

Theorien für die Darstellung von Unsicherheit

**Ein Vergleich der Wahrscheinlichkeits-, Möglichkeits- und
Dempster-Shafer Theorie**

Johannes Leitner

Inhalt

- I Modellierung von Unschärfe
 - Unscharfe Mengen
 - Unscharfe Maße
- II Dempster-Shafer-Theorie
- III Wahrscheinlichkeitstheorie
- IV Möglichkeitstheorie
- V Zusammenfassung und Vergleich

Modellierung von Unschärfe

Zwei Arten unscharfer Information

- *Unpräzise* Information

- Inhaltliche unscharf, aber „wahr-oder-falsch-Aussagen“
- z.B. „Tim ist groß“
- Modellierung durch *Unscharfe Mengen*

- *Unsichere* Information

Thema dieses Vortrags

- Inhaltlich eindeutig, aber assoziiert mit „Wahrheitsgrad“
- z.B. „Möglicherweise ist Tim 194,23 cm groß“
- Modellierung durch *Unscharfe Maße*

Modellierung von Unschärfe

Erläuterung



- wir wissen genau, was die Flasche enthält
- was das bedeutet, ist unklar

Unpräzise Information



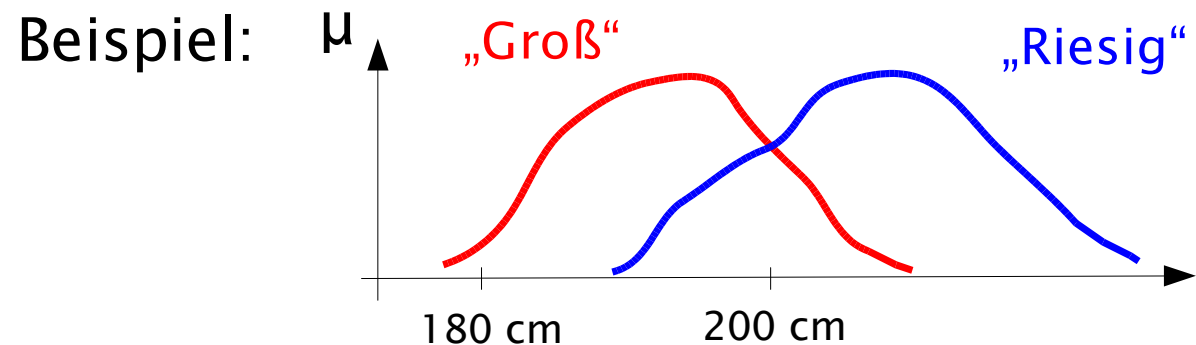
- die Flasche enthält entweder Gift oder nicht.
- das lässt sich aber nicht entscheiden

Unsichere Information

Unscharfe Mengen

Modellierung unpräziser Information

- Normale („scharfe“) Menge M und Funktion $\mu: M \rightarrow [0,1]$
- μ drückt den Grad der Zugehörigkeit aus.
- Mengenoperationen \cap und \cup definiert als $\min(\mu_1, \mu_2)$ bzw. $\max(\mu_1, \mu_2)$



Unscharfe Maße

Definition

- **Definition:** $M: 2^U \rightarrow [0,1]$, mit
 - $M(\emptyset) = 0$
 - *Monotonie:* $A \subseteq B \Rightarrow M(A) \leq M(B)$
 - *Stetigkeit:* $\lim M(X_i) = M(\lim X_i)$, mit $X_{i-1} \subseteq X_i$
[unwesentlich bei endlichen Grundmengen]
- $M_1(X) = 1 - M_2(\neg X) \Leftrightarrow$
 M_1 heisst **duales Maß** zu M_2
- **Ableitbare Gesetze:**
 $M(A \cap B) \leq \min \{ M(A), M(B) \}$
 $M(A \cup B) \geq \max \{ M(A), M(B) \}$

Unscharfe Maße

Erläuterung

- Also:
 - *Keine* Beziehung zwischen $M(x)$ und $M(\neg x)$
 - Zusammenhang zwischen $M(A)$, $M(B)$ und $M(A \cap B)$ wesentlich schwächer als bei Wahrscheinlichkeiten
- Haupt“vorteil“ der unscharfen Maße bei der Modellierung unscharfer Information.

Maßbasis


Konstruktion unscharfer Maße

- Wie kann man solche Maße konstruieren?

Idee: „Verteile das Gesamt-Zutrauen“
auf bestimmte Teilmengen X

Also $\sum m(X) = \text{const}$ (meist 1)
m heißt Maßbasis, die X mit
 $m(X) \neq 0$ fokale Elemente

		O		OO
1 to 18	1st 12	1	2	3
		4		6
		7	8	9
		10	11	12
	2nd 12	13	14	15
		16		18
		19	20	21
		22	23	24
	3rd 12	25	26	27
		28	29	
19 to 36		31	32	33
		34		36
		2 to 1	2 to 1	2 to 1

 = const

Dempster–Shafer–Theorie

Glaubwürdigkeit und Plausibilität

- Daraus werden unscharfe Maße durch:

- Glaubwürdigkeitsmaß $GI(X) := \sum_{Y \subseteq X} m(Y)$

„Summe der Ereignisse, die dafür sprechen“

- Plausibilitätsmaß $PI(X) := \sum_{X \cap Y \neq \emptyset} m(Y)$

„Summe der Ereignisse, die nicht dagegen sprechen“

- Es lässt sich zeigen:

$$GI(X) \leq PI(X)$$

Diese Maße sind zueinander dual: $GI(X) = 1 - PI(\neg X)$

$$GI(A \cup B) \geq GI(A) + GI(B) - GI(A \cap B)$$

$$PI(A \cup B) \leq PI(A) + PI(B) - PI(A \cap B)$$

Dempster-Shafer-Theorie

Beispiel

- Zwei binäre Sensoren von denen manchmal nur einer ein Ergebnis liefert:

	Beobachtung				N	m(A)	GI(A)	PI(A)
	(0,0)	(0,1)	(1,0)	(1,1)				
(0,-)	■	■			212	0,21		
(1,-)			■	■	128	0,13		
(-,0)	■		■		315	0,32		
(-,1)		■		■	151	0,15		
(0,0)	■				106	0,11		
(0,1)		■			55	0,06		
(1,0)			■		25	0,03		
(1,1)				■	8	0,01		

The diagram illustrates the aggregation of evidence from two sensors. The mass function values $m(A)$ are circled in red. Blue arrows point from $0,21$ and $0,11$ to $0,37$. Red arrows point from $0,21$, $0,13$, $0,32$, $0,15$, $0,11$, and $0,06$ to $0,84$.

Dempster–Shafer–Theorie

Dempsters Rule Of Combination

- Zwei Beobachtungen zusammenfassen:

- Geg. 2 Maßbasen m_1 und m_2

- Dann ist

$$m_1 \oplus m_2(X) := \frac{\sum_{A \cap B = X} m_1(A) m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B)}$$

- Dies nennt man *Dempsters Rule Of Combination*

Dempster-Shafer-Theorie

Schlussfolgern: Beispiel

- Zwei Zeugenbefragungen finden statt, über die Verdächtigen A, B, C

	{A,B}	{B,C}	{A,C}	{A,B,C}	{A}	{B}	{C}
$.m_1$	0.5	0.2	-	0.3	-	-	-
$.m_2$	-	-	-	0.4	0.6	-	-

⇒ Normalisierungsfaktor: $1 - 0.2 * 0.6 = 0.82$

	{A,B}	{B,C}	{A,C}	{A,B,C}	{A}	{B}	{C}
$.m_1 \oplus m_2$	0.23	0.09	-	0.14	0.54	-	-

Wahrscheinlichkeitstheorie

- Überlegung: Was ist wenn fokale Elemente \triangleq atomare Aussagen $\{x\}$?
 - Dann ist $PI=GI$.
 - $\Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - $\Rightarrow P(X) = 1 - P(\neg X)$
- \Rightarrow Es ist ein **Wahrscheinlichkeitsmaß!**
- Interpretation: Wissen ist „von maximaler Granularität“

Wahrscheinlichkeitstheorie

Kombinationen

- Geg. $P^1_x(u)$ und $P^2_x(u)$:

$$\text{Dann ist } P_x(u) := (P^1_x(u) + P^2_x(u)) / 2$$

- Geg. $P_x(u)$ und $P_y(v)$:

$$\text{Dann ist } P_{X,Y}(u,v) := P_x(u) P_y(v)$$

bei „voneinander unabhängigen“ X und Y

- Satz von Bayes

$$P(B|A) := P(B)P(A|B) / P(A)$$

Spezialfall der
Dempsters Rule



Wahrscheinlichkeitstheorie

Beispiel: Kombination

- Geg. $P_{\text{Wetter}}(u)$ und $P_{\text{Klausurnote}}(v)$

- $P_{\text{Klausurnote}}(u) =$

U	1,0	2,0	3,0	4,0	5,0
P(u)	0,2	0,3	0,3	0,1	0,1

- $P_{\text{Wetter}}(v) =$

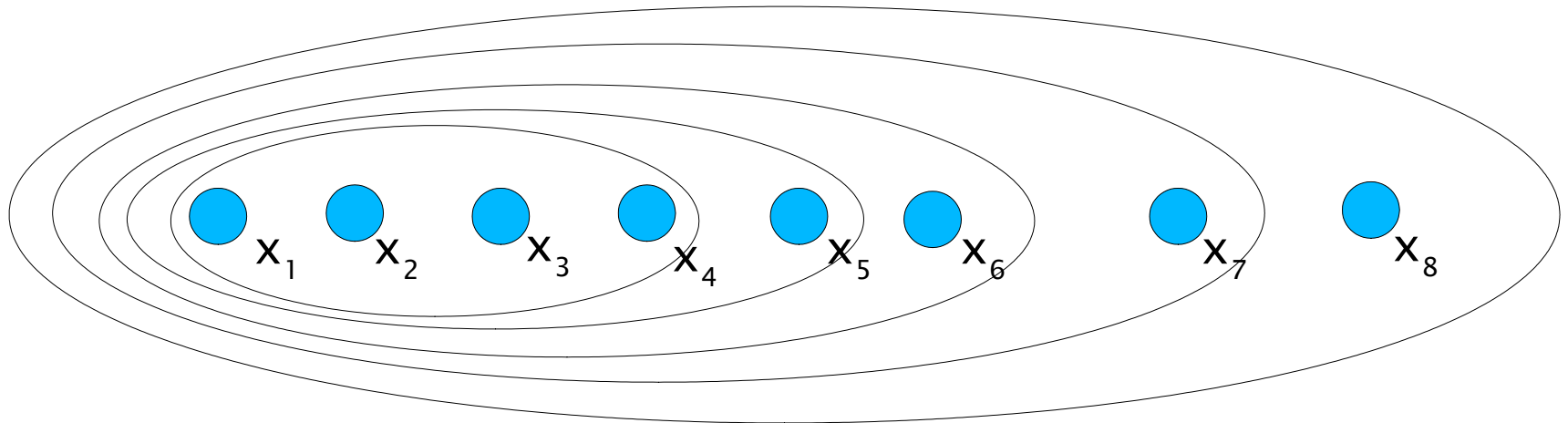
V	Regen	Sonnenschein
P(v)	0,7	0,3

- „ \Rightarrow “ $P_{\text{Note, Wetter}}(u,v) =$

(u,v)	(1,R)	(2,R)	(3,R)	(4,R)	(5,R)	(1,S)	(2,S)	(3,S)	(4,S)	(5,S)
P(u,v)	0,14	0,21	0,21	0,07	0,07	0,06	0,09	0,09	0,03	0,03

Möglichkeitstheorie

- Andere Überlegung: Die fokalen Elemente bilden eine „Inklusions-Kette“:



- Diese Maße nennt man Möglichkeitsmaß $\Pi(X)$ und Notwendigkeitsmaß $N(X)$
- Interpretation: „*Wissen ist komplett frei von Widersprüchen*“

Möglichkeitstheorie

- Eigenschaften:
 - $\Pi(X \cup Y) = \max \{ \Pi(X), \Pi(Y) \}$
 - $N(X \cap Y) = \min \{ N(X), N(Y) \}$
 - Π ist durch Werte auf „Atomen“ $\{x\}$ eindeutig festgelegt. Diese Einschränkung heißt *Möglichkeitsverteilung* $\pi(x) := \Pi(\{x\})$.
- Hier wird eine andere Art von Kombinationsmechanismen verwendet:

Möglichkeitstheorie

Kombinationsregeln

- Kombination

- $\pi^1_x(u), \pi^2_x(u)$

- $\Rightarrow \pi(u) := \min \{ \pi^1_x(u), \pi^2_x(u) \}$

- Bei unterschiedlichen Variablen dasselbe!

- Projektion

- $\pi_{x_1, \dots, x_n}(u_1, \dots, u_n)$

- $\Rightarrow \pi_{x_i}(u) := \max \{ \pi(\dots, u_i, \dots) \}$

- Das sind genau die *Standardmengenoperationen der Fuzzymengen!*

Möglichkeitstheorie

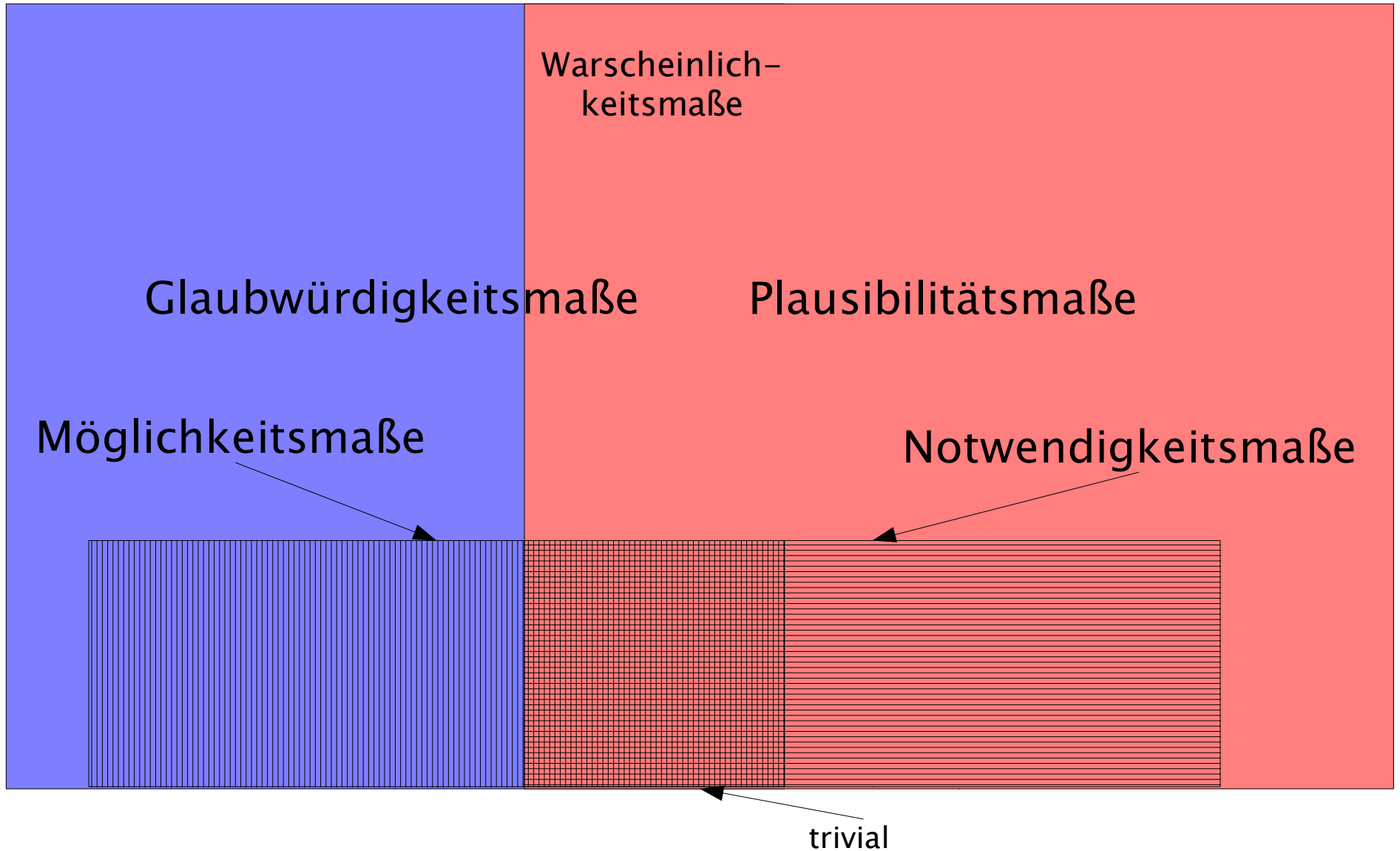
Schlussfolgern – Beispiel

- Geg:
 - $\pi_{W,H}(x,y)$ die Möglichkeit, daß ein Mensch x kg wiegt und y cm groß ist
 - $\pi_H(x)$ die Möglichkeit, daß Tim x cm groß ist
- Wieviel könnte Tim wiegen?
- $f(x) = \max_W \min \{ \pi_H(x), \pi_{W,H}(x,y) \}$
- Dies heißt *min-max-Komposition* oder Prinzip von Kombination & Projektion

Übersicht

	Dempster-Shafer	Möglichkeit	Wahrscheinlichkeit
Interpretation	Wie sehr wird $X=k$ <i>unterstützt</i> ?	Wie <i>möglich</i> ist es, daß $X=k$?	Wie häufig war $X=k$ bisher?
„ $F_X(k)=1$ “	„Ereignis wird voll unterstützt“	„Ereignis ist möglich“	„Ereignis ist sicher“
„ $F_X(k)=0$ “	„Keine Aussage treffbar“	„Ereignis ist nicht möglich“	„Ereignis kann nicht eintreten“
Unwissen	0	1	-
Regelwerk	Dempsters Rule Of Combination	Min-max-Komposition	Bayes Rule, etc.

Klassifikation



Interpretation

- Mit *Plausibilität* kann lückenhaftes, ungenaues Wissen modelliert werden
- Mit *Möglichkeit* widerspruchsfreies Wissen
- *Warscheinlichkeitstheorie* befasst sich mit vollständigem, eindeutig zugeordnetem Wissen

Anwendungen / Vergleich

- **Möglichkeitstheorie**
 - Wegen der engen Verwandtschaft alle Anwendungen der *fuzzy set theory*
 - Unscharfe Prädikate in menschlicher Sprache, KI
- **Dempster–Shafer–Theorie**
 - Meistbenutzte Methode in physikalischen Anwendungen
 - Steuerungen, Fehlertoleranz, Intervallanalyse, Sensorabgleich
- **Wahrscheinlichkeitstheorie**
 - Statistische Daten aus „zuverlässigen“ Quellen