



Vorlesung Künstliche Intelligenz Wintersemester 2008/09

Teil III: Wissensrepräsentation und Inferenz

Kap.7: Unsicherheit und Vagheit

Dieses Kapitel basiert auf Material von A. Abecker, welches auf Material basiert, das auf der nächsten Folie erwähnt wird.



- Die Darstellung zum Theorem von Bayes folgt weitgehend der von Dr. Andreas Tolk, Uni der Bundeswehr, München
- Die Folien zur Fuzzy-Logik fassen ein Tutorial der Uni Linz, Österreich, zusammen (Autoren Peter Bauer, Stephan Nouak, Roman Winkler, <http://www.flll.uni-linz.ac.at/>).
- Socher-Ambrosius / Heinsohn geben eine gute Einführung in Fuzzy-Logik, Bayes-Theorie, Sicherheitsfaktoren, etc.
- M.M. Richter: Prinzipien der Künstlichen Intelligenz. Teubner-Verlag, Stuttgart. Gibt einen umfassenden Einblick in existierende Ansätze zur Verarbeitung von Unsicherheit & Vagheit.



Subjektive Unschärfe

Vagheit

- **Fuzzy-Logik**
- Rough Sets

Meinung

- epistemische Logik

Gerade bei mathematisch fundierten Modellen muss man beachten, ob die theoretischen (z.B. Unabhängigkeitsannahmen) und praktischen Voraussetzungen (z.B. relative Häufigkeiten als Schätzung für bedingte Häufigkeiten) jeweils gegeben sind.

Objektive Unsicherheit

Unwissenheit

- gewisse Fakten nicht bekannt
 - Default-Reasoning
- Häufigkeitsverteilung der möglichen Werte kann bekannt sein
 - Stochastik, z.B. **Bayes-Theorem**
- Häufigkeitsverteilung kann auch unbekannt sein
 - Sicherheitsfaktoren
 - Evidenztheorie

Ungenauigkeit, z.B. Messungenauigkeiten

- Intervallarithmetik
- Qualitative Abstraktion
- **Fuzzy-Logik**



- Bekannt ist die *a-priori Wahrscheinlichkeit* einer Hypothese:

Wie häufig kommt die Hypothese in der Regel vor?

- Zudem ist die *bedingte Wahrscheinlichkeit* bekannt, dass eine Beobachtung gemacht wird, wenn eine Hypothese zutrifft.
- Gesucht ist die *wahrscheinlichste Hypothese*, wenn entsprechende Beobachtungen vorliegen!



- Wir wissen von einem Fachbereich, dass an ihm $1/3$ der Studenten Mathematik und $2/3$ Elektrotechnik studieren.
- Wir wissen die jeweiligen Durchfallquoten in der Vorlesung Analysis I: $1/5$ für die Mathematiker, $5/6$ für die E-Techniker.
- Dies ergibt die bedingten Wahrscheinlichkeiten dafür, dass ein Student besteht, wenn sein Studiengang bekannt ist.
- Nun betrachten wir einen einzelnen Studenten. Gesucht ist sein Studiengang.



Aus der Wahrscheinlichkeitstheorie ist bekannt:

$$P(H/B) = \frac{|B \cap H|}{|B|} = \frac{|H|}{|H|} \cdot \frac{|B \cap H|}{|B|} = |H| \cdot \frac{|H \cap B|}{|H|} \cdot \frac{1}{|B|} = P(H) \cdot \frac{P(B/H)}{P(B)}$$

- $P(H/B)$ Wahrscheinlichkeit für Hypothese H ,
wenn Beobachtung B vorliegt.
- $P(H)$ à priori Wahrscheinlichkeit von Hypothese H
- $P(B)$ à priori Wahrscheinlichkeit von Beobachtung B
- $P(B/H)$ Wahrscheinlichkeit, dass B beobachtet wird,
wenn die Hypothese H gilt.

Anwendung der Formel von Bayes im Beispiel



- Hypothese: Ein Mathematiker steht uns gegenüber.
- Wir wissen, daß $1/3$ der Studenten von diesem Typ sind, d.h. $P(H)=0.33$.
- Nun beobachten wir: Der Student hat bestanden!
- Was können wir daraus schließen?
- Aufgrund der Bestehensquote der Mathematiker gilt $P(B/H) = 4/5 = 0.8$.
- Aufgrund der Bestehensquoten und der $1/3 : 2/3$ - Verteilung wissen wir, dass $P(B) = 1/3 \cdot 4/5 + 2/3 \cdot 1/6 \approx 0.35$ für das Bestehen gilt.
- Schluss gemäß obiger Formel:
$$P(H/B) = P(H)/P(B) \cdot P(B/H) = 0.33 / 0.35 \cdot 0.8 = 0,75$$
- Das heißt: Wenn wir wissen, dass der Student bestanden hat, ist die Wahrscheinlichkeit, dass uns ein Mathematiker gegenübersteht, mehr als doppelt so hoch, als dies a priori der Fall ist!



Das Theorem lässt sich mit den folgenden beiden Sätzen auf mehrere Hypothesen und mehrere Beobachtungen erweitern:

Satz: Wenn die Ereignisse B_1, \dots, B_n stochastisch unabhängig sind, dann gilt $P(B_1 \wedge \dots \wedge B_n) = P(B_1) \cdot \dots \cdot P(B_n)$.

Satz: Sind die Hypothesen H_1, \dots, H_m paarweise disjunkt und decken sie den ganzen Wahrscheinlichkeitsraum ab, dann gilt $P(B) = \sum_{i=1..m} P(B | H_i) \cdot P(H_i)$.



Satz von Bayes: Wenn

- die Hypothesenmenge vollständig ist (Closed World Assumption),
- die Hypothesen paarweise disjunkt sind (Single Fault Assumption),
- die Beobachtungen unabhängig sind
- und die Beobachtungen nur von den Hypothesen abhängen,

dann gilt für die relative Wahrscheinlichkeit einer Hypothese :

$$P_{\text{rel}}(H_j/B_1 \wedge \dots \wedge B_n) = \frac{P(H_j) \cdot P(B_1/H_j) \cdot \dots \cdot P(B_n/H_j)}{\sum_{j=1}^m P(H_j) \cdot P(B_1/H_j) \cdot \dots \cdot P(B_n/H_j)}$$



In der Praxis kommen folgende Forderungen hinzu:

- Es müssen entsprechend vollständige Statistiken vorliegen, um die bedingten Wahrscheinlichkeiten aufstellen zu können.
- Die Wahrscheinlichkeiten müssen konstant sein oder der zeitliche Verlauf muss bekannt sein (inkl. der bereits geforderten Statistiken).



Ein Bayessches Netz ist ein gerichteter Graph:

- Knoten: Zufallsvariablen
- gerichtete Kanten: bedingte Wahrscheinlichkeiten

Eine Kante von A nach B existiert, wenn A eine kausale Ursache für B ist. (D.h. insbesondere dass die bedingte Wahrscheinlichkeit nicht Null ist.)

Ist die Wahrscheinlichkeit eines Knotens bekannt, kann die Wahrscheinlichkeit aller Folgeknoten berechnet werden.

Wenn bestimmte Knoten initialisiert wurden (weil der zugörige Fall eintrat), können Folgewahrscheinlichkeiten durch das Netz propagiert werden.



Fuzzy Logic ist im Kern eine mehrwertige Logik, die es erlaubt, Zwischenwerte zwischen konventionellen Bewertungen wie „wahr/falsch“, „ja/nein“, „schwarz/weiß“ zu definieren.

Sie erlaubt es, Begrifflichkeiten wie „ziemlich warm“ oder „etwas kühl“ mathematisch zu formulieren und mit Rechnern zu verarbeiten.

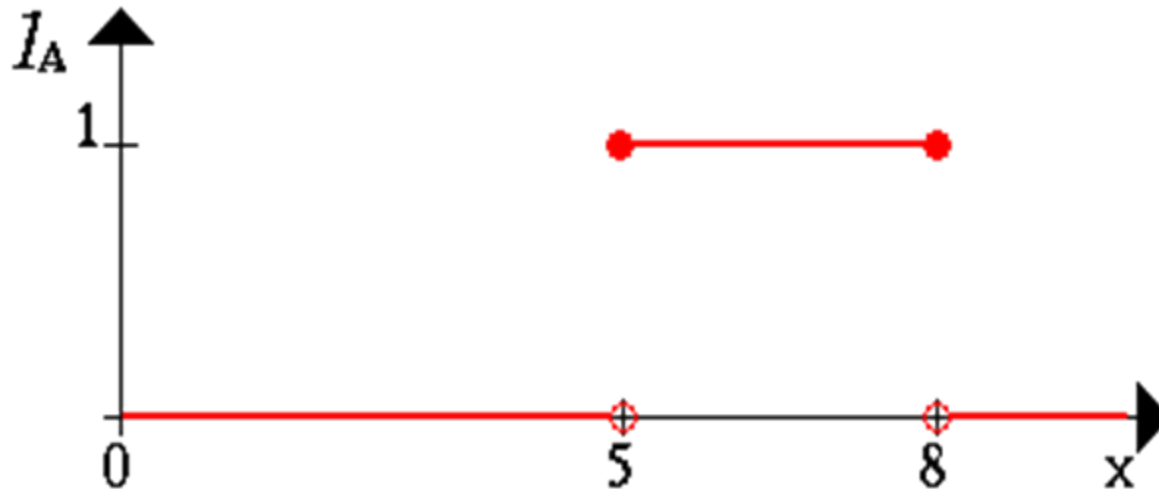
Auf diese Weise soll die Symbolverarbeitung des Computers der des Menschen ähnlicher werden.

Fuzzy Logic wurde 1965 von Lotfi A. Zadeh (University of California, Berkeley) eingeführt.

Klassische Mathematik: „crisp sets“



- Sei unser Universum X die Menge aller reellen Zahlen zwischen 0 und 10, und sei $A := [5, 8] \subseteq X$.
- Die Menge A kann durch ihre charakteristische Funktion beschrieben werden:



Noch ein Beispiel: die Menge aller jungen Leute



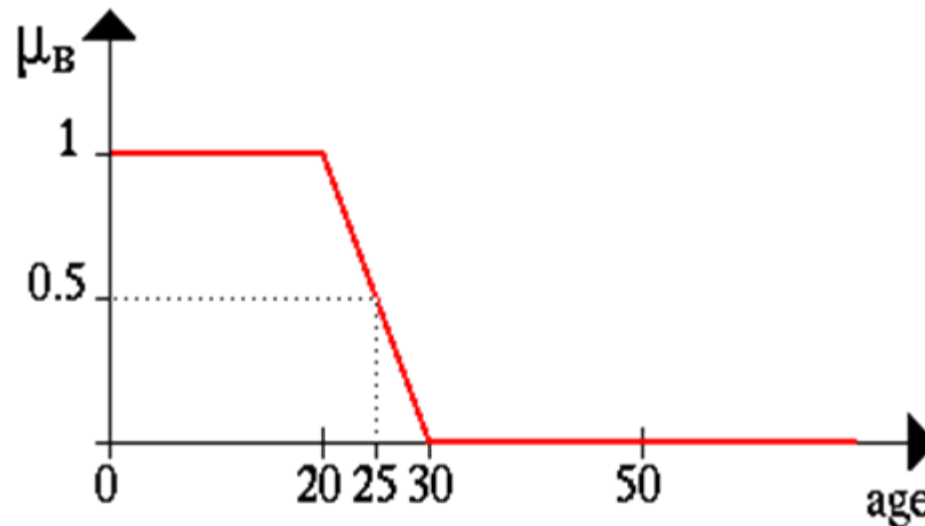
- Mögliche Modellierung: $B = [0,20]$???
- ABER: warum ist jemand an seinem 20 Geburtstag noch jung und genau einen Tag später überhaupt nicht mehr?
- Das Verschieben der Grenze verschiebt nur das Problem: dies ist ein strukturelles Problem der Modellierung mit harten Intervallen.
- Natürlicherer Weg: Aufweichen der strengen Trennung zwischen „jung“ und „nicht mehr jung“.
- Ansatz:
 - Erlaube nicht mehr nur die harte Aussage „JA! Er/sie ist in der Menge der jungen Leute enthalten.“ oder „NEIN! Er/sie ist nicht in der Menge enthalten.“,
 - stattdessen auch Aussagen wie „Na ja, er/sie gehört ein wenig mehr zur Menge der jungen Leute.“ oder „Nein, er/sie gehört fast gar nicht mehr dazu.“

Fuzzy Membership durch allgemeinere charakteristische Funktionen



- Statt 0 und 1 erlauben wir beliebige reelle Zahlen im Intervall von 0 bis 1 als Werte der charakteristischen Funktion einer Menge $[0, 1]$.
- Wir interpretieren 1 als „Element von B“ und 0 als „kein Element von B“. Alle anderen Werte bezeichnen eine graduelle Mitgliedschaft in der Menge B.

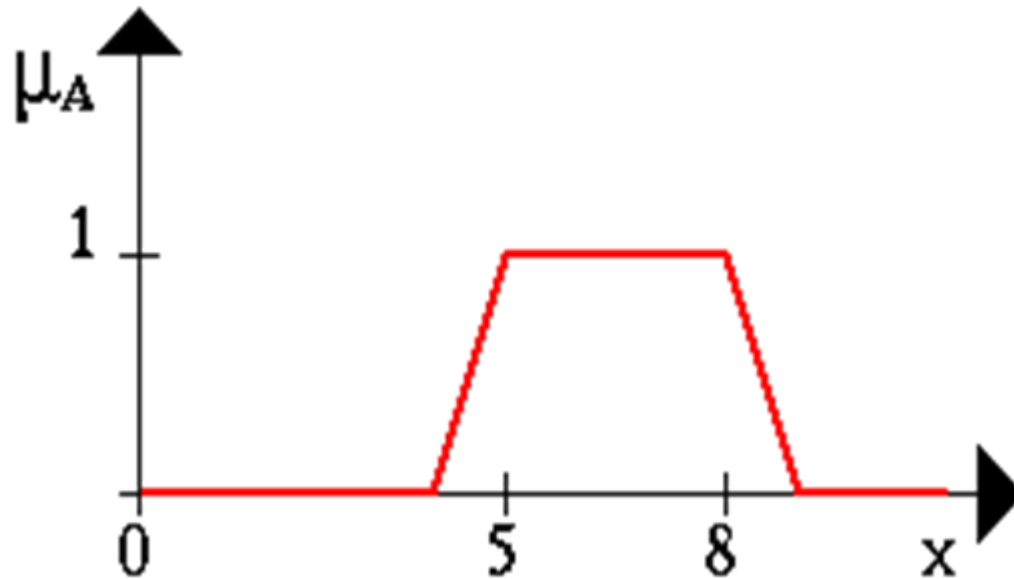
- Im Beispiel:



Wie definiert man Mengenoperationen auf Fuzzy Sets?

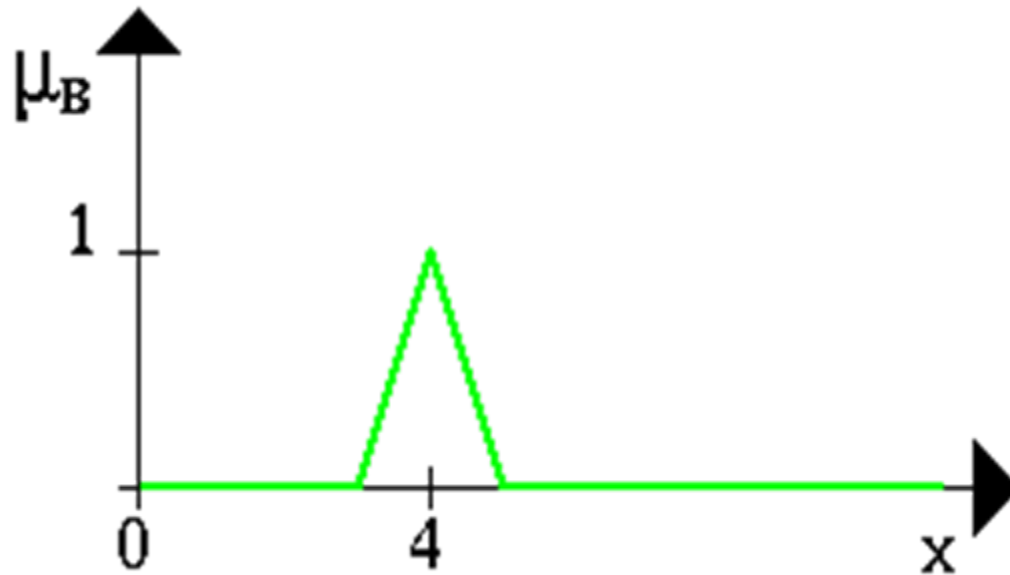


„etwa zwischen 5 und 8“



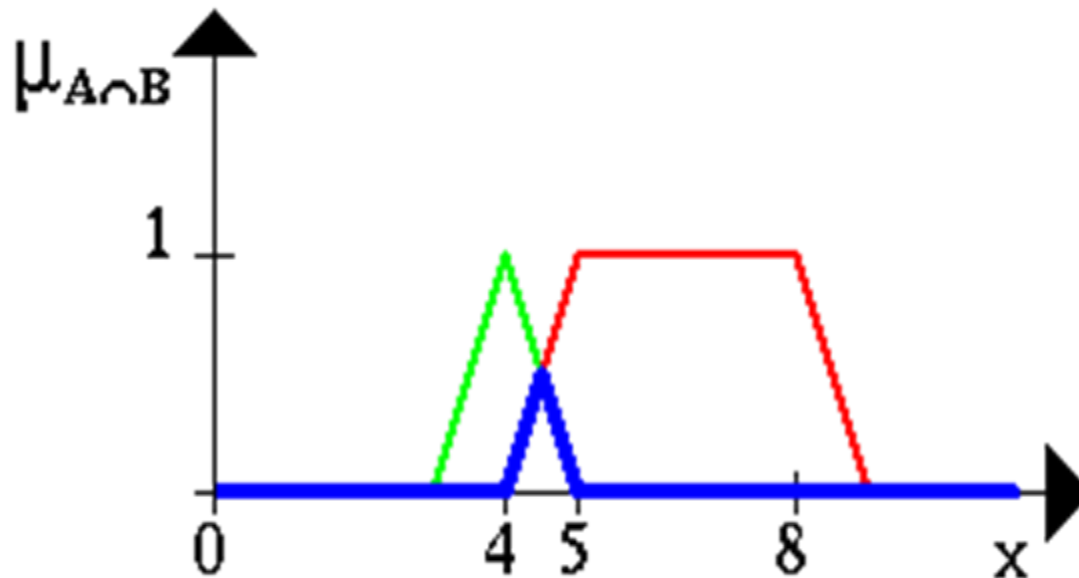


„ungefähr 4“



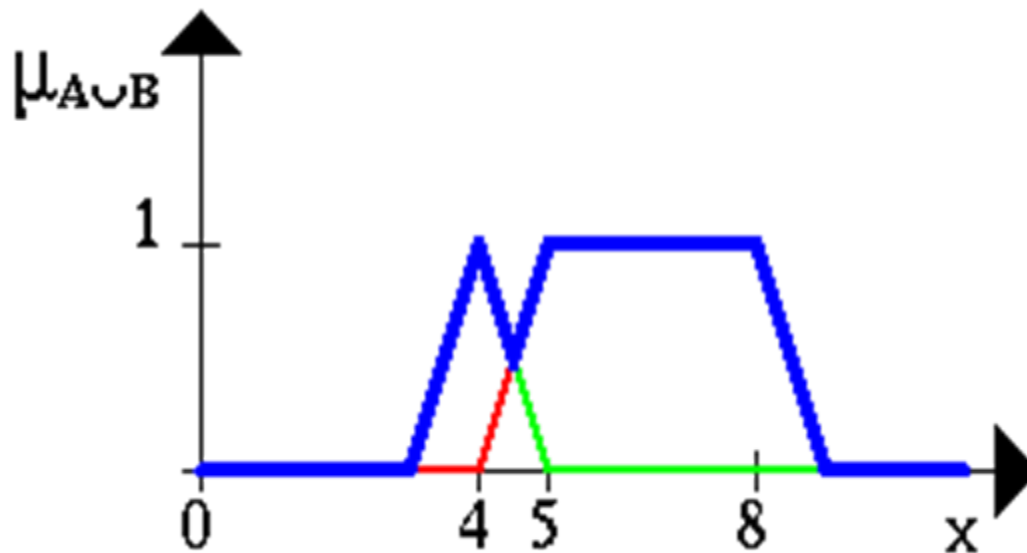


„(etwa zwischen 5 und 8) UND (ungefähr 4)“



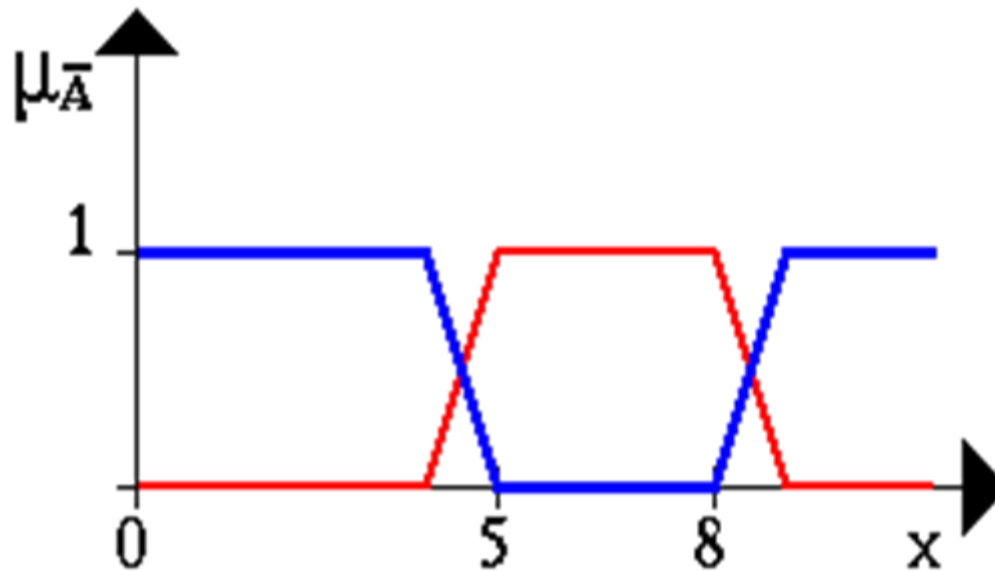


„(etwa zwischen 5 und 8) ODER (ungefähr 4)“





„NICHT (etwa zwischen 5 und 8)“



Und was kann man jetzt damit machen?



- Fuzzy Controller sind die wichtigste Anwendung der Fuzzy-Theorie.
- Anstelle von Differentialgleichungen benutzen sie Expertenwissen zur Prozess-Steuerung.
- Zur Formulierung dieses Expertenwissens benutzt man **linguistische Variablen**, beschrieben als Fuzzy-Mengen.

Beispiel: Stabbalance (inverses Pendel):

Ein Stab steht senkrecht auf der Ladefläche eines kleinen Wagens, der sich nur nach links und rechts bewegen darf. Der Wagen ist so zu bewegen, dass der Stab nicht umkippt.



Relevante Größen zur Beschreibung des Steuerverhaltens:

- Winkel zwischen Stab und Ladefläche
- Winkelgeschwindigkeit des Stabes beim Umkippen
- Geschwindigkeit des Wagens nach links bzw. rechts

Beispielregeln zur Kontrolle des Experiments:

- „Wenn der Stab aufrecht steht (Winkel gleich 0) und sich gar nicht bewegt (Winkelgeschwindigkeit gleich 0), dann muss ich offensichtlich nichts tun.“
- „Wenn der Stab in aufrechter Position ist, sich aber mit geringer Geschwindigkeit in eine bestimmte Richtung bewegt, dann sollte ich den Wagen ebenfalls mit geringer Geschwindigkeit in dieselbe Richtung bewegen.“



Zuerst einmal definieren wir für die linguistischen Werte der drei betroffenen Variablen die Wertebereiche:

negative high, negative low, zero, positive low, positive high

Die obigen Regeln werden dann zu einfachen Produktionen

- If angle is zero and angular velocity is zero then speed shall be zero.
- If angle is zero and angular velocity is pos. low then speed shall be positive low.
- ...

Zur Vereinfachung nehmen wir an, dass zu Beginn der Stab in nahezu senkrechter Position ist, so dass ein größerer Auslenkwinkel als etwa 45 Grad per Definition niemals auftauchen kann.



Winkel zwischen Stab und Wagen

	<i>NH</i>	<i>NL</i>	<i>Z</i>	<i>PL</i>	<i>PH</i>
<i>NH</i>			<i>NH</i>		
<i>NL</i>			<i>NL</i>	<i>Z</i>	
<i>Z</i>	<i>NH</i>	<i>NL</i>	<i>Z</i>	<i>PL</i>	<i>PH</i>
<i>PL</i>		<i>Z</i>	<i>PL</i>		
<i>PH</i>			<i>PH</i>		

Winkel-
geschwin-
digkeit des
Stabes

Abkürzungen:

NH: negative high

NL: negative low

Z: zero

PL: positive low

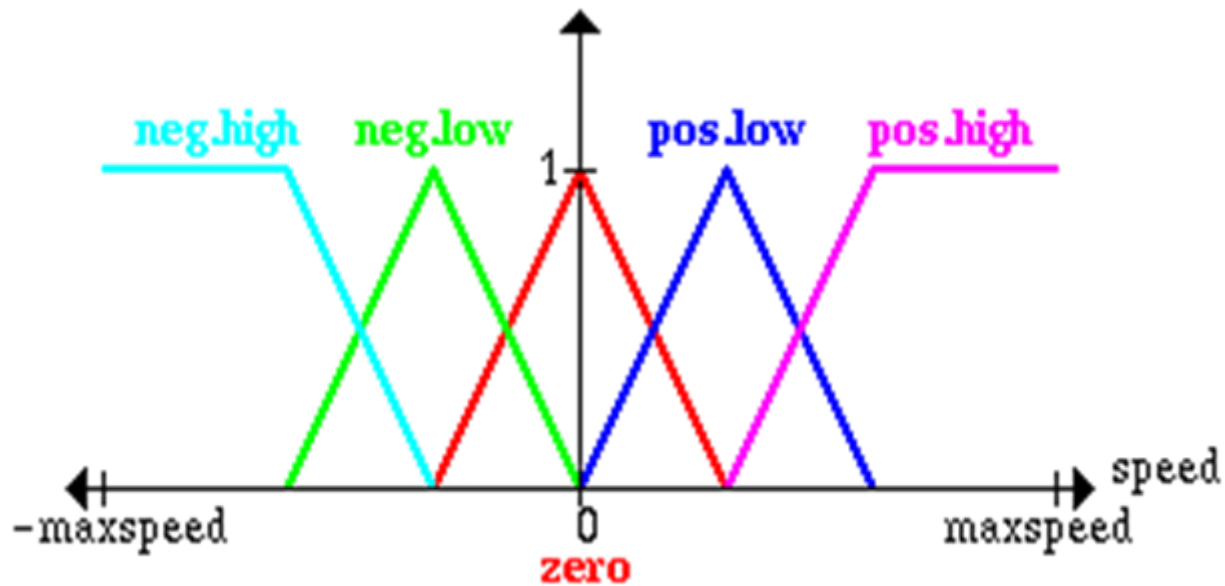
PH: positive high

Reaktion:
Wagen-
geschwindigkeit

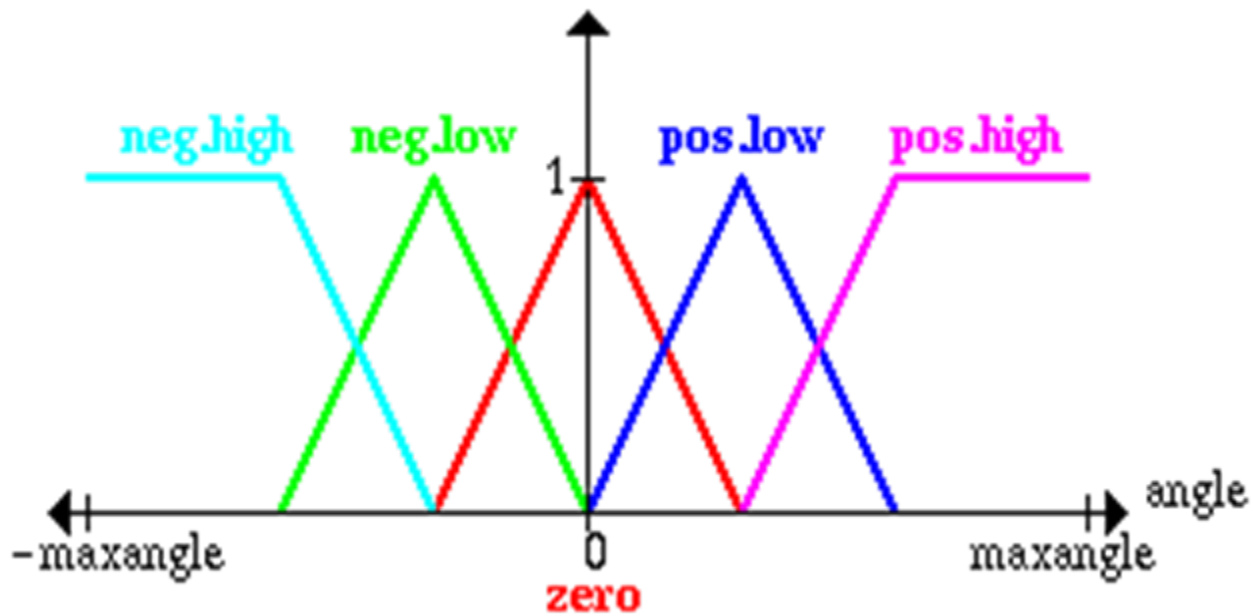
Fuzzy-Mengen für die Geschwindigkeit des Wagens



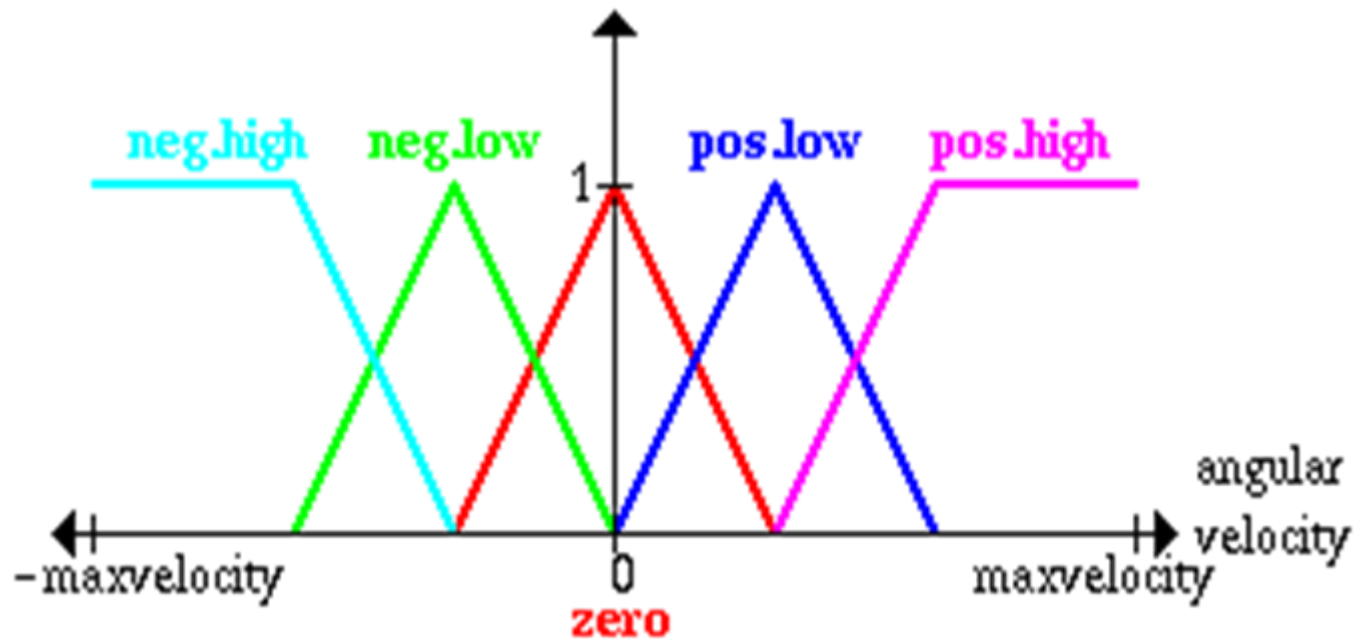
negative high (cyan)
negative low (green)
zero (red)
positive low (blue)
positive high (magenta)



Fuzzy-Mengen für den Winkel zwischen Ladefläche und Stab

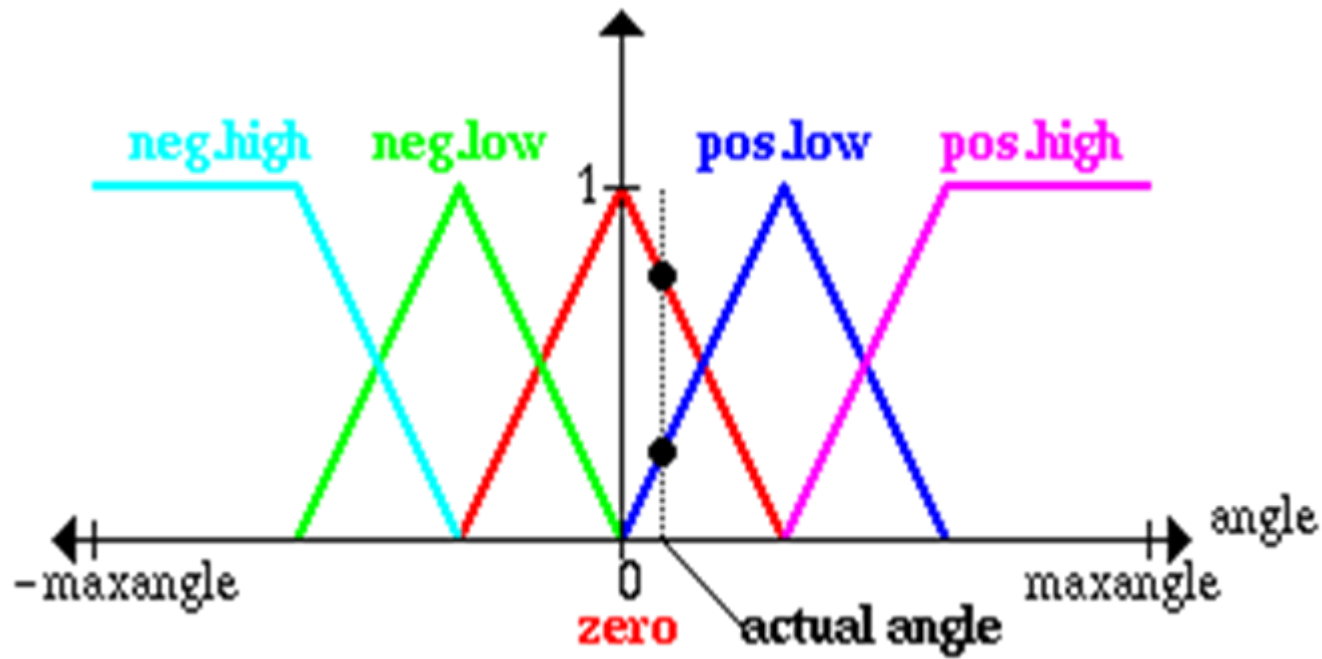


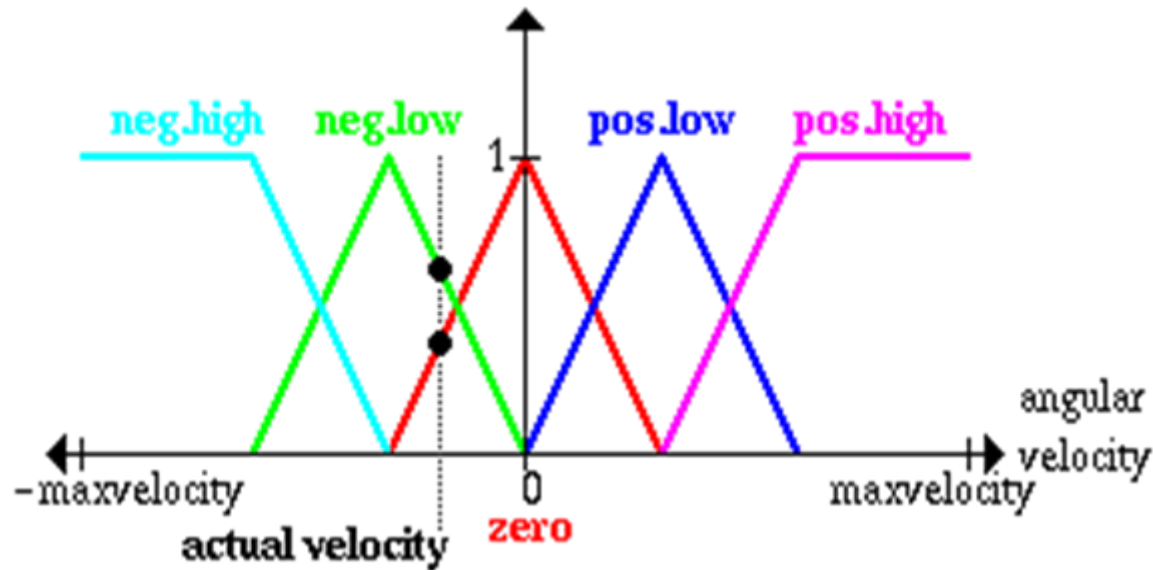
Fuzzy-Mengen für die Winkelgeschwindigkeit des Stabes



Wie interpretiere ich nun einen konkreten aktuellen Wert für den Winkel/die Winkelgeschwindigkeit in der Fuzzy-Welt?

Regelanwendung für einen konkreten Winkelwert

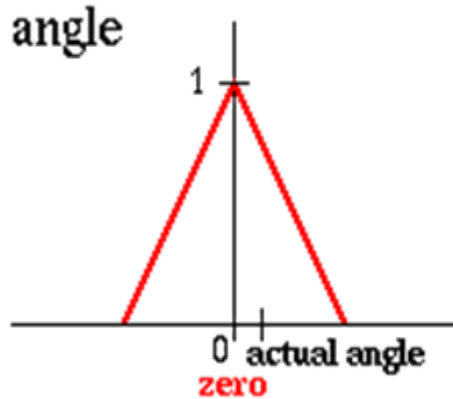




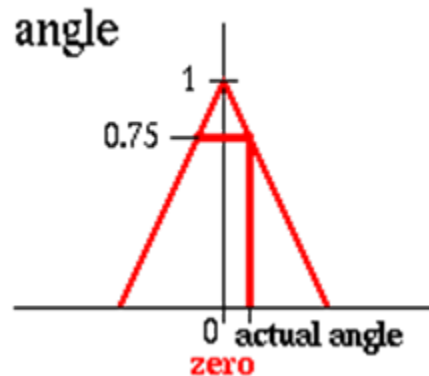
Wie werden jetzt komplexe Regelprämissen (mit mehreren konjunktiv verknüpften Vorbedingungen) interpretiert?



If angle is zero and angular velocity is zero then speed is zero.



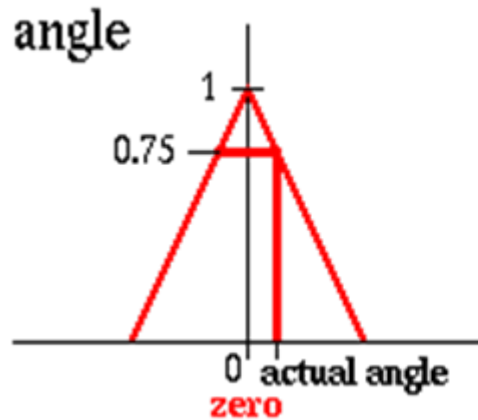
This is the linguistic variable "angle" where we zoom in on the fuzzy set "zero" and the actual angle.



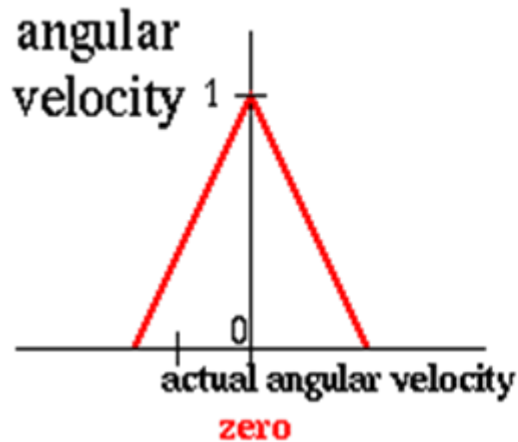
We realize that our actual value belongs to the fuzzy set "zero" to a degree of 0.75.



If angle is zero and angular velocity is zero then speed is zero



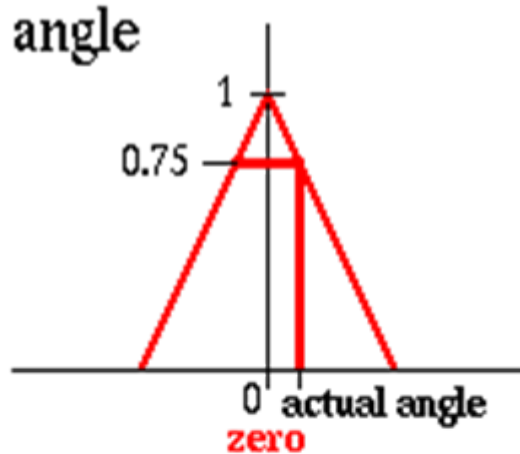
We realize that our actual value belongs to the fuzzy set "zero" to a degree of 0.75.



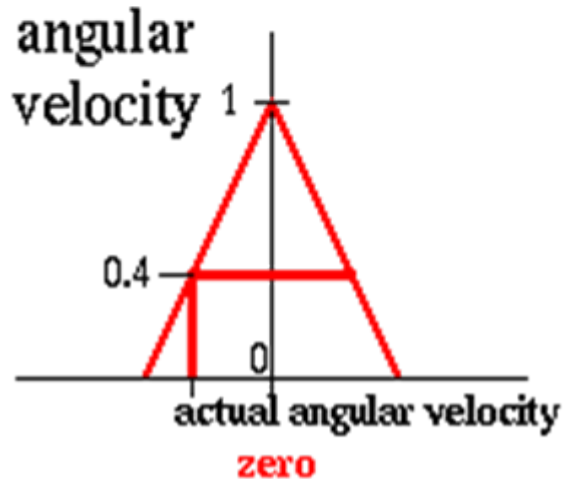
This is the linguistic variable "angular velocity" where we zoom in on the fuzzy set "zero" and the actual angular velocity.



If angle is zero and angular velocity is zero then speed is zero



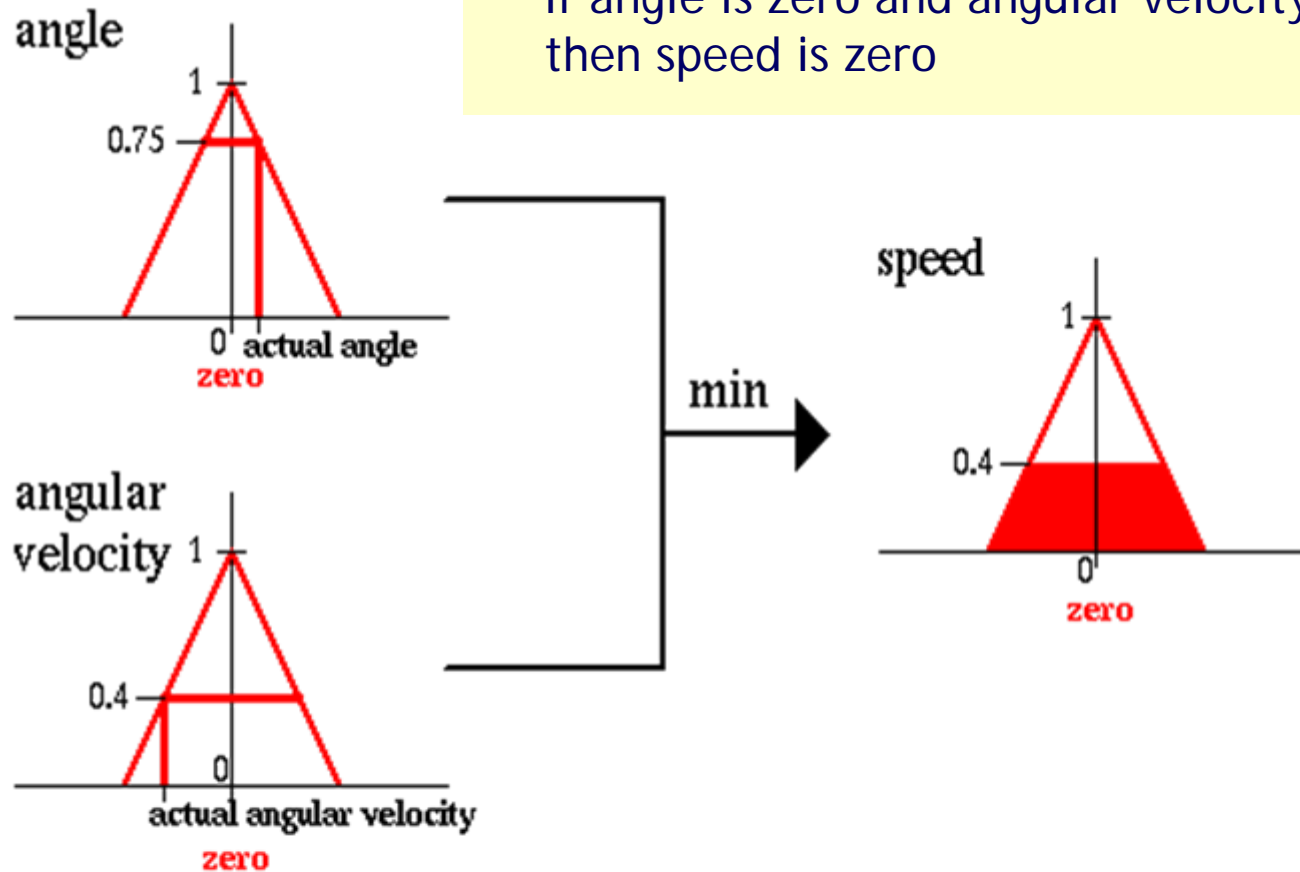
We realize that our actual value belongs to the fuzzy set "zero" to a degree of 0.75.



We realize that our actual value belongs to the fuzzy set "zero" to a degree of 0.4.



If angle is zero and angular velocity is zero then speed is zero

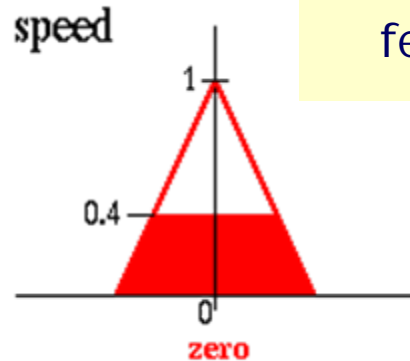


Since the two parts of the condition of our rule are connected by an AND we calculate $\min(0.75, 0.4) = 0.4$ and cut the fuzzy set "zero" of the variable "speed" at this level (according to our rule).

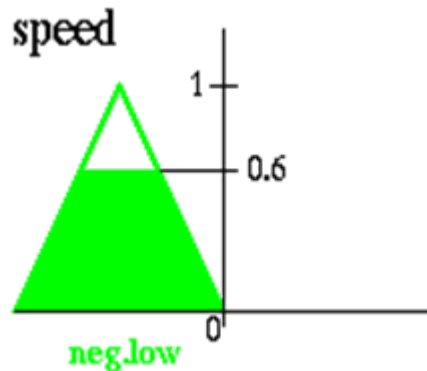
Kombination der Ergebnisbeiträge aller (vier) feuernenden Regeln:



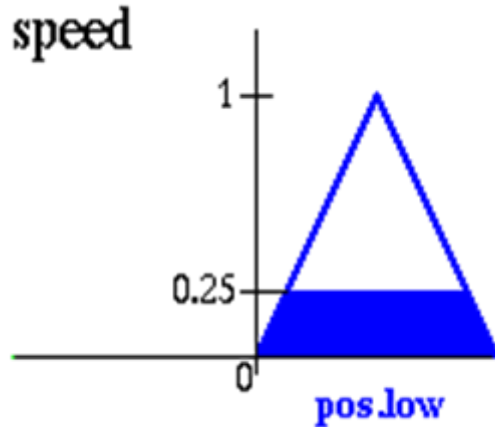
Als nächstes müssen die Ergebnisbeiträge aller vier feuernenden Regeln zusammengefasst werden:



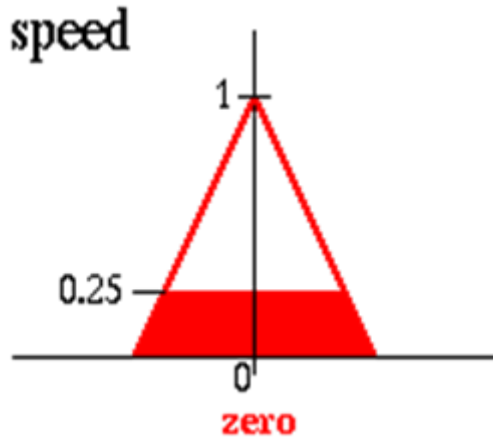
If angle is zero and angular velocity is zero then speed is zero.



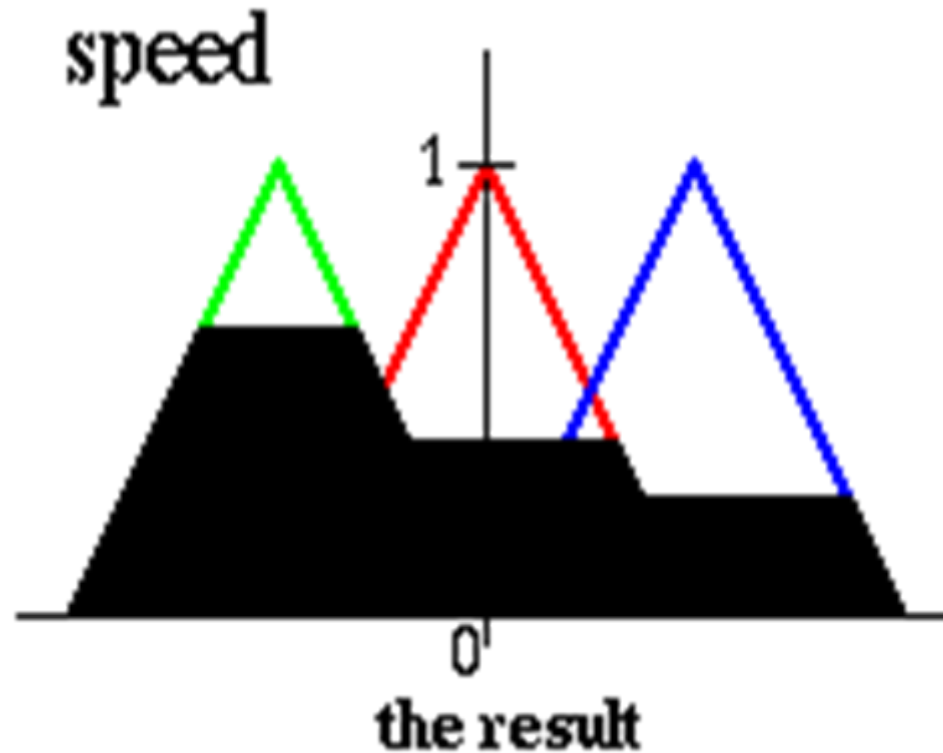
If angle is zero and angular velocity is negative low then speed is negative low.



If angle is positive low and angular velocity is zero then speed is positive low.



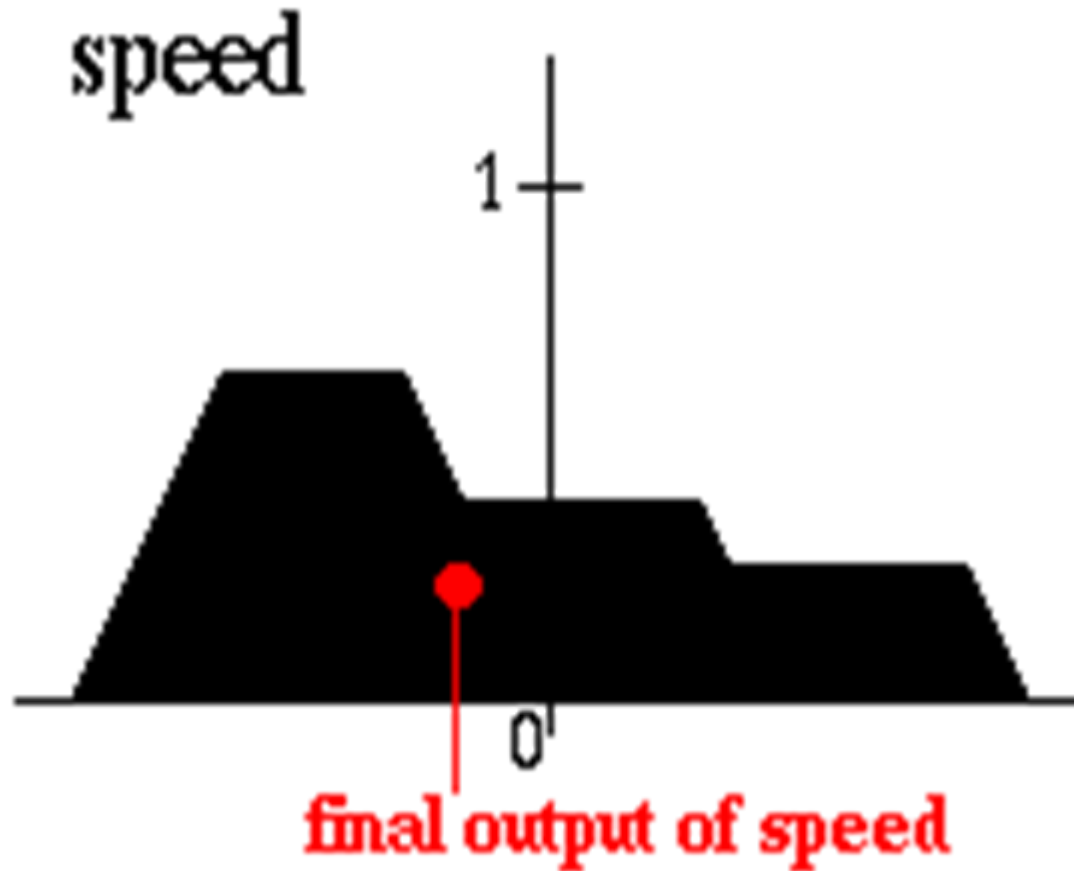
If angle is positive low and angular velocity is negative low then speed is zero.

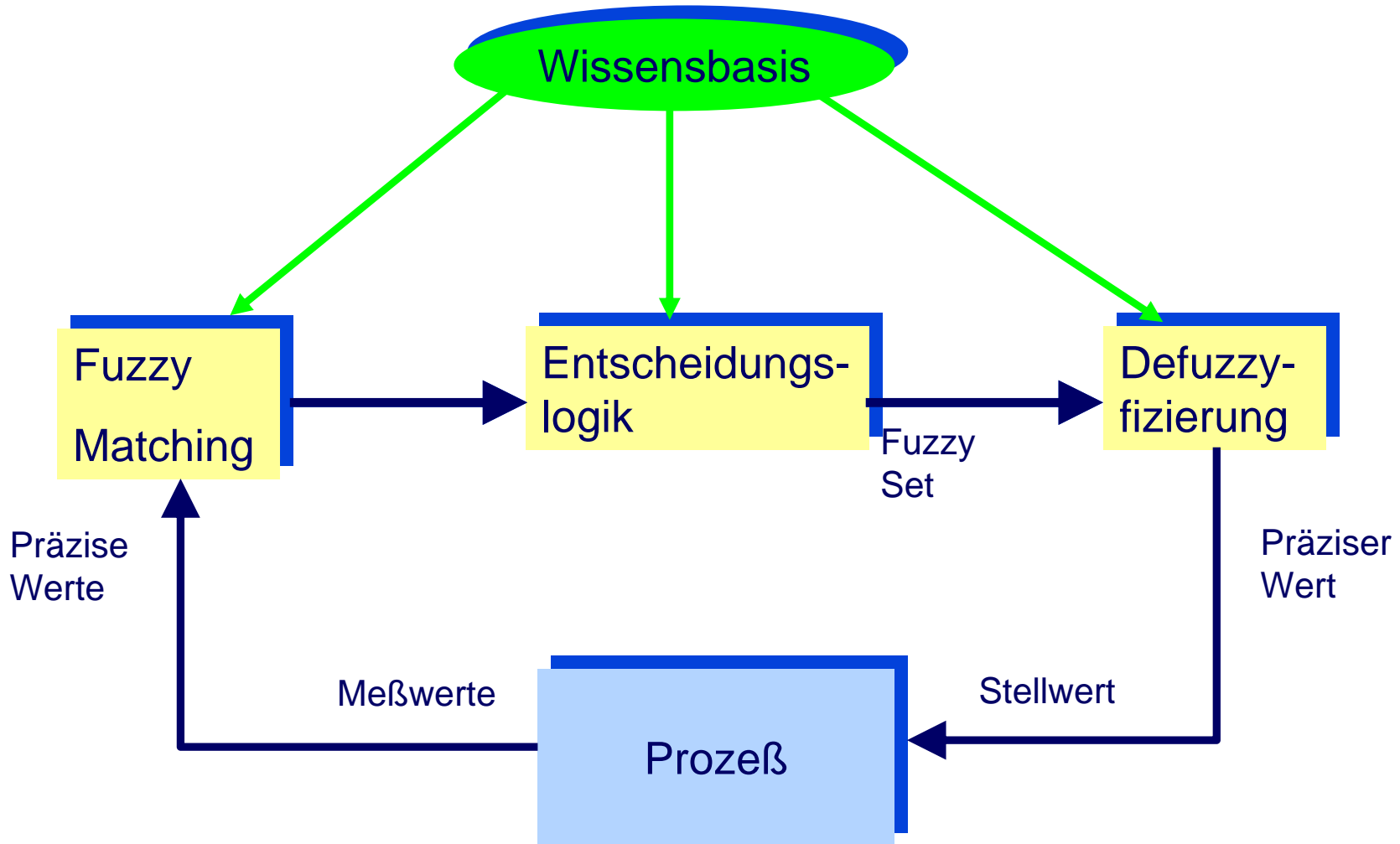


Wie kommt man nun von dieser Fuzzy-Menge zur Ergebnisbeschreibung zu einer konkreten quantitativen Handlungsanweisung?



Eine Standardmethode betrachtet den Schwerpunkt der Form der Fuzzy-Membership-Funktion:







Fuzzy Control ist empfehlenswert ...

- für sehr komplexe Prozesse, für die kein mathematisches Modell existiert.
- für hochgradig nichtlineare Prozesse.
- wenn man sprachlich formuliertes Expertenwissen verarbeiten muss.

Die Verwendung von Fuzzy Control mag eine schlechte Idee sein, wenn ...

- die konventionelle Kontrolltheorie bereits zufriedenstellende Ergebnisse liefert.
- ein leicht lösbares mathematisches Modell bereits existiert.
- das Problem nicht lösbar ist.

Einige Anwendungen der Fuzzy Logic



- Automatic control of dam gates for hydroelectric-powerplants (Tokio Electric Pow.)
- Preventing unwanted temperature fluctuations in air-conditioning systems (Mitsubishi, Sharp)
- Efficient and stable control of car engines (Nissan)
- Positioning of wafer-steppers in the production of semiconductors (Canon)
- Optimized planning of bus timetables (Toshiba, Nippon-System, Keihan-Express)
- Archiving system for documents (Mitsubishi Elec.)
- Prediction system for early recognition of earthquake's (Inst. of Seismology Bureau of Metrology, Japan)
- Medicine technology: cancer diagnosis (Kawasaki Medical School)
- Recognition of handwritten symbols with pocket computers (Sony)
- Automatic motor-control for vacuum cleaners with recognition of surface condition and degree of soiling (Matsushita)
- Back light control for camcorders (Sanyo)
- Compensation against vibrations in camcorders (Matsushita)
- Single button control for washing-machines (Matsushita, Hitachi)
- Recognition of handwriting, objects, voice (CSK, Hitachi, Hosai Univ., Ricoh)
- Flight aid for helicopters (Sugeno)
- Controlling of machinery speed and temperature for steel-works (Kawasaki Steel, New-Nippon Steel, NKK)
- Controlling of subway systems in order to improve driving comfort, precision of halting and power economy (Hitachi)
- Improved fuel-consumption for automobiles (NOK, Nippon Denki Tools)
- Improved sensitiveness and efficiency for elevator control (Fujitec, Hitachi, Toshiba)
- Improved safety for nuklear reactors (Hitachi, Bernard, Nuclear Fuel div.)