IIT Jodhpur

Biological Vision and Applications
Module 03-04: Bayesian Networks

Hiranmay Ghosh

Bayesian reasoning revisited

- Bayesian framework of reasoning
  - Create a system model in terms of n stochastic (random) variables
    - $\mathcal{X} = \{X_1, X_2, X_3, \ldots, X_n\}$
    - A variable X_i can have k_i states. $X_i : \{x_i^1, x_i^2, \ldots x_i^{k_i}\}$
  - Some variables are observable, some are hidden (to be inferred)
  - Inference is a result of probability updates based on the observed data
- The joint probability distribution table will contain $\prod_i k_i - 1$ **independent** entries
- A trivial system with 10 binary variables will have $2^{10} - 1 = 1023$ entries
  - That is a big number !

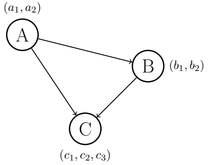
Joint probability and conditional probability

Joint probabilities

$P(a_1, b_1, c_1)$	$P(a_2, b_1, c_1)$
$P(a_1, b_1, c_2)$	$P(a_2, b_1, c_2)$
$P(a_1, b_1, c_3)$	$P(a_2, b_1, c_3)$
$P(a_1, b_2, c_1)$	$P(a_2, b_2, c_1)$
$P(a_1, b_2, c_2)$	$P(a_2, b_2, c_2)$
$P(a_1, b_2, c_3)$	$P(a_2, b_2, c_3)$

Conditional probabilities

$P(a_1)$	
$P(b_1 \mid a_1)$	$P(b_1 \mid a_2)$
$P(c_1 \mid a_1, b_1)$	$P(c_1 \mid a_2, b_1)$
$P(c_1 \mid a_1, b_2)$	$P(c_1 \mid a_2, b_2)$
$P(c_2 \mid a_1, b_1)$	$P(c_2 \mid a_2, b_1)$
$P(c_2 \mid a_1, b_2)$	$P(c_2 \mid a_2, b_2)$



Non-circular depencency between variables assumed

- Consider three variables
  - A: $\{a_1, a_2\}$, B: $\{b_1, b_2\}$, C: $\{c_1, c_2, c_3\}$
- Joint probability table will have 11 independent entries
- Equivalently, they can be expressed with 11 conditional probabilities
- The joint probabilities can be computed from the conditional probabilities, e.g.

$$\boxed{\begin{aligned} P(a_1, b_1, c_1) &= P(a_1, b_1).P(c_1|a_1, b_1) \\ &= P(a_1).P(b_1 \mid a_1).P(c_1|a_1, b_1) \end{aligned}}$$

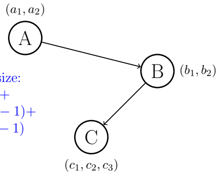
Conditional Independence

Variables A and B are conditionally independent of each other, iff $P(A.B) = P(A).P(B)$

Variables A and B are conditionally independent of each other given C, iff $P(A.B \mid C) = P(A \mid C).P(B \mid C)$

Conditional probabilities

$P(a_1)$	
$P(b_1 \mid a_1)$	$P(b_1 \mid a_2)$
$P(c_1 \mid b_1)$	$P(c_1 \mid b_2)$
$P(c_2 \mid b_1)$	$P(c_2 \mid b_2)$



CPT size:
$k_a - 1+$
$k_a.(k_b - 1)+$
$k_b.(k_c - 1)$

This topology assumes A and C are conditionally independent, given B

- Many of the variables in real world are conditionally independent of each other given the state of some other variables ☺, e.g.,
  - Color of a fruit and its shape, given the fruit
  - ...
- Conditional independence simplifies probability computations
  - Another reason to work with conditional probabilities

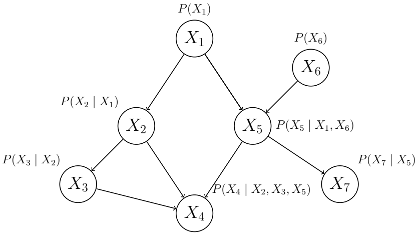
Probabilistic Graphical Models

- Graphical models exploit conditional independence
- The variables are depicted as nodes in the graph
- Only the variables that are not conditionally independent are connected with edges
- Generally a graph is sparse
  - Size of the CPT is much smaller than exhaustive joint distribution table
- There are many probabilistic graphical models
  - Markov Field, Hidden Markov Model, Bayesian Network, ...

See Koller. Probabilistic Graphical Models (book) / Cousera course

Bayesian Networks (BN)

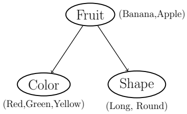
Models a probabilistic reasoning problem with cause-effect relationship



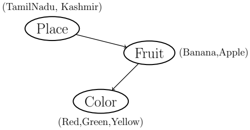
CPT Size:
$k_1 - 1+$
$k_1.(k_2 - 1)+$
$k_2.(k_3 - 1)+$
$k_2 k_3 k_5.(k_4 - 1)+$
$\ldots$

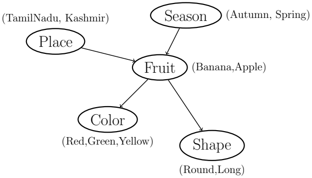
- A Directed Acyclic Graph (DAG)
- Nodes represent events in a system
  - $X_i = (x_i^1, x_i^2, \ldots, x_i^{k_i})$
  - Some nodes are observable
  - Others need to be inferred
- Edges represent **causal** relations between the events
- Conditional probabilities $P(X_i \mid \mathrm{Pa}(X_i))$are associated with every node
  - where $\mathrm{Pa}(X_i)$ represents the parent set of node X_i

Examples of BN

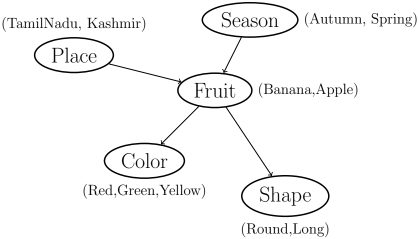


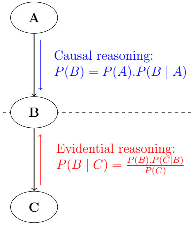
- Naïve Bayesian Network





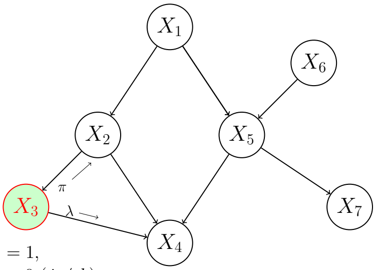
Causal inference and Evidential inference





- Fruit is inferred from
  - Causal reasoning: Where you are, what is the season (contextual cues)
  - Evidential reasoning: It's color and shape (visual cues)
- Bayesian network supports both types of reasoning

Inferencing with Bayesian Networks



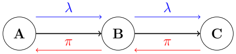
$P(x_3^k) = 1,$
$P(x_3^j) = 0 \ (j \neq k)$

Pearl's algorithm

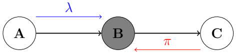
- Hand compute probabilities
  - There can be multiple (undirected) paths between a pair on nodes
  - Extremely complex
- Pearl's belief propagation algorithm
  - π and λ messages
    - Probabilities of neighboring nodes updated
  - Traverses recursively in the network
    - Till no more nodes left / blocked

Network topology and Belief propagation

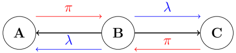
D-Separation



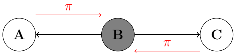
A causes **B**, **B** causes **C**, State of **B** is unknown
Belief flows between **A** and **C**



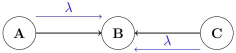
A causes **B**, **B** causes by **C**, State of **B** is known
The path between **A** and **C** is blocked by **B**



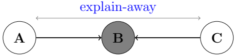
B is the cause of **A** and **C**, state of **B** is unknown
Belief flows between **A** and **C**



A and **C** are causes of **B**, state of **B** is known
The path between **A** and **C** is blocked by **B**



A and **C** are causes of **B**, state of **B** is unknown
A and **C** are conditionally independent

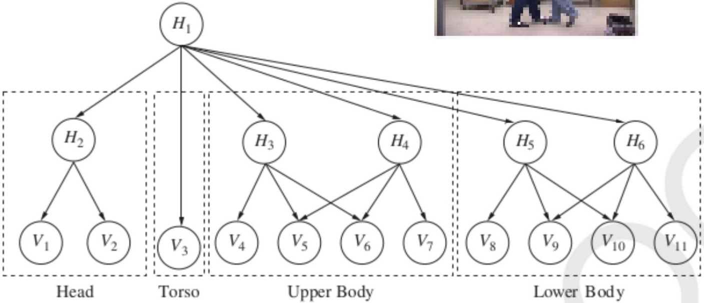


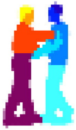
A and **C** are causes of **B**, state of **B** is known
A explains away **C**, and vice-versa

- Belief flows between two nodes in a network if there is an unblocked path between them
- If there are no unblocked path between two nodes, they are said to be d-separated

Hierarchical organization in Bayesian Network

Example

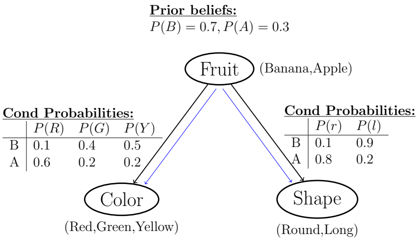




Park & Aggarwal. A hierarchical Bayesian network for event recognition of human actions and interactions

Simple Bayesian Network Example

Prior beliefs, Conditional probabilities and likelihoods



- We need to evaluate the two hypotheses
  - Fruit is either Banana or Apple
- From the given data, we can find the marginal probabilities (likelihoods)

Colors:

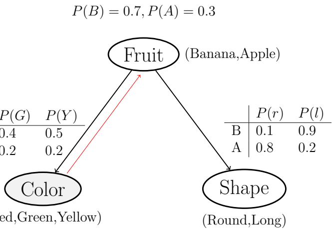
$$
\begin{aligned}
P(Red) &= P(R \mid B) \times P(B) + P(R \mid A) \times P(A) = 0.25 \\
P(Green) &= 0.34 \\
P(Yellow) &= 0.41
\end{aligned}
$$

Shapes:

$$
\begin{aligned}
P(Round) &= 0.31 \\
P(Long) &= 0.69
\end{aligned}
$$

Simple Bayesian Network Example

Posteriors



	$P(R)$	$P(G)$	$P(Y)$
B	0.1	0.4	0.5
A	0.6	0.2	0.2

	$P(r)$	$P(l)$
B	0.1	0.9
A	0.8	0.2

$P(R) = 1$
$P(G) = P(Y) = 0$

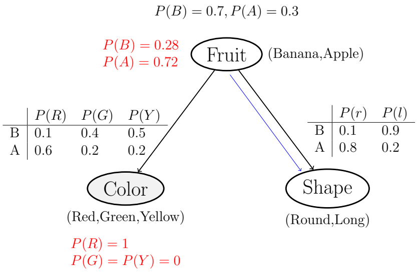
- We see a fruit to be red

$$\begin{aligned}
&\underline{\textit{Fruits:}} && \text{(un-normalized)}\\
&P(Banana \mid Red) && = P(R \mid B) \times P(B)\\
& && = 0.1 \times 0.7 = 0.07\\
&P(Apple \mid Red) && = 0.18
\end{aligned}$$

$$\begin{aligned}
&\underline{\textit{Fruits:}} && \text{(normalized)}\\
&P(Banana \mid Red) && = 0.28\\
&P(Apple \mid Red) && = 0.72
\end{aligned}$$

Simple Bayesian Network Example

Posteriors (contd.)



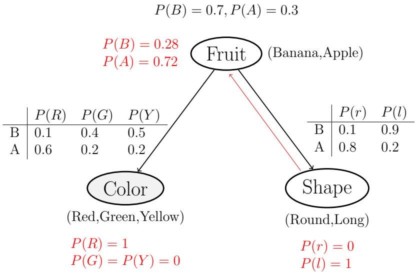
- Change in probability of fruits changes posterior probability of shapes

Shapes:

$$P(Round) = 0.60$$
$$P(Long) = 0.40$$

Simple Bayesian Network Example

Posteriors (contd.)



- Now we see the fruit to be **long**

Fruits: (un-normalized)

$$
\begin{aligned}
P(Banana \mid Red, Long) &= P(l \mid B) \times P(B) \\
&= 0.9 \times 0.28 = 0.25 \\
P(Apple \mid Red, Long) &= 0.14
\end{aligned}
$$

Fruits: (normalized)

$$
\begin{aligned}
P(Banana \mid Red, Long) &= 0.63 \\
P(Apple \mid Red, Long) &= 0.37
\end{aligned}
$$

Quiz

Quiz 03-04

End of Module 03-04