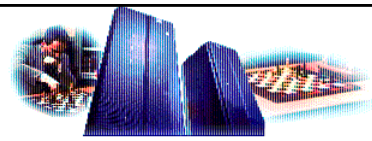
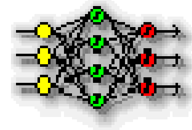
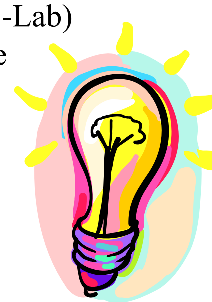




Sistemi Intelligenti
I fuzzy system

Alberto Borghese

Università degli Studi di Milano
Laboratory of Applied Intelligent Systems (AIS-Lab)
Dipartimento di Scienze dell'Informazione
borghese@dsi.unimi.it



Overview

I fuzzy set

I fuzzy system

θ' \ θ	NL	NM	NS	ZE	PS	PM	PL
NL				PL			
NM				PM			
NS				PS	NS		
ZE	PL	PM	PS	ZE	NS	NM	NL
PS			PS	NS			
PM				NM			
PL				NL			

B. Kosko, Neural Networks and Fuzzy Systems, Prentice Hall.

Per la logica basata su fuzzy set: corso di Logica Fuzzy della laurea specialistica in informatica tenuto dai Proff. Aguzzoli e Marra.

La logica classica

Logica a 2 valori: A = {T,F}.

La funzione verita' T: {proposizione} -> {0, 1} definisce la logica classica e può essere implementata come tabella della verità: descrizione esaustiva del funzionamento della funzione per **tutti** i possibili (discreti) valori in ingresso.

T(A,B,C,D) = {T,F} nel caso di una funzione a più valori.

Inoltre valgono le proprietà:

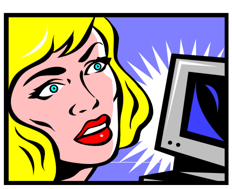
$A \cap A^c = \emptyset$.

$A = T$ (True) $\Leftrightarrow$ $A^c = F$ (False).

I Fuzzy set

Nella logica fuzzy, tutto è questione di gradazione.

Dal punto di vista matematico, fuzzy e' equivalente ad una proposizione logica che può assumere un numero infinito di valori tra vero e falso.



La funzione verita' generalizzata T: {proposizione} -> [0, 1] definisce la logica fuzzy o $\mathbf{L_1}$.

La funzione verità fuzzy

La funzione verità generalizzata T: {proposizione} -> [0, 1] definisce la logica fuzzy o $\mathbf{L_1}$.

$A \cap A^c \neq \emptyset$ Viene violata la legge di non contraddizione.

$A \cup A^c \neq X$ Viene violata la legge del terzo escluso.

Esempio di Fuzzy set

Costruiamo una montagnetta di granelli di sabbia S(.) è l'affermazione "è una montagnetta di sabbia" ed è funzione del numero di granelli, d.

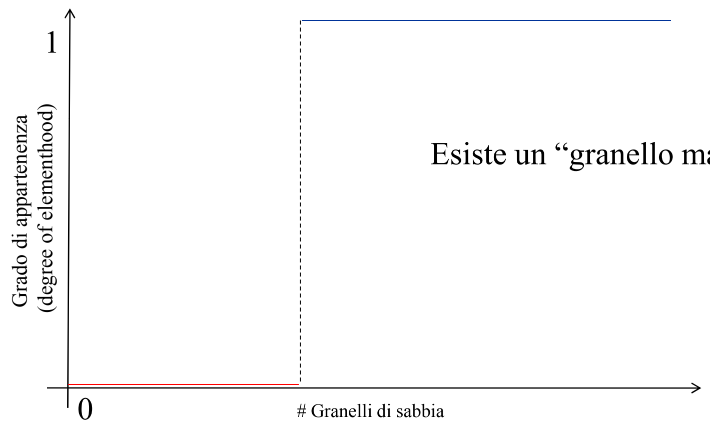
Non esiste un particolare numero di granelli, n*, per cui $S_{n*} = T$ diventi $S_{n*-1} = F$.

Supponendo di avere una montagnetta di n granelli di sabbia, possiamo scrivere: $T(S(_{n-m}) = 1 - d_{n-m}$. $T(S_n) = 1$ $T(0) = 0$. d esprime un margine di dubbio. (cf. l'esperimento di Moravec: the prosthetic brain).

Vale la relazione $0 = d_0 \leq d_1 \leq d_2 \leq d_3 \leq \ldots.. \leq d_{n-1} \leq d_n = 1$.

Nulla e' detto ancora sulla forma della funzione che passa da 0 ad 1.

Esempio di funzione di appartenenza (logica classica)



Esiste un “granello magico”

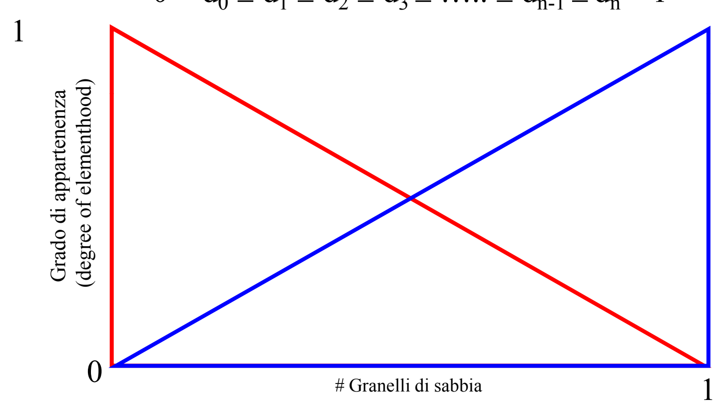
m(X) castello_false
m(X) castello_true

Sono 2 proposizioni antitetiche. In questo caso vale:

$$m(X)_{castello\_false} = 1 - m(X)_{castello\_true}$$

Esempio di funzione di appartenenza (logica fuzzy)

$$0 = d_0 \leq d_1 \leq d_2 \leq d_3 \leq \ldots\ldots \leq d_{n-1} \leq d_n = 1$$

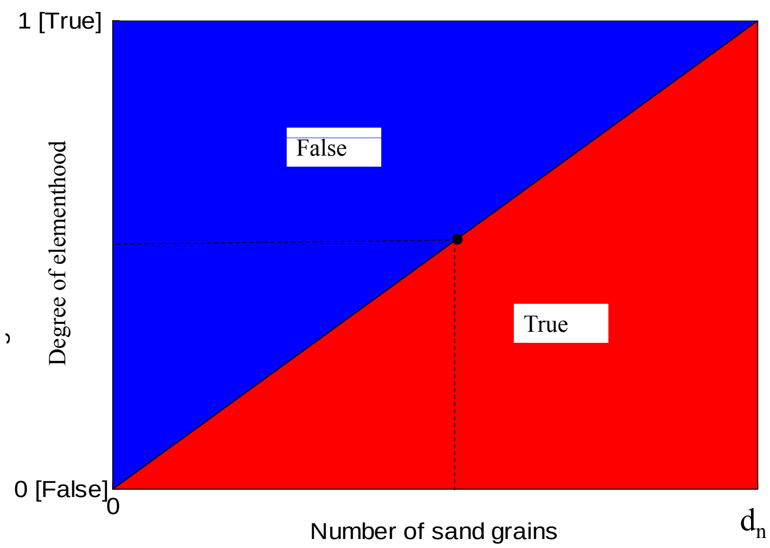


m(X) castello_false
m(X) castello_true

Sono 2 proposizioni antitetiche. In questo caso vale:

$$m(X)_{castello\_false} = 1 - m(X)_{castello\_true}$$

Esempio di funzione di appartenenza (membership function, $m_A(x)$)

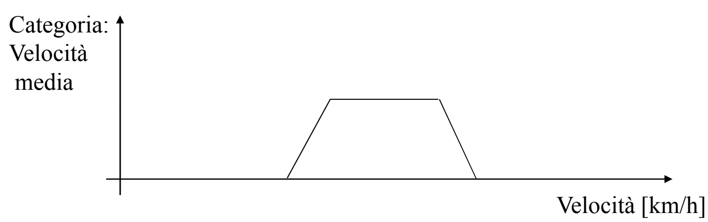


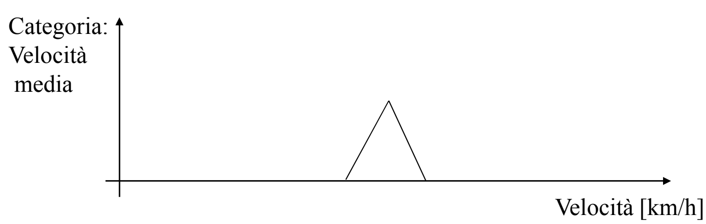
Insieme A:
montagnetta
di sabbia

$S(d_n) = T$
$S(0) = F$

$m_A(x)$ = Degree(x $\in$ A) è membership value or fit value

Altre forme di funzioni di appartenenza (membership function, $m_A(x)$)





Le funzioni di membership si possono sovrapporre?

Una variabile per più classi - I

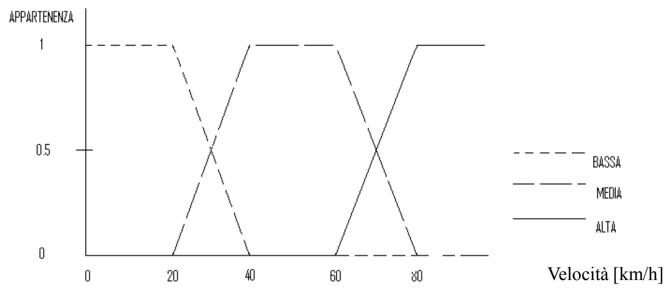
Velocità (km/h)	Grado di appartenenza a :(membership)		
	BASSA	MEDIA	ALTA
10	1	0	0
20	1	0	0
30	0.5	0.5	0
40	0	1	0
50	0	1	0
60	0	1	0
70	0	0.5	0.5
80	0	0	1

Differenti idee su cosa sia velocità bassa, media o alta.

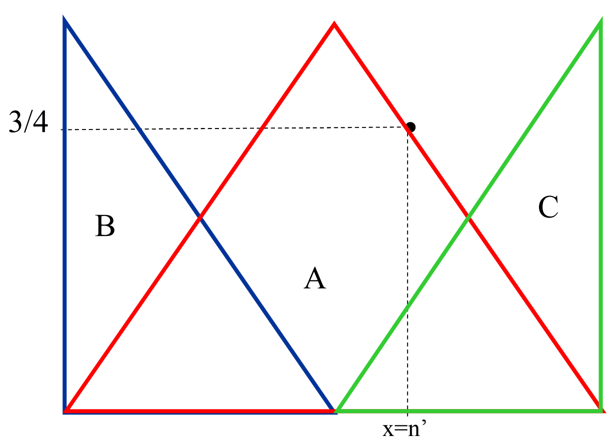
In generale le funzioni di membership hanno forma trapezoidale o triangolari.

Sovrapposizione tra le funzioni circa 25%

T(Alta) = F(Media)
T(Media) + F(Media) = 1



Una variabile, più classi - II



Sta piovendo in modo leggero?

Tre insiemi fuzzy
(e.g. pioggia):

A) Pioggia leggera.
B) Qualche goccia.
C) Pioggia intensa.

Esempio:
$m_A(x=n') = 3/4$
$m_C(x=n') = 1/4$

Fuzzyness versus probability

La **fuzzyness** descrive l'ambiguità insita in un evento.
La **probabilità** descrive l'incertezza che l'evento avvenga.

Che un evento accada o meno è una questione di probabilità, in quale misura accada è un questione di fuzzyness.

Esempio: "C'è il 20% di probabilità che domani ci sia pioggia leggera".

Ma anche: "Errori piccoli", "Clienti soddisfatti", "paesi in via di sviluppo", "segnali affetti da rumore",….

Una soluzione potrebbe essere utilizzare delle misture di probabilità e quindi test statistici sull'ipotesi nulla, vedremo più avanti quando tratteremo l'apprendimento statistico.

Fuzzyness versus Probabilità (esempi)

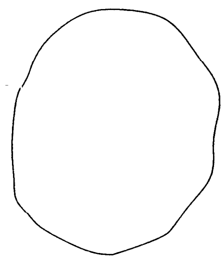
•Propongo un libro ad un editore (che può rispedirlo al mittente o chiederne la revisione, problema di royalities!).

•Apro il frigorifero in cui trovo una mezza mela (bicchiere mezzo pieno e mezzo vuoto).

•Parcheggio la mia automobile (non è previsto che venga parcheggiata a cavallo delle strisce: si rischia la contravvenzione).

Dopo che un evento è avvenuto, la sua probabilità svanisce.

Fuzzy versus probabilità (formalmente)



Ha più senso dire che questo è un ellisse fuzzy o che è proabilmente un ellisse?

Pur essendoci tutti gli elementi dell'evento, l'incertezza rimane. A cosa è dovuta?

Ha senso scrivere che: $m_A(x) = \text{Prob}\{x \in A\}$ true?

La fuzzyness è un'incertezza deterministica.

Gli operatori logici nella logica fuzzy

Introduzione di una norma, detta T-norm

T(A **AND** B) = ***min***(**T**(A), **T**(B)) $0 \leq T(A) \leq 1$

T(A **OR** B) = ***max***(**T**(A), **T**(B))

T(**NOT**-A) = 1 – **T**(A)

Zadeh, 1969.

Si noti che gli operatori min e max valgono anche per la logica classica.

Segue che: **T**(A **AND** A^c) + **T**(A **OR** A^c) = 1 (dimostrare!).

Rapporto con gli operatori della logica classica? Cosa succede se B = !A?

Valgono se vale il principio di mutua esclusione e del terzo escluso.

Come combinare operatori logici fuzzy

Eventi: temporale forte (A), temporale leggero (B).
Classi fuzzy: pioggia pesante, molti fulmini, tuoni, vento forte.

Logica fuzzy

$A = (1.\ .8\ 0.\ .5)$
$B = (.4\ .4\ .7\ .7)$

OR $A \cup B = (1.\ .8\ .7\ .7)$
AND $A \cap B = (.4\ .4\ 0.\ .5)$

$A^c = (0\ .2\ 1.\ .5)$
$A \cap A^c = (0\ .2\ 0.\ .5)$
$A \cup A^c = (1\ .8\ 1.\ .5)$

$A \cap A^c + A \cup A^c = 1$

$\mathbf{A \cap A^c \neq \varnothing}$
$\mathbf{A \cup A^c \neq X}$

Logica classica

$A = (1\ 0\ 0\ 1)$
$B = (0\ 0\ 1\ 1)$

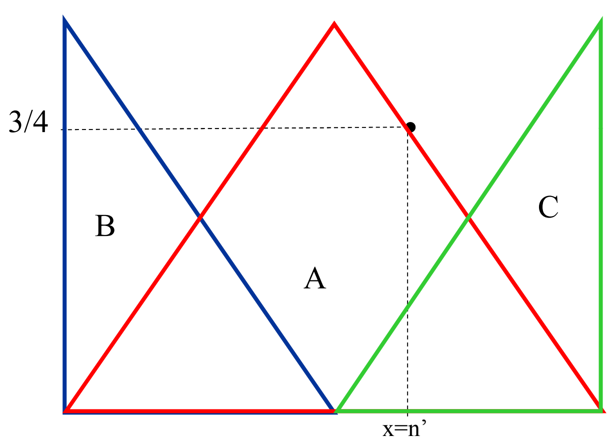
$A \cup B = (1\ 0\ 1\ 1)$
$A \cap B = (0\ 0\ 0\ 1)$

$A^c = (0\ 1\ 1\ 0)$
$A \cap A^c = (0\ 0\ 0\ 0)$
$A \cup A^c = (1\ 1\ 1\ 1)$

$A \cap A^c + A \cup A^c = 1$

$A \cup B$ Caratteristica presente al minimo grado in entrambi gli insiemi.
$A \cap B$ Caratteristica presente al massimo grado in entrambi gli insiemi.

Una variabile, più classi



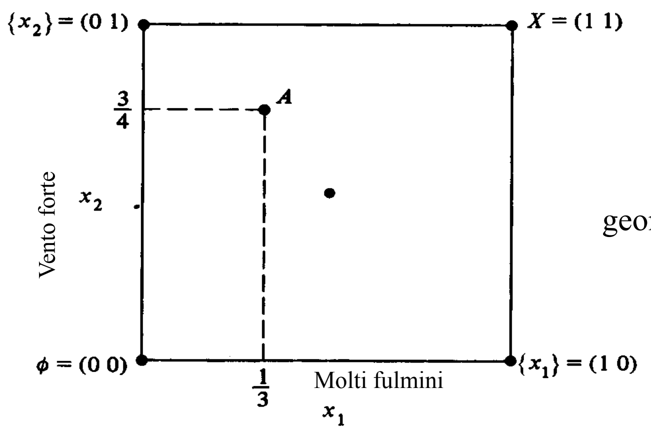
Sta piovendo in modo leggero?

Tre insiemi fuzzy
(e.g. pioggia):

A) Pioggia leggera.
B) Qualche goccia.
C) Pioggia intensa.

Esempio:
$m_A(x=n') = 3/4$
$m_C(x=n') = 1/4$

Più varibili, una classe



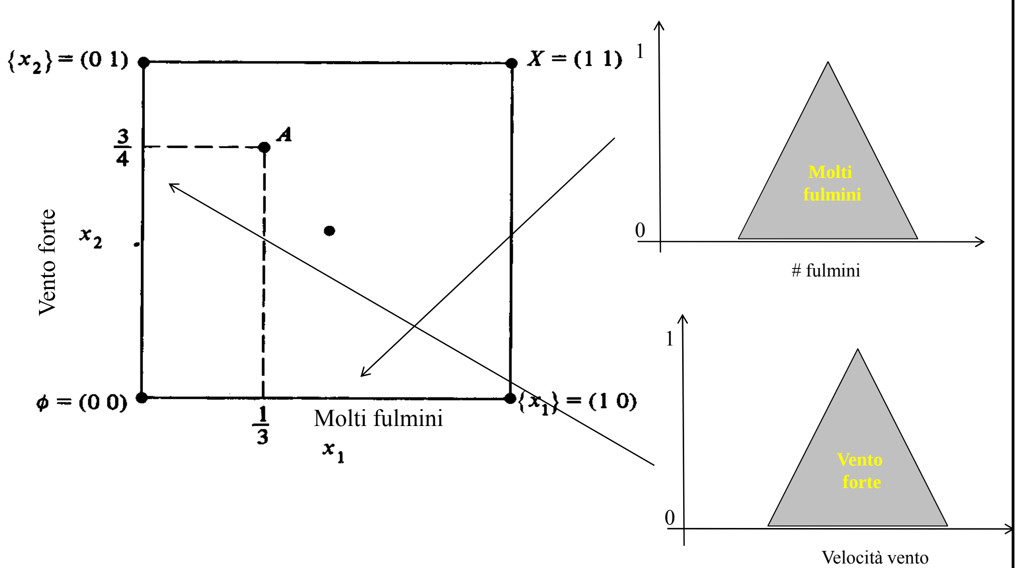
Su ogni asse il grado di appartenenza ad una classe.

Rappresentazione geometrica di 2 classi fuzzy.

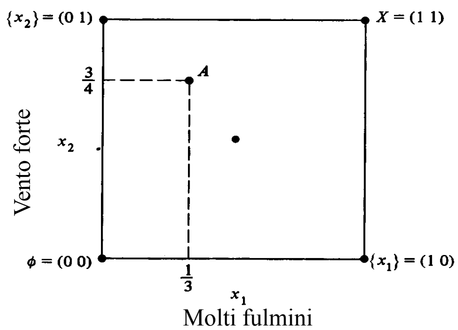
Valore di **fit** di A = (1/3 3/4). E' un vettore a n dimensioni.

Un certo insieme fuzzy (di fulmini e di vento forte) è un punto in un ipercubo di dimensioni $I_2 = [0,1]x[0,1] = [0,1]^2$, *in generale,* $I_n = [0,1]^n$.

Rappresentazione e classi



Proprietà dello spazio fuzzy (I)

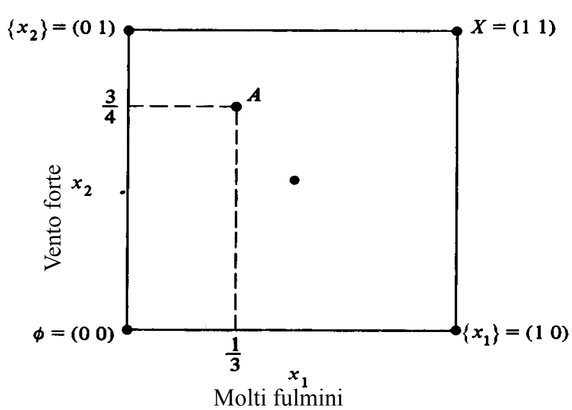


I vertici dell'ipercubo cosa rappresentano?

NF = {Ø, X, $\{x_1\}$, $\{x_2\}$}.

Questi 4 elementi rappresentano i 4 insiemi non fuzzy dello spazio fuzzy definito dalle 2 classi: vento forte e molti fulmini.

Proprietà dello spazio fuzzy (II)



Dove troviamo la massima fuzzyness?

Al centro: A U A_c = A $\cap A_c$

Paradossi associati al centro

Il bugiardo di Creta che diceva che tutti a Creta erano bugiardi.

Il barbiere di Bertrand Russell. Fuori dal suo negozio era appeso un cartello che diceva: “il barbiere taglia la barba solamente a tutti quelli che non se la tagliano da soli”. Chi taglia la barba al barbiere?

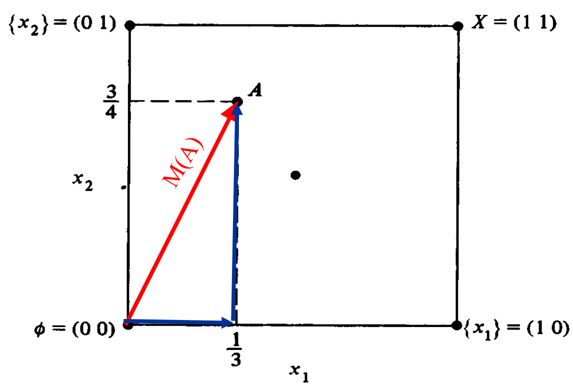
Prendiamo una carta che dice su un lato: “Quello che è riportato sull’altro lato è vero” e sull’altro: “Quello che è riportato sull’altro lato è falso”.

Bicchiere mezzo pieno e mezzo vuoto.

Al centro: A U A_c = A $\cap A_c$

Esempio A: barbiere, A_c : non_barbiere

Misure in un insieme fuzzy



Norma di un vettore:

$$M(A) = \sqrt[p]{\sum_{i=1}^{n} |m_A(x_i)|^p}$$

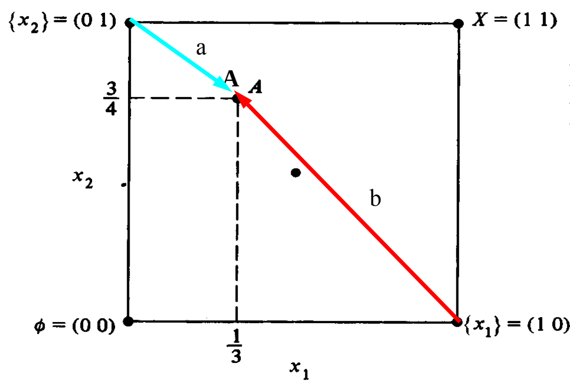
n: numero di classsi fuzzy

Per p = 2 misuriamo la lunghezza del vettore A – O.
Per p = 1 cosa misureremmo in logica classica?

Distanza fuzzy di Hamming ***-> l_1*** $0 \leq M(A) = 1/3+3/4 = 13/12 \leq n$
Distanza euclidea. ***-> l_2*** $0 \leq M(A) = \sqrt[2]{97/144} \leq \sqrt[2]{n}$

Entropia fuzzy

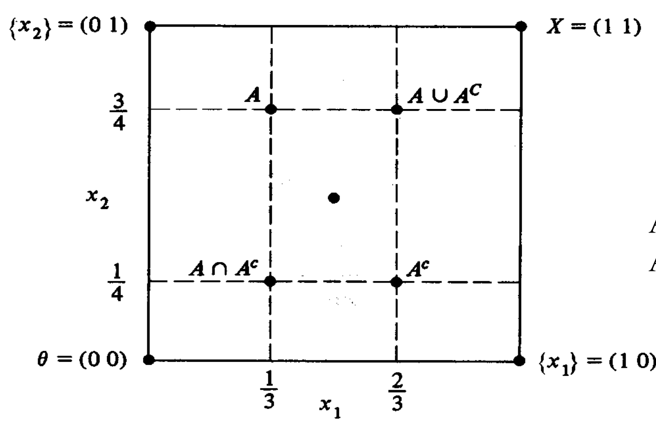
How (much) fuzzy is a fuzzy set?



E(A) = 0 nei vertici
E(A) = 1 nel centro

$$E(A) = \frac{a}{b} = \frac{l^1(A, A_{vicino})}{l^1(A, A_{lon\tan o})} \qquad 0 \leq E(A) = \frac{1/3 + 1/4}{2/3 + 3/4} = 7/17 \;\; 0 \leq 1$$

Ipercubo fuzzy interno.



A = (1/3 3/4)
A^c = (2/3 1/4)

$A \cup A^c$ = (2/3 3/4) Max
$A \cap A^c$ = (1/3 1/4) min

- Dipende da A.
- Può degenerare se una proprietà è massimamente fuzzy. Se sono fuzzy sia x_1 che x_2: $A \cup A^c = A \cap A^c = (1/2 \quad 1/2)$.

Riassunto

•Fuzzyness descrive l'ambiguità di un evento.

•La funzione di appartenenza descrive il grado di appartenenza di un evento ad una o più classi (non mutuamente esclusive, le funzioni di membership sono parzialmente sovrapposte).

•La fuzzyness si distingue dalla probabilità che misura la probabilità che un evento, nella sua interezza si verifichi o meno.

•Si possono definire attraverso la T-norm, le operazioni logiche tra insiemi fuzzy: AND, OR e NOT.

•Un oggetto fuzzy si può rappresentare come un punto all'interno di un ipercubo di dimensionalità pari al numero di elementi che caratterizzano l'oggetto.

•E' possibile associare ad un insieme fuzzy una misura di entropia che rappresenta il rapporto tra la misura della violazione della legge di non contraddizione e del terzo escluso.

Overview

I fuzzy set

I fuzzy system

θ' \ θ	NL	NM	NS	ZE	PS	PM	PL
NL				PL			
NM				PM			
NS				PS	NS		
ZE	PL	PM	PS	ZE	NS	NM	NL
PS			PS	NS			
PM				NM			
PL				NL			

Breve storia

Max Black (1937). Applico' i principi della logica fuzzy a liste di elementi o simboli. Sviluppo' la prima funzione di appartenenza fuzzy e chiamo' l'incertezza dell'appartenenza **vaghezza** (vagueness).

Lofti Zadeh (1965). "Fuzzy sets" sviluppo' una teoria sugli insiemi fuzzy ed una logica (fuzzy) in grado di lavorare su questi insiemi.

Ebrahim Mamdani (1974) dimostrò come utilizzare la logica fuzzy per risolvere problemi applicativi (inferenza fuzzy).

Dal 1985 ad oggi applicazioni tecnologiche con successo crescente in campi più diversi, dal controllo, alla ricerca nei data base, all'elaborazione delle immagini alla classificazione. Vediamo di capirne i fondamenti del successo.

$S : I^n \rightarrow I^p$.

Sistemi esperti

•E' basato su regole in parallelo (legate da AND e OR) che rappresentano la conoscenza.
•Funziona mediante un motore inferenziale che lavora sulle regole. Il motore è del tipo: IF ..... THEN ..... ELSE (*reasoning engine*). Si parla anche di intelligenza artificiale.

•La risposta T / F (modalità del tipo winner-take-all). Eventualmente più di una risposta Crisp.

•Sistemi di analisi di guasti, sistemi di diagnosi automatica (Computer-Aided Diagnosis), ragionamento automatico….

$S : I^n \to I^p$

Funzionamento di un sistema fuzzy

• L'uscita di ogni regola (attivata) è considerata con un certo grado di membership o verosimiglianza.

•Vengono utilizzate più regole in parallelo (legate da AND e OR) che rappresentano la conoscenza.

•L'uscita e' ottenuta pesando il grado di attivazione delle regole, ovverosia quanto ciascuna regola pesa nella formazione dell'output.

Dal punto di vista geometrico, mappa un ipercubo n-dimensionale in un inpercubo p-dimensionale:

➔ **FAM** (*Fuzzy Associative Memories*).

$$S : I^n \to I^p$$

Fuzzy Associative Memory (FAM)

Una FAM **trasforma** una spazio di input in uno spazio di output.

1) Spazi continui.
2) Spazi fuzzy.

Matematicamente una FAM opera la seguente trasformazione:
$I^n \to I^p$ dove n è il numero di classi della variabile fuzzy di ingresso e p è il numero di classi della variabile fuzzy di uscita.

Dato un insieme fuzzy A $\in I^n$, viene definito un insieme B $\in I^p$, che gli corrisponde (notazione alternativa (A,B))

E.g. Traffico:	A $\in I^n$ = [Assente, leggero, medio, pesante]	n = 4
Durata semaforo verde:	B $\in I^p$ = [Breve, media, lunga]	p = 3

Esempio di regola: (Traffico_leggero, Durata_semaforo_media)

Un input può attivare più classi di uscita.

Descrizione di una FAM

Una FAM **trasforma** una spazio di input in uno spazio di output.

1) **Spazi continui.**
2) Spazi fuzzy.

Dato un insieme fuzzy A $\in I^n$, viene definito un insieme B $\in I^p$, che gli corrisponde (notazione alternativa (A,B)).

La trasformazione **continua** da A a B, può essere descritta mediante l'insieme **finito** delle trasformazioni (***regole***) tra gli insiemi che costituiscono A e B $\{(A_i,B_j)\}$; $1 \le i \le n$; $1 \le j \le p$.

E.g. Traffico: $A \in I^n$ = {Assente, leggero, medio, pesante}.
Durata semaforo verde: $B \in I^p$ = {Breve, media, lunga}

Esempio di regola:
(Traffico_leggero, Durata_semaforo_breve) -- (A_2,B_3) + grado di fit della regola

Come si definisce la trasformazione codificata in una FAM?

La FAM traduce delle definizioni linguistiche.

Le associazioni sono tra A e B. Ovverosia ad ogni insieme (fuzzy) di A (input), viene associato un insieme (fuzzy) di B (output). E.g. Se il traffico è leggero, regola la durata del verde a media.

Fino qui siamo nel dominio della logica classica. Qual'è il problema?

•L'associazione è tra 2 classi (A_i,B_j). Ma l'evento (densità di traffico) può appartenere a due classi diverse, A_i, che sono associate a due o più classi tra le possibili classi, B_i, di uscita. Cosa facciamo?
•Le due classi rappresentano un intervallo di valori, come lo trattiamo?

Esempio di costruzione di una FAM

Controllore di un semaforo.

Ingresso: misura del traffico. 3 classi fuzzy: SCARSO, MEDIO, PESANTE.

Uscita: durata verde: 2 classi fuzzy: BREVE, LUNGO.

$I^3 \rightarrow I^2$

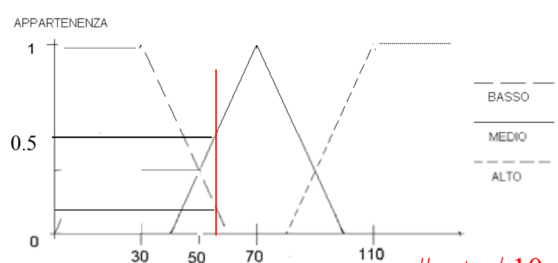
(regola 1) IF (SCARSO) THEN (BREVE)
(regola 2) IF (MEDIO) THEN (LUNGO)
(regola 3) IF (ALTO) THEN (LUNGO)

FAM

Fuzzyficazione: dalla misura alle classi fuzzy

Misuriamo la portata del traffico e vogliamo un controllore che generi una lunghezza in secondi.

Necessità di tradurre l'approccio linguistico in numeri.



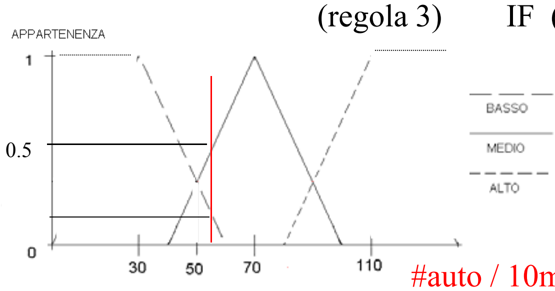
Fuzzyficazione viene risolta con le funzioni di membership: 55 auto/10min è un traffico *Scarso* con grado di fit 0.2 *Medio* con grado di fit 0.5.

#auto / 10m : densità di traffico

m(A) = (0.2, 0.5, 0)

Utilizzo della FAM

(regola 1) IF (BASSO) THEN (BREVE)
(regola 2) IF (MEDIO) THEN (LUNGO)
(regola 3) IF (ALTO) THEN (LUNGO)



m(A) = (0.2, 0.5, 0)

#auto / 10m : densità di traffico.

Solamente le prime 2 regole (associate a traffico *Scarso* e *Medio*) vengono attivate.

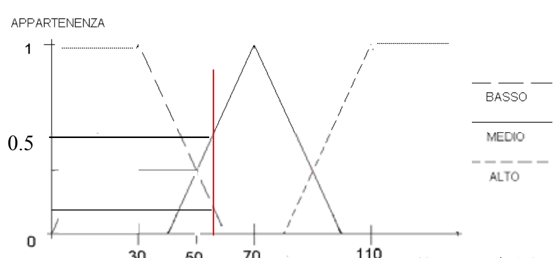
If (A_1) then B_1 Se è basso allora breve, grado 0.2
If (A_2) then B_2 Se è medio allora lungo, grado 0.5

Durata?

Fuzzyficazione: dalla misura alle classi fuzzy

Misuriamo la portata del traffico e vogliamo un controllore che generi una lunghezza in secondi.

Necessità di tradurre l'approccio linguistico in numeri.



Fuzzyficazione viene risolta con le funzioni di membership: 55 auto/10min è un traffico *Scarso* con grado di fit 0.2 *Medio* con grado di fit 0.5.

#auto / 10m : densità di traffico

m(A) = (0.2, 0.5, 0)

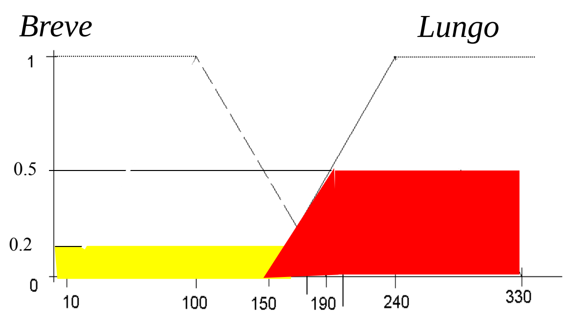
Dalle classi fuzzy alla generazione dell'output

L'output viene mappato in classi fuzzy analogamente all'input.

Breve = 80s

Lungo = 260s

Quanto deve durare il semaforo?



Defuzzyficazione mediante massimo

$$y = \max_{1 \le j \le k} m_B(y_j)$$

Viene scelta l'uscita proveniente da una proposizione linguistica

=

Viene utilizzata un'unica regola.

La scelta della regola dipende dal valore di fit della classe in ingresso.

Defuzzyficazione mediante media pesata

$$Uscita = y = \frac{\sum F_i * y_i}{\sum F_i}$$

F_i peso della regola i attivata, fit della regola
y_i azione associata alla regola (A_i,B_i)

L'uscita di ciascuna regola viene pesata con il grado di fit della classe in ingresso alla regola.

Tanto maggiore è il grado di fit, di verosimiglianza, della variabile in ingresso, tanto maggiore sarà il peso dell'azione intrapresa in funzione di quella variabile.

Defuzzyficazione mediante media pesata con le aree

La tecnica della media pesata non tiene conto della forma delle classi associate alle variabili di uscita: una classe molto ampia ha lo stesso peso di una classe molto stretta.

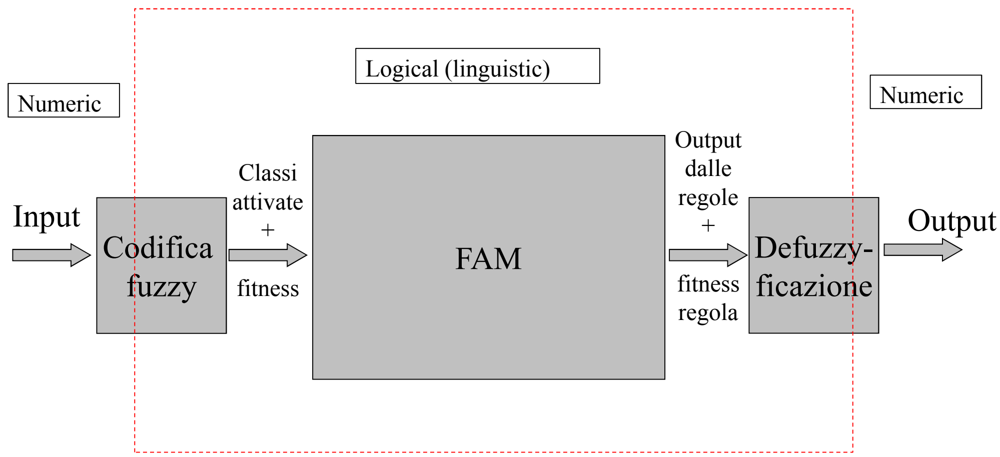
Si preferisce perciò utilizzare il criterio di fitness della variabile in ingresso per individuare un'area nella classe di uscita.

$$Y = \frac{\int y m_B(y) dy}{\int m_B(y) dy}$$

L'integrale dà il peso della regola.

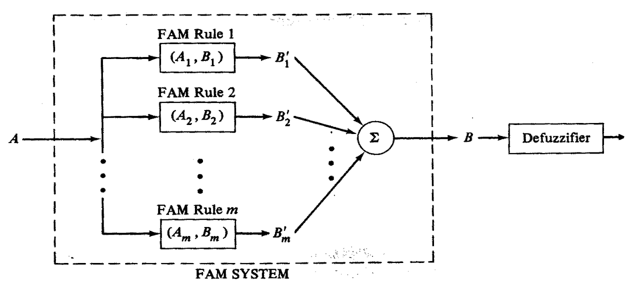
Gli integrali diventano una sommatoria per membership function semplici.

Struttura di un sistema fuzzy



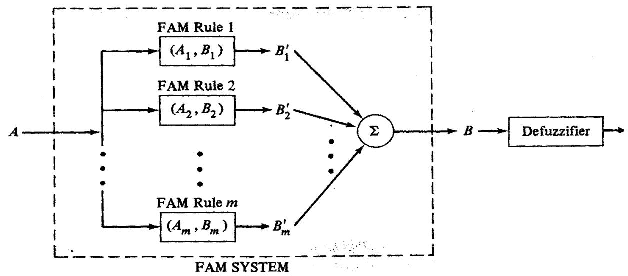
Tutte le regole della FAM ricevono un input, sono effettivamente attivate quelle che hanno un input con fitness > 0

Progettazione di un sistema fuzzy - I



1) Identificazione delle variabili del sistema e del loro range (A e B).
2) Identificazione delle classi fuzzy in cui le variabili sono da suddividere.
3) Definizione della trasformazione I/O come insieme di regole fuzzy: per ogni combinazione di classi fuzzy di input è possibile definire una classe di output (FAM).

Progettazione di un sistema fuzzy - II



4) Identificazione delle regole attivate da un certo insieme fuzzy in ingresso.
5) Valutazione del grado di fit della regola.
6) Unione degli insiemi fuzzy di output risultanti.
7) Calcolo del centroide degli insiemi risultanti (defuzzyficazione).

Overview

I fuzzy set

I fuzzy system

θ' \ θ	NL	NM	NS	ZE	PS	PM	PL
NL				PL			
NM				PM			
NS				PS	NS		
ZE	PL	PM	PS	ZE	NS	NM	NL
PS			PS	NS			
PM				NM			
PL				NL			