Graphical Models for Belief Functions

Prakash P. Shenoy

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

October 30, 2019

Outline

1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Outline

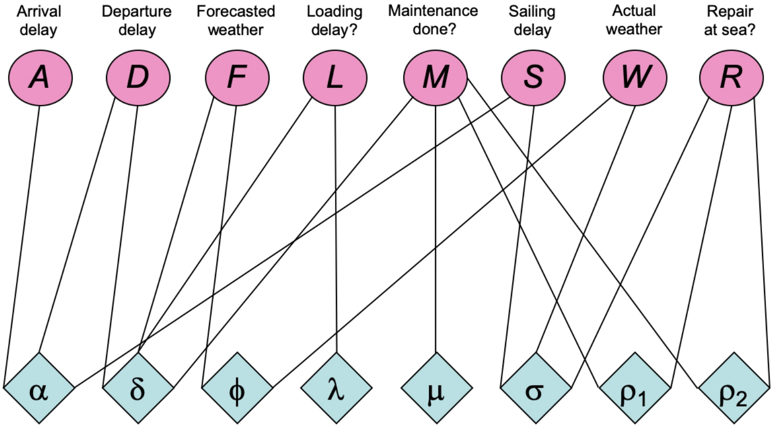
1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Valuation-based Systems

- A valuation-based System (VBS) is an abstract framework for representation of, and reasoning with, knowledge.
- It has two parts. A static part that is concerned with representation of knowledge, and a dynamic part that is concerned with reasoning with knowledge.
- The static part consists of:
  - Variables: A finite set Φ of variables $\{X, Y, Z, \ldots\}$. Subsets of Φ will be denoted by $r, s, t, \ldots$.
  - Valuations: A finite set Ψ of valuations $\{\rho, \sigma, \tau, \ldots\}$. Each valuation encodes knowledge about a subset of variables. Thus, we say, ρ is a valuation for r, where $r \subseteq \Phi$.
- A graphical representation of a VBS is called a valuation network.

Valuation-based Systems

- The valuation network for the Captain's problem: A bipartite graph with variables and valuations as nodes. Each valuation is linked to the variables in its domain.



Valuation-based Systems

- The dynamic part consists of several operators:
  - Combination: $\oplus : \Psi \times \Psi \to \Psi$ that enables us to aggregate knowledge.
  - The combination operator has the following properties:
    - (Domain) If ρ is a valuation for r, and σ is a valuation for s, then $\rho \oplus \sigma$ is a valuation for $r \cup s$
    - (Commutativity) $\rho \oplus \sigma = \sigma \oplus \rho$
    - (Associativity) $\rho \oplus (\sigma \oplus \tau) = (\rho \oplus \sigma) \oplus \tau$
  - The sequence in which knowledge is aggregated should make no difference.
  - The combination of all valuations, $\oplus\Psi$, is called the joint valuation.

Valuation-based Systems

- Another operator is marginalization
- Marginalization: $-X : \Psi \to \Psi$ that allows us to coarsen knowledge marginalizing X out of the domain of a valuation.
- Properties of Marginalization
  - (Domain) If ρ is a valuation for r, and $X \in r$, then ρ^{-X} is a valuation for $r \setminus \{X\}$.
  - (Order does not matter) If ρ is a valuation for r, $X, Y \in r$, then $(\rho^{-X})^{-Y} = (\rho^{-Y})^{-X} = \rho^{-\{X,Y\}}$
  - (Local computation) If ρ and σ are valuations for r and s, respectively, $X \in r$, and $X \notin s$, then $(\rho \oplus \sigma)^{-X} = (\rho^{-X}) \oplus \sigma$
- We will sometimes denote $\rho^{-\{X\}}$ by $\rho^{\downarrow r \setminus \{X\}}$

Valuation-based Systems

- Making inference means finding marginals of the joint valuation for the variables of interest
- Thus, if X is a variable of interest, we compute $(\oplus\Psi)^{\downarrow X} = (\oplus\Psi)^{-(\Phi\setminus\{X\})}$ by marginalizing all variables in $\Phi \setminus \{X\}$ out of the joint valuation $\oplus\Psi$.

Valuation-based Systems

- VBS is an abstraction of several uncertainty calculi
  - propositional calculus
  - probability theory
  - belief function theory
  - Spohn's epistemic belief calculus
  - possibility theory
  - ...
- VBS can also be considered as an abstraction of
  - Optimization
  - Bayesian decision theory
  - Solving systems of equations
  - Relational database theory
  - ...

Outline

1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Basics of D-S Belief Function Theory

- Static: We represent knowledge using either:
  - basic probability assignment (bpa) μ
  - belief function β
  - plausibility function π
  - commonality function χ
- Dynamic: We make inferences using:
  - Dempster's rule of combination
  - Marginalization rule
- Inference: Given a set of belief functions (bpa, plausibility, belief, or commonality) representing knowledge of the domain, and all evidence, we would like to find the marginals of the joint 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

- Suppose Φ denotes a finite set of variables
- For each $X \in \Phi$, Ω_X denotes a finite set of states of X
- 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 non-empty subsets of Ω_s
- A basic probability assignment (bpa) μ for s is a function $\mu : 2^{\Omega_s} \to [0, 1]$ such that:

$$\sum_{\mathsf{a} \in 2^{\Omega_s}} \mu(\mathsf{a}) = 1 \tag{1}$$

- Subsets $\mathsf{a} \in 2^{\Omega_s}$ such that $\mu(\mathsf{a}) > 0$ are called focal elements of μ.

Basics of D-S Belief Function Theory

- (Vacuous bpa) Consider μ_0 for X such that $\mu_0(\Omega_X) = 1$. This represents vacuous knowledge of X. This is distinct from the equi-probable (Laplacian) distribution μ_L for X such that $\mu_L(\{x_i\}) = \frac{1}{|\Omega_X|}$ for each $x_i \in \Omega_X$.
- Smets' [2003] Example: Peter, Paul, or Mary?
  - The Godfather has decided to assassinate Mr. Jones.
  - He has three assassins on his payroll: Peter, Paul, and Mary
  - He will flip a fair coin. If heads, he will pick either Peter or Paul to do the job (we know nothing about how the Godfather will choose between Peter and Paul). If tails, he will pick Mary.
  - Mr. Jones is found dead. Who is the killer?

Basics of D-S Belief Function Theory

- Suppose K is a variable with states $\Omega_K = \{Pe, Pa, Ma\}$.
- Let μ_{K_1} denote the bpa for $\{K\}$ as follows:

$$\begin{aligned} \mu_{K_1}(\{Pe, Pa\}) &= 0.5, \\ \mu_{K_1}(\{Ma\}) &= 0.5. \end{aligned}$$

Basics of D-S Belief Function Theory

- The combination rule in 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.

Basics of D-S Belief Function Theory

- Let μ_{K_1} denote the bpa for $\{K\}$ as follows:

$$\begin{aligned} \mu_{K_1}(\{Pe, Pa\}) &= 0.5, \\ \mu_{K_1}(\{Ma\}) &= 0.5. \end{aligned}$$

- Evidence: Peter has an air-tight alibi. Let μ_{K_2} denote the bpa for $\{K\}$ as follows:

$$\mu_{K_2}(\{Pa, Ma\}) = 1$$

- After combining evidence using Dempster's rule, we have:

$$\begin{aligned} (\mu_{K_1} \oplus \mu_{K_2})(\{Pa\}) &= 0.5, \\ (\mu_{K_1} \oplus \mu_{K_2})(\{Ma\}) &= 0.5. \end{aligned}$$

Basics of D-S Belief Function Theory

- In general $\mu \oplus \mu \neq \mu$.
- Thus, in combining, e.g., μ_1 and μ_2 by Dempster's rule, it is important that μ_1 and μ_2 are distinct pieces of evidence, and there is no double-counting of uncertain knowledge.

Basics of D-S Belief Function Theory

- Dempster's rule satisfies all properties of combination
  - (Domain) If μ_1 is a bpa for s_1 and μ_2 is a bpa for s_2, then $\mu_1 \oplus \mu_2$ is a bpa for $s_1 \cup s_2$
  - (Commutativity) $\mu_1 \oplus \mu_2 = \mu_2 \oplus \mu_1$
  - (Associativity) $\mu_1 \oplus (\mu_2 \oplus \mu_3) = (\mu_1 \oplus \mu_2) \oplus \mu_3$

Basics of D-S Belief Function Theory

- Marginalization in belief function theory is addition.
- Projection of states: If $\mathbf{x} \in \Omega_s$, and $X \in s$, then $\mathbf{x}^{\downarrow s \setminus \{X\}}$ (or $\mathbf{x}^{-X}$) 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 a^{-X} (or $\mathsf{a}^{\downarrow s \setminus \{X\}}$) is

$$\mathsf{a}^{-X} = \{\mathbf{x}^{-X} \,:\, \mathbf{x} \in \mathsf{a}\}$$

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

$$\mu^{-X}(\mathsf{a}) = \sum_{\mathsf{b} \in 2^{\Omega_s} \,:\, \mathsf{b}^{-X} = \mathsf{a}} \mu(\mathsf{b})$$

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

Basics of D-S Belief Function Theory

- Suppose M and R are variables with $\Omega_M = \{t, f\}$ and $\Omega_R = \{t, f\}$,
- Suppose ρ is a bpa for $\{M, R\}$ such that

$$\begin{aligned} \rho(\{(t,t),(f,t),(f,f)\}) &= 0.1, \\ \rho(\{(t,f),(f,t),(f,f)\}) &= 0.7, \\ \rho(\Omega_{\{M,R\}}) &= 0.2. \end{aligned}$$

- Then ρ^{-M} is a bpa for $\{R\}$ such that:

$$\rho^{-M}(\{t, f\}) = 1.$$

- And ρ^{-R} is a bpa for $\{M\}$ such that:

$$\rho^{-R}(\{t, f\}) = 1.$$

- Thus, ρ by itself tells us nothing about M or R.

Basics of D-S Belief Function Theory

- The definition of marginalization of bpa function satisfies the properties of marginalization:
  - (Domain) If ρ is a bpa for r, and $X \in r$, then ρ^{-X} is a bpa for $r \setminus \{X\}$.
  - (Order does not matter) If ρ is a bpa for r, $X, Y \in r$, then $(\rho^{-X})^{-Y} = (\rho^{-Y})^{-X} = \rho^{-\{X,Y\}} = \rho^{\downarrow r \setminus \{X,Y\}}$.
  - (Local computation) If ρ and σ are bpa's for r and s, respectively, $X \in r$, and $X \notin s$, then $(\rho \oplus \sigma)^{-X} = (\rho^{-X}) \oplus \sigma$.

Basics of D-S Belief Function Theory

Conditional bpa's

- A conditional bpa for Y given $X = x$, denoted by $\mu_{Y|x}$ is a bpa for Y, i.e., $\mu_{Y|x} : 2^{\Omega_Y} \to [0,1]$ such that

$$\sum_{\mathsf{a} \in 2^{\Omega_Y}} \mu_{Y|x}(\mathsf{a}) = 1.$$

- The knowledge of Y encoded in $\mu_{Y|x}$ is valid only in the case $X = x$.
- Using Smets' conditional embedding, we convert the conditional bpa $\mu_{Y|x}$ for Y to an unconditional bpa $\mu_{x,Y}$ for (X, Y) as follows:

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

  for all focal elements b of $\mu_{Y|x}$.

Basics of D-S Belief Function Theory

An Example

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

$$\begin{aligned} \mu_{Y|x}(\{y\}) &= 0.8, \\ \mu_{Y|x}(\Omega_Y) &= 0.2. \end{aligned}$$

- Then $\mu_{x,Y}$ is as follows:

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

Basics of D-S Belief Function Theory

- Conditional bpa $\mu_{Y|x}$ for Y given x is only well-defined if $\mu_X(\{x\}) > 0$.
- The (unconditional) bpa $\mu_{x,Y}$ for (X, Y) has the following three properties:
  1. $\mu_{x,Y}^{\downarrow X}$ is a vacuous bpa for X, i.e., $\mu_{x,Y}^{\downarrow X}(\Omega_X) = 1$. $\mu_{x,Y}$ by itself tells us nothing about X.
  2. $\mu_{x,Y}^{\downarrow Y}$ is a vacuous bpa for Y, i.e., $\mu_{x,Y}^{\downarrow Y}(\Omega_Y) = 1$. $\mu_{x,Y}$ by itself tells us nothing about Y.
  3. Suppose μ_x is a bpa for X as follows: $\mu_x(\{x\}) = 1$. Then, $(\mu_{x,Y} \oplus \mu_x)^{\downarrow Y} = \mu_{Y|x}$.

Basics of D-S Belief Function Theory

- Consider $\mu_{x,Y}$:

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

- It is clear that $\mu_{x,Y}^{\downarrow X}$ is vacuous for X, and $\mu_{x,Y}^{\downarrow Y}$ is vacuous for Y.
- Consider $\mu_{x,Y} \oplus \mu_x$:

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

Outline

1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Captain's Problem

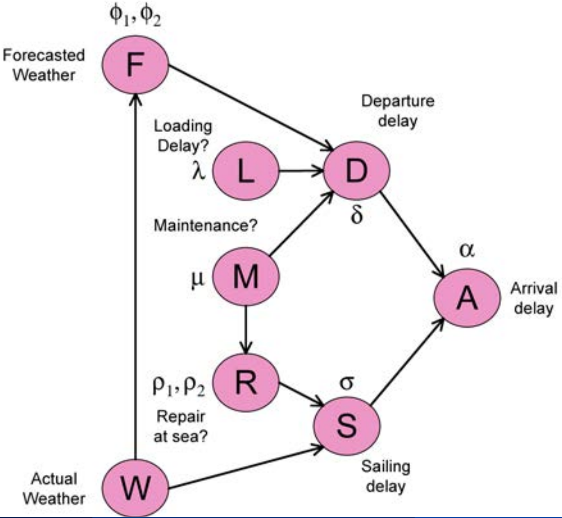
- Captain's Problem (R. Almond, *Graphical Belief Modeling*, Chapman and Hall, 1995)
  - A ship's captain is concerned about how many days his ship may be delayed before arrival at a destination.
  - The delay in arrival may be a result of delay in departure and/or delay in sailing.
  - Delay in departure may be a result of maintenance (at most 1 day), delay in loading (at most 1 day) or due to forecast of bad weather (at most I day).
  - Delay in sailing may be a result of bad weather (at most 1 day) and/or whether repairs may be needed at sea (at most 1 day).
  - If maintenance is done before sailing, chances of repairs at sea is less likely.
  - Weather forecast says small chance of bad weather (.2), good chance of good weather (0.6). Forecast is 80% reliable.
  - Captain has some knowledge of loading delay, and whether maintenance is done before departure.

Captain's Problem

- Variables
  - A (arrival delay), $\Omega_A = \{0, 1, 2, 3, 4, 5\}$.
  - D (departure delay), $\Omega_D = \{0, 1, 2, 3\}$.
  - S (sailing delay), $\Omega_S = \{0, 1, 2\}$.
  - L (is loading delayed?), $\Omega_L = \{t, f\}$.
  - F (weather forecast), $\Omega_F = \{b, g\}$.
  - W (actual weather), $\Omega_W = \{b, g\}$.
  - M (is maintenance done before sailing?), $\Omega_M = \{t, f\}$.
  - R (is repair at sea needed?), $\Omega_R = \{t, f\}$.

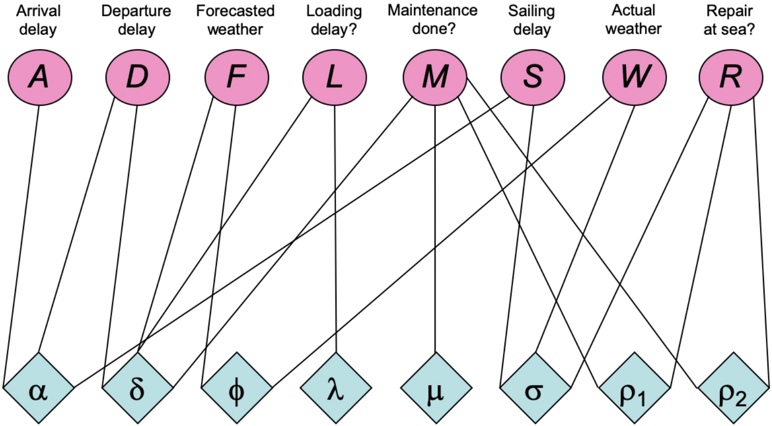
Captain's Problem

- The Captain problem can be described by a causal directed acyclic graph (DAG) as follows:



Captain's Problem

- Valuation Network: A bipartite graph with variables and valuations as nodes. Each valuation is linked to the variables in its domain.



Captain's Problem

- Consider the piece of knowledge: Arrival delay is sum of departure delay and sailing delay
- We model this piece of knowledge by a bpa α for $\{A, D, S\}$ such that

$$\alpha(\{(0,0,0),(1,1,0),(2,2,0),(3,3,0),\\ (1,0,1),(2,1,1),(3,2,1),(4,3,1),\\ (2,0,2),(3,1,2),(4,2,2),(5,3,2)\}) = 1.$$

- α has one focal set. Such bpa are called deterministic.

Captain's Problem

- Loading delay, bad weather forecast, and maintenance each adds one day to departure delay
- We model this piece of knowledge by a bpa δ for $\{D, L, F, M\}$ such that

$$\delta(\{(0, f, g, f), (1, t, g, f), (1, f, b, f), (1, f, g, t), \\ (2, f, b, t), (2, t, g, t), (2, t, g, f), (3, t, b, t)\}) = 1.$$

Captain's Problem

- At least 90% of the time, bad weather and repair at sea each adds 1 day to sailing delay
- We model this by bpa σ for $\{S, W, R\}$ such that

$$\begin{aligned} \sigma(\{(0,g,f),(1,b,f),(1,g,t),(2,b,t)\}) &= 0.9, \\ \sigma(\Omega_{\{S,A,R\}}) &= 0.1 \end{aligned}$$

Captain's Problem

- Forecast is 80% reliable
- This piece of knowledge is represented by bpa ϕ_1 for $\{F, W\}$ such that

$$\begin{aligned} \phi_1(\{(b,b),(g,g)\}) &= 0.8, \\ \phi_1(\Omega_{\{F,W\}}) &= 0.2. \end{aligned}$$

- Forecast predicts bad weather with chance 0.2 and good weather with chance 0.6
- This piece of knowledge is represented by bpa ϕ_2 for $\{F\}$ such that

$$\begin{aligned} \phi_2(\{b\}) &= 0.2, \\ \phi_2(\{g\}) &= 0.6, \\ \phi_2(\Omega_{\{F\}}) &= 0.2. \end{aligned}$$

Captain's Problem

- Loading is delayed with chance 0.3, and on schedule with chance 0.5.
- This piece is model by bpa λ for $\{L\}$ such that

$$\begin{aligned}\lambda(\{t\}) &= 0.3,\\ \lambda(\{f\}) &= 0.5,\\ \lambda(\Omega_{\{L\}}) &= 0.2.\end{aligned}$$

- No maintenance was done on the ship prior to departure
- This piece of knowledge is represented by bpa μ for $\{M\}$ such that

$$\mu(\{f\}) = 1.$$

Captain's Problem

- If maintenance was done prior to sailing, then chances of repair at sea is between 10 and 30%. This is represented by conditional bpa $\rho_{R|M=t}$ as follows:

$$\begin{aligned}\rho_{R|M=t}(\{t\}) &= 0.1,\\ \rho_{R|M=t}(\{f\}) &= 0.7,\\ \rho_{R|M=t}(\{t,f\}) &= 0.2.\end{aligned}$$

- After conditional embedding, ρ_1 is a bpa for (M, R) as follows:

$$\begin{aligned}\rho_1(\{(t,t),(f,t),(f,f)\}) &= 0.1,\\ \rho_1(\{(t,f),(f,t),(f,f)\}) &= 0.7,\\ \rho_1(\Omega_{\{M,R\}}) &= 0.2.\end{aligned}$$

Captain's Problem

- If maintenance was not done prior to sailing, then chances of repair at sea is between 20 and 80%. This is represented by conditional bpa $\rho_{R|M=f}$ as follows:

$$\begin{aligned} \rho_{R|M=f}(\{t\}) &= 0.2, \\ \rho_{R|M=f}(\{f\}) &= 0.2, \\ \rho_{R|M=f}(\{t,f\}) &= 0.6. \end{aligned}$$

- After conditional embedding, ρ_2 is a bpa for (M, R) as follows:

$$\begin{aligned} \rho_2(\{(f,t),(t,t),(t,f)\}) &= 0.2, \\ \rho_2(\{(f,f),(t,t),(t,f)\}) &= 0.2, \\ \rho_2(\Omega_{\{M,R\}}) &= 0.6. \end{aligned}$$

Outline

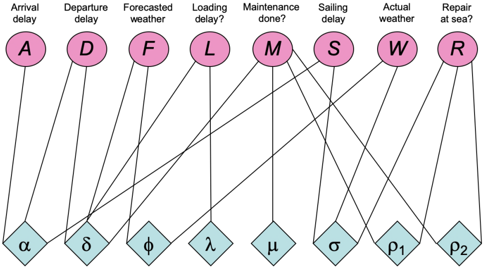
1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Local Computation

- Making inference means finding marginals of the joint valuation $\oplus\Psi$ for the variables of interest.
- If there are many variables in Φ, computing the joint valuation $\oplus\Psi$ for Φ is intractable.
- However, one can compute the marginal of the joint for X, $(\oplus\Psi)^{\downarrow X}$, without computing the joint explicitly, using so-called local computation.
- The axiom that allows local computation is the local computation axiom: If ρ and σ are bpa's for r and s, respectively, $X \in r$, and $X \notin s$, then $(\rho \oplus \sigma)^{-X} = (\rho^{-X}) \oplus \sigma$.

Local Computation

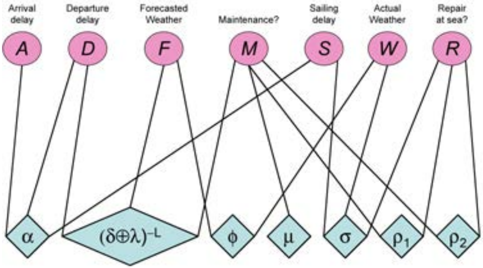
- Consider the Captain's problem. We would like to compute the marginal of the joint for A. So we have to marginalize all other variables from the joint.



- Consider L. It is in the domain of δ and λ only. The local computation axiom guarantees that if we replace δ and λ by $(\delta \oplus \lambda)^{-L}$, then the product of all valuations will give us $(\oplus\Psi)^{-L}$.

Local Computation

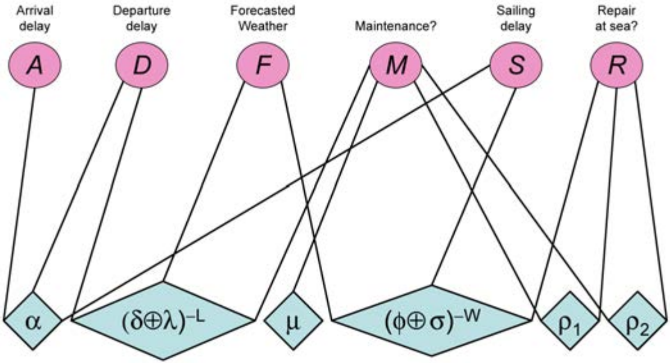
- The reduced VN is as follows:



- We can recursively remove all but A from the VN.
- Consider W. It is in the domain of ϕ for $\{F, W\}$ and σ for $\{S, W, R\}$. Thus, $(\phi \oplus \sigma)^{-W}$ will be a bpa for $\{F, S, R\}$.

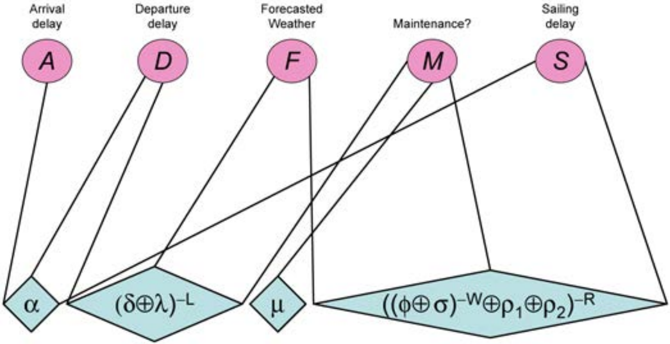
Local Computation

- After deletion of $\{L, W\}$:



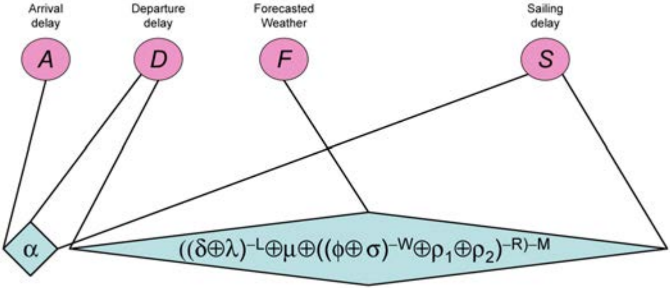
Local Computation

- After deletion of $\{L, W, R\}$:



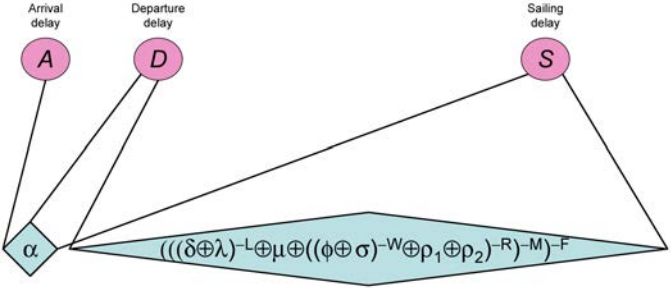
Local Computation

- After deletion of $\{L, W, R, M\}$:



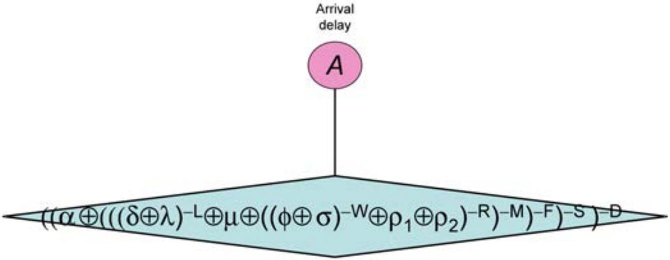
Local Computation

- After deletion of $\{L, W, R, M, F\}$:



Local Computation

- After deletion of $\{L, W, R, M, F, S, D\}$ in this order:



Local Computation

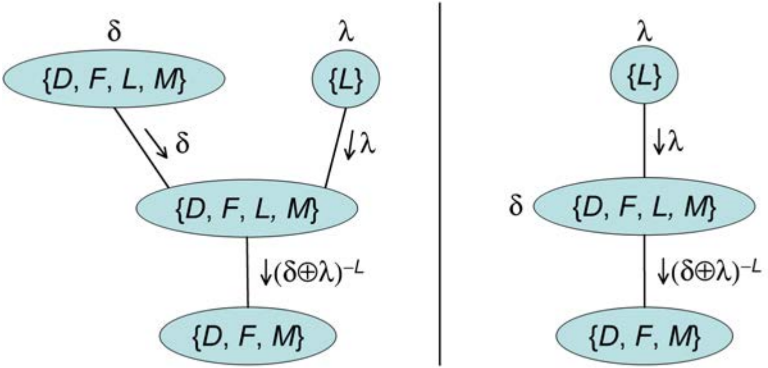
- In finding the marginal for A, we used deletion sequence $L\,W\,R\,M\,F\,S\,D$.
- The order does not matter axiom allows us to use any deletion sequence (and obtain the same marginal).
- Some deletion sequences involve less computation than others.
- Finding an optimal deletion sequence is a hard problem.
- So we use heuristics to select a sequence such as one-step-look-ahead: The variable to be marginalized next is the one that leads to combination on the smallest domain.
- A local computation algorithm for finding marginals is implemented in Belief Function Machine (to be demonstrated later).

Local Computation

- If we can find the marginal for A, we can find the marginal for any other variable in a similar manner.
- However, there may be duplication of computations.
- We can avoid duplication by saving intermediate computations in a data structure called a "join tree."
- A join tree is a tree with subsets of variables as nodes having the property that if a variable appears in two different nodes, then it appears in all nodes in the path between the two nodes.

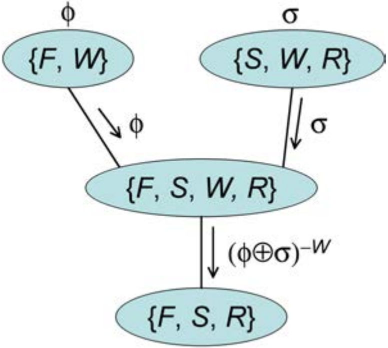
Local Computation

- Consider deletion of L. We can describe the computation as messages between nodes as follows:



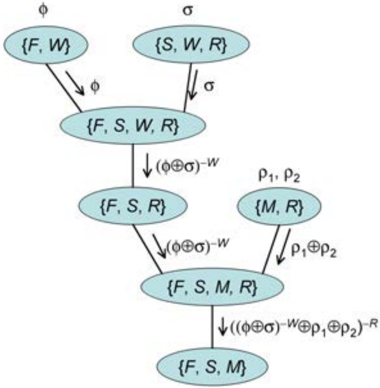
Local Computation

- Similarly, deletion of W can be described as follows:



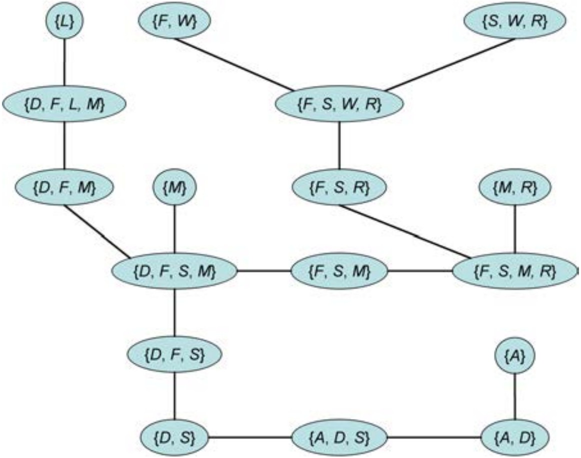
Local Computation

- Similarly, deletion of R can be described as follows:



Local Computation

- The tree thus constructed is a join tree.

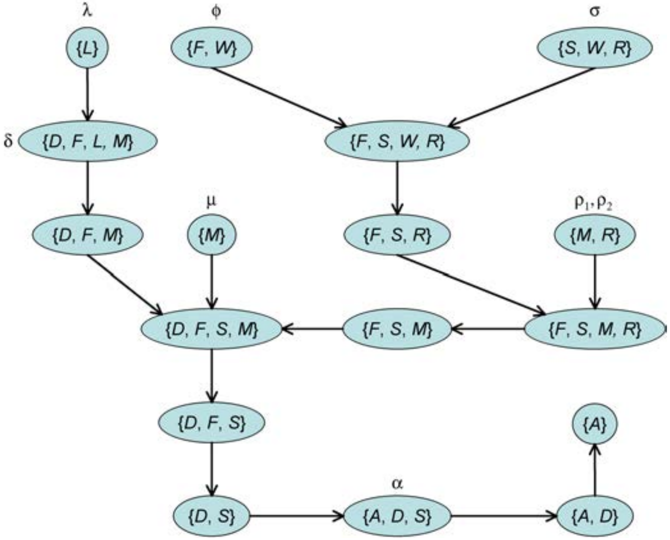


Local Computation

- To find the marginal for a variable, say A, orient all edges toward A, which is now the root of the tree.
- Each node sends a message to its inward neighbor, which is the combination of all messages it receives from its outward neighbors plus what it has suitably marginalized.
- Timing: Leaves (nodes with no outward neighbors) can send messages right away. Non-leaves have to wait till they receive a message from all its outward neighbors.
- Process is finished when the root has received a message from its outward neighbors. The root then combines all the messages it receives plus what it has.

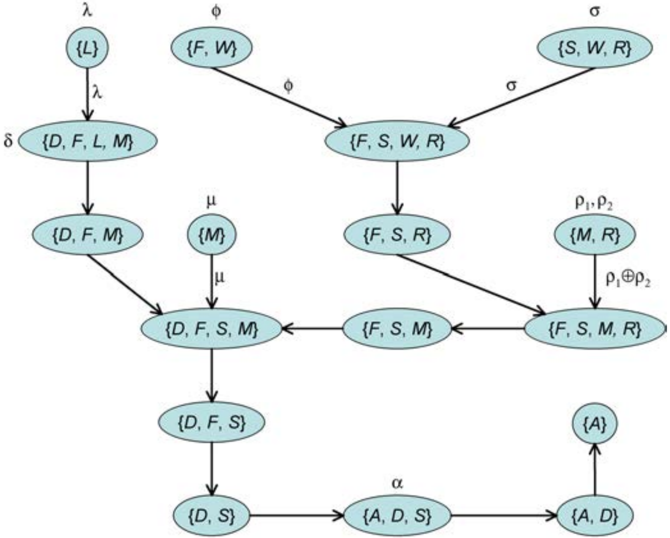
Local Computation

- At the beginning:



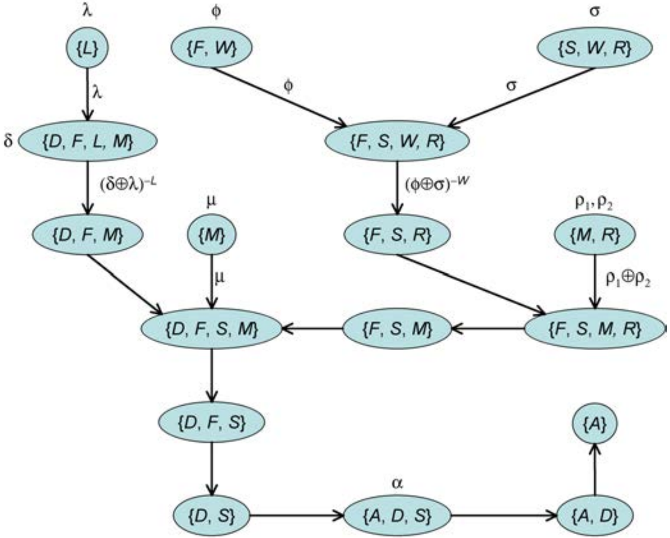
Local Computation

- Time step 1:



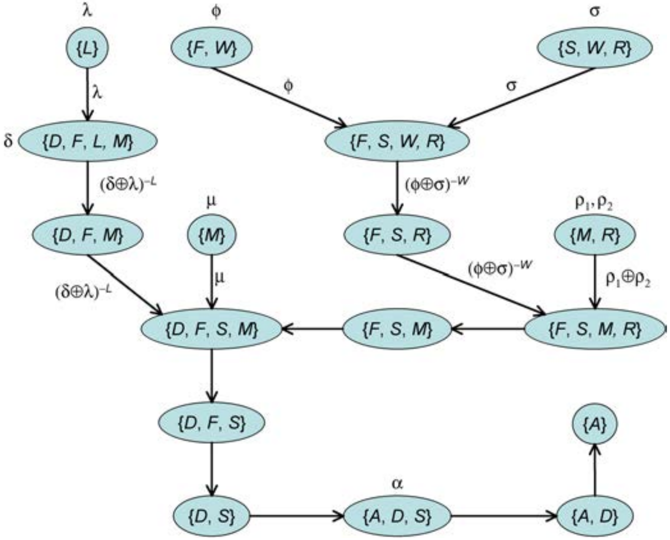
Local Computation

- Time step 2:



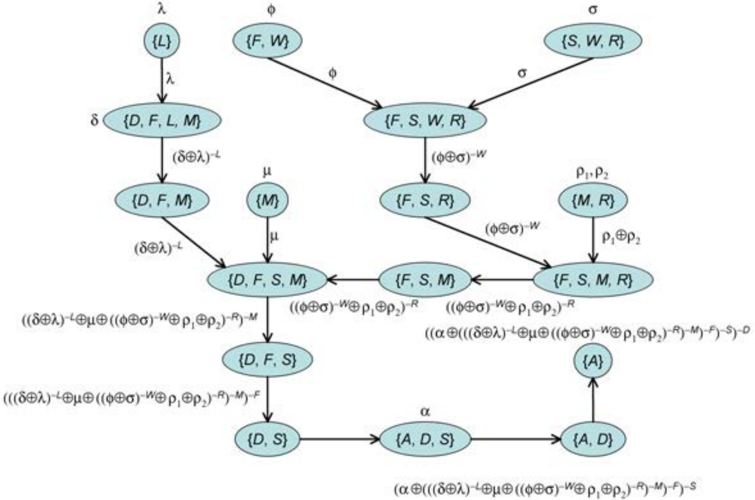
Local Computation

- Time step 3:



Local Computation

- Computing a marginal can be described as propagation of messages in a join tree:

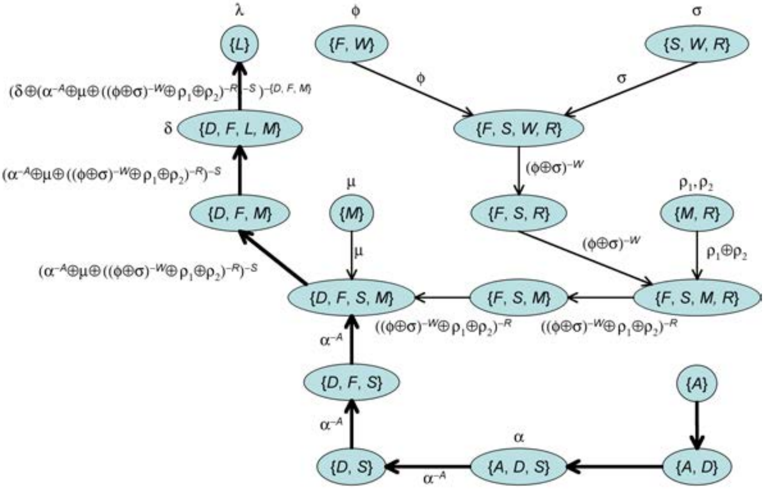


Local Computation

- Suppose we now wish to find the marginal for L.
- We now re-orient some of the edges so L is the root, and compute new messages for the re-oriented edges.

Local Computation

- End of propagation for L:



Outline

1. Valuation-based Systems
2. Basics of Dempster-Shafer belief function theory
3. Captain's Problem
4. Local Computation
5. Belief Function Machine
   - Captain's Problem
   - Chest Clinic
   - Communication Network
6. References

Belief Function Machine

- A software 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 reduce the marginal belief function to probabilities
  - Can do sensitivity analysis
- Can be downloaded for free from http://pshenoy.faculty.ku.edu/Papers/BFM072503.zip

Belief Function Machine

- Suppose we wish to solve the Captain's Problem
- Input the problem as a UIL file "captain.txt"
  - Define variables and their state spaces
  - Define valuations and their domains
  - Describe the details of each valuations as bpa's or as conditional bpa's
  - Conditional bpa's are converted to regular bpa's using Smets' conditional embedding

Belief Function Machine

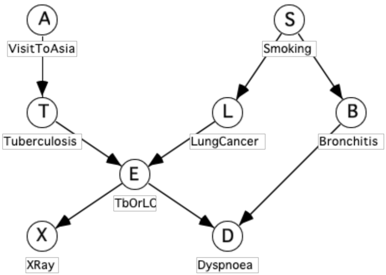
Demo solution of Captain's problem using BFM in Matlab.

Script:

```
uil2bm('Captain.txt','bmcaptain');
global BELIEF VARIABLE ATTRIBUTE STRUCTURE FRAME QUERY
BELTRACE NODE BJTREE TRANPROTOCOL;
belall = condiembed([1:9]);
keepbel = belall;
showbel(solve(belall,'A'));
showbel(bel2prob(26));
```

Belief Function Machine

- BFM can solve 'complete' models using belief functions
- Consider Chest Clinic example from [Lauritzen-Spiegelhalter 1988]
- BFM gives exactly the same answers as a Bayes net software would.



$P(A)$:	$p(a) = .01$	$P(E \mid L,T)$:	$p(e \mid l, t) = 1$ $p(e \mid l, \sim t) = 1$ $p(e \mid \sim l, t) = 1$ $p(e \mid \sim l, \sim t) = 0$
$P(T \mid A)$:	$p(t \mid a) = .05$ $p(t \mid \sim a) = .01$		
$P(S)$:	$p(s) = .50$	$P(X \mid E)$:	$p(x \mid e) = .98$ $p(x \mid \sim e) = .05$
$P(L \mid S)$:	$p(l \mid s) = .10$ $p(l \mid \sim s) = .01$	$P(D \mid E,B)$:	$p(d \mid e, b) = .90$ $p(d \mid e, \sim b) = .70$ $p(d \mid \sim e, b) = .80$ $p(d \mid \sim e, \sim b) = .10$
$P(B \mid S)$:	$p(b \mid s) = .60$ $p(b \mid \sim s) = .30$		

Belief Function Machine

Demo solution of Chest Clinic using BFM in Matlab.
Script:

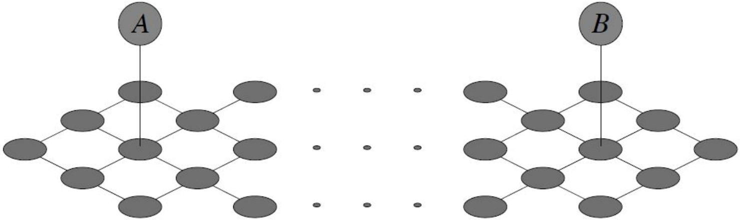
```
uil2bm('lschestclin.txt', 'bmlschest');
global BELIEF VARIABLE ATTRIBUTE STRUCTURE FRAME QUERY
BELTRACE NODE BJTREE TRANPROTOCOL;
belall = condiembed([1:15]);
keepbel(belall);
showbel(solve(belall, 'T'));
bjtbuild('T', belall)
showbel(solvetreeall(1));
belupdate(1, 15, observe('A', 'a'));
```

Comment: 1 is tree #, 15 is old belief function for A, observe('A','a') is new belief function for A

```
showbel(solvetreeall(1));
```

Belief Function Machine

- Communication network [Haenni-Lehmann 2002]
- We have a grid of $44 = 8 + 9 + 10 + 9 + 8$ communication nodes arranged in 5 rows
- There are 68 links, and each link has 90% reliability
- Nodes A and B are connected to the grid with links having 80% reliability
- What is the reliability of the connection between A and B? (Ans: $\approx 64\%$)



Belief Function Machine

Demo solution of Communication network using BFM in Matlab.

Script:

```
uil2bm('comm8.txt', 'bmcomm8');
global BELIEF VARIABLE ATTRIBUTE STRUCTURE FRAME QUERY
BELTRACE NODE BJTREE TRANPROTOCOL;
showbel(solve([1:70], [1,2]));
```

Outline

1 Valuation-based Systems

2 Basics of Dempster-Shafer belief function theory

3 Captain's Problem

4 Local Computation

5 Belief Function Machine
- Captain's Problem
- Chest Clinic
- Communication Network

6 References

References

- VBS: Shenoy, PP, "A Valuation-Based Language for Expert Systems," *Int. J. of Approx. Reas.*, 3(2), 1989, 383–411
- Local Computation: Shenoy, PP & G Shafer, "Axioms for Probability and Belief-Function Propagation," in Shachter, RD, T Levitt, JF Lemmer and LN Kanal (eds.), *Uncert. in Art. Int.*, 4, 1990, 169–198
- Captain's Problem: Almond, R, *Graphical Belief Modeling*, 1995, Chapman & Hall
- Chest Clinic: SL Lauritzen and DJ Spiegelhalter, "Local computations with probabilities on graphical structures and their application to experts systems" *JRSS*, ser. B, 50(2), 1988, 157–224.
- Communication network: R Haenni and N Lehmann, "Resource bounded and anytime approximation of belief function computations," *Int. J. of Approx. Reas.*, 31(1–2), 2002, 103–154.

Questions

Questions?