Wissensbasierte Systeme Semester 6 Skript

Anmerkung:

In der Aussagen- und Prädikatenlogik wurden folgende Symbole für Operatoren verwendet:

∀ als Allquantor
∃ als Existenzquantor
& als logisches UND

Inhaltsverzeichnis

1. Einführung in KI (künstliche Intelligenz) und XPS (Expertensysteme)

1.1. <u>Informations- und Wissensverarbeitung</u>

1.1.1. <u>Informationsmodelle</u>

<u>Nachrichten technische Information (Shannon, 1949).</u>

Es handelt sich um rein mathematische Definition. Im Wesentlichen drei Anforderungen:

- Je kleiner die Wahrscheinlichkeit p(x) für ein Ereignis x, um so größer ist der Informationsgehalt der Nachricht
- Nachricht x mit p(x) = 1 bedeutet Informationsgehalt I(x) = 0
- Der Informationsgehalt von unabhängigen Ereignissen ist additiv : $I(x_1,x_2) = I(x_1) + I(x_2)$

Der Ausdruck $I(x) = \ln(1 / p(x))$

Entropie ist $H := p(x_i) * \ln(1 / p(x_i))$ = Gesamtinformation

Die Definition hat folgende Merkmale:

- Es handelt sich um eine reine Informationsmenge
- Es gibt keine semantische Bedeutung
- Es gibt weder wahr noch falsch
- Es gibt kein neues Wissen

<u>Informationswissenschaftliche Definition (Wersig)</u>

Außenwelt
Rezeptoren (Sinnesorgane)
Internes Außenweltmodell (Sichtweisen)
Innenwelt

Wissen:	Struktur des internen Außenweltmodells.
Denken:	Logische Operationen im internen Außenweltmodell.
Redundanz:	Information ohne Korrektur des internen Außenweltmodells.
Ungewissheit:	Keine Lösung für ein Problem im internen Außenweltmodell.
↓	
Information:	Reduktion der Ungewissheit

1.1.2. Wissensmodelle (Rauch, Meadow, 1992)

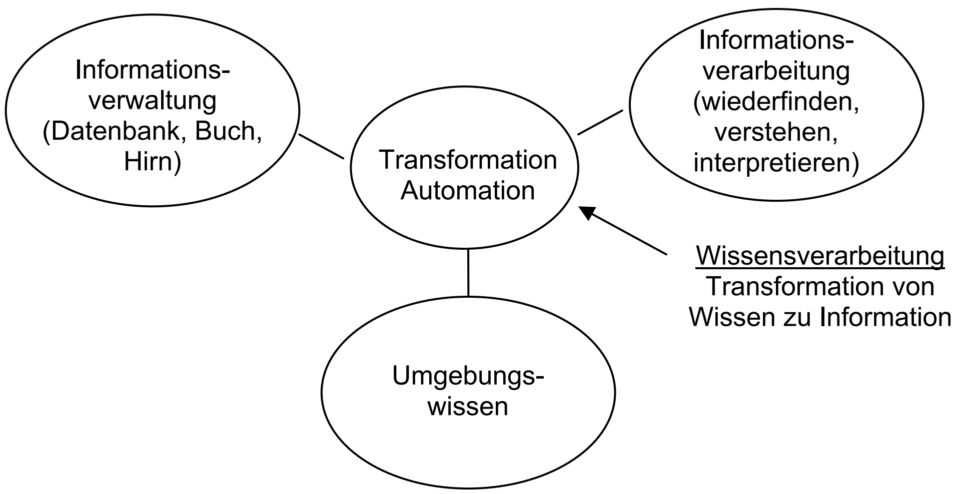
Daten: Menge von Symbolen $\{x_1, \ldots, x_n\}$ mit festgelegter Syntax.

Informationen: Daten, die eine Zustandsänderung im Empfängersystem (Mensch, Computer) generieren.

Informationen haben Eigenschaften:
- Neuigkeitswert
- Kontextabhängigkeit
- Zeitabhängigkeit
- Redundanz mit Evidenzverstärkung bzw. Evidenzabschwächung

Wissen: Die im System gespeicherte Informationsmenge mit Eigenschaften.
- wahr oder falsch
- Wahrheitswert variabel durch neue Informationen
- Widerspruch möglich

1.1.3. Wissensverarbeitung



Sender → Speicher → Empfänger

Dokumente
Internet
Datenbanken
Massenmedien

1.2. Einsatz von Methoden der KI und XPS

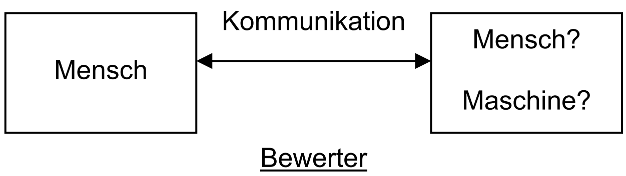
1.2.1. Was ist KI

AI Artificial Intelligence
„Moderne Informationsverarbeitung“

Paradigma: KI beschäftigt sich mit Simulation und Nachahmung menschlicher kognitiver Fähigkeit. Ziel ist die effiziente Nachbildung der Struktur intelligenter Leistung durch Computerprogramme.

Beispiel:

a) Turing Test



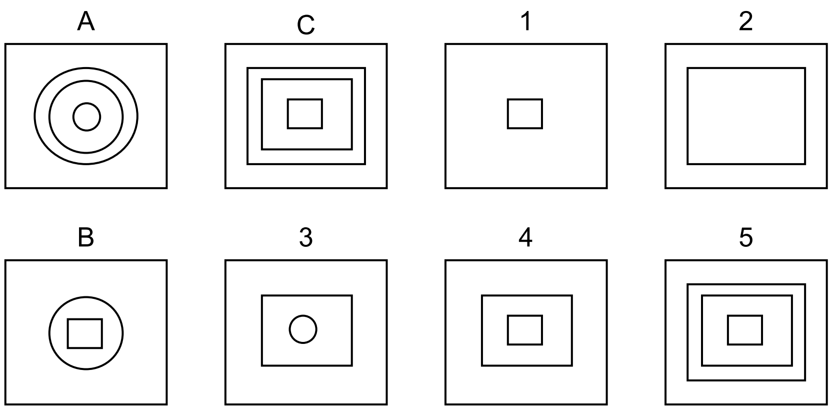
b) Zahlenfolgen

2 8 26 80 242 728 $x_{neu} = 3 * x_{alt} + 2$

2 3 5 7 11 15 23 alternierend $x_{neu} = 2 * x_{alt} + 1$

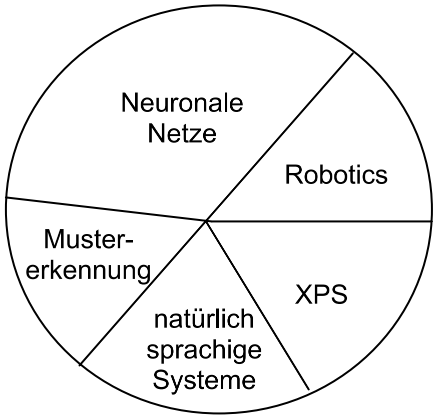
2 1 4 4 8 7 16 20 32 13 $x_{gerade} + 3$, $x_{ungerade} * 2$
x_1 x_2 x_3 x_4 …

c) Verhalten



Es gilt A zu B
wie C zu 3

1.2.2. Einsatzgebiete



Robotics
Nachahmung der menschlichen Motorik.
Steuerung von technischen Prozessen.

Natürlich sprachige Systeme
Erkennung, Analyse, Interpretation von natürlicher Sprache.
(z.B. Auskunftsysteme)

Mustererkennung
Erkennung, Analyse von typischen Datenstrukturen (Mustern).
- optische Mustererkennung (Bilderkennung) (2 dimensionale Mustererkennung)
- 1 dimensionale Mustererkennung (akustische Mustererkennung, Signalanalyse, EKG, Sprachanalyse, Sprechererkennung)
- 3 dimensionale Mustererkennung (Bewegungsverfolgung, Szenenanalyse z.B. bei Verkehrskreuzungen)

Beispiel einer 2 dimensionale Mustererkennung

Forderung: Programm um 1 von 6 zu unterscheiden

Merkmal 1: Pixelanzahl	1 :	30	typische 1 :	30 Pixel
	1 :	29		
	1 :	31		
	6 :	50	typische 6 :	50 Pixel
	6 :	52		
	6 :	48		
Merkmal 2: Quotient	9 :	52		50 Pixel
	9 :	50		2 Quot.
	9:	48		

Repräsentanten der Klasse „9“

Unterscheidungen von Klassen

10 Ziffern
26 Großbuchstaben
26 Kleinbuchstaben
10 Sonderzeichen

$\downarrow$

100 Klassen

Man benötigt viele Merkmale m1, m2, …, mn um die Klassen zu unterscheiden.

$\begin{pmatrix} m_1 \\ \dots \\ \dots \\ m_n \end{pmatrix}$ heißt Merkmalsvektor $\in R^n$ und jede Komponente gibt ein vom Lehrer festgelegtes Merkmal an.

Klasse 1: $\begin{pmatrix} m_1 \\ \dots \\ \dots \\ m_n \end{pmatrix}^1 = \vec{m}^1$

…

…

…

Klasse k: $\begin{pmatrix} m_1 \\ \dots \\ \dots \\ m_n \end{pmatrix}^k = \vec{m}^k$

Beispiel:

$$\left.\begin{aligned} 1 &= \begin{pmatrix} \text{Anz. Pixel} \\ \text{Quotient} \end{pmatrix} = \begin{pmatrix} 3.0 \\ 1.5 \end{pmatrix} \\ 6 &= \qquad\qquad\quad\; = \begin{pmatrix} 5.0 \\ 0.5 \end{pmatrix} \\ 9 &= \qquad\qquad\quad\; = \begin{pmatrix} 5.0 \\ 2.0 \end{pmatrix} \end{aligned}\right\}\ \text{Klassenrepräsentanten für „1“, „6“, „9“}$$

Wie lässt sich ein Klassenrepräsentant berechnen aus vielen Trainingsmustern?? → Mittelwertbildung

<u>Bemerkung:</u>

- System benötigt zum Training viele Trainingsmuster (um Mittelwert zu berechnen)
- Merkmale werden vom Trainer = Lehrer festgelegt
- Alle Merkmale sind numerische Vektorgrößen

➔ <u>numerische explizite Wissensverarbeitung</u>

Situation: Es kommt ein unbekanntes Muster

: Zu welcher Klasse gehört dieses Muster?

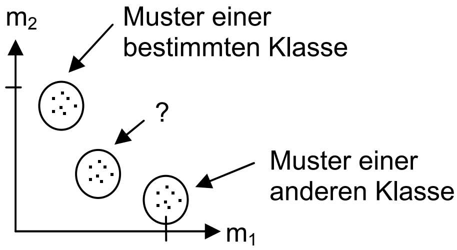
Muster gehört in die Klasse, zu dessen Vektor es minimalen Abstand hat. Muster hat den Merkmalsvektor $\vec{x}$ Є Rn (die n Merkmale werden als Komponenten berechnet)

$$d_j(\vec{x}, \vec{m}^j) = \min_{i = 1 \dots k} d_i(\vec{x}, \vec{m}^i) \rightarrow \text{x gehört in Klasse j}$$

$$[(x_1 - m_1^i)^2 + \dots (x_n - m_n^i)^2]^{1/2} \qquad \text{Beispiel für Euklid Abstand}$$

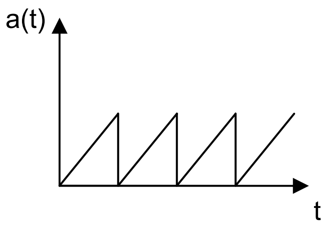
<u>Bemerkung:</u> bzgl. Normierung: Euklid nur sinnvoll bei gleicher Größenordnung der Komponenten. → sonst → Makalanobis Abstand (berücksichtigt Streuung).

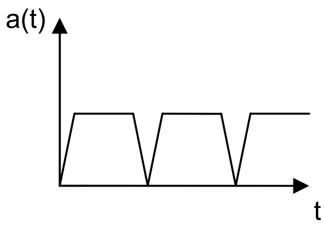
Problem bei unscharfen Mustern:



<u>Clusteranalyse:</u>
1. Zuordnung von Mustern in vorgegebene Klassen $k_1 \dots k_k$
2. Zuordnung in eine unbekannte Menge von Klassen
3. Klassen selbst sind unbekannt

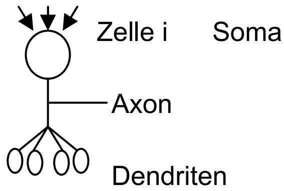
Akustischer Bereich



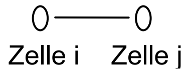


<u>Neuronale Netze</u>
Mensch als biologisches Vorbild

Neuron



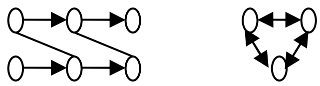
$10^{10} – 10^{11}$ Neuronen
Jedes mit 10.000
anderen verknüpft



Ein Neuronales Netz ist charakterisier durch folgende 3 Größen:

1) Signalverarbeitung in Zelle $\quad net_i = \sum_{j=1}^{n} w_j, o_j \quad \begin{array}{l} \geq c_i \rightarrow \text{Zelle aktiv} \\ < c_i \rightarrow \text{Zelle inaktiv} \end{array}$

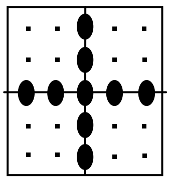
2) Topologie des neuronalen Netzes



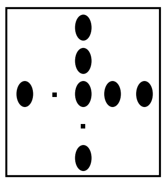
2) assoziatives Lernen. Änderung der Gewichtsmatrix bei neuen Beispielen

➔ Neuronen werden modelliert über Prozessoren. Simple Signalverarbeitung, aber massive Parallelität.

Verfahren sind robust.



Neuronen Netz 5x5
Kreuz



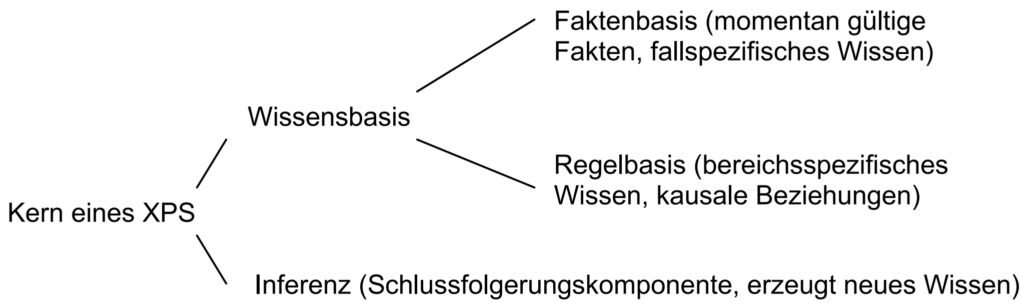
Implizite Wissensverarbeitung

<u>XPS (Experten System)</u>
Methode um Vorgehensweise eines menschlichen Experten bei der Problemlösung in einer bestimmten Domäne (Wissensgebiet) nachzubilden.

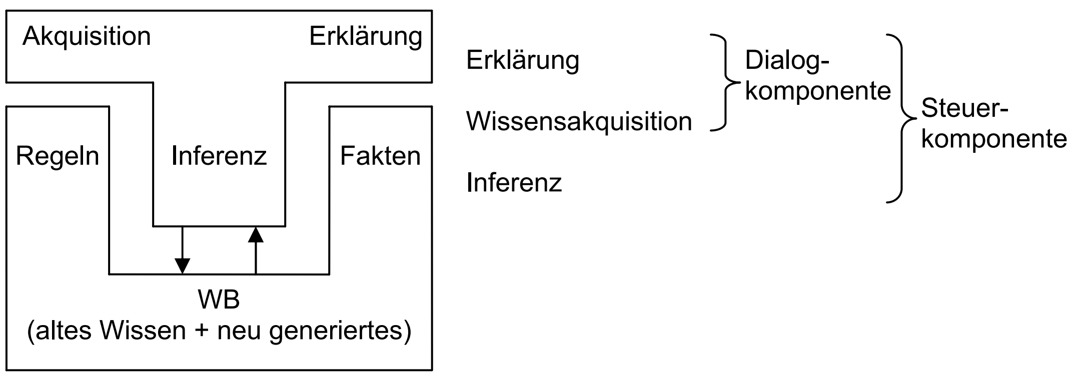
Menschlicher Experte	**Maschine (Computer)**
Problem verstehen	-
Problem lösen	+
Problemlösung erklären	+
Randgebiete überblicken	-
Kompetenzeinschätzung	+

Definition: Ein XPS ist ein wissensbasiertes Computerprogramm. Es verarbeitet große Wissensmengen und benutzt Heuristiken (Erfahrungswerte), um aus vorhandenem Wissen neues zu generieren.

1.3. Architektur eines Expertensystems



<u>Knowledge Engineering:</u>



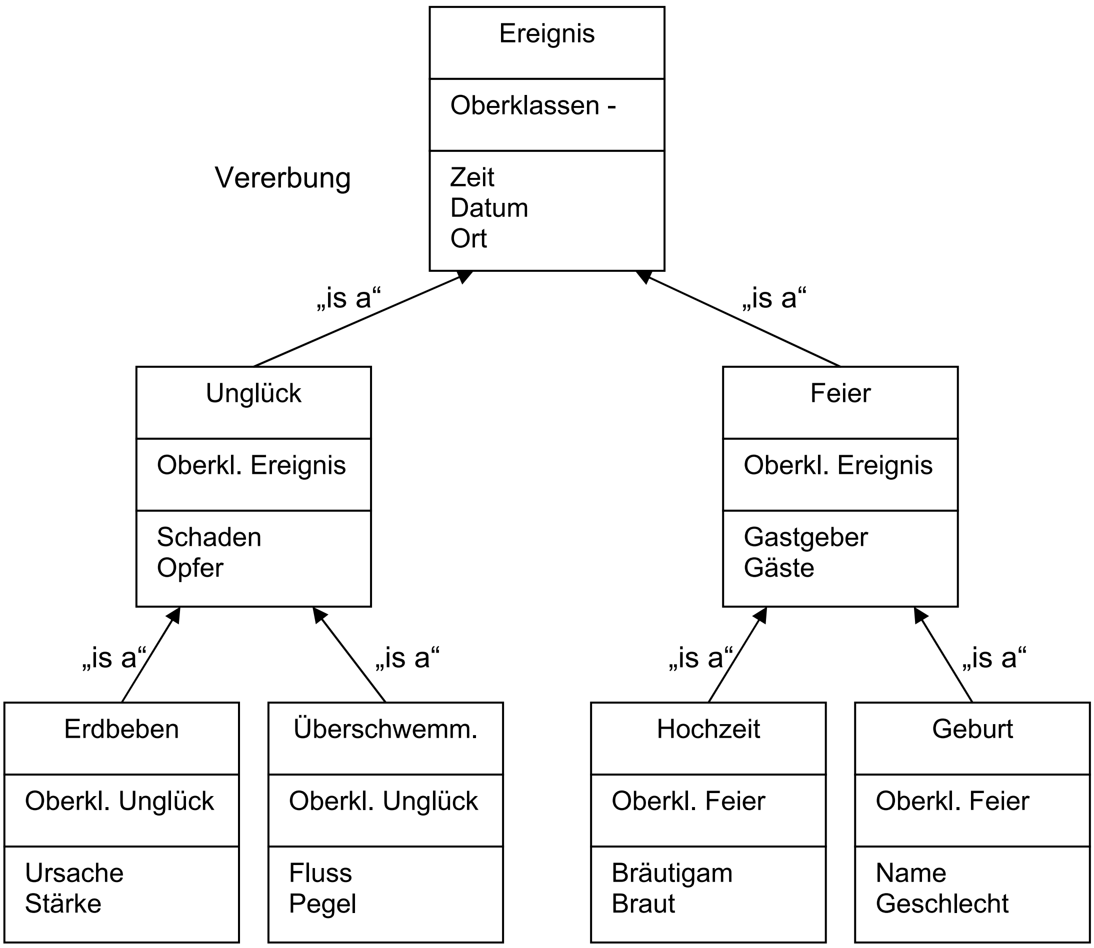
Erklärung: jede Schlussfolgerung kann nachvollzogen werden
Wissensakquisition: Wissen muss analysiert, strukturiert formalisiert werden.

<u>Bemerkung:</u> Wissen selbst und Strategien, das Wissen zu verarbeiten, sind logisch getrennt.

➔ <u>explizite symbolische Wissensverarbeitung</u>

2. Wissensbasis und Wissensrepräsentationen

2.1. Semantische Netze und Framestrukturen



- In der Baumstruktur werden entlang der Relation „is a“, die Attribute an die Unterklassen vererbt.
- Multiple Vererbung möglich
- Attribute können ab bestimmter Unterklasse ausgeblendet werden.

Instanzenframe

Geburt	
Zeit	14.00 Uhr
Datum	22.12.2008
Ort	FN
Gastgeber	Eltern
Gäste	Oma
Name	Uli
Geschlecht	m

Ein Baum (s.o.) ist ein spezielles Semantisches Netz.
Ein semantisches Netz besteht aus Knoten (Klassen/Objekte) und Kanten (frei definierbare Relationen). Eine Vererbung findet nur entlang der Relation „is a“ statt.

2.2. Prädikatenlogische Ausdrücke

Umkehrung des frameorientierten Ansatzes.

klein (Haus) := Das Haus ist klein (einstellige Relation)
teurer(Daimler, VW) := Daimler ist teurer als VW (zweistellige Relation)

allgemein: $\text{relation}\underbrace{(x_1, \ldots, x_2)}_{\text{Argumentliste, Reihenfolge!!!}}$

Prädikat
einstellige Relation

Man spricht auch von atomaren Formeln.
Sie haben den Wahrheitswert Є {wahr, falsch}.

Komplexe Formeln sind atomare Formeln, Verknüpft mit Operatoren

Operator	Bezeichnung	wahr wenn
^	Konjunktion (UND)	A wahr und B wahr
v	Disjunktion (ODER)	A wahr oder B wahr
¬, ~	Negation	A falsch
→	Implikation	B wahr oder A falsch
←→	Äquivalenz	A und B jeweils gleiche Wahrheitswerte

Bemerkung: (A → B) ←→ (B v ¬A)

Beispiel für komplexe Formel

Diff_pos(x, y) := $x - y > 0$
Diff_neg(x, y) := $x - y < 0$
Diff_0(x, y) := $x = y$
Zahl_pos(x) := $x > 0$
Zahl_pos(y) := $x < 0$

∀x ∀y [Zahl_pos(x) & Zahl_pos(y) → Diff_neg(x, y) v Diff_pos(y, x) v Diff_0(x, y)]

Diese komplexe Formel ist wahr für alle Belegungen x, y. Man spricht von einer wahren Interpretation.

2.3. Regeln

Wenn $B_1 \ldots B_n$ ← Prämisse, Vorbedingung, Antezedenz

dann gilt $H_1 \ldots H_m$
und führe aus $A_1 \ldots A_e$
} Konklusion, Schlussfolgerung < deklarativ (Hyp) / prozedural (Aktio)

2.4. Constraints

Ein Constraint definiert eine Relation über Randbedingungen. Es sind ungerichtete Zusammenhänge.

Beispiel $x = y + z$

$F = P * A$
Kraft = Druck * Fläche

1. if (Druck = x) AND (Fläche = y)
   then Kraft = x * y
2. if (Druck = x) AND (Kraft = y)
   then Fläche = y / x
3. if (Fläche = x) AND (Kraft = y)
   then Druck = y / x

2.5. Fakten-Regel Systeme

Definition: Ein Fakten-Regel System S ist eine endliche Menge von Fakten F(S) und Regeln R(S).

$S = F(S) \vee R(S)$

falls nun A aus S logisch hervorgeht, schreibt man $S \rightarrow A$

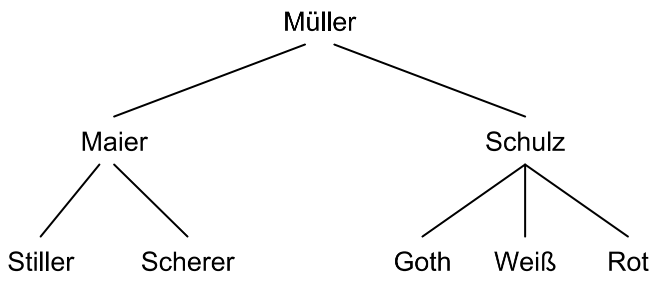
cons(S) $:= \{A | S \rightarrow A\}$ = Menge aller logischen Schlussfolgerungen

Beispiel 1

$S_1 =$ { ∀x Mensch(x) → sterblich(x),
∀x, y verwandt(x, y) & Mensch(x) → Mensch(y)
Mensch(Hans)
verwandt(Hans, Maria)
verwandt(Maria, Susi) }

cons S_1 = { Mensch(Hans)
verwandt(Hans, Maria)
verwandt(Maria, Susi)
sterblich(Hans)
Mensch(Maria)
Mensch(Susi)
sterblich(Maria)
sterblich(Susi) }

Beispiel 2



Angestellter(x)	:	ist Angestellter
Vorgesetzter(x, y)	:	x ist unmittelbarer Vorgesetzter von y
gleiche(x, y)	:	x und y sind auf gleicher Ebene

Beschreibung von gleicher Ebene

S_2 = { Angestellter(…) Vorgesetzter(…, …);
$\forall x$ Angestellter(x) → gleiche(x, x)
$\forall x, y, x_1, y_1$ gleiche(x, y) & Vorgesetzter(x, x_1)
& Vorgesetzter(y, y_1) → gleich(x_1, y_1) }

Bemerkung: Mit Fakten-Regel Systemen können Fakten aus den Regeln je nach Bedarf hergeleitet werden und brauchen nicht a priori in der Faktenbasis (Teil der Wissensbasis) gespeichert werden. (Reduktion der Faktenmenge durch Regelformulierung)

3. Inferenzmechanismen und Problemlösungsstrategien

3.1. Semantische Arten der Inferenz

A → B mathematisch logisch korrekt; deduktiv

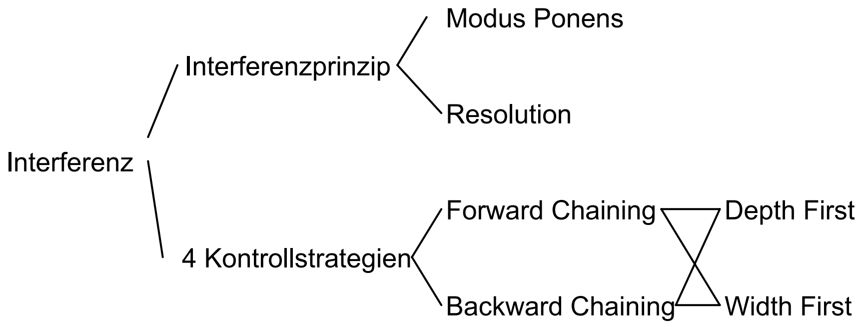
1) $\frac{A}{B}$ ← Logische Inferenz

2) C → B
A → B
$\frac{B}{A, C}$ Induktiv; Möglichkeit

3) A → B
$\frac{A}{B}$ probabistisches Schließen

4) A → B
A' → B' analoges Schließen

3.2. Logische Inferenzmechanismen



3.2.1. Logische Inferenz des Modus Ponens

$A \rightarrow B$ ← Regel Є Wissensbasis

$\dfrac{A}{B}$

Fakt Є Wissensbasis, Prämisse
Logische Inferenzregel
Geschlussfolgertes Fakt Є Wissensbasis
(wird in die Wissensbasis eingetragen)

Definition Inferenzregel:

Eine Inferenzregel ist eine Vorschrift, wie aus zwei Formeln eine neue Formel generiert werden kann.

3.2.2. Logische Inferenz der Resolution

Ausgangsform: ODER Formen; Klauselform

Aus	A v B	Vorgängerklauseln
und	¬B v C	
	———	← Logische Inferenzregel der Resolution
folgt	A v C	← Nachfolger

Bemerkung: 1) Klauseln entsprechen Regeln in der Wissensbasis
2) eventuell Reduzierung der Regelmenge

Beweis der Resolution:

A	B	C	AvB	¬BvC	AvC	AvB & ¬BvC	AvB & ¬BvC → AvC
0	0	0	0	1	0	0	1
0	1	0	1	0	0	0	1
0	0	1	0	1	1	0	1
0	1	1	1	1	1	1	1
1	0	0	1	1	1	1	1
1	1	0	1	0	1	0	1
1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	1

Varianten der Resolution:

1)
$$\begin{array}{l} A \vee B \\ \neg B \vee C \\ \hline A \vee C \end{array}$$

2)
$$\begin{array}{l} A_1 \vee A_2 \dots \vee A_n \\ B_1 \vee B_2 \dots \vee B_m \vee \neg A_1 \\ \hline B_1 \vee \dots B_m \vee A_2 \vee \dots A_n \end{array}$$

3)
$$\begin{array}{l} A \vee B \\ \neg A \vee B \\ \hline B \end{array}$$

4)
$$\begin{array}{l} A \vee B \\ \neg B \vee B \\ \hline A \vee B \end{array}$$

5)
$$\begin{array}{l} A \vee B \\ \neg A \vee \neg B \\ \hline \text{wahr} \end{array}$$

6)
$$\begin{array}{l} A \\ \neg A \\ \hline \lightning \end{array}$$

Beispiel für den Beweis einer Hypothese:

Regel1:	Aktie tief → kaufen
Regel2:	kaufen → Depot
Fakt:	Aktie tief
Frage:	Depot?
Strategie:	Formuliere ein Gesamtsystem aus der Wissensbasis und negierter Hypothese Zeige: Gesamtsystem ist inkonsistent → negative Hypothese muss falsch sein

1. Schritt: Klauselform

kaufen v ¬Aktie tief (1)
Depot v ¬kaufen (2)
Aktie tief (3)
¬Depot (4)

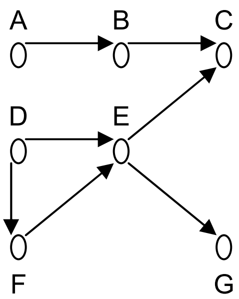
(1) & (2) = ¬Aktie tief (5)

(3) & (5) = Depot (6)

(4) & (6) = ↯ = Beweis der Hypothese

3.3. Kontrollstrategien

Beispiel:



Reihenfolge nicht beliebig

Forward Chaining: Initiierung durch Eingangsdaten (data driven)

Backward Chaining: Ist Zielaussage gültig? (goal driven)

3.3.1. Algorithmische Darstellung der Vorwärtsverkettung

Engl.: Forward Chaining

Es seien { R_1, …, R_n } Regeln und { F } Fakten.

while Fakten erfüllen nicht die Terminierung do

begin Konfliktmenge := Menge aller anwendbaren Regeln
(Prämisse durch Fakten erfüllt)
Auswahl einer Regel aus Konfliktset
Fakten: Aktionsteil von R auf Prämissen
end

3.3.2. Algorithmische Darstellung der Rückwärtsverkettung

Engl.: Backward Chaining

Seien { R_1, …, R_n } Regeln aus einer Wissensbasis und Ziel Z

1. bestimmte Teilmenge { R_i } $\subseteq$ { R_1, …, R_n } der Regeln, deren Aktionsteil das Ziel Z wahr machen
2. Führe für jedes R_i folgende Schritte aus

3. Prüfe alle Vorbedingungen B_j von R_i nach folgenden Kriterien
   a) Ist B_j bereits Fakt in der Faktenbasis
   b) Wird B_j durch Frage an Benutzer erfüllt
   c) Ist B_j der Aktionsteil einer anderen Regel, wähle B_j als neues Ziel

3.3.3. <u>Tiefensuche</u>

1. Verfolgung eines Lösungspfades bis Sackgasse
2. Rückkehr zur letzten Abgabelung
3. Einschlagen des nächsten Weges (siehe 1.)

3.3.4. <u>Breitensuche</u>

1. Verfolgung eines Lösungspfades um einen Schritt (1-Schritt-Verfahren)
2. Rückkehr zur letzten Abgabelung

4. Formale Logik bei Wissensrepräsentation und Inferenz

4.1. Kalküle

Definition: Kalküle sind Wissensrepräsentationsformalismus und Ableitungsregeln K = (F, A, R)
mit F = Menge aller bildbaren Formeln
A = Menge der Axiome
R = Menge der Regeln R: $M \rightarrow \alpha$
$\sqsubseteq F \quad \sqsubseteq F$

<u>Beispiel1:</u> $M = \{ \alpha, \alpha \rightarrow \beta \}$
mit $\alpha \sqsubseteq F$, $\alpha \rightarrow \beta \sqsubseteq F$
r_1: $M \rightarrow \beta$ ist Modus Ponens

<u>Beispiel2:</u> $M = \{ \alpha \vee \beta, \neg\beta \vee \gamma \}$
mit $\alpha \vee \beta \sqsubseteq F$, $\neg\beta \vee \gamma \sqsubseteq F$
r_2: $M \rightarrow \alpha \vee \gamma \sqsubseteq F$ (Resolution)

<u>Beispiel3:</u> $M = \{ \alpha, \alpha \rightarrow \beta, \beta \rightarrow \gamma, \gamma \rightarrow \delta \}$
mit … $\sqsubseteq F$
r_3: $M \rightarrow \delta$

Aussagenlogik	- atomare Aussagen - komplexe Aussagen mit Operatoren &, v, ¬, →, ↔
↓	
Prädikatenlogik 1. Ordnung	∀, ∃ Variable Menge von Individuen
↓	
Prädikatenlogik 2. Ordnung	∀, ∃ für Mengen und Eigenschaften von Mengen

4.2. Syntax und Semantik der Aussagenlogik

Sei $\Sigma = \{ A_1, \ldots, A_n \}$ eine endliche Menge von einfachen atomaren Aussagen.

Seien α, β Formeln über Σ durch Operatoren verknüpft.

z.B.: &, v, ¬, →, ↔

Beispiel: $\alpha \rightarrow \alpha$
$\alpha \rightarrow A$
$\alpha \rightarrow \beta \leftrightarrow \beta \vee \neg\alpha$

Definition: atoms(α) = Menge aller in α vorkommenden Atome.

Beispiel: atoms(A & B → C) = { A, B, C }
atoms(¬α) = atoms(α)
atoms(α v β) = atoms (α & β) = atoms(α → β)
= atoms(α ↔ β) = atoms(α) u atoms(β)

Es gelten Prioritätsregeln für die Operatoren:

1. ¬
2. &
3. v
4. →, ↔

Beispiel1:

A & ¬B → C v D & F

Beispiel2: ¬A → B ≠ ¬(A → B)

Es gilt weiterhin: Für die Operatoren &, v, $\leftrightarrow$ gilt die Assoziationsregel:

$$(\alpha * \beta) * \gamma = \alpha * (\beta * \gamma)$$

Es gilt nicht:

$$(\alpha \rightarrow \beta) \rightarrow \gamma = \alpha \rightarrow (\beta \rightarrow \gamma)$$

Definition: $|\ \alpha\ |$:= Formellänge (Anzahl der Atome)

<u>Beispiel:</u> $|\ A\ | = 1$
$|\ \alpha \ \&\ \beta \ \&\ \gamma\ | = |\ \alpha\ | + |\ \beta\ | + |\ \gamma\ |$

Es gelten folgende Äquivalenzen und Formeltransformationen:

$$\alpha \rightarrow \beta \quad \leftrightarrow \quad \beta \vee \neg\alpha$$
$$\alpha \leftrightarrow \beta \quad \leftrightarrow \quad (\alpha \rightarrow \beta) \ \&\ (\beta \rightarrow \alpha) \quad \leftrightarrow \quad (\alpha \ \&\ \beta) \vee (\neg\alpha \ \&\ \neg\beta)$$
$$\alpha = \beta \quad \leftrightarrow \quad \neg\alpha \rightarrow \neg\beta$$

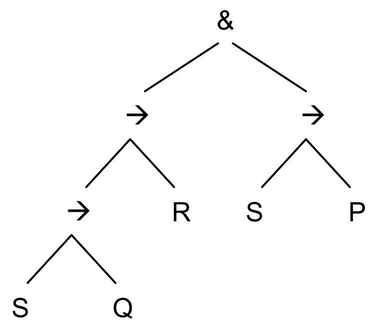
$$\alpha = \beta \Rightarrow \left\{ \begin{array}{l} \gamma \ \&\ \alpha = \gamma \ \&\ \beta \\ \gamma \vee \alpha = \gamma \vee \beta \\ \gamma \rightarrow \alpha = \gamma \rightarrow \beta \\ \alpha \rightarrow \gamma = \beta \rightarrow \gamma \\ \alpha \leftrightarrow \gamma = \beta \leftrightarrow \gamma \end{array} \right. \quad \text{Für beliebige Formel}$$

Weiterhin gilt:

Kommutativität	$\alpha \vee \beta = \beta \vee \alpha$ $\alpha \ \&\ \beta = \beta \ \&\ \alpha$
Assoziativität	$(\alpha \vee \beta) \vee \gamma = \alpha \vee (\beta \vee \gamma)$ $(\alpha \ \&\ \beta) \ \&\ \gamma = \alpha \ \&\ (\beta \ \&\ \gamma)$
Distributivgesetz	$(\alpha \ \&\ \beta) \vee \gamma = (\alpha \vee \gamma) \ \&\ (\beta \vee \gamma)$ $(\alpha \vee \beta) \ \&\ \gamma = (\alpha \ \&\ \gamma) \vee (\beta \ \&\ \gamma)$
De Morgan Gesetz	$\neg(\alpha \ \&\ \beta) = \neg\alpha \vee \neg\beta$ $\neg(\alpha \vee \beta) = \neg\alpha \ \&\ \neg\beta$

Repräsentation der Formel in Baumstrukturen

A: $((P \rightarrow Q) \rightarrow R) \,\&\, (S \rightarrow P)$

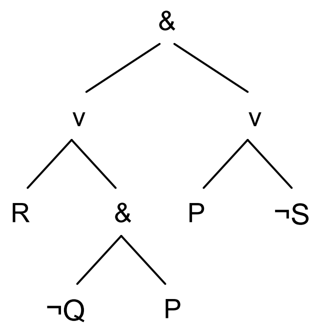


Blätter entsprechen atomaren Formeln. Knoten sind Operatoren.

<u>Bemerkung:</u> Im Baum können die atomaren Formeln durch komplexe ersetzt werden (ganzer Teilbaum).

Baumstruktur nur mit &, v, ¬ heißt kanonisch.

$[R \vee \neg(Q \vee \neg P)] \,\&\, (P \vee \neg S)$
$(R \vee \neg Q \,\&\, P) \,\&\, (P \vee \neg S)$



4.3. Übergang zur Prädikatenlogik

4.3.1. Formalisierung eines Sachverhaltes mit Prädikaten

<u>Beispiel1:</u>

Umgangssprache:
1. Jedes Auto hat mindestens ein Reserverad
2. Opel ist Auto und hat nur Notrad

Formalisierung:

Auto(x):	x ist Auto
Reserverad(y, x):	y ist Reserverad in x
Notrad(y):	y ist Notrad

Mit Prädikaten:
1. $\forall x \; Auto(x) \rightarrow \exists y \; Reserverad(y, x)$
2. $Auto(Opel) \,\&\, \forall y \; Reserverad(y, Opel) \rightarrow Notrad(y)$

$$\text{Umgangssprache} \xrightarrow{\text{nicht eindeutig}} \text{Prädikatenlogik}$$

$$\text{Umgangssprache} \xleftarrow[\text{eindeutig}]{} \text{Prädikatenlogik}$$

Beispiel2:

$\forall x\, [\, G(x, f, t) \rightarrow I(t, x)\,]$

a)

t	:=	Hauskatze
f	:=	Futter
I(a, b)	:=	a liebt b
G(a, b, c)	:=	a gibt das b an c

→ wahr;

b)

t	:=	10
f	:=	5
I(a, b)	:=	a > b
G(a, b, c)	:=	a + b > c

→ falsch z.B. für x = 11;

➔ Quintessenz: Der Wahrheitswert einer kompletten Formel hängt ab von den speziellen Prädikaten und den Parametern.

4.3.2. Einsatz von Inferenzregeln in der Prädikatenlogik

Inferenzregel	Ausgangsformel	Neu generierte Formel
Modus Ponens	$\alpha(x) \rightarrow \beta(x)$ $\alpha(x)$	$\beta(x)$
Modus Tollens	$\alpha(x) \rightarrow \beta(x)$ $\neg\beta(x)$	$\neg\alpha(x)$
Resolution	$\alpha(x) \vee \beta(x)$ $\neg\beta(x) \vee \gamma(x)$	$\alpha(x) \vee \gamma(x)$
Reductio ad Absurdum	$\alpha(x) \rightarrow \beta(x)$ $\alpha(x) \rightarrow \neg\beta(x)$	$\neg\alpha(x)$
Spezialisierung	$\forall x\, \alpha(x)$	$\alpha(x_0)$

Einsatz der Resolution in der Prädikatenlogik zum Beweis einer Aussage:

Beispiel: Sei X = { x | x ist Lebewesen }

I(x)	:=	x ist intelligent
D(x)	:=	x ist ein Delfin
G(x)	:=	x ist gebildet
L(x)	:=	x kann lesen

Aussagen: Wer lesen kann, ist gebildet.
Delfine sind nicht gebildet
Es gibt intelligente Delfine

Behauptung: Es gibt intelligente Lebewesen, die nicht lesen können.

1. Schritt: Formalisierung der Umgangsprache inkl. neg. Behauptung

Aussagen: $\forall x\ L(x) \rightarrow G(x)$
$\forall x\ D(x) \rightarrow \neg G(x)$
$\exists x\ D(x)\ \&\ I(x)$

Behauptung: $\exists x\ I(x)\ \&\ \neg L(x)$

¬Behauptung: $\forall x\ \neg I(x)\ v\ L(x)$

2. Schritt: Formulierung in Klauselform

Aussagen: $G(x)\ v\ \neg L(x)$
$\neg G(x)\ v\ \neg D(x)$
$D(x_0)$
$I(x_0)$

¬Behauptung: $\neg I(x)\ v\ L(x)$

3. Schritt: Resolution

$G(x)\ v\ \neg L(x)$ und $\neg G(x)\ v\ \neg D(x) = \neg L(x)\ v\ \neg D(x)$

$\neg L(x)\ v\ \neg D(x)$ und $D(x_0) = \neg L(x_0)$

$\neg I(x)\ v\ L(x)$ und $I(x_0) = L(x_0)$

- ➔ $L(x_0)$ und $\neg L(x_0)$ ist eine Inkonsistenz.
- ➔ Da die Wissensdatenbank zuvor konsistent war, kann die Inkonsistenz nur von der negierten Behauptung herführen.
- ➔ Behauptung ist wahr!

Satz über die Entscheidbarkeit in der Aussagenlogik:

Es gibt algorithmische Verfahren, die alle Ergebnisse in endlicher Zeit berechnen.

Satz über die Entscheidbarkeit in der Prädikatenlogik:

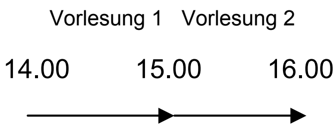
Es gibt algorithmische Verfahren, die jede richtige Hypothese in endlicher Zeit validieren (beweisen, bestätigen).
Das technische Verfahren ist die Resolution.

4.4. Temporales Schließen

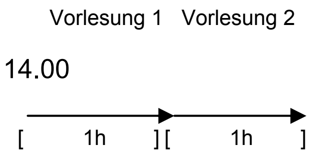
Schließen (Schlussfolgerungen ziehen) mit Zeitangaben.

Basisrepräsentationen:

- punktbezogen (Anfangszeit, Endzeit)

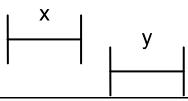
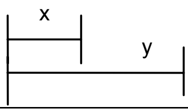
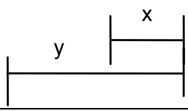
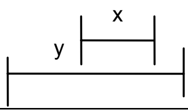
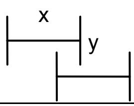
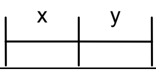
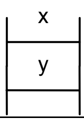


- intervallbezogen (Anfangszeit, Intervalllänge)



Es gibt konkrete Zeitangaben in min, std, Tage, … absolute Zeiten.
Es gibt relative Zeitangaben, die Bezug auf Referenzereignisse nehmen.

Mc Allen: Festlegung von „wenigen“ Relationen, um Ereignisse zu anderen Ereignissen (Referenzereignisse) in Beziehung zu setzen; Seien x, y zeitliche Ereignisse

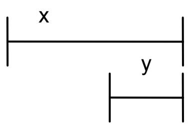
Relation	Grafische Repräsentation	Bedingung
x before, < y		$x^- < x^+ < y^- < y^+$
x starts y		$x^- = y^- < x^+ < y^+$
x finishes y		$y^- < x^- < x^+ = y^+$
x during y		$y^- < x^- < x^+ < y^+$
x overlaps y		$x^- < y^- < x^+ < y^+$
x meets y		$x^- < x^+ = y^- < y^+$
x equal = y		$x^- = y^- < x^+ = y^+$

Umkehrrelation:

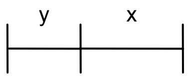
Sei r eine Relation zwischen zwei Ergebnissen a, b dann gilt für die sogenannte Umkehrrelation r^{-1} folgendes:

$a\ r\ b \leftrightarrow b\ r^{-1}\ a$

Beispiel:

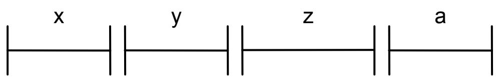


$x\ f^{-1}\ y$



$x\ m^{-1}\ y$

Situation:



x b y
y b z
x b z
x b a
y b a
z b a

Definition: Transitivität

Seien a, b, c Zeitereignisse, sei r eine Relation,
r heißt transitiv, wenn gilt
$a\ r\ b\ \&\ b\ r\ c \Rightarrow a\ r\ c$

Folgende Zeitrelationen sind transitiv (siehe Tabelle für volle Bezeichnung):

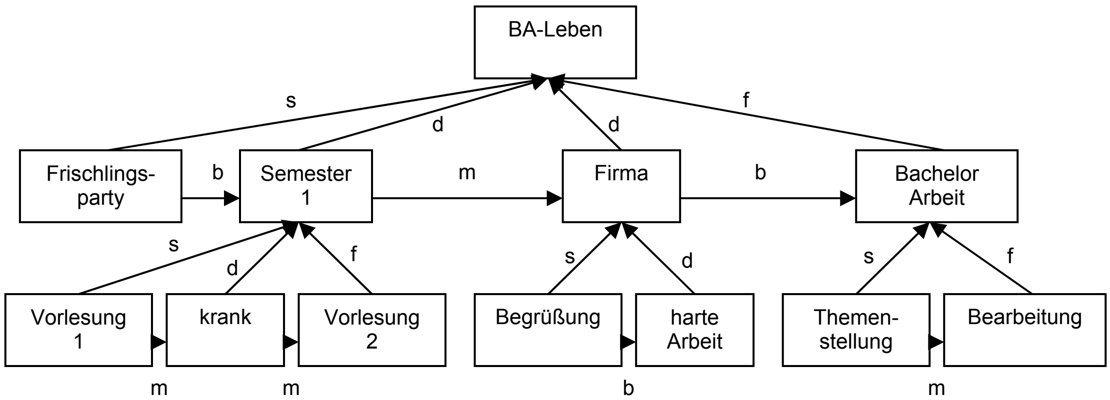
b, b^{-1}, s, s^{-1}, f, f^{-1}, d, d^{-1}, e

Folgende Zeitrelationen sind nicht transitiv:

o, o^{-1}, m, m^{-1},

Bemerkung: XPS benötigt nicht alle explizit möglichen Relationen, sondern es genügen oftmals wenige unter Berücksichtigung der Transitivität.

Beispiel:



Abschätzung:

12 Ereignisse
= 1 + 2 + … + 12
= ½ * 12 * 13 = 78

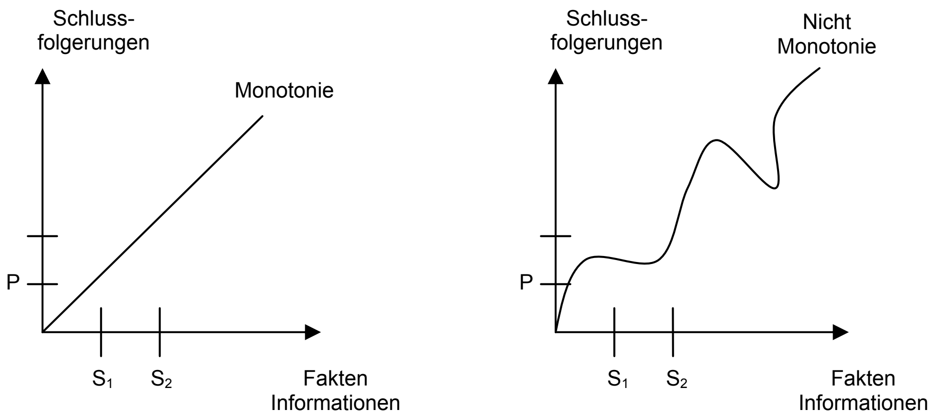
↓

<u>156</u>

Angabe: 14 Referenzen
Mögliche Relationsmenge: 156

4.5. Monotones, nicht monotones Schließen

Berücksichtigung des zeitabhängigen Wissensstandes, insbesondere der zeitabhängigen Schlussfolgerungen.



Systeme, welche falsche Informationen bis zu einem gewissen Punkt wieder aus der Wissensdatenbank entfernen können, nennt man Belief Revision oder Truth Maintaince.

Definition: Monotone Wissenserschließung

S_1 S_2 & $P_1 < P_2$
$\forall S_1 S_2 P: (S_1 \rightarrow P) \rightarrow (S_1 \sqcup S_2 \rightarrow P)$

d.h.: Es kann bei einer größeren Ausgangsmenge auf jeden Fall auf P geschlossen werden.

Definition: Nicht monotone Wissenserschließung

$\exists S_1 S_2 P: (S_1 \rightarrow P) \not\rightarrow (S_1 \sqcup S_2 \rightarrow P)$

4.6. Constraints, Constraintnetze, Constraintpropagierung

Definition1: Sei $V = \{v_1, \dots, v_n\}$ Variablen mit zugehörigem Wertebereich $D(v_1), \dots, D(v_n)$; Sei R eine Relation zwischen den Variablen $R \in D(v_1) \times D(v_2) \times \dots \times D(v_n)$, dann ist ein k-stelliger Constraint C über $V' = \{v_1', \dots, v_k'\} \sqsubseteq V$ eine Teilmenge $C \sqsubseteq D(v_1') \times \dots \times D(v_k')$.

Definition2: Sei C eine endliche Menge von Constraints
$C_1 = (V', D')$
$C_2 = (V'', D'')$
$C_3 = (V''', D''')$
dann heißt (V, C) ein Constraintnetz.

Definition3: Eine Lösung eines Constraint Satisfaction Problems (CSP) ist eine Belegung $\{v_1, \dots, v_n\}$, so dass alle $v_i \in D^i$ und alle C_i erfüllt sind.

Beispiel1:

Auswahl eines Autos nach Kosten.

Sei $V = \{v_1, v_2, v_3\}$ {Leasingkosten, Tankkosten, Unterhalt}
$D(v_1) = [0, 1000]$
$D(v_2) = [100, 400]$
$D(v_3) = [0, 500]$

$R \sqsubseteq [500, 800] \times [100, 400] \times [100, 500]$

Formulierung der Constraints (Randbedingungen):

$C_1 = (\{v_1\}, v_1 = 600)$	einstelliger Constraint
$C_2 = (\{v_2,v_3\}, v_2 + v_3 \leq 900)$	zweistelliger Constraint
$C_3 = (\{v_1,v_2,v_3\}, v_1 + v_2 + v_3 \leq 900)$	dreistelliger Constraint

Constraintpropagierung ist die Einschränkung des Wertebereichs D der Variablen.

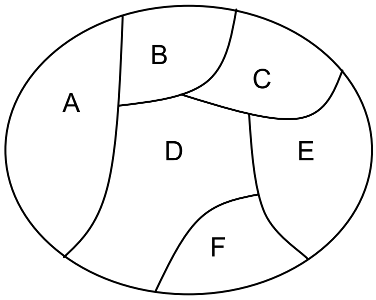
$v_1 = 600 \rightarrow v_3 \in [100, 200] \rightarrow v_2 \in [100, 200]$

d.h. durch Constraintpropagierung erhält man als mögliche Lösung den Bereich (600, [100, 200], [100, 200]).

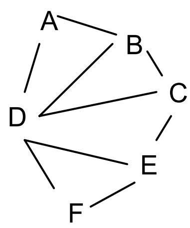
Beispiel2:

Färbeproblem bei Landkarten
Problem: mit 3 Farben sollen Länder eingefärbt werden, so dass keine angrenzenden Flächen die gleiche Farbe haben.
Länder seien A, B, C, D, E, F
Farben seien {rot, grün, blau} für jedes Land.

vgl. oben: v_1 = A, v_2 = B, …
$D(v_1)$ = {r, g, b}, $D(v_2)$ = {r, g, b}



Kantendarstellung



Modellierung über Knoten und Kanten

1. Anfangsfarbe für A sei rot (willkürlich gewählt)
2. Für B (r, g, b) fällt rot aus → B (g, b), wähle als Farbe willkürlich g
3. Für C(r, g, b) stehen auf Grund der Nachbarschaft nur r und b zur Verfügung C (r, b), Wahl fällt auf rot

A = rot

B (~~r~~, g, b) → B = g

C (r, ~~g~~, b) → C = r

D (~~r~~, ~~g~~, b) → D = b

E(~~r~~, g, ~~b~~) → E = g

F(r, ~~g~~, ~~b~~) → F = r

(A, B, C, D, E, F)
= (r, g, r, b, g, r)

als eine mögliche Lösung.
d.h. diese Lösung erfüllt
den Constraint

Anwendungsbeispiele:

- Stundenplanerstellung
- Färbeproblem
- Transportprobleme (Logistikbereich)

Fragestellungen:

- Existieren eine/mehrere Längen?
- Wie sieht eine Lösung aus?
- Falls es keine Lösung gibt, gibt es eine geringe Abweichung, sodass eine Lösung entsteht?

Definition4: Sei Ci ein Constraint über $V_i \sqsubseteq V = \{v_1, \ldots, v_n\}$
Falls eine Belegung $\beta(v_1), \ldots, \beta(v_n)$ eine Constraint Ci über $\{v_1, \ldots, v_n\}$ erfüllt, heißt β lokale Lösung
Erfüllt β alle C_i, dann heißt β globale Lösung.

Beispiel3:

Finde zwei Zahlen x, y mit x + y = 7 und x > y > 2

$V \{x, y\}$ $D(x) = D(y) = V$

$C_1 = \{y \mid y > 2\}$
$C_2 = \{(x, y) \mid x > y\}$
$C_3 = \{(x, y) \mid x + y = 7\}$

Constraintnetz $C = \{C_1, C_2, C_3\}$

z.B. x = 2, y = 1 ist lokale Lösung des Constraints C_2 aber keine globale Lösung, da C_3 nicht erfüllt ist.

z.B. x = 4, y = 3 ist globale Lösung, da alle C_i erfüllt.

Beispiel4:

Sei $V = \{x, y, z\}$ mit $D(x) = D(y) = D(z) = [0,1]$

Gegeben sei ein Constraintnetz $C = \{C_1, C_2, C_3, C_4\}$
mit $C_1 = \{(x, y) \mid x > y\}$ über $V_1 = \{x, y\}$
$C_2 = \{y \mid y > 0,5\}$ über $V_2 = \{y\}$
$C_3 = \{(x, z) \mid x + z = 1\}$ über $V_3 = \{x, z\}$
$C_4 = \{(x, z) \mid x < z\}$ über $V_4 = \{x, z\}$

Lokale Lösung bezüglich C_2 ist β = (0,5; 0,7; 0,5).
Lokale Lösung bezüglich C_3 ist β = (0,5; 0,7; 0,5).

Gibt es eine globale Lösung?

Nein, weil $\left.\begin{matrix} C_1 \\ C_2 \\ C_3 \end{matrix}\right\}$ $z > x > y > 0,5 \rightarrow x + z > 1$

Widerspruch zu C_3

Keine globale Lösung.

Generelle Bemerkung: sog. Backtracking:

> Existiert zu einem Zeitpunkt keine weitere Belegung der restlichen Variablen, muss ein Schritt zurück gegangen werden, um eine neue Belegung zu testen.
> Falls alle Belegungen ohne Erfolg sind, gibt es keine globale Lösung.

5. Repräsentation unscharfen Wissens

Ausgangspunkt: Binärlogik: Wahrheitsmenge = {wahr, falsch}

Erweiterung: Wahrheitsmenge
= {wahr, falsch, vielleicht,
mit_hoher_Wahrscheinlichkeit}

Erweiterung: Wahrheitsmenge = [0, 1]
mit 0 absolut falsch, 1 absolut wahr

Bemerkung: [0, 1] ist keine Einschränkung, da unendlich
viele Zahlen in [0, 1]

5.1. Statistischer Ansatz nach Bayes

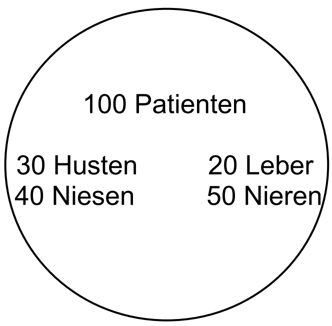
Aussagen sind mit Wahrscheinlichkeiten behaftet.
Wahrscheinlichkeitswert Є [0, 1]
Bayes benutzt die Wahrscheinlichkeit, um Ereignisse zu berechnen

Ausgangspunkt für ein Diagnosesystem:
Symptome → Diagnose
Hypothese

S → H
→ D

Es gelten folgende Vorraussetzungen:

1) Sei p(S) bzw. p(D) die Wahrscheinlichkeit für Auftreten eines Symptoms S bzw. einer Diagnose D. (Man nennt diese auch a priori Wahrscheinlichkeiten.)



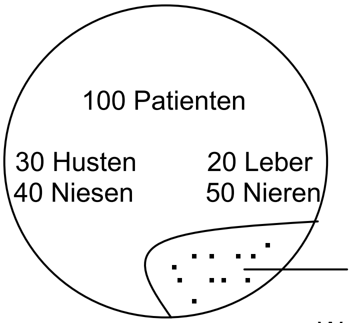
p(Husten) = 0,3
p(Nieren) = 0,5

2) Die Symptome S_i und die Diagnosen D_i seien jeweils unabhängig; d.h.:

$$p(S_1 \,\&\, S_2) = p(S_1) * p(S_2); \qquad p(D_1 \,\&\, D_2) = p(D_1) * p(D_2)$$

$$p(\text{Husten} \,\&\, \text{Niesen}) = 0{,}3 * 0{,}5 = 0{,}15$$

3) Seien $p(S_i \mid D_j)$ und $p(D_i \mid S_j)$ die bedingten Wahrscheinlichkeiten für das Auftreten des Symptoms S_i unter der Bedingung D_j bzw. für D_i unter Bedingung S_j.



Nierenkranke;
Davon 10 Husten;
Wahrscheinlichkeit, dass einer hustet, falls er an Niesen erkrankt ist,

p(Husten) = 0,3
p(Nieren) = 0,5

p(Husten | Nieren) = 0,2
bedingte Wahrscheinlichkeit

4) Die bedingten Wahrscheinlichkeiten für Si unter einer Diagnose D seien unabhängig

$$p(S_1 \,\&\, S_2 \mid D) = p(S_1 \mid D) * p(D_2 \mid S)$$

analog ist die bedingte Wahrscheinlichkeit für D_i unter der Vorraussetzung eines Symptoms S auch unabhängig.

Dann gilt:

$$p(D_1 \,\&\, D_2 \,\&\, \ldots \,\&\, D_m \mid S_1 \,\&\, S_2 \,\&\, \ldots \,\&\, S_n)$$

$$\begin{aligned} = \; & p(S_1 \mid D_1) * p(S_2 \mid D_1) * \ldots * p(S_n \mid D_1) * \\ & p(S_1 \mid D_2) * \ldots \\ & \ldots \\ & p(S_1) \mid D_m) * \ldots * p(S_n \mid D_m) \\ & * \frac{p(D_1)^n * \ldots * p(D_m)^n}{p(S_1)^m * \ldots * p(S_n)^m} \end{aligned}$$

Bemerkung:

1) Terme rechts von „=“ sind wohl definiert, d.h. Symptomverteilung für jede Krankheit ist bekannt a priori auch bekannt, Nenner von Bruch ≠ 0
2) Ausdruck links vom „=“ ist die interessierende Größe

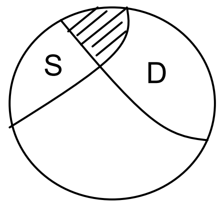
3) Die Gleichung ist korrekt

Beweis:

1. (einfacher) Fall:

$$p(D \mid S) = p(S \mid D) * \frac{p(D)}{p(S)} \qquad | * p(S)$$

$$p(D \mid S) * p(S) = p(S \mid D) * p(D)$$ trivial, da die Reihenfolge der Selektierung beliebig



2. Beweis der Formel durch <u>vollständige Induktion</u> über n (über m ganz analog)

Die Formel ist bereits gültig für n = 1
Es genügt zu zeigen: wenn Formel für n gültig, dann auch für n + 1.

$$p(D \mid S_1 \,\&\, \dots \,\&\, S_n \,\&\, S_{n+1}) = p(D \mid S_1 \,\&\, \dots \,\&\, S_n) * p(D \mid S_{n+1}) =$$

$$p(S_1 \mid D) * \dots p(S_n \mid D) * \frac{p(D)^n}{p(S_1) * \dots * p(S_n)} * p(S_{n+1} \mid D) * \frac{p(D)}{p(S_{n+1})}$$

Sind die a priori Wahrscheinlichkeiten $p(S_i)$ für die einzelnen Symptome S_i gegeben und unabhängig, ebenso bei D_j und hat man auch die bedingten Wahrscheinlichkeiten $p(S_i \mid D_j)$ und ihre Unabhängigkeiten, dann lässt sich die sogenannte posteriori Wahrscheinlichkeit $p(D_i \mid S_j)$ gemäß obiger Formel berechnen

5.2. <u>Gewissheitsfaktor, Glaubwürdigkeitsmaße, Unglaubwürdigkeit</u>

Ausgangspunkt:

S → H
(Symptom) (Hypothese)

Ich hab Vertrauen in Hypothese H bei Beobachtung von S.

MB (H, S) ist Maßzahl und gibt Stärke des positiven Vertrauens an.
MB = Measure of beliefe

MD (H, S) ist Maßzahl und gibt Stärke des Misstrauens an.
MD = Measure of disbelief

Normierung: Es soll gelten: $0 \leq MB(H, S), MD(H, S) \leq 1$
Es gilt für <u>ein</u> H und <u>ein</u> S folgendes: $MB(H, S) > 0 \rightarrow MD(H, S) = 0$
$MD(H, S) > 0 \rightarrow MB(H, S) = 0$
d.h.: Man kann nicht gleichzeitig einer Hypothese H trauen und nicht trauen.

Bei mehreren Symptomen (im Laufe der Zeit) ist
MB eine monoton wachsende Funktion Є [0, 1]
MD eine monoton wachsende Funktion Є [0, 1]

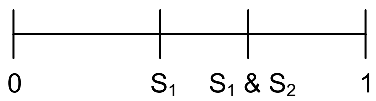
CF(H, S) := MB(H, S) – MD(H, S)
CF = Certainty factor (Gewissheitsgrad)

CF ist nicht monoton wachsend
CF Є [-1, 1] mit
CF = -1: Aussage ist absolut unglaubwürdig
CF = 1: Aussage ist absolut glaubwürdig
CF = 0: Vertrauen so groß wie Misstrauen

5.2.1. Berechnung von CF bei mehreren Symptomen S_i

Seien zwei Symptomen S_1, S_2 und eine Hypothese H gegeben mit $MB(H, S_1)$, $MB(H, S_2)$, dann gilt:

1) $MB(H, S_1 \& S_2) = MB(H, S_1) + [1 – MB(H, S_1)] * MB(H, S_2)$



Bemerkung: $MB(H, S_2)$ geht nur mit Restintervall als Belichtung ein
→ MB bleibt stets in [0, 1]

2) $MD(H, S_1 \& S_2) = MD(H, S_1) + [1 – MD(H, S_1)] * MD(H, S_2)$

→ CF = MB – MD

S_1 → H (ergibt MB oder MD)
S_2 → H (ergibt MB oder MD)
S_3 → H (ergibt MB oder MD)
… …
S_n → H (ergibt MB oder MD)

Vertrauenszuwachs MB oder Vertrauensschwund MD

5.2.2. Berechnung von CF bei komplexen Hypothesen $H_1 \& H_2$, $H_1 \vee H_2$

Seien zwei Hypothesen H1, H2 und eine Symptommenge S gegeben

mit $MB (H_1, S)$, $MB(H_2, S)$
bzw. $MD(H_1, S)$, $MD(H_2, S)$

dann gilt für die UND-Verknüpfung der Hypothesen:

$MB(H_1 \& H_2, S) = \min[MB(H_1, S) MB(H_2, S)]$
$MD(H_1 \& H_2, S) = \max[MD(H_1, S) MD(H_2, S)]$

Weiterhin gilt für die ODER-Verknüpfung der Hypothesen:

$$MB(H_1 \vee H_2, S) = \max[\, MB(H_1, S)\ MB(H_2, S)\,]$$
$$MD(H_1 \vee H_2, S) = \min[\, MD(H_1, S)\ MD(H_2, S)\,]$$

$CF(H, S) = MB(H, S) – MD(H, S)$
$CF(H_1, H_2, \ldots, H_n, S) = MB(H_1, H_2, \ldots, H_n, S) – MD(H_1, H_2, \ldots, H_n, S)$

CF ist groß bei MB groß und MD klein, d.h. bei großem positiven Vertrauen in alle Hypothesen und kleinem Misstrauen in alle Hypothesen.

5.2.3. Beispiel zur Berechnung von CF bei mehreren Symptomen S_i

Husten → Erkältung MB(Erkältung, Husten) = 0,3
(S_1) (H)

Niesen → Erkältung MB(Erkältung, Niesen) = 0,4
(S_2) (H)

Appetit → Erkältung MD(Erkältung, Appetit) = 0,7
(S_3) (H)

Fitness → Erkältung MD(Erkältung, Fitness) = 0,5
(S_4) (H)

CF(Erkältung, Husten & Niesen & Appetit & Fitness)
= MB(Erkältung, Husten & Niesen) – MD (Erkältung, Appetit & Fitness)
= [0,3 + 0,7 * 0,4] – [0,7 + 0,3 * 0,5] =
= 0,58 – 0,85 =
= <u>– 0,27</u>

5.3. Numerische Schärfe und Unschärfe

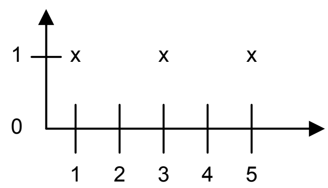
Sei $X = \{x\}$ Menge von Objekten
Sei $A \subseteq X$

betrachte $\mu_A(x) = \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases}$ Ist eine scharf definierte Funktion

$\mu_A : X \rightarrow \{0, 1\}$

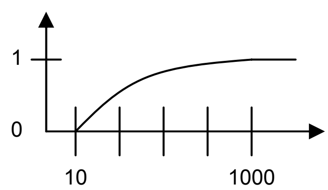
Beispiel1: (scharf) Sei $X = N$
$A \subseteq X$ mit $A = \{x \mid x \text{ ist gerade}\}$

$$\mu_A(x) = \begin{cases} 1 & \text{falls } x \text{ gerade} \\ 0 & \text{falls } x \text{ ungerade} \end{cases}$$



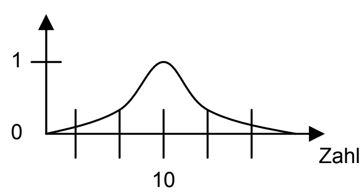
Beispiel2: (unscharf) $X = R^+$
$A \subseteq R+$ mit $A = \{x \mid x$ sehr viel größer als 10 $(x >> 10)\}$

$\mu_A(0) = \ldots \mu_A(10) = 0$, $\mu_A(11) = 0{,}1$, $\mu_A(1000) = 1$



Beispiel3: unscharfe Zahlen
„ungefähr 10"

Modellierung dieser unscharfen Zahl



1) symmetrisch
2) $\mu_A(0) = 1$ nur für diese Zahl x_0

x	1	2	3	4…	10	11…
$\mu_A(x)$	0,1	0,2	0,3	0,4…	1	0,9…

Beispiel4: unscharfe Relationen
„größer als“

Seien X, Y zwei Mengen von Elementen
$R = \{(x, y) \mid \mu_R(x, y)\}$ ist unscharfe Relation

für X = {Hans, Willi, Manfred, Joachim}
Y = {Hans, Willi, Manfred, Joachim}

Die Personen haben folgende Größen:

Hans	|	1,90
Willi	|	1,75
Manfred	|	1,65
Joachim	|	1, 85

Die Zugehörigkeit wird über die Funktion $\mu_R(x, y)$ modelliert.

$$\mu_R(x, y) = \begin{cases} 0 & \text{für } x \leq y \\ x\text{-}y/0{,}1y & \text{für } y \leq x \leq 1{,}1\,y \\ 1 & \text{für } x > 1{,}1y \end{cases}$$

x > y	Hans	Willi	Manfred	Joachim
Hans	0	0,857	1	0,27
Willi	0	0	0,606	0
Manfred	0	0	0	0
Joachim	0	0,571	1	0

5.4. Semantische Arten der Unsicherheit (Unschärfe)

- Stochastische Unsicherheit

  „Ziel zu treffen, hat die Wahrscheinlichkeit 0,8“
  Inhalt scharf definiert; numerische Wahrscheinlichkeit

- Lexikalische Unsicherheit (sprachliche)

  „große Männer“, „warme Temperaturen“
  ∃Variablen (linguistische Variablen), mit denen die Unsicherheit modelliert wird;

- Informationale Unsicherheit

  „Attraktivität“, „formschön“
  Es gibt Informationsüberfluss

6. Fuzzy Sets

Fuzzy: weich, pflaumig;

1920: erste Fuzzy Ansätze von Lukasiewicz.

Beobachtung: Terme wie „groß“, „klein“, „schwer“ sind nicht unter Wahrheitsbegriff wahr, falsch einzuordnen

→ Erweiterung des Systems der Logik auf reelle Zahlen [0, 1]

„Eine Zahl Є [0, 1] beschreibt die Möglichkeit (possibility) ob eine Aussage wahr oder falsch ist.“

1960 – 1965 Zadeh

Weiterentwicklung der Possibilitätstheorie zu einem formalen, logischen System.
System zur Beschreibung und Verarbeitung von natürlicher Sprache.

<u>Beispiel:</u>

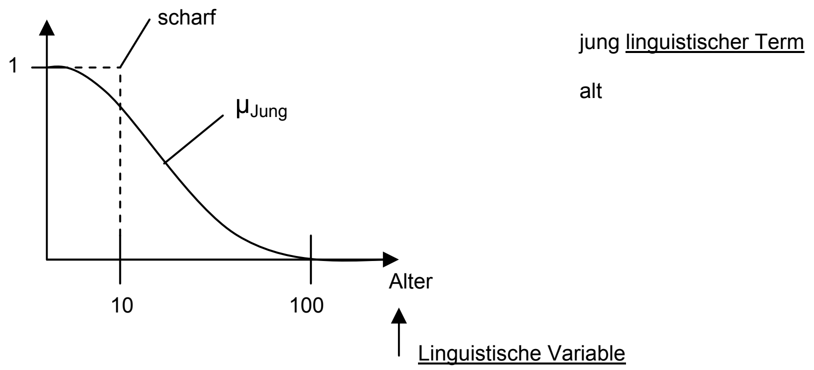
Linguistische Variable	**Linguistischer Term**
Temperatur	„warm“, „kalt“
Höhe	„niedrig“, „mittel“, „hoch“
Geschwindigkeit	„langsam“, „schnell“

Der Bereich der quantitativen Werte heißt Diskurs, bzw. Grundbereich.
Jeder linguistische Term (z.B. „kalt“) wird auf eine unscharfe Menge abgebildet; „kalt“ ist die unscharfe Aussage.

6.1. Fuzzy Mengen

Die Graduierung der Zugehörigkeit zu einer unscharfen Menge (wie gültig die unscharfe Aussage ist) wird dargestellt durch eine Zugehörigkeitsfunktion (Membershipfunktion) μ.

Beispiel:



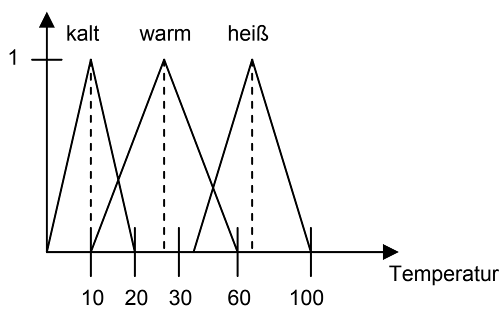
Grundbereich, Diskurs [0,100]

Definition:

Sei X ein Grundbereich mit Elementen x.
Eine Fuzzy Menge A von X ist beschrieben durch eine charakteristische Funktion μ_A, die jedem Element x Є X einen Zugehörigkeitswert μ_A (x) Є [0,1] für A zuordnet.

$$\mu_A : X \rightarrow [0, 1]$$
$$x \rightarrow \mu_A(x)$$

Beispiel: für Raumtemperatur



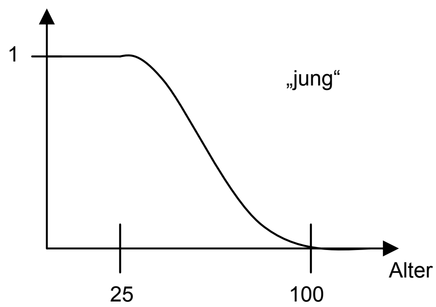
Linguistische Variable Temperatur
Linguistische Terme „kalt“, „warm“, „heiß“
Jeder linguistische Term wird modelliert über einen Fuzzy Set im gleichen Grundbereich.

Sei x 15 ➔ kalt & warm
$\mu_{kalt}(15) = 0,8$
$\mu_{warm}(15) = 0,4$

<u>Parametrische Ansätze für Fuzzy Sets:</u>

$\mu(x, c_1, p) = [1 + c_1 |x - x_0|^p]^{-1}$ $\qquad c_1 > 0, p > 1$

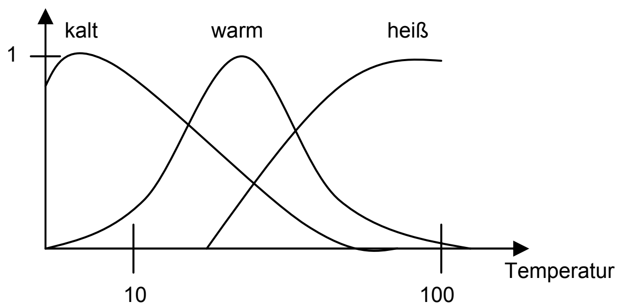
<u>Beispiel:</u> Alter von Menschen im Diskurs [0, 100]



$$\mu_{Jung} = \begin{cases} 1 & \text{falls } 0 \le x \le 25 \\ [1 + (\frac{x-25}{5})^2]^{-1} & \text{falls } 25 \le x \le 100 \end{cases}$$

6.2. <u>Schreibweise von Fuzzy Sets</u>

Sei $X = \{x_1, \ldots, x_n\}$ Grundbereich und A eine hierauf definierte Fuzzy Menge ($\mu_A(x)$).



kalt = (0.9/10, 0.8/20, 0.4/30, 0.1/60) = (0.9, 0.8, 0.4, 0.1)

warm = (0.1/10, 0.2/20, 0.6/30, 0.4/60) = (0.1, 0.2, 0.6, 0.4)

heiß = (0.0/10, 0.0/20, 0.1/30, 0.8/60) = (0.0, 0.0, 0.1, 0.8)

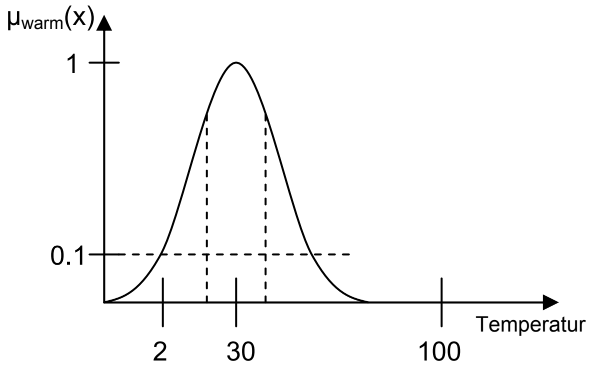
1) $A = (a_1, a_2, \ldots, a_n)$ $\qquad$ mit $a_i = \mu_A(x_i)$
Vektorschreibweise

2) $A = (a_1/x_1, a_2/x_2, \ldots a_n/x_n)$ $\qquad$ mit $a_i = \mu_A(x_i)$
Vektorschreibweise mit Stützstellen

Beispiel: “groß” = (0/170, 0.25/185, 0.5/190, 0.75/195, 1/200)
= (0, 0.25, 0.5, 0.75, 1)

3) $A = \mu_1/x_1 + \mu_2/x_2 + \ldots + \mu_n/x_n = \sum_{i=1}^{n} \mu_A(x_i)/x_i$
Zadeh Schreibweise

6.3. Charakterisierung von Fuzzy Sets



Definitionen:

1. $T(A)\ \{ x \mid x \in X, \mu_A(x) > 0 \}$ Träger des Fuzzy Sets A

2. α-Schnitt einer Fuzzy Menge

   $A_\alpha = \{ x, x \in X, \mu_A(x) \geq \alpha \}$

   Der 1-Schnitt d.h. $A_1 = \{ x, x \in X, \mu_A(x) = 1 \}$ heißt Kern

3. Höhe H(A) eines Fuzzy Sets

   $H(A) = \sup_{x \in X} \mu_A(x)$

   Fuzzy Mengen mit H(A) = 1 heißen normalisiert (normal)
   Fuzzy Mengen mit H(A) < 1 heißen subnormal

4. Kardinalität C(A) einer Fuzzy Menge

   $C(A) = \sum_{x \in X} \mu_A(x)$ (richtige Summe)

Beispiel:

X= [0, 100]

Träger
T(warm) = { x | 0 ≤ x ≤ 50 }

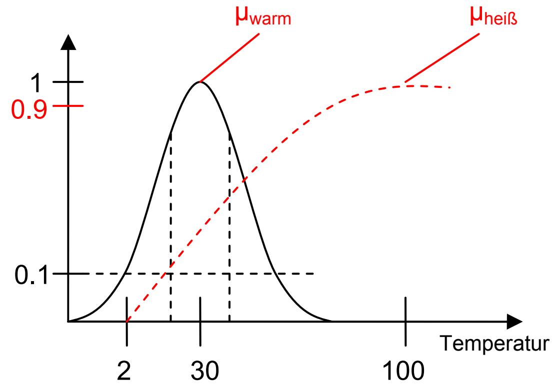
α-Schnitt:
warm$_{0.9}$ = { x | 28 ≤ x ≤ 32 }
Kern(warm) = {30}
Höhe(warm) = 1

C(warm) =

T(heiß) = [0, 100]

α-Schnitt:
heiß$_{0.9}$ = [60, 100]
Kern(heiß) = Ø

C(heiß) =



Beispiel:

8 Studenten $x_1, \dots, x_8$ haben in einer Klausur folgende Punktzahlen (max 75):

Student	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Punkte	69	26	52	55	60	41	46	53

Fuzzy Menge ←→ Fuzzy Aussage „gut bestanden"

Student	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Punkte	0.92	0.34	0.69	0.73	0.8	0.54	0.61	0.7

T(A) = $\{x_1, \dots, x_8\}$

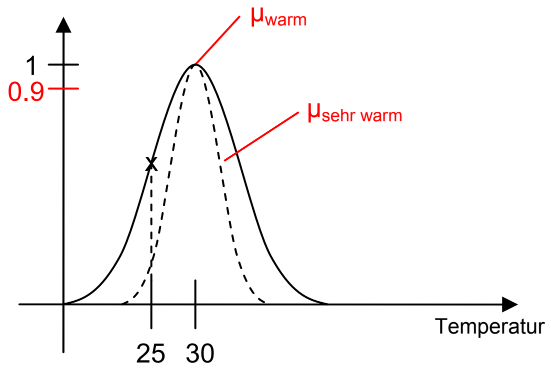
α-Schnitt: $A_{\alpha > 0.9} = \{x_1\}$
$A_{\alpha > 0.65} = \{x_1, x_3, x_4, x_5, x_8\}$

Höhe(A) = sup { } = 0.92

C(A) = 0.92 + … = 5,35

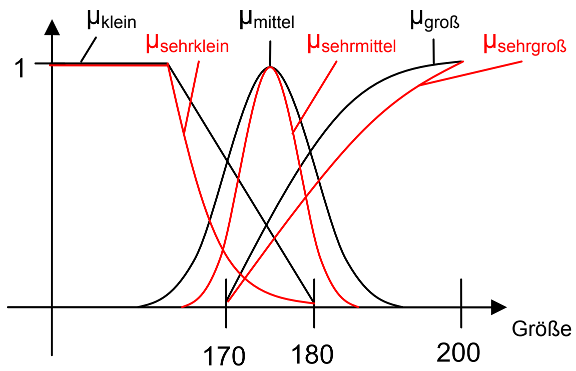
6.4. Modifier von Fuzzy Sets

Beispiel:



Beispiel: Größe von Menschen X = [150, 200]

Linguistische Variable Größe
Linguistische Terme „klein“, „mittel“, „groß“



„sehr klein“, „ziemlich klein“

„ziemlich mittel“ 3 Modifikationen von Grundtermen

„ganz groß“

Definition:

Sei X der Grundbereich und A ein Fuzzy Set mit $\mu_A(x)$

1. Verstärkung (Concentration) „sehr“

$\mu_{conc(A)}(x) = \mu_A(x)^2$

2. Aufweichung (Dilation) „ein bisschen“, „etwas“

   $\mu_{Dilation(A)}(x) = \sqrt{(\mu_A(x))}$

   $\mu_{Dil(A)}(x) = \sqrt{(\mu_A(x))}$

3. extra Verstärkung (Power) „sehr gut“, „ganz ganz groß“

   $\mu_{Pow(A)}(x) = \mu_A(x)^n \qquad n > 2$

<u>Beispiel:</u> linguistischer Term „alt“

$$\mu_{alt}(x) = \begin{cases} [1 + (\frac{x-50}{5})^{-2}]^{-1} & \text{für } 50 \leq x \leq 100 \\ 0 & \text{sonst} \end{cases}$$

$$A_{alt} = \sum_{x=51}^{100} [1 + (\frac{x-50}{5})^{-2}]^{-1} / x$$

$$A_{sehralt} = \sum_{x=51}^{100} [1 + (\frac{x-50}{5})^{-2}]^{-2} / x$$

$$A_{wenigeralt} = \sum_{x=51}^{100} [1 + (\frac{x-50}{5})^{-2}]^{-0.5} / x$$

6.5. Mengenoperationen auf Fuzzy Sets

Klassische Mengen $A \sqsubset X$

$x \in A$: x hat Eigenschaften von A

Durchschnitt $A \sqcap B$, Vereinigung $A \sqcup B$, Komplement $\neg A$ $\qquad x \in A, x \in B$

$A \sqsubset B \quad \leftrightarrow \quad \forall x \in A \ \ x \in B$

<u>Fuzzy Mengen</u>

X hat die Eigenschaften von A mit einem Grad μ_A

Frage: wie werden Fuzzy Sets aussagenlogisch verknüpft?

Definition:

Sei X der Grundbereich und A, B seien Fuzzy Sets mit μ_A, μ_B.

1. Es gilt: Teilmengenbeziehung: $A \sqsubseteq B : \mu_A(x) \leq \mu_B(x)$
2. Gleichheit von Fuzzy Sets

   $A = B \leftrightarrow \mu_A(x) = \mu_B(x) \quad \forall x \in X$
3. Durchschnitt von Fuzzy Sets

   A, B mit μ_A, μ_B.

   $\mu_{A \sqcap B}(x) = \min[\mu_A(x), \mu_B(x)]$

   Schreibweise: $\mu_A(x)$ & $\mu_B(x)$
4. Vereinigung von Fuzzy Sets

   A, B mit $\mu_A(x)$, $\mu_B(x)$.

   $\mu_{A \sqcup B}(x) = \max[\mu_A(x), \mu_B(x)]$

   Schreibweise: $\mu_A(x) \vee \mu_B(x)$
5. Komplement

   A mit $\mu_A(x)$.

   $\mu_{\neg A}(x) = 1 - \mu_A(x)$

Beispiel für Mengenoperationen:

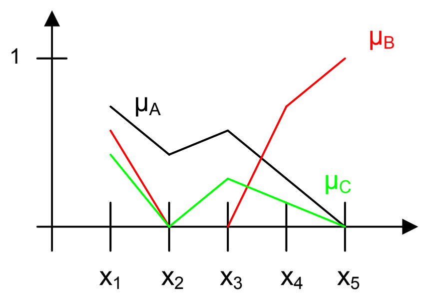
Grundbereich $X = \{x_1, x_2, x_3, x_4, x_5\}$

Es gebe drei Fuzzy Sets

$A = 0.7 / x_1 + 0.3 / x_2 + 0.4 / x_3 + 0.2 / x_4$

$B = 0.5 / x_1 + 0.6 / x_4 + 1.0 / x_5$

$C = 0.3 / x_1 + 0.2 / x_3 + 0.1 / x_4$



Dann gilt: $C \sqsubset A$, $C \not\sqsubset B$

$\neg A = 0.3 / x_1 + 0.7 / x_2 + 0.6 / x_3 + 0.8 / x_4 + 1.0 / x_5$

$A \sqcap B = 0.5 / x_1 + 0.2 / x_4$

$= \sum_{i=1}^{5} \min[\mu_A(x_i), \mu_A(x_i)] / x_i$

$A \sqcup B = 0.7 / x_1 + 0.3 / x_2 + 0.4 / x_3 + 0.6 / x_4 + 1.0 / x_5$

$= \sum_{i=1}^{5} \max[\mu_A(x_i), \mu_A(x_i)] / x_i$

$\mu_{\neg(A \sqcup B)}(x) = 1 - \mu_{A \sqcup B}(x)$

$\neg(A \sqcup B) = [1 - \max(\mu_A(x), \mu_A(x))]$

$= \min[(1 - \mu_A(x)), (1 - \mu_A(x))]$

$= \mu_{\neg A \sqcap \neg B)}(x)$

$\neg(A \sqcup B) = \neg A \sqcap \neg B$

<u>Mengenoperationen auf Fuzzy Sets zusammengefasst:</u>

Sei X die Grundmenge, A, B, C Fuzzy Mengen

Kommutativ: $A \sqcup B = B \sqcup A$
$A \sqcap B = B \sqcap A$

Assoziativität: $(A \sqcup B) \sqcup C = A \sqcup (B \sqcup C)$
$(A \sqcap B) \sqcap C = A \sqcap (B \sqcap C)$

Idempotenz: $A \sqcup A = A$
$A \sqcap A = A$

Distributivität: $A \sqcup (B \sqcap C) = (A \sqcup B) \sqcap (A \sqcup C)$

Neutrals Element: $A \sqcup 0 = A$

De Morgan: $\neg(A \sqcup B) = \neg A \sqcap \neg B$
$\neg(A \sqcap B) = \neg A \sqcup \neg B$

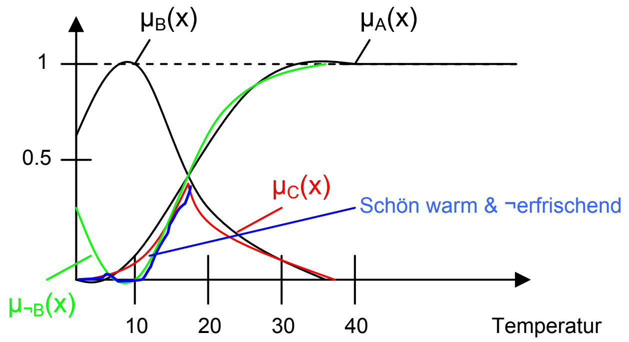
Beispiel für Fuzzy Modellierung:

(z.B. Einstellen einer Schwimmbadwassertemperatur)

A: „schön warm“
B: „erfrischend“
C: „angenehm“ := „schön warm“ & „erfrischend“

Drei linguistische Terme mit Bezug auf eine linguistische Variable (Temperatur).

Modellierung über Membershipfunktionen



Temperatur: x mit $\mu_C(x)$ = max = max min(μ_A, μ_B)

min(μ_A, 1 – μ_B) schön warm & ¬erfrischend

6.6. Fuzzy Inferenz

Prädikatenlogik: Regel a → b
Bezug zwischen zwei Aussagen
a, b und a → b wird durch Wahrheitswert repräsentiert

Fuzzy Regel: A → B
Bezug zwischen zwei Aussagen
A, B sind Fuzzy Mengen,
„→“ d.h. Implikation ist auch unscharf

→ syntaktisch kein Unterschied

Beispiel:

Höhe = { niedrig, mittel, hoch }
Gewicht = { leicht, normal, schwer }

A: Höhe ist hoch = {a_1, a_2, a_3} = {0.6/170, 0.8/180, 0.9/190}
B: Gewicht ist schwer = {b_1, b_2, b_3} = {0.5/60, 0.7/70. 0.9/80}

Regel: wenn Höhe = hoch dann Gewicht = schwer

A → B

Problem: Zuordnung

$\mu_A(x_i) \rightarrow \mu_B(x_i)$

➔ Matrixorganisation

Die Fuzzy Regel A → B wird über eine Matrixrelation repräsentiert, welche in den einzelnen Komponenten den Grad der Relation von x_i nach y_i angibt.

Hierbei ist $A = (\mu_A(x_1), \ldots, \mu_A(x_n))$
$B = (\mu_B(y_1), \ldots, \mu_B(y_m))$

$$M_R = \begin{pmatrix} \mu_A(x_1, y_1) & \mu_A(x_1, y_2) & \ldots & \mu_A(x_1, y_m) \\ \mu_A(x_2, y_1) & \mu_A(x_2, y_2) & \ldots & \mu_A(x_2, y_m) \\ \ldots & & & \\ \mu_A(x_n, y_1) & \mu_A(x_n, y_2) & \ldots & \mu_A(x_n, y_m) \end{pmatrix}$$

<u>Beispiel:</u>

Sei A = (0.2, 0.4, 0.6, 1) B = ?

$a_1 \rightarrow b_1$ $a_1 \rightarrow b_3$

$$M_R = \begin{pmatrix} 0.1 & 0.6 & 0.8 \\ 0.6 & 0.8 & 0.6 \\ 0.8 & 0.6 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

$a_4 \rightarrow b_1$ $a_4 \rightarrow b_3$

b_1 = max[min(0.2, 0.1), min(0.4, 0.6), min(0.6, 0.8), min(1, 0)]
= [0.1, 0.4, 0.6, 0] = 0.6

b_2 = [0.2, 0.4, 0.6, 0.5] = 0.6

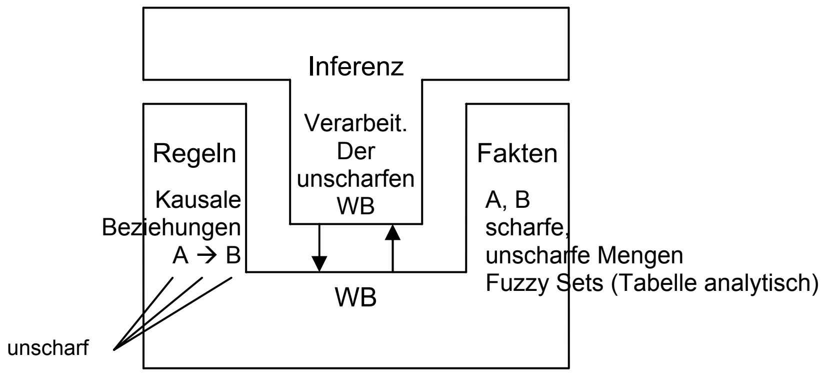
b_3 = [0.2, 0.4, 0.5, 0.5] = 0.5

➔ B = (0.6, 0.6, 0.5)

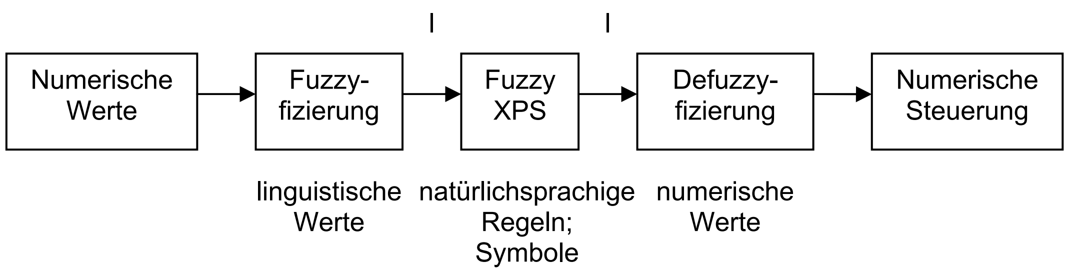
Fuzzy Inferenz als <u>max min Operator</u>

6.7. Fuzzy XPS

Struktur und Architektur



6.7.1. Prinzip des Fuzzy Control



6.7.2. Einteilung von Fuzzy XPS

1. Fuzzy Logic (im engeren Sinne)

$$\frac{\begin{array}{c} A \\ A \rightarrow B \end{array}}{B}$$

A, B Fuzzy Mengen

→ Fuzzy Relation

2. Approximate Reasoning; Plausible Reasoning;

   Aufweichung der strengen Fuzzy Logic

$$\frac{\begin{array}{c} A \rightarrow B \\ A' \end{array}}{B'}$$

7. Übungsaufgaben

7.1. Aufgabe zur Resolution

Zeigen Sie, dass	aus	A → B	
	und	B & C → D	
	die Formel	A & C → D	gilt.

Beweisen Sie mit Hilfe der Resolution für Klauseln.

Lösung1:

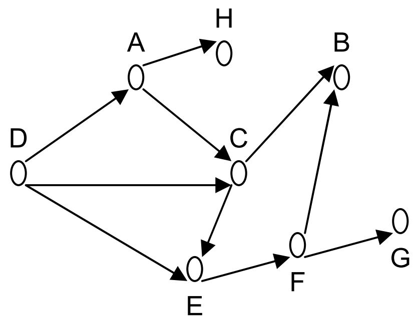
B v ¬A (1)
D v ¬(B & C) = D v ¬B v ¬C (2)

(1) & (2) = D v ¬A & ¬C = D v ¬(A & C) = A & C → D

Lösung2:

$$\frac{\begin{array}{c} A \rightarrow B \\ B \,\&\, C \rightarrow D \end{array}}{A \,\&\, C \rightarrow B} \quad \leftrightarrow \quad \frac{\begin{array}{c} \cancel{B} \vee \neg A \\ D \vee \cancel{\neg B} \vee \neg C \end{array}}{D \vee \neg A \vee \neg C}$$

7.2. Aufgabe zu Kontrollstrategien



VT
VB
RT
RB

Start ist D bzw. B

Lösung VT: D→A→H→ C→E→F→G→B

Lösung VB: D→A→C→E→F→H→B→G

Lösung RT: B→F→E→C→A→D

Lösung RB: B→C→F→E→A→D

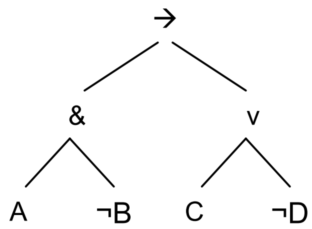
7.3. Aufgabe zur Aussagenlogik

Geben Sie folgende Formel als Baumstruktur und kanonische Baumstruktur an:

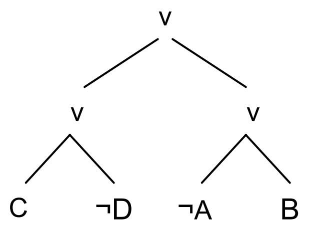
(A & ¬B) → (C v ¬B)

Lösung:

Baum (normal):



kanonisch: (C v ¬D) v ¬(A & ¬B) = (C v ¬D) v (¬A v B)



7.4. Aufgabe zu Inferenzregeln in der Prädikatenlogik

Zeigen Sie, dass der Modus Tollens eine logische Inferenzregel ist (mit Wahrheitstabelle).

Lösung:

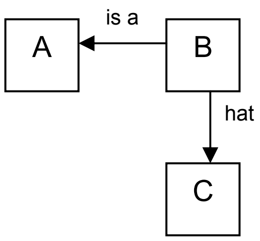
α(x)	**β(x)**	**β(x) v ¬α(x)**	**¬β(x)**	**¬α(x)**	**β(x) v ¬α(x) & ¬β(x)**	**β(x) v ¬α(x) & ¬β(x) → ¬α(x)**
0	0	1	1	1	1	1
0	1	1	0	1	0	1
1	0	0	1	0	0	1
1	1	1	0	0	0	1

7.5. Aufgabe zu Wissensrepräsentationen

Geben Sie die Ihnen bekannten Repräsentationsformen an, anhand eines kleinen Beispiels (formal).

Lösung:

1. Semantisches Netz



A, B, C Є R

2. Regel:

A & B → C

3. Prädikatenlogischer Ausdruck:

schnell(Auto)

4. Fakten-Regel System:

Mensch(x)
Mensch(Hans)
∀ x Mensch(x) → sterblich

5. Constraints:

y + x = 7 * z

7.6. Denkaufgabe

Ausgangsform: a a a _ b b b

as dürfen auf ein freies Feld um eins nach rechts vorrücken
bs dürfen um ein freies Feld um eins nach links vorrücken
as dürfen ein einzelnes b nach rechts auf ein freies Feld überspringen
bs dürfen ein einzelnes a nach links auf ein freies Feld überspringen

tauschen Sie die Positionen aller as und bs, sodass das Ergebnis
b b b _ a a a ist.
Bilden sie die daraus die minimale Anzahl an benötigten, logischen Regeln.

Lösung:

1)	a a a _ b b b	
2)	a a a b _ b b	b vor
3)	a a _ b a b b	a über b
4)	a _ a b a b b	a vor
5)	a b a _ a b b	b über a
6)	a b a b a _ b	b über a
7)	a b a b a b _	b vor
8)	a b a b _ b a	a über b
9)	a b _ b a b a	a über b
10)	_ b a b a b a	a über b
11)	b _ a b a b a	b vor
12)	b b a _ a b a	b über a
13)	b b a b a _ a	b über a
14)	b b a b _ a a	a vor
15)	b b _ b a a a	a über b
16)	b b b _ a a a	b vor

Regeln:

Sei x eine leere Position, 7 Positionen gesamt

R1): a: n $\rightarrow$ n + 1 wenn x = n + 1 und n $\leq$ 6
R2): a: n $\rightarrow$ n + 2 wenn b = n + 1 und x = n + 2 und $\leq$n $\leq$ 5
R3): b: n $\rightarrow$ n – 1 wenn x = n – 1 und n $\geq$ 2
R4): b: n $\rightarrow$ n – 2 wenn a = n – 1 und x = n – 2 und n $\geq$ 3

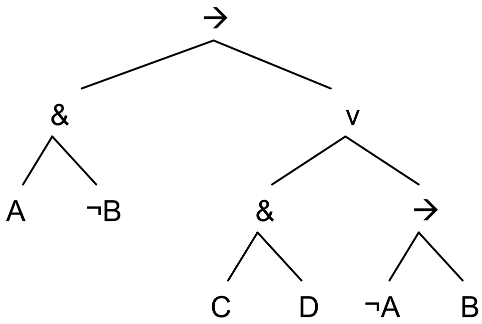
4 komplexe Regeln, die alle Spielkombinationen erzeugen können.

7.7. Sonstige Aufgaben

7.7.1. Aufgabe1

Geben Sie die Baumstruktur an für A & ¬B → C & D v (¬A → B)

Lösung:



7.7.2. Aufgabe2

Zeigen Sie: ¬A(x) & B(x) → ¬A(x) v C(x) & D(x) ist stets wahr
d.h. enthält keine spezifische Information (keine Wahrheitstabelle)

Lösung:

¬A(x) & B(x) → ¬A(x) v C(x) & D(x)

¬A(x) v C(x) & D(x) v A(x) v ¬B(x)

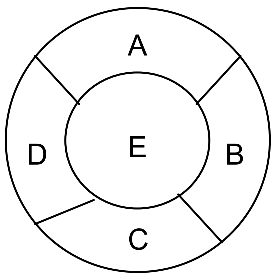
1 v C(x) & D(x) v ¬B(x) = immer wahr da ODER-Verknüpft

7.7.3. Aufgabe3

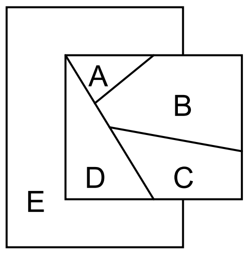
Färbeproblem bei 2 Weltkarten

Länder A, B, C, D, E sollen mit r, g, b so eingefärbt werden, dass Nachbarn ungleiche Farben haben

1)



2)



Gibt es eine Constraintlösung?! Wenn ja, welche?

Lösung:

1)

A = rot

B (~~r~~, g, b) → B = g

C (r, ~~g~~, b) → C = r

D (~~r~~, g, b) → D = g

E(~~r~~, ~~g~~, b) → E = b

(A, B, C, D, E)
= (r, g, r, g, b)

Nicht eindeutig!

2)

A = rot

B (~~r~~, g, b) → B = g

C (r, ~~g~~, b) → C = r

D (~~r~~, ~~g~~, b) → D = b

E(~~r~~, ~~g~~, ~~b~~) → E = ?

keine Lösung möglich

E grenzt an alle Länder, daher müssten alle anderen Länder mit zwei Farben abgedeckt werden. Das ist aber nicht möglich, da D an drei Länder grenzt und die dritte Farbe benötigt.

7.7.4. Aufgabe4

Zeigen Sie anhand der Definition:

Für Fuzzy Aussagen gilt das Assoziativgesetz wie in der Prädikatenlogik

Lösung:

(A v B) v C = A v (B v C)

$\mu_{(A \vee B) \vee C} = \max\,[\mu_{A \vee B}, \mu_C] = \max[\max(\mu_A, \mu_B), \mu_C] = \max(\mu_A, \mu_B, \mu_C)$

$\updownarrow$

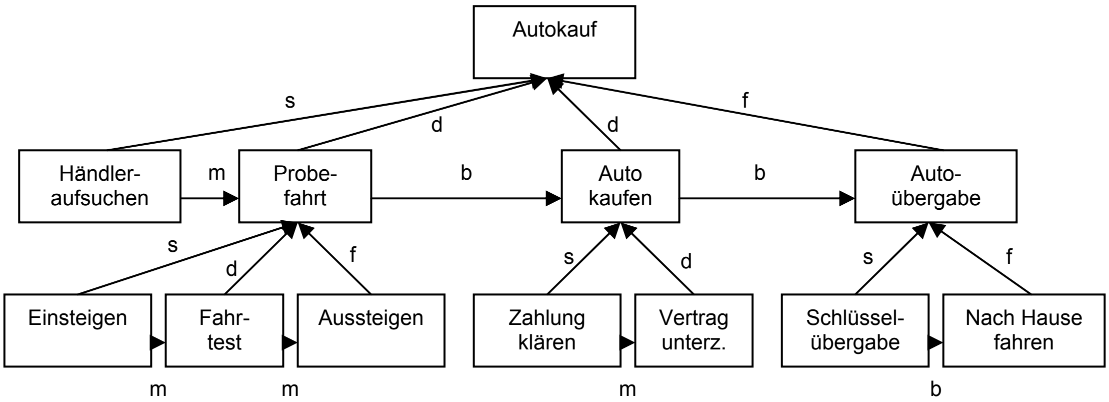
$\mu_{A \vee (B \vee C)} = \max\,[\mu_A, \mu_{B \vee C}] = \max[\mu_A, \max(\mu_B, \mu_C)] = \max(\mu_A, \mu_B, \mu_C)$

7.7.5. Aufgabe5

Modellieren Sie nach Mc Allen ein Beispiel für Referenzereignisse mit einer Baumtiefe von mindestens 3.

Lösung:

a)



b)

before	=	transitiv
starts	=	transitiv
finishes	=	transitiv
meets	=	nicht transitiv
equal	=	transitiv
during	=	transitiv
overlaps	=	nicht transitiv

7.8. Klausuraufgaben

7.8.1. Aufgabe 1 (15P)

Es gibt drei Symptome S_1, S_2, S_3, die für eine Hypothese sprechen (also jeweils positives Glaubwürdigkeitsmaß). Berechnen Sie die Gesamtkonfidenz und zeigen Sie, dass es bei dieser Berechnung der Konfidenz nicht auf die Reihenfolge der ankommenden Symptome ankommt.

Lösung:

S_1 für $MB(H, S_1)$
S_2 für $MB(H, S_2)$
S_3 für $MB(H, S_3)$

$$
\begin{aligned}
CF(H, S_1 \& S_2 \& S_3) &= MB(H, S_1 \& S_2 \& S_3) - 0 \\
&= S_1 + (1 - S_1) * S_2 + (1 - (S_1 + (1 - S_1) * S_2) * S_3 \\
&= S_1 + S_2 - S_1S_2 + S_3 - (S_1 + S_2 - S_1S_2) * S_3 \\
&= S_1 + S_2 + S_3 - S_1S_2 - S_2S_3 - S_1S_3 - S_1S_2S_3
\end{aligned}
$$

Bemerkung: beliebig permutierbar (austauschbar)!

7.8.2. Aufgabe 2 (15P)

Zur Detektion von Erkältung (E) und Infektion (I) werden als Symptome Müdigkeit (M) und erhöhte Temperatur (T) benutzt. Die Statistik belegt: Von allen Patienten haben 50% Erkältung und 50% Infektion. Von ersteren sind 20% müde, von letzteren 50%. Von ersteren haben 30% erhöhte Temperatur, von letzteren 40%. Sie wissen weiterhin: 30% aller Patienten sind müde, 40% haben erhöhte Temperatur.

a) Geben Sie alle a priori Wahrscheinlichkeiten und bedingte Wahrscheinlichkeiten an
b) Mit welcher Wahrscheinlichkeit schließt das XPS bei einem Bayesschen Ansatz auf
   - eine Infektion bei Beobachtung von erhöhter Temperatur?
   - eine Erkältung bei Beobachtung von Müdigkeit?
   - eine Erkältung bei Beobachtung von erhöhter Temperatur und Müdigkeit?

Lösung:

a)

$p(E) = 0,5$
$p(I) = 0,5$
$p(M) = 0,3$
$p(T) = 0,4$
$p(M \mid E) = 0,2$
$p(T \mid E) = 0,3$
$p(M \mid I) = 0,5$

p(T | I) = 0,4

b)

p(I | T) = p(T | I) * p(I) / p(T) = 0,4 * 0,5 / 0,4 = 0,5
p(E | M) = p(M | E) * p(E) / p(M) = 0,2 * 0,5 / 0,3 = 1/3
p(E | M & T) = …

7.8.3. Aufgabe 3 (10P)

Ihr XPS, welches als Inferenzregeln nur den „Modus Ponens“ und den „Modus Tollens“ kennt, besitz folgendes Regelwissen:

(1) C → ¬E
(2) A → B
(3) B → E
(4) F & ¬A → C

a) Welche Schlüsse zieht das System, wenn Sie anfangs nur F eingeben, was passiert, wenn dann noch ¬B eingegeben wird.
b) Gibt es einen Widerspruch, oder wann sollte das System mit der Abarbeitung stoppen?
c) Formulieren Sie alle Regeln in Klauseln

Lösung:

a)

Bei Eingabe von nur F passiert nichts, da weder Modus Ponens, noch Modus Tollens angewendet werden kann.

Bei Eingabe von F und dann ¬B:

(¬B & (A → B)) → A	Modus Tollens
(F & ¬A) → C	Modus Ponens
C → ¬E	Modus Ponens
(¬E & B → E) → ¬B	Modus Tollens

b)

Es gibt keinen Widerspruch. Das System sollte stoppen, wenn keine neuen Schlüsse gezogen werden können

c)

(1) ¬C v ¬E
(2) ¬A v B
(3) ¬B v E
(4) ¬F v A v C

7.8.4. Aufgabe 4 (10P)

Zeigen Sie, dass man mit den beiden Ausgangsformeln A → B und der Formel ¬B logisch korrekt auf ¬A schließen kann. Beweis sollte über eine Wahrheitstabelle erfolgen.

Lösung:

A	B	¬B	A → B	(A → B) & ¬B	¬A	((A → B) & ¬B) → ¬A
0	0	1	0	1	1	1
0	1	0	0	0	1	1
1	0	1	1	0	0	1
1	1	0	1	0	0	1

7.8.5. Aufgabe 5 (10P)

a) Was bedeutet „Nicht monotones Schließen“, „Temporales Schließen“ und „Unscharfes Schließen“?

b) Was ist der charakteristischer Unterschied zwischen der Bayes Anwendung und Gewissheitsfaktoren

Lösung:

a)

Nicht Monotones Schließen: Mit zunehmenden Informationen können auch bereits gültiges Wissen an seiner Gültigkeit verlieren.

Temporales Schließen: Zeitliche Informationen werden in den Aussagen mit berücksichtigt.

Unscharfes Schließen: Die Wahrheitswerte sind größer als 0 und 1, also nicht mehr Binär.

b)

Bayes arbeitet mit Wahrscheinlichkeiten, benötigt diese also auch. Gewissheitsfaktoren benötigen keine Wahrscheinlichkeit sondern benutzen subjektive Glaubwürdigkeiten.