# Einführung in die Dempster-Shafer-Theorie

## Umgang mit unsicherem Wissen

Dominik Elsbroek    Fynn Feldpausch

Universität Bremen, Fachbereich 3

14. Mai 2009

# Überblick

# Einführung

# Entwicklungsgeschichte

1960er Arthur Jeffrey **Dempster** verfasst Artikel zu einer Theorie für die Verrechnung von Evidenzen

# Entwicklungsgeschichte

1960er Arthur Jeffrey **Dempster** verfasst Artikel zu einer Theorie für die Verrechnung von Evidenzen

1970er Glenn **Shafer** entwickelt die Theorie weiter

# Entwicklungsgeschichte

1960er Arthur Jeffrey **Dempster** verfasst Artikel zu einer Theorie für die Verrechnung von Evidenzen

1970er Glenn **Shafer** entwickelt die Theorie weiter

1976 Veröffentlichung der Theorie
$\Rightarrow$ *The Dempster-Shafer Theory of Evidence*

# Kritik an der Bayes'schen Wahrscheinlichkeitstheorie

- Keine Unterscheidung zwischen…
  1. …dem Wissen, dass etwas nicht wahr
  2. …Unwissenheit

# Kritik an der Bayes'schen Wahrscheinlichkeitstheorie

- Keine Unterscheidung zwischen...
  1. ...dem Wissen, dass etwas nicht wahr
  2. ...Unwissenheit
- Verwendung einer einzigen Datenerhebungsgrundlage
  $\Rightarrow$ objektive Einschätzungen sind oft subjektive Meinungen

# Kritik an der Bayes'schen Wahrscheinlichkeitstheorie

- Keine Unterscheidung zwischen…
  1. …dem Wissen, dass etwas nicht wahr
  2. …Unwissenheit
- Verwendung einer einzigen Datenerhebungsgrundlage
  $\Rightarrow$ objektive Einschätzungen sind oft subjektive Meinungen
- keine Plausibilität von Wahrscheinlichkeiten

# Kritik an der Bayes'schen Wahrscheinlichkeitstheorie

- Keine Unterscheidung zwischen...
  1. ...dem Wissen, dass etwas nicht wahr
  2. ...Unwissenheit
- Verwendung einer einzigen Datenerhebungsgrundlage
  $\Rightarrow$ objektive Einschätzungen sind oft subjektive Meinungen
- keine Plausibilität von Wahrscheinlichkeiten
- keine Konfliktdarstellung unterschiedlicher Datenerhebungen

# Kritik an der Bayes'schen Wahrscheinlichkeitstheorie

- Keine Unterscheidung zwischen...
  1. ...dem Wissen, dass etwas nicht wahr
  2. ...Unwissenheit
- Verwendung einer einzigen Datenerhebungsgrundlage
  $\Rightarrow$ objektive Einschätzungen sind oft subjektive Meinungen
- keine Plausibilität von Wahrscheinlichkeiten
- keine Konfliktdarstellung unterschiedlicher Datenerhebungen
- vollständiges Wissen aller Wahrscheinlichkeiten notwendig

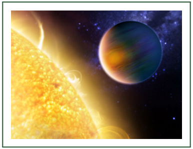
# Vorstellung des Beispiels

- ein fortlaufendes Beispiel durch den gesamten Vortrag
- Erläuterung der Konzepte anhand dieses Beispiels

# Vorstellung des Beispiels

- ein fortlaufendes Beispiel durch den gesamten Vortrag
- Erläuterung der Konzepte anhand dieses Beispiels

## Beispiel: *Leben in Siriusnähe?*



Gibt es erdähnliche Planeten in der Nähe des Sterns Sirius?
Gibt es Leben auf diesen Planeten?

# Theorie

# Der Betrachtungsrahmen (*frame of discernment*)

- Menge $\Theta$ sich gegenseitig ausschließender Elemente
- Interpretation der Elemente als Hypothesen, Aussagen

# Der Betrachtungsrahmen (*frame of discernment*)

- Menge $\Theta$ sich gegenseitig ausschließender Elemente
- Interpretation der Elemente als Hypothesen, Aussagen

**Beispiel:** *Leben in Siriusnähe?*

Festlegung verschiedener Aussagen:

A Es existieren Planeten und es existiert Leben.
B Es existieren Planeten, aber kein Leben.
C Es existieren keine Planeten.

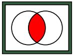
Betrachtungsrahmen $\Theta = \{A, B, C\}$

# Aussageklassen

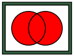
- Potenzmenge des Betrachtungsrahmens $2^{\Theta}$ ergibt hierarchisch angeordnete Aussageklassen

# Aussageklassen

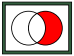
- Potenzmenge des Betrachtungsrahmens $2^{\Theta}$ ergibt hierarchisch angeordnete Aussageklassen
- Korrespondenz zwischen Aussagen und (Unter-)Mengen ermöglicht Zusammenhang zwischen logischen und mengentheoretischen Operatoren:
  - Konjunktion $(A \wedge B) \Rightarrow$ Schnitt $(A \cap B)$

  - Disjunktion $(A \vee B) \Rightarrow$ Vereinigung $(A \cup B)$

  - Negation $(\neg A) \Rightarrow$ Komplement $(\overline{A})$

# Aussageklassen

- Potenzmenge des Betrachtungsrahmens $2^\Theta$ ergibt hierarchisch angeordnete Aussageklassen

Beispiel: *Leben in Siriusnähe?*

Bedeutung verschiedener Aussageklassen:

Potenzmenge $2^\Theta = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \Theta\}$

# Aussageklassen

- Potenzmenge des Betrachtungsrahmens $2^{\Theta}$ ergibt hierarchisch angeordnete Aussageklassen

**Beispiel:** *Leben in Siriusnähe?*

Bedeutung verschiedener Aussageklassen:

{A} Es existieren Planeten und es existiert Leben.
{B} Es existieren Planeten, aber kein Leben.
{C} Es existieren keine Planeten.

Potenzmenge $2^{\Theta} = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \Theta\}$

# Aussageklassen

- Potenzmenge des Betrachtungsrahmens $2^\Theta$ ergibt hierarchisch angeordnete Aussageklassen

## Beispiel: *Leben in Siriusnähe?*

Bedeutung verschiedener Aussageklassen:

| | |
|---|---|
| {A} | Es existieren Planeten und es existiert Leben. |
| {B} | Es existieren Planeten, aber kein Leben. |
| {C} | Es existieren keine Planeten. |
| {A,B} | Entspricht $A \vee B \Rightarrow$ Es existieren Planeten. |
| {Θ} | Unwissenheit. |

Potenzmenge $2^\Theta = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A,B\}, \{A,C\}, \{B,C\}, \Theta\}$

# Die Wahrscheinlichkeitsfunktion (*mass function*)

- Funktion misst die Unterstützung für eine Aussage
- Werte stammen z.B. von
  - statistischen Erhebungen
  - Messungen
  - Expertenbefragungen
- Funktion muss bestimmte Kriterien erfüllen:

# Die Wahrscheinlichkeitsfunktion (*mass function*)

- Funktion misst die Unterstützung für eine Aussage
- Werte stammen z.B. von
  - statistischen Erhebungen
  - Messungen
  - Expertenbefragungen
- Funktion muss bestimmte Kriterien erfüllen:

## Theorem

*Ist $\Theta$ ein Betrachtungsrahmen, dann heißt eine Funktion $m : 2^{\Theta} \rightarrow [0,1]$* ***mass function****, wenn gilt:*

1. $m(\emptyset) = 0$
2. $\sum_{X \subseteq \Theta} m(X) = 1$

# Die Wahrscheinlichkeitsfunktion (*mass function*)

- Summe der Wahrscheinlichkeiten muss 1 ergeben
- Unwissenheit (*uncommitted belief*) wird $\Theta$ zugeordnet

# Die Wahrscheinlichkeitsfunktion (*mass function*)

- Summe der Wahrscheinlichkeiten muss 1 ergeben
- Unwissenheit (*uncommitted belief*) wird $\Theta$ zugeordnet

## Beispiel: *Leben in Siriusnähe?*

Mögliche Belegungen der Wahrscheinlichkeitsfunktion:

$$\begin{aligned} m_1(\{C\}) &= 0,6 \\ m_1(\{A, B\}) &= 0,3 \\ m_1(\{\Theta\}) &= 0,1 \end{aligned}$$

$$\begin{aligned} m_2(\{A\}) &= 0,8 \\ m_2(\{\Theta\}) &= 0,2 \end{aligned}$$

# Dempsters Kombinationsregel

- Regel zur Verrechung zweier Wahrscheinlichkeitsfunktionen
- ermöglicht so das (normalisierte) Zusammentragen verschiedener Wissensquellen
  $\Rightarrow$ z.B. die Einschätzungen verschiedener Experten

# Dempsters Kombinationsregel

- Regel zur Verrechung zweier Wahrscheinlichkeitsfunktionen
- ermöglicht so das (normalisierte) Zusammentragen verschiedener Wissensquellen
  $\Rightarrow$ z.B. die Einschätzungen verschiedener Experten
- Kombination durch:

$$m(X) = \begin{cases} 0, & \text{wenn } X = \emptyset \\ \frac{1}{1-m'(\emptyset)} \cdot m'(X), & \text{wenn } X \neq \emptyset \end{cases}$$

- Dabei gilt für $m'(X)$:

$$m'(X) = \sum_{A_i \cap B_j = X} m_1(A_i) \cdot m_2(B_j)$$

# Dempsters Kombinationsregel

## Beispiel: *Leben in Siriusnähe?*

1. Verrechung der Wahrscheinlichkeiten
   - Menge ergibt sich aus Schnittmenge
   - Wert ergibt sich durch Multiplikation

| | $m_1(\{C\}) = 0,6$ | $m_1(\{A,B\}) = 0,3$ | $m_1(\{\Theta\}) = 0,1$ |
|---|---|---|---|
| $m_2(\{A\}) = 0,8$ | | | |
| $m_2(\{\Theta\}) = 0,2$ | | | |

# Dempsters Kombinationsregel

## Beispiel: *Leben in Siriusnähe?*

1. Verrechung der Wahrscheinlichkeiten
   - Menge ergibt sich aus Schnittmenge
   - Wert ergibt sich durch Multiplikation

| | $m_1(\{C\}) = 0,6$ | $m_1(\{A,B\}) = 0,3$ | $m_1(\{\Theta\}) = 0,1$ |
|---|---|---|---|
| $m_2(\{A\}) = 0,8$ | $m'(\{\emptyset\}) = 0,48$ | $m'(\{A\}) = 0,24$ | $m'(\{A\}) = 0,08$ |
| $m_2(\{\Theta\}) = 0,2$ | $m'(\{C\}) = 0,12$ | $m'(\{A,B\}) = 0,06$ | $m'(\{\Theta\}) = 0,02$ |

# Dempsters Kombinationsregel

## Beispiel: *Leben in Siriusnähe?*

2. Summierung der Ergebnisse
   - Addition der Ergebnisse aus der Tabelle

$$\begin{aligned} m'(\emptyset) &= 0,48 \\ m'(\{A\}) &= 0,24 + 0,08 = 0,32 \\ m'(\{C\}) &= 0,12 \\ m'(\{A,B\}) &= 0,06 \\ m'(\Theta) &= 0,02 \end{aligned}$$

# Dempsters Kombinationsregel

## Beispiel: *Leben in Siriusnähe?*

3. Normalisierung der Wahrscheinlichkeiten
   - Anwenden der Kombinationsregel

$$\begin{aligned} m(\emptyset) &= 0 \\ m(\{A\}) &= 0,62 \\ m(\{C\}) &= 0,23 \\ m(\{A,B\}) &= 0,11 \\ m(\Theta) &= 0,04 \end{aligned}$$

# Dempsters Kombinationsregel

## Beispiel: *Leben in Siriusnähe?*

3. Normalisierung der Wahrscheinlichkeiten
   - Anwenden der Kombinationsregel

$$
\begin{aligned}
m(\emptyset) &= 0 \\
m(\{A\}) &= 0,62 \\
m(\{C\}) &= 0,23 \\
m(\{A,B\}) &= 0,11 \\
m(\Theta) &= 0,04
\end{aligned}
$$

- Wie groß ist die **gesamte** Unterstützung für $\{A, B\}$?

# Die Vertrauensfunktion (*belief function*)

- Funktion misst die **gesamte** Unterstützung für eine Aussage
- Funktion muss bestimmte Kriterien erfüllen

# Die Vertrauensfunktion (*belief function*)

- Funktion misst die **gesamte** Unterstützung für eine Aussage
- Funktion muss bestimmte Kriterien erfüllen

## Theorem

*Ist $\Theta$ ein Betrachtungsrahmen, dann heißt eine Funktion $Bel : 2^{\Theta} \to [0, 1]$* ***belief function****, wenn gilt:*

1. $Bel(\emptyset) = 0$
2. $Bel(\Theta) = 1$
3. *Für jede natürliche Zahl $n$ und jede Sammlung $A_1, \ldots, A_n$ von Untermengen von $\Theta$,*

$$Bel(X_1 \cup \cdots \cup X_n) \geq \sum_{I \subset \{1,\ldots,n\}, I \neq \emptyset} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} X_i\right)$$

# Die Vertrauensfunktion (*belief function*)

- Funktion misst die **gesamte** Unterstützung für eine Aussage
- Mögliche Funktionen sind:
  1. Addition (*Aufsammeln*) aller Teilmengen:

$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

# Die Vertrauensfunktion (*belief function*)

- Funktion misst die **gesamte** Unterstützung für eine Aussage
- Mögliche Funktionen sind:
  1. Addition (*Aufsammeln*) aller Teilmengen:

$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

  2. ausdruckslose Vertrauensfunktion (*vacuous belief function*):
     $\Rightarrow$ totale Unsicherheit
     $\Rightarrow$ es liegen keine Beweise vor

$$Bel(X) = \begin{cases} 1, & \text{wenn } X = \Theta \\ 0, & \text{wenn } X \neq \Theta \end{cases}$$

# Die Glaubhaftigkeitsfunktion (*upper probability*)

- Funktion misst die **maximale** Unterstützung für ein X
- daher auch maximale Wahrscheinlichkeit (*upper probability*) genannt.
- dient zur Ermittlung des Widerspruchs

# Die Glaubhaftigkeitsfunktion (*upper probability*)

- Funktion misst die **maximale** Unterstützung für ein X
- daher auch maximale Wahrscheinlichkeit (*upper probability*) genannt.
- dient zur Ermittlung des Widerspruchs

## Theorem

*Sei m eine mass function. Eine Funktion* $P*: 2^{\Theta} \rightarrow [0,1]$ *heißt* ***upper probability function****, wenn gilt:*

$$P^*(X) = \sum_{A \cap X \neq \emptyset} m(A)$$

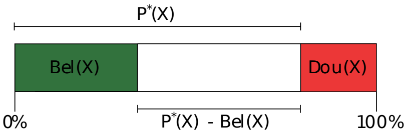
# Die Doubt-Funktion (*degree of doubt*)

## Theorem

*Sei Bel eine belief function, dann gilt:*

1. $P^*(A) = 1 - Dou(A)$
2. $Dou(A) = Bel(\overline{A})$

- drückt den Zweifel an einer Aussage aus



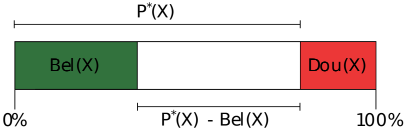
# Die Unwissenheit des System

## Theorem

*Ist Bel eine belief function und $P^*$, dann gilt:*

$$0 \leq Bel \leq P^* \leq 1$$

- vollständige Unwissenheit wird durch [0,1] gekennzeichnet
- Je breiter das Intervall, desto weniger vertrauenswürdig ist die Aussage. Je schmaler das Intervall, desto mehr Sicherheit bietet eine Aussage.



# Die Unwissenheit des System

## Beispiel: *Leben in Siriusnähe?*

- Berechnung der Unsicherheit:

| | $Bel$ | $P^*$ | $P^* - Bel$ |
|---|---|---|---|
| $\emptyset$ | | | |
| $\{A\}$ | | | |
| $\{B\}$ | | | |
| $\{C\}$ | | | |
| $\{A, B\}$ | | | |
| $\{A, C\}$ | | | |
| $\{B, C\}$ | | | |
| $\Theta$ | | | |

# Die Unwissenheit des System

## Beispiel: *Leben in Siriusnähe?*

- Berechnung der Unsicherheit:

| | $Bel$ | $P^*$ | $P^* - Bel$ |
|---|---|---|---|
| $\emptyset$ | 0 | 0 | 0 |
| $\{A\}$ | 0,62 | 0,62 + 0,11 + 0,04 = 0,77 | 0,15 |
| $\{B\}$ | 0 | 0,11 + 0,04 = 0,15 | 0,15 |
| $\{C\}$ | 0,23 | 0,23 + 0,04 = 0,27 | 0,04 |
| $\{A, B\}$ | 0,62 + 0,11 = 0,73 | 0,62 + 0,11 + 0,04 = 0,77 | 0,04 |
| $\{A, C\}$ | 0,62 + 0,23 = 0,85 | 0,62 + 0,11 + 0,23 + 0,04 = 1 | 0,15 |
| $\{B, C\}$ | 0,23 | 0,11 + 0,23 + 0,04 = 0,38 | 0,15 |
| $\Theta$ | 0,62 + 0,23 + 0,11 + 0,04 = 1 | 0,62 + 0,11 + 0,23 + 0,04 = 1 | 0 |

# Zusammenfassung der Theorie

# Funktionen und deren Zusammenhänge

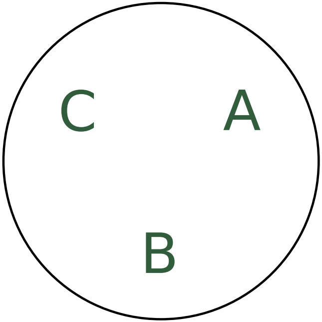
- Aussagen / Aussageklassen

C A

B

# Funktionen und deren Zusammenhänge

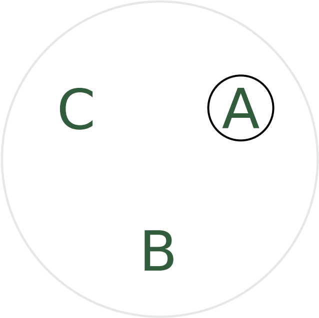
- Betrachtungsrahmen (*frame of discernment*)




- $\Theta = \{A, B, C\}$

# Funktionen und deren Zusammenhänge

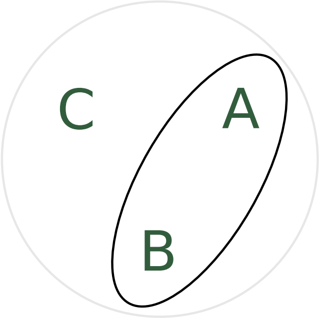
- Wahrscheinlichkeitsfunktion (*mass function*)




- $m(\{A\}) = 0,62$

# Funktionen und deren Zusammenhänge

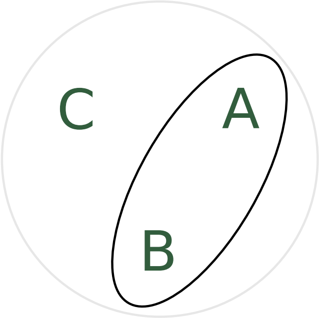
- Wahrscheinlichkeitsfunktion (*mass function*)




- $m(\{A\}) = 0,62$
- $m(\{A, B\}) = 0,11$

# Funktionen und deren Zusammenhänge

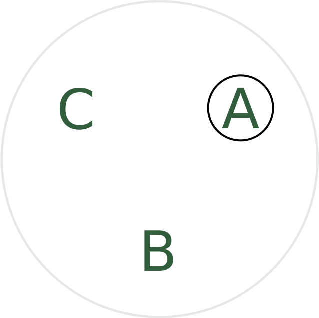
- Vertrauensfunktion (*belief function*)




$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

# Funktionen und deren Zusammenhänge

- Vertrauensfunktion (*belief function*)

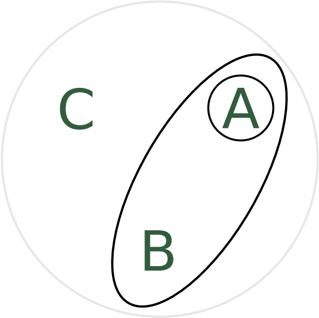



$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

- $Bel(\{A\}) = 0,62$

# Funktionen und deren Zusammenhänge

- Vertrauensfunktion (*belief function*)

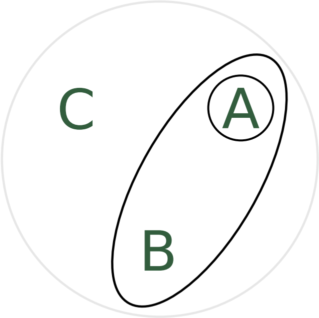



$$Bel(X) = \sum_{Y \subseteq X} m(Y)$$

- $Bel(\{A\}) = 0,62$
- $Bel(\{A,B\}) = 0,62 + 0,11 = 0,73$

# Funktionen und deren Zusammenhänge

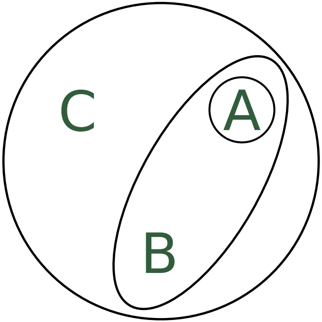
- Plausibilitätsfunktion (*upper probability function*)




$$P^*(X) = \sum_{Y \cap X \neq \emptyset} m(Y)$$

# Funktionen und deren Zusammenhänge

- Plausibilitätsfunktion (*upper probability function*)

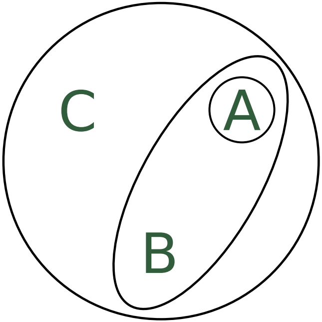



$$P^*(X) = \sum_{Y \cap X \neq \emptyset} m(Y)$$

- $P^*(\{A\}) = 0,62 + 0,11 + 0,04 = 0,77$

# Funktionen und deren Zusammenhänge

- Plausibilitätsfunktion (*upper probability function*)




$$P^*(X) = \sum_{Y \cap X \neq \emptyset} m(Y)$$

- $P^*(\{A\}) = 0,62 + 0,11 + 0,04 = 0,77$
- $P^*(\{A, B\}) = 0,62 + 0,11 + 0,04 = 0,77$

# Kritik an der Theorie

- durch fehlerhafte Abschätzungen entstehen noch immer falsche Ergebnisse

# Kritik an der Theorie

- durch fehlerhafte Abschätzungen entstehen noch immer falsche Ergebnisse
- viele verschiedene Interpretationen der *belief function* möglich

# Kritik an der Theorie

- durch fehlerhafte Abschätzungen entstehen noch immer falsche Ergebnisse
- viele verschiedene Interpretationen der *belief function* möglich
- Information wird bei der Normalisierung verworfen
  - $m(\emptyset)$ muss 0 sein
  - $m'(\emptyset)$ meist ungleich 0
    $\Rightarrow$ Normalisierung

# Quellen


- Glenn Shafer: *A mathematical Theory of Evidence*.
  Copyright © 1976 by Princeton University Press.
- Michael Höhl: *Dempster-Shafer-Theorie*.
  `www.rvs.uni-bielefeld.de/~mhoehl/dempster_shafer.ps`, 10.05.09


Vielen Dank für
eure Aufmerksamkeit