KR-IST Lecture 9a
Bayesian networks

Chris Thornton

November 22, 2013

Bayes rule again

Given evidence E and some conclusion C, it's always the case that

$$P(C|E) = \frac{P(C)P(E|C)}{P(E)}$$

We can plug any values we like into this formula to infer the probability of the conclusion given the evidence.

Lecture attendance example

So, from

$$P(\text{attendLectures}|\text{passExam}) = 0.8$$
$$P(\text{passExam}) = 0.6$$
$$P(\text{attendLectures}) = 0.5$$

we can use Bayes rule to calculate probability of passing the exam:

$$P(\text{passExam}|\text{attendLectures}) \quad = \quad = \frac{0.8 \times 0.6}{0.5} \quad = 0.96$$

The probability of passing the exam given you attend lectures is 0.96.

Complex Bayesian reasoning

Bayes' rule provides a single step of probabilistic backwards reasoning.

This works for simple scenarios, e.g., where we have a lot probabilities relating diseases to symptoms, and want a rule that produces a diagnosis from the symptoms shown.

But in more complex cases, we may have a networks or chain of probabilistic relationships to deal with.

For example,

```
cheapMoney =&gt; consumerBorrowing =&gt; highDemand =&gt; inflation
```

How do we represent and perform inference with complex chains of this sort?

Answer: **Bayesian networks**

Probabilistic representation

Let's say we want to work out if it's going to be humid tomorrow.

We have a tangle of relationships to take into account. Sunshine increases the humidity but so does rain and temperature, both of which are themselves affected by sunshine.

Probabilistic representation

Let's say we want to work out if it's going to be humid tomorrow.

We have a tangle of relationships to take into account. Sunshine increases the humidity but so does rain and temperature, both of which are themselves affected by sunshine.

- Represent each element of the domain as a *variable* which takes certain values (e.g., sun=yes, sun=no).

Probabilistic representation

Let's say we want to work out if it's going to be humid tomorrow.

We have a tangle of relationships to take into account. Sunshine increases the humidity but so does rain and temperature, both of which are themselves affected by sunshine.

- Represent each element of the domain as a *variable* which takes certain values (e.g., sun=yes, sun=no).
- Represent the relationships between variables in terms of conditional probabilities, e.g., probabilities like P(temp=high | sun=yes) = 0.8, P(temp=high | sun=no) = 0.2, P(rain=yes | temp=low) = 0.6 etc.

Probabilistic representation

Let's say we want to work out if it's going to be humid tomorrow.

We have a tangle of relationships to take into account. Sunshine increases the humidity but so does rain and temperature, both of which are themselves affected by sunshine.

- Represent each element of the domain as a *variable* which takes certain values (e.g., sun=yes, sun=no).
- Represent the relationships between variables in terms of conditional probabilities, e.g., probabilities like P(temp=high|sun=yes) = 0.8, P(temp=high|sun=no) = 0.2, P(rain=yes|temp=low) = 0.6 etc.

Apply the laws of probability

To find out if it's going to be humid, start with some variables that we know the value of and work forwards, establishing the probability distribution on one variable by taking into account all its conditional probabilities and all the (distributions on) all the variables which those probabilities depend on.

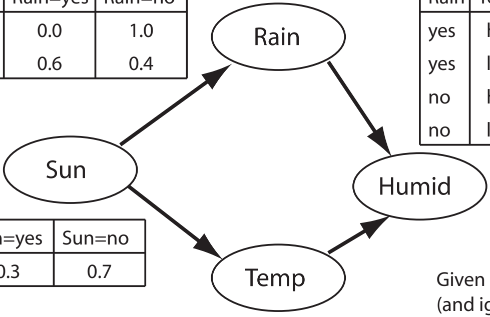
This is also known as forwards propagation.

Bayesian network

Draw out the network of conditional relationships and annotate nodes with the CPTs (conditional probability tables).

Sun	Rain=yes	Rain=no
yes	0.0	1.0
no	0.6	0.4

Rain	Temp	Humid=yes	Humid=no
yes	high	0.9	0.1
yes	low	0.4	0.6
no	high	0.5	0.5
no	low	0.3	0.7



Sun=yes	Sun=no
0.3	0.7

Sun	Temp=high	Temp=low
yes	0.8	0.2
no	0.5	0.5

Given information that P(Sun=yes) = 1.0
(and ignoring zero-valued terms)

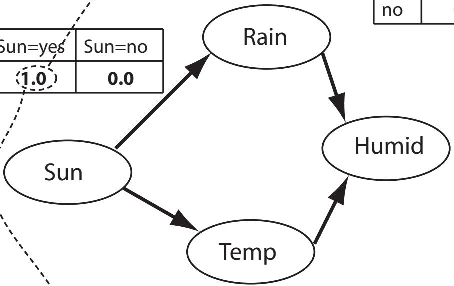
P(Humid=yes) = (0.5 x 0.8 x 1.0) + (0.3 x 0.2 x 1.0) = 0.46

P(Humid=no) = (0.5 x 0.8 x 1.0) + (0.7 x 0.2 x 1.0) = 0.54

Reasoning as propagation

Sun=yes	Sun=no
1.0	**0.0**

Sun	Rain=yes	Rain=no
yes	0.0	1.0
no	0.6	0.4



Rain	Temp	Humid=yes	Humid=no
yes	high	0.9 x 0.0 x 0.8	0.1 x 0.0 x 0.8
yes	low	0.4 x 0.0 x 0.2	0.6 x 0.0 x 0.2
no	high	0.5 x 1.0 x 0.8	0.5 x 1.0 x 0.8
no	low	0.3 x 1.0 x 0.2	0.7 x 1.0 x 0.2
		0.46	0.54

Sun	Temp=high	Temp=low
yes	0.8	0.2
no	0.5	0.5

Given information that P(Sun=yes) = 1.0
(and ignoring zero-valued terms)

P(Humid=yes) = (0.5 x 0.8 x 1.0) + (0.3 x 0.2 x 1.0) = 0.46

P(Humid=no) = (0.5 x 0.8 x 1.0) + (0.7 x 0.2 x 1.0) = 0.54

Terminology and assumptions

In Bayesian networks, any variable Y which has a direct influence on variable X is said to be X's *parent*.

An arrow points from parent Y to X

Variable X is then said to be Y's *child*, while X and all X's children are Y's *descendants*.

When reasoning is done using probability propagation, the assumption is made that two variables are conditionally independent of all *non-descendants* given their parents.

This is another way of saying that variables are only influenced by their parents.

Top-down/forwards propagation

To find the probability of the ith value of variable X, use

$$P(X_i) = \prod_{Y \in CPT(X)} P(X_i | Y) \prod_j P(Y_j)$$

This defines the distribution on X recursively. Each value is obtained by iterating over the combinations of parental values taking the product of the combination's probability and the probability of the value which is *conditional* on the combination.

Termination

Termination is achieved by providing a non-conditional distribution for some root variable.

$$P(X_i) = \quad P(R_i) \quad \text{if } X = R$$

Note that distributions must sum to 1 (so normalization may be required).

How well does it work?

Reasoning using Bayesian nets works perfectly in the sense that probabililies are consistently propagated.

But depending on how variables are related, we can easily end up with very uncertain conclusions.

The key factor which affects performance is the level of uncertainty we have about conditioned variables.

This is the termed **equivocation**.

Equivocation formula

To calculate the equivocation of a conditioned variable relative to a conditioning variable, derive the weighted average of the uncertainties (entropies) of conditional distributions.

$$\sum_j P_j \times -\sum_i P_i \log_2 P_i$$

P_i is the conditioned probability of the ith value of the conditioned variable, and P_j is the probability of the conditioning value.

Equivocation as a weighted uncertainty sum

Equivocation is really just a weighted sum of the uncertainties of the conditional distributions.

Other things being equal, higher equivocation will mean less successful Bayesian reasoning, i.e., less certain conclusions.

.

Summary

Summary

- Bayes rule again

Summary

- Bayes rule again
- Probabilistic representation

Summary

- Bayes rule again
- Probabilistic representation
- Use of Bayesian networks

Summary

- Bayes rule again
- Probabilistic representation
- Use of Bayesian networks
- Reasoning as propagation

Summary

- Bayes rule again
- Probabilistic representation
- Use of Bayesian networks
- Reasoning as propagation
- Top-down propagation

Summary

- Bayes rule again
- Probabilistic representation
- Use of Bayesian networks
- Reasoning as propagation
- Top-down propagation
- Equivocation

Summary

- Bayes rule again
- Probabilistic representation
- Use of Bayesian networks
- Reasoning as propagation
- Top-down propagation
- Equivocation

Questions

Questions

- Let's say the university communicates your degree result to you using either a tick or a cross. What is the level of equivocation?

Questions

- Let's say the university communicates your degree result to you using either a tick or a cross. What is the level of equivocation?

Exercises

Exercises

- Use Bayes' rule to work out P(east | sun) given that P(sun)= 0.3, P(east)=0.4 and P(sun | east)=0.6.

Exercises

- Use Bayes' rule to work out P(east | sun) given that P(sun)= 0.3, P(east)=0.4 and P(sun | east)=0.6.
- Use the frequency interpretation of probability to explain why Bayes rule works.

Exercises

- Use Bayes' rule to work out P(east | sun) given that P(sun)= 0.3, P(east)=0.4 and P(sun | east)=0.6.
- Use the frequency interpretation of probability to explain why Bayes rule works.