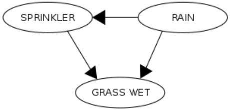
Deduction from Conditional Knowledge on Bayesian Networks with Interval Probability Parameters

Tanmay Goyal - AI20BTECH11021

Bayesian networks

RAIN	SPRINKLER T	SPRINKLER F
F	0.4	0.6
T	0.01	0.99



RAIN T	RAIN F
0.2	0.8

SPRINKLER	RAIN	GRASS WET T	GRASS WET F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

1. Directed Acyclic Graph (DAG)
2. Represents Probabilistic Relationships

Find Probability that it is raining, given grass is wet?

$$\Pr(R = T | G = T) \quad (1)$$

$$= \frac{TTT + TFT}{TTT + TFT + FTT + FFT} \quad (2)$$

Logic Vs Probability

1. Boolean Logic: $b \to a$, "if b, then a", $\Pr(a|b)$
2. Classical Logic: $\Pr(b \to a) = \Pr(a|b) + \Pr(b')\Pr(a'|b) \geq \Pr(a|b)$

$$= \frac{\Pr(a,b) + \Pr(b')\left(\Pr(b) - \Pr(a,b)\right)}{\Pr(b)} \tag{3}$$

$$= \Pr(a,b) + \Pr\left(b'\right) \tag{4}$$

3. Lewis' Theorem: cannot equate the conditional probability $\Pr(A|B)$ with the probability of conditional event, $B \to A$, except for a trivial set of events
4. Conditional Event Algebra (CEA): Unlike Boolean Algebra, allows defining $\Pr(B \to A) = \Pr(A|B)$ over a broad range of conditions
5. CEA allows generalisation of Boolean Algebra to conditional events, while maintaining compatibility of standard and conditional probability

Interval Probability and Higher Order Logic

1. A p-probability interval for θ is an interval $[a, b]$ with $\Pr(a \leq \theta \leq b) = p$, which can be $\sum_{a \leq \theta_i \leq b} \Pr(\theta_i)$ or $\int_a^b f(\theta)\, d\theta$
2. Interval Parameters: a and b
3. Higher- Order Logic: if b_1 then a_1,if b_2 then $a_2 \ldots$ then if c then d
4. $\Pr(c|(a|b)) \neq \Pr(c|(a' \cup b))$
5. The existing definitions and concepts of interval probability do not apply to conditional probability, and cannot be used to draw inferences from higher- order conditionals using BNs.
6. Proposal: Bayesian higher- order probabilistic reasoning approach with interval probability parameters, combining CEA and weak conditional probability.

Conditional Event Algebra

1. Mathematically modelling conditionals as conditional events. Express probability weight of "$b \to a$" in terms of $\Pr(a|b)$
2. Let (Ω, B, P) be a standard probability space such that Ω represents the Sample Space, B represents a definite event, and P a certain probability measure:
    1. $P(\Omega) = 1$
    2. $P(b_1 \cup b_2 \cup \ldots + b_j + \ldots) = P(b_1) + P(b_2) \ldots + P(b_j) + \ldots$ for mutually exclusive b_is
    3. The set of all conditional events is Boolean Algebra with connectives $\wedge, \vee$ and '
    4. Any probability measure P defined on events of Ω can be extended a probability measure P_0 on conditional events such that $P_0(a|b) = P(a|b)$, i.e conditional probability is a true probability measure on conditional events.

Conditional Event Algebra

1. Given normal events a and b, we define an extended probability space (Ω_0, B_0, P_0), such that, $\Omega_0 = \Omega \times \Omega \times \ldots$, $B_0 = B \times B \times \ldots$, and B_0 being a Boolean Algebra extension. Then,

$$(a|b) = ((a \wedge b) \times \Omega_0) \vee (b' \times (a \wedge b) \times \Omega_0) \vee (b' \times b' \times (a \wedge b) \times \Omega_0) \ldots \tag{5}$$

2. Define $f : B \times B \to B_0$ as:

$$f(a, b) = ab \vee (b' \times ab) \vee (b' \times b' \times ab) \ldots \tag{6}$$

$$P_0(f(a, b)) = P_0(ab) + P_0(b' \times ab) + P_0(b' \times b' \times ab) \ldots \tag{7}$$

$$P_0(f(a, b)) = P(ab) \sum P(b')^j \tag{8}$$

$$= P(a|b) \tag{9}$$

Probability logic reasoning on Bayesian Networks

If if b_1 then a_1 and if b_2 then $a_2 \dots$ then if d then c can be expressed as:

$$G = [(a|b)_j; (c|d)] \tag{10}$$

$$A(G) = \cap\{a_j b_j, a_j' b_j, b_j'\} \cap \{cd, c'd, d'\} \tag{11}$$

$$= \{\omega_1, \omega_2, \dots \omega_m\} \tag{12}$$

From (7)

$$f(a_j, b_j, c, d) = \vee(\omega_j) \tag{13}$$

$$\Pr((a|b)_j; (c|d)) = \Pr(A(G)) = P_0(f(a_j, b_j, c, d)) \tag{14}$$

$$= \sum \Pr(w_j) \tag{15}$$

Interval Probability

1. U- universe, S-subset of U. S is called an interval value if $v_1, v_2, v_3 \in U$ and $v_1 \leq v_2 \leq v_3$, and $v_1, v_3 \in S$ implies $v_2 \in S$

2. Let b be a random event, then lower and upper probabilities, written as $L(b)$ and $U(b)$, are given by: (x_i denotes a sample from sample space Ω)
    1. $L(b) = \sum \Pr(x_i)$ for x_i in b
    2. $U(b) = \sum \Pr(x_i)$ for $x_i \cap b \neq \phi$

3. $L(a) \leq \Pr(a) \leq U(a)$

Interval Probability

1. $L(a|b) = inf \frac{\Pr(a \cap b)}{\Pr(b)}$ and $U(a|b) = sup \frac{\Pr(a \cap b)}{\Pr(b)}$ where $pr(b) \neq 0$
2. Bound limited weak conditional probabilities: $BL(a|b)$ and $BU(a|b)$
   1. If $L(a) \leq L(a|b) \leq U(a|b) \leq U(a)$, then $BL(a|b) = max\{EL(a|b), L(a)\}$ and $BU(a|b) = min\{EU(a|b), U(a)\}$
   2. If $L(a|b) \leq L(a) \leq U(a) \leq U(a|b)$, then $BL(a|b) = min\{EL(a|b), L(a)\}$ and $BU(a|b) = max\{EU(a|b), U(a)\}$
3. $BL(a|b) \leq \Pr(a|b) \leq BU(a|b)$

Multiplication rules

1. $L(abc) \approx BL(abc) = L(a)BL(b|a)BL(c|ab)$
2. $U(abc) \approx BU(abc) = U(a)BU(b|a)BU(c|ab)$

Probability logic reasoning on Bayesian Networks

$$BL(\omega_j) \leq \Pr(\omega_j) \leq BU(\omega_j) \tag{16}$$

$$\sum BL(\omega_j) \leq \sum \Pr(\omega_j) \leq \sum BU(\omega_j) \tag{17}$$

$$\sum BL(\omega_j) \leq \Pr((a|b)_j; (c|d)) \leq \sum BU(\omega_j) \tag{18}$$

<u>Logic:</u> Conditional Events $\longrightarrow$ Joint events $\longrightarrow$ Multiplication rule

Example

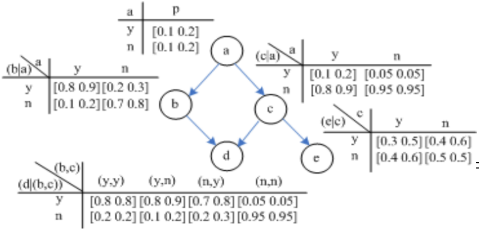
$$\Pr\left(d|(b|a),(c|a)\right) \tag{19}$$

$$G = [(b|a),(c|a);(d|\Omega)] \tag{20}$$

$$A(G) = \{ba, b'a, a'\} \wedge \{ca, c'a, a'\} \wedge \{d\Omega, d'\Omega, \phi\} \tag{21}$$

$abcd$	ω_1
$ab'cd$	ω_2
$abc'd$	ω_3
$ab'c'd$	ω_4
$a'd$	ω_5
$abcd'$	ω_6
$ab'cd'$	ω_7
$abc'd'$	ω_8
$ab'c'd'$	ω_9
$a'd'$	ω_{10}

Example



$$
\begin{aligned}
BL(\omega_1) &= BL(abcd) && (22)\\
&= L(a)BL(b|a)BL(c|a)BL(d|bc) && (23)\\
&= 0.1 \times 0.8 \times 0.1 \times 0.8 && (24)\\
&= 0.0064 && (25)
\end{aligned}
$$

$$
\begin{aligned}
BU(\omega_1) &= BU(abcd) && (26)\\
&= U(a)BU(b|a)BU(c|a)BU(d|bc) && (27)\\
&= 0.2 \times 0.8 \times 0.2 \times 0.9 && (28)\\
&= 0.0288 && (29)
\end{aligned}
$$

Example

		BL(.)	BU(.)
$abcd$	ω_1	0.0064	0.0288
$ab'cd$	ω_2	0.0007	0.0064
$abc'd$	ω_3	0.0512	0.1458
$ab'c'd$	ω_4	0.0004	0.0018
$a'd$	ω_5	0.0217	0.0677
$abcd'$	ω_6	0.0016	0.0072
$ab'cd'$	ω_7	0.0020	0.0024
$abc'd'$	ω_8	0.0064	0.0324
$ab'c'd'$	ω_9	0.0076	0.0342
$a'd'$	ω_{10}	0.0660	0.1588
	$\sum \omega_j$	0.1766	0.4855

The probability of G is estimated to be between [0.1766,0.4855]