Making Inferences in Incomplete Bayesian Networks: A Dempster-Shafer Belief Function Approach

Prakash P. Shenoy

University of Kansas, School of Business, Department of Analytics, Information, and Operations

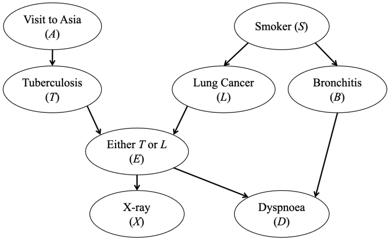
September 16, 2022

Outline

Outline

Introduction

- Suppose we have an incomplete Bayesian network (BN), i.e., a BN with missing priors and conditional distributions. How do we make inferences? i.e., compute marginals of the incompletely specified joint probability distribution for variables of interest?
- An example:



$P_A(a) = \mathbf{?}$, $P_A(\bar{a}) = \mathbf{?}$	$P_{E\|L,T}(l, t, e) = 1$, $P_{E\|L,T}(l, t, \bar{e}) = 0$ $P_{E\|L,T}(l, \bar{t}, e) = 1$, $P_{E\|L,T}(l, \bar{t}, \bar{e}) = 0$
$P_S(s) = \mathbf{?}$, $P_A(\bar{s}) = \mathbf{?}$	$P_{E\|L,T}(\bar{l}, t, e) = 1$, $P_{E\|L,T}(\bar{l}, t, \bar{e}) = 0$ $P_{E\|L,T}(\bar{l}, \bar{t}, e) = 0$, $P_{E\|L,T}(\bar{l}, \bar{t}, \bar{e}) = 1$
$P_{T\|A}(a, t) = 0.05$, $P_{T\|A}(a, \bar{t}) = 0.95$ $P_{T\|A}(\bar{a}, t) = \mathbf{?}$, $P_{T\|A}(\bar{a}, \bar{t}) = \mathbf{?}$	$P_{X\|E}(e, x) = 0.98$, $P_{X\|E}(e\ \bar{x}) = 0.02$ $P_{X\|E}(\bar{e}, x) = \mathbf{?}$, $P_{X\|E}(\bar{e}, \bar{x}) = \mathbf{?}$
$P_{L\|S}(s, l) = 0.10$, $P_{L\|S}(s, \bar{l}) = 0.90$ $P_{L\|S}(\bar{s}, l) = 0.01$, $P_{L\|S}(\bar{s}, \bar{l}) = 0.99$	$P_{D\|E,B}(e, b, d) = \mathbf{?}$, $P_{D\|E,B}(e, b, \bar{d}) = \mathbf{?}$ $P_{D\|E,B}(e, \bar{b}, d) = 0.70$, $P_{D\|E,B}(e, \bar{b}, \bar{d}) = 0.30$
$P_{B\|S}(s, b) = 0.80$, $P_{B\|S}(s, \bar{b}) = 0.20$ $P_{B\|S}(\bar{s}, b) = 0.30$, $P_{B\|S}(\bar{s}, \bar{b}) = 0.70$	$P_{D\|E,B}(\bar{e}, b, d) = 0.80$, $P_{D\|E,B}(\bar{e}, b, \bar{d}) = 0.20$ $P_{D\|E,B}(\bar{e}, \bar{b}, d) = 0.10$, $P_{D\|E,B}(\bar{e}, \bar{b}, \bar{d}) = 0.90$

Introduction

- The Dempster-Shafer (D-S) belief function theory is a generalization of probability theory
- Our proposal: Embed the incomplete BN in a D-S belief function graphical model, omit the missing information, and make inferences from the D-S model using belief function theory.
- We will see that a belief function analysis of an incomplete BN is more satisfying than a corresponding Bayesian analysis using Laplacian equally-likely distributions for missing data.

Outline

Basics of D-S Belief Function Theory

Static: We represent knowledge using either:

- basic probability assignment (BPA) m
- belief function (BF) Bel
- plausibility function (PL) Pl

Dynamic: We make inferences using:

- Marginalization rule
- Dempster's rule of combination
- Inference: Given a set of belief functions (BPA, BF, or PF) representing knowledge of the domain and all evidence, we would like to find the marginals of the joint belief function for some variables of interest.
- The joint belief function is obtained by combining all belief functions using Dempster's rule of combination.

Basics of D-S Belief Function Theory

Notation:

- Suppose Φ denotes a finite set of variables
- For each $X \in \Phi$, Ω_X denotes a finite set of states of X
- Let r, s, t, etc., denote subsets of Φ
- For every non-empty subset $s \subseteq \Phi$,

$$\Omega_s = \prod_{X \in s} \Omega_X$$

  denotes the states of s
- Let 2^{Ω_s} denote the set of all subsets of Ω_s

Basics of D-S Belief Function Theory

Basic Probability Assignment:

- A basic probability assignment (BPA) m for s is a function $m : 2^{\Omega_s} \to [0,1]$ such that:

$$\begin{aligned} m(\emptyset) &= 0, && (1) \\ \sum_{\emptyset \neq \mathsf{a} \in 2^{\Omega_s}} m(\mathsf{a}) &= 1. && (2) \end{aligned}$$

- s is called the domain of m.
- Subsets $\mathsf{a} \in 2^{\Omega_s}$ such that $m(\mathsf{a}) > 0$ are called focal elements of m.

Basics of D-S Belief Function Theory

- Consider prior probability mass function for A: $P_A(a) = 0.01$, $P_A(\bar{a}) = 0.99$.
- P_A can be represented by BPA m_A for A as follows: $m_A(\{a\}) = 0.01$, $m_A(\{\bar{a}\}) = 0.99$
- If all focal elements of m are singleton subsets, then m is called Bayesian. BPA m_A for A is Bayesian.
- If we have only one focal element (with probability 1), then we say m is deterministic
- Consider E (either T or L). $E = e$ iff $T = t$ or $L = l$.
- Consider a BPA m for $\{T, L, E\}$ as follows:

$$m(\{(t, l, e), (t, \bar{l}, e), (\bar{t}, l, e), (\bar{t}, \bar{l}, \bar{e})\}) = 1$$

Then, m is deterministic.

Basics of D-S Belief Function Theory

- Suppose m is a BPA for s. We say m is vacuous if m is deterministic with focal set Ω_s.
- Consider BPA m_A for A such that $m_A(\Omega_A) = 1$. m is vacuous.
- Consider BPA m_A for A such that $m_A(\{a\}) = 0.5$, $m_A(\{\bar{a}\}) = 0.5$. This is a Bayesian, non-vacuous BPA for A.
- In D-S belief function theory, we can distinguish between equally-likely states and vacuous knowledge.

Basics of D-S Belief Function Theory

Belief Functions:

- A belief function Bel_m for s corresponding to BPA m is a function $Bel_m : 2^{\Omega_s} \to [0, 1]$ such that

$$Bel_m(\mathsf{a}) = \sum_{\mathsf{b} \in 2^{\Omega_s} : \mathsf{b} \subseteq \mathsf{a}} m(\mathsf{b}) \tag{3}$$

- $Bel_m(\mathsf{a})$ is the measure of total belief committed to a.
- Example: Suppose m for T is as follows:

$$m(\emptyset) = 0, m(\{t\}) = 0.0005, m(\{\bar{t}\}) = 0.9405, m(\Omega_T) = 0.0590.$$

Then,

$$Bel_m(\emptyset) = 0, Bel_m(\{t\}) = 0.5, Bel_m(\{\bar{t}\}) = 0.1, Bel_m(\Omega_T) = 1.$$

Basics of D-S Belief Function Theory

Plausibility Function:

- A plausibility function Pl_m for s corresponding to BPA m is a function $Pl_m : 2^{\Omega_s} \to [0, 1]$ such that

$$Pl_m(\mathsf{a}) = \sum_{\mathsf{b} \in 2^{\Omega_s} : \mathsf{b} \cap \mathsf{a} \neq \emptyset} m(\mathsf{b}) \tag{4}$$

- $Pl_m(\mathsf{a})$ is the measure of the extent to which one finds proposition a plausible.
- Example: Suppose m for T is as follows:

$$m(\emptyset) = 0, m(\{t\}) = 0.0005, m(\{\bar{t}\}) = 0.9405, m(\Omega_T) = 0.0590.$$

  Then,

$$Pl_m(\emptyset) = 0, Pl_m(\{t\}) = 0.0595, Pl_m(\{\bar{t}\}) = 0.9995, Pl(\Omega_T) = 1.$$

- In general, $0 \leq Bel(\mathsf{a}) \leq Pl(\mathsf{a}) \leq 1$, and $Pl(\mathsf{a}) = 1 - Bel(\Omega_s \setminus \mathsf{a})$.

Basics of D-S Belief Function Theory

- BPA m, BF Bel_m, and PF Pl_m have the same information.
- Given any one, we can recover the others.

Basics of D-S Belief Function Theory

Marginalization rule

- Marginalization in belief function theory is addition.
- Projection of states: If $\mathbf{x} \in \Omega_s$, and $X \in s$, then $\mathbf{x}^{\downarrow s \setminus \{X\}}$ is the state of $s \setminus \{X\}$ obtained from $\mathbf{x}$ by dropping the state of X.
- Projection of subset of states: If $\mathsf{a} \in 2^{\Omega_s}$, then $\mathsf{a}^{\downarrow s \setminus \{X\}}$ is

$$\mathsf{a}^{\downarrow s \setminus \{X\}} = \{\mathbf{x}^{\downarrow s \setminus \{X\}} \ : \ \mathbf{x} \in \mathsf{a}\}$$

Definition (Marginalization rule)

If m is a bpa for s, and $X \in s$, then $m^{\downarrow s \setminus \{X\}}$ is a bpa for $s \setminus \{X\}$ defined as follows:

$$m^{\downarrow s \setminus \{X\}}(\mathsf{a}) = \sum_{\mathsf{b} \in 2^{\Omega_s} \,:\, \mathsf{b}^{\downarrow s \setminus \{X\}} = \mathsf{a}} m(\mathsf{b})$$

for all $\mathsf{a} \in 2^{s \setminus \{X\}}$.

Basics of D-S Belief Function Theory

- Consider the deterministic BPA m for $\{T, L, E\}$ as follows:

$$m(\{(t, l, e), (t, \bar{l}, e), (\bar{t}, l, e), (\bar{t}, \bar{l}, \bar{e})\}) = 1$$

  Then, $m^{\downarrow\{T,L\}}$ is the BPA for $\{T, L\}$ given by

$$m^{\downarrow\{T,L\}}(\{(t, l), (t, \bar{l}), (\bar{t}, l), (\bar{t}, \bar{l})\}) = m^{\downarrow\{T,L\}}(\Omega_{\{T,L\}}) = 1.$$

- Notice that $m^{\downarrow\{T,L\}}$ is the vacuous BPA for $\{T, L\}$.

Basics of D-S Belief Function Theory

- The definition of marginalization of BPA functions has the following properties:
  - (Domain) If m is a BPA for s, and $X \in s$, then $m^{\downarrow s \setminus \{X\}}$ is a BPA for $s \setminus \{X\}$.
  - (Order does not matter) If m is a BPA for s, $X, Y \in s$, then

$$(m^{\downarrow s \setminus \{X\}})^{\downarrow s \setminus \{X,Y\}} = (m^{\downarrow s \setminus \{Y\}})^{\downarrow s \setminus \{X,Y\}}.$$

Basics of D-S Belief Function Theory

- The combination rule in the D-S theory of belief functions is Dempster's rule, which Dempster called the "product-intersection" rule.
- The product of the BPA masses is assigned to the intersection of the focal elements, any mass assigned to the empty set is discarded, and the remaining masses re-normalized.

Definition (Dempster's combination rule)

Suppose m_1 is a BPA for s_1, m_2 is a BPA for s_2, and m_1 and m_2 are *distinct*. Then, $m_1 \oplus m_2$ is a BPA for $s_1 \cup s_2$ such that for all $\mathsf{a} \in 2^{\Omega_{s_1 \cup s_2}}$

$$(m_1 \oplus m_2)(\mathsf{a}) = K^{-1} \sum_{\mathsf{a}_1 \in 2^{\Omega_{s_1}}, \mathsf{a}_2 \in 2^{\Omega_{s_2}} : (\mathsf{a}_1 \times \Omega_{s_2 \setminus s_1}) \cap (\mathsf{a}_2 \times \Omega_{s_1 \setminus s_2}) = \mathsf{a}} m_1(\mathsf{a}_1)\, m_2(\mathsf{a}_2), \tag{5}$$

where K is a normalization constant given by

$$K = \sum_{\mathsf{a}_1 \in 2^{\Omega_{s_1}}, \mathsf{a}_2 \in 2^{\Omega_{s_2}} : (\mathsf{a}_1 \times \Omega_{s_2 \setminus s_1}) \cap (\mathsf{a}_2 \times \Omega_{s_1 \setminus s_2}) \neq \emptyset} m_1(\mathsf{a}_1)\, m_2(\mathsf{a}_2). \tag{6}$$

Basics of D-S Belief Function Theory

- In general $m \oplus m \neq m$.
- Thus, in combining, e.g., BPAs m_1 and m_2 by Dempster's rule, it is important that m_1 and m_2 are distinct pieces of evidence (to avoid double-counting of uncertain knowledge).
- Dempster's rule satisfies the following properties:
  - (Domain) If m_1 is a BPA for s_1 and m_2 is a BPA for s_2, then $m_1 \oplus m_2$ is a BPA for $s_1 \cup s_2$.
  - (Commutative) $m_1 \oplus m_2 = m_2 \oplus m_1$
  - (Associative) $m_1 \oplus (m_2 \oplus m_3) = (m_1 \oplus m_2) \oplus m_3$
- (Local computation) Marginalization and Dempster's rules satisfy the following property: If m_1 is a BPA for s_1, m_2 is a BPA for s_2, $X \in s_1$, and $X \notin s_2$, then:

$$(m_1 \oplus m_2)^{\downarrow (s_1 \cup s_2) \setminus \{X\}} = m_1^{\downarrow s_1 \setminus \{X\}} \oplus m_2$$

Basics of D-S Belief Function Theory

In directed graphical models, we have conditional BPAs.

Definition (Conditional BPA)

Suppose r and s are disjoint subsets of variables, and $m_{s|r}$ is a BPA for $r \cup s$. We say $m_{s|r}$ is a conditional BPA for s given r if and only if

1. $(m_{s|r})^{\downarrow r}$ is a vacuous BPA for r, and
2. for any BPA m_r for r, m_r and $m_{s|r}$ are distinct. Thus, $m_r \oplus m_{s|r}$ is a BPA for $r \cup s$.

We say s is the head of the conditional $m_{s|r}$, and r the tail.

Basics of D-S Belief Function Theory

How do we obtain conditional BPAs? One way is to use Smets' conditional embedding.

- Suppose we have some knowledge about Y in the context $X = x$ encoded as BPA m_Y for Y.
- The knowledge of Y encoded in BPA m_Y for Y is valid only in the case $X = x$.
- Using Smets' conditional embedding, we convert the BPA m_Y for Y to a BPA $m_{Y|x}$ for $\{X, Y\}$ as follows:

Definition (Smets' conditional embedding)

$m_{Y|x}$ for $\{X, Y\}$ is defined as follows:

$$m_{Y|x}((\{x\} \times \mathsf{b}) \cup ((\Omega_X \setminus \{x\}) \times \Omega_Y)) = m_Y(\mathsf{b})$$

for each focal element b of m_Y.

Basics of D-S Belief Function Theory

An Example

- Suppose X and Y are variables with $\Omega_X = \{x, \bar{x}\}$ and $\Omega_Y = \{y, \bar{y}\}$.
- If $X = x$, assume m_Y is as follows:

$$\begin{aligned} m_Y(\{y\}) &= 0.8, \\ m_Y(\Omega_Y) &= 0.2. \end{aligned}$$

- Then $m_{Y|x}$ is a BPA for $\{X, Y\}$ as follows:

$$\begin{aligned} m_{Y|x}(\{(x, y), (\bar{x}, y), (\bar{x}, \bar{y})\}) &= 0.8, \\ m_{Y|x}(\Omega_{X,Y}) &= 0.2. \end{aligned}$$

Basics of D-S Belief Function Theory

- The BPA $m_{Y|x}$ for $\{X, Y\}$ obtained from m_Y by Smets' conditional embedding has the following property.
  1. $(m_{Y|x})^{\downarrow X}$ is a vacuous bpa for X. Thus, it is a conditional BPA for Y given X.
  2. Suppose $m_{X=x}$ is a bpa for X as follows: $m_{X=x}(\{x\}) = 1$. Then,

$$(m_{Y|x} \oplus m_{X=x})^{\downarrow Y} = m_Y.$$

Basics of D-S Belief Function Theory

- Consider $m_{Y|x}$:

$$\begin{aligned} m_{Y|x}(\{(x,y),(\bar{x},y),(\bar{x},\bar{y})\}) &= 0.8, \\ m_{Y|x}(\Omega_{X,Y}) &= 0.2. \end{aligned}$$

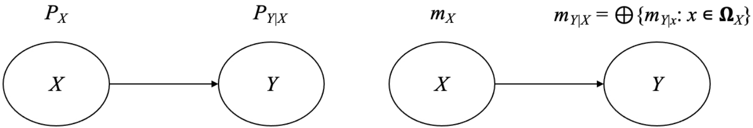
- It is clear that $m_{Y|x}^{\downarrow X}$ is vacuous for X.
- Consider $m_{X=x} \oplus m_{Y|x}$:

$m_{X=x} \oplus m_{Y\|x}$	$\{(x,y),(\bar{x},y),(\bar{x},\bar{y})\}$ 0.8	$\Omega_{\{X,Y\}}$ 0.2
$\{(x,y),(x,\bar{y})\}$ 1	$\{(x,y)\}$ 0.8	$\{(x,y),(x,\bar{y})\}$ 0.2

- Thus, $(m_{X=x} \oplus m_{Y|x})(\{(x,y)\}) = 0.8$, $(m_{X=x} \oplus m_{Y|x})(\{(x,y),(x,\bar{y})\}) = 0.2$.
- Thus, $(m_{X=x} \oplus m_{Y|x})^{\downarrow Y} = m_Y$.

Outline

Bayesian Networks as Belief Function Models

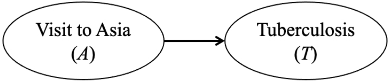


Theorem

Consider a two-variable BN with variables X and Y with state spaces Ω_X and Ω_Y, respectively. Let P_X and $P_{Y|X}$ denote the prior probability mass function (PMF) for X and CPT for $Y|X$, respectively. Let $P_{X,Y}$ denote the joint PMF of (X, Y) such that $P_{X,Y}(x, y) = P_X(x)\, P_{Y|X}(x, y)$ for all $(x, y) \in \Omega_{X,Y}$. Let BPA m_X for X denote the Bayesian BPA for X corresponding to PMF P_X. Let m_Y denote the Bayesian BPA for Y corresponding to $P_{Y|X}(x, y)$ when $X = x$. Let BPA $m_{Y|x}$ for (X, Y) denote the conditional BPA for Y given $X = x$ obtained from m_Y by Smets' conditional embedding. Let $m_{Y|X}$ denote $\oplus\{m_{Y|x} : x \in \Omega_X\}$ Then, $m_X \oplus m_{Y|X}$ is the Bayesian BPA for (X, Y) corresponding to PMF $P_{X,Y}$.

Bayesian Networks as Belief Function Models

Bayesian analysis:



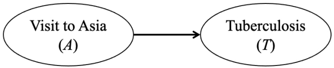
$P_A(a) = 0.01$
$P_A(\bar{a}) = 0.99$

$P_{T|A}(a, t) = 0.05$
$P_{T|A}(a, \bar{t}) = 0.95$
$P_{T|A}(\bar{a}, t) = 0.01$
$P_{T|A}(\bar{a}, \bar{t}) = 0.99$

$\Omega_{\{A,T\}}$	$(P_A)^{\uparrow\{A,T\}}$	$P_{\{T\|A\}}$	$P_{\{A,T\}}$
(a, t)	0.01	0.05	0.0005
$(a, \bar{t})$	0.01	0.95	0.0095
$(\bar{a}, t)$	0.99	0.01	0.0099
$(\bar{a}, \bar{t})$	0.99	0.99	0.9801

Bayesian Networks as Belief Function Models

Belief Function Analysis:



$m_A(\{a\}) = 0.01$
$m_A(\{\bar{a}\}) = 0.99$

$m_{T|a}(\{(a, t), (\bar{a}, t), (\bar{a}, \bar{t})\}) = 0.05$
$m_{T|a}(\{(a, \bar{t}), (\bar{a}, t), (\bar{a}, \bar{t})\}) = 0.95$
$m_{T|\bar{a}}(\{(a, t), (a, \bar{t}), (\bar{a}, t)\}) = 0.01$
$m_{T|\bar{a}}(\{(a, t), (a, \bar{t}), (\bar{a}, \bar{t})\}) = 0.99$

- There are 3 BPAs, m_A, $m_{T|a}$, and $m_{T|\bar{a}}$.
- First, we combine the two conditionals as follows:

$m_{T\|A} =$ $m_{T\|a} \oplus m_{T\|\bar{a}}$	$\{(a,t),(a,\bar{t}),(\bar{a},t)\}$ 0.01	$\{(a,t),(a,\bar{t}),(\bar{a},\bar{t})\}$ 0.99
$\{(a,t),(\bar{a},t),(\bar{a},\bar{t})\}$ 0.05	$\{(a,t),(\bar{a},t)\}$ 0.0005	$\{(a,t),(\bar{a},\bar{t})\}$ 0.0495
$\{(a,\bar{t}),(\bar{a},t),(\bar{a},\bar{t})\}$ 0.95	$\{(a,\bar{t}),(\bar{a},t)\}$ 0.0095	$\{(a,\bar{t}),(\bar{a},\bar{t})\}$ 0.9405

Bayesian Networks as Belief Function Models

Belief Function Analysis:(continued)

$m_{A,T} =$ $m_{T\|A} \oplus (m_A)^{\uparrow\{A,T\}}$	$\{(a,t),(a,\bar{t})\}$ 0.01	$\{(\bar{a},t),(\bar{a},\bar{t})\}$ 0.99
$\{(a,t),(\bar{a},t)\}$ 0.0005	$\{(a,t)\}$ 0.000005	$\{(\bar{a},t)\}$ 0.000495
$\{(a,t),(\bar{a},\bar{t})\}$ 0.0495	$\{(a,t)\}$ 0.000495	$\{(\bar{a},\bar{t})\}$ 0.049005
$\{(a,\bar{t}),(\bar{a},t)\}$ 0.0095	$\{(a,\bar{t})\}$ 0.000095	$\{(\bar{a},t)\}$ 0.009405
$\{(a,\bar{t}),(\bar{a},\bar{t})\}$ 0.9405	$\{(a,\bar{t})\}$ 0.009405	$\{(\bar{a},\bar{t})\}$ 0.931095

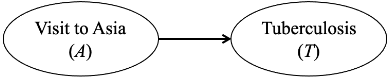
Thus,

$$\begin{aligned}
m_{A,T}(\{(a,t)\}) &= 0.000005 + 0.000495 = 0.0005,\\
m_{A,T}(\{(a,\bar{t})\}) &= 0.000095 + 0.009405 = 0.0095,\\
m_{A,T}(\{(\bar{a},t)\}) &= 0.000495 + 0.009405 = 0.0099,\\
m_{A,T}(\{(\bar{a},\bar{t})\}) &= 0.049005 + 0.931095 = 0.9801.
\end{aligned}$$

Outline

Incomplete BNs

- Suppose we have a BN where we are missing some information.
- For example, consider the Asia-Tuberculosis example as follows:



$$P_A(a) = ?$$
$$P_A(\bar{a}) = ?$$

$$P_{T|A}(a, t) = 0.05$$
$$P_{T|A}(a, \bar{t}) = 0.95$$
$$P_{T|A}(\bar{a}, t) = 0.01$$
$$P_{T|A}(\bar{a}, \bar{t}) = 0.99$$

- What can we do? Replace missing information with Laplacian equally-likely probabilities.

Incomplete BNs

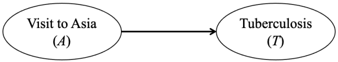
Bayesian analysis with incomplete information:

$\Omega_{\{A,T\}}$	$(P_A)^{\uparrow\{A,T\}}$	$P_{\{T\|A\}}$	$P_{\{A,T\}}$	P_T
(a,t)	0.50	0.05	0.025	0.030
$(a,\bar{t})$	0.50	0.95	0.475	0.970
$(\bar{a},t)$	0.50	0.01	0.005	
$(\bar{a},\bar{t})$	0.50	0.99	0.495	

- Bayesian analysis cannot distinguish between unknown prior and known equally-likely prior for A.

Incomplete BNs

Belief Function Analysis:



$m_A(\{a, \bar{a}\}) = 1$

$m_{T|a}(\{(a,t),(\bar{a},t),(\bar{a},\bar{t})\}) = 0.05$
$m_{T|a}(\{(a,\bar{t}),(\bar{a},t),(\bar{a},\bar{t})\}) = 0.95$
$m_{T|\bar{a}}(\{(a,t),(a,\bar{t}),(\bar{a},t)\}) = 0.01$
$m_{T|\bar{a}}\big(\{(a,t),(a,\bar{t}),\}(\bar{a},\bar{t})\big) = 0.99$

- The joint BPA is $m_{A,T} = m_{T|A} = m_{T|a} \oplus m_{T|\bar{a}}$

$$\begin{aligned} m_{A,T}(\{(a,t),(\bar{a},t)\}) &= 0.0005 \\ m_{A,T}(\{(a,t),(\bar{a},\bar{t})\}) &= 0.0495 \\ m_{A,T}(\{(a,\bar{t}),(\bar{a},t)\}) &= 0.0095 \\ m_{A,T}(\{(a,\bar{t}),(\bar{a},\bar{t})\}) &= 0.9405 \end{aligned}$$

- $m_{A,T}$ is not Bayesian.

Incomplete BNs

Belief Function Analysis:(continued)

- The marginal of the joint for T is

$$\begin{aligned} m_T(\{t\}) &= 0.0005 \\ m_T(\{\bar{t}\}) &= 0.9405 \\ m_T(\{t, \bar{t}\}) &= 0.0495 + 0.0095 = 0.0590 \end{aligned}$$

- $Bel_T(\{t\}) = 0.0005$, $Pl(\{t\}) = 0.0005 + 0.0590 = 0.0595$
- $Bel_T(\{\bar{t}\}) = 0.9405$, $Pl(\{\bar{t}\}) = 0.9405 + 0.0590 = 0.9995$
- Thus, $0.0005 \leq P_T(t) \leq 0.0595$, $0.9405 \leq P_T(\bar{t}) \leq 0.9995$.
- The belief function analysis seems more satisfying than the Bayesian analysis

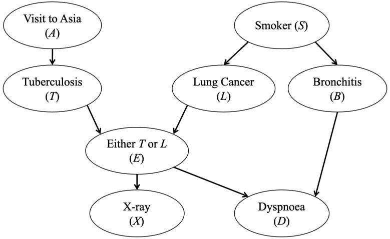
Outline

Belief Function Machine

- A collection of routines to build belief function models and compute marginals using local computation
- Implemented in MATLAB
- Written in 2002 by Phan Giang under the supervision of Philippe Smets, Thierry Denoeux, and I. Further developed by Sushila Shenoy in 2003.
- Features:
    - Belief function model is input as a text file using a language called UIL (unified input language)
    - Can solve "large" models
    - Solve means finding marginal of the joint for variables of interest
- Can be downloaded for free from http://pshenoy.faculty.ku.edu/Papers/BFM072503.zip

Belief Function Machine

- Consider the complete Chest Clinic example from [Lauritzen-Spiegelhalter 1988]



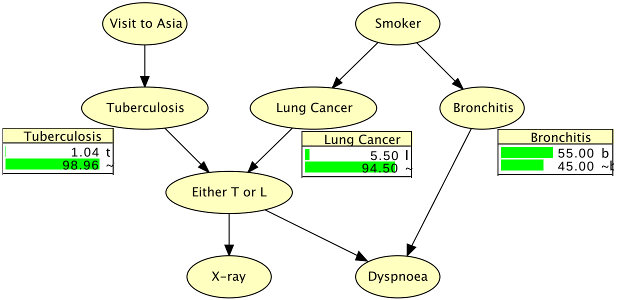
$P_A(a) = 0.01$, $P_A(\bar{a}) = 0.99$	$P_{E\|L,T}(l, t, e) = 1$, $P_{E\|L,T}(l, t, \bar{e}) = 0$ $P_{E\|L,T}(l, \bar{t}, e) = 1$, $P_{E\|L,T}(l, \bar{t}, \bar{e}) = 0$
$P_S(s) = 0.50$, $P_A(\bar{s}) = 0.50$	$P_{E\|L,T}(\bar{l}, t, e) = 1$, $P_{E\|L,T}(\bar{l}, t, \bar{e}) = 0$ $P_{E\|L,T}(\bar{l}, \bar{t}, e) = 0$, $P_{E\|L,T}(\bar{l}, \bar{t}, \bar{e}) = 1$
$P_{T\|A}(a, t) = 0.05$, $P_{T\|A}(a, \bar{t}) = 0.95$ $P_{T\|A}(\bar{a}, t) = 0.01$, $P_{T\|A}(\bar{a}, \bar{t}) = 0.99$	$P_{X\|E}(e, x) = 0.98$, $P_{X\|E}(e\ \bar{x}) = 0.02$ $P_{X\|E}(\bar{e}, x) = 0.05$, $P_{X\|E}(\bar{e}, \bar{x}) = 0.95$
$P_{L\|S}(s, l) = 0.10$, $P_{L\|S}(s, \bar{l}) = 0.90$ $P_{L\|S}(\bar{s}, l) = 0.01$, $P_{L\|S}(\bar{s}, \bar{l}) = 0.99$	$P_{D\|E,B}(e, b, d) = 0.90$, $P_{D\|E,B}(e, b, \bar{d}) = 0.10$ $P_{D\|E,B}(e, \bar{b}, d) = 0.70$, $P_{D\|E,B}\ (e, \bar{b}, \bar{d}) = 0.30$
$P_{B\|S}(s, b) = 0.80$, $P_{B\|S}(s, \bar{b}) = 0.20$ $P_{B\|S}(\bar{s}, b) = 0.30$, $P_{B\|S}(\bar{s}, \bar{b}) = 0.70$	$P_{D\|E,B}(\bar{e}, b, d) = 0.80$, $P_{D\|E,B}(\bar{e}, b, \bar{d}) = 0.20$ $P_{D\|E,B}\ (\bar{e}, \bar{b}, d) = 0.10$, $P_{D\|E,B}(\bar{e}, \bar{b}, \bar{d}) = 0.90$

Complete Chest Clinic

Table: The marginal BPAs m_T for T, m_L for L, and m_B for B, in the complete chest clinic example.

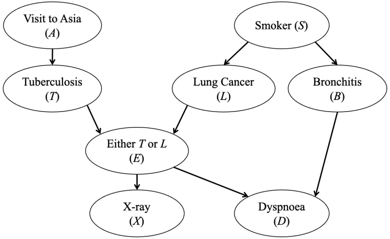
2^{Ω_T}	m_T	2^{Ω_L}	m_L	2^{Ω_B}	m_B
$\{t\}$	0.0104	$\{l\}$	0.0550	$\{b\}$	0.5500
$\{\bar{t}\}$	0.9896	$\{\bar{l}\}$	0.9450	$\{\bar{b}\}$	0.4500

Figure: The results from a Bayesian propagation for the complete chest clinic example.



Incomplete Chest Clinic

Consider the incomplete Chest Clinic example where we are missing five pieces of information:



$P_A(a) = \textbf{?}$, $P_A(\bar{a}) = \textbf{?}$

$P_S(s) = \textbf{?}$, $P_A(\bar{s}) = \textbf{?}$

$P_{T|A}(a, t) = 0.05$, $P_{T|A}(a, \bar{t}) = 0.95$
$P_{T|A}(\bar{a}, t) = \textbf{?}$, $P_{T|A}(\bar{a}, \bar{t}) = \textbf{?}$

$P_{L|S}(s, l) = 0.10$, $P_{L|S}(s, \bar{l}) = 0.90$
$P_{L|S}(\bar{s}, l) = 0.01$, $P_{L|S}(\bar{s}, \bar{l}) = 0.99$

$P_{B|S}(s, b) = 0.80$, $P_{B|S}(s, \bar{b}) = 0.20$
$P_{B|S}(\bar{s}, b) = 0.30$, $P_{B|S}(\bar{s}, \bar{b}) = 0.70$

$P_{E|L,T}(l, t, e) = 1$, $P_{E|L,T}(l, t, \bar{e}) = 0$
$P_{E|L,T}(l, \bar{t}, e) = 1$, $P_{E|L,T}(l, \bar{t}, \bar{e}) = 0$
$P_{E|L,T}(\bar{l}, t, e) = 1$, $P_{E|L,T}(\bar{l}, t, \bar{e}) = 0$
$P_{E|L,T}(\bar{l}, \bar{t}, e) = 0$, $P_{E|L,T}(\bar{l}, \bar{t}, \bar{e}) = 1$

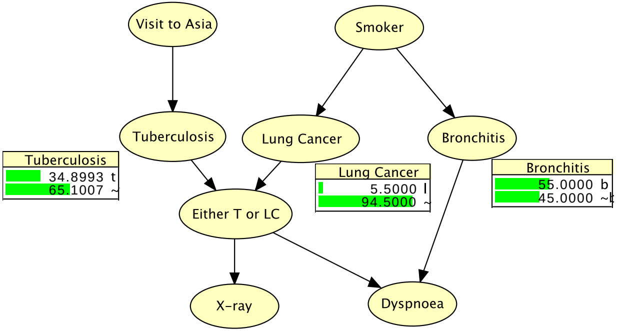
$P_{X|E}(e, x) = 0.98$, $P_{X|E}(e\ \bar{x}) = 0.02$
$P_{X|E}(\bar{e}, x) = \textbf{?}$, $P_{X|E}(\bar{e}, \bar{x}) = \textbf{?}$

$P_{D|E,B}(e, b, d) = \textbf{?}$, $P_{D|E,B}(e, b, \bar{d}) = \textbf{?}$
$P_{D|E,B}(e, \bar{b}, d) = 0.70$, $P_{D|E,B}(e, \bar{b}, \bar{d}) = 0.30$
$P_{D|E,B}(\bar{e}, b, d) = 0.80$, $P_{D|E,B}(\bar{e}, b, \bar{d}) = 0.20$
$P_{D|E,B}(\bar{e}, \bar{b}, d) = 0.10$, $P_{D|E,B}(\bar{e}, \bar{b}, \bar{d}) = 0.90$

Incomplete Chest Clinic

- To do a Bayesian analysis, we have to replace missing information with Laplacian equally-likely distributions.
- Results are as follows:

Figure: The results from a Bayesian propagation for the incomplete chest clinic example.



Incomplete Chest Clinic

- For a D-S belief function analysis, we can either replace the missing information with vacuous belief functions or omit the missing information.

Table: The marginal BPAs m_T for T, m_L for L, and m_B for B in the incomplete chest clinic example.

2^{Ω_T}	m_T	2^{Ω_L}	m_L	2^{Ω_B}	m_B
$\{t\}$	0	$\{l\}$	0.001	$\{b\}$	0.18
$\{\sim t\}$	0	$\{\sim l\}$	0.891	$\{\sim b\}$	0.28
$\{t, \sim t\}$	1	$\{l, \sim l\}$	0.108	$\{b, \sim b\}$	0.54

- Thus,

$$0 \leq P_T(t) \leq 1, 0.001 \leq P_L(l) \leq 0.109, 0.18 \leq P_B(b) \leq 0.72$$

$$0 \leq P_T(\bar{t}) \leq 1, 0.891 \leq P_L(\bar{l}) \leq 0.999, 0.28 \leq P_B(\bar{b}) \leq 0.82$$

Outline

Summary & Conclusions

- Given an incomplete BN (with missing parameters), one solution is to embed the incomplete BN in a corresponding D-S belief function model and then make inferences using the D-S belief function theory.
- D-S belief function theory is a generalization of Bayesian probability theory.
- For complete BNs, we get the same results (marginals of variables of interest).
- For incomplete BNs, we get upper and lower bounds on probabilities of interest.
- This is a more satisfying solution than replacing missing information with Laplacian equally-probable probabilities and getting the point estimates of probabilities of interest from a Bayesian analysis.
- Belief Function Machine software can be used to solve large graphical belief-function models.

Questions

Questions?