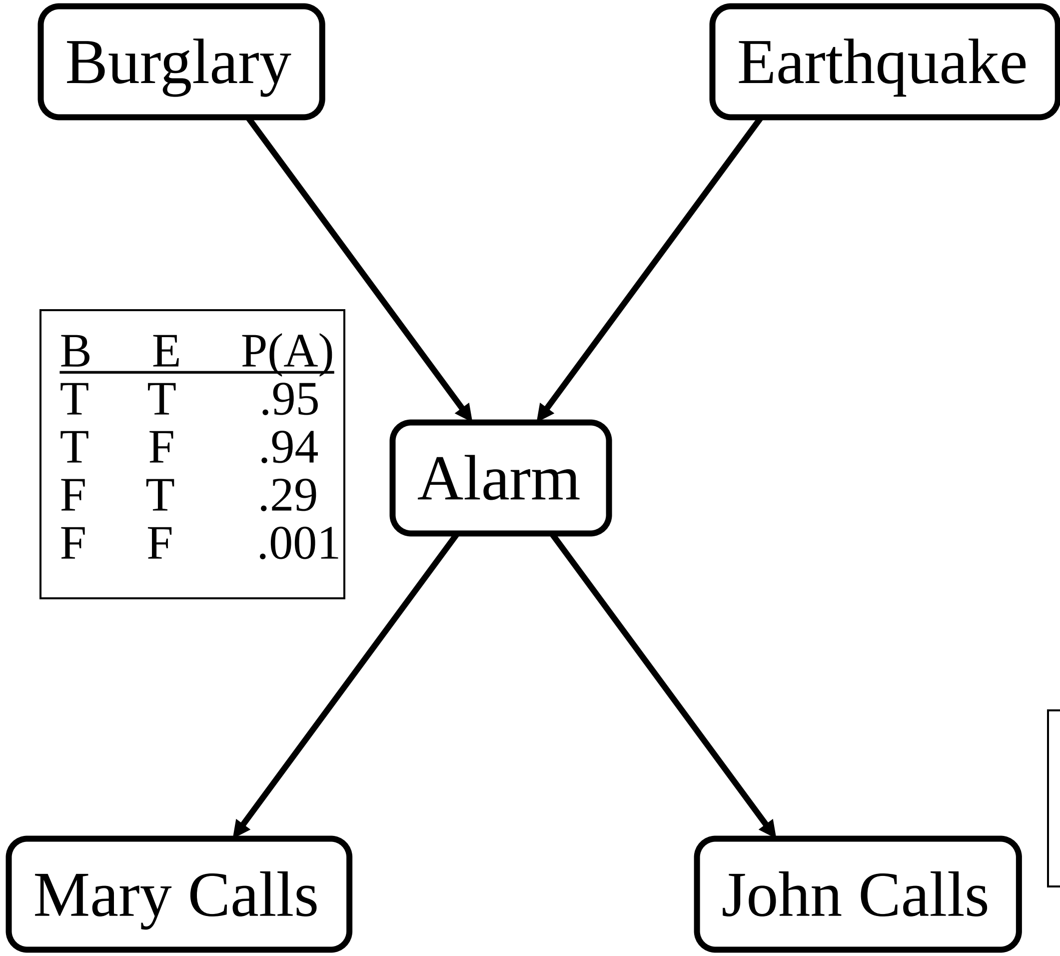
Bayesian Networks

- Joint Distribution allows us to see inside the Black Box
- Bayesian network contains:
  - Causal Structure with interconnected Nodes (Directed Acyclic Links)
  - Conditional distributions at each node
- Joint Distribution formed conditional distributions with independence assumptions based on graph

Bayesian Network Example



Bayesian Network Example



P(B)
.001

P(E)
.002

B	E	P(A)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(M)
T	.70
F	.01

A	P(J)
T	.90
F	.05

Retrieving Probabilities from the Conditional Distributions

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(x_i))$$

e.g. Probability of John (J) and Mary Calling (M), Alarm sounding (A), no Burglary (¬B) and no Earthquake (¬E):

P(J & M & A & ¬B & ¬E)

= P(J|A)P(M|A)P(A|¬B,¬E)P(¬B)P(¬E)

= 0.9 x 0.7 x 0.001 x 0.999 x 0.998

= 0.00062

Inference using WEKA – Alarm

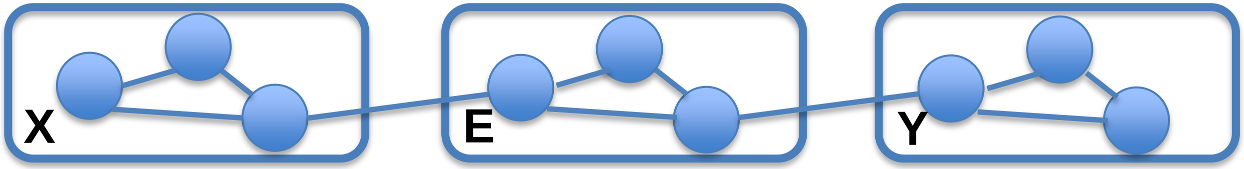
Conditional Independence - D-Separation

- To do inference in a Belief Network we have to know if two sets of variables are conditionally independent given a set of evidence.

- Method to do this is called Direction-Dependent Separation or D-Separation.

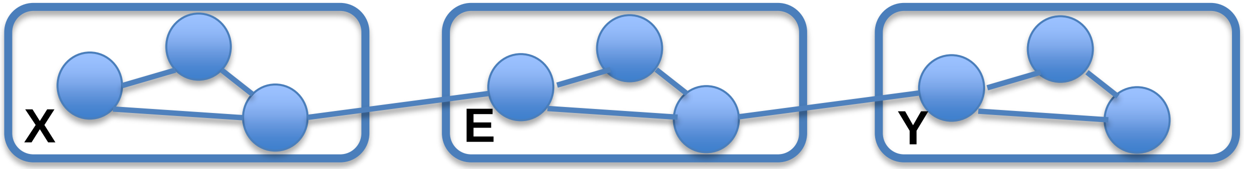
Conditional Independence - D-Separation

- If every undirected path from a node in X to a node in Y is d-separated by E, then X and Y are conditionally independent given E.
  - X is a set of variables with unknown values
  - Y is a set of variables with unknown values
  - E is a set of variables with known values.



Conditional Independence - D-Separation

- A set of nodes, E, d-separates two sets of nodes, X and Y, if every undirected path from a node in X to a node in Y is *Blocked* given E.



- A path is blocked given a set of nodes, E if:

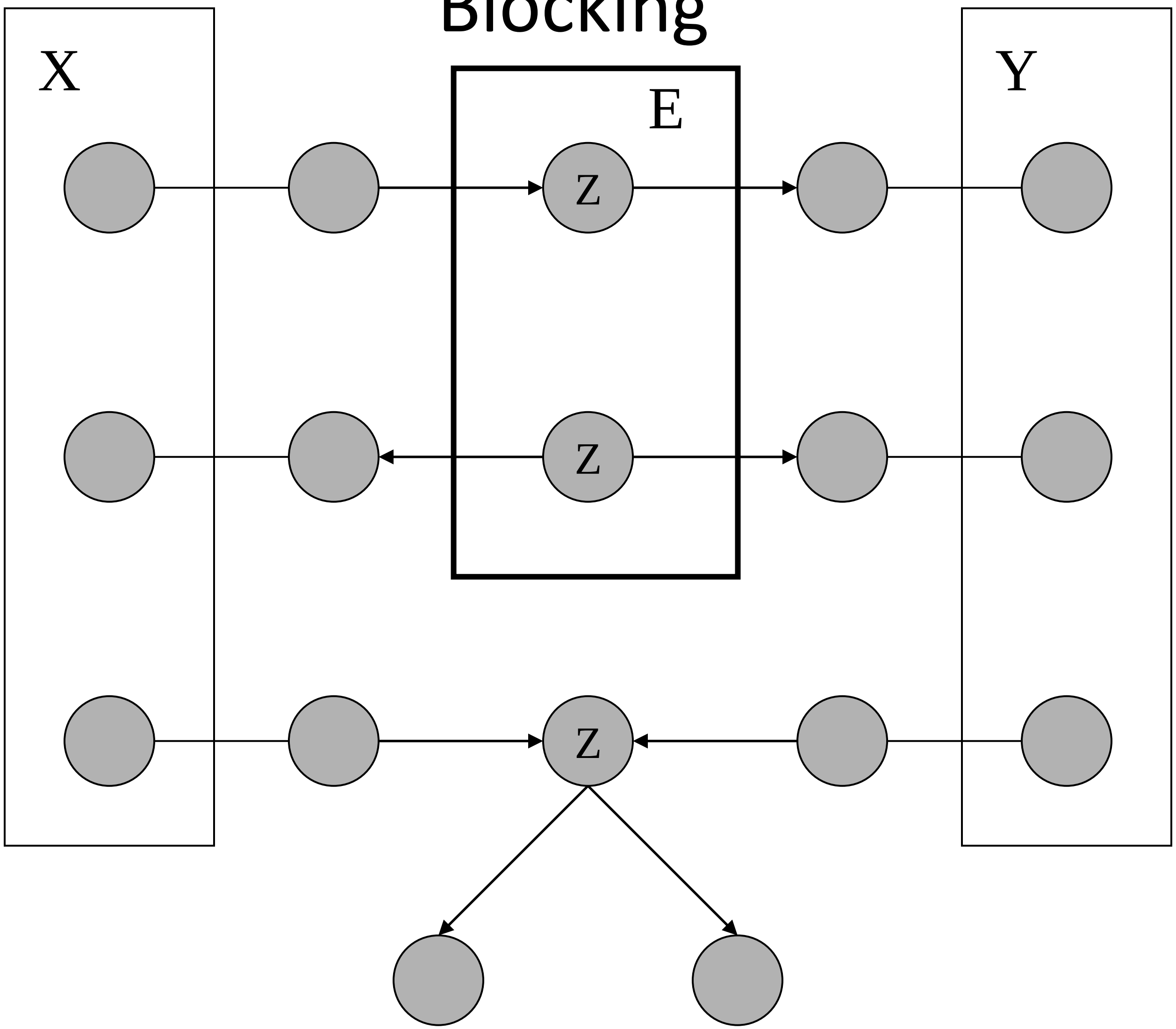
  1) Z is in E and Z has one arrow leading in and one leading out.

  2) Z is in E and has both arrows leading out.

  3) Neither Z nor any descendant of Z is in E and both path arrows lead in to Z.

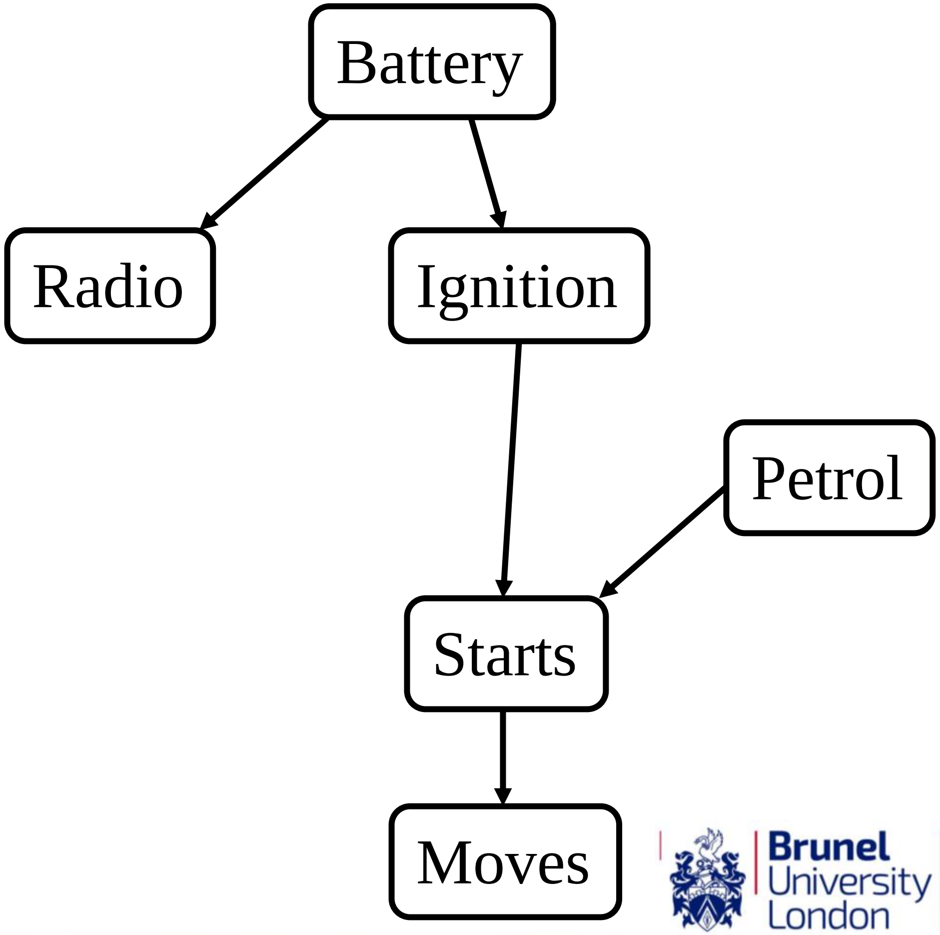
Brunel University London

Blocking



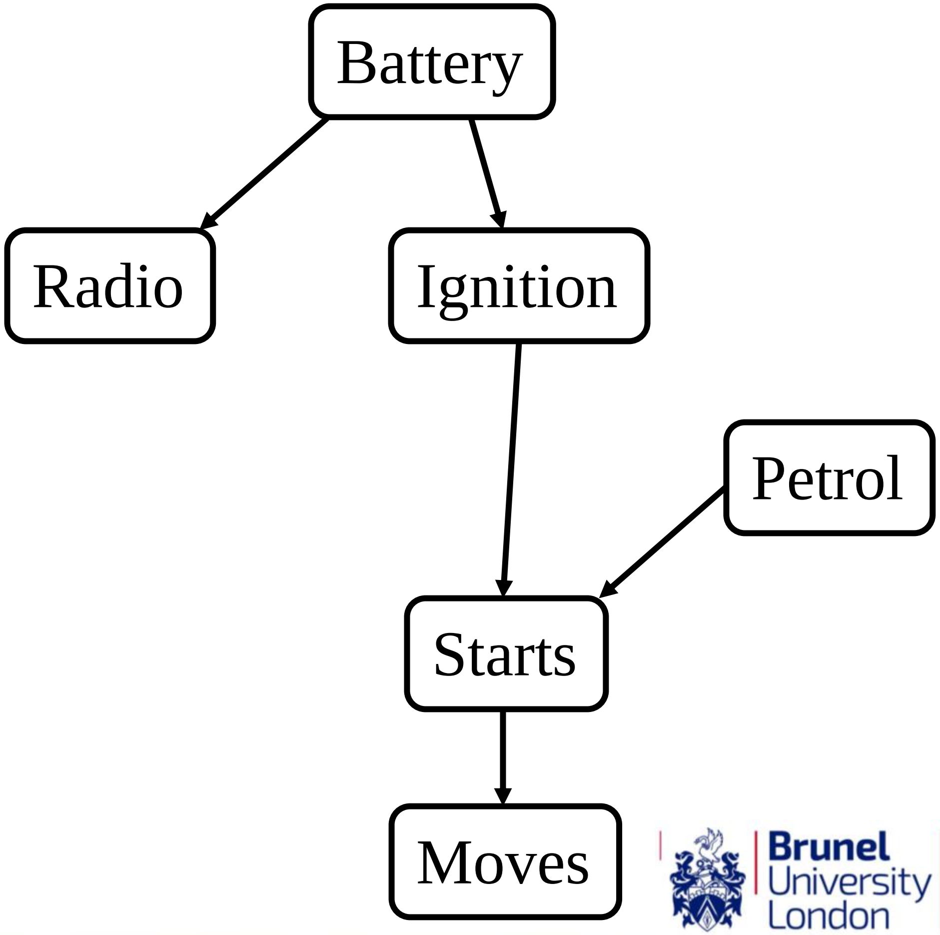
D-Separation - Example

- Moves and Battery are independent given it is known about Ignition
- Moves and Radio are independent if it is known that Battery works
- Petrol and Radio are independent given no evidence. But are dependent given evidence of Starts



Markov Blanket

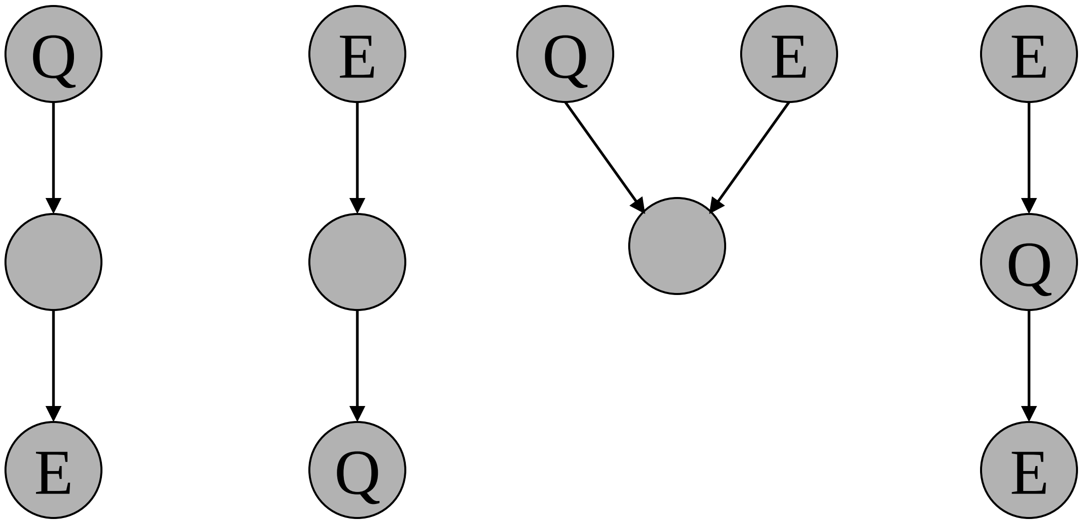
- All parents of x_i, children of x_i and parent of the children of x_i
- Markov Blanket of x_i renders all other nodes independent of x_i



Inference

- Diagnostic Inferences (effects to causes)
- Causal Inferences (causes to effects)
- Intercausal Inferences - or ‘Explaining Away’ (between causes of common effect)
- Mixed Inferences (combination of two or more of the above)

Inference



Diagnostic Causal Intercausal Mixed

Example using GENIE – Lung Cancer - Asia

Example using GENIE – Monty Hall

Learning from data

- Like Feature Selection it is NP-Hard
- Typically involves:
  - A search method
    - Hill climb
    - Greedy search
    - Genetic Algorithm
  - A scoring metric
    - Maximum likelihood
    - Minimum Description Length

Learning from data – K2

- Quick and easy but suffers from local optima:

```
Start with a network with no links
rep = true
Do
   Insert the link that increases the likelihood most
   If no link improves the likelihood Then rep = false
While rep = true
```

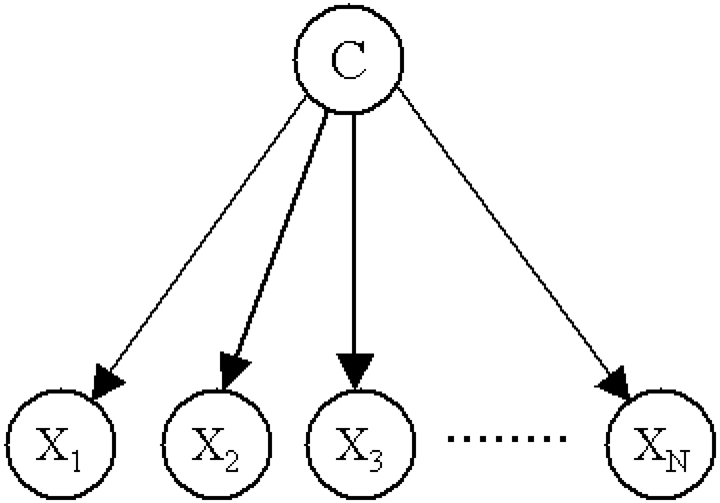
- Can also use optimisation / constraint based methods

Bayesian Classifiers

- We can use Bayesian networks to classify data
- Use nodes to represent variables as well as a class node

Bayesian Classifiers

- Naïve Bayes:
  - Assumes independence between variables given the class: p(X|C)
  - Can calculate p(C|X) using Bayes Rule or inference



Example using GENIE – NBC

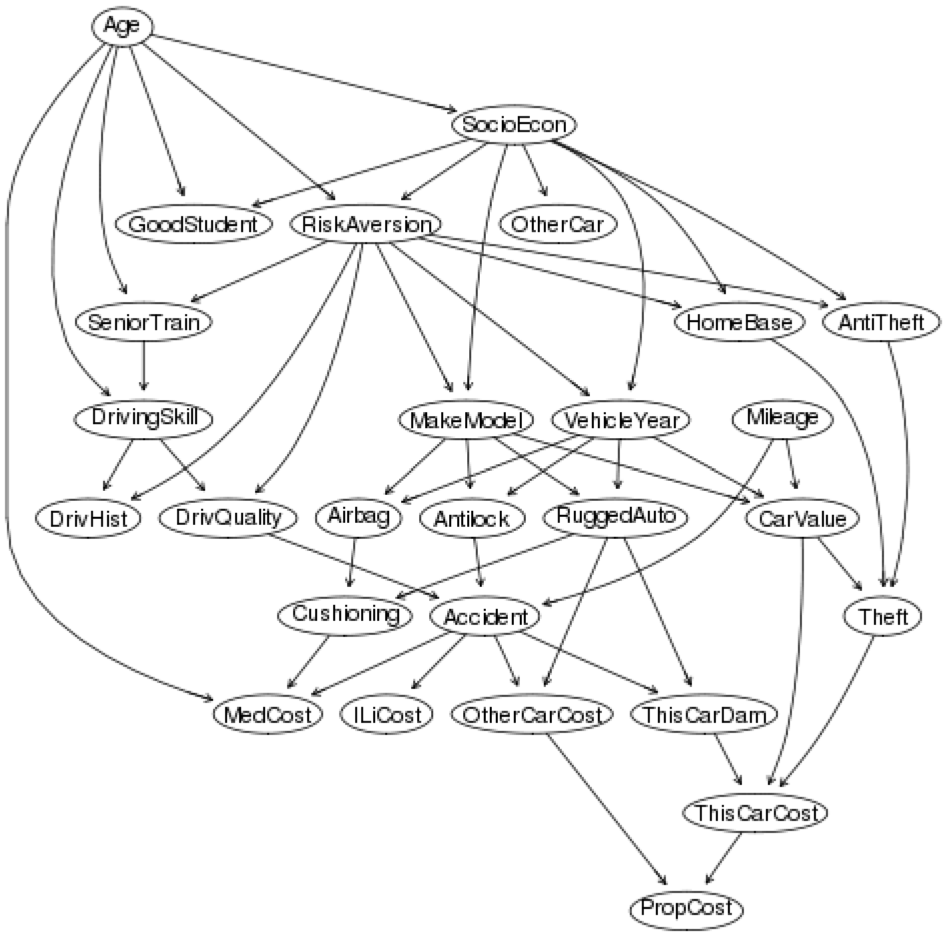
Applications

Many can be found in the BN repository:

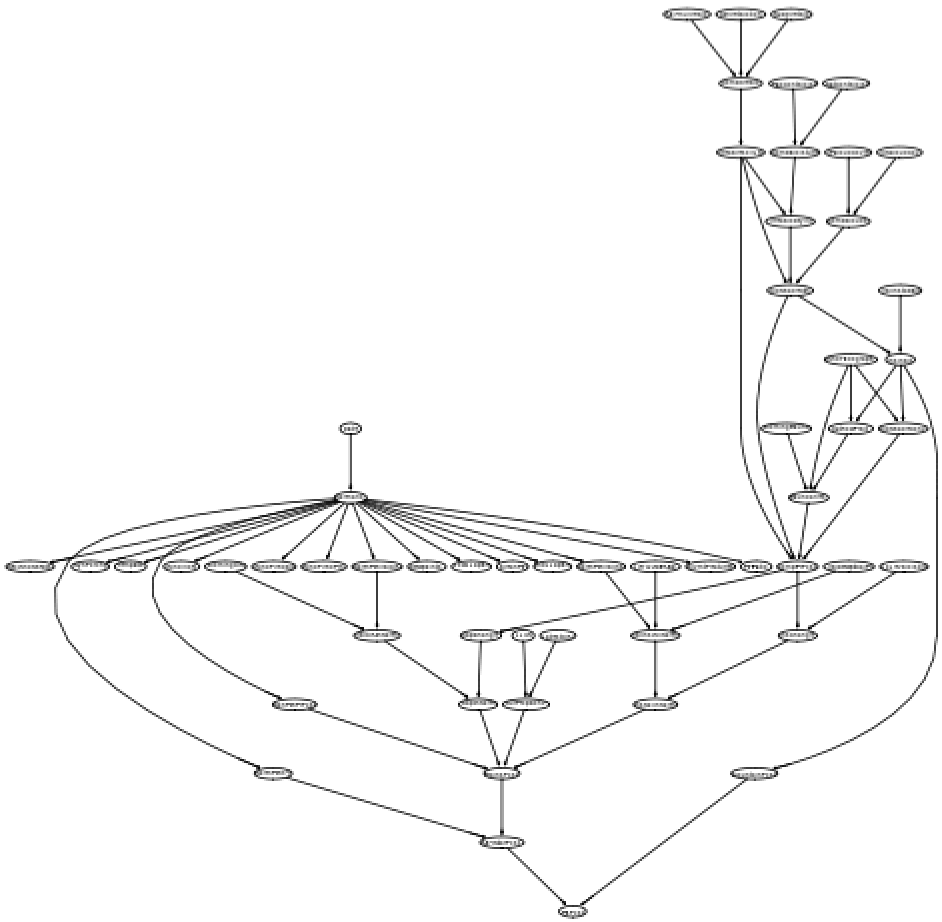
http://www.bnlearn.com/bnrepository/

Many spam filters
Waste Water Treatment Process
Pathfinder: Medical Expert System
Protein Unfolding Model
Hailfinder: Severe Weather Forecasting

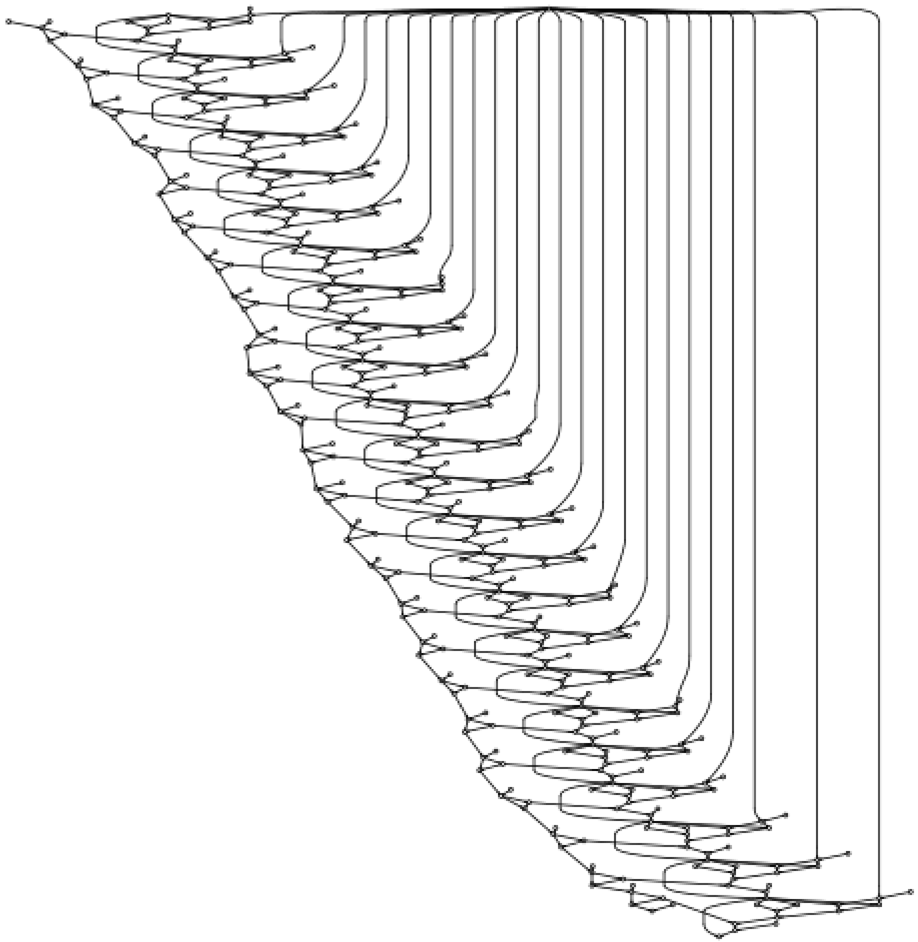
Applications: Car Insurance



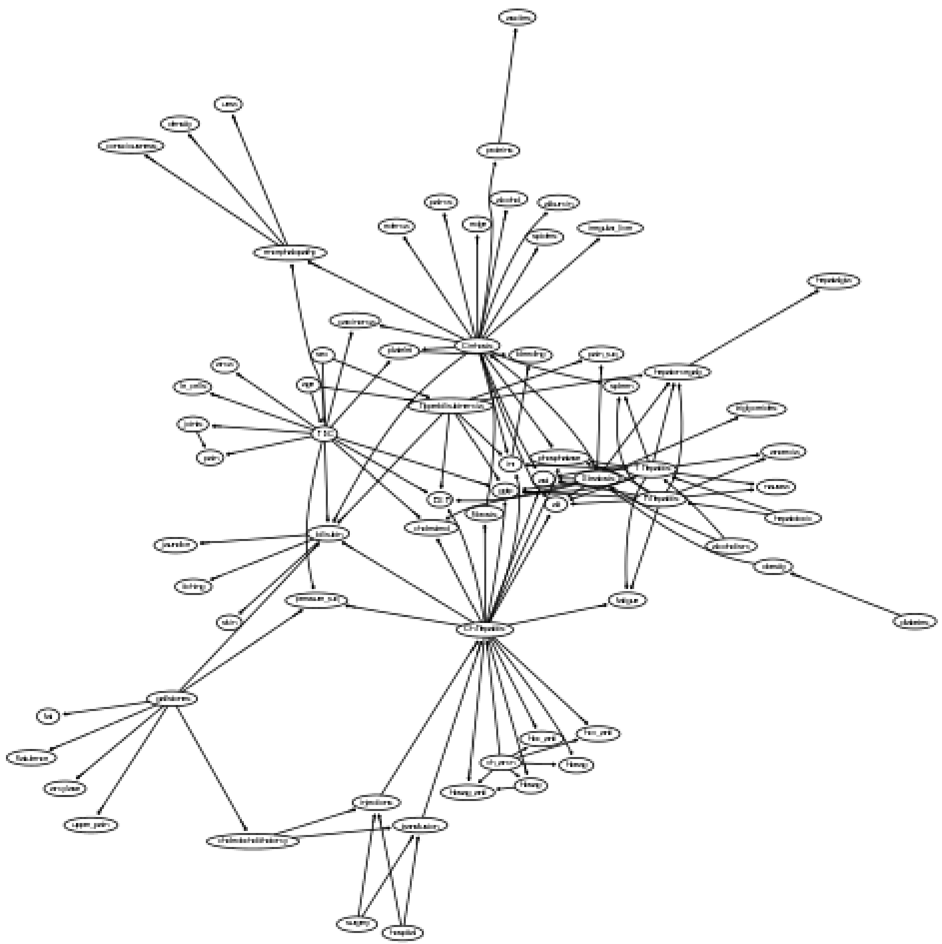
Applications: Hailfinder



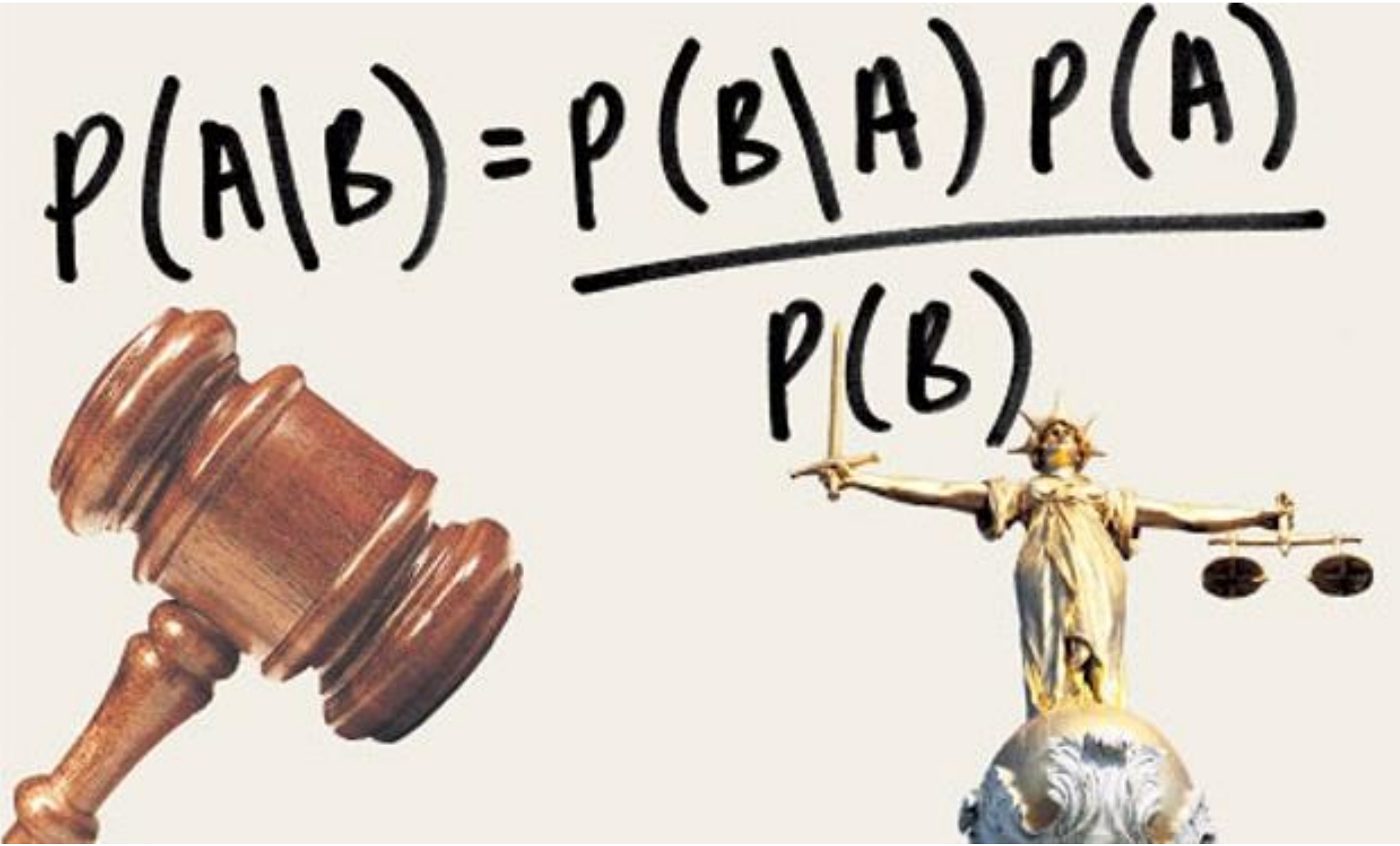
Applications: Diabetes Control



Applications: Liver Disorder



Brunel University London



Brunel University London

Summary

- Definition of a BN
- How to retrieve the joint distribution
- D-separation to model conditional independence
- Markov Blankets & Inference
- Learning BNs from data
- BNs for clustering & classification

Lab after the break

- Build some simple Bayesian networks
- Use them to perform simple reasoning tasks
- Build a Naive Bayes Classifier

Reading

- Russell & Norvig:
  - Chapter 14 (Sections 1 & 2), Chapter 20
- Korb and Nicholson
  - Chapter 2 (pdf on blackboard)
- BNlearn in R:
  - https://www.bnlearn.com/
- GENIE software:
  - https://www.bayesfusion.com/genie/

Brunel University London

Next Week

- Dynamic Bayesian Networks
- Hidden Markov Models