# Artificial Intelligence:

Summary of the Module

Summary of Exam

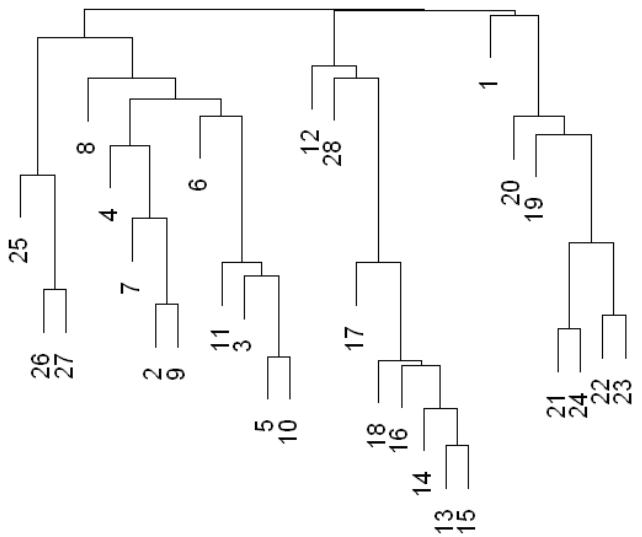
# The Topics

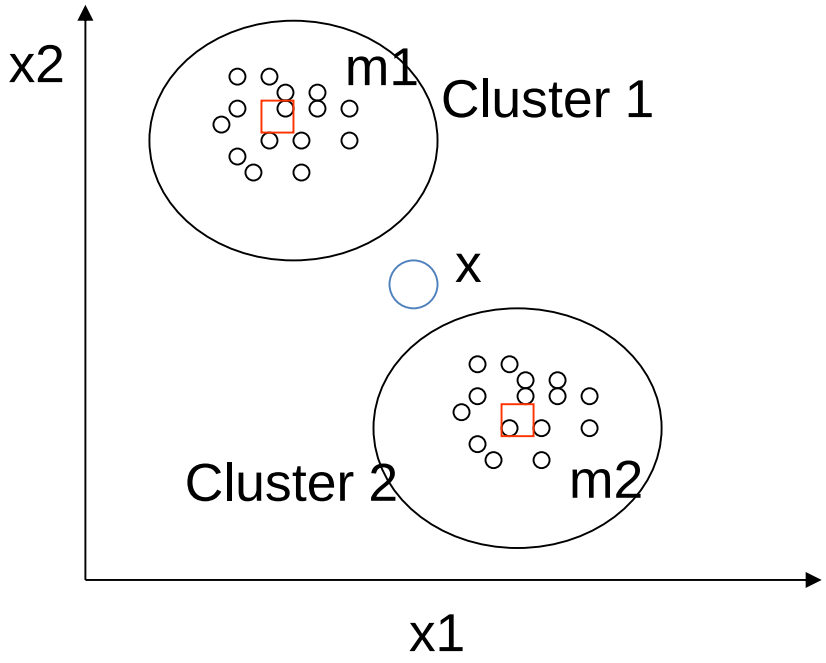
- Clustering & Association Rules
  - Distance metrics
  - 2 key clustering algorithms in detail
- Classification
  - 2 key classification algorithms in detail
  - Sensitivity Analysis – TPs vs FPs
- Neural Networks
  - Forward propagation in Perceptron and Multilayer NNs
  - General form of Backpropagation

# The Topics

- Clustering & Association Rules
  - Distance metrics
  - 2 key clustering algorithms in detail

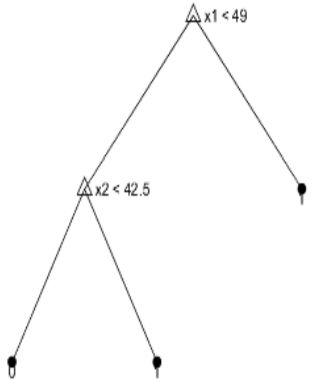
| | | | | | |
|---|---|---|---|---|---|
| x1 | 5.5 | 2.9 | 4.8 | 6.7 | 0.6 |
| x2 | 0.2 | 1.0 | 4.8 | 3.8 | 9.2 |

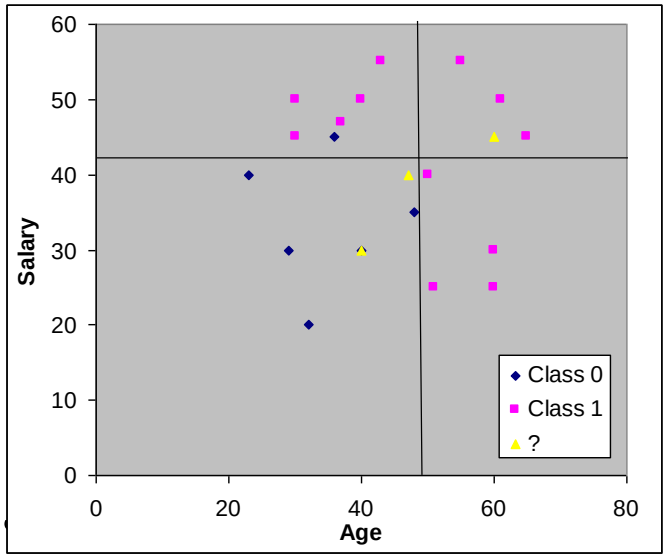


# The Topics

- Clustering & Association Rules



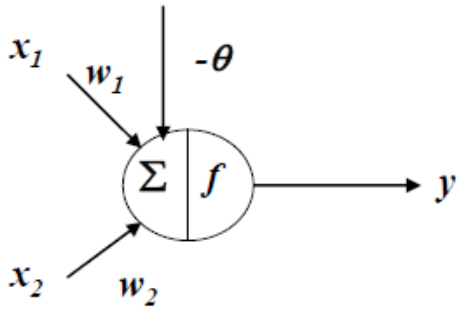

- Classification
  - 2 key classification algorithms in detail
  - Sensitivity Analysis – TPs vs FPs
  - Resampling & Overfitting

| | Class Pos. (C+) | Class Neg. (C-) |
|---|---|---|
| Predict Pos. (P+) | True Positive | False Positive |
| Predict Neg. (P-) | False Negative | True Negative |

# The Topics

Think about running a bath…

- $Y_d$: desired temperature
- Y: current temperature
- w1: opening of hot water
- w2: opening of cold water




Adjust strategy:

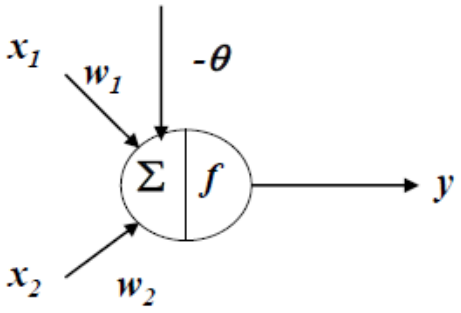
- If $Y_d < Y$, decrease w1 and increase w2
- If $Y_d > Y$, increase w1 and decrease w2

- Neural Networks
  - Forward propagation in Perceptron and Multilayer NNs
  - General form of Backpropagation

# The Topics

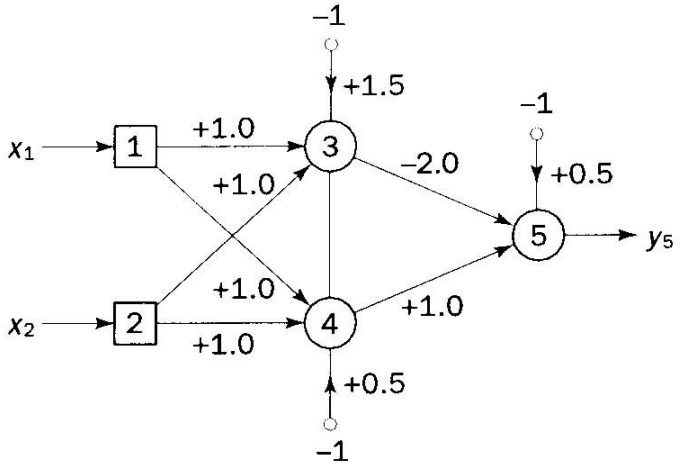
Think about running a bath…

- $Y_d$: desired temperature
- Y: current temperature
- w1: opening of hot water
- w2: opening of cold water



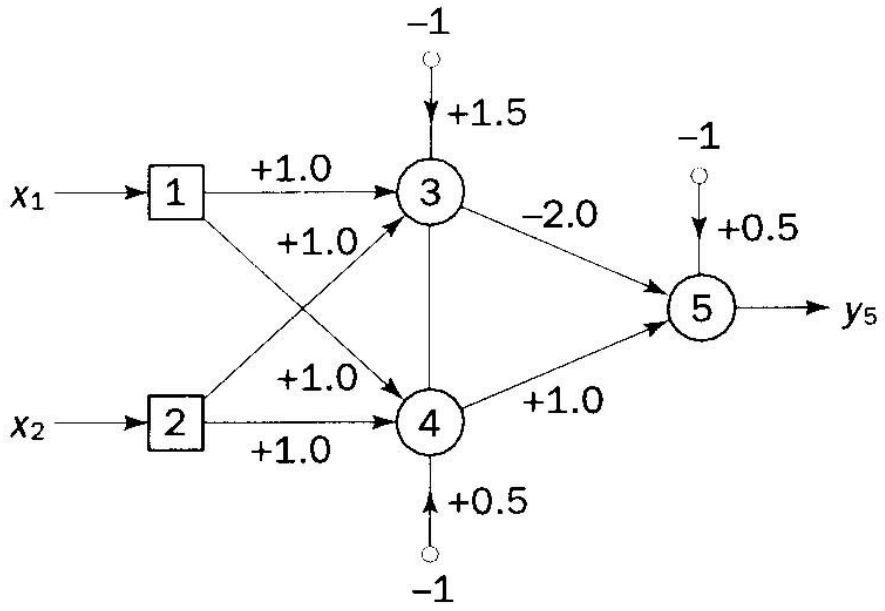

Adjust strategy:

- If $Y_d < Y$, decrease w1 and increase w2
- If $Y_d > Y$, increase w1 and decrease w2



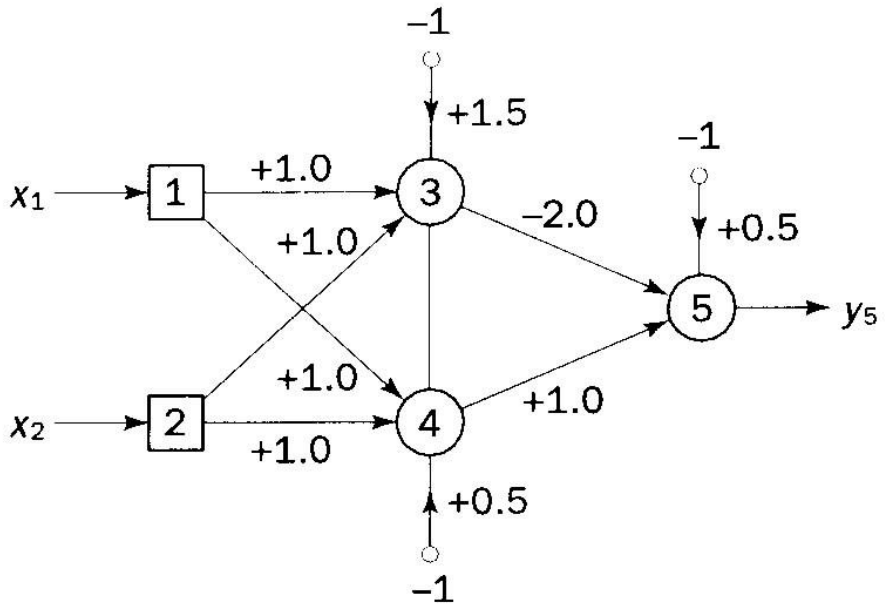

- Neural Networks
  - Forward propagation in Perceptron and Multilayer NNs
  - General form of Backpropagation
  - Deep Learning – general concept

# The Topics



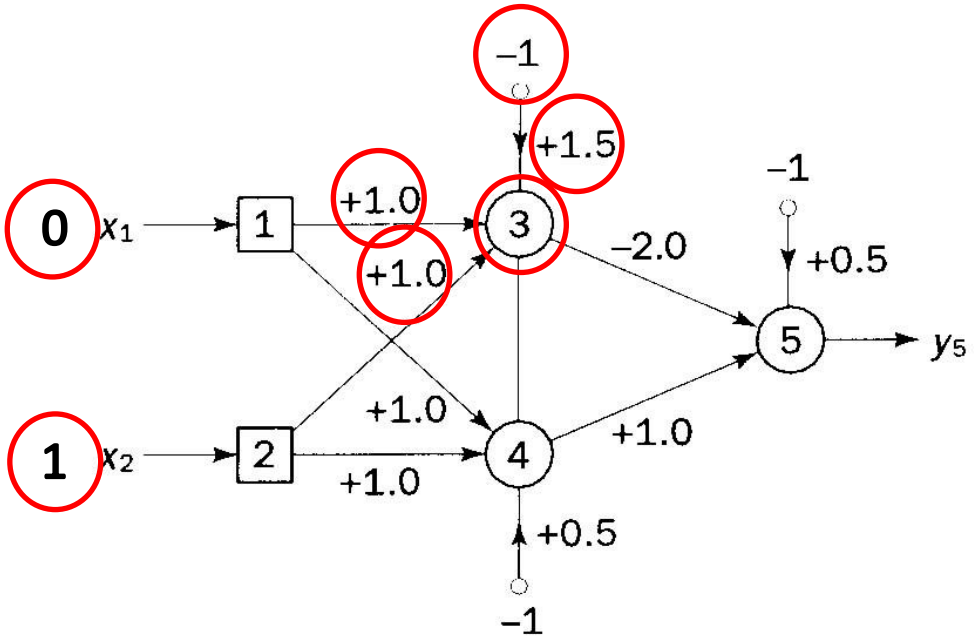

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 0 | 0 | -1.5 | -0.5 | 2 |
| 0 | 1 | -0.5 | 0.5 | 1 |
| 1 | 0 | -0.5 | 0.5 | 1 |
| 1 | 1 | 0.5 | 1.5 | 0 |

# The Topics



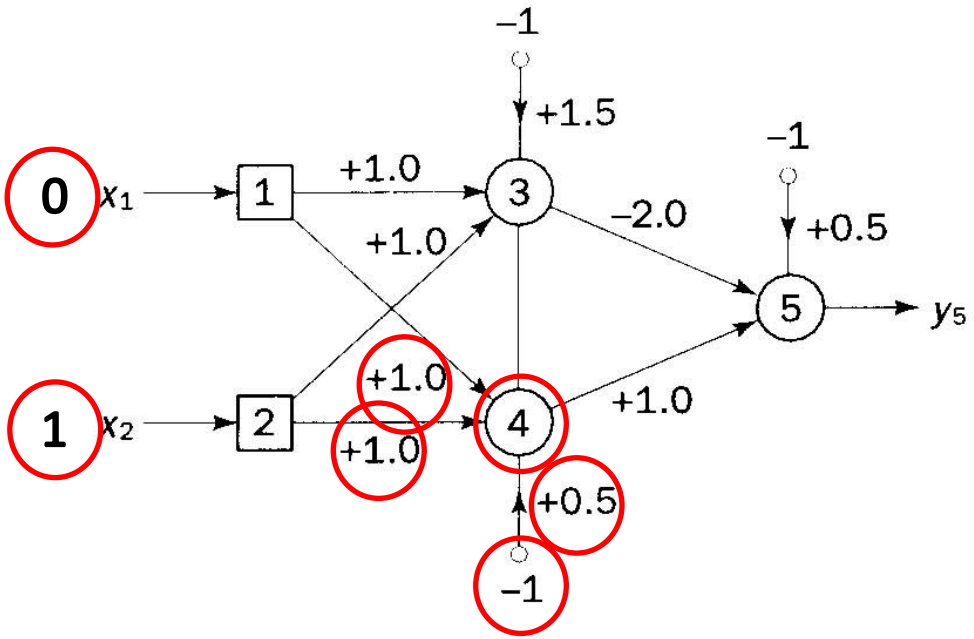

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 0 | 0 | -1.5 | -0.5 | 2 |
| 0 | 1 | -0.5 | 0.5 | 1 |
| 1 | 0 | -0.5 | 0.5 | 1 |
| 1 | 1 | 0.5 | 1.5 | 0 |

# The Topics



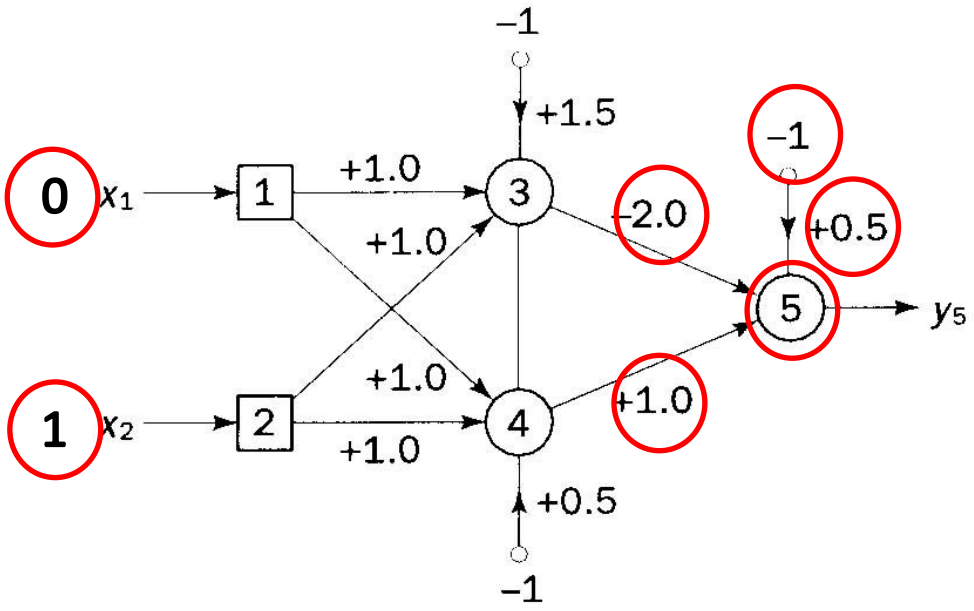

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 0 | 0 | -1.5 | -0.5 | 2 |
| 0 | 1 | -0.5 | 0.5 | 1 |
| 1 | 0 | -0.5 | 0.5 | 1 |
| 1 | 1 | 0.5 | 1.5 | 0 |

# The Topics




| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 0 | 0 | -1.5 | -0.5 | 2 |
| 0 | 1 | -0.5 | 0.5 | 1 |
| 1 | 0 | -0.5 | 0.5 | 1 |
| 1 | 1 | 0.5 | 1.5 | 0 |

# The Topics



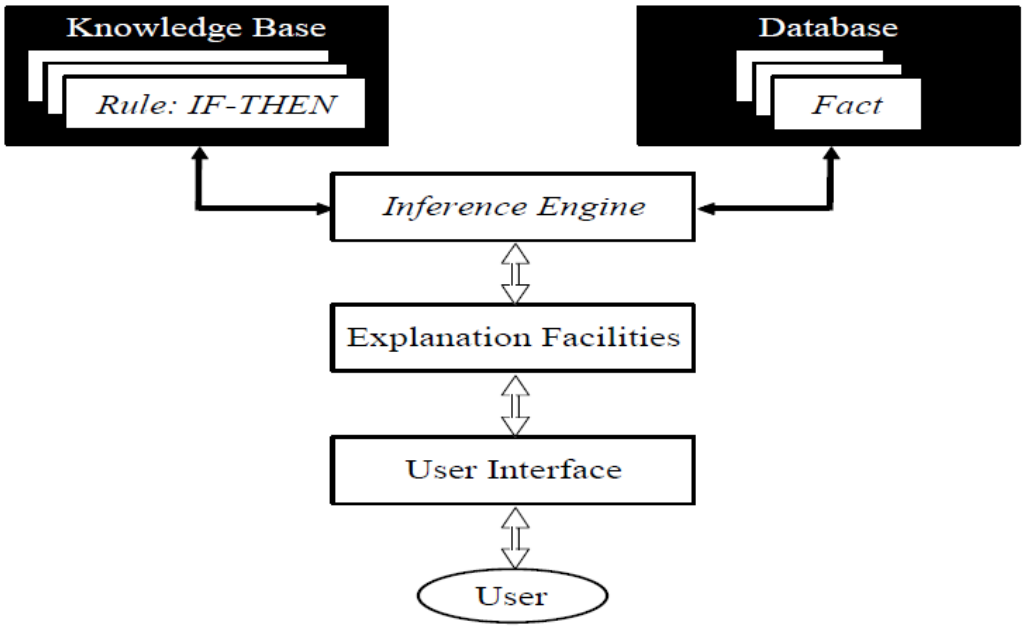

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 0 | 0 | -1.5 | -0.5 | 2 |
| 0 | 1 | -0.5 | 0.5 | 1 |
| 1 | 0 | -0.5 | 0.5 | 1 |
| 1 | 1 | 0.5 | 1.5 | 0 |

# The Topics

- Expert Systems
  - Knowledge Representation & Definition of Expert System
  - Rule based ES & Inference (forward / backward chaining & conflict resolution)
- Bayesian Networks
  - Definition and how to retrieve the joint probability
  - D-separation & Markov Blanket
- Time Series / Sequence Models
  - Markov chains and calculating probabilities of sequences
  - Hidden Markov Models and the key algorithms

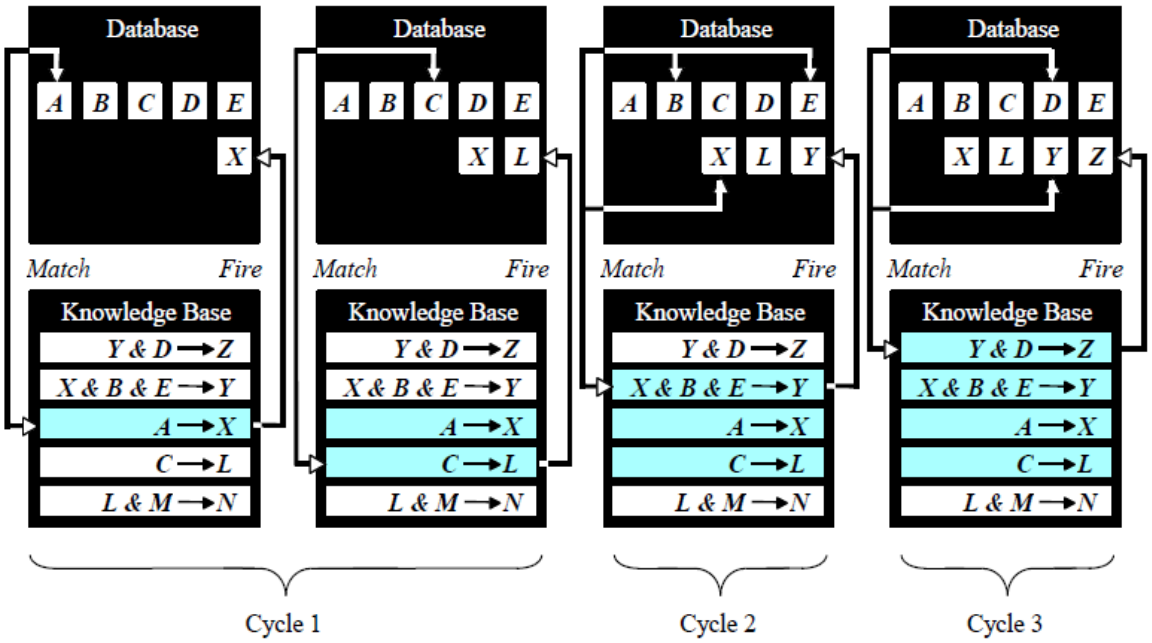
# The Topics

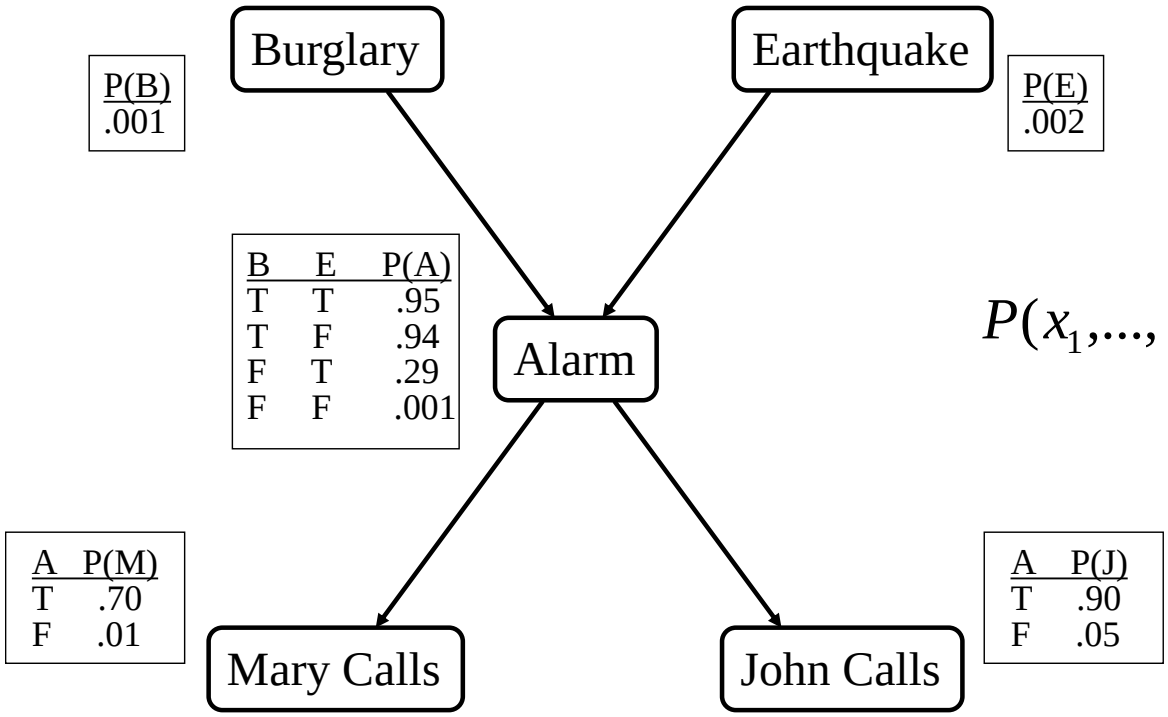
- Expert Systems
  - Knowledge Representation & Definition of Expert System
  - Rule based ES & Inference (forward / backward chaining & conflict resolution)



# The Topics

- Expert Systems
  - Knowledge Representation & Definition of Expert System
  - Rule based ES & Inference (forward / backward chaining & conflict resolution)






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

- Bayesian Networks
  - Definition and how to retrieve the joint probability
  - D-separation & Markov Blanket

e.g. Probability of John and Mary Calling, Alarm sounding, no Burglary and no Earthquake:

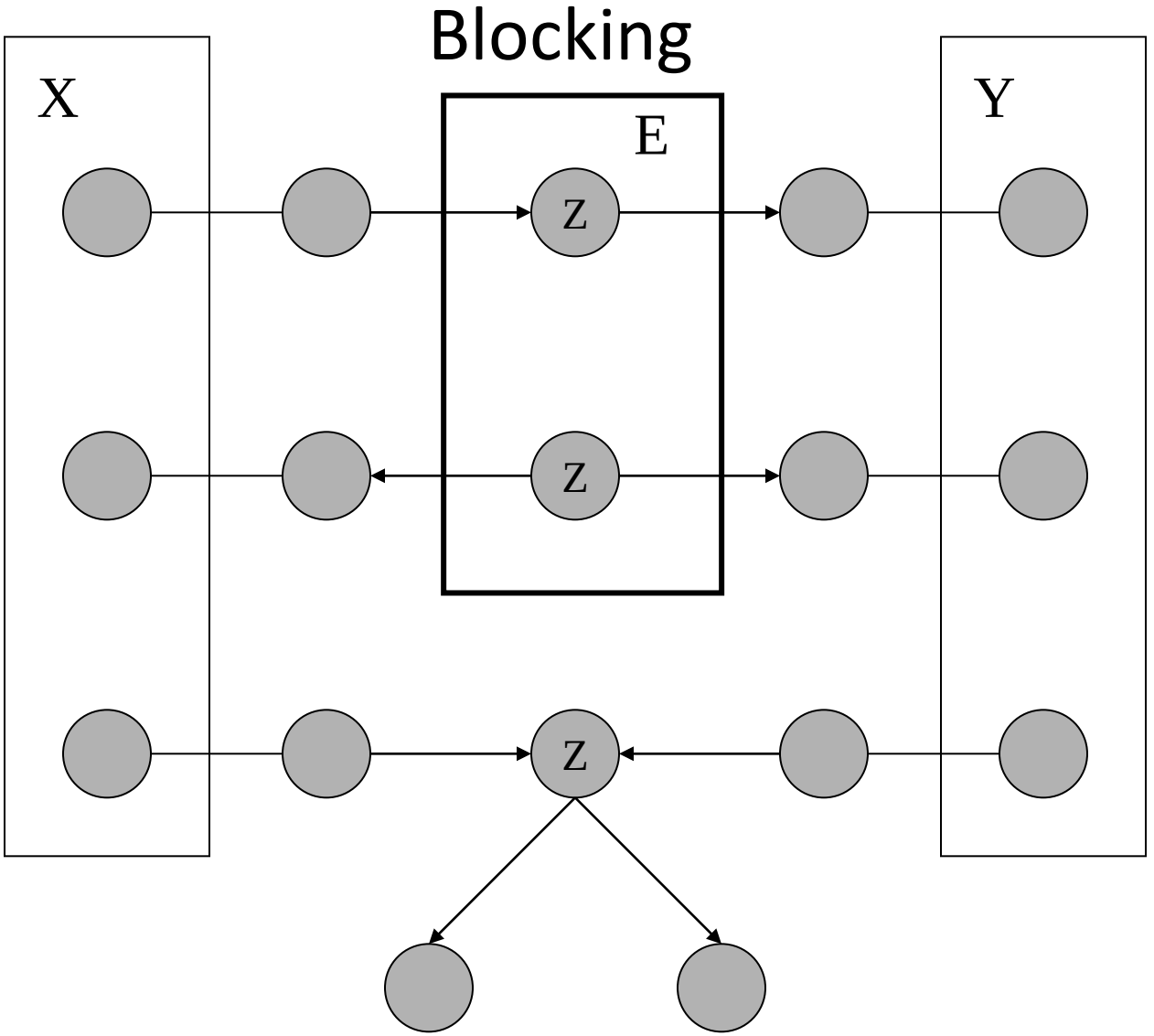
P(J & M & A & ¬B & ¬E)

= 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

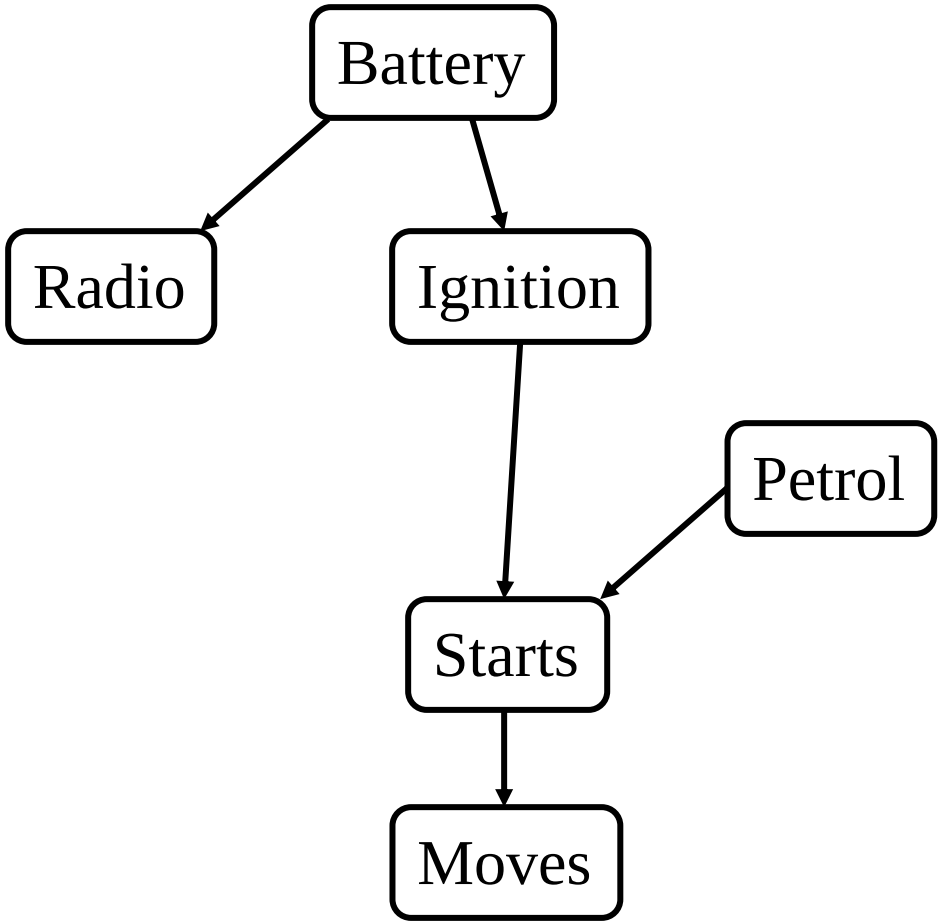
= 0.00062

# D-Separation



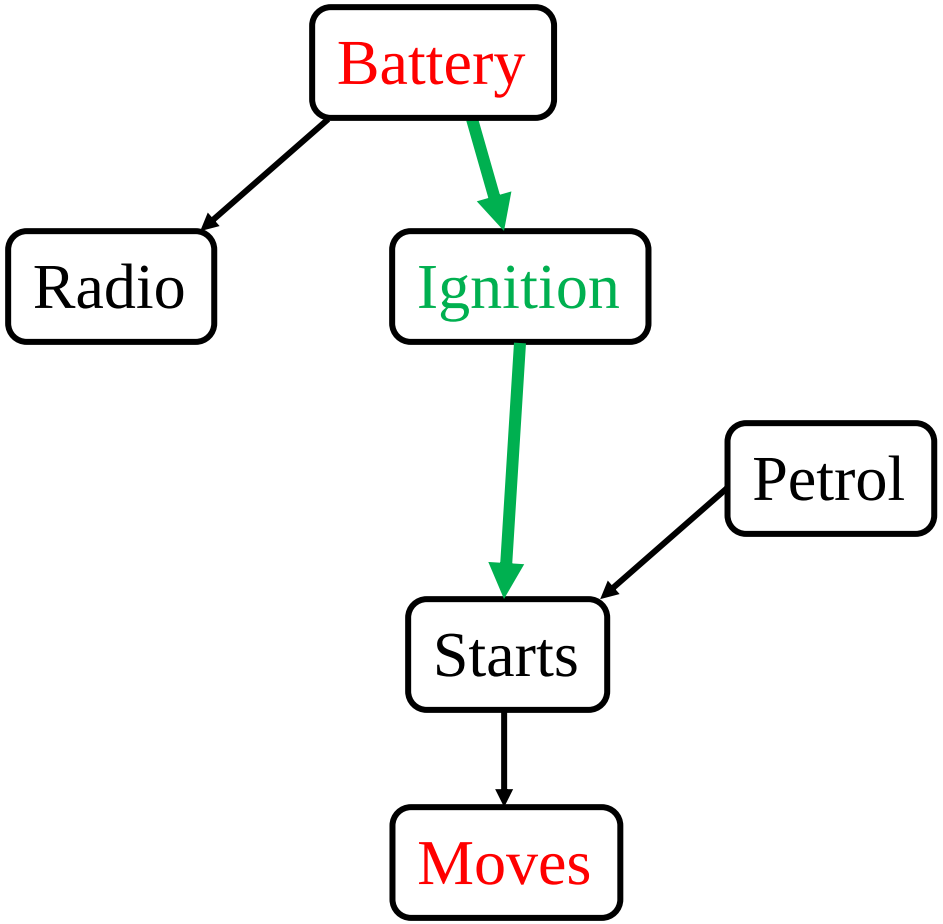
# D-Separation - Example

- 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

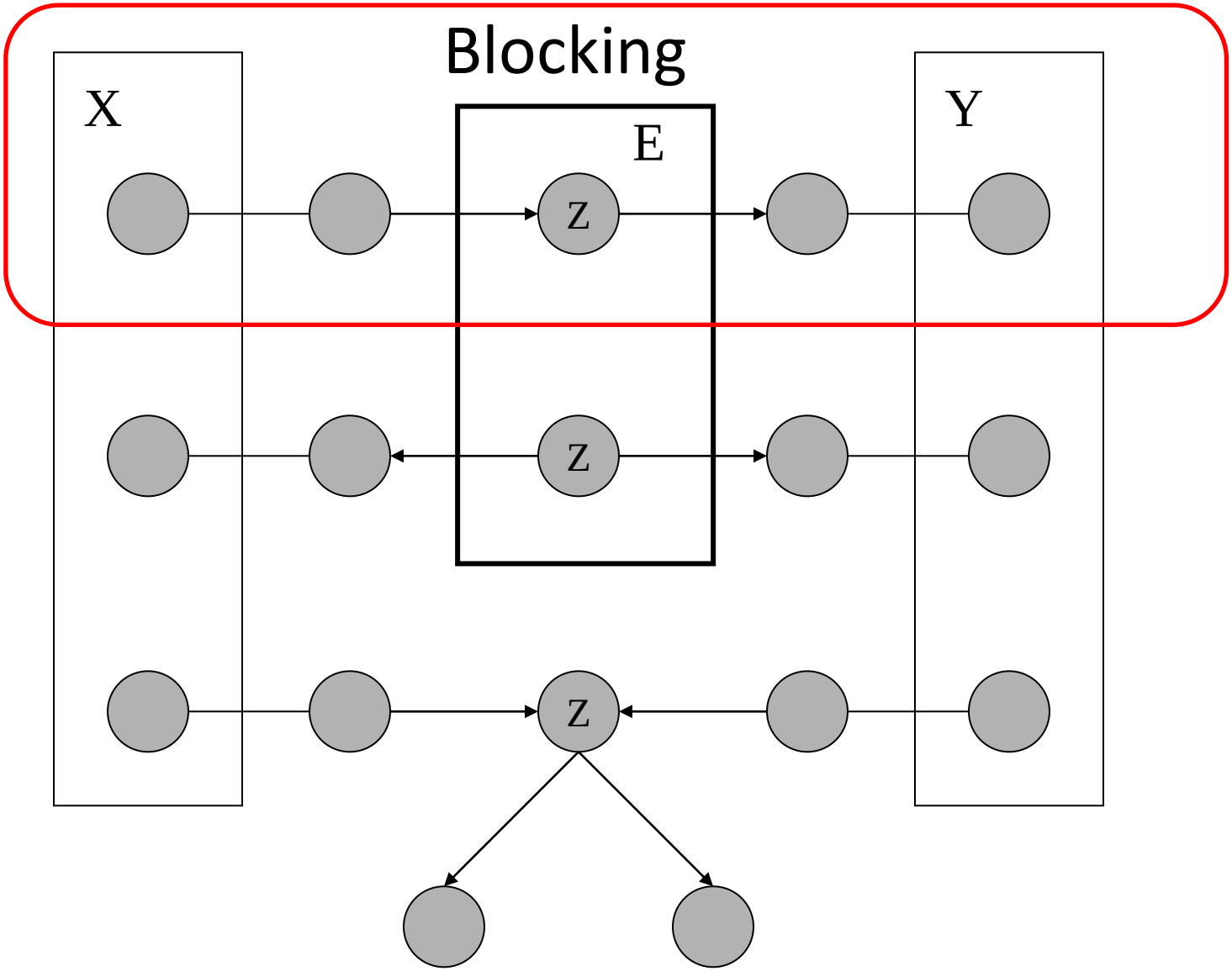


# D-Separation - Example

- 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

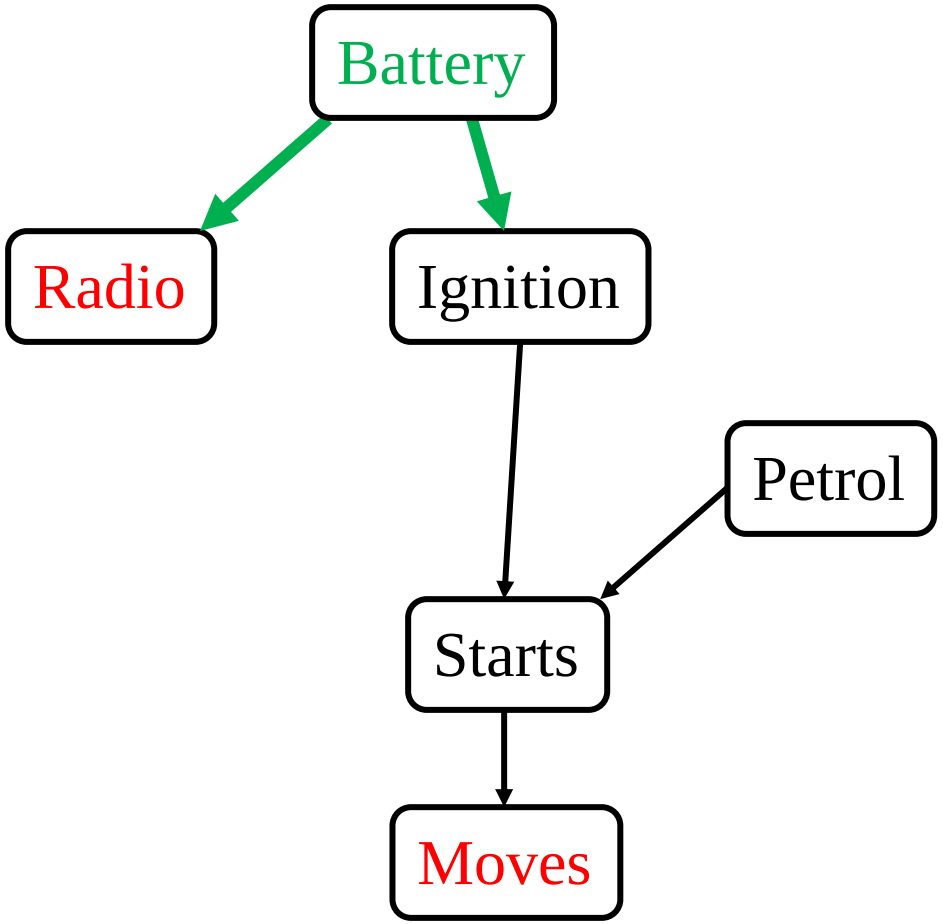


# D-Separation

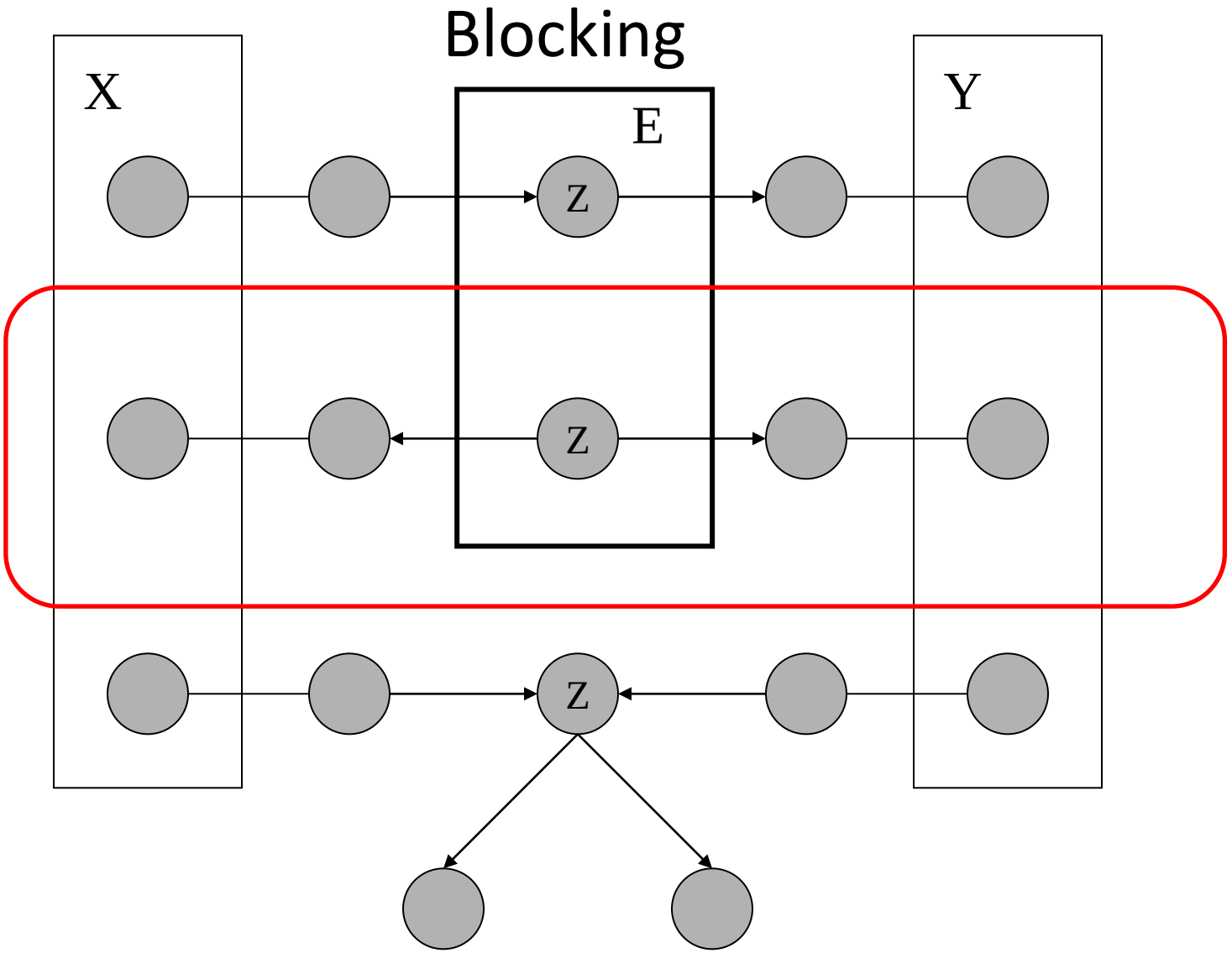


# D-Separation - Example

- 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

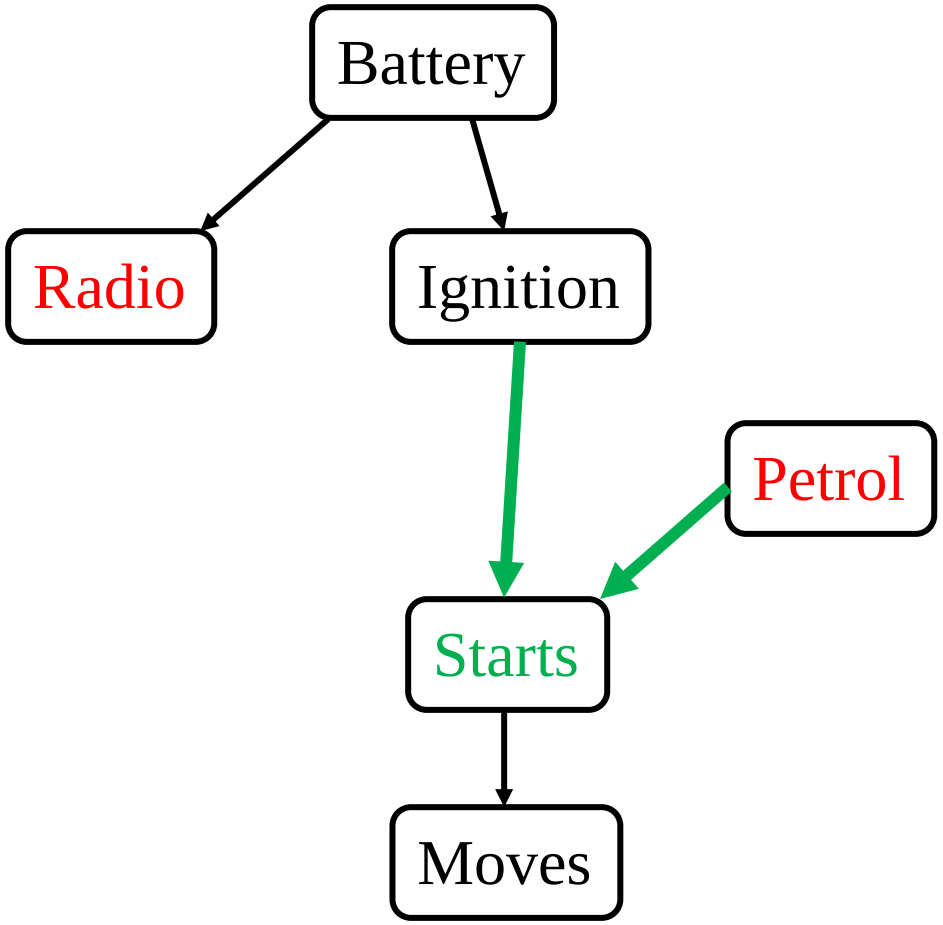


# D-Separation

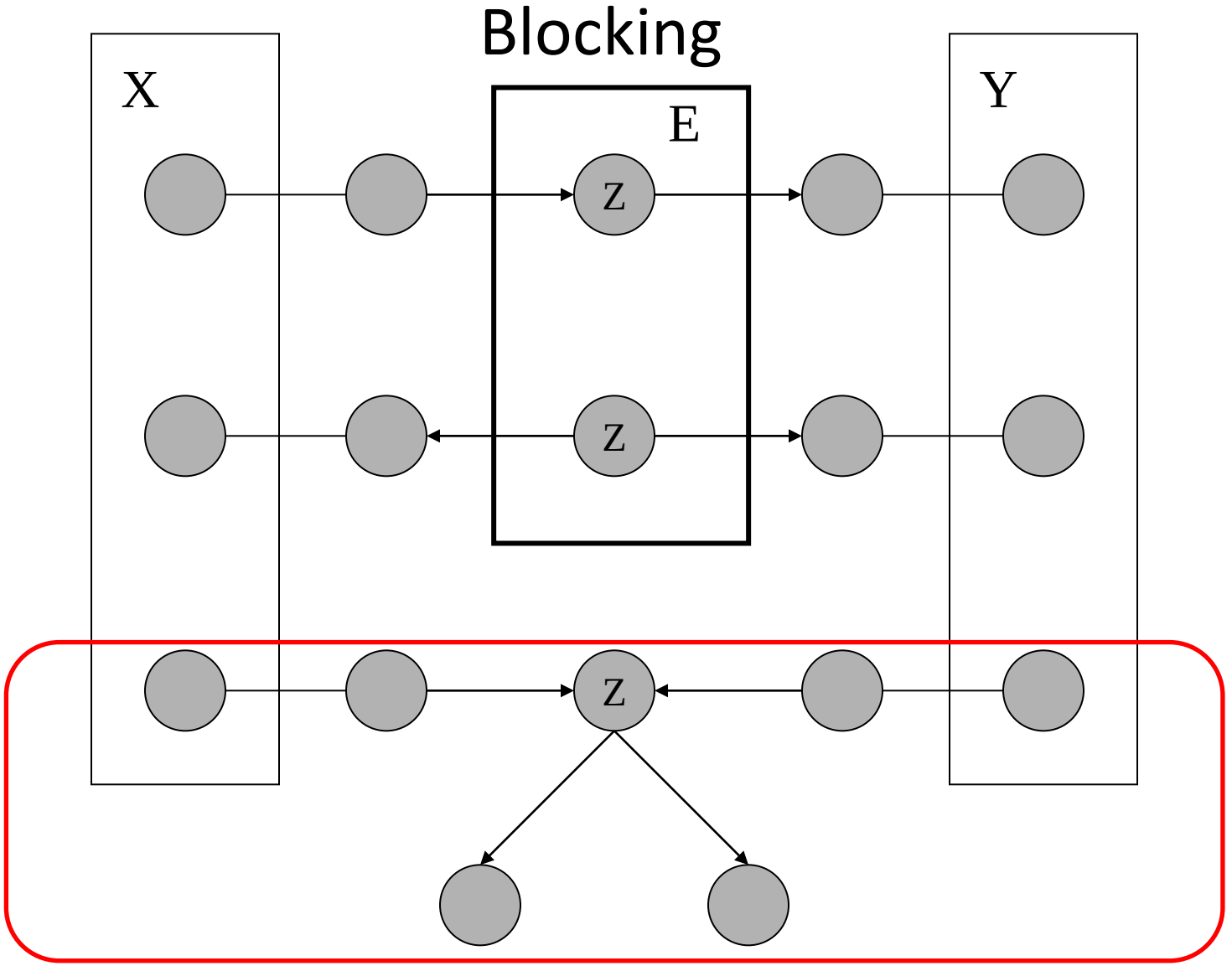


# D-Separation - Example

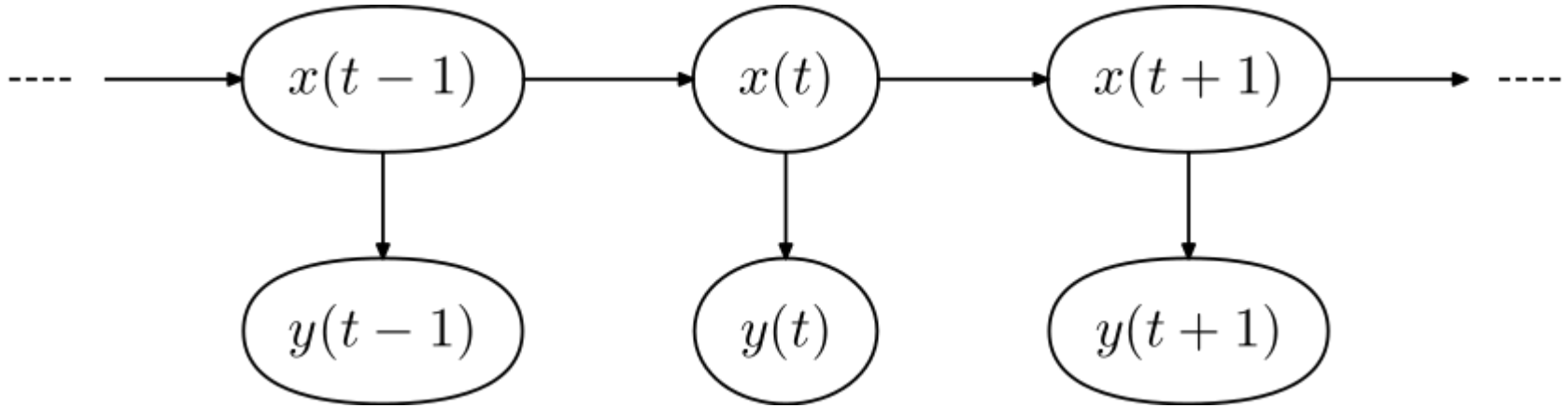
- Moves and Battery are independent given it is known about Ignition
- Moves and Radio are independent if it is known that Battery works
- Petrol and Radio are independent given no evidence. But are dependent given evidence of Starts



# D-Separation



# The Topics




- Time Series / Sequence Models
  - Markov chains and calculating probabilities of sequences
  - Hidden Markov Models and the key algorithms

# The Topics

- Initial State Distribution - P($H^1 = i$ ): $\pi = \begin{bmatrix} 1 & 0 \end{bmatrix}$
- Transition Probability Distribution – P($H^t = j \mid H^{t-1} = i$):

$$A = \begin{matrix} \\ t-1 \end{matrix} \overset{t}{\begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}}$$

- We also need an Emission (or Sensor) Distribution:

$$B = \begin{matrix} \\ H \end{matrix} \overset{X}{\begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix}}$$

- Time Series / Sequence Models
  - Markov chains and calculating probabilities of sequences
  - Hidden Markov Models and the key algorithms

# The Topics

- Deep Learning (Alina Miron)
  - Image Analysis and CNNs
  - Natural Language Processing and LSTMs
- Philosophy & Social Aspects:
  - What has been easy and what has been hard (examples)
  - Language / consciousness – Searle's Chinese room
  - Impact on Society / Automation / Trolley problem / Black Box vs Explanation

# Convolutional Neural Networks

- Convolution
- Pooling / MaxPooling
- Dropout

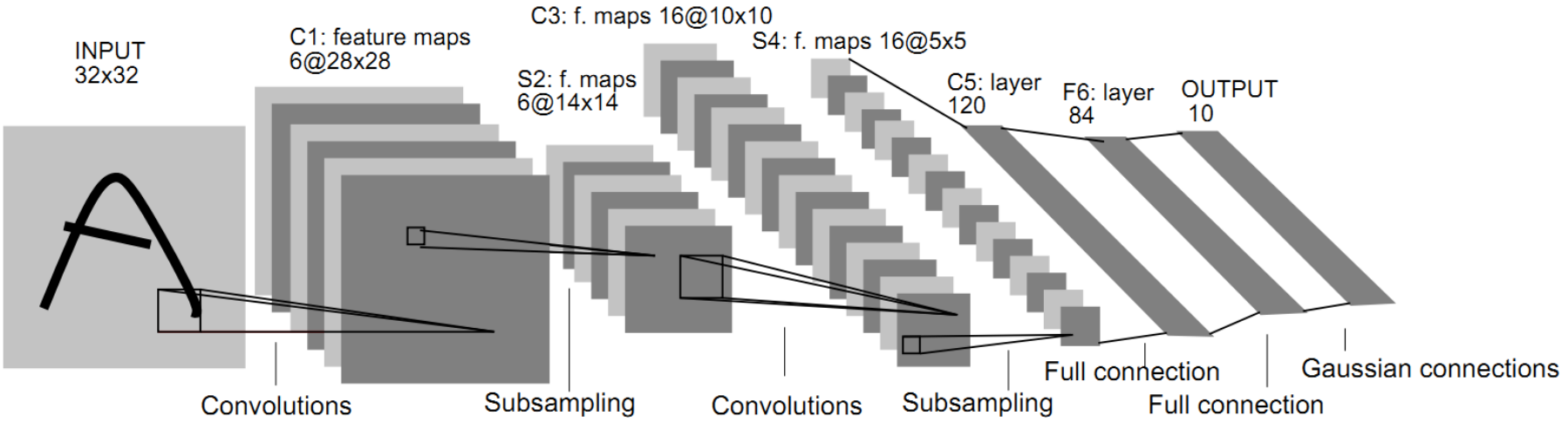
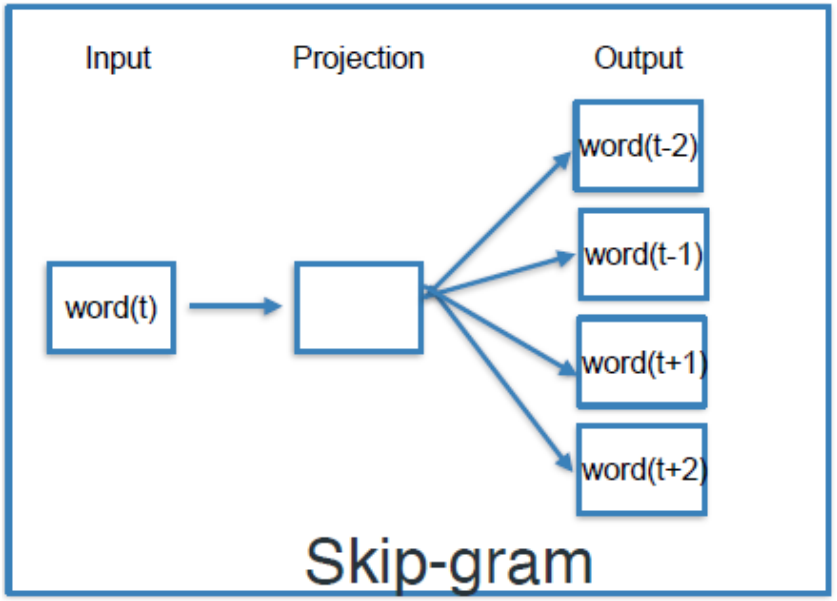



Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

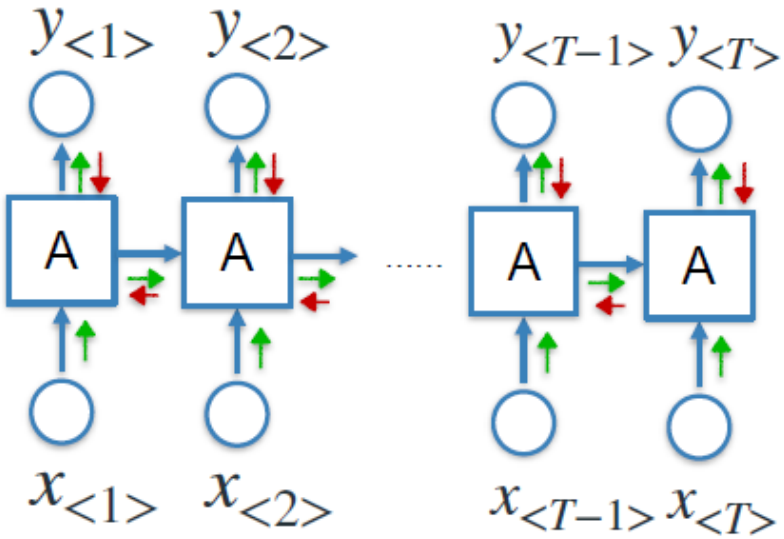
# Natural Language Processing

- One Hot Encoding
- Term frequency / Inverse Document Frequency
- Word Embedding (similar words – similar vectors)
- Relu function
- AlexNet



# Natural Language Processing

- One Hot Encoding
- Term frequency / Inverse Document Frequency
- Recurrent Neural Networks & LSTMs



# The Topics

- Philosophy & Social Aspects:
  - What has been easy and what has been hard (examples)
  - Language / consciousness – Searle's Chinese room
  - Impact on Society / Automation / Trolley problem / Black Box vs Explanation

# The Exam

- In-person exam on Wiseflow (NOT open book)
- 3 hours
- 100 marks
- Short answer questions

# Finally – Revise & Good Luck!

anorak.co.uk