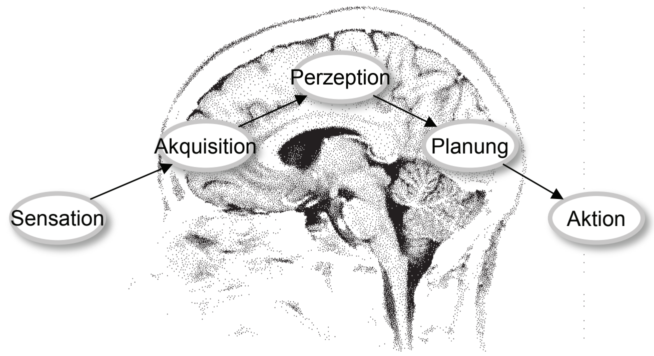
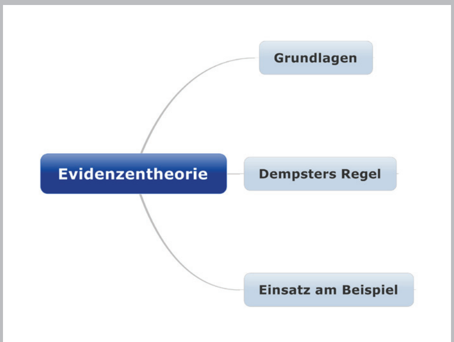
DHBW Stuttgart

Wissensbasierte Systeme



Kapitel 6

Evidenzen und Dempsters Regel



Prof. Dr. Dirk Reichardt

Vorlesungsunterlagen
Studiengang Angewandte Informatik

INHALTSVERZEICHNIS

VORWORT 2

GRUNDLAGEN UND DEFINITIONEN 2

DEFINITION : DIE ALTERNATIVENMENGE Ω 2

DEFINITION: DAS BASISMASS M 3

DEFINITION: DIE BELIEF-FUNKTION B 3

DEFINITION: DIE ZWEIFEL-FUNKTION ZW 3

DEFINITION:
DIE PLAUSIBILITÄTS-FUNKTION PL 3

DEMPSTERS REGEL 4

DEFINITION: KONFLIKT K 4

DEFINITION: DEMPSTERS REGEL 4

BEISPIEL: KRIMINALFALL 4

VORWORT

Neben der Wahrscheinlichkeitsrechnung und den daraus abgeleiteten Repräsentationsmethoden wie den Probabilistischen Netzen gibt es weitere, teils weiter gehende Repräsentatiosmethoden für unscharfe Informationen und Zusammenhänge.

Eine dieser Methoden ist die Modellierung von Evidenzen. Die Evidenztheorie betrachtet insgesamt unbekannte Sachverhalte für die jedoch objektiv richtige Teilinformationen gegeben sind. Dies können Beobachtungen sein, die gemacht werden. Aus diesen Beobachtungen soll nun auf ein Gesamtes geschlossen werden, welches nicht vollständig beobachtbar ist.

Eine Motivation für die Nutzung von Evidenzen ist ein beobachteter Nachteil der Wahrscheinlichkeitsmodelle. Dieser Nachteil wird wie folgt beschrieben: Nehmen Sie an, sie müssten eine Wahrscheinlichkeit für das Eintreten des Ereignisses „A“ angeben, haben jedoch keine Information über dieses Ereignis. Gleichzeitig sollen Sie eine Wahrscheinlichkeit für ein Ereignis „B“ angeben von dem Sie ebenso nichts wissen, nur dass es sich um das Komplementärereignis von „A“ handelt. Da es keinen Grund gibt ohne Information irgendeines der Ereignisse zu bevorzugen, ist eine Wahrscheinlichkeit von jeweils 0.5 anzunehmen. Das ist jedoch eine sehr hohe Wahrscheinlichkeit für etwas von dem nicht das Geringste bekannt ist! Was man hier also nicht modellieren kann ist das Unwissen. Genau an diesem Punkt setzt die Evidenztheorie an und erlaubt es nun gerade dieses Unwissen explizit in das Modell aufzunehmen.

GRUNDLAGEN UND DEFINITIONEN

Die grundlegende Aufgabe ist es, eine Entscheidung zu treffen. Vo-raussetzung für die Anwendbarkeit der Evidenztheorie ist eine ab-geschlossene Alternativenmenge aus der eine mögliche Lösung eines gegebenen Problems gewählt werden kann.

DEFINITION : DIE ALTERNATIVENMENGE Ω

Sei Ω = {A1, …, An} die Menge der Lösungsalternativen, dann gilt für Ω :

(1) Die Alternativenmenge ist vollständig

(2) Die Alternativen A1, … An schließen sich gegenseitig aus

Wie kann dies interpretiert werden? Ein beliebtes Beispiel ist die Lösung eines Kriminalfalls. Sie sind in der Rolle des Kommissars und haben eine Situation vor sich, wie in Agatha Christies „Zehn kleine Negerlein“, nämlich eine fest vorgegebene Menge von möglichen Tätern. Der Täter muss aus dem bekannten Kreis von Alternativen kommen. Zudem wird nun ausgeschlossen, dass es ein Gemeinschaftswerk war. Es war also genau einer. Damit sind die beiden obigen Bedingungen erfüllt. Sollte man durch Hinweise darauf schließen, dass keine der Alternativen zutrifft, so verlässt man die Grenzen des Systems. Dies kann nur durch „Bezweifeln" der Richtigkeit eines solchen Hinweises gelöst werden. Wird nun ein falscher Hinweis aus dem System entfernt, ist die Lösung wieder möglich. Dies entspricht etwa der Konsistenzprüfung von Mengen logischer Aussagen.

Bisher wurde von Hinweisen gesprochen, ohne diese formal zu definieren. In der Evidenztheorie wird ein Hinweis (Evidence) mit Hilfe eines so genannten Basismaßes definiert.

DEFINITION: DAS BASISMASS M

Sei Ω = {A1, …, An} die Menge der Lösungsalternativen, dann ist ein Basismaß m definiert als eine Abbildung m: Pot(Ω) →[0,1]. Für m gilt:

(1) $m(\Omega) = 0$ (2) $\sum_{X \subseteq \Omega} m(X) = 1$

Durch ein Basismaß wird jeder Teilmenge der Alternativenmenge ein Evidenzwert zugeordnet. Wie ist dies zu interpretieren? Nehmen wir wieder einen Kriminalfall an. Es gibt einen Hinweis, der besagt, dass sich der Mörder zum fraglichen Zeitpunkt im Salon aufhielt. Das macht alle Personen gleichermaßen verdächtig, die sich im Salon aufhielten. In anderen Worten: der Teilmenge der Alternativen (Personen im Salon) wird nun eine Evidenz zugeordnet. Wenn man sich absolut sicher ist, so wird diese Evidenz auf 1 gesetzt. Ist man sich nicht so sicher so wird das verbleibende Unwissen dem Basismaß für die Gesamtmenge der Alternativen (m(Ω)) zugeordnet. Der Wert von m(Ω) ist entsprechend als das Unwissen zu interpretieren.

Auf der Grundlage dieses Basismaßes lässt sich nun eine „Glaubensfunktion" realisieren, welche angeben soll, wie glaubwürdig eine Lösungsmenge bei Vorlage einer Messung (bzw. Beobachtung) ist.

DEFINITION: DIE BELIEF-FUNKTION B

Eine Glaubensfunktion (Belief-Funktion) B ist einem Basismaß m zugeordnet und ist definiert als eine Funktion B = Bm: Pot(Ω) → [0,1]. Dabei ist der Glaube an die Richtigkeit einer Teilmenge von Alternativen definiert durch die Formel:

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

Für die weitere Arbeit mit Evidenzen sind noch die Begriffe „Zweifel", „obere Wahrscheinlichkeit" und (auch zur einfacheren Handhabung) die „fokale Menge" einzuführen:

DEFINITION: DIE ZWEIFEL-FUNKTION ZW

Der Zweifel kann als der Glaube an die Richtigkeit einer beliebigen anderen Lösung interpretiert werden. D.h. man glaubt an die Komplementärmenge der Alternativen. Also wird der Zweifel Zw(X) definiert als Zw(X) = B(Ω \ X).

Die Wahrscheinlichkeitsrechnung ist in gewisser Weise wieder abbildbar, indem man eine so genannte Bayes-Belief-Funktion definiert. Diese hat die Eigenschaft B(X) + B(Ω \ X) = 1 für alle Teilmengen X. Was hier wegfällt ist genau das Unwissen, welches sich als „Glaube – Glaube an das Komplement" bestimmen lässt. Die obere Wahrscheinlichkeit wird in diesem Zusammenhang als B*(X) = 1 – Zw(X) definiert. In einer Bayes-Belief-Funktion entspricht die Funktion B gerade dieser oberen Wahrscheinlichkeit.

Ein weiterer Begriff, der im Sprachgebrauch häufig Verwendung findet, ist die Plausibilität. Wann gilt etwas als plausibel? Für unsere Zwecke soll die Plausibilität für eine Alternativenmenge X definiert sein als die Summe der Ereignisse, die nicht dagegen sprechen. Dies ist etwas weiter gefasst als die Glaubensfunktion, die alles sammelt, was dafür sprechen kann.

DEFINITION:
DIE PLAUSIBILITÄTS-FUNKTION PL

Eine Plausibilitätsfunktion Pl für ein gegebenes Basismaß m ist definiert als:

$$Pl(X) = \sum_{X \cap Y \neq \emptyset} m(Y) .$$

Der Begriff fokal wird für Mengen von Alternativen X verwendet, für die m(X) > 0 gilt. Bei der expliziten Definition eines Basismaßes beschränkt man sich in der Regel darauf die Werte für solche fokalen Mengen anzugeben. Diese Gesamtmenge aller fokalen Mengen eines Basismaßes wird auch als Kern bezeichnet.

Wir haben nun eine formale Möglichkeit geschaffen, aus Beobachtungen und Messungen eine Evidenz für Teilmengen einer vorgegebenen Alternativenmenge zu bestimmen. In unserem Kriminalfall entspricht das z.B. einer Beobachtung, z.B. einer Fußspur. Eine Fußspur am Tatort könnte auf den Täter hinweisen. D.h. alle Personen, deren Schuhe eine entsprechende Größe aufweisen werden etwas verdächtiger sein als andere.

DEMPSTERS REGEL

Nachdem nun ein einzelner Hinweis modelliert werden kann, ist es von Interesse, mehrere Hinweise zu kombinieren – die typische Aufgabe des Kommissars. Eine Möglichkeit dazu ist Dempsters Regel. Betrachten wir zunächst zwei Beobachtungen, d.h. zwei Zeugenaussagen, Messungen oder ähnliches. Wir erhalten zwei Basismaße m1 und m2, die für die gleiche Alternativenmenge Ω definiert sind.

Intuitiv können diese nun wie folgt verknüpft werden: Nehmen wir an Zeuge 1 sagt, er habe einen dunkelhaarigen Täter gesehen, Zeuge 2 behauptet, es sei wohl eine Frau gewesen. Insgesamt erhöht sich also die Evidenz für die Schnittmenge, d.h. für die dunkelhaarigen Frauen.

Allgemein ließe sich das wie folgt formulieren:

$$m_1 + m_2(Z) = \sum_{Z = X \cap Y} m_1(X) \cdot m_2(Y)$$

Es ergibt sich damit eine Evidenz für alle Alternativenmengen, die Schnittmenge zweier Elemente der fokalen Mengen der beiden Basismaße sind. In obigem Beispiel wäre eine Ausprägung von Z gerade die Menge der dunkelhaarigen Frauen, die sich aus dem Schnitt der Menge der Frauen (X) und der Menge der Dunkelhaarigen (Y) ergeben. Die Evidenz wird durch das Produkt der beiden Einzelevidenzen bestimmt.

Hier ist zu beachten, dass nur die Kerne der beiden Basismaße betrachtet werden müssen, da alle anderen Produkte Null als Ergebnis haben.

Nehmen wir nun an, dass es in der Alternativenmenge zwar Frauen gibt und auch Dunkelhaarige gibt, jedoch keine dunkelhaarigen Frauen! Was nun? Die beiden Aussagen passen also in der gegebenen Form nicht zusammen. Entweder hat sich einer der Zeugen getäuscht oder bewusst gelogen. Diese Konstellation nennt man einen Konflikt.

DEFINITION: KONFLIKT K

Der Konflikt zwischen zwei Basismaßen m1 und m2 wird bestimmt durch die Menge der Elemente der Kerne von m1 und m2 für die gilt: m1(X) > 0 und m2(Y) > 0 und X ∩ Y = Ø. Er berechnet sich als

K = . $\sum_{X \cap Y = \varnothing} m_1(X) \cdot m_2(Y)$

Sollte der Konflikt K = 1 sein, so sind die beiden Basismaße nicht vereinbar. In allen anderen Fällen kann trotzdem weiter gearbeitet werden. Allerdings enthält die Anfangs aufgestellte Formel zur Kombination von Evidenzen nun ein kleines Problem: sie weist der leeren Menge eine Evidenz zu! Diese ist nach Definition nicht zulässig. Dempsters Regel berücksichtigt dies durch eine Korrektur.

DEFINITION: DEMPSTERS REGEL

Die direkte Summe (Akkumulation) der Evidenzen zweier Basismaße m1 und m2 berechnet sich wie folgt:

(1) $m_1 \oplus m_2(\phi) = 0$

(2) $m_1 \oplus m_2(Z) = \left(\sum_{Z = X \cap Y} m_1(X) \cdot m_2(Y) \right) \cdot \frac{1}{1-K}$

Durch Akkumulation zweier Basismaße entsteht bei Anwendung von Dempsters Regel ein neues Basismaß. Dadurch ist eine Akkumulation beliebig vieler Basismaße zu einer Gesamtaussage möglich. Als Analogie können im Kriminalfall beliebig viele Zeugenaussagen, Messungen und Beobachtungen aufgenommen und verarbeitet werden, um den Täter schließlich einzukreisen.

BEISPIEL: KRIMINALFALL

Als Beispiel soll nun ein solcher Kriminalfall einmal durchgespielt werden. Als Täter kommen die Personen Markus, Thomas, David, Stephanie, Birgit und Heike in Frage. Zwei Zeugen stehen zur Verfügung um den Täter näher einzukreisen.

Doch starten wir ganz zu Anfang – ohne jede Information. Diese Situation kann durch das leere Basismaß m0 beschrieben werden. Das Basismaß m0 ist damit definiert als m0(Ω) = 1 und spiegelt das absolute Unwissen wider.

Kommen wir nun zu den Zeugen. Diese konnten durch Beobachtungen Aussagen über die mögliche Größe und Haarfarbe des Täters machen. Für die Alternativenmenge wird daher eine Übersicht angefertigt, welche diese Informationen enthält.

Person	Haarfarbe	Größe	Geschlecht
Markus	Blond	1,73 m	M
Thomas	Braun	1,80 m	M
David	Schwarz	1,76 m	M
Stephanie	Blond	1,76 m	W
Birgit	Blond	1,64 m	W
Heike	Braun	1,67 m	W

Beschreibung:

Zeuge 1 behauptet, dass der Täter dunkle Haare hatte. Schwarz oder Braun konnte er leider nicht unterscheiden. Auch ist er sich nicht absolut sicher, dass die gesehene Person dann auch wirklich der Täter war, da er diese nur vom Tatort weglaufen sah.

Diese Aussage muss nun in ein Basismaß umgewandelt werden. Wie sieht die fokale Menge (bzw. der Kern) aus? Man kann die Personen nur aufgrund von Haarfarbengruppen unterscheiden. Daher ergibt sich eine Gruppe {Thomas, David, Heike} für die eine Beobachtung gemacht wurde. Aufgrund der Unsicherheit des Zeugen wird dieser eine Evidenz zwischen 0 und 1 zugeordnet, die in diesem Fall – sagen wir – bei 0.6 liegt. Für das Unwissen bleibt 0.4 womit sich folgender Kern für m1 ergibt:

m1({Thomas, David, Heike}) = 0.6 m1(Ω) = 0.4

Die Akkumulation mit dem nichts sagenden Basismaß m0 ergibt trivialerweise wieder m1, so dass wir diesen Schritt an dieser Stelle überspringen. Ein zweiter Zeuge wird befragt.

Beschreibung:

Zeuge 2 behauptet, dass der Täter ca. 1,75-1,80 m groß gewesen sein muss. Er behauptet ein gutes Augenmaß zu haben und schließt andere Körpergrößen definitiv aus. Allerdings ist es auch hier nicht sicher, dass es sich bei der gesehenen Person tatsächlich um den Täter handelt.

Auch diese Aussage muss nun in ein Basismaß umgewandelt werden. Wie sieht die fokale Menge aus? Aufgrund der Attribute der Personen ergibt sich eine Gruppe {Thomas, David, Stephanie}. Die Evidenz für diese Aussage sei mit 0.7 festgelegt. Für das Unwissen bleibt somit 0.3 und folgender Kern für m2:

m2({Thomas, David, Stephanie}) = 0.7 m2(Ω) = 0.3

Nun sollen diese beiden Basismaße akkumuliert werden. Betrachtet man die für Dempsters Regel aufgestellte Formel, so wird klar, dass die entstehenden fokalen Mengen sich aus allen möglichen Kombinationen der beiden Kerne ergeben. Um die direkte Summe zu bestimmen kann man sich hilfsweise eine Tabelle der folgenden Form aufstellen:

	m_4({Birgit,Stephanie,Heike}) 0.4	$m_4(\Omega)$ 0.6
m_3({Thomas, David}) 0.42	m_5({}) 0.168	m_5({Thomas,David}) 0.252
m_3({Thomas,David,Heike}) 0.18	m_5({Heike}) 0.072	m_5({Thomas,David,Heike}) 0.108
m_3({Thomas,David, Stefanie}) 0.28	m_5({Stephanie}) 0.112	m_5({Thomas,David,Stepha 0.168
$m_3(\Omega)$ 0.12	m_5({Birgit,Stephanie,Heike}) 0.048	$m_5(\Omega)$ 0.072

Zur Probe können wir die Evidenzen der Elemente des Kerns aufsummieren um zu sehen, dass sich wieder der Gesamtwert 1 ergibt, wie es in der Definition des Basismaßes festgelegt wurde. Was ist zu beobachten?

- Das Unwissen sinkt, d.h. m3(Ω) < m2(Ω) und m3(Ω) < m1(Ω)
- Der Kreis der Verdächtigen wird teilweise enger gezogen : {Thomas, David}

Wie plausibel ist nun die Täterschaft von Thomas oder David? Und wie groß ist der Glaube an diese Täterschaft, wie groß dagegen der Zweifel? Diese Werte lassen sich wie folgt bestimmen:

Plausibilität :

$$\mathrm{Pl}(\{\text{Thomas, David}\}) = \sum_{\{Thomas,David\}\cap Y\neq\emptyset} m(Y) = 1$$

Glaube :

$$\mathrm{B}(\{\text{Thomas, David}\}) = \sum_{Y\subseteq\{Thomas,David\}} m(Y) = 0.42$$

Zweifel :

$$\mathrm{Zw}(\{\text{Thomas, David}\}) = \sum_{Y\subseteq\{Stefanie,Heike,Birgit,Markus\}} m(Y) = 0$$

Hier sieht man ein Beispiel für eine Belief-Funktion, die nicht nach Bayes definiert ist. Hier ist der Zweifel Null, d.h. es gibt kleine Hinweise auf einen anderen Täter, die nicht auch Thomas und David betreffen. Allerdings ist der Glaube noch lange nicht bei Eins.

Da taucht ein dritter Zeuge auf, der am Gang der Person erkannt haben will, dass es sich um eine Frau handelte. Diese Aussage wird zwar nur mit einer Evidenz von 0.4 versehen, könnte jedoch dennoch Änderungen in der Gesamtsituation mit sich bringen. Das zugehörige Basismaß lautet: m4({Birgit, Heike, Stephanie}) = 0.4 m4(Ω) = 0.6

Betrachtet man sich die Akkumulation von m4 und m3, so ergibt sich erstmals ein Konflikt K, da der Schnitt zwischen {Thomas, David} und {Birgit, Heike, Stephanie} leer ist. Daher muss mit einem Korrekturfaktor gerechnet werden. Betrachten wir zunächst wieder die Akkumulationsmatrix:

Diese Tabelle liefert noch nicht die korrekten Evidenzen. Um diese zu erhalten müssen die Werte noch mit dem Korrekturfaktor 1/(1-K) multipliziert werden, hier also mit 1/0.832 :

m5({Thomas, David}) = 0.252 / 0.832 = 0.3029

m5({Thomas, David, Heike}) = 0.108 / 0.832 = 0.1298

m5({Heike}) = 0.072 / 0.832 = 0.0865

m5({Stephanie}) = 0.112 / 0.832 = 0.1346

m5({Birgit, Stephanie, Heike}) = 0.048 / 0.832 = 0.0577

m5({Thomas, David, Stephanie}) = 0.168 / 0.832 = 0.2019

m5(Ω) = 0.072 / 0.832 = 0.0865

Zur Probe wollen wir noch einmal die Werte für Plausibilität, Glaube und Zweifel bestimmen. Diesmal ergeben sich folgende Werte:

Plausibilität :

$$Pl(\{Thomas, David\}) = \sum_{\{Thomas, David\} \cap Y \neq \emptyset} m(Y) = 0.7211$$

Glaube :

$$B(\{Thomas, David\}) = \sum_{Y \subseteq \{Thomas, David\}} m(Y) = 0.3029$$

Zweifel :

$$Zw(\{Thomas, David\}) = \sum_{Y \subseteq \{Stefanie, Heike, Birgit, Markus\}} m(Y)$$

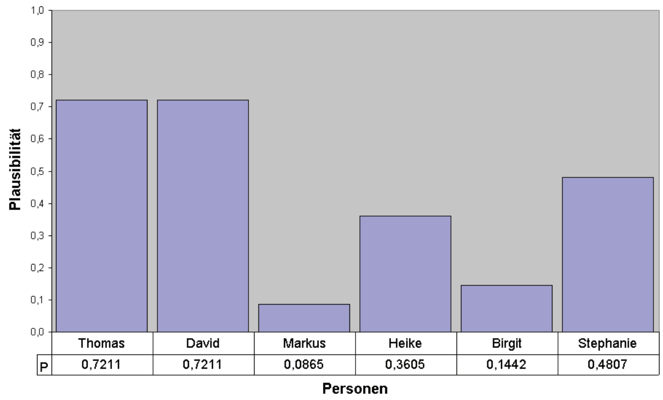
$$= 0.2788$$

Nach diesem Hinweis geht die Plausibilität von {Thomas, David} etwas runter, da der Hinweis die beiden Personen nicht enthält. Ebenso steigt der Zweifel an, so dass dieser nur noch knapp unter dem Glauben an die Täterschaft von {Thomas, David } liegt.

Da uns nun eigentlich Einzelpersonen interessieren, können wir uns auch die Werte Plausibilität, Glaube und Zweifel für alle Personen berechnen lassen und gegenüberstellen. In der nachfolgenden Grafik ist dies am Beispiel der Plausibilität nach den drei vorgelegten Zeugenaussagen durchgeführt worden.

Um noch einen weiteren Fall aufzunehmen, den es bei der Bestimmung von akkumulierten Evidenzen zu beachten gilt, sei das folgende Basismaß m6 gegeben, welches einen weiteren Hinweis darstellt:

m6({Thomas, Birgit, Stephanie}) = 0.7   m6(Ω) = 0.3



Akkumuliert man dieses Maß mit dem bisher berechneten, so ergibt sich ein nicht seltener „Sonderfall“, der in der nachfolgenden Tabelle aufgeführt wird. An dieser Stelle wird nur die für diesen Fall relevante Information in der Tabelle eingetragen, die restlichen Felder können Sie zur Übung ausfüllen.

	m_6({Thomas, Birgit, Stephanie}) 0.7	m_6(Ω) 0.3
m_5({Thomas, David}) 0.3029 m_5({Thomas, David, Heike}) 0.1298	m_7({Thomas}) 0.21 m_7({Thomas}) 0.09	
m_5({Heike}) 0.0865 m_5({Stephanie}) 0.1346		
m_5({Birgit, Stephanie, Heike}) 0.0577 m_5({Thomas, David, Stephanie}) 0.2019		
m_5(Ω) 0.0865		

Man sieht, dass m7 zwei Einträge für die Teilmenge {Thomas} enthält. In einem solchen Fall berechnet sich die Evidenz als Summe aller in der Tabelle befindlichen Teilevidenzen. D.h. in obigem Fall ergibt sich für m7({Thomas}) = 0.21 + 0.09 = 0.3.

Übungsaufgabe zum Verständnis:

1. Die Glaubensfunktion B und die Plausibilität Pl stehen in einem speziellen Verhältnis. Zeigen Sie dass **Pl(X) = 1 - B(Ω\X)** gilt!

2. Geben Sie eine Konstellation an, in der - analog zur Wahrscheinlichkeitsrechnung - B(X) = 1-B(Ω\X) gilt.

Weitere Informationen zu Evidenzen finden Sie u.a. hier:

[Richter 89] M. M. Richter, „Prinzipien der Künstlichen Intelligenz“, Teubner, Stuttgart, 1989