Intelligente Systeme

Unsicheres Wissen

Prof. Dr. R. Kruse   C. Braune

{kruse,cbraune}@iws.cs.uni-magdeburg.de

Institut für Wissens- und Sprachverarbeitung

Fakultät für Informatik

Otto-von-Guericke Universität Magdeburg

Motivation

In vielen Fällen ist unser Wissen über die Welt

unvollständig: **Wissensumfang ist nicht hinreichend**
z.B.: „Ich weiß nicht, wie der Bus an Feiertagen fährt, da ich nur werktags fahre!“

unscharf: **Wissen ist nicht präzise genug**
z.B.: „Der Bus fährt ungefähr zu jeder vollen Stunde!“

oder unsicher: **Wissen ist unzuverlässig**
z.B.: „Der Bus fährt wahrscheinlich um $9:00$ ab.“

Wir müssen trotzdem agieren!

Schließen unter Unsicherheit

Schließen über Eintretenswahrscheinlichkeiten

... und Kosten/Nutzen

Beispiel

Ziel: Um 9:15 Uhr in der Uni sein, um an einer Vorlesung teilzunehmen.
Es gibt mehrere **Pläne**, um das Ziel zu erreichen:

P_1: 8:00 Uhr aufstehen, 8:55 Uhr Haus verlassen, 9:00 Uhr Bus...

P_2: 7:30 Uhr aufstehen, 8:25 Uhr Haus verlassen, 8:30 Uhr Bus...

...

Alle Pläne sind *korrekt*, aber

sie implizieren verschiedene **Kosten** und verschiedene **Wahrscheinlichkeiten** Ziel *tatsächlich* zu erreichen.

P_2 am besten, denn Vorlesung ist wichtig, und Erfolgsrate bei P_1 liegt nur bei ca. 85–95%.

Grade der Überzeugung

- Wir (oder andere Agenten) sind von Regeln und Fakten nur bis zu gewissem Grad überzeugt
- Eine Möglichkeit, **Grad der Überzeugung auszudrücken**: Benutzung von **Wahrscheinlichkeiten**
- Die Aussage „Agent ist von Sensorinformation zu 0,9 überzeugt." bedeutet: Der Agent glaubt in 9 von 10 Fällen an Richtigkeit der Information
- Wahrscheinlichkeiten fassen „Unsicherheit" bedingt durch Unwissen zusammen
- Unsicherheit nicht mit *Unschärfe (Impräzision)* verwechseln ($\rightarrow$ Fuzzy-Logik)
- Das Prädikat *groß* ist **unscharf**, wohingegen die Aussage *„Das könnte Peters Uhr sein."* **unsicher** ist

Rationale Entscheidungen unter Unsicherheit

Verschiedene *Aktionen* oder *Pläne* zur Auswahl können mit verschiedenen *Wahrscheinlichkeiten* zu verschiedenen Ergebnissen führen

Aktionen verursachen verschiedene – ggf. subjektive – **Kosten**

Ergebnisse haben verschiedenen – ggf. subjektiven – **Nutzen**

Rational: wähle Aktion, die größten **zu erwartenden Gesamtnutzen** hat

Entscheidungstheorie = Nutzentheorie + Wahrscheinlichkeitstheorie

Entscheidungstheoretischer Agent

Algorithmus 1 DT-Agent

Eingabe: eine *Wahrnehmung*
Ausgabe: eine *Aktion*

1: static: eine Menge probabilistischer Überzeugungen über den Zustand der Welt
2: Berechne aktualisierte Wahrscheinlichkeiten für den aktuellen Zustand auf Basis vorliegender Evidenzen inkl. der aktuellen Wahrnehmung und vorheriger Aktionen
3: Berechne Ausgabewahrscheinlichkeiten für Aktionen, gegebene Aktionsbeschreibungen und Wahrscheinlichkeiten der aktuellen Zustände
4: Wähle *Aktion* mit höchstem erwarteten Nutzen, gegeben die Ausgabewahrscheinlichkeiten und Informationen über die Nützlichkeit
5: **return** *Aktion*

Axiomatische Wahrscheinlichkeitstheorie

Gegeben eine Ereignisalgebra $\mathcal{E}$, wird jedem Ereignis $E \in \mathcal{E}$ seine Wahrscheinlichkeit durch eine **Wahrscheinlichkeitsfunktion** $P: \mathcal{E} \to [0,1]$ zugewiesen.

P muss dabei die sogenannten **Kolmogorov Axiome** erfüllen:

- $\forall E \in \mathcal{E}: \ 0 \leq P(E) \leq 1$
- $P(\Omega) = 1$
- Für paarweise disjunkte Ereignise $E_1, E_2, \ldots \in \mathcal{E}$ gilt:

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

Aus diesen Axiomen lassen sich u.a. die folgenden Eigenschaften ableiten:

- $\forall E \in \mathcal{E}: \ P(\overline{E}) = 1 - P(E)$
- $P(\emptyset) = 0$
- Wenn $E_1, E_2 \in \mathcal{E}$ sich gegenseitig ausschließen, dann $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Wieso sind die Axiome sinnvoll?

- Wenn P objektiv beobachtbare Wahrscheinlichkeiten beschreibt (z.B. Würfelexperiment), dann machen die Axiome natürlich Sinn
- Aber wieso sollte ein Agent Axiome beachten, wenn er den *Grad seiner (subjektiven) Überzeugung* modelliert?
- Axiome schränken Menge der Überzeugungen ein, die Agent aufrechterhalten kann
- Eines der überzeugendsten Argumente, warum subjektive Überzeugungen den Axiome genügen sollten, stammt von de Finetti (1931):
  Verhält sich ein Agent in einer unsicheren Umgebung nicht gemäßden Axiomen, so hat er auf lange Sicht in der Umgebung Nachteile.

Unbedingte Wahrscheinlichkeiten

$P(A)$ bezeichnet **unbedingte** oder **a priori** Wahrscheinlichkeit, dass Ereignis A eintritt, falls *keine* zusätzliche Information verfügbar, wobei A eine Menge ist

Man kann Wahrscheinlichkeiten auch für Aussagen nutzen:

$$P(\text{Karies} = \text{wahr}) = 0.1, P(\text{Karies} = \text{falsch}) = 0.9$$

Karies ist eine binäre Zufallsvariable

Wahrscheinlichkeiten werden durch statistische Analysen oder aus allgemeinen Regeln gewonnen

Unbedingte Wahrscheinlichkeiten

Eine Zufallsvariable kann nicht nur die Werte wahr und falsch annehmen, sondern mehrere Werte

Zufallsvariable Wetter mit $P(\text{Wetter} = \text{sonnig})$ als Kurzform für $P(\{\omega \mid \text{Wetter}(\omega) = \text{sonnig}\})$

$$
\begin{aligned}
P(\text{Wetter} = \text{sonnig}) &= 0.7 \\
P(\text{Wetter} = \text{Regen}) &= 0.2 \\
P(\text{Wetter} = \text{bewölkt}) &= 0.08 \\
P(\text{Wetter} = \text{Schnee}) &= 0.02
\end{aligned}
$$

Verbundwahrscheinlichkeit I

$P(X)$: Vektor der Wahrscheinlichkeiten für (geordneten) Wertebereich der Zufallsvariable X:

$$\boldsymbol{P}(\text{Schmerz}) = \langle 0.1, 0.9 \rangle$$
$$\boldsymbol{P}(\text{Wetter}) = \langle 0.7, 0.2, 0.08, 0.02 \rangle$$

Vektoren definieren Wahrscheinlichkeitsverteilung der Zufallsvariablen *Schmerz* und *Wetter*

$P(\text{Schmerz}, \text{Wetter})$ ist (4×2)-Tabelle von Wahrscheinlichkeiten aller Kombinationen der Werte einer Zufallsvariablen

	Schmerz = wahr $\hat{=}$ S	Schmerz = falsch $\hat{=} \neg$ S
Wetter = sonnig	$P(\text{Wetter} = \text{sonnig} \wedge S)$	$P(\text{Wetter} = \text{sonnig} \wedge \neg S)$
Wetter = Regen	$P(\text{Wetter} = \text{Regen} \wedge S)$	$P(\text{Wetter} = \text{Regen} \wedge \neg S)$
Wetter = bewölkt	$P(\text{Wetter} = \text{bewölkt} \wedge S)$	$P(\text{Wetter} = \text{bewölkt} \wedge \neg S)$
Wetter = Schnee	$P(\text{Wetter} = \text{Schnee} \wedge S)$	$P(\text{Wetter} = \text{Schnee} \wedge \neg S)$

Verbundwahrscheinlichkeit II

Verbundwahrscheinlichkeitsverteilung $P(X_1, \ldots, X_n)$ weist jedem atomaren Ereignis eine Wahrscheinlichkeit zu:

	Schmerz=wahr	Schmerz=falsch
Karies=wahr	0.04	0.06
Karies=falsch	0.01	0.89

Da alle atomaren Ereignisse disjunkt, ist die Summe über alle Felder 1 (Disjunktion der Ereignisse)

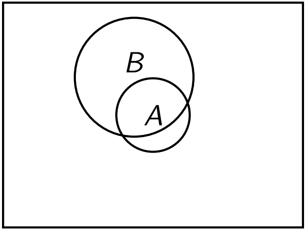
Alle interessanten Wahrscheinlichkeiten sind aus dieser Tabelle berechenbar, z.B.:

$$P(\text{Karies}{=}\text{wahr}) =$$

$$P(\text{Karies}{=}\text{wahr} \wedge \text{Schmerz}{=}\text{wahr}) + P(\text{Karies}{=}\text{wahr} \wedge \text{Schmerz}{=}\text{falsch})$$

Bedingte Wahrscheinlichkeiten I

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$



Produktregel: $P(A \wedge B) = P(A \mid B)P(B)$
analog: $P(A \wedge B) = P(B \mid A)P(A)$
A und B heißen **unabhängig** voneinander, falls $P(A \mid B) = P(A)$ und $P(B \mid A) = P(B)$
A und B unabhängig $\iff P(A \wedge B) = P(A)P(B)$

Interpretation bedingter Wahrscheinlichkeiten

Die Interpretation:

„$P(A \mid B) = 0.8$ bedeutet, dass $P(A) = 0.8$ ist
falls B wahr ist“

ist aus mindestens zwei Gründen falsch!

$P(A)$ ist immer die a priori Wahrscheinlichkeit, nicht die a posteriori Wahrscheinlichkeit $P(A \mid B)$, gegeben eine Evidenz (B wahr)

$P(A \mid B) = 0.8$ ist nur dann anwendbar, falls keine andere Evidenz außer B vorhanden ist! Wenn C bekannt ist, dann muss $P(A \mid B \wedge C)$ berechnet oder geschätzt werden. I.A. gilt $P(A \mid B \wedge C) \neq P(A \mid B)$; z.B. könnte $C \rightarrow A$ gelten (C impliziert A)

Bedingte Wahrscheinlichkeiten II

Neue Information kann die (bedingte) Wahrscheinlichkeit eines Ereignisses ändern.

Beispiel: Wahrscheinlichkeit von Zahnlöchern erhöht sich, wenn man weiß, dass Patient Zahnschmerzen hat.

Bei Zusatzinformation: nicht mehr mit a priori Wahrscheinlichkeiten rechnen!

$P(A \mid B)$ bezeichnet **bedingte** oder **a posteriori** Wahrscheinlichkeit von A, sofern alleinige Beobachtung (Evidenz) B gegeben.

$P(X \mid Y)$: Tabelle aller bedingten Wahrscheinlichkeiten über alle Werte von X und Y

Bedingte Wahrscheinlichkeiten III

Beispiel: $P(\text{Schmerz}, \text{Karies})$ ist die gemeinsame Wahrscheinlichkeitsverteilung von beiden Zufallsvariablen und besteht aus der folgenden Tabelle:

	S$\hat{=}$Schmerz = wahr	$\neg$S$\hat{=}$Schmerz = falsch
K$\hat{=}$Karies = wahr	$P(\text{K} \wedge \text{S})$	$P(\text{K} \wedge \neg\text{S})$
$\neg$K$\hat{=}$Karies = falsch	$P(\neg\text{K} \wedge \text{S})$	$P(\neg\text{K} \wedge \neg\text{S})$

Alle Zellen summieren sich zu 1 auf.

Die Bayes'sche Regel

Es gilt

$$P(A \wedge B) = P(A \mid B)P(B) \quad \text{und} \quad P(A \wedge B) = P(B \mid A)P(A)$$

Durch Gleichsetzen der rechten Seiten folgt:

$$P(A \mid B)P(B) = P(B \mid A)P(A)$$
$$\Longrightarrow P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Für mehrwertige Variablen erhält man mehrere Gleichungen:

$$\mathbf{P}(Y \mid X) = \frac{\mathbf{P}(X \mid Y)\mathbf{P}(Y)}{\mathbf{P}(X)}$$

Verallgemeinerung (bzgl. Hintergrundevidenzen):

$$\mathbf{P}(Y \mid X, E) = \frac{\mathbf{P}(X \mid Y, E)\mathbf{P}(Y \mid E)}{\mathbf{P}(X \mid E)}$$

Anwendung der Bayes'schen Regel I

$$P(\text{Schmerz}{=}\text{wahr} \mid \text{Karies}{=}\text{wahr}) = 0.4$$
$$P(\text{Karies}{=}\text{wahr}) = 0.1$$
$$P(\text{Schmerz}{=}\text{wahr}) = 0.05$$
$$P(\text{Karies}{=}\text{wahr} \mid \text{Schmerz}{=}\text{wahr}) = \frac{0.4 \cdot 0.1}{0.05} = 0.8$$

Warum nicht gleich $P(\text{Karies}{=}\text{wahr} \mid \text{Schmerz}{=}\text{wahr})$ schätzen?

Anwendung der Bayes'schen Regel II

Kausales Wissen wie $P(\text{Schmerz}=\text{wahr} \mid \text{Karies}=\text{wahr})$ ist i.A. robuster als diagnostisches Wissen $P(\text{Karies}=\text{wahr} \mid \text{Schmerz}=\text{wahr})$.

Wahrscheinlichkeit $P(\text{Schmerz}=\text{wahr} \mid \text{Karies}=\text{wahr})$ ist unabhängig von den a priori Wahrscheinlichkeiten $P(\text{Schmerz}=\text{wahr})$ und $P(\text{Karies}=\text{wahr})$.

Nähme $P(\text{Karies}=\text{wahr})$ bei Karies-Epidemie zu, so bliebe Kausalität $P(\text{Schmerz}=\text{wahr} \mid \text{Karies}=\text{wahr})$ unverändert, während sich $P(\text{Schmerz}=\text{wahr})$ proportional mit $P(\text{Schmerz}=\text{wahr})$ und $P(\text{Karies}=\text{wahr})$ änderte.

Ein Arzt, der $P(\text{Karies}=\text{wahr} \mid \text{Schmerz}=\text{wahr})$ geschätzt hätte, wüsste keine Regel zum Update.

Ein Arzt, der Bayes'sche Regel benutzt hätte, wüsste um proportionales Verhalten zwischen $P(\text{Schmerz}=\text{wahr} \mid \text{Karies}=\text{wahr})$ und $P(\text{Karies}=\text{wahr})$.

Normierung I

Wenn absolute Wahrscheinlichkeit von $P(K \mid S)$ zu bestimmen und $P(S)$ nicht bekannt, dann vollständige Fallanalyse durchführen (z.B. für K und $\neg K$) und Zusammenhang $P(K \mid S) + P(\neg K \mid S) = 1$ (hier boole'sche Variable) ausnutzen:

$$P(K \mid S) = \frac{P(S \mid K)P(K)}{P(S)}$$

$$P(\neg K \mid S) = \frac{P(S \mid \neg K)P(\neg K)}{P(S)}$$

$$P(K \mid S) + P(\neg K \mid S) = \frac{P(S \mid K)P(K)}{P(S)} + \frac{P(S \mid \neg K)P(\neg K)}{P(S)}$$

$$P(S) = P(S \mid K)P(K) + P(S \mid \neg K)P(\neg K)$$

Normierung II

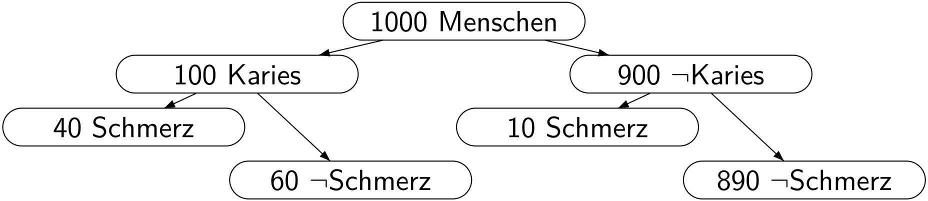
Durch Einsetzen in oberste Gleichung folgt

$$P(K \mid S) = \frac{P(S \mid K)P(K)}{P(S \mid K)P(K) + P(S \mid \neg K)P(\neg K)}$$

Bayes Theorem — mit ganzen Zahlen

$$P(\text{Schmerz} \mid \text{Karies}) = 0.4 \qquad P(\text{Karies}) = 0.1$$

$$P(\text{Schmerz} \mid \neg\text{Karies}) = \frac{1}{90}$$



$$P(\text{Karies} \mid \text{Schmerz}) = \frac{40}{40 + 10} = 0.8$$

Multiple Evidenzen I

Patient bejaht Frage nach Zahnschmerzen.
Aus dieser ersten Evidenz schließt Zahnarzt:

$$P(\text{Karies} \mid \text{Schmerz}) = 0.8$$

Zahnarzt ertastet bei Untersuchung mit Haken Loch im Zahn.
Aus dieser zweiten Evidenz schließt er:

$$P(\text{Karies} \mid \text{Loch}) = 0.95$$

$$P(\text{Karies} \mid \text{Schmerz} \wedge \text{Loch}) = \frac{P(\text{Schmerz} \wedge \text{Loch} \mid \text{Karies}) \cdot P(\text{Karies})}{P(\text{Schmerz} \wedge \text{Loch})}$$

Multiple Evidenzen

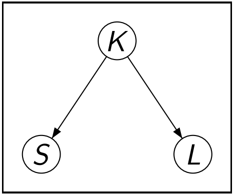
Problem: Er braucht $P(\text{Schmerz} \wedge \text{Loch} \mid \text{Karies})$, d.h. Diagnosewissen für alle Kombinationen von Symptomen. Besser: Evidenzen mit Hilfe der Evidenzregel schrittweise aufnehmen

$$P(Y \mid X, E) = \frac{P(X \mid Y, E) P(Y \mid E)}{P(X \mid E)}$$

Abkürzungen:

K — Karies, S — Schmerz

L — Loch



Ziel: Berechne $P(K \mid S, L)$ nur mit Hilfe von kausalen Aussagen der Form $P(\,\cdot \mid K)$ und unter Ausnutzung von Unabhängigkeiten zwischen den Variablen.

Multiple Evidenzen

A priori: $P(K)$

Evidenz Schmerz: $P(K \mid S) = P(K)\,\dfrac{P(S \mid K)}{P(S)}$

Evidenz Loch: $P(K \mid S, L) = P(K \mid S)\,\dfrac{P(L \mid K, S)}{P(L \mid S)}$

S und L unabhängig von K $\Leftrightarrow$ $P(L \mid K, S) = P(L \mid K)$

Multiple Evidenzen

Zusammengesetzte Gleichung von vorheriger Folie:

$$P(K \mid S, L) = P(K) \frac{P(S \mid K)\ P(L \mid K)}{P(S)\ P(L \mid S)}$$

$$= P(L) \frac{P(S \mid K)\ P(L \mid K)}{P(L, S)}$$

$P(L, S)$ ist eine Normalisierungskonstante und kann berechnet werden, wenn $P(L \mid \neg K)$ und $P(S \mid \neg K)$ bekannt sind:

$$P(L, S) = \underbrace{P(L, S \mid K)}_{P(L|K)P(S|K)} P(K) + \underbrace{P(L, S \mid \neg K)}_{P(L|\neg K)P(S|\neg K)} P(\neg K)$$

Multiple Evidenzen

$$P(K \mid L, S) = \frac{P(K)\ P(S \mid K)\ P(L \mid K)}{P(L \mid K)\ P(S \mid K)\ P(K) + P(L \mid \neg K)\ P(S \mid \neg K)\ P(\neg K)}$$

Man beachte, dass wir nur die bedingten Wahrscheinlichkeiten $P(\,\cdot \mid K)$ zusammen mit den Apriori-(Marginal-)Wahrscheinlichkeiten $P(K)$ and $P(\neg K)$ verwendet werden.

Multiple Evidenzen — Zusammenfassung

Mehrfache Evidenzen sind berechenbar durch Reduktion auf

a priori Wahrscheinlichkeiten und

(kausale) bedingte Wahrscheinlichkeiten für Evidenz.

Zusätzlich werden gewisse bedingte Unabhängigkeiten vorausgesetzt.

Allgemeine Kombinationsregel:

$$\boldsymbol{P}(Z \mid X, Y) = \alpha \boldsymbol{P}(Z)\boldsymbol{P}(X \mid Z)\boldsymbol{P}(Y \mid Z)$$

für X und Y bedingt unabhängig bei gegebenem Z und mit Normierungskonstante α

Zusammenfassung

Unsicherheit ist unvermeidbar in komplexen und dynamischen Welten, in denen Agenten zur Ignoranz gezwungen sind

Wahrscheinlichkeiten formulieren Unfähigkeit eines Agenten, definitive Entscheidung zu fällen und drücken Grad seiner Überzeugung aus

Falls ein Agent die wahrscheinlichkeitstheoretischen **Axiome** verletzt, so wird er unter Umständen irrationales Verhalten zeigen

Bayessche Regel ermöglicht es, unbekannte Wahrscheinlichkeiten aus bekannten Wahrscheinlichkeiten zu berechnen

Multiple Evidenzen können bei bedingter Unabhängigkeit effektiv in Berechnung einbezogen werden

Beispiel I

Expertenwissen:
„Metastatischer Krebs ist eine mögliche Ursache von Hirntumor, und ist auch eine Erklärung für erhöhte Kalziumwerte im Blut. Beide Phänomene zusammen können erklären, dass ein Patient ins Koma fällt. Schwere Kopfschmerzen sind auch möglicherweise verbunden mit einem Hirntumor."

Spezieller Fall:
"Der Patient hat schwere Kopfschmerzen."

Anfrage:
Wird der Patient ins Koma fallen?

Wahl des Zustandsraums

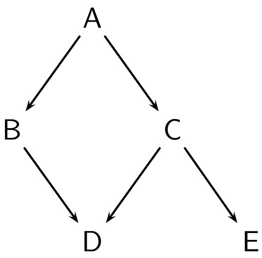
Ω ist endliche Menge von erschöpfenden, sich gegenseitig ausschließenden Werten

Beispiel:
Explizite Spezifikation der Zufallsvariablen und ihrer Werte

Zufallsvariable		**Mögliche Werte**
A	metastatischer Krebs	$a, \neg a$
B	erhöhter Calciumwert im Blut	$b, \neg b$
C	Hirntumor	$c, \neg c$
D	Koma	$d, \neg d$
E	schwere Kopfschmerzen	$e, \neg e$

Wahrscheinlichkeit für geg. Abhängigkeiten I

$$
\begin{aligned}
& P(A = a, B = b, C = c, D = d, E = e) \\
& = P(E = e \mid A = a, B = b, C = c, D = d) \\
& \cdot P(D = d \mid A = a, B = b, C = c) \\
& \cdot P(C = c \mid A = a, B = b) \\
& \cdot P(B = b \mid A = a) \cdot P(A = a) \\
& \\
& = P(E = e \mid C = c) \\
& \cdot P(D = d \mid B = b, C = c) \\
& \cdot P(C = c \mid A = a) \\
& \cdot P(B = b \mid A = a) \\
& \cdot P(A = a)
\end{aligned}
$$



Was muss vom Experten spezifiziert werden?

A priori Wahrscheinlichkeiten der Knoten ohne Eltern

Bedingte Wahrscheinlichkeiten eines Knotens bezogen auf Menge von elementaren Knoten

Wahrscheinlichkeit für geg. Abhängigkeiten II

$P(e \mid c) = 0.8$ $P(e \mid \neg c) = 0.6$	Kopfschmerzen sind häufig, aber viel häufiger, falls Tumor vorhanden ist
$P(d \mid b, c) = 0.8$ $P(d \mid b, \neg c) = 0.8$ $P(d \mid \neg b, c) = 0.8$ $P(d \mid \neg b, \neg c) = 0.05$	Koma ist selten, aber häufig, falls einer der beiden Gründe vorliegt
$P(b \mid a) = 0.8$ $P(b \mid \neg a) = 0.2$	erhöhtes Kalzium ist ungewöhnlich, aber eine häufige Konsequenz von Metastasen
$P(c \mid a) = 0.2$ $P(c \mid \neg a) = 0.05$	Hirntumor ist selten, aber häufige Konsequenz von Metastasen
$P(a) = 0.2$	Vorkommen von metastatischem Krebs in relevanten Studien

11 Werte (anstatt 31) spezifiziert, andere berechenbar

Wahrscheinlichkeiten anhand von Fallstudien, Lehrbuchinformationen und massiven Tests

Lösung des Problems

Gegeben:	Wahrscheinlichkeit P
Evidenz:	Patient hat starke Kopfschmerzen
Frage:	Wie groß ist die Wahrscheinlichkeit, dass der Patient ins Koma fallen wird?

Evidenz:	$E = e$
Bedingte Wahrscheinlichkeit:	$P(\bullet \mid E = e)$
Frage:	$D = d$?
Randwahrscheinlichkeit:	$P(D = d \mid E = e)$

Analyse dieses speziellen Falles

„Der Patient hat starke Kopfschmerzen“

Idee: Beschreibung der bedingten Wahrscheinlichkeiten $P(\bullet \mid E = e)$

$$P(A = a', B = b', C = c', D = d', E = e' \mid E = e)$$
$$= \begin{cases} \frac{P(A=a', B=b', C=c', D=d', E=e')}{\sum_{a \in A} \sum_{b \in B} \sum_{c \in C} \sum_{d \in D} P(a,b,c,d,e)} & \text{falls } e' = e \\ 0 & \text{falls } e' = \neg e \end{cases}$$

Auswertung der Daten des Patienten

Wird der Patient ins Koma fallen?

Idee: berechne Randwahrscheinlichkeiten $P(\{d\} \mid \{e\})$

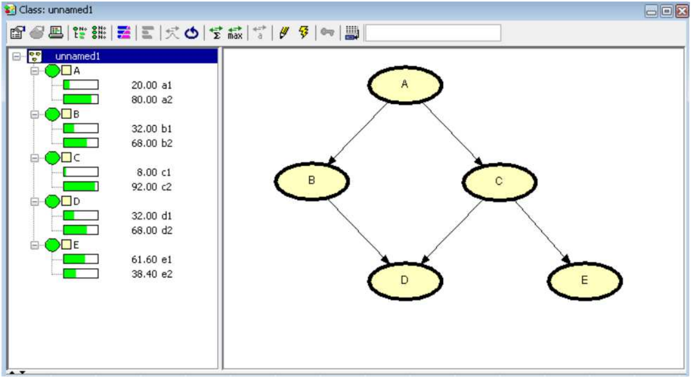
$$P(\{d\} \mid \{e\}) = \sum_{a' \in A} \sum_{b' \in B} \sum_{c' \in C} \sum_{e' \in E} P(a', b', c', d, e' \mid e)$$

Probleme:

Komplexität der Berechnungen der bedingten Randwahrscheinlichkeiten (100 Zufallsvariablen mit 3 Merkmalen: 3^{99} Summationen von kleinen Werten)

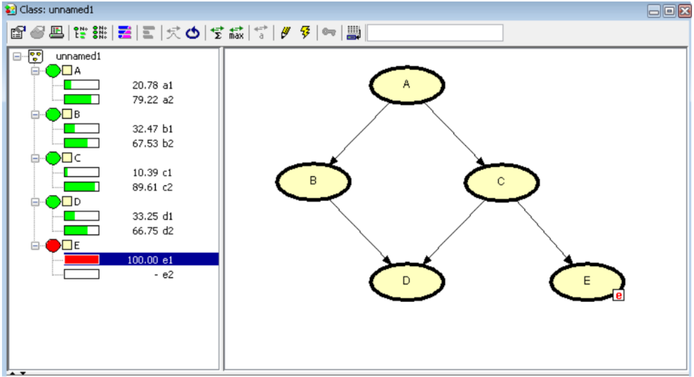
Können (oder müssen) wir erwarten, dass Nutzer mit Wahrscheinlichkeiten im Produktraum zurechtkommen?

Randwahrscheinlichkeiten im HUGIN-Tool



Bedingte Randwahrscheinlichkeiten mit Evidenz

$E = e_1$



Bayes'sche Netze

Ein Bayes'sches Netz dient dazu, die gemeinsame Wahrscheinlichkeitsverteilung aller beteiligten Variablen unter Ausnutzung bekannter bedingter Unabhängigkeiten möglichst kompakt zu repräsentieren. Dabei wird die bedingte (Un)abhängigkeit von Untermengen der Variablen mit dem A-priori-Wissen kombiniert.

Beispiel

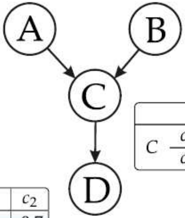
gerichteter, azyklischer Graph

Knoten bedeuten Zufallsvariable

Kanten symbolisieren (prob.) Abhängigkeiten

beinhalten **qualitative** und **quantitative** Informationen

		ϕ
A	a_1	0.4
	a_2	0.6



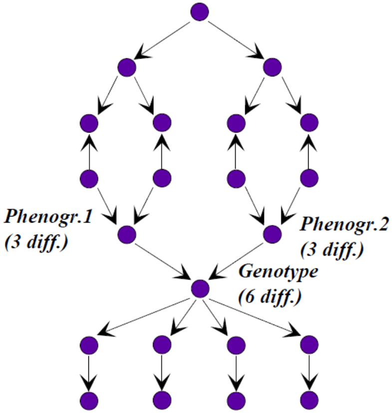
		ϕ
B	b_1	0.7
	b_2	0.3

		a_1, b_1	a_1, b_2	a_2, b_1	a_2, b_2
C	c_1	0.9	0.5	0.2	0.4
	c_2	0.1	0.5	0.8	0.6

		c_1	c_2
D	d_1	0.2	0.7
	d_2	0.8	0.3

Beispiel: Genotyp-Bestimmung des Jersey-Rinds

22 Variablen, $6 \cdot 10^{13}$ Zustände, 324 Parameter



graphisches Modell

Knoten: Zustandsvariablen

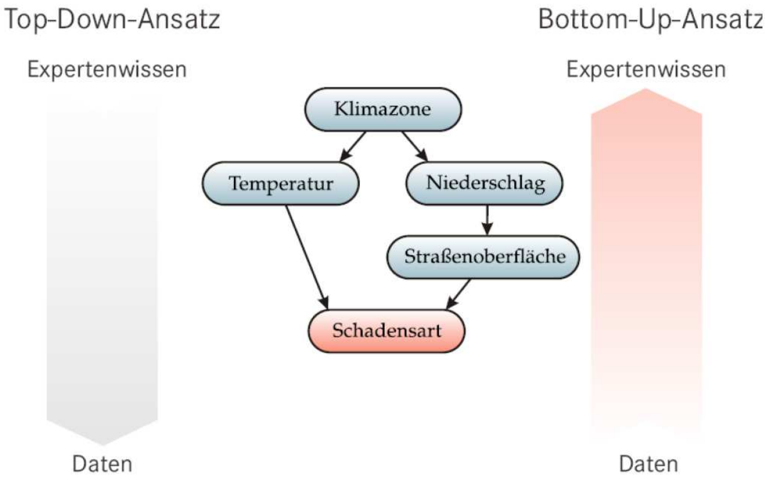
Kanten: Bedingte Abhängigkeiten

Zerlegung:

$$P(X_1, \dots, X_{22}) = \prod_{i=1}^{22} P(X_i \mid \text{Eltern}(X_i))$$

Diagnose: $P(\bullet \mid \text{Evidenz})$

Bayes'sche Netze



Beispiel IV: Planungssystem für Fahrzeugkonfigurationen

Etwa **200 item families** (Variablen, Merkmale)

Von **2 bis 50 items pro item family** (Werte pro Variable)

Mehr als 2^{200} mögliche Fahrzeug-Spezifikationen

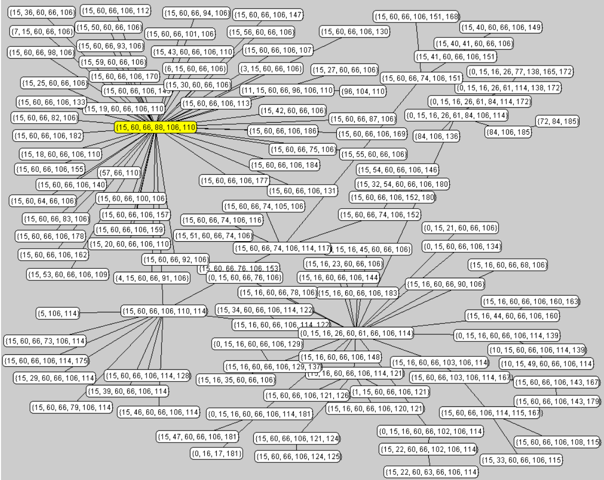
Wahl der gültigen Spezifikationen ist beschränkt durch **Regelsysteme** (**10.000 technische Regeln**, noch mehr marketing- und produktionsorientierte Regeln)

Beispiel:

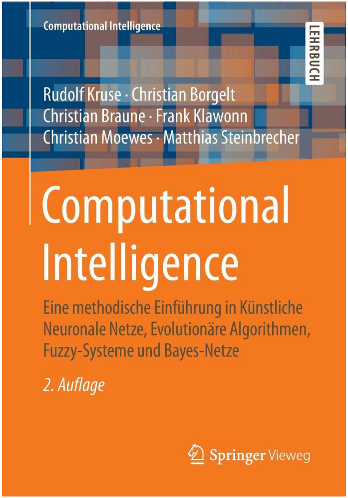
$$\textbf{if } \textit{engine} = e_1 \textbf{ then } \textit{transmission} = t_3$$

$$\textbf{if } \textit{engine} = e_4 \textbf{ and } \textit{auxiliary heater} = h_1 \\ \textbf{then } \textit{generator} \in \{g_3, g_4, g_5\}$$

Beispiel IV: Ungerichteter Graph der Abhängigkeiten der Merkmale



Literatur


Computational Intelligence
LEHRBUCH
Rudolf Kruse · Christian Borgelt
Christian Braune · Frank Klawonn
Christian Moewes · Matthias Steinbrecher
Computational Intelligence
Eine methodische Einführung in Künstliche Neuronale Netze, Evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze
2. Auflage
Springer Vieweg