Bayesian Networks

Lecture 1: Uncertainty and Probability

Jingpeng Li

Outline

- 5 Lectures
  - Lecture 1: Uncertainty and Probability
  - Lecture 2: Bayes Theorem
  - Lectures 3, 4, 5: Bayesian Networks
- Introduction and background
  - Decision support systems
  - Bayesian networks
- Uncertainty
- Probability

2

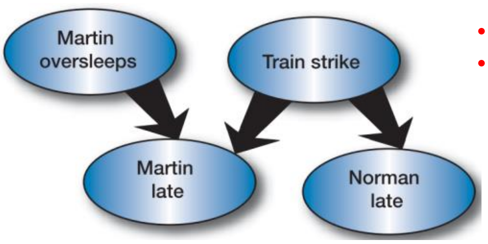
What is a Decision Support System (DDS)?

- A computer-based system that collects, organises and analyses data to facilitate quality business decision-making for management, operations and planning
- Systems that help people make decisions based on data that is combined from a wide range of sources

3

What is a Bayesian Network?

- A Bayesian Network (BN) is a way of describing the relationships between causes and effects, and is made up of nodes and arcs



- Nodes: Random Variables
- Arcs: Causal or influential relationships

False	0.9
True	0.1

NPT for “Train strike”

For each node Tables give the *the conditional probability* of each possible outcome, given combinations of outcomes from parent nodes

Norman Late

Train Strike	False	True
False	0.9	0.2
True	0.1	0.8

4

Applications of Bayesian networks

- Main uses: to support decision making and to find strategies to solve tasks under uncertainty
- Initially (80s and 90s) focus on diagnostic problems: medicine and fault diagnosis
- Nowadays: fault diagnosis, pharmaceutical trials, air traffic management, terrorist treat assessment, spatial mission control, etc.

Commercial tools & companies

- Agenarisk: Agena www.agenarisk.com
- HUGIN Expert: https://www.hugin.com/

5

Uncertainty & Probability

- The key feature of BNs is that they enable us to model and reason about uncertainty
- We need probability to quantify uncertainty
- To understand and implement Bayesian networks, we need two mathematical tools
  - Probability theory
  - Propositional logic

6

Some definitions

- **Uncertainty**: The lack of certainty. A state of having limited knowledge where it is impossible to exactly describe the existing state, a future outcome, or more than one possible outcome.
- **Measurement of uncertainty**: A set of possible states or outcomes where probabilities are assigned to each possible state or outcome
- **Risk**: A state of uncertainty where some possible outcomes have an undesired effect or significant loss.
- **Measurement of Risk**: A set of measured uncertainties where some possible outcomes are losses, and the magnitudes of those losses

Dealing with uncertainty

- The real-world is imperfect and knowledge based systems are expected to operate in the real-world
- Why the difficulty? Uncertain information
  - missing data
  - ambiguous data
- 3 methods of dealing with uncertain reasoning
  - Confidence Factors/certainty Factors
  - Fuzzy logic
  - **Bayesian theory** (Probability, Propositional Logic)

Propositional (symbolic) logic

- **Logic**: the study of the principles of correct reasoning
- **Propositional logic**: study the ways of joining and/or modifying entire statements or propositions to form more complex ones
- ***Statement****:* a declarative sentence, a complete grammatical sentence that makes a claim
  - *Snow is white*
  - *The moon is made of cheese*
- Statements are considered as indivisible units. They can be True or False
- Letters or simple words (*A, B, Martin late, green*) are used to represent statements

(Operator connectives)

Symbol	Meaning	Notes
¬	negation (NOT)	The tilde (˜) is also often used
∧	conjunction (AND)	The ampersand (&) or dot (·) are also often used
∨	disjunction (OR)	This is the inclusive disjunction, equivalent to “or” in English

9

Introduction to Probability

- Consider a chance *experiment* with a finite number of possible *outcomes.* Examples
  - Toss a coin: H(heads), T(tails)
  - Roll a dice: 1, 2, 3, 4, 5, 6
- Mathematical notation
  - The sample space *W* is a set with all possible outcomes
  - *w* in *W* are the possible outputs
- A *probability model* is a sample space with an assignment *P(w)* for each possible outcome
  - P(w) is a number in [0, 1]
  - Sum of all P(w) = 1

10

How to calculate probabilities?

Experiment: extract a single marble from the bag

What is the probability of

- Pick a yellow marble, P(yellow) = ?
- Pick a green marble, P(green) = ?
- Pick a yellow or green marble, P(yellow or green) = ?
- Pick not a green marble, P(not green) = ?

$$\text{Probability of an event} = \frac{\text{Number of positive outcomes}}{\text{Total number of possible outcomes}}$$

Math Goodies
http://www.mathgoodies.com/lessons/vol6/intro_probability.html

11

Probabilities – OR condition "V"

Determine the probability of either one or more events occurring

Example from bags of marble: P(yellow or green)

Whenever the **OR** condition is used we simply add the probabilities together to find out the probability of one or the other occurring.

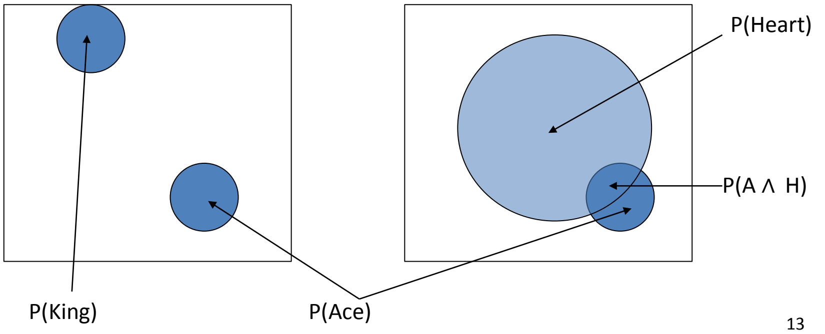
P(yellow or green) = P(yellow) + P (green) = 1/10 + 3/10 = 4/10 = 1/5

Note: order is not important

P(yellow or green) = P(green or yellow)

Probabilities – OR condition "V"

- There are **52** cards in a deck, **4** suits, each with one of **13** possible values from 2 to Ace
- Consider the Venn diagram
  - The area of the rectangle is 1, and the area of each region gives the probability of the event(s) associated with that region
- P(A ∨ B) means "the probability of observing event A or event B"



13

Probabilities – OR condition "V"

- The probability of picking an Ace or a King is

  P(A or K) = (4 + 4)/52 = 8/52 = 2/13 = 0.15 = 15%

- How is the probability of picking an Ace or a King related to the probability of picking an Ace and to the probability of picking a King?

  P(A or K) = P(A) + P(K) (4/52 +4/52)

- What about the probability of picking a Heart **or** an Ace **?**

  P(H or A) = (13 + **3**)/52 = ... ≠ P(H) + P(A) = (4 + 13) /52

  P(A or B) = P(A ∨ B) = P(A) + P(B) with A, B disjoint.

  P(A or B) = = P(A) + P(B) - P(A ∧ B) otherwise

  (what about A or B or C ?)

  P(A or H) = P(A) + P(H) - P(A ∧ H) = (4 + 13 **- 1**) / 52 = ...

  (event / all outcomes !)

14

Probabilities – AND condition "∧"

Determine the probability of 2 or more events happening together

Independent events: the occurrence of one event does not impact on the occurrence of another event

Examples

- Getting a 6 with a roll of a die and flipping a coin and it landing on heads
- The likelihood of passing an exam and the probability I choose to write my exam answers in blue ink
- The probability Jane wins the lottery and the probability Jane lives in Glasgow

Dependent events: the occurrence of one event does impact on the probability of the occurrence of another event

Examples

- Having clouds in the sky and the likelihood of rain
- Smoking 40 cigarettes a day and the likelihood of contracting lung cancer
- Doing no revision and the likelihood of passing an exam

15

Probabilities – AND condition "∧": Independent events

If 2 (say, event A and event B) or more events are independent (not connected) then

P(A AND B) = p(event A) * p(eventB) (Product Rule)

Worked example: Rolling a die and flipping a coin

- P(rolling a 6 with die) = **1/6**
- P(heads) = **1/2**
- P(rolling a 6 AND coin landing on heads) = **(1/6)*(1/2) = 1/12**
- P(rolling a 6 OR coin landing on heads) = **1-(5/6)*(1/2) = 7/12**

NB: P(A ∧ B) = P(B ∧ A)

Order is not important here. WHY?

16

Probabilities – AND condition "∧": Dependent events

- If the occurrence of one event does affect the probability of the other occurring, then the events are dependent.
- This dependency is used in **Conditional Probability**
- The probability of event B occurring given that event A has already occurred is read "*the probability of B given A*" and is written: P(B|A)
- With dependent events,

  P(A and B) = P(A) * P(B|A)

  which can also be written as

  $P(A \wedge B) = P(A) * P(B|A)$

NB: $P(A|B) \neq P(B|A)$

Order is important here WHY?

17

Probabilities – AND condition "∧": Dependent events

- The question, "Do you smoke?" was asked of 100 people. Results are shown in the table.
- The table allows us to work out both joint probabilities for dependent events and conditional probabilities for dependent events.
- Why are the events dependent?

Smoker

	Yes	No	Total
Male	19	41	60
Female	12	28	40
Total	31	69	100

18

Probabilities – AND condition "Λ": Dependent events

Smoker	Yes	No	Total
Male	19	41	60
Female	12	28	40
Total	31	69	100

- What is the probability of a randomly selected individual being a male who smokes?
  - This is just a joint probability. The number of "Male and Smoke" divided by the total = 19/100 = 0.19

    P(male Λ smoker) = 0.19
- What is the probability of a randomly selected individual being a male?
  - This is the total for male divided by the total = 60/100 = 0.60. Since no mention is made of smoking or not smoking, it includes all the cases.

    P(male) = 0.6
- What is the probability of a randomly selected individual smoking?
  - Again, since no mention is made of gender, this is a marginal probability, the total who smoke divided by the total = 31/100 = 0.31.

    P(smoker) = 0.31

19

Probabilities – AND condition "Λ": Dependent events

Smoker	Yes	No	Total
Male	19	41	60
Female	12	28	40
Total	31	69	100

- Having selected a male, what is the probability of him being a smoker?
  - This time, you're told that you have a male - think of stratified sampling. What is the probability that the male smokes? Well, 19 males smoke out of 60 males, so 19/60 = 0.3167

    P(smoker | male) = p(male Λ smoker) / p(male)

    = 0.19/ 0.6 = 0.3167
- What is the probability that a randomly selected smoker is male?
  - This time, you're told that you have a smoker and asked to find the probability that the smoker is also male. There are 19 male smokers out of 31 total smokers, so 19/31 = 0.6129

    P(male | smoker) = p(smoker Λ male) / p(smoker)

    = 0.19/ 0.31 = 0.6129

20

Probability: summary

- The probability of all outcomes add to 1
- If we want the probability of either event A **or** event B occurring we '**+**' the two independent probabilities together
- If we want the probability of event A **and** event B occurring
  - If the two events are independent '*****' the two probabilities together
  - If the two events are not independent we must use the conditional rule:

$$P(A \land B) = P(B|A)\ P(A)$$

21

Lecture: summary

- Decision support Systems: help people make decisions based on data that is combined from a wide range of sources
- A Bayesian Network (BN) is a way of describing the relationships between causes and effects
- BNs used to support decision making and to find strategies to solve tasks under *uncertainty*
- BNs use Bayesian probability theory
- Next lecture we will study the Bayes Theorem

22