COMPUTATIONAL INTELLIGENCE

10. FUZZY LOGIC

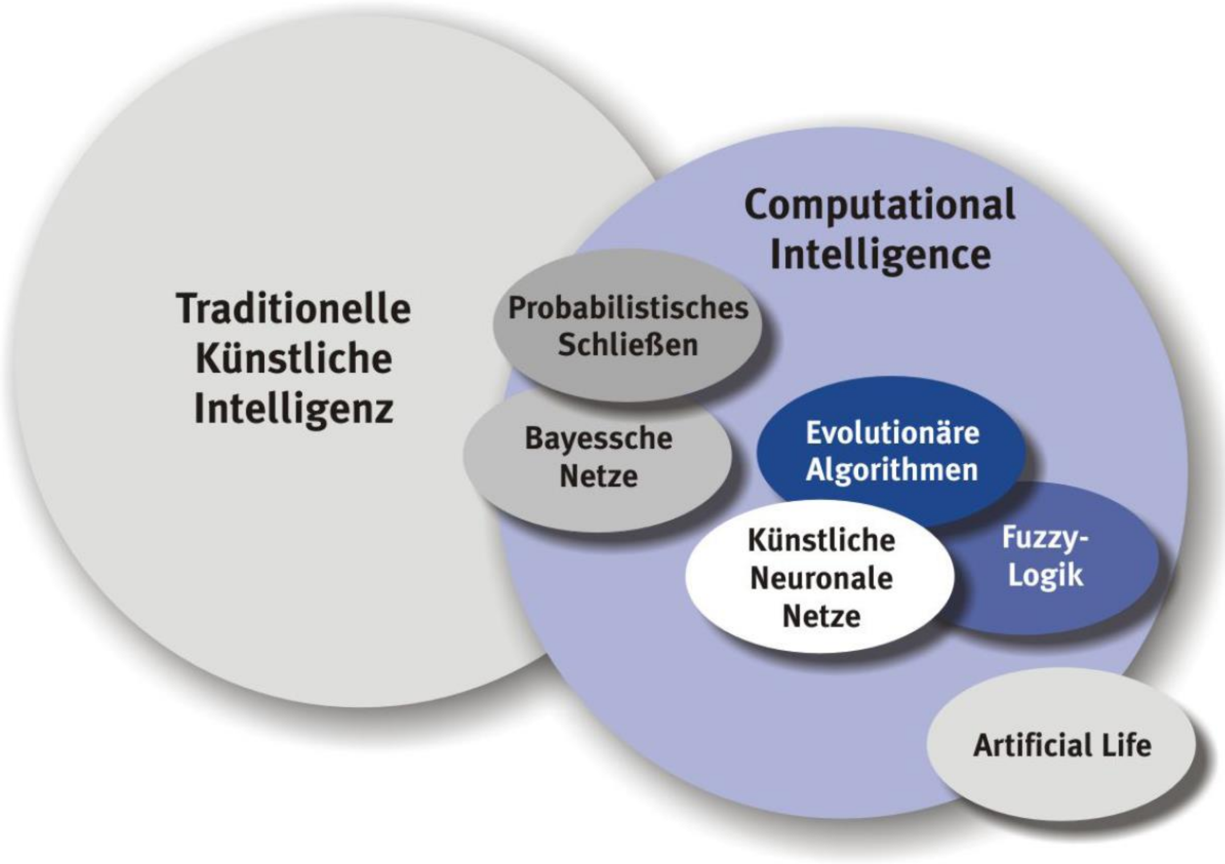
Prof. Dr. Sven Behnke

KLAUSUR

- 20. Juli ~~15:00-17:00 im HS 1+2 (Hörsaalzentrum, Endenicher Allee 19c)~~

  Geändert: 12:00-14:00 im Wolfgang Paul-Hörsaal, Kreuzbergweg 28
- 90 Minuten Bearbeitungszeit
- Closed Book

BAUSTEINE DER COMPUTATIONAL INTELLIGENCE



MOTIVATION

- Entwickelt von Lotfi Zadeh seit 1965
- Modellierung unscharfer Konzepte:
  - Bsp: Alter, Gewicht, Größe, …
- Modellierung unscharfer Abhängigkeiten (z.B. Regeln):
  - Wenn die Temperatur niedrig ist und Öl wenig kostet, drehe die Heizung auf
- Herkunft der Information:
  - Modellierung von Expertenwissen
  - Repräsentation von Informationen, die aus inhärent ungenauen Daten extrahiert wurden

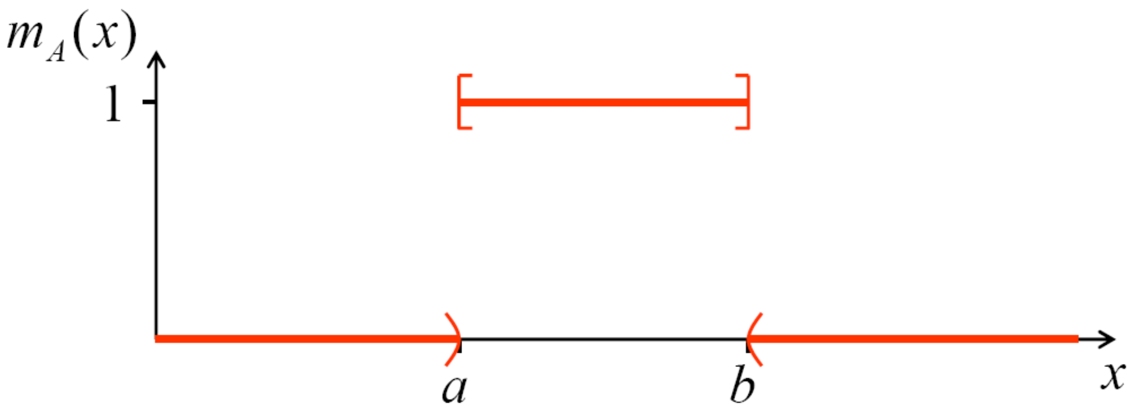
CHARAKTERISTISCHE FUNKTIONEN KLASSISCHER (SCHARFER) MENGEN

- Klassische Mengen können durch eine charakteristische Funktion beschrieben werden:

$$m_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \qquad m_A(x) \in \{0,1\}$$

- Beispiel:

$$A = \{x \mid a \le x \le b\}$$

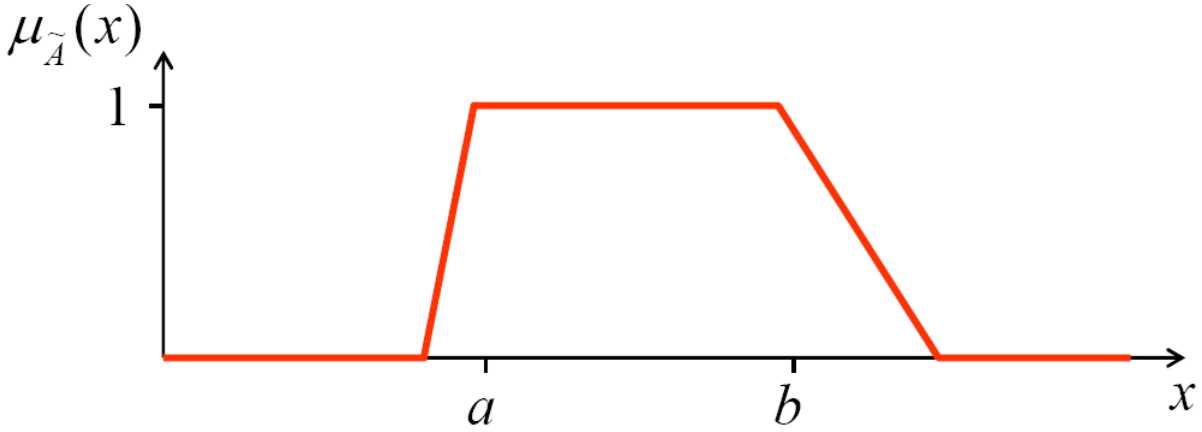


CHARAKTERISTISCHE FUNKTIONEN VON FUZZY-MENGEN

- Fuzzy-Mengen werden durch eine Mitgliedschaftsfunktion beschrieben:

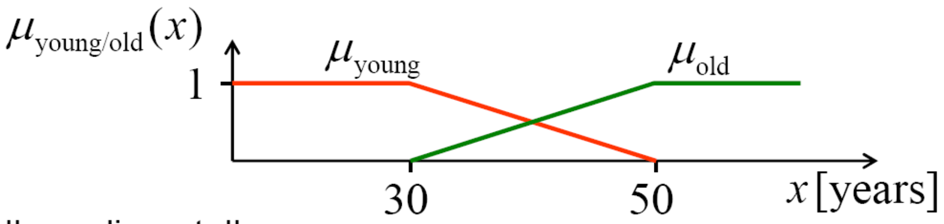
$$\mu_{\tilde{A}}(x) \in [0,1]$$

- Beispiel: $\tilde{A} = x \text{ is roughly in } [a, b]$

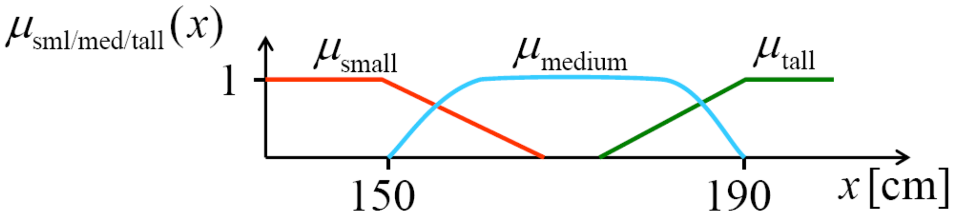


LINGUISTISCHE VARIABLEN UND WERTE

- Zuschreibung von **Semantik** zu Fuzzy-Mengen
  - **Linguistische Variablen**: Domäne der Fuzzy-Mengen
  - **Linguistische Werte**: „Diskretisierung“ dieser Domäne durch Fuzzy-Mengen
- Bsp:
  - Age: young, old

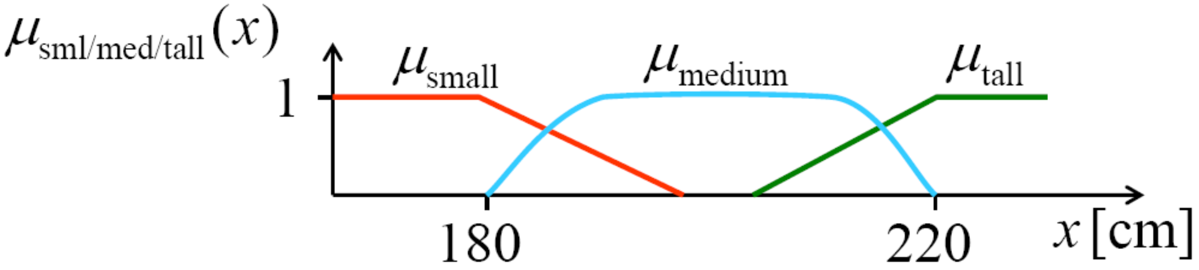


  - Size: small, medium, tall

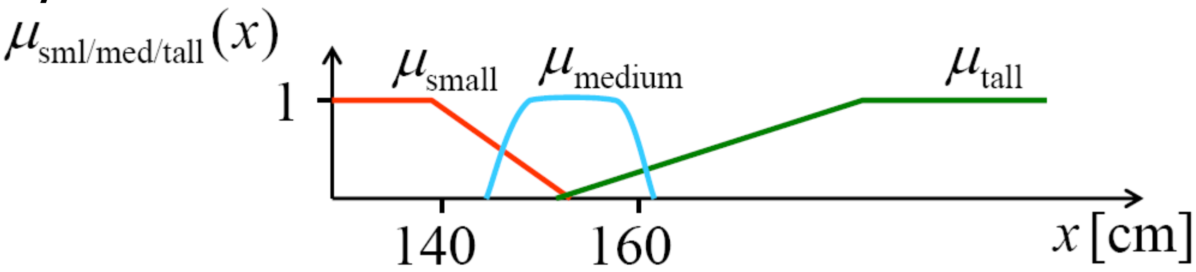


LINGUISTISCHE WERTE UND KONTEXT

- Basketball-Spieler:



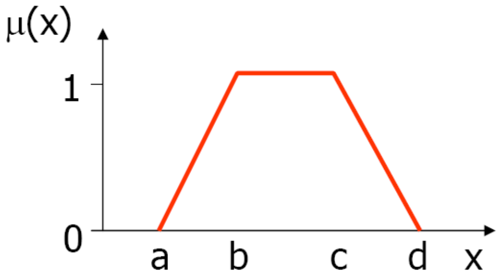
- Jockeys:



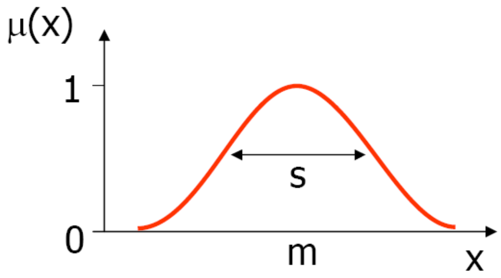
➔ Linguistische Werte sind kontextabhängig!

BELIEBTE MITGLIEDSCHAFTSFUNKTIONEN

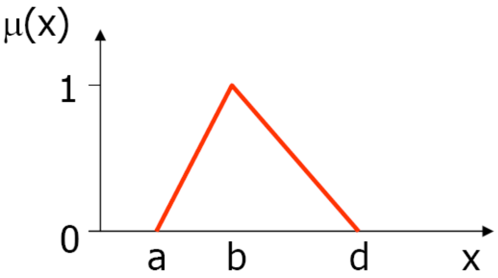
- Trapezoid <a,b,c,d>



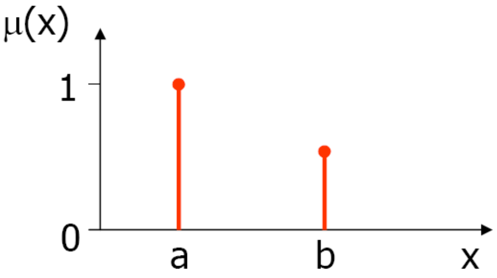
- Gaußglocke N(m,s)



- Dreieck <a,b,b,d>

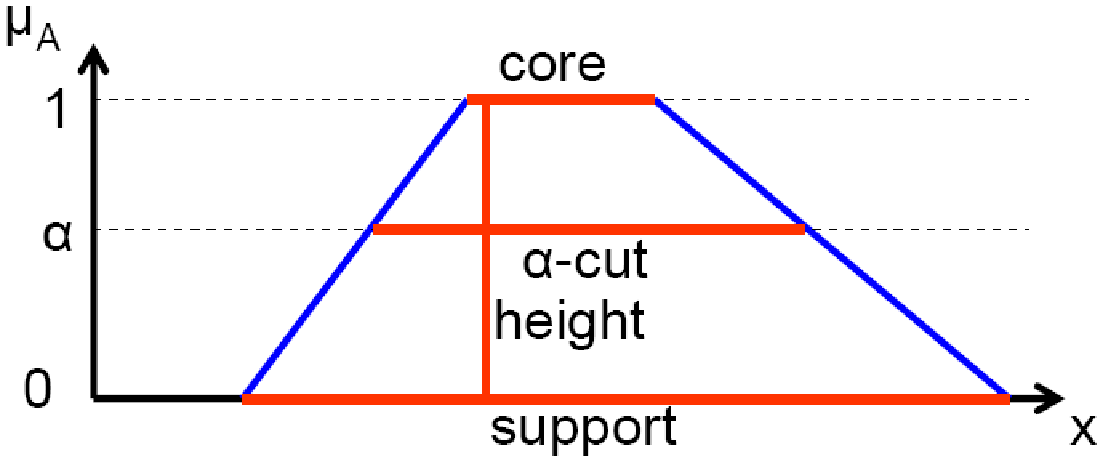


- Diskrete Werte {(a,1), (b,0.5)}



FUZZY-MITGLIEDSCHAFTSFUNKTIONEN: GRUNDLEGENDE BEGRIFFE

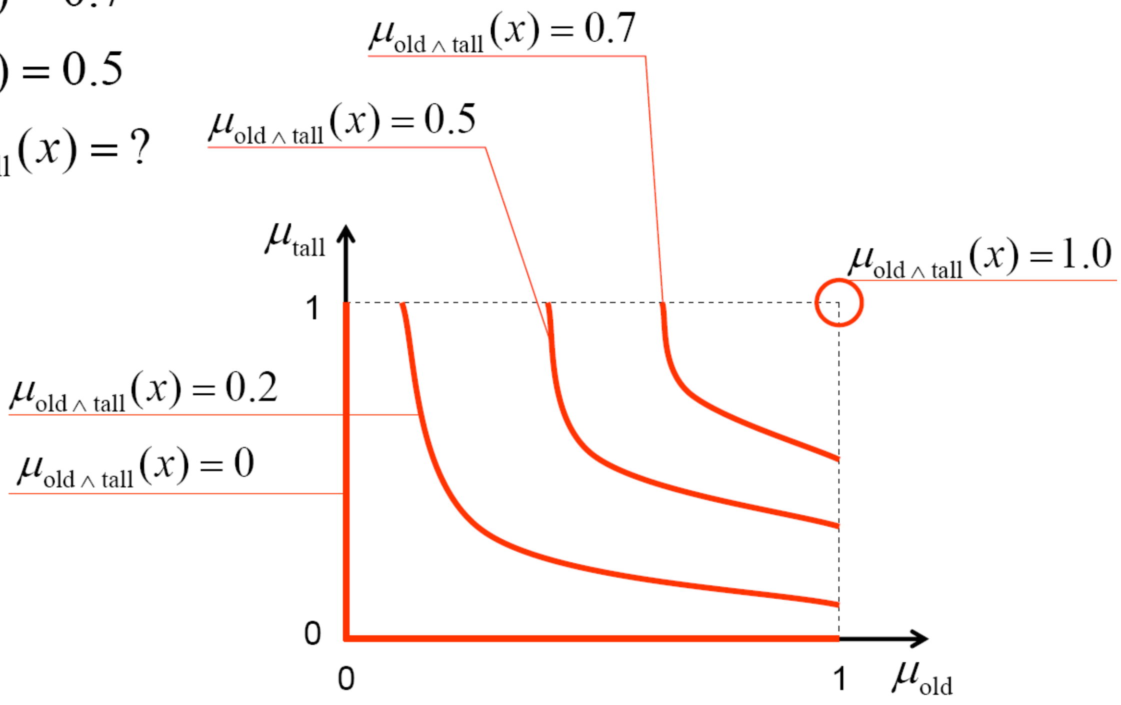
- **Support**: Menge mit Mitgliedschaft μ(x)>0
- **Core**: Menge mit Mitgliedschaft μ(x)=1
- **α-cut**: Menge mit Mitgliedschaft μ(x)>=α
- **Height**: Maximale Mitgliedschaft



OPERATOREN AUF FUZZY-MENGEN: KONJUNKTION (UND)

- Menge von Personen, die sowohl alt als auch groß sind (Konjunktion)

$$\mu_{\text{old}}(x) = 0.7$$
$$\mu_{\text{tall}}(x) = 0.5$$
$$\mu_{\text{old} \wedge \text{tall}}(x) = ?$$



Grenzfälle der Booleschen Logik müssen weiterhin gelten!

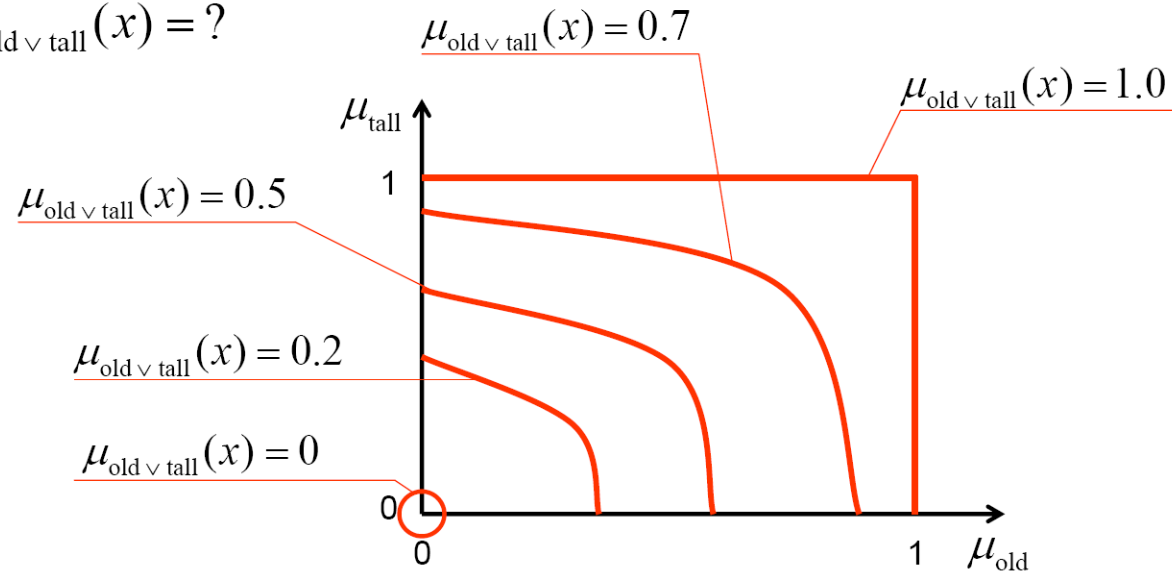
OPERATOREN AUF FUZZY-MENGEN: DISJUNKTION (ODER)

- Menge von Personen, die alt oder groß sind (Disjunktion)

$$\mu_{\mathrm{old}}(x) = 0.7$$

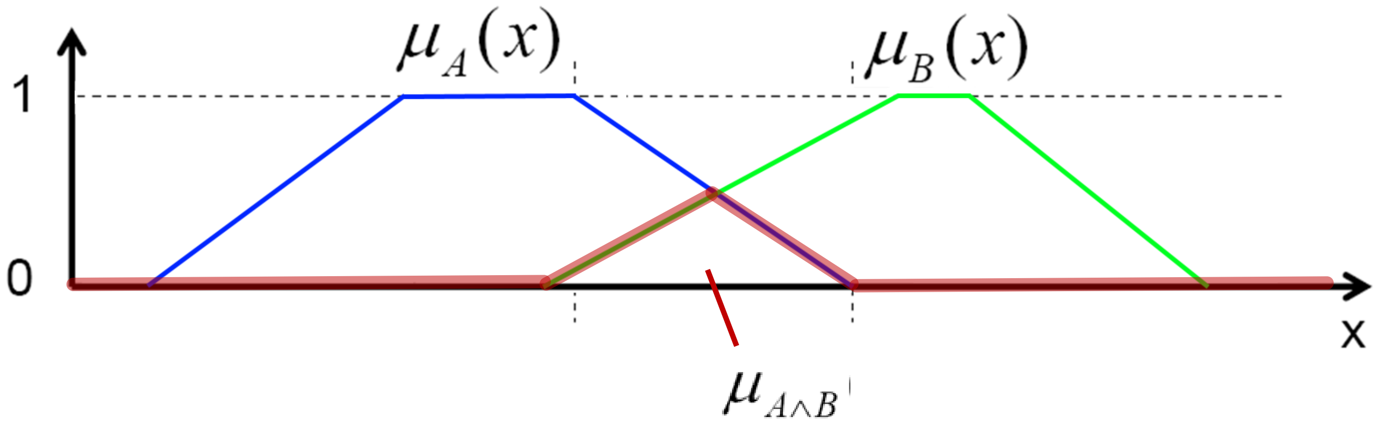
$$\mu_{\mathrm{tall}}(x) = 0.5$$

$$\mu_{\mathrm{old} \vee \mathrm{tall}}(x) = ?$$



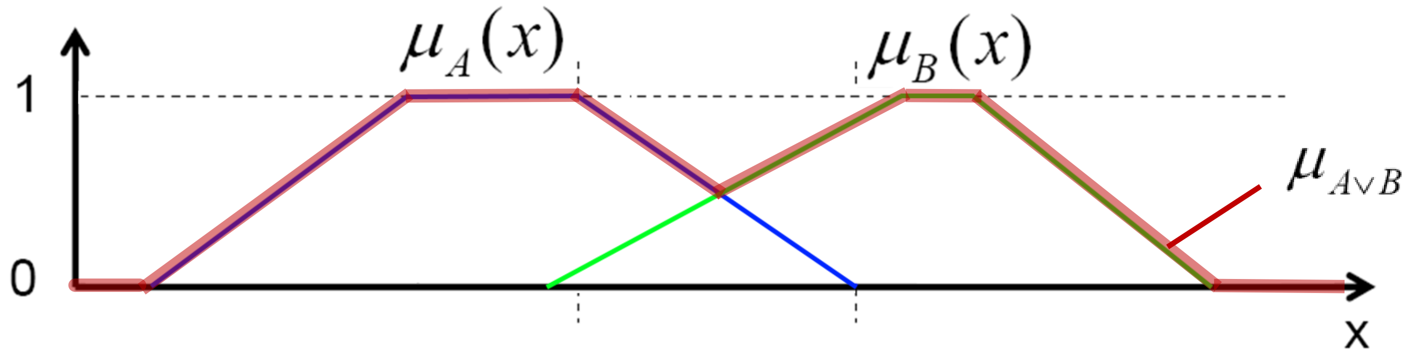
OPERATOREN: MIN/MAX - NORM

- Klassische Fuzzy-Operatoren: Min/Max-Norm
  - Konjunktion: $\mu_{A\wedge B}(x) := \min\{\mu_A(x), \mu_B(x)\}$



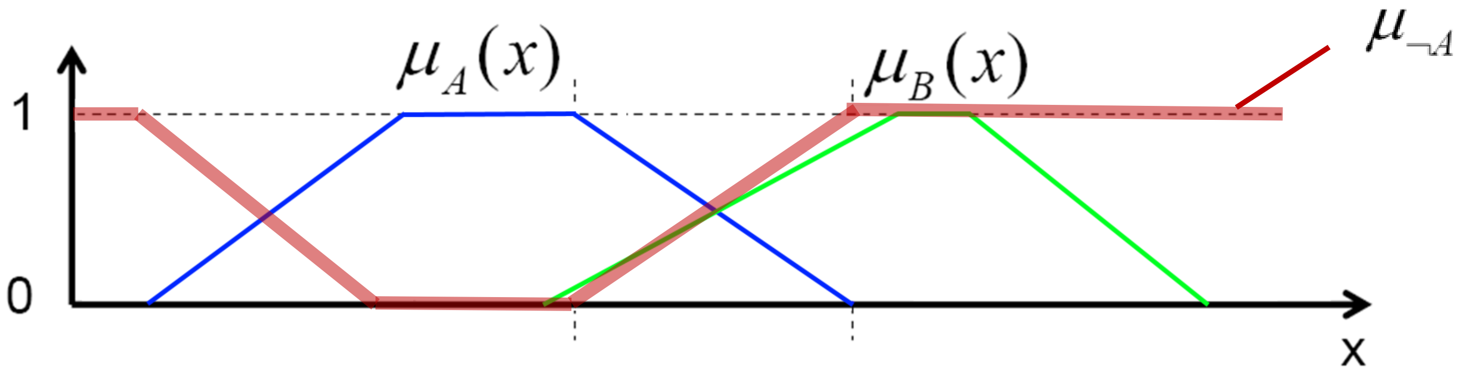
OPERATOREN: MIN/MAX - NORM

- Klassische Fuzzy-Operatoren: Min/Max-Norm
  - Konjunktion: $\mu_{A \wedge B}(x) := \min\{\mu_A(x), \mu_B(x)\}$
  - Disjunktion: $\mu_{A \vee B}(x) := \max\{\mu_A(x), \mu_B(x)\}$



OPERATOREN: MIN/MAX - NORM

- Klassische Fuzzy-Operatoren: Min/Max-Norm
  - Konjunktion: $\mu_{A \wedge B}(x) := \min\{\mu_A(x), \mu_B(x)\}$
  - Disjunktion: $\mu_{A \vee B}(x) := \max\{\mu_A(x), \mu_B(x)\}$
  - Negation: $\mu_{\neg A}(x) := 1 - \mu_A(x)$



T-NORM UND S-NORM

- Operatoren für Konjunktion (T-Norm) und Disjunktion (S-Norm) müssen zueinander passen

- Können mit De Morganscher Regel ineinander überführt werden:

$$(A \wedge B = \neg(\neg A \vee \neg B))$$

$$S(u,v) = 1 - T(1-u,1-v)$$
$$T(u,v) = 1 - S(1-u,1-v)$$

- Beispiel Min/Max-Norm:

$$\max(u,v) = 1 - \min(1-u,1-v)$$

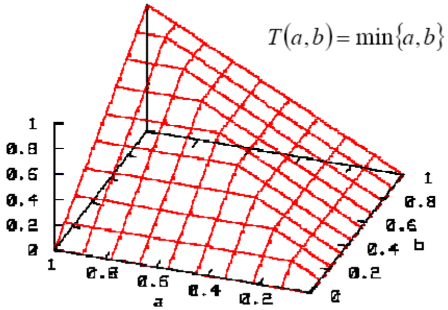
$$\ldots = 1-1-\min(-u,-v) = \max(u,v)$$

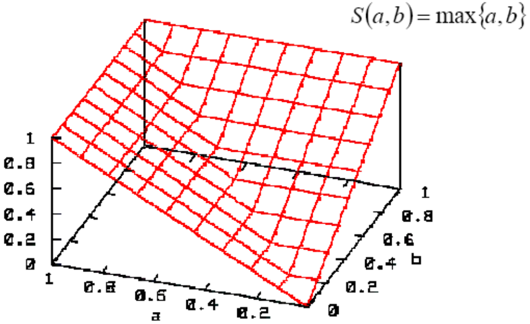
FUZZY T- UND S-NORMEN

T-Norm

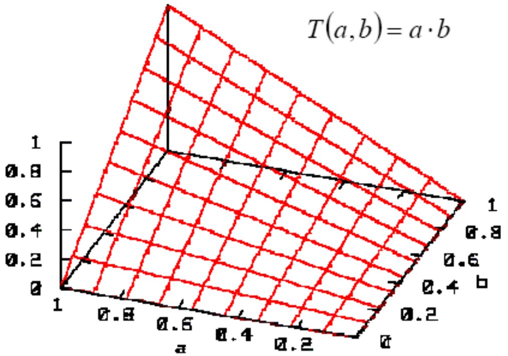
S-Norm

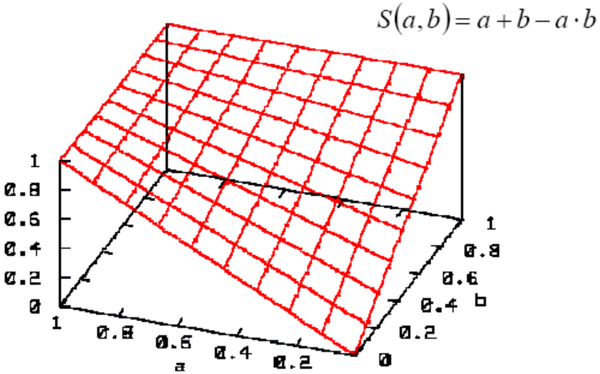
Min-Max





Product-Sum





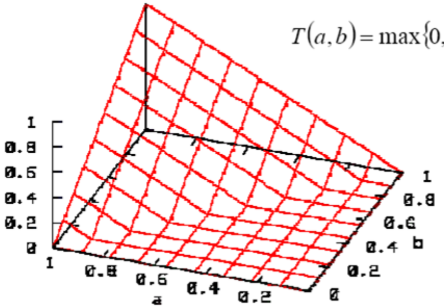
EXTREME FUZZY T- UND S-NORMEN

T-Norm

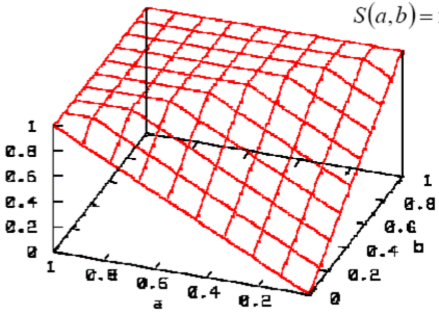
S-Norm

Lukasiewicz Norm

$$T(a,b) = \max\{0, a+b-1\}$$

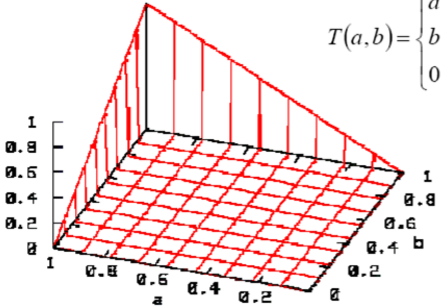


$$S(a,b) = \min\{1, a+b\}$$

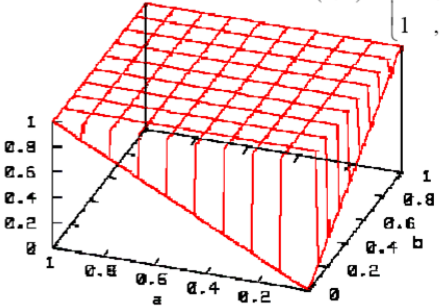


Drastic Product/Sum

$$T(a,b) = \begin{cases} a & ,\text{if } b=1 \\ b & ,\text{if } a=1 \\ 0 & ,\text{else} \end{cases}$$

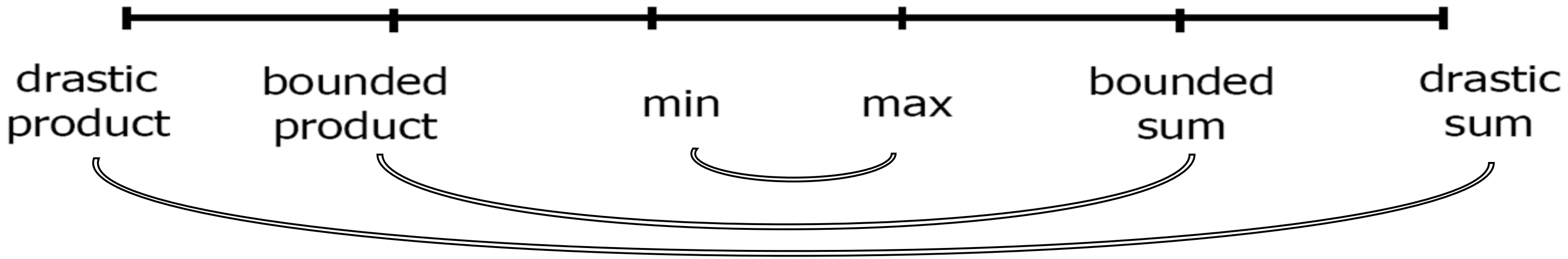


$$S(a,b) = \begin{cases} a & ,\text{if } b=0 \\ b & ,\text{if } a=0 \\ 1 & ,\text{else} \end{cases}$$



OPERATOR-SPEKTRUM

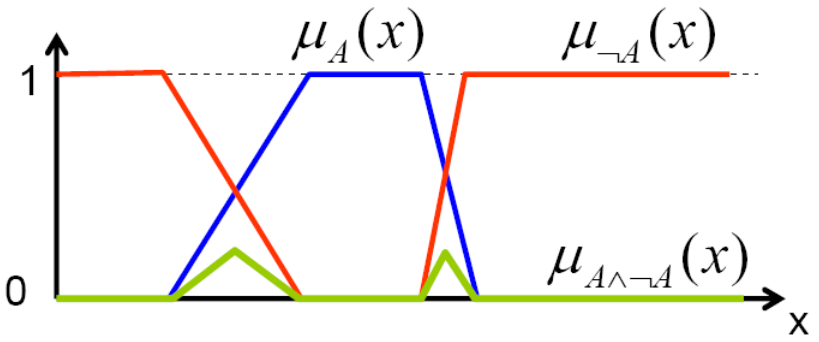
- T-Norm
- T-Konorm (S-Norm)



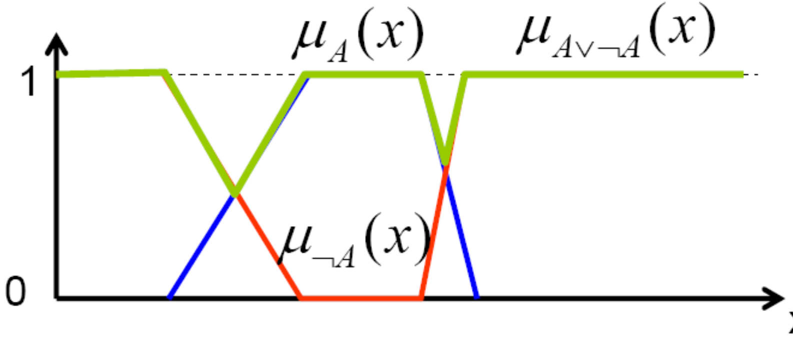
- Es gibt unterschiedlich extreme Operatoren
- T-Norm und S-Norm müssen zusammen passen

FUZZY NORMEN: PROBLEME

- Interessante Effekte:
  - $A \wedge \neg A = ?$



  - $A \vee \neg A = ?$



MITGLIEDSCHAFT VS. WAHRSCHEINLICHKEIT

- Wahrscheinlichkeit:
  - Häufigkeit von Zufallsereignissen
- Mitgliedschaft:
  - Grad der Zugehörigkeit zu einem Konzept
- Bsp: Krankheitswahrscheinlichkeit vs.
  Grad der Zugehörigkeit zu den Kranken

FUZZY-IMPLIKATION

- Eine Möglichkeit:
  - Ableitung von Tautologie $A \rightarrow B = \neg A \vee (A \wedge B)$
    und Min/Max-Norm

$$\mu_{A\rightarrow B}(x) := \underbrace{\max}_{*\vee*}\{\underbrace{1-\mu_A(x)}_{\neg A}, \underbrace{\min\{\mu_A(x), \mu_B(x)\}}_{A\wedge B}\}$$

FUZZY-IMPLIKATION

- Oder umgekehrt:
  - Starte mit Lukasiewicz-Implikation

$$\mu_{A\to B}(x) := \min\{1, 1-\mu_A(x)+\mu_B(x)\}$$

und leite Lukasiewicz-Disjunktion und -Konjunktion ab

- A∨B = ¬A→B

$$\mu_{A\vee B}(x) := \min\{1, \mu_A(x)+\mu_B(x)\}$$

- A∧B = ¬ (¬A∨¬B) // de Morgan

$$\mu_{A\wedge B}(x) := 1-\min\{1, 1-\mu_A(x)+1-\mu_B(x)\} \quad \text{// mit Lukasiewicz-Disjunktion}$$

$$= \max\{0, \mu_A(x)+\mu_B(x)-1\}$$

FUZZY-FOLGERUNGEN

- Klassischer Modus Ponens:

$$\frac{\begin{array}{l} A \\ A \rightarrow B \end{array}}{B} \qquad\qquad \frac{\begin{array}{l} \neg A \\ A \rightarrow B \end{array}}{?}$$

- Fuzzy: Generalisierter Modus Ponens

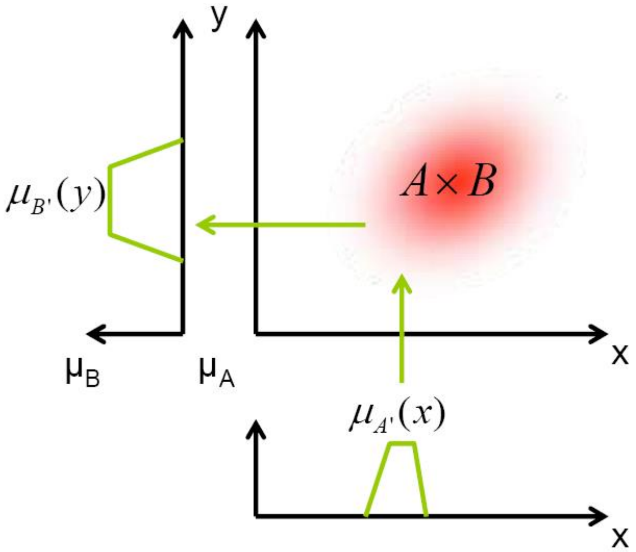
$$\frac{\begin{array}{l} \mu_{A'}(x) \\ A \rightarrow B \end{array}}{\mu_{B'}(y)}$$

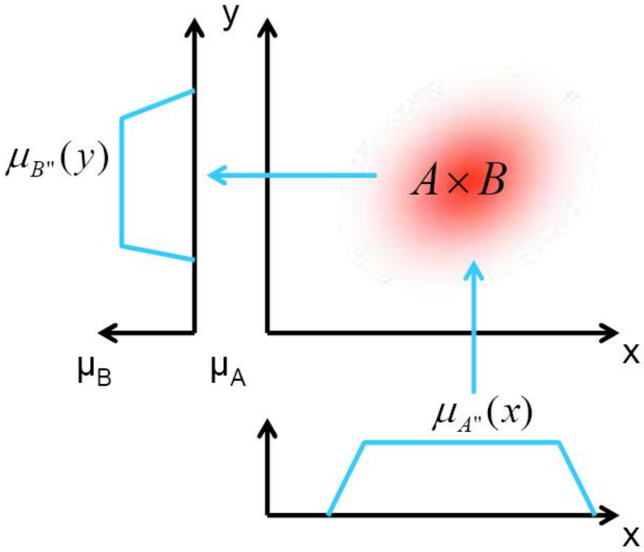
GEMEINSAMER CONSTRAINT: SUPPORT-VERTEILUNG

- Mit Min/Max-Norm führt Implikation A->B zu Einschränkungen des Kartesischen Produkts A×B

$$\mu_{A\times B}(x,y)=\min\{\mu_A(x),\mu_B(y)\}$$

$$\mu_{B'}(y)=\sup_x\{\min\{\mu_{A'}(x),\mu_{A\times B}(x,y)\}\}$$





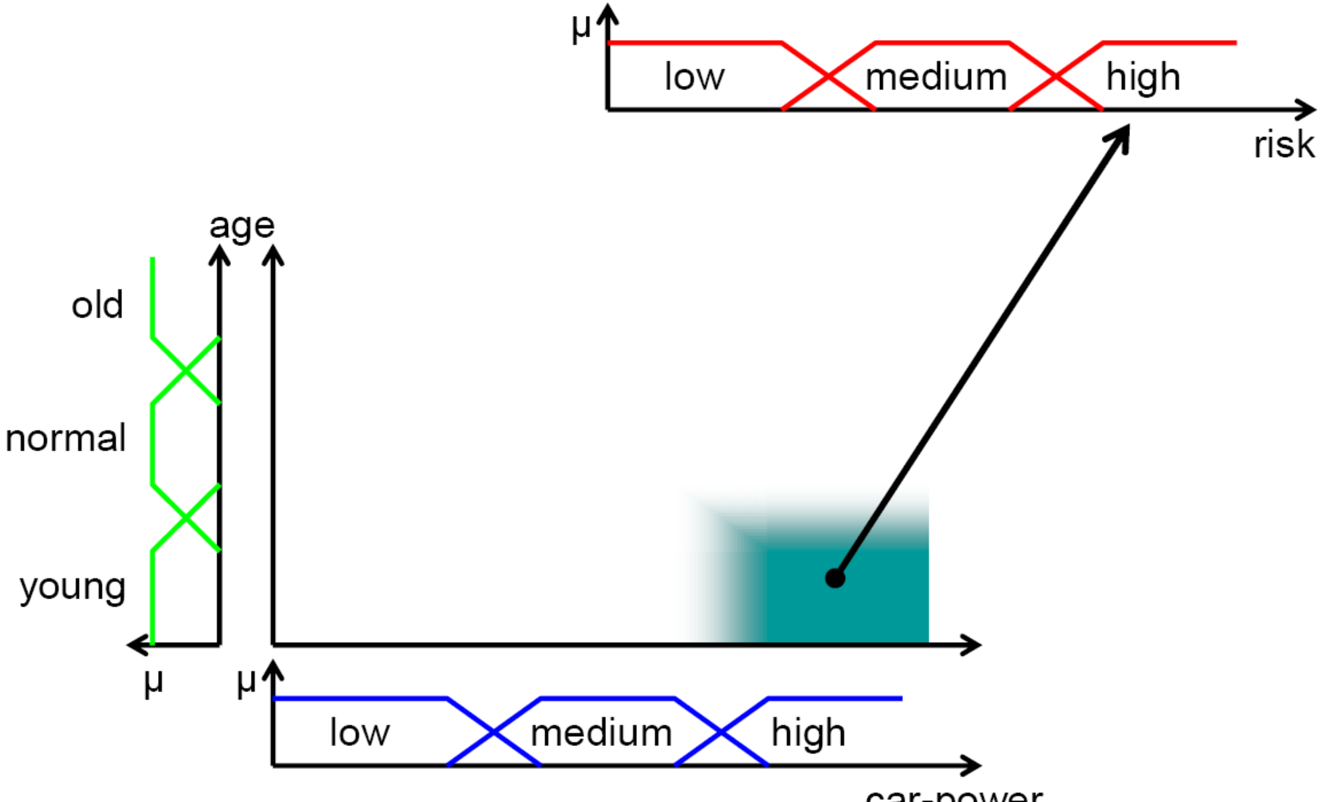
FUZZY-REGELN

- Regel:
  IF <Prämisse> THEN <Konklusion>

- Fuzzy-Version 1: Mamdani-Regeln
  - Prämisse: Konjunktion (Und) von Fuzzy-Mitgliedschaften
  - Konklusion: Fuzzy-Menge

BEISPIEL: MAMDANI-REGEL

IF age IS young AND car-power IS high THEN risk IS high



FUZZY-REGELN

- Regel:
  IF <Prämisse> THEN <Konklusion>

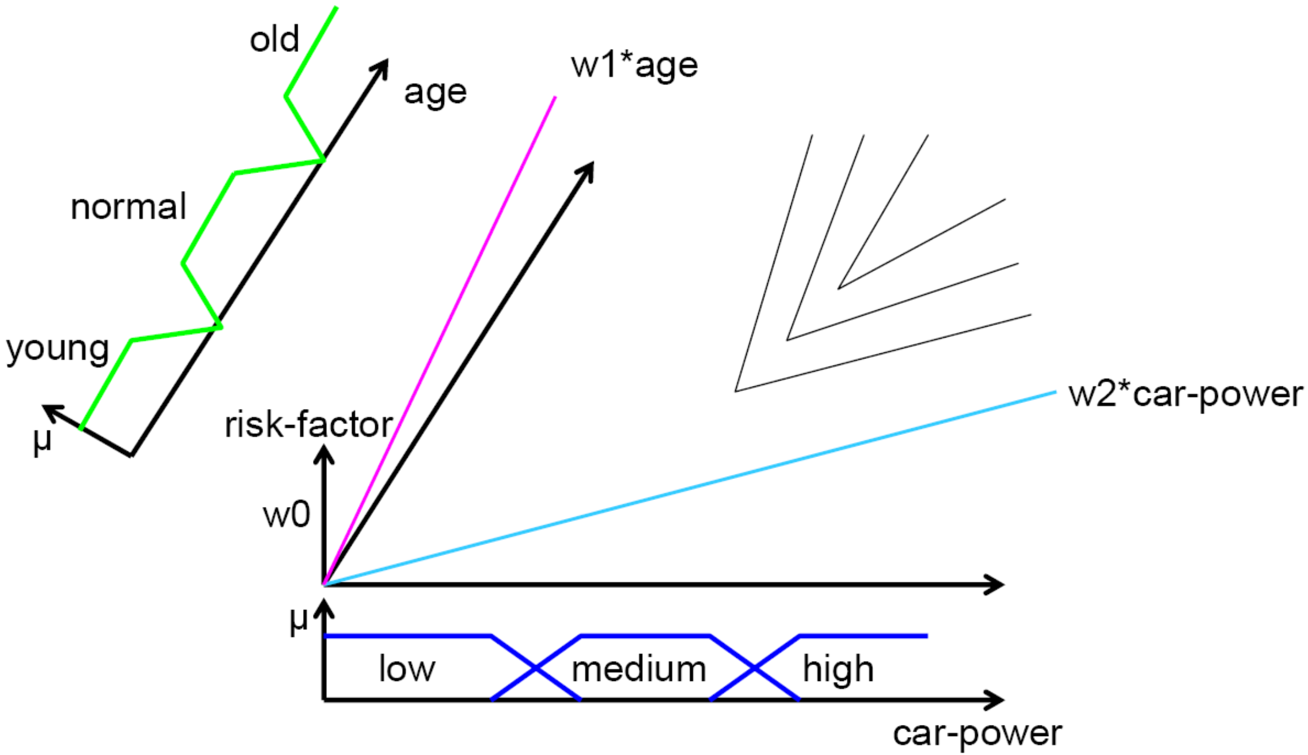
- Fuzzy-Version 1: Mamdani-Regeln
  - Prämisse: Konjunktion (Und) von Fuzzy-Mitgliedschaften
  - Konklusion: Fuzzy-Menge

- Fuzzy-Version 2: Takagi-Sugeno-Regeln
  - Prämisse: Konjunktion (Und) von Fuzzy-Mitgliedschaften
  - Konklusion: (normalerweise) reellwertige Funktion vom Grad 0-2.

BEISPIEL: TAKAGI-SUGENO-REGEL

IF age IS young AND car-power IS high

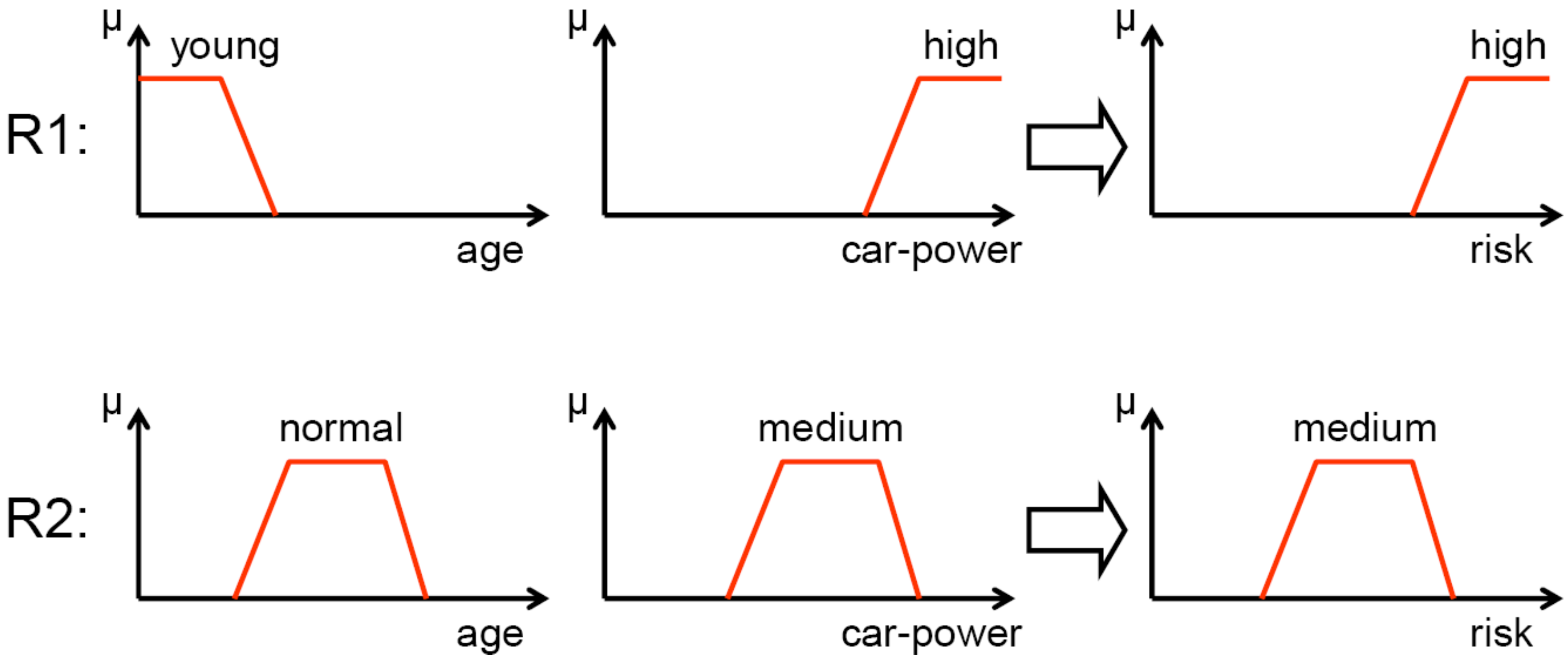
THEN risk-factor = w0+w1*age+w2*car-power



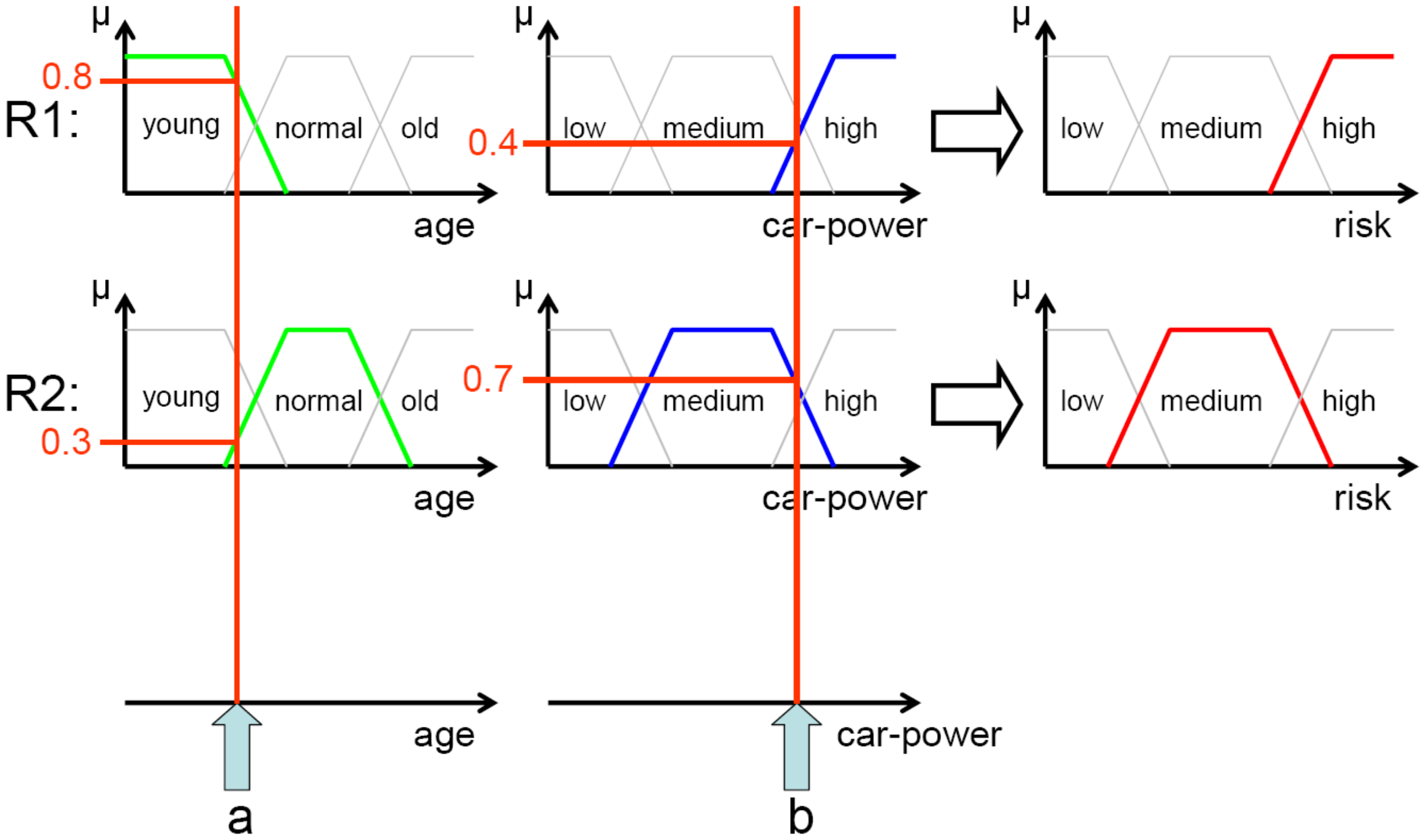
FUZZY-REGELSYSTEM

R1: IF age IS young AND car-power IS high THEN risk IS high

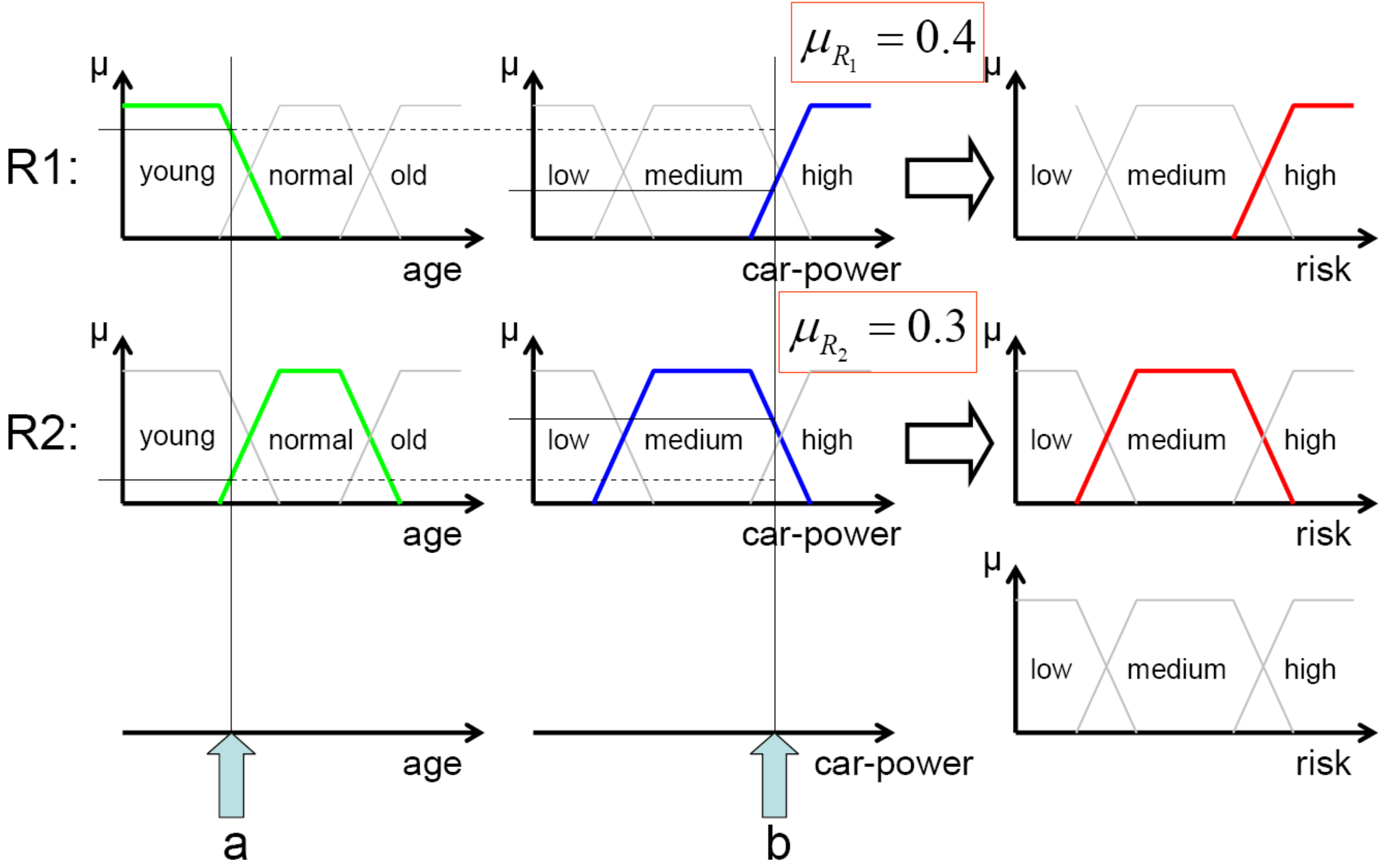
R2: IF age IS normal AND car-power IS medium THEN risk IS medium



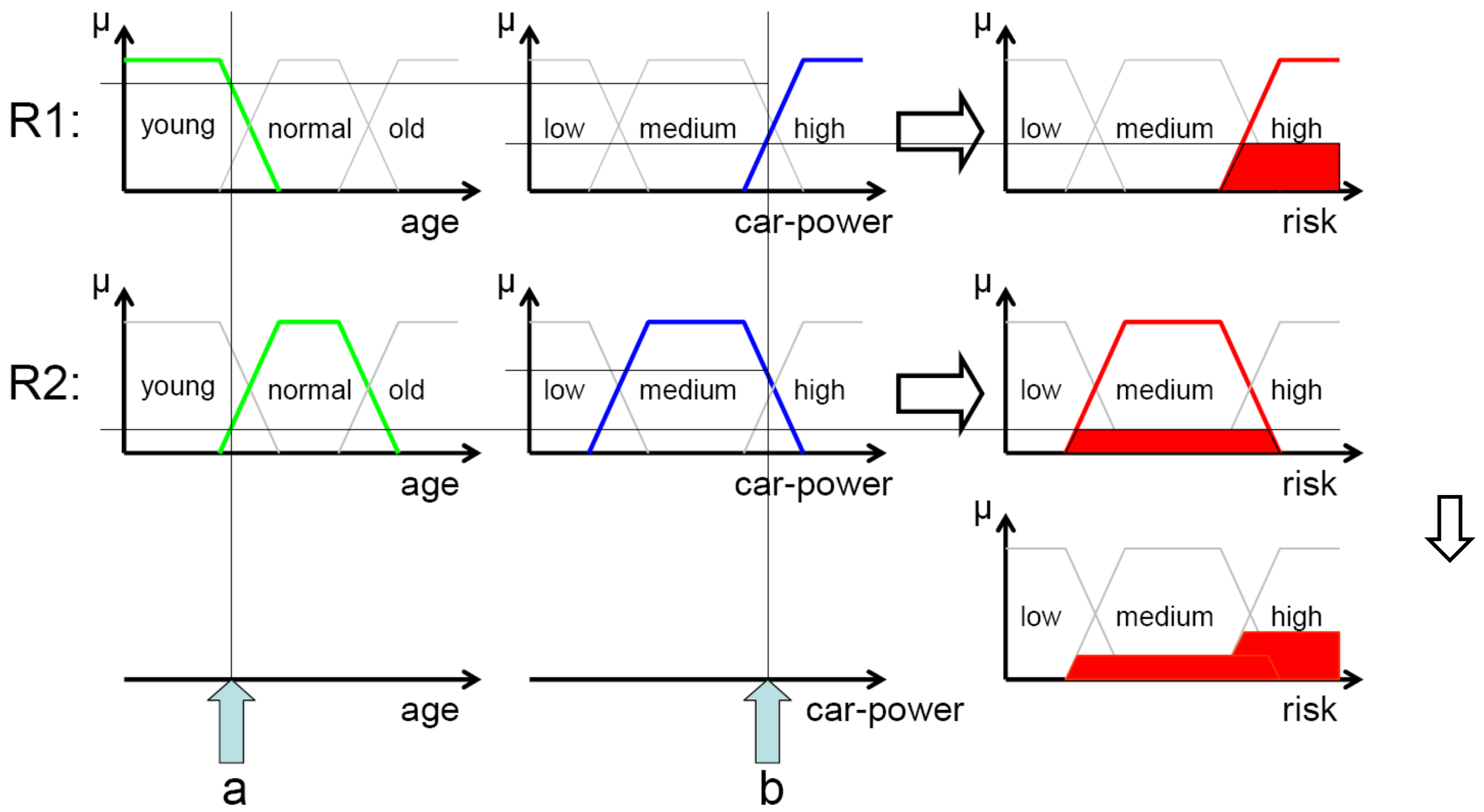
FUZZYFIZIERUNG SCHARFER EINGABEN



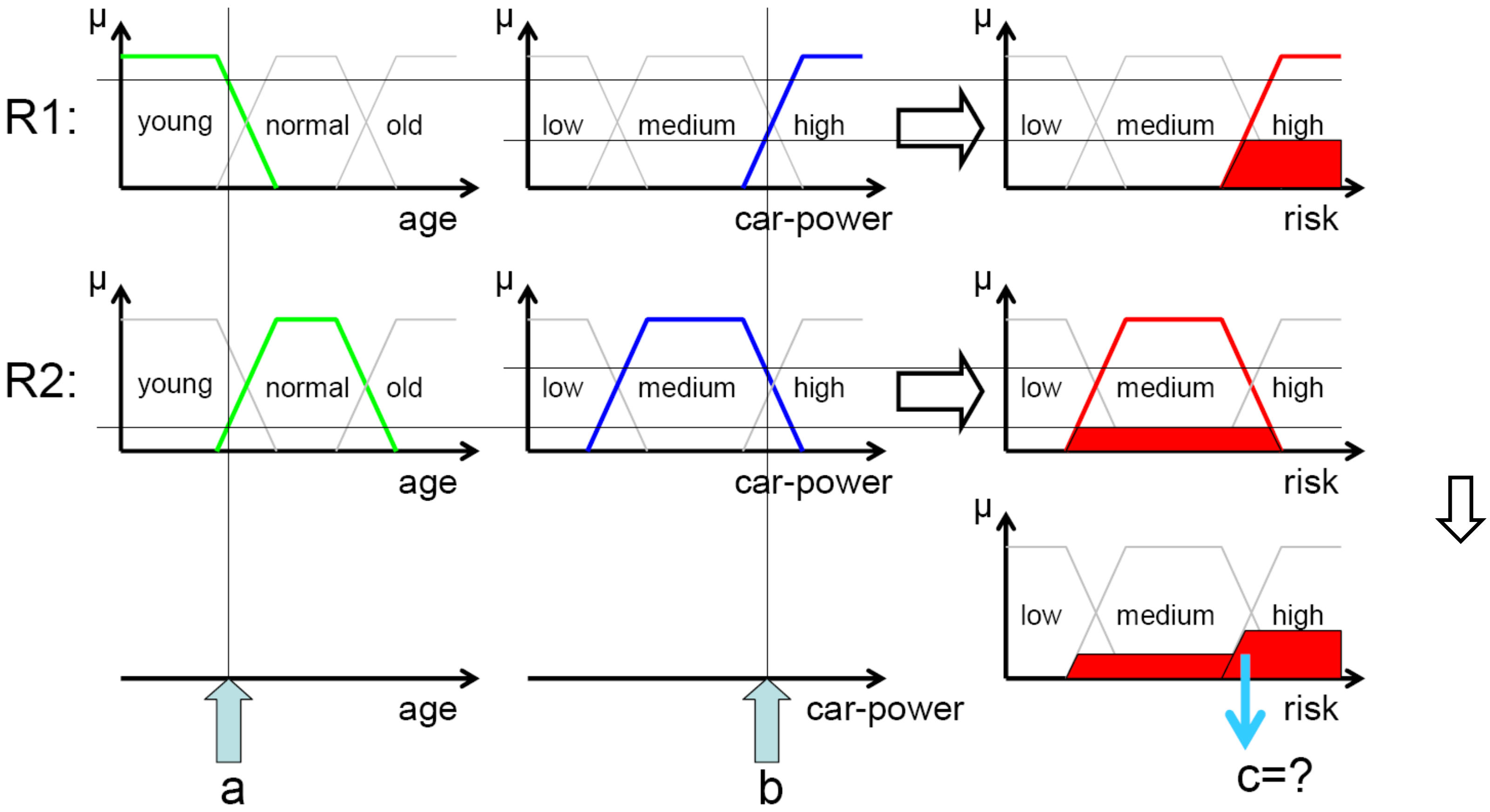
INFERENZ (Z.B. VIA MIN/MAX-NORM)



INFERENZ (Z.B. VIA MIN/MAX-NORM)



DEFUZZYFIZIERUNG



DEFUZZYFIZIERUNG

- Schwerpunkt:

$$y = \frac{\int y \cdot \mu_{\mathrm{risk}}(y)\mathrm{d}y}{\int \mu_{\mathrm{risk}}(y)\mathrm{d}y}$$

- Approximation: Gewichtete Summe der einzelnen Ausgabe-Fuzzymengen

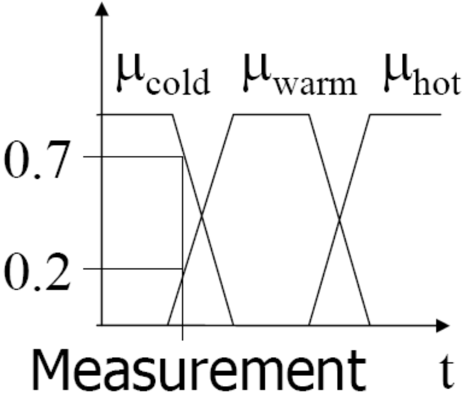
Aktivierung der Regel j

Schwerpunkt der Ausgabe-Fuzzymenge j

$$y = \frac{\sum_{j=1}^{r} \mu_j \cdot s_j}{\sum_{j=1}^{r} \mu_j}$$

FUZZY INFERENZ (MAMDANI)

→ Fuzzification → Inference → Defuzzification →



Fuzzy Rules

if temp is cold
then valve is open

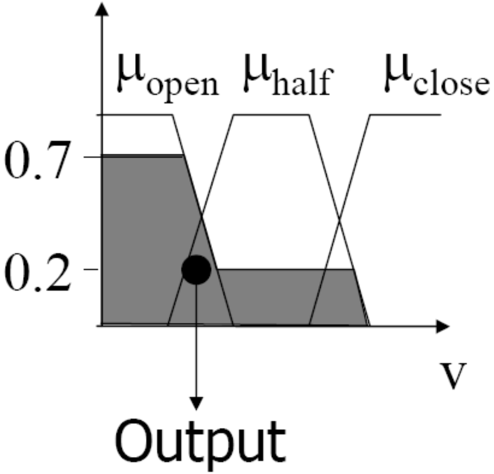
$\mu_{cold} = 0.7$

if temp is warm
then valve is half

$\mu_{warm} = 0.2$

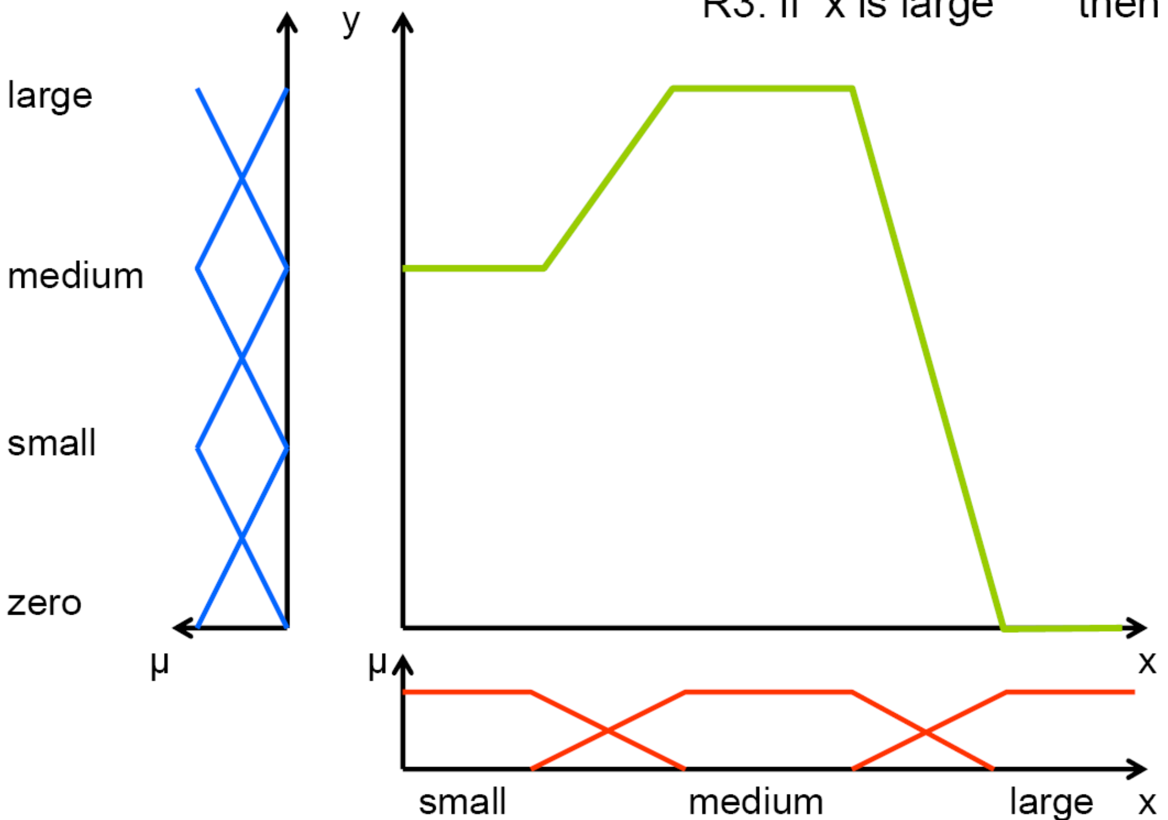
if temp is hot
then valve is close

$\mu_{hot} = 0.0$



FUZZY-REGELSYSTEM (MAMDANI)

R1: if x is small then y is medium
R2: if x is medium then y is large
R3: if x is large then y is zero



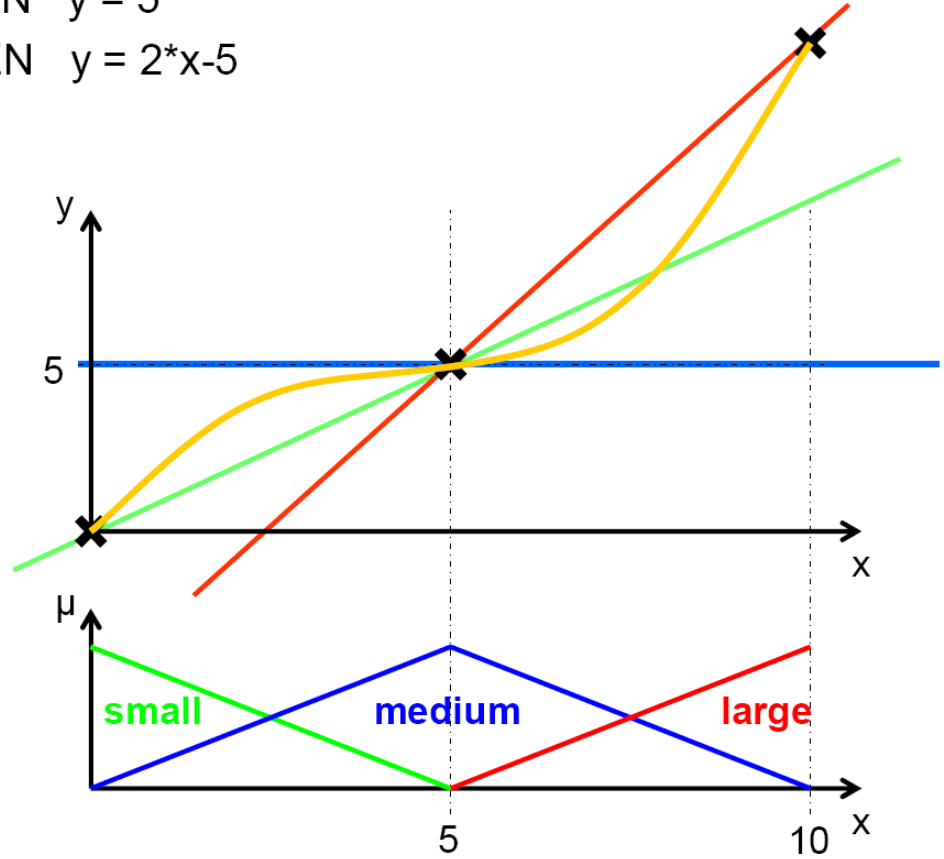
FUZZY-REGELSYSTEM (TAKAGI-SUGENO)

R1: IF x IS small THEN y = x

R2: IF x IS medium THEN y = 5

R3: IF x IS large THEN y = 2*x-5

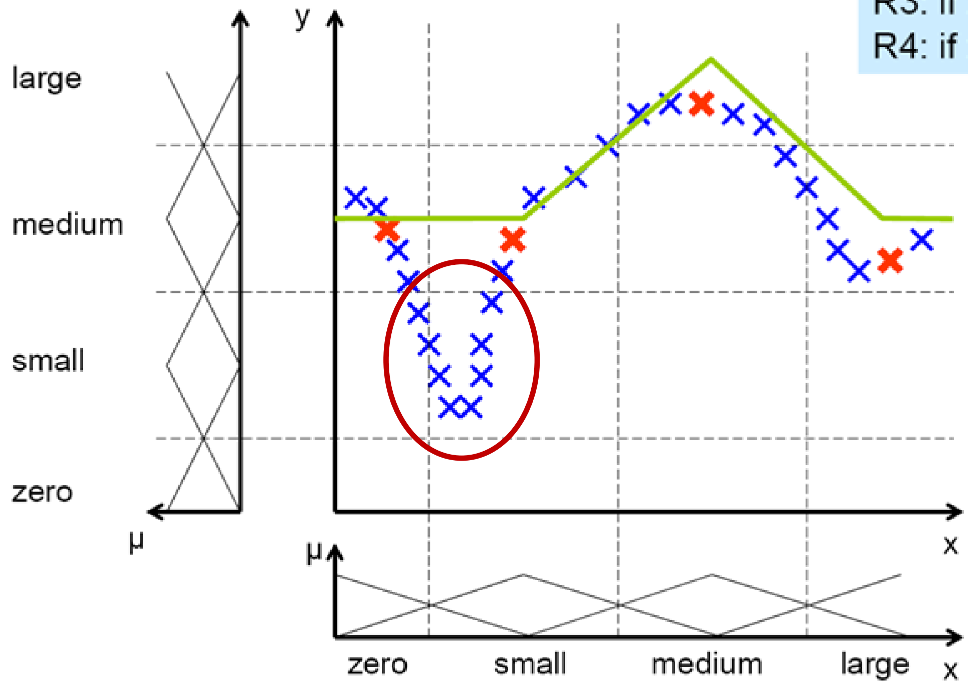
$$y = \frac{\sum_{i=1}^{r} \mu_{R_i} \cdot y_i(\vec{x})}{\sum_{i=1}^{r} \mu_{R_i}}$$



KONSTRUKTION VON FUZZY-REGELSYSTEMEN
WANG & MENDEL - ALGORITHMUS

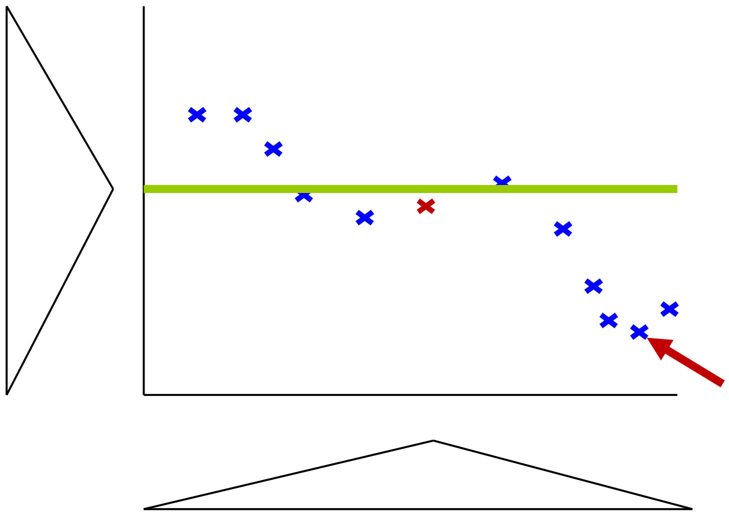
- Gegeben: Menge von Beispielen, vordefinierte Fuzzy-Mengen
- Beispiele mit höchster Mitgliedschaft bestimmen die Ausgabe

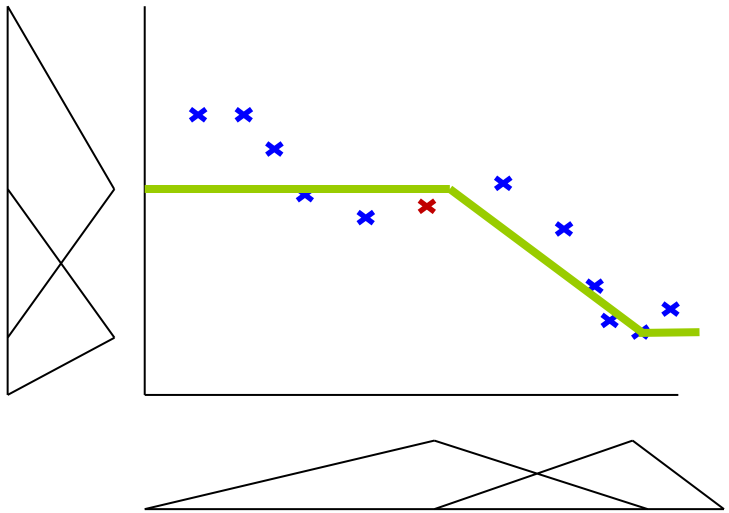
R1: if x is zero then y is medium
R2: if x is small then y is medium
R3: if x is medium then y is large
R4: if x is large then y is medium



- Exponentiell viele Regeln in hochdimensionalen Räumen
- Beliebig gute Approximation wenn Grid-Auflösung fein genug (aber zu welchen Kosten!)
- Falsche Wahl des Grids kann Extrema verpassen

HIGGINS & GOODMAN - ALGORITHMUS





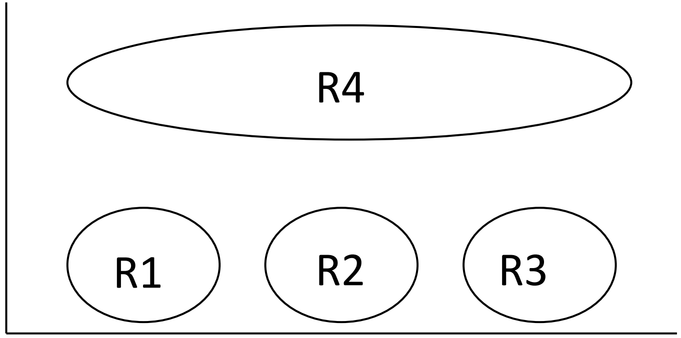
- Starte mit nur einer Mitgliedschaftsfunktion
- Füge neue ein wo der Fehler am größten ist
- Vordefinierter Schwellwert für erwünschten Approximationsfehler
- Findet “beste” Partitionierung (=Grid)

- Nachteile:
  - Konzentration auf Outlier
  - Interpretation schwierig: Granularität hängt nur von Daten ab

ANDERE KONSTRUKTIVE ALGORITHMEN

- Lokale Mitgliedschaftsfunktionen
- Teilweise vordefinierte Mitgliedschaftsfunktionen/Granularitäten möglich
- Teilweise Toleranz gegen Outlier durch Relevanz-Filter

FREIE FUZZY-REGELN



- Keine globale Granularität:
  - Individuelle (pro Regel) Mitgliedschafts-Funktionen
  - Bessere Modellierung lokaler Eigenschaften
- Nicht alle Attribute in jeder Regel verwendet:
  - Individuelle Auswahl der Prämisse durch Beschränkung auf wenige Attribute
  - Bessere Interpretierbarkeit in hochdimensionalen Räumen
  - Kein exponentielles Wachstum der Regelanzahl mit der Dimensionalität

ERZEUGUNG FREIER FUZZY-REGELN

- Ziel: Klassifikation des Eingaberaums entsprechend Trainingsmenge $\{(x_i, c_i)\}$

Algorithmus FRL zur Klassifikation:

FORALL training examples (**x**, c) DO

IF correct rule of class c exists:

- COVERED:
  - increase weight +1
  - adjust core region of rule to cover **x**

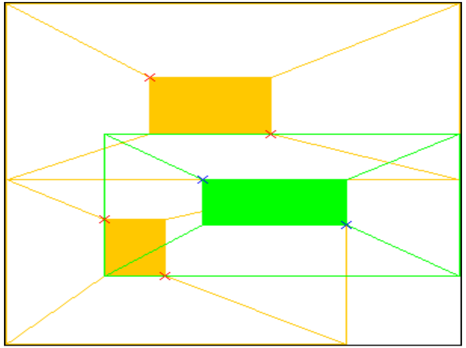
ELSE:

- COMMIT:
  - insert new rule with core=**x**
  - Support = infinite (i.e. rule is not constrained)

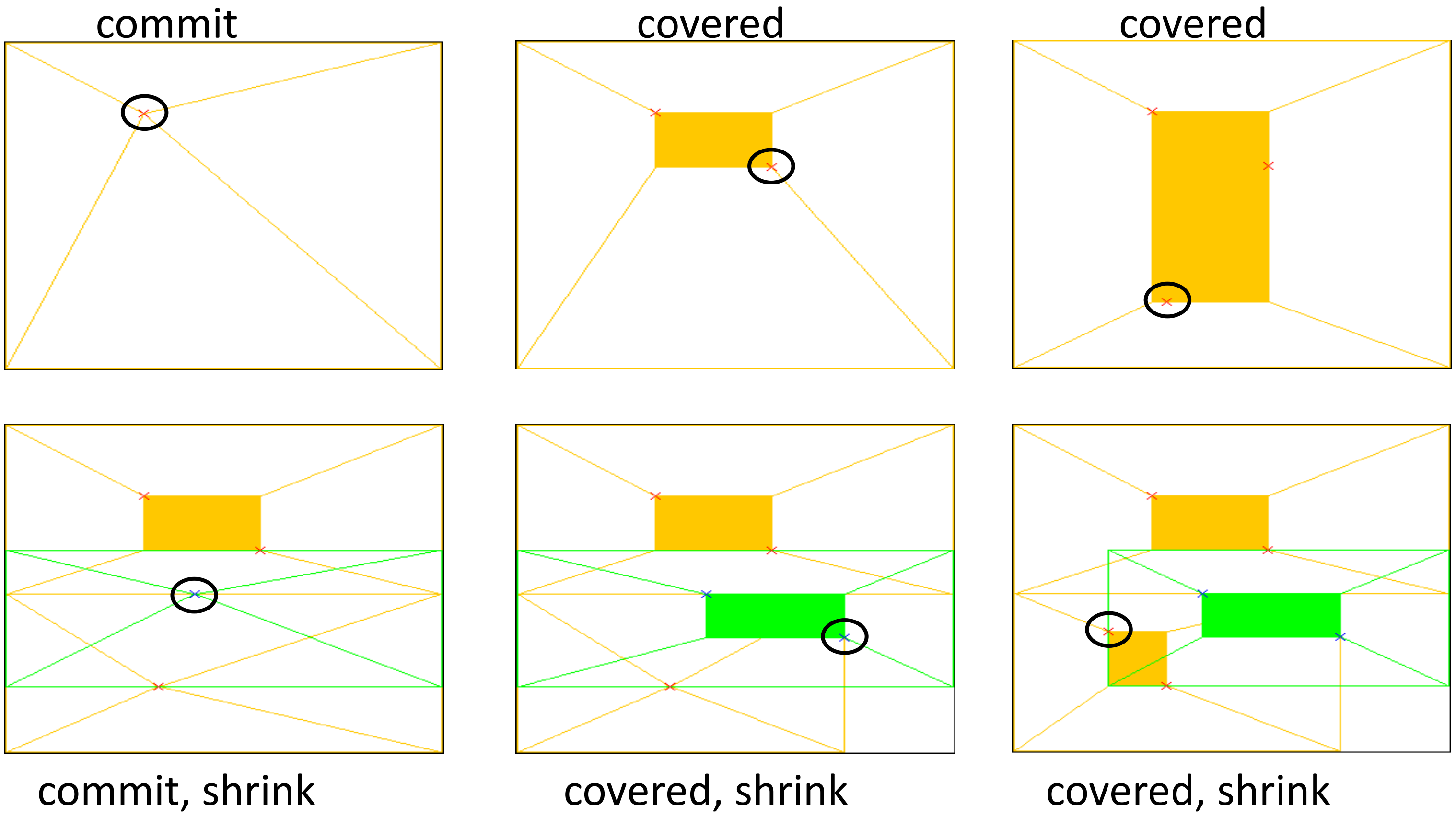
SHRINK:

- Reduce support of all rules of conflicting class that cover **x** (e.g. try to keep largest volume).

Until no more changes occurred.



BEISPIEL: FUZZY-KLASSIFIKATOR



ERZEUGUNG FREIER FUZZY-REGELN

Beobachtungen:

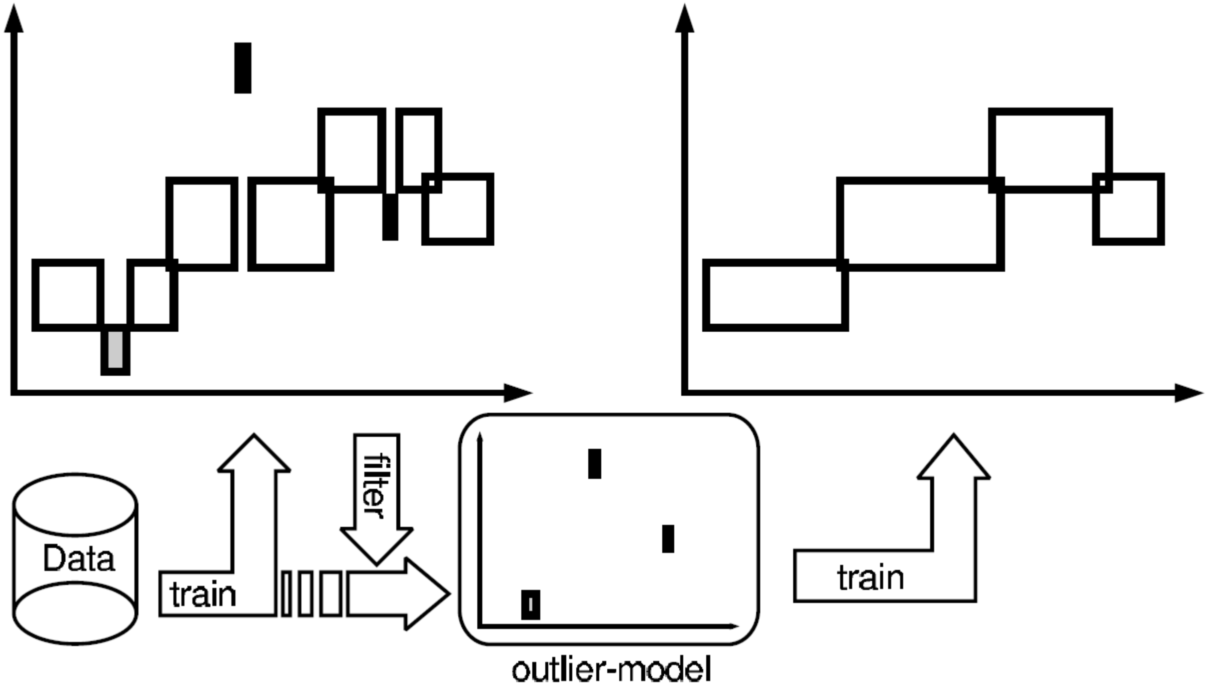
- FRL findet Regelmenge, die die Daten vollständig beschreibt (wenn Daten ohne Konflikte sind …)

- Jede Regel ist eine *partielle Hypothese für eine Teilmenge der Trainingsdaten*
  - **Core**: Am meisten spezifische Hypothese, die eine Teilmenge der Daten abdeckt
  - **Support**: eine der generellsten Hypothesen, die eine Teilmenge der Daten abdeckt
  - Support **is_more_general_than** Core.

- Core und Support-Regionen können interpretiert werden als:
  - Kleinstes Gebiet mit höchster Konfidenz (Wir haben Evidenz)
  - Größtes Gebiet ohne Konflikte (Wir haben noch keine Gegenbeispiele gesehen)

OUTLIER-FILTER

- Problem: Verrauschte Daten erzeugen viele Regeln mit wenigen Datenpunkten
- Ansatz: Entferne Datenpunkte, die von Regeln mit geringem Gewicht beschrieben werden und trainiere erneut



[Berthold 2000]

ANDERE FUZZY-LERNMETHODEN

- Konstruktive Methoden:
  - Erzeuge Fuzzy-Regeln durch Wachstum aus Singleton-Regeln
  - FRL verkleinert Regeln ausgehend von den generellsten Regeln, bis diese zu den Daten passen
- GRID-basierte Methoden:
  - Verschmelze Gridzellen (oder Zeilen/Spalten) wenn keine Datenpunkte enthalten sind oder dieselbe Klasse vorhergesagt wird
- Adaptive Methoden:
  - Initialisiere Regeln zufällig (oder mit Expertenwissen) und optimiere Regel-Parameter (Orte, manchmal auch die Anzahl der Mitgliedschaftsfunktionen) iterativ (Gradientenabstieg, heuristische Optimierung wie z.B. Hill Climbing).
- Neuro-Fuzzy-Methoden:
  - Initialisiere neuronales Netz mit den Regeln und trainiere dieses mit Backpropagation

BEISPIEL: AUTOMATIK-GETRIEBE

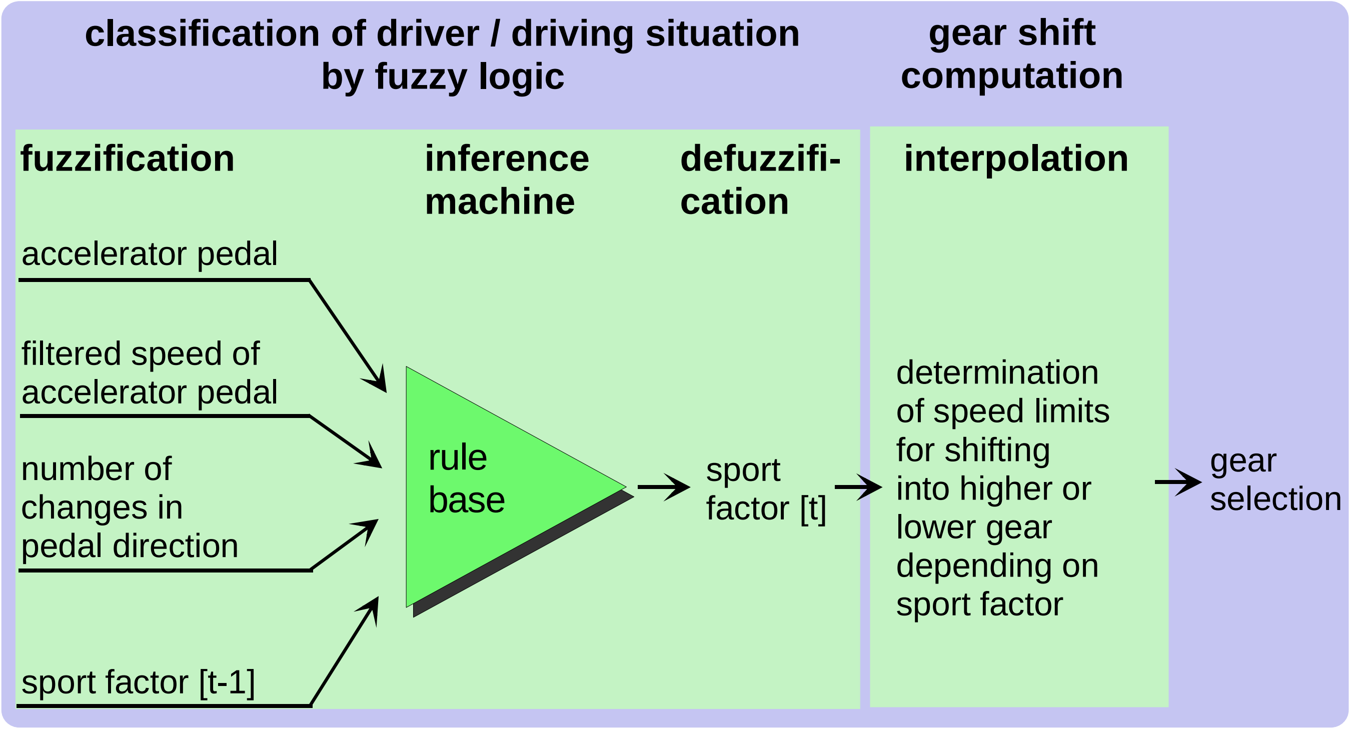
- Aufgabe: Verbesserung des VW-Automatik-Getriebes
  - Keine zusätzlichen Sensoren
  - Individuelle Anpassung des Schaltverhaltens

- Idee (1995):
  - Das Fahrzeug “beobachtet” und klassifiziert den Fahrer nach Sportlichkeit
    - ruhig, normal, sportlich => Bestimmung eines Sport-Faktors aus [0, 1]
    - nervös => Beruhigung des Fahrers

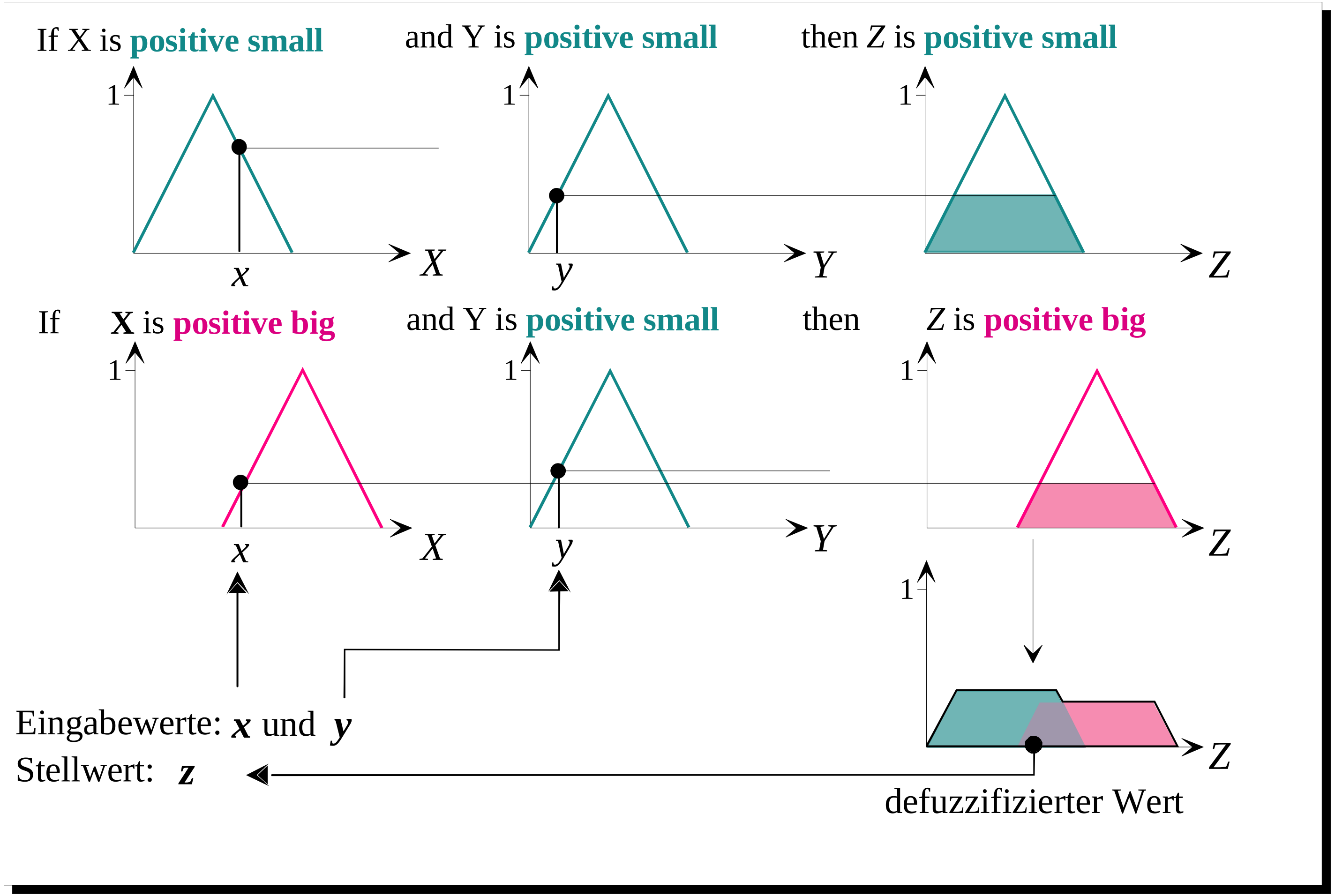
- Datenerfassung im Testfahrzeug:
  - Verschiedene Fahrer, Klassifikation durch Experten (Mitfahrer)
  - Gleichzeitige Messungen:
    - Geschwindigkeit,
    - Position,
    - Geschwindigkeit des Gaspedals,
    - Winkel des Lenkrades, ...

BEISPIEL : AUTOMATIK-GETRIEBE

- Überblick über das Gesamtsystem



FUZZY-INFERENZ

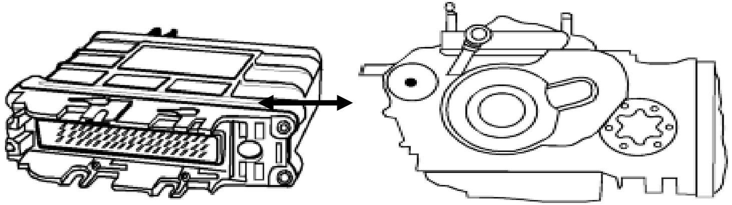


ERGEBNIS: ADAPTIVES SCHALTVERHALTEN

- Fuzzy-Regler mit 7 Regeln
- Optimiertes Programm

$$\left.\begin{array}{l} -\ \ 24\ \text{Byte RAM} \\ -\ 702\ \text{Byte ROM} \end{array}\right\}\ \text{auf Digimat}$$

AG 4



- Laufzeit 80 ms,
  12 mal pro Sekunde wurde ein neuer Sportfaktor bestimmt
- In Serie im VW Konzern
- Erlernen von Regelsystemen mit Hilfe von Künstlichen Neuronalen Netzen, Optimierung mit evolutionären Algorithmen