Topic 4

Representation and Reasoning with Uncertainty

Contents

4.3 Dempster-Shafer Theory

- Dempster-Shafer theory is an approach to combining evidence
- Dempster (1967) developed means for combining degrees of belief derived from independent items of evidence.
- His student, Glenn Shafer (1976), developed method for obtaining degrees of belief for one question from subjective probabilities for a related question
- People working in Expert Systems in the 1980s saw their approach as ideally suitable for such systems.

4.3 Dempster-Shafer Theory

- Each fact has a degree of support, between 0 and 1:
  - 0 No support for the fact
  - 1 full support for the fact
- Differs from Bayesian approah in that:
  - Belief in a fact and its negation need not sum to 1.
  - Both values can be 0 (meaning no evidence for or against the fact)

4.3 Dempster-Shafer Theory

Set of possible conclusions: Θ

$$\Theta = \{ \theta_1, \theta_2, \ldots, \theta_n \}$$

Where:

- Θ is the set of possible conclusions to be drawn
- Each θ_i is **mutually exclusive**: at most one has to be true.
- Θ is **Exhaustive**: At least one θ_i has to be true.

4.3 Dempster-Shafer Theory

Frame of discernment :

$$\Theta = \{ \theta_1, \theta_2, \ldots, \theta_n\}$$

- Bayes was concerned with evidence that supported single conclusions (e.g., evidence for each outcome θ_i in Θ):
  - $p(\theta_i \mid E)$

- D-S Theoryis concerned with evidences which support subsets of outcomes in Θ, e.g.,

  $\theta_1 \vee \theta_2 \vee \theta_3$ or $\{\theta_1, \theta_2, \theta_3\}$

4.3 Dempster-Shafer Theory

Frame of discernment :

- The “frame of discernment” (or “Power set”) of Θ is the set of all possible subsets of Θ:
  - E.g., if $\Theta = \{ \theta_1, \theta_2, \theta_3\}$
- Then the frame of discernment of Θ is:

  $(\varnothing, \theta_1, \theta_2, \theta_3, \{\theta_1, \theta_2\}, \{\theta_1, \theta_3\}, \{\theta_2, \theta_3\}, \{ \theta_1, \theta_2, \theta_3\})$
- $\varnothing$, the empty set, has a probability of 0, since one of the outcomes has to be true.
- Each of the other elements in the power set has a probability between 0 and 1.
- The probability of $\{ \theta_1, \theta_2, \theta_3\}$ is 1.0 since one has to be true.

4.3 Dempster-Shafer Theory

Mass function m(A):

(where A is a member of the power set)

= proportion of all evidence that supports this element of the power set.

"The mass *m*(*A*) of a given member of the power set, *A*, expresses the proportion of all relevant and available evidence that supports the claim that the actual state belongs to *A* but to no particular subset of *A*." (wikipedia)

"The value of *m*(*A*) pertains *only* to the set *A* and makes no additional claims about any subsets of *A*, each of which has, by definition, its own mass.

4.3 Dempster-Shafer Theory

Mass function m(A):

- Each m(A) is between 0 and 1.
- All m(A) sum to 1.
- m(Ø) is 0 - at least one must be true.

4.3 Dempster-Shafer Theory

Mass function m(A): Interpetation of m({AvB})=0.3

- means there is evidence for {AvB} that cannot be divided among more specific beliefs for A or B.

4.3 Dempster-Shafer Theory

Mass function m(A): example

- 4 people (B, J, S and K) are locked in a room when the lights go out.
- When the lights come on, K is dead, stabbed with a knife.
- Not suicide (stabbed in the back)
- No-one entered the room.
- Assume only one killer.
- $\Theta = \{ B, J, S\}$
- $P(\Theta) = (\emptyset, \{B\}, \{J\}, \{S\}, \{B,J\}, \{B,S\}, \{J,S\}, \{B,J,S\})$

4.3 Dempster-Shafer Theory

Mass function m(A): example (cont.)

- Detectives, after reviewing the crime-scene, assign mass probabilities to various elements of the power set:

Event	Mass
No-one is guilty	0
B is guilty	0.1
J is guilty	0.2
S is guilty	0.1
either B or J is guilty	0.1
either B or S is guilty	0.1
either S or J is guilty	0.3
One of the 3 is guilty	0.1

4.3 Dempster-Shafer Theory

Belief in A:

The **belief** in an element A of the Power set is the sum of the masses of elements which are subsets of A (including A itself).

E.g., given $A=\{q_1, q_2, q_3\}$

$$\begin{aligned}Bel(A) = \; & m(q_1)+m(q_2)+m(q_3) \\ & + m(\{q_1, q_2\})+m(\{q_2, q_3\})+m(\{q_1, q_3\}) \\ & +m(\{q_1, q_2, q_3\})\end{aligned}$$

4.3 Dempster-Shafer Theory

Belief in A: example

- Given the mass assignments as assigned by the detectives:

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1

- bel({B}) = m({B}) = 0.1
- bel({B,J}) = m({B})+m({J})+m({B,J}) =0.1+0.2+0.1=0.4
- Result:

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
bel(A)	0.1	0.2	0.1	0.4	0.3	0.6	1.0

4.3 Dempster-Shafer Theory

Plausibility of A: pl(A)

The plausability of an element A, pl(A), is the sum of all the masses of the sets that intersect with the set *A*:

E.g. pl({B,J}) = m(B)+m(J)+m(B,J)+m(B,S) +m(J,S)+m(B,J,S)

= 0.9

All Results:

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
pl(A)	0.4	0.7	0.6	0.9	0.8	0.9	1.0

4.3 Dempster-Shafer Theory

Disbelief (or Doubt) in A: dis(A)

The disbelief in A is simply bel(¬A).

It is calculated by summing all masses of elements which do not intersect with A.

The plausibility of A is thus 1-dis(A):

$$pl(A) = 1- dis(A)$$

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
dis(A)	0.6	0.3	0.4	0.1	0.2	0.1	0
pl(A)	0.4	0.7	0.6	0.9	0.8	0.9	1.0

4.3 Dempster-Shafer Theory

Belief Interval of A:

The certainty associated with a given subset A is defined by the belief interval:

$$[\ bel(A)\ \ pl(A)\ ]$$

E.g. the belief interval of {B,S} is: [0.1 0.8]

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
bel(A)	0.1	0.2	0.1	0.4	0.3	0.6	1.0
pl(A)	0.4	0.7	0.6	0.9	0.8	0.9	1.0

4.3 Dempster-Shafer Theory

Belief Intervals & Probability

The probability in A falls somewhere between bel(A) and pl(A).

- bel(A) represents the evidence we have for A directly. So prob(A) cannot be less than this value.
- pl(A) represents the maximum share of the evidence we could possibly have, if, for all sets that intersect with A, the part that intersects is actually valid. So pl(A) is the maximum possible value of prob(A).

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
bel(A)	0.1	0.2	0.1	0.4	0.3	0.6	1.0
pl(A)	0.4	0.7	0.6	0.9	0.8	0.9	1.0

4.3 Dempster-Shafer Theory

Belief Intervals:

Belief intervals allow Demspter-Shafer theory to reason about the degree of certainty or certainty of our beliefs.

- A small difference between belief and plausibility shows that we are certain about our belief.
- A large difference shows that we are uncertain about our belief.

- However, even with a 0 interval, this does not mean we know which conclusion is right. Just how probable it is!

A	{B}	{J}	{S}	{B,J}	{B,S}	{J,S}	{B,J,S}
m(A)	0.1	0.2	0.1	0.1	0.1	0.3	0.1
bel(A)	0.1	0.2	0.1	0.4	0.3	0.6	1.0
pl(A)	0.4	0.7	0.6	0.9	0.8	0.9	1.0