331 – Intro to Intelligent Systems
Week09
Bayesian Belief Networks

T.J. Borrelli

Bayesian Network

- Causality – if *this*, then *that*
- But we do not know the whole picture
- Describe events in terms of probabilities
- Allows us to reverse direction of probabilistic statement

Uses of BBNs

- Predict what will happen
- Or to infer causes from observed events
- Allow calculation of conditional probabilities of the nodes in the network, given that the values of some of the nodes have been observed

Types of Inference for BBNs

- Causal Inference (abduction, prediction) - Given that A implies B, and we know the value of A, we can compute the value of B
- Diagnostic Inference (induction) - Given that A implies B, and we know the value of B, we can compute the probability of A

Types of Inference

- Intercausal Inference (“explaining away”) - Given that both A and B imply C, and we know C, we can compute how a change in the probability of A will affect the probability of B, even though A and B were originally independent

Bayesian Belief Network (BBN)

- Directed Acyclic Graph
- Causal relations between variables
- Models cause-and-effect relations
- Allows us to make inferences based on limited knowledge of environment and known probabilities
- Allows us to reverse direction of probabilistic statement

BBN

- Certain Independence assumptions must hold between random variables
- To specify the probability distribution of a Bayesian network
  - Need to know the *a priori* probabilities of all root nodes (nodes with no predecessors)
  - Conditional probabilities of all non-root nodes

Logically

- Only part of the story is considered in the causal (cause-and-effect) relationship between phenomena
- It is $P \rightarrow Q$ not $P \leftrightarrow Q$
- Meaning there could be other potential causes of Q

Directed and Acyclic Graph

- The directions show the causal relations between events – thus somewhat important
- If cycles are present we have a Causal Loop

Bayes' paper

- Bayes described in his famous paper (in 1763) “degree-of-belief” as opposed to classical probability (or fuzzy logic which has partial set membership)
- As new information is added to our defined system of beliefs, the probabilities of events occurring should be updated
- Also known as “Bayesian Inference”

Bayes says a lot of things ...

- However in his original paper, didn't explicitly state the rule attributed to him (conveniently called Bayes' Rule or Bayes' Theorem). So I will state it here:

- $P(A \mid B) = (P(B \mid A) * P(A)) / P(B)$

Explanation of Bayes' Rule

- $P(A \mid B) = (P(B \mid A) * P(A)) / P(B)$
- Means that the Probability of event A occurring given that event B has occurred is equal to the probability of event B occurring given that A has occurred times the independent probability of A divided by the independent probability of B

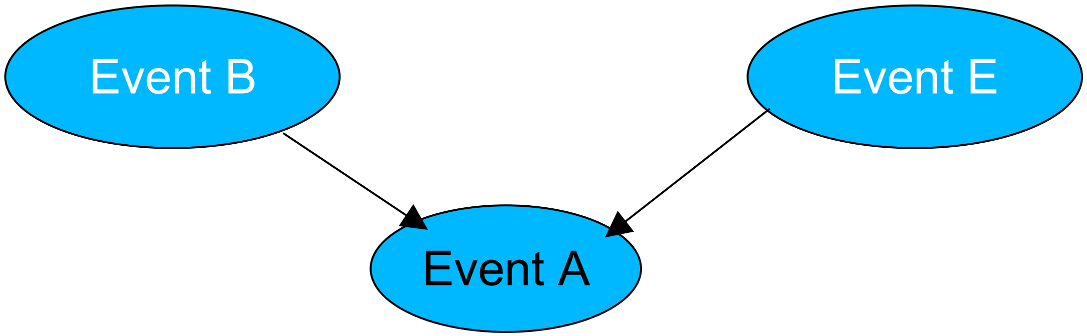
Derivation of Bayes' Rule

- By definition of Conditional Probabilities:
- $P(A \mid B) = P(A \cap B) / P(B)$
- By the same note: $P(B \mid A) = P(B \cap A) / P(A)$
- By the commutative law of probability and set theory: $P(A \cap B) \equiv P(B \cap A)$
- $P(A \cap B) = P(A \mid B) * P(B) = P(B \mid A) * P(A)$
- Thus: $P(A \mid B) = (P(B \mid A) * P(A)) / P(B)$

Definitions

- Bayes' Rule: P(A | B) = (P(B | A) * P(A)) / P (B)
- P(A) is "prior probability" or "marginal probability" of A
- P(A|B) is "conditional probability of A, given B. Also called "posterior probability"
- P(B) is also a "prior" or "marginal" probability and acts as a normalizing constant so event probabilities add to 1

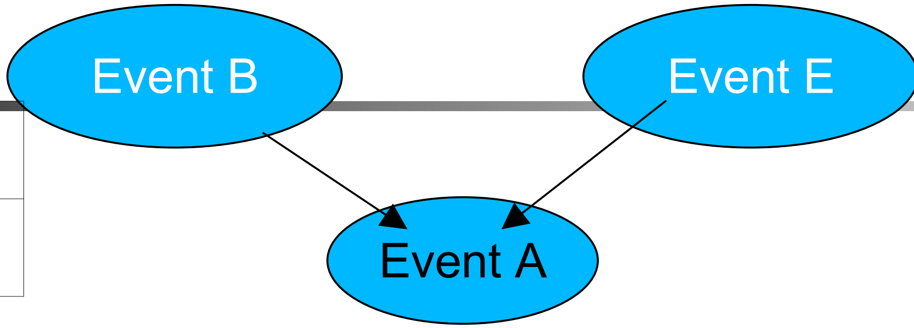
Belief Network



- Event A is “Alarm”
- Event B is “Burglary”
- Event E is “Earthquake”

Alarm Example

t	.001
f	.999



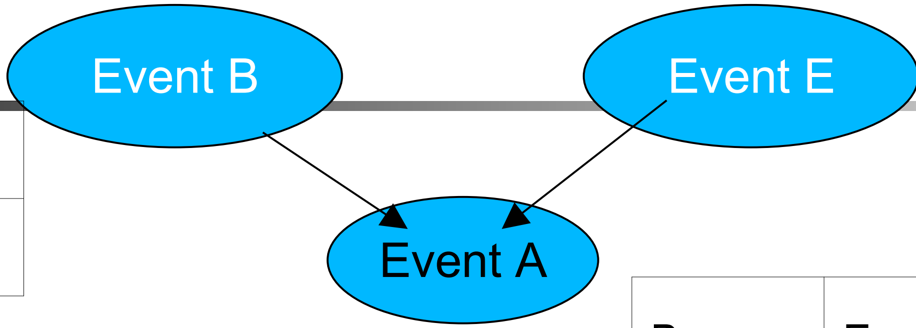
t	.002
f	.998

- Given
  - Initial Truth Tables
  - Causal relationships

Alarm Example-2

T	.002
F	.998

T	.001
F	.999

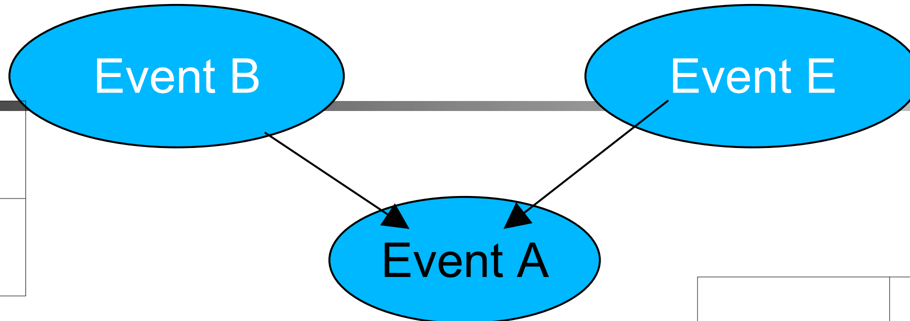


B	E	A
T	T	.95
T	F	.94
F	T	.29
F	F	.001

- Given
  - Initial Truth Tables
  - Causal relationships
  - Conditional Probabilities
- Can calculate conditional probabilities based on new information

Alarm Example-3

T	.002
F	.998



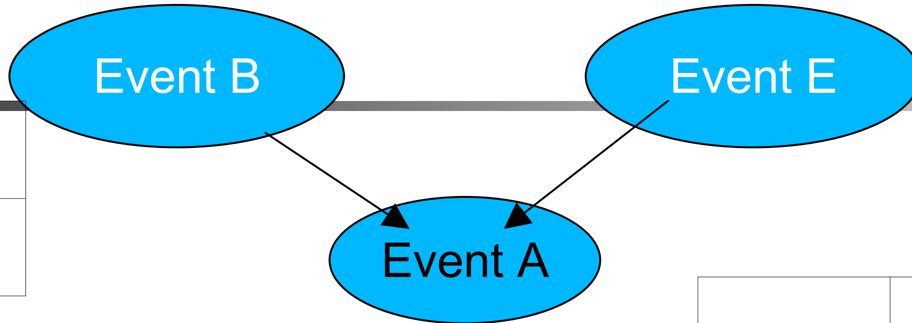
T	.001
F	.999

- $P(A) = P(A|B,E)*P(B)*P(E) + P(A|\neg B,E)*P(\neg B)*P(E) + P(A|B,\neg E)*P(B)*P(\neg E)+ P(A|\neg B,\neg E)*P(\neg B)*P(\neg E)$
  $= .002516$

B	E	A
T	T	.95
T	F	.94
F	T	.29
F	F	.001

Alarm Example-4

T	.002
F	.998



T	.001
F	.999

B	E	A
T	T	.95
T	F	.94
F	T	.29
F	F	.001

- Now Suppose We Observe

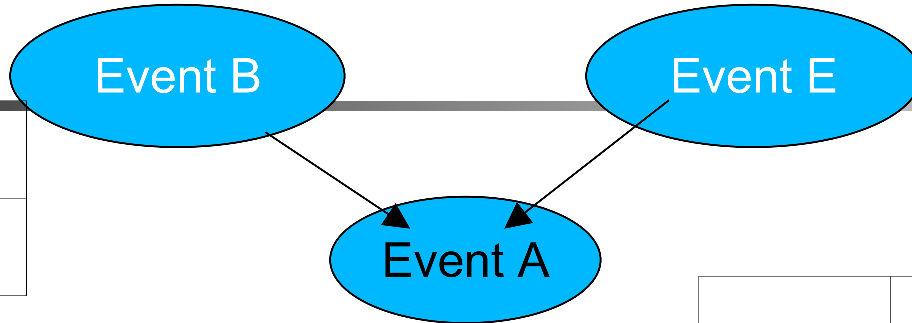
That E is True!

….

What does this change in the previous equation?

Alarm Example-5

T	.002
F	.998



T	.001
F	.999

- E is True!
- What does this change in the previous equation?
- $P(A) = P(A|B,E)*P(B)*P(E) +$

$P(A|\neg B,E)*P(\neg B)*P(E) +$

$P(A|B,\neg E)*P(B)*P(\neg E)+$

$P(A|\neg B,\neg E)*P(\neg B)*P(\sim E) =$

$P(A|B,E)*P(B) + P(A|\neg B,E)*P(\neg B) = .95 * .001 + .29 * .999 = .291$

B	E	A
T	T	.95
T	F	.94
F	T	.29
F	F	.001

Example

- Suppose we want to know the chance that it rained yesterday if we suddenly find out that the lawn is wet
- By Bayes' rule we can calculate this probability from it's inverse: the probability that the lawn would be wet if it *had* rained yesterday

$P(A|B) = (P(B|A) * P(A)) / P(B)$

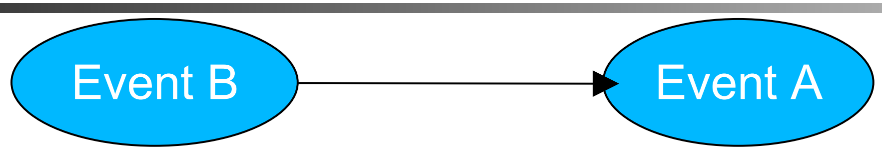
- $P(A|B)$ = “the probability of event A, given that we know B” - for example, the probability that it rained yesterday given that the lawn is wet”
- $P(A)$ = the probability of rain, all other things being equal
- $P(B)$ = the probability of the lawn being wet, all other things being equal

What did we just do?

- We were able to rephrase the probability in terms of the probability of B given A

(P(B|A)) and independent probabilities P(A) and P(B)

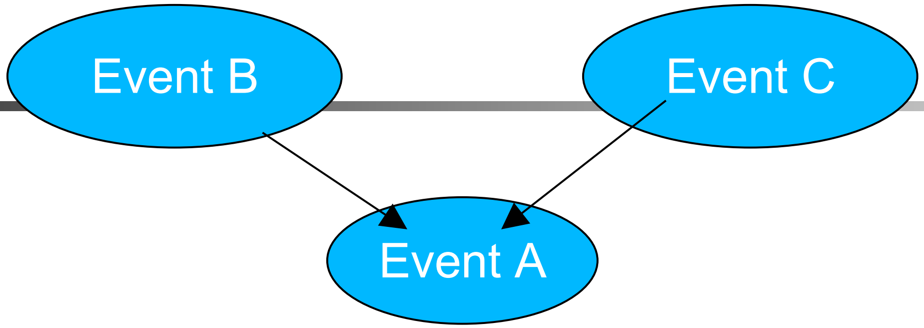
More Probability



- As an extension of Bayes' Theorem:
- Given that event A is dependent upon event B
- $P(A) = P(A|B) * P(B) +$

  $P(A|\neg B) * P(\neg B)$

Similarly



- Event A is dependent upon event B and C
- $P(A) = P(A|B,C) * P(B) * P(C) +$

  $P(A|B, \neg C) * P(B) * P(\neg C) +$

  $P(A|\neg B, C) * P(\neg B) * P(C) +$

  $P(A| \neg B, \neg C) * P(\neg B) * P(\neg C)$

Problems of Inconsistency

- If we look at the following “reasonable-looking” situation:
  - $P(A|B) = .8$
  - $P(B|A) = .2$
  - $P(B) = .6$
- Bayes’ Rule: $P(A|B) = (P(B|A)*P(A))/ P(B)$

Inconsistency

- $P(A|B) = (P(B|A)*P(A))/ P(B)$
- $.8 = .2 * P(A) / .6$
- $.48 = .2 * P(A)$
- Thus, $P(A) = 2.4$
- So, What's Wrong With This?

Skunk in the Woodpile

- You can’t see it, but you know it’s there
- The probability for an event cannot exceed 1
- Problem: failure to keep probabilities consistent
- This is sometimes difficult to see at first
- Need to check for this

Some Uses

- Used in medical problem diagnosis (PATHFINDER) when only limited information and events is known
- Used in map learning and language processing and understanding
- Used in Games to provide human-like reasoning based on incomplete information. (First Person Sneaker)

A Good Paper to Read

- Charniak, E. *Bayesian Networks without Tears.* American Association for Artificial Intelligence. AI Magazine. 1991.