Reasoning with Uncertainty

Dempster-Shafer

Dempster-Shafer

- Dempster-Shafer is a belief system that deals with the evidence available for a hypothesis
  - Uncertainty is represented as:
    - $Bel(U)$ Belief (evidence for a hypothesis)
    - $Plaus(U)$ Plausibility (evidence that does not contradict a hypothesis)
  - Belief and Plausibility can be viewed as providing a lower and upper bound, respectively, on the likelihood of U

Dempster-Shafer

- Dempster-Shafer belief and plausibility functions are defined over the set of all subsets of possible states (or worlds)
  - $Bel(x): 2^W \rightarrow [0..1]$, $W=\{w_1, w_2, \ldots, w_n\}$
  - $Plaus(x): 2^W \rightarrow [0..1]$, $W=\{w_1, w_2, \ldots, w_n\}$
  - Since belief and plausibility encode evidence, they can not be defined solely on individual states

$$\sum_i Bel(w_i) \leq 1 \, , \; \sum_i Plaus(w_i) \geq 1$$

Properties of Dempster-Shafer Belief Functions

- Dempster-Shafer Belief functions have the following properties:
  - $Bel(\emptyset) = 0$
  - $Bel(W) = 1$
  - $Bel\left(\bigcup_{i=1}^{n} U_i\right) \geq \sum_{i=1}^{n} \sum_{\{I \subseteq \{1,\ldots,n\} : |I| = i\}} (-1)^{i+1} Bel\left(\bigcap_{j \in I} U_j\right)$
  - $Plaus(U) = 1 - Bel(\overline{U})$
  - $Plaus\left(\bigcap_{i=1}^{n} U_i\right) \geq \sum_{i=1}^{n} \sum_{\{I \subseteq \{1,\ldots,n\} : |I| = i\}} (-1)^{i+1} Plaus\left(\bigcup_{j \in I} U_j\right)$
- From the properties follows:
  - $Bel(U) \leq Plaus(U)$

Dempster-Shafer Belief Functions

- Belief encodes the sum of all support for any subset of U provided by the available evidence
- Support is encoded as a mass function m

  $m: 2^W \rightarrow [0..1]$

  - $m(\emptyset) = 0$
  - $\sum_{U \subseteq W} m(U) = 1$

- $Bel_m(U) = \sum_{U' \subseteq U} m(U')$
- $Plaus_m(U) = \sum_{U' \cap U \neq \emptyset} m(U')$

Combining Evidence

- "Independent" Evidence for a belief can be combined using Dempster's Rule of Combination
  - "Independence" of evidence implies that the sources of the evidence are unrelated
  - Combination requires at least two sets U_1, U_2 with:
    - $U_1 \cap U_2 \neq \varnothing$
    - $m_1(U_1) * m_2(U_2) \neq 0$

Combining Evidence

- Dempster's Rule of Combination

$$(m_1 \oplus m_2)(\emptyset) = 0$$

$$(m_1 \oplus m_2)(U) = \sum_{U_1 \cap U_2 = U} m_1(U_1) m_2(U_2) / c$$

$$c = \sum_{U_1 \cap U_2 \neq \emptyset} m_1(U_1) m_2(U_2)$$

  - Properties:
    - Commutative: $m_1 \oplus m_2 = m_2 \oplus m_1$
    - Associative: $(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3)$

Conditioning Belief Functions

- The way to condition belief functions depends on their interpretation.
  - $[Bel(U), Plaus(U)]$ as a likelihood interval
  - $Bel(U)$ and $Plaus(U)$ as supporting evidence

- It is important to be careful when deciding which interpretation is appropriate for a particular problem.

Conditioning Belief Functions

- Interpreting Belief and Plausibility as a likelihood interval
  - $[Bel(V|U), Plaus(V|U)]$ is an interval for the potential value of $P(V|U)$

$$Bel(V \mid U) = \begin{cases} \dfrac{Bel(V \cap U)}{Bel(V \cap U) + Plaus(\overline{V} \cap U)} & \text{if } Plaus(\overline{V} \cap U) > 0 \\ 1 & \text{otherwise} \end{cases}$$

$$Plaus(V \mid U) = \begin{cases} \dfrac{Plaus(V \cap U)}{Plaus(V \cap U) + Bel(\overline{V} \cap U)} & \text{if } Plaus(V \cap U) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Conditioning Belief Functions

- Interpreting Belief and Plausibility as supporting evidence (DS conditioning)
  - $Bel(V\,||U)$ is the sum of all support for V if U has support 1
  - $Plaus(V\,||U)$ is the sum of all support that is not against V if U has support 1
    - $m_{||U}(V) = (m \oplus m_U)(V)$, $m_U(U) = 1$, $m_U(V) = 0$ for all $V \neq U$

$$Bel(V \,\|\, U) = \frac{Bel(V \cup \overline{U}) - Bel(\overline{U})}{1 - Bel(\overline{U})}$$

$$Plaus(V \,\|\, U) = \frac{Plaus(V \cap U)}{Plaus(U)}$$

Conditioning Beliefs: Example

- Three Prisoner Problem:
  - Of three prisoners, A, B, and C, only one will not be executed the next day.
  - Prisoner A has the option to ask the guard to give him the name of one of the two other prisoners who will be executed.
  - Question: Given that the guard says that B will be executed, what should As belief be in not being executed ?
- Formulation:
  - Let L_A, L_B, L_C indicate that A, B, or C will not be executed, respectively.
  - Let GsB, and GsC indicate that the guard said that B or C are going to be executed, respectively.
  - Then the problem is to determine the belief in L_A given GsB

Conditioning Beliefs: Example

- Probabilistic answer:

$$P(L_A) = P(L_B) = P(L_C) = \frac{1}{3}$$

$$P(GsB) = P(GsC) = P(GsB \mid L_A) = P(GsC \mid L_A) = \frac{1}{2}$$

$$P(L_A \mid GsB) = \frac{P(GsB \mid L_A)P(L_A)}{P(GsB)} = \frac{\frac{1}{2} * \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

- Dempster-Shafer Formulation:

$$W = \{(L_A, GsB), (L_A, GsC), (L_B, GsC), (L_C, GsB)\}$$

$$m(L_A) \equiv m(\{(L_A, GsB), (L_A, GsC)\}) = \frac{1}{3}$$

$$m(L_B) \equiv m((L_B, GsC)) = \frac{1}{3}$$

$$m(L_C) \equiv m((L_C, GsB)) = \frac{1}{3}$$

$$m(X) = 0 \;\; \textit{for all other sets } X$$

$$Bel(L_A) = Bel(L_B) = Bel(L_C) = \frac{1}{3}$$

$$\begin{aligned} Bel(GsB) &= Bel(\{(L_A, GsB), (L_C, GsB)\}) \\ &= m((L_A, GsB)) + m((L_C, GsB)) + m(\{(L_A, GsB), (L_C, GsB)\}) \\ &= 0 + \frac{1}{3} + 0 \end{aligned}$$

Conditioning Beliefs: Example

- Likelihood Interval Interpretation:

$$Bel(L_A \mid GsB) = \frac{Bel(L_A \cap GsB)}{Bel(L_A \cap GsB) + Plaus(\overline{L_A} \cap GsB)}$$

$$= \frac{Bel(\{(L_A, GsB), (L_A, GsC)\} \cap \{(L_A, GsB), (L_C, GsB)\})}{Bel(\{(L_A, GsB), (L_A, GsC)\} \cap \{(L_A, GsB), (L_C, GsB)\}) + Plaus(\{(L_B, GsC), (L_C, GsB)\} \cap \{(L_A, GsB), (L_C, GsB)\})} = \frac{0}{0 + \frac{1}{3}} = 0$$

$$Plaus(L_A \mid GsB) = \frac{Plaus(L_A \cap GsB)}{Plaus(L_A \cap GsB) + Bel(\overline{L_A} \cap GsB)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{3}} = \frac{1}{2}$$

- Evidence Interpretation:

$$Bel(L_A \| GsB) = \frac{Bel(L_A \cup \overline{GsB}) - Bel(\overline{GsB})}{1 - Bel(\overline{GsB})}$$

$$= \frac{Bel(\{(L_A, GsB), (L_A, GsC)\} \cup \{(L_A, GsC), (L_B, GsC)\}) - Bel(\{(L_A, GsC), (L_B, GsC)\})}{1 - Bel(\{(L_A, GsC), (L_B, GsC)\})} = \frac{\frac{2}{3} - \frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{2}$$

$$Plaus(L_A \| GsB) = \frac{Plaus(L_A \cap GsB)}{Plaus(GsB)} = \frac{Plaus(\{(L_A, GsB), (L_A, GsC)\} \cap \{(L_A, GsB), (L_C, GsB)\})}{Plaus(\{(L_A, GsB), (L_C, GsB)\})} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

Dempster-Shafer

- Advantages
  - Allows differentiation between unknown information and uncertainty
  - Represents a likelihood interval
  - Permits to deal with more subjective interpretations of evidence
- Disadvantages
  - High complexity due to the representation on subsets of the state space
  - Sensitive to the correct interpretation during conditioning