BITS Pilani
Pilani Campus

Artificial & Computational Intelligence

DSECSZG557

M5 : Probabilistic Representation and Reasoning

Dr. Sudheer Reddy

Course Plan

M1 Introduction to AI

M2 Problem Solving Agent using Search

M3 Game Playing

M4 Knowledge Representation using Logics

M5 Probabilistic Representation and Reasoning

M6 Reasoning over time, Reinforcement Learning

M7 Ethics in AI

Reasoning

- Monotonic Reasoning

- Non- Monotonic Reasoning

  Dependency Directed Backtracking: when a statement is deleted as “ no more valid”, other related statements have to be backtracked and they should be either deleted or new proofs have to be found for them. This is called dependency directed backtracking (DDB)

Reasoning

- Monotonic Reasoning
- Non- Monotonic Reasoning

Monotonic	Non-Monotonic
Consistent	Relaxed Consistency
Complete Knowledge	Incomplete Knowledge
Static	Dynamic
Discrete	Continuous & Learning Agent
Predicate Logic	Probabilistic Model

Bayesian Net???

Wumpus World Problem

- Wumpus Ghost traces of scent in the visited cell
- Earlier visited cell may become unsafe!!!
- **Problem**: Given the information that there is a possibility of apparition of Wumpus anywhere in the cave, AI agent needs to be safely travel with more caution!!

Uncertainty

You can reach Bangalore Airport from MG Road within 90 mins if you go by route A.

- There is uncertainty in this information due to partial observability and non determinism
- Agents should handle such uncertainty

Previous approaches like Logic represent all possible world states

Such approaches can't be used as multiple possible states need to be enumerated to handle the uncertainty in our information

Uncertainty

You can reach Bangalore Airport from MG Road within 90 mins if you go by route A.

Road Block	Festival Season	Weekend	Observation (20)	Prob
F	F	F	12	0.6
F	F	T	3	0.15
F	T	F	2	0.1
F	T	T	2	0.1
T	F	F	0	0
T	F	T	0	0
T	T	F	1	0.05
T	T	T	0	0
				=1

Probability Theory

Basics
Conditional Probability
Chain Rule
Independence
Conditional Independence
Belief Nets
Joint Probability distribution

Probability Basics - Refresher

Sample Space: Set of all possible outcomes.

- Ex: After tossing 2 coins, the set of all possible outcomes are
- {HH, HT, TH, TT}

Event: A subset of a sample space.

- An event of interest might be - {HH}

Probability Basics - Model

A fully specified **probability model** associates a numerical probability P(ω) with each possible world.

The **basic axioms**

1. Every possible world has a probability between 0 and 1
2. Sum of probabilities of possible worlds is 1

   P(True) = 1

   P(False) = 0

3. P(a ∨ b) = P(a) + P(b) - P(a ∧ b)

E.g., P(HH) = 0.25; P(HT) = 0.25; P(TT) = 0.25, P(TH) = 0.25

$$0 \leq P(\omega) \leq 1 \text{ for every } \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1$$

Probability Basics – Exclusive / Exhaustive events

Mutually Exclusive Events:

- Two events (or propositions) are mutually exclusive or disjoint if they cannot both occur at the same time(be true).
- A clear example is the set of outcomes of a single coin toss, which can result in either heads or tails, but not both.

Exhaustive Events:

- A set of events is jointly or collectively exhaustive if at least one of the events must occur.
- E.g., when rolling a six-sided die, the events 1, 2, 3, 4, 5, and 6 are collectively exhaustive.

Probability Basics - Propositions

Probabilistic assertions (Propositions)

- Usually not a particular world event but about a set of them
- E.g., two dice when rolled, a proposition φ can be “the sum of dice is 11”

For any proposition φ,

$P(\varphi) = P(\text{sum} = 11) \qquad = P((5, 6)) + P((6, 5))$

$= 1/36 + 1/36 = 1/18$

Probability Basics – Unconditional/Prior

Unconditional / Prior probabilities:

Propositions like P(sum = 11) or P(two dices rolling equals) are called Unconditional or Prior probabilities

They refer to degree of belief in absence of any other information

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

$$P(a \wedge b) = P(a \mid b)P(b)$$

Probability Basics - Conditional

However, most of the time we have some information, we call it **evidence**

E.g., we can interested in two dice rolling a double (i.e., 1,1 or 2,2, etc)

When one die has rolled 5 and the other die is still spinning

Here, we not interested in unconditional probability of rolling a double

Instead, the **conditional** or **posterior** probability for rolling a double given the first die has rolled a 5

$P(doubles \mid Die_1 = 5)$ where | is pronounced “given”

E.g., if you are going for a dentist for a checkup, P(cavity) = 0.2

- If you have a toothache, then P(cavity | toothache) = 0.6

Independence

If we have two random variables, TimeToBnlrAirport and HyderabadWeather

P(TimeToBnlrAirport, HyderabadWeather)

To determine their relation, use the product rule

= P(TimeToBnlrAirport | HyderabadWeather) / P(HyderabadWeather)

However, we would argue that HyderabadWeather and TimeToBnlrAirport doesn't have any relation and hence

P(TimeToBnlrAirport | HyderabadWeather) = P(TimeToBnlrAirport)

This is called Independence or Marginal Independence

Independence between propositions a and b can be written as

$$P(a \mid b) = P(a) \quad \text{or} \quad P(b \mid a) = P(b) \quad \text{or} \quad P(a \wedge b) = P(a)P(b)$$

Bayes Rule

Using the product rule for propositions a and b

$$P(a \wedge b) = P(a \mid b)P(b) \qquad \text{and} \qquad P(a \wedge b) = P(b \mid a)P(a)$$

Equating the right hand sides and dividing by P(a)

$$P(b \mid a) = \frac{P(a \mid b)P(b)}{P(a)}$$

This is called the Bayes Rule

Conditional Independence

2 random variables A and B are conditionally independent given C iff

$$P(a, b \mid c) = P(a \mid c)\, P(b \mid c) \text{ for all values } a, b, c$$

More intuitive (equivalent) conditional formulation

- A and B are conditionally independent given C iff

  $$P(a \mid b, c) = P(a \mid c) \text{ OR } P(b \mid a, c) = P(b \mid c), \text{ for all values } a, b, c$$

- Intuitive interpretation:

  P(a | b, c) = P(a | c) tells us that learning about b, given that we already know c, provides no change in our probability for a, i.e., b contains no information about a beyond what c provides

  P(R | F, P) = P(R | P)

Joint Probability Distributions

Instead of distribution over single variable, we can model distribution over multiple variables, separated by comma

E.g., **P**(A, B) = **P**(A | B) . **P**(B)

P(A, B) is the probability distribution over combination of all values of A and B

E.g., if A = Weather and B = Cavity

$$
\begin{aligned}
P(W = sunny \wedge C = true) &= P(W = sunny|C = true)\, P(C = true)\\
P(W = rain \wedge C = true) &= P(W = rain|C = true)\, P(C = true)\\
P(W = cloudy \wedge C = true) &= P(W = cloudy|C = true)\, P(C = true)\\
P(W = snow \wedge C = true) &= P(W = snow|C = true)\, P(C = true)\\
P(W = sunny \wedge C = false) &= P(W = sunny|C = false)\, P(C = false)\\
P(W = rain \wedge C = false) &= P(W = rain|C = false)\, P(C = false)\\
P(W = cloudy \wedge C = false) &= P(W = cloudy|C = false)\, P(C = false)\\
P(W = snow \wedge C = false) &= P(W = snow|C = false)\, P(C = false)\,.
\end{aligned}
$$

Probabilistic Inference

Computation of posterior probabilities given observed evidence, i.e., full joint probability distribution

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

Query: P(cavity ∨ **toothache)**

0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28

Conditional Probability

Towards Chain Rule:
P(a | b) = P(a,b) / P(b)

P(a, b) = P(a | b) P(b)
P(a, b , c) = P(a , x) where x= b,c
P(a,x) = P(a | x) . P(x)
 = P(a | bc) . P(b, c)
 = P(a | bc) . P(b | c). P(c)

Hence : P(a,b,c) = P(a | bc) . P(b | c). P(c)
Chain Rule : Generalization

$$\mathbf{P(X_1, X_2, \ldots X_K) = \prod P(X_i \mid X_{i-1}, \ldots X_1)}$$

Where i = k to 1 (reverse)

Probability Theory

<u>Independence</u>

P(a | b) = P(a)

Implication:

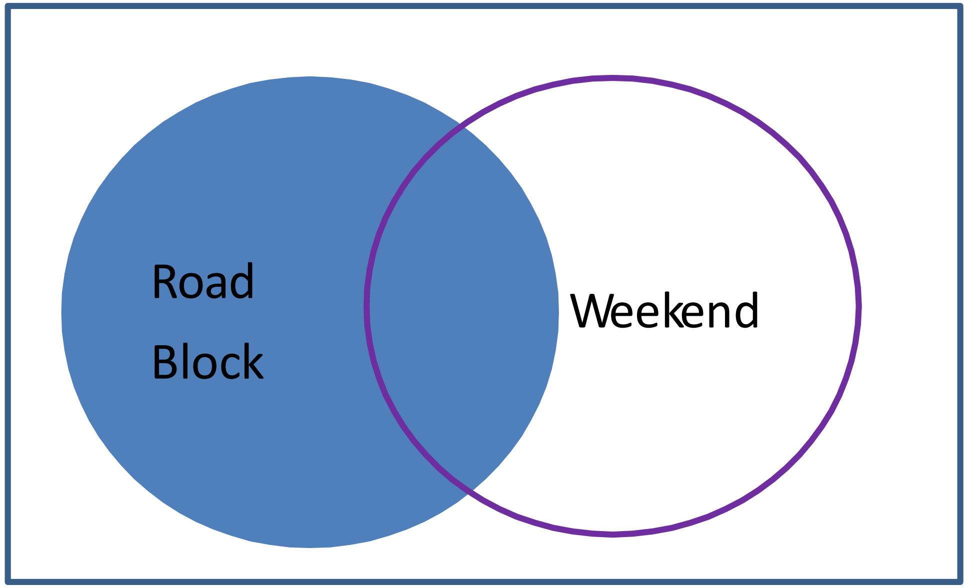
P(a | b) = P(a,b) / P(b)

P(a) = P(a,b) / P(b)

P(a,b) = P(a) . P(b)

<u>Conditional Independence</u>

P(a | b c) = P(a | c)



Probability Theory

<u>Conditional Independence</u>

P(a | b c) = P(a | c)

Extension:

P(a b | c) = P(a | c) . P(b | c)

Required Reading: AIMA - Chapter # 13

Thank You for all your Attention

Note : Some of the slides are adopted from AIMA TB materials