Default-Typen und ihre Eigenschaften

Es kann von 0 bis beliebig viele Extensionen geben!

- Beispiele
  keine: $(\{:\neg\alpha / \alpha\}, \{\})$
  mehrere: $(\{:\alpha /\alpha,\ :\neg\alpha /\neg\alpha\}, \{\})$: $Th(\{\alpha\})$, $Th(\{\neg\alpha\})$

Sind *alle* Defaults einer Default-Theorie (D,W) von einer der folgenden Gestalten, so gelten die entsprechenden Aussagen (hier ohne Beweise):

- $\alpha : \beta / \beta$ (normale Defaults): Existenz von Extensionen sicher!
  (Außerdem Semimonotonie: Wenn $\mathcal{E}$ Extension von (D,W), dann gibt es Extension $\mathcal{E}'$ von $(D'\cup D,W)$ mit $\mathcal{E} \subseteq \mathcal{E}'$.)
- $:\beta / \beta$ (super-normale Defaults): Extensionen sind W -Theoreme von W plus max. konsistenten Teilmengen von D-Konklusionen
- $\alpha : \beta\wedge\gamma / \beta$ (semi-normale Defaults): Existenz v. Ext. unsicher!

Weitere Formalismen zum Normalfallschließen

- ***Circumscription*** („Umschreibung"):
  Präferiere Modelle, die ausgewählte Prädikate für möglichst wenige Objekte wahr machen. Beispiel:

  $$Vogel(x) \land \neg Abnormal_1(x) \Rightarrow Kann\_fliegen(x)$$

  $$Vogel(Tweety)$$

  erlaubt den Schluss $Kann\_fliegen(Tweety)$, es sei denn, es ist ableitbar, dass $Abnormal_1(Tweety)$

- **Inkonsistenztolerantes** Schließen:
  Definiere Ableitbarkeit bzgl. konsistenter Teilmengen aus einer im Allgemeinen inkonsistenten Prämissenmenge

Nichtwissensrepräsentation

- Die Effekte von Handlungen in der Welt unterliegen einer Vielzahl von Arten von Unsicherheit.
- Beispiel: Aktion (bzw. Plan) $F(n)$ bringe Agenten in n Minuten zum Flughafen. Unsicherheiten:
  - Begrenzte Wahrnehmbarkeit der Welt (Straßenzustand?)
  - Sensorrauschen (Verlässlichkeit von Staumeldungen?)
  - Ausnahmeeffekte in Aktionen (Platter? Kartoffel im Auspuff?)
  - Komplexität des zu modellierenden Weltausschnitts
- Logik-Ansätze zwingen dazu, sie „auszumodellieren“, wenn man sie berücksichtigen können will
- Das will man nicht, aber Unsicherheit soll repräsentierbar sein

Verfügbare Formalismen

- Allerlei ad-hoc-Ansätze („unscharfe“ Regeln, …):
  keine ordentliche Theoriebasis
- Default-Schließen: s. voriger Abschnitt:
  oft zu genau, zu feinteilig, zu mächtig, daher zu aufwändig
- *Fuzzy logic*: Modellierung von „gradueller Wahrheit“:
  Groß(x) ist absolut wahr (1,0) für *Peter* (der 2,01m groß ist), mittelwahr (0,4) für *Fritz* (1,75m), unwahr (0,05) für *Karl* (1,52m)
  andere Ontologie als Logik, daher andere Art Modellierung mit anders interpretierbaren Ergebnissen;
  in vielen Fällen erfolgreich angewendet
- Wahrscheinlichkeitstheoretische Ansätze:
  … um die es im Folgenden geht

Wahrscheinlichkeit (Sicht der KI)

… drückt (subjektive) Erwartung über die relative Häufigkeit von Fakten oder Zusammenhängen aus

… beruht auf der Logik-Ontologie, dass Dinge „objektiv“ „eigentlich“ genau wahr oder falsch sind

… modelliert unterschiedliche Formen von Nichtwissen (theoretisches, praktisches, pragmatisches) in demselben Maß

Beispiel: Korrelation zwischen Loch im Zahn und Zahnschmerzen

Weder ist korrekt $\forall p.\ Symptom(p, Zahnschmerz) \Rightarrow Krankheit(p, Zahnloch)$

noch $\forall p.\ Krankheit(p, Zahnloch) \Rightarrow Symptom(p, Zahnschmerz)$

Ob Patient ein Loch im Zahn hat und ob er Zahnschmerzen hat, ist feststellbar.

Genaue Zusammenhänge können unbekannt oder zu komplex sein.

Verwendung der W‘keitsinformation

Unterschiedliche Aktionen eines Agenten erzeugen mit unterschiedlichen W‘keiten Effekte.

Beispiel:

$P(F(25)$ bringt mich rechtzeitig zum Flug | …) = 0,04
$P(F(90)$ bringt mich rechtzeitig zum Flug | …) = 0,7
$P(F(120)$ bringt mich rechtzeitig zum Flug | …) = 0,95
$P(F(1440)$ bringt mich rechtzeitig zum Flug | …) = 0,9999

Und nun?

Handle so, dass Du den maximalen erwarteten Nutzen erzielst!
(s. Kapitel 1: Rationaler Agent)

Entscheidungstheorie = W‘keitstheorie + Nutzentheorie

Exkurs zum Weltbild von Logik-Verwendern

Logik verwenden heißt nicht, die Existenz und Relevanz von Unsicherheit zu leugnen!

Es heißt lediglich, damit anders umzugehen (kümmere Dich intensiver um die Standardfälle!)

Logik und W'theorie können sich perfekt ergänzen (z.B. *probabilistic logic*: definiere W'verteilungen auf Modellen logischer Theorien)

W'theorie Grundbegriffe (I)

Ergebnisraum Ω: Mögliche Ergebnisse eines Zufallsexperiments
z.B.: {„1", „2", „3", „4", „5", „6"} beim Würfeln
{[„1",*sonnig*], [„1",*regnet*], [„2",*sonnig*], …, [„6",*regnet*]}

Ereignis: Teilmenge von Ω. z.B.: {„2", „4", „6"} beim Würfeln („Gerade")

Elementarereignis: $\omega\in\Omega$. (eigentlich: einelementiges Ereignis)
z.B.: [„1",sonnig]

Zufallsvariable: Funktion von Elementarereignissen $\omega\in\Omega$ in einen Wertebereich $\mathcal{D}$. (*Notation*: beginnt mit Großbuchstabe)
$\mathcal{D}$={true, false}: Boolesche Zufallsvariable, z.B.: *Zahnloch*
$\mathcal{D}$ diskret: Diskrete ZV, z.B.: *Augenzahl* (auf einem Würfel)
$\mathcal{D}$ stetig (hier: $\Re$): kontinuierliche ZV

W'theorie Grundbegriffe (II)

Wahrscheinlichkeitsraum (Ω, P), $\Omega \neq \emptyset$: P ist eine auf Ereignissen aus Ω definierte reellwertige Funktion („Wahrscheinlichkeit"), sodass die Kolmogorow-Axiome gelten:

1. $0 \leq P(e) \leq 1$ für alle $e \subseteq \Omega$
2. $P(\Omega) = 1$
3. Für $d, e \subseteq \Omega$, sodass $d \cap e = \emptyset$: $P(d \cup e) = P(d) + P(e)$

z.B.: $P(Gerade) = P(\text{„2"}) + P(\text{„4"}) + P(\text{„6"}) = 1/6+1/6+1/6=1/2$

W'raum präziser (Ω, Σ, P) mit Ereignisraum Σ – hier verkürzt dargestellt)

Elementare Folgerungen aus den K.-Axiomen

- $P(\emptyset)=0$
- $P(\bigcup_i e_i)=\sum_i P(e_i)$ falls e_i paarweise disjunkt
- $P(\neg e) = 1-P(e)$ ($\neg$: Komplement oder bei booleschen Var‘n Negation)
- falls $e \subseteq f$, so $P(e) \leq P(f)$
- $P(e \cup f)= P(e) + P(f) - P(e \cap f)$
  bzw. für Boolesche Var‘n: $P(e \vee f)= P(e) + P(f) - P(e \wedge f)$

W'verteilungen und W'dichten

P induziert Funktionen auf den Zufallsvariablen:

Wahrscheinlichkeitsverteilung $\mathbf{P}$ einer diskreten Zufallsvariablen: Vektor aus W'keiten der möglichen Werte der Zufallsvariablen, wobei $\sum_\Omega \mathbf{P}(\omega)=1$

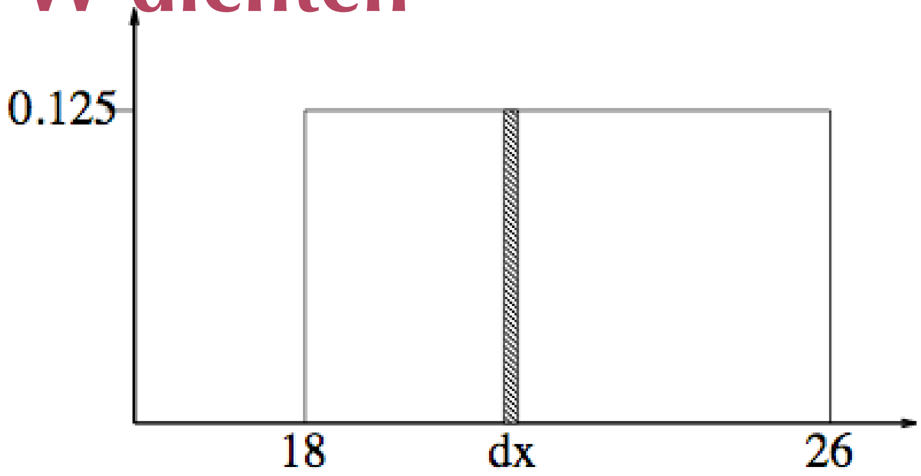
z.B.: $\mathbf{P}(Augenzahl) = \langle P(Augenzahl=„1“), \ldots, P(Augenzahl=„6“) \rangle = \langle 1/6, 1/6, 1/6, 1/6, 1/6, 1/6 \rangle$

Wahrscheinlichkeitsdichte $\mathbf{P}$ einer kontinuierlichen Zufallsvariablen: nicht-negative Fkt. auf dem Wertebereich der Zufallsvariablen, wobei $\int_\Omega \mathbf{P}(x)\mathrm{d}x = 1$

Es gilt: $P(a \leq X \leq b) = \int_a^b \mathbf{P}(x)dx$

z.B.: ...

Beispiele für W'dichten



P(*Temperatur*) = *U*[18,26] =
Gleichverteilung auf [18,26]

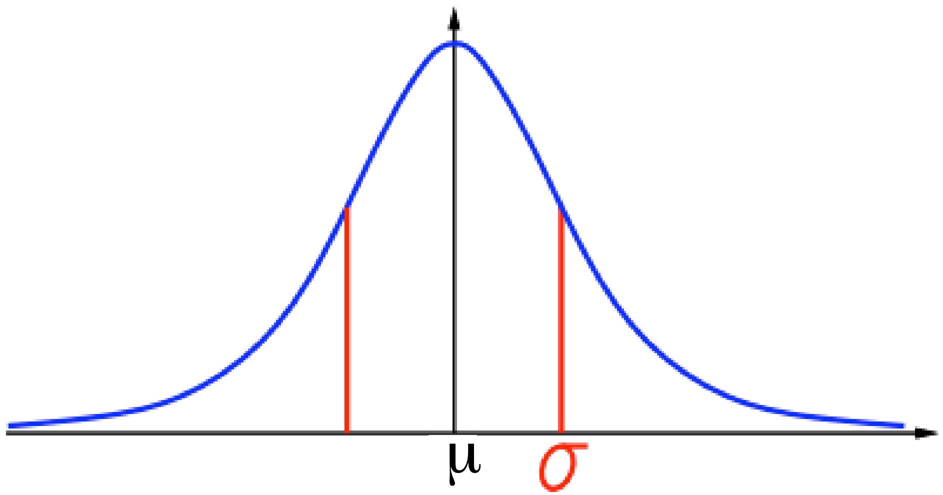
$$\int_{-\infty}^{\infty} U[18,26](x)\mathrm{d}x = 1$$

P(*Temperatur*=20,5) = 0,125

$$P(20 \leq Temperatur \leq 22) = \int_{20}^{22} U[18,26](x)dx = 0,25$$

Gaußverteilung mit Mittelwert μ
und Standardabweichung σ:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Unbedingte W‘keiten

W‘keit P „an sich“, ohne weitere Information:

Unbedingte Wahrscheinlichkeit (**a priori** W‘keit)

z.B.: $P(OSWetter{=}regnet) = 0{,}3$

Verteilung für unbedingte W‘keiten wie beschrieben

z.B.: $\mathbf{P}(OSWetter) = \langle P(OSWetter{=}sonnig), {=}wolkig, {=}regnet, {=}frostig\rangle$

$= \langle 0{,}25,\ 0{,}4,\ 0{,}3,\ 0{,}05\rangle$

Gemeinsame Verteilung von n ZVen entspricht n-dim. Matrix:

OSWetter= *Zahnloch=*	*sonnig*	*wolkig*	*regnet*	*frostig*
false	0,225	0,36	0,27	0,045
true	0,025	0,04	0,03	0,005

Vollständige gemeinsame Verteilung gibt W‘keit aller Elementarereignisse

Bedingte W'keiten

W'keit P bei gegebener Information:

Bedingte Wahrscheinlichkeit (**a posteriori** W'keit)

z.B.: $P(OSWetter{=}regnet \mid Luftdruck{=}hoch) = 0{,}1$

Definiert (falls $P(b)>0$) als $P(a \mid b) = \dfrac{P(a \cap b)}{P(b)}$

… oder als („Produktregel"): $P(a \cap b) = P(a|b)P(b) = P(b|a)P(a)$

z.B.: $P(sonnig|\neg zahnloch)=0{,}225/0{,}9=0{,}25$

entspricht $P(sonnig)$, weil unabhängig!

Entsprechend für W'verteilungen:

$$\mathbf{P}(OSWetter, Zahnloch) = \mathbf{P}(OSWetter|Zahnloch)\ \mathbf{P}(Zahnloch)$$

Inferenz über (vollständige) gemeinsame W'keiten

Angenommen, wir suchen:
die (vollständige) gemeinsame Posteriori-Verteilung
der Variablen **Y**, gegeben Werte **e** der Variablen **E** („Evidenz")

$$\mathbf{P}(\mathbf{Y}|\mathbf{E}=\mathbf{e}) = \alpha\mathbf{P}(\mathbf{Y}, \mathbf{E}=\mathbf{e}) = \alpha\sum_{\mathbf{h}}\mathbf{P}(\mathbf{Y}, \mathbf{E}=\mathbf{e}, \mathbf{H}=\mathbf{h})$$

(α Normalisierung, **H** verborgene Variablen)

Ginge über Ausrechnen aus der Matrix der (vollständigen) gemeinsamen Verteilung, aber …

… das sind ziemlich viele Gleichungen

… **woher sollen die ganzen W'keiten kommen??**

Unabhängigkeit

A und B unabhängig, gdw.

$\mathbf{P}(A|B) = \mathbf{P}(A)$ oder $\mathbf{P}(B|A) = \mathbf{P}(B)$ oder $\mathbf{P}(A,B) = \mathbf{P}(A)\,\mathbf{P}(B)$

z.B. $\mathbf{P}(Sonnig|Zahnloch) = \mathbf{P}(Sonnig)$

$\mathbf{P}(A,B,C,D) = \mathbf{P}(A,B,C)\,\mathbf{P}(D)$

Exponentielle Reduktion des Aufwandes
(Rechnen und Repräsentieren),

… doch volle Unabhängigkeit ist selten!

Bedingte Unabhängigkeit

Bspl: Heiße *Zahnloch*: Ich habe ein Loch im Zahn
Schmerz: Ich habe Zahnschmerzen
Haken: Der Stahlhaken beim Zahnarzt greift

↪ gemeinsam *nicht unabhängig*!

Doch es gilt:

$$\mathbf{P}(Haken \mid Schmerz, Zahnloch) = \mathbf{P}(Haken \mid Zahnloch)$$

(ob der Haken greift, hängt nicht vom Zahnschmerz ab)

Haken ist **bedingt unabhängig** von *Schmerz*, gegeben *Zahnloch*

entsprechend:

$$\mathbf{P}(Schmerz \mid Haken, Zahnloch) = \mathbf{P}(Schmerz \mid Zahnloch)$$

$$\mathbf{P}(Schmerz, Haken \mid Zahnloch) = \mathbf{P}(Schmerz \mid Zahnloch)\mathbf{P}(Haken \mid Zahnloch)$$

↪ **Exponenzielle Reduktion trotz genereller Abhängigkeit!!**

Die Bayessche Regel (Satz von Bayes)

Erinnerung Produktregel: $P(a \cap b) = P(a|b)P(b) = P(b|a)P(a)$

Satz von Bayes

$$P(b \mid a) = \frac{P(a \mid b)P(b)}{P(a)}$$

… oder für Verteilungen:

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

… oder bei vorhandener Evidenz **e**:

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

Anwendung von Bayes: Diagnoseregeln

$$P(Ursache \mid Wirkung) = \frac{P(Wirkung \mid Ursache)P(Ursache)}{P(Wirkung)}$$

Oft sind kausale Zusammenhänge besser bekannt als diagnostische

Bedingte Unabhängigkeit und die Bayessche Regel

$$\mathbf{P}(Zahnloch, Schmerz, Haken)$$
$$= \alpha\, \mathbf{P}(Schmerz, Haken \mid Zahnloch)\, \mathbf{P}(Zahnloch)$$
$$= \alpha\, \mathbf{P}(Schmerz \mid Zahnloch)\, \mathbf{P}(Haken \mid Zahnloch)\, \mathbf{P}(Zahnloch)$$

allgemeiner: **Naives Bayessches Modell**:

$$\mathbf{P}(Ursache, Effekt_1, \ldots, Effekt_n)$$
$$= \alpha\, \mathbf{P}(Ursache) \prod_i \mathbf{P}(Effekt_i \mid Ursache)$$

Notation in **Bayes-Netzen**: