LA LOGICA FUZZY E IL PARADOSSO DEL MUCCHIO DI GRANO

Giangiacomo Gerla
Facoltà di Scienze, Università di Salerno,
Via S. Allende Baronissi (*SA*).
e-mail gerla@matna2.dma.unina.it

Introduzione.

Allo scopo di fornire modelli matematici dei concetti vaghi, cioè concetti come quelli di *"alto"*, *"piccolo"*, *"scuro"*, *"vicino"*, L. A. Zadeh in un famoso articolo del 1965 ha proposto la nozione di *sottoinsieme fuzzy*. Successivamente, a partire da un articolo di J. A. Goguen del 1968, da un fondamentale scritto di Zadeh del 1975 e dalla rigorosa formalizzazione proposta da J. Pavelka nel 1979, si sono sviluppate una serie di ricerche volte a formalizzare i ragionamenti che coinvolgono concetti vaghi. L'insieme di tali ricerche è compreso sotto il nome di *logica fuzzy* o di *ragionamento approssimato.* Tale logica si ricollega alla più antica tradizione di studi sulle logiche a più valori iniziata, fin dal 1930, all'interno della scuola polacca di logica. Tuttavia, mentre dal punto di vista semantico non esistono sostanziali differenze, lo scopo che si pone l'apparato deduttivo in logica fuzzy è completamente diverso. Infatti, nella tradizione degli studi sulle logiche a più valori l'apparato inferenziale deve mirare a produrre l'insieme (classico) di formule che per qualche motivo si ritengono valide. Invece nella logica fuzzy l'apparato inferenziale deve produrre, per ogni fissato fuzzy sottoinsieme di ipotesi, il fuzzy sottoinsieme di formule che da esso si possono dedurre. In altri termini, in logica fuzzy sia l'essere ipotesi che l'essere deducibile da date ipotesi sono concetti vaghi.

Lo scopo di questo articolo non è di dare una precisa e completa esposizione della logica fuzzy (che d'altra parte è ancora ben lontana dall'avere una sistemazione definitiva). Esso vuole solo evidenziare il rapporto tra le idee che ne sono alla base ed alcuni paradossi che sono riconducibili a quelli famosi "del mucchio di grano" e "dell'uomo calvo".

1. Due famosi paradossi dell'antichità.

Eubulo di Mileto, un filosofo greco della seconda metà del IV secolo, propose un famoso paradosso detto "del mucchio di grano". Il paradosso consiste nel provare il seguente teorema.

Teorema 1.1. Tutti i mucchi di grano sono piccoli.

Dim. Accettiamo, come ci sembra naturale, le due seguenti asserzioni:

(a) un mucchio con un solo chicco è piccolo;

(b) se ad un mucchio piccolo si aggiunge un chicco allora il mucchio rimane piccolo.

Da (a) e (b) segue che un mucchio con due chicchi è piccolo. D'altra parte, una volta provato che un mucchio con due chicchi è piccolo, tramite (b) possiamo inferire che un mucchio con tre chicchi è piccolo ... e così via.

Per analizzare un tale paradosso procediamo ad una formalizzazione del ragionamento ora esposto. Indicando con $P(n)$ l'asserzione "un mucchio con n chicchi è piccolo", possiamo riscrivere le due assunzioni (a) e (b) come segue:

(a) $P(1)$; (b) $P(n) \rightarrow P(n+1)$,

dove, per evitare di coinvolgere inutilmente la logica del primo ordine, intendiamo (b) come *schema*, cioè come l'insieme di formule che si ottengono fissando in tutti i modi possibili il valore di n. Utilizziamo poi come regola di inferenza il *Modus Ponens*, che indicheremo in breve con *MP*, il quale, come è noto, permette di inferire da due premesse del tipo α e $\alpha \rightarrow \beta$ la conclusione β. Schematizziamo una applicazione di MP al modo seguente:

$$\frac{\alpha \ ; \ \alpha \rightarrow \beta}{\beta}$$

Fissiamo allora un qualunque numero m e proviamo che vale $P(m)$. Infatti, a partire dalla coppia di formule $P(1)$ e $P(1) \rightarrow P(2)$ ed applicando *MP,*

$$\frac{P(1) \ ; \ P(1) \rightarrow P(2)}{P(2)}$$

si ottiene $P(2)$. A partire da $P(2)$ e $P(2) \rightarrow P(3)$, applicando una seconda volta *MP*,

$$\frac{P(2) \ ; \ P(2) \rightarrow P(3)}{P(3)}$$

si ottiene $P(3)$,

. . .

A partire da $P(m\text{-}1)$ e $P(m\text{-}1) \rightarrow P(m)$, applicando la $(m\text{-}1)$-esima volta *MP,*

$$\frac{P(m\text{-}1) \ ; \ P(m\text{-}1) \rightarrow P(m)}{P(m)}$$

possiamo concludere che vale $P(m)$. Abbiamo allora provato il seguente teorema.

Teorema 1.2. Assumiamo a) e b), allora per ogni modo di fissare m esiste una dimostrazione di $P(m)$. Tale dimostrazione utilizza m-1 volte la regola *MP.*

Un paradosso simile è quello dell'uomo calvo, dovuto ad Euclide di Megara.

Teorema 1.3. Un uomo cui mancano completamente i capelli non è calvo.

Dim. Indichiamo con $NC(n)$ l'affermazione per cui un uomo con n capelli non è calvo e consideriamo le asserzioni

$NC(10.000)$

$NC(n) \rightarrow NC(n\text{-}1)$.

La prima asserzione esprime il fatto che un uomo con 10.000 capelli non è calvo, la seconda che se un uomo non calvo perde un capello allora continua ad essere non calvo. Allora, procedendo in maniera analoga a quanto fatto nel paradosso del mucchio di grano, è possibile dedurre la formula $NC(0)$.

Si noti che mentre abbiamo presentato il paradosso del mucchio di grano tramite un ragionamento *"dal basso all'alto"* (da $P(1)$ si giunge a $P(m)$), per quello dell'uomo calvo invece abbiamo esposto un ragionamento *dall'alto verso il basso* (da $NC(m)$ si giunge a $NC(1)$). In realtà i due tipi di ragionamento sono equivalenti. Ad esempio, indicando con *NP* la negazione di "essere un mucchio piccolo", si sarebbe potuto partire da $NP(m)$ e provare $NP(1)$. Infatti il fatto stesso che $P(m)$ appaia assurdo significa che posso ragionevolmente accettare l'assunzione $NP(m)$, inoltre le implicazioni (b) possono essere riscritte nella forma $NP(n) \rightarrow NP(n\text{-}1)$. È evidente allora che ciò consente di provare $NP(1)$, e quindi, nel nostro caso, giungere alla paradossale conclusione che un mucchio con un solo chicco non è piccolo.

2. Discussione del paradosso.

Se si vuole bloccare il paradosso due sono le possibilità:

a) si può mettere in discussione la correttezza del ragionamento

b) si possono mettere in discussione le premesse del ragionamento.

Per quanto riguarda *a*), la formalizzazione ora esposta mostra la struttura estremamente semplice del ragionamento che consiste in una lunga, ripetitiva catena di applicazioni del *MP*. Per quanto riguarda la possibilità b), non si può certo negare la validità di $P(1)$, cioè, che un mucchio con un solo chicco di grano sia piccolo. D'altra parte, negare che le implicazioni $P(n)\rightarrow P(n+1)$ siano tutte vere significa (nell'ambito della logica classica) ammettere l'esistenza di un intero n tale che $P(n)\rightarrow P(n+1)$ sia falsa e quindi, per la tavola di verità dell'implicazione, tale che $P(n)$ risulti vera e $P(n+1)$ falsa. Ciò condurrebbe ad ammettere l'esistenza di un *numero magico n* tale che un mucchio con n chicchi di grano è (ancora) piccolo, ma che tale mucchio cessa di essere piccolo non appena gli si aggiunge un solo chicco! Questo sembra poco accettabile specialmente per il fatto che contrasta con la flessibilità della nozione di essere piccolo. Su tale questione esamineremo quale è il punto di vista *"normativo"*, quello del *"finitismo stretto"*, quello delle *"supervalutazioni"* e, nei paragrafi successivi, quello della *"logica fuzzy"*.

Punto di vista *normativo*. Di fatto, è facile convincersi che il paradosso sussiste poiché abbiamo a che fare con un concetto come quello di "piccolo" che per sua natura è vago; esistono mucchi che sono sicuramente piccoli, mucchi che sicuramente non sono piccoli e mucchi per i quali non si saprebbe se considerarli piccoli o meno. D'altra parte che il problema abbia origine nella vaghezza viene dimostrato dal fatto che, come vedremo nel seguito, non è difficile riscrivere un analogo paradosso per ogni concetto vago. Ne segue che i diversi punti di vista circa il paradosso del mucchio di grano corrispondono ai diversi modi di considerare la questione dei concetti vaghi. Naturalmente, una soluzione radicale potrebbe allora consistere semplicemente nel dichiarare la insensatezza di ogni ragionamento che coinvolga nozioni vaghe. In tale senso il paradosso sarebbe una sorta di dimostrazione della impossibilità di trattare tali nozioni. Questo punto di vista è abbracciato ad esempio da G. Frege e B. Russell nella prospettiva della costruzione di un linguaggio "ideale" la cui caratteristica principale dovrebbe essere, appunto, la precisione. Non sarebbe un difetto della logica matematica quello di non riuscire a spiegare e formalizzare l'uso dei concetti vaghi; piuttosto è un difetto dei linguaggi naturali il fatto che in essi spesso siano presenti tali concetti. Ad esempio afferma Frege:

... ogni concetto deve avere confini netti, così che sia determinato per ogni oggetto se esso cade o no sotto il concetto (...) *La nota fallacia conosciuta come Acervus poggia sulla circostanza che viene trattato come concetto qualcosa* (*ad esempio il mucchio*) *che per la sua imperfetta delimitazione non può venire riconosciuto come tale in logica* (*Scritti postumi pag. 268*).

Naturalmente tale punto di vista è formalmente inattaccabile, tuttavia è un fatto che una larga parte del comportamento razionale dell'uomo nella vita di tutti i giorni coinvolge la trasmissione e l'elaborazione di informazioni relative a nozioni vaghe come quelle di *maturo*, *piccolo*, *grande*, *alto*, *bello* e così via. Dichiarare che informazioni di tale tipo vanno tenute fuori da un rigoroso discorso scientifico vuol dire accettare che la razionalità abbia spazio solo nei campi tradizionali della ricerca scientifica (la biologia, la chimica, la fisica, la matematica). Soluzione, questa, non solo, per così dire, rinunciataria, ma anche illusoria se, come vedremo in seguito, nemmeno tali campi possono evitare di coinvolgere nozioni vaghe ed i paradossi che ne conseguono.

Comunque è bene osservare che il punto di vista normativo è abbastanza diffuso quando si debba passare dalla vita di tutti i giorni ad una regimentazione del comportamento umano. In tali casi si è portati a sostituire i concetti vaghi con una loro, necessariamente rozza, traduzione in un concetti precisi. Ad esempio, a volte ci si trova nella necessità di formulare una legge che fa scattare una possibilità o una sanzione (non vaga) non appena una condizione (vaga) risulti verificata. Potrebbe essere il caso di una legge del tipo:

"tutti quelli che hanno fatto abbastanza precariato saranno immessi in ruolo".

Ovviamente tale legge non può essere applicata se non si stabilisce che cosa si intende per "avere fatto abbastanza precariato". Allora si deve procedere ad una determinazione "rigida" del tipo: "avere dieci anni un mese ed un giorno di servizio". In altri casi una legge può coinvolgere la nozione di "essere maturo" che viene tradotta in quella di "essere maggiorenne" che a sua volta si caratterizza con la proprietà di avere una età maggiore di 18 anni.

Ma tali traduzioni in termini classici presentano sempre qualcosa di paradossale e di insoddisfacente. Ad esempio comportano che una persona che ha solo dieci anni ed un mese di servizio, per un solo giorno non può entrare in ruolo. Ancora, il reato di seduzione di minorenne potrebbe scattare se compiuto un secondo prima della mezzanotte di un dato giorno e non scattare un secondo dopo.

Il finitismo stretto. Un modo per negarne la validità del ragionamento con cui si prova il paradosso, almeno che non si voglia giungere a rifiutare il *MP*, sarebbe quello di negare che una dimostrazione i cui passi siano singolarmente validi sia sempre una dimostrazione valida. È questo il punto di vista del *finitismo stretto* che attribuisce significato solo a quei concetti matematici (o a quelle dimostrazioni) che concretamente possono essere verificati o costruiti. Ne segue che solo le dimostrazioni "non troppo grandi" possono essere considerate accettabili e quindi che il ragionamento che conduce al paradosso, essendo lunghissimo, non risulterebbe essere valido. Per una interessante ed estesa esposizione di una tale soluzione si rimanda al saggio *Il paradosso di Wang* di M. Dummett. Qui mi limito solo ad osservare che se così si riesce a risolvere il paradosso del mucchio di grano non è difficile trovare paradossi analoghi che, realizzandosi in un numero molto piccolo di passi, non possono essere risolti in questo modo. Ad esempio consideriamo l'implicazione

$$(molto\_vicino(x,y) \wedge molto\_vicino(y,z)) \rightarrow molto\_vicino(x,z)$$

o, se si vuole, quella equivalente

$$molto\_vicino(x,y) \rightarrow (molto\_vicino(y,z) \rightarrow molto\_vicino(x,z)),$$

e supponiamo che $s_1,...,s_{20}$ siano i nomi di venti stazioni ferroviarie successive tali che ogni stazione s_i sia a distanza di venti chilometri dalla successiva s_{i+1}. Allora poiché possiamo accettare la validità delle formule $molto\_vicino(s_i,s_{i+1})$, con un ragionamento di venti passi possiamo pervenire all'assurdo per cui s_1 è molto vicino a s_{20} pur essendo queste due stazioni a 400 chilometri di distanza. D'altra parte un ragionamento di 20 passi, potendosi facilmente scrivere o leggere, è difficilmente criticabile anche dal più stretto finitismo.

La teoria delle supervalutazioni. Un modo diverso di risolvere il paradosso è tentare di stabilire delle regole formali che consentano di trattare i concetti vaghi. Un proposta in tale senso, che utilizza la teoria delle supervalutazioni, è stata delineata da Dummett nell'articolo citato e sviluppata da K. Fine. Per Fine le nozioni vaghe si caratterizzano per il fatto di avere molte possibili "determinazioni", ciascuna determinazione dipendendo dal contesto in cui ci si muove. Ad esempio la nozione di "mucchio piccolo" può essere resa precisa assumendo che siano piccoli tutti i mucchi al di sotto dei 1000 chicchi, oppure tutti quelli al di sotto di 1001 chicchi e così via. In un certo senso il significato un concetto vago è dato proprio dall'insieme delle sue possibili determinazioni. In accordo con tale punto di vista un'asserzione α coinvolgente la nozione vaga R deve essere considerata vera (falsa) solo quando α risulta vera (falsa) per ciascuna delle possibili determinazioni di R. Per evitare confusioni terminologiche in questo caso diremo che α è *SV-vera* (rispettivamente *SV-falsa*), mentre diremo che α è *indeterminata* se non risulta essere né *SV*-vera né *SV*-falsa. Ad esempio $P(1000)$ è da considerare indeterminata essendoci dei contesti in cui un mucchio di 1000 chicchi è grande e contesti in cui è piccolo. Invece $P(1)$ è vera e $P(10^{100})$ è falsa. È immediato verificare che *MP* è una regola di inferenza valida rispetto alla nozione di *SV*-verità. Infatti se α è vera per ogni determinazione classica e $\alpha\rightarrow\beta$ è vera per ogni determinazione classica allora β è vera per ogni determinazione classica. Ne segue che il ragionamento che conduce al paradosso è corretto anche se si accetta la logica delle supervalutazioni. Tuttavia molte delle implicazioni $P(n)\rightarrow P(n+1)$ non sono *SV*-vere e ciò non consente di concludere che $P(n)$ sia *SV*-vera. Consideriamo infatti una possibile determinazione della nozione di "essere un mucchio piccolo", e quindi l'insieme M di mucchi di grano che si è deciso di considerare piccoli (ovviamente questo può essere fatto in vari modi). Sia m la più grande delle cardinalità dei mucchi appartenenti ad M, allora essendo $P(m)$ vera e $P(m+1)$ falsa, l'mplicazione $P(m)\rightarrow P(m+1)$ risulta falsa per questa determinazione. Si osservi che non può essere asserita l'esistenza di un intero m tale che $P(m)\rightarrow P(m+1)$ sia *SV*-falsa e ciò ci solleva dalla difficoltà a dover accettare l'esistenza di un numero magico m per cui $P(m)$ sia vera e $P(m+1)$ falsa. Infatti se $P(m)\rightarrow P(m+1)$ fosse falsa per una determinazione M non lo sarebbe tuttavia nella determinazione M' che considera piccoli tutti i mucchi che hanno meno di $m+1$ chicchi. È interessante ancora osservare che mentre la formula $\exists n(P(n)\wedge\neg P(n+1))$ risulta *SV*-vera, per nessun intero n possiamo affermare che $P(n)\wedge\neg P(n+1)$ sia *SV*-vera. Infatti in ogni determinazione di P esiste un intero n per cui $P(n)$ è vera e $P(n+1)$ è falsa, tuttavia non esiste un unico intero n per cui ciò si verifichi qualunque sia la determinazione.

Anche una tale soluzione al paradosso mi sembra insoddisfacente. Infatti con la teoria delle supervalutazioni tutte le argomentazioni coinvolgenti concetti vaghi vengono di fatto annullate per l'indeterminatezza delle premesse e ciò rende non molto dissimile il punto di vista delle supervalutazioni da quello normativo. Inoltre il passare dalla categoria vero-falso della logica classica a quella vero-falso-indeterminato della teoria delle supervalutazioni non sembra rendere conto della ricchezza e flessibilità dei concetti vaghi. Quello che è discutibile è l'assunzione per cui il significato di un concetto vago possa essere identificato con l'insieme delle sue possibili determinazioni. In alcuni casi una tale punto di vista è forse giustificato: ad esempio la nozione giuridica di “maggiorenne” può assumere in epoche e paesi differenti significati diversi, ognuno dei quali esattamente determinato dall’aver fissato una età-soglia. Tuttavia, a differenza dell’analogo concetto di “essere maturo”, sarebbe meglio dire che il significato di “maggiorenne” più che essere vago dipende dal contesto. Comunque nella maggioranza dei casi una tale riduzione non sembra molto giustificabile e ciò viene messo in rilievo in modo particolare quando si rapportino tra loro più proprietà vaghe. Supponiamo ad esempio di volere definire la nozione di “x è un guidatore prudente” e che per fare questo si debba considerare, tra le altre, la proprietà

se x è vicino ad un casello autostradale allora x procede con velocità piccola.

In tale espressione non solo sono coinvolti due concetti vaghi, *vicino* e *piccolo,* viene espresso anche un qualche collegamento tra i gradi con cui tali concetti sono verificati.

Il significato sottointeso è che tanto più x è vicino ad un casello quanto più la sua velocità è piccola. D’altra parte se determiniamo il concetto di vicino con “a meno di 500 metri ” e quello di “velocità piccola” con “a meno di 30 chilometri orari” otteniamo due determinazione dei concetti vaghi in gioco ciascuna ragionevole se presa a se stante. La corrispondente determinazione della frase sopra scritta sarebbe allora

se x è a meno di 500 metri da un casello autostradale allora x procede a meno di 30 chilometri orari.

Pertanto si perviene al paradosso per cui una persona che arriva ad un metro da un casello alla velocità di 30 chilometri orari debba essere considerata prudente.

3. Soluzione proposta dalla logica fuzzy.

Riprendiamo in esame le implicazioni $P(n)\rightarrow P(n+1)$ sulle quali, a differenza della formula $P(1)$, si è portati a nutrire un qualche sospetto, e scegliamo non di considerare alcune di queste come false e nemmeno come indeterminate ma piuttosto di accettare che siano tutte “quasi completamente vere”. Per formalizzare poi la modalità “quasi completamente vero”, appare naturale riferirsi a valori di verità diversi da 0 ed 1 (denotanti rispettivamente il falso ed il vero) che siano quindi capaci di rappresentare livelli diversi di verità. Ad esempio possiamo considerare come possibili valori i numeri dell’intervallo [0,1] e quindi assumere:

$P(1)$	con grado di verità 1
$P(n)\rightarrow P(n+1)$	con grado di verità 0.9.

A suo volta, l’introduzione di tali valori di verità rende necessario affiancare a *MP* una regola che dica come il grado con cui si può ritenere provata la conclusione dipenda dai gradi con cui sono state provate le premesse. Il *MP* allora assumerà la forma:

$$\frac{\alpha\ ;\ \alpha\rightarrow\beta}{\beta} \quad , \qquad \frac{x\ ;\ y}{x\otimes y}$$

con $\otimes$ opportuna operazione binaria in [0,1]. Una applicazione di tale regola sarà letta al modo seguente: se α è stata provata con grado x e $\alpha\rightarrow\beta$ con grado y allora β si può ritenere provata con grado $x\otimes y$. L’operazione $\otimes$ corrisponde alla valutazione di quanto sia lecito accettare la congiunzione della premessa α e della premessa $\alpha\rightarrow\beta$, pertanto deve essere una ragionevole interpretazione della congiunzione in una logica a più valori. In letteratura usualmente ci si riferisce ad una particolare classe di operazioni chiamate *T-norme continue*. Due esempi che considereremo in questo paragrafo sono l’usuale moltiplicazione tra reali ed il prodotto di Lukasievicz che si definisce ponendo

$x\otimes y=x+(y-1)$ se $x+y-1>0$

$x \otimes y = 0$ altrimenti.

Nel paragrafo 7 considereremo il caso in cui $\otimes$ è il minimo, cioè $x \otimes y = min\{x,y\}$. La seguente proposizione evidenzia alcuni aspetti di tali T-norme il cui interesse apparirà nel seguito. La nozione di potenza n-esima di un elemento γ di [0,1] rispetto ad una T-norma $\otimes$ viene definita al solito modo tramite le equazioni $\gamma^1 = \gamma$, $\gamma^{n+1} = \gamma^n \otimes \gamma$.

Proposizione 3.1. Sia γ un elemento di [0,1]diverso da 0 ed 1, allora
- se $\otimes$ è l'usuale prodotto allora $\gamma^n \neq 0$ per ogni intero n e $\lim_{n\to\infty} \gamma^n = 0$;
- se $\otimes$ è il prodotto di Lukasievicz allora $\gamma^n = 0$ per ogni $n \geq 1/(1-\gamma)$ e quindi $\lim_{n\to\infty} \gamma^n = 0$;
- se $\otimes$ è il minimo allora $\gamma^n = \gamma \neq 0$ per ogni $n \geq 1/(1-\gamma)$ e quindi $\lim_{n\to\infty} \gamma^n \neq 0$.

Dim. Ci limitiamo ad osservare che se $\otimes$ è il prodotto di Lukasievicz allora ponendo per comodità $\gamma = 1-\delta$

$(1-\delta)^2 = 1 - \delta + 1 - \delta - 1 = 1-2\cdot\delta$

$(1-\delta)^3 = 1 - 2\cdot\delta + 1 - \delta - 1 = 1 - 3\cdot\delta$,

e, più in generale

$(1-\delta)^n = 1-n\cdot\delta$ se $1 - n\cdot\delta > 0$

$(1-\delta)^n = 0$ se $1 - n\cdot\delta \leq 0$.

Ritorniamo ora al ragionamento che conduce al paradosso del mucchio di grano, e formalizziamolo al modo seguente:

Poiché

$P(1)$ [vale con grado 1]

e

$P(1)\to P(2)$ [è accettato con grado 0.9]

allora, applicando *MP*,

$P(2)$ [è provato con grado $1\otimes 0.9=0.9$].

Poiché

$P(2)$ [è provato con grado 0.9]

e

$P(2)\to P(3)$ [è accettato con grado 0.9]

allora, applicando una seconda volta *MP*,

$P(3)$ [è provato con grado 0.9^2].

. . .

Poiché

$P(m-1)$ [è provato con grado 0.9^{m-2}]

e

$P(m-1)\to P(m)$ [è accettato con grado 0.9]

allora, applicando per la (m-1)-esima volta *MP*,

$P(m)$ [è provato con grado 0.9^{m-1}].

Teorema 3.2. Per ogni modo di fissare m esiste una dimostrazione del fatto che un mucchio di m chicchi è piccolo ma tale dimostrazione prova tale asserzione solo con grado pari a 0.9^{m-1}.

Da tale teorema segue che sia la logica in cui $\otimes$ è il prodotto che quella in cui $\otimes$ è il prodotto di Lukasievicz consentono di risolvere il paradosso. Infatti, supposto *m* sufficientemente grande, nel primo caso 0.9^{m-1} è quasi uguale a zero, nel secondo caso 0.9^{m-1} è uguale a zero. La scelta tra le due logiche dipende dall'idea che si ha di "piccolo" o, se si vuole, della sua negazione "grande". Infatti l'usuale prodotto è preferibile se si ritiene che l'attributo di grande non sia mai pienamente realizzato da nessun mucchio. Tale interpretazione consente meglio la comparazione se si considera che dato un mucchio grande se ne può trovare sempre uno più grande. Il prodotto di Lukasievicz corrisponde invece ad una nozione di "grande" per cui sono grandi con grado uno tutti i mucchi al di sopra di una certa dimensione.

Se invece si considera il paradosso dell'uomo calvo, il prodotto di Lukasievicz appare il più adatto. Supponiamo infatti di assegnare all'implicazione $NC(n)\rightarrow NC(n\text{-}1)$ grado di verità γ, e che quindi il grado con cui è possibile provare "un uomo con zero capelli è non calvo" sia $(\gamma)^{10.000-1}$. Ora nel caso del prodotto, per quanto si voglia attribuire una basso grado di verità all'asserzione $NC(n)\rightarrow NC(n\text{-}1)$, il numero $(\gamma)^{10.000-1}$ è comunque diverso da zero come sarebbe ragionevole aspettarsi. Invece, nel caso del prodotto di Lucasievicz, se assegniamo all'implicazione $NC(n)\rightarrow NC(n\text{-}1)$ grado di verità $\gamma = 1-\delta$ con $\delta = 1/(10.000\text{-}1)$, allora, perveniamo alla conclusione $NC(0)$ con grado $(\gamma)^{10.000-1}$ $=0$, cosa che sembra più corretta.

4. Modelli per le nozioni vaghe: i fuzzy sottoinsiemi.

Procediamo ora ad una maggiore precisazione mostrando come si possa modellizzare la nozione di concetto vago. Il discorso può essere impostato facendo riferimento al famoso principio di comprensione della teoria degli insiemi. Ricordiamo che una proprietà P si dice *ben definita* in un insieme S se per ogni $x\in S$ è possibile dire se x verifica o meno P. In questo caso *l'assioma di comprensione* (per meglio dire l'assioma di isolamento) assicura che la collezione di elementi di S che verifica P è un insieme. Consideriamo ora proprietà vaghe come "grande", . . . "alto" . . . "calvo"… che non sono ben definite.

Problema: è possibile applicare un assioma di comprensione a tali proprietà ?

In altre parole: hanno senso espressioni come *"l'insieme dei numeri grandi"*, *"l'insieme degli uomini alti"*, *"l'insieme degli uomini calvi"* ? Per affrontare tale problema ricordiamo che ogni sottoinsieme X di S può essere identificato con la sua funzione caratteristica $c : X\rightarrow\{0,1\}$, dove

$$c(x) = \begin{cases} 1 & se \quad x \in X \\ 0 & se \quad x \notin X. \end{cases}$$

Ciò suggerisce di rappresentare sottoinsiemi corrispondenti a proprietà non ben definite con "*funzioni caratteristiche generalizzate*".

Definizione 4.1. (Zadeh [1965], Goguen [1968/69]) Sia $L=(V, \leq, 0, 1)$ un insieme ordinato il cui massimo elemento è 1 e minimo *0* e chiamiamo *valori di verità* i suoi elementi. Un *L-sottoinsieme* di S (*sottoinsieme fuzzy*) è una funzione $s{:}S\rightarrow V$.

Gli elementi 0 ed 1 rappresentano il falso ed il vero. Il valore $s(x)$ è interpretato come il *grado di appartenenza* di x ad s. Inoltre, diciamo che s è *incluso* in s' se $s(x) \leq s'(x)$ per ogni $x \in S$. Ad esempio al concetto di "essere un mucchio piccolo" corrisponderà una funzione $s:M \rightarrow [0,1]$ dove M è l'insieme dei mucchio di grano e dove abbiamo assunto che $V=[0,1]$. Naturalmente esistono molti modi di scegliere s e quindi di modellizzare la nozione di mucchio piccolo. Tuttavia tale scelta non è certo arbitraria; ad esempio s deve essere decrescente rispetto alla cardinalità di x e tendente a zero quando tale cardinalità tende all'infinito. In tale senso la teoria dei fuzzy sets come la teoria delle supervalutazioni continua ad accettare che ci siano diverse determinazioni di un concetto vago. Tuttavia tali determinazioni non sono in generale degli insiemi classici. Naturalmente non è chiaro fino a quale punto la nozione di sottoinsieme fuzzy costituisce uno strumento adeguato a rappresentare i concetti vaghi. Si vedano ad esempio le interessanti considerazioni di S. Termini nel suo "Su alcuni sviluppi delle scienze dell'informazione".
Per definire nella classe L^S degli L-sottoinsiemi di S le operazioni di unione, intersezione e complemento ricordiamo che nella teoria degli insiemi classica tali operazioni sono il corrispondente dei connettivi logici ET, OPPURE, NON. Infatti, dati due insiemi A e B, si pone
$A \cap B = \{ x \mid x \in A \text{ ET } x \in B\}$
$A \cup B = \{ x \mid x \in A \text{ OPPURE } x \in B\}$
$-A = \{ x \mid \text{ NON } x \in A\}$.
Ne segue che abbiamo bisogno di avere tre operazioni in L che interpretino rispettivamente la congiunzione, la disgiunzione e la negazione.

Definizione 4.2. Una *struttura di valutazione* è una struttura del tipo $L=\{V, \leq, \otimes, \oplus, \neg ,0,1\}$ con $\leq$ relazione d'ordine con minimo 0 e massimo 1 e tale che le operazioni $\otimes$, $\oplus$, $\neg$ siano una estensione delle corrispondenti operazioni nell'algebra di Boole $2=\{0,1\}$.

Ad esempio possiamo considerare la struttura $L=([0,1], \leq, \wedge, \vee, \neg, 0, 1\}$ dove, per ogni $x,y \in [0,1]$

$$x \wedge y = \min\{x,y\} \ , \ x \vee y = \{x,y\} \ , \ \neg(x) = 1-x.$$

Definizione 4.3. Sia S un insieme ed L una struttura di valutazione, allora dati due L-sottoinsiemi s e s', definiamo $s \cup s'$, $s \cap s'$ e $-s$ ponendo, per ogni $x \in S$

$$(s \cup s')(x) = s(x) \oplus s'(x) \quad ; \quad (s \cap s')(x) = s(x) \otimes s'(x) \quad ; \quad (-s)(x) = \neg s(x).$$

In altri termini, $(s \cup s')(x)$ è una valutazione nella logica a più valori dell'affermazione
x appartiene ad s oppure x appartiene ad s' ed analogo discorso vale per $(s \cap s')(x)$ e $(-s)(x)$. Con riferimento alla struttura di valutazione che abbiamo ora esposto, le operazioni di unione di intersezione e di complemento sono definite al modo seguente:

$$(s \cup s')(x) = s(x) \vee s'(x) \quad ; \quad (s \cap s')(x) = s(x) \wedge s'(x) \quad ; \quad (-s)(x) = 1-s(x).$$

L'aver dato la nozione di fuzzy sottoinsieme consente di dimostrare in maniera rigorosa la seguente proposizione.

Proposizione 4.4. È possibile avere un paradosso simile a quello del mucchio di grano in corrispondenza di ogni concetto vago (purché possa variare gradualmente dal vero al falso).

Dim. Sia P un concetto vago rappresentabile da un fuzzy sottoinsieme s di S e sia $(x_n)_{n\in N}$ una successione di elementi di S tale che

(i) $s(x_1)=1$; (ii) $s(x_n)_{n\in N}$ è decrescente e tende a zero ; (iii) $s(x_n)-s(x_{n+1})<\varepsilon$

con ε tanto piccolo da rendere non percettibile la differenza tra $s(x_n)$ e $s(x_{n+1})$.
Allora sono accettabili gli assiomi $P(x_1)$ e $P(x_n)\to P(x_{n+1})$ e quindi si può riprodurre il paradosso del mucchio di grano. Infatti si può provare $P(x_n)$ per n arbitrariamente grande, mentre la valutazione $s(x_n)$ di tale asserzione è prossima a zero.

Ci si può divertire a costruire paradossi assumendo ad esempio che P significhi *essere magro*, *essere pulito, essere vicino* . . . . Ad esempio si può provare che per quanti cioccolatini una persona *magra* possa mangiare la persona rimane magra, oppure che due punti dello spazio sono sempre *vicini,* e così via.

5. Principio di induzione e insiemi che sono finiti e infiniti.
È possibile formalizzare il paradosso del mucchio di grano facendo riferimento al principio di induzione

$$\forall P\ [P(1)\to((\forall n(P(n)\to P(n+1))\to \forall nP(n))]$$

invece che al solo *MP*. Infatti proviamo ad accettare, insieme alle due formule
(i) $P(1)$
(ii) $\forall n(P(n)\to P(n+1))$
anche il principio di induzione applicato al predicato P
(iii) $P(1)\to((\forall n(P(n)\to P(n+1))\to \forall nP(n))$.
Allora due semplici applicazioni del *MP* permettono di provare la formula $\forall nP(n)$ e quindi, per particolarizzazione, $P(m)$ per ogni prefissato m. Da notare che la lunghezza di una tale dimostrazione non dipende dalla grandezza di m come invece avveniva nelle dimostrazioni esposte nel primo paragrafo.

In questo caso la logica fuzzy non sembra poter bloccare il paradosso se non con un completo rifiuto del principio di induzione (la stessa cosa si può dire del finitismo stretto). Infatti, sulla linea di quanto fatto nel paragrafo 3, proviamo ad attribuire a (ii) valore di verità 0.9 ed a (iii) un valore x da stabilirsi. Allora la conclusione $\forall nP(n)$ può essere provata con grado $1\otimes 0.9\otimes x$ e da ciò segue, se si vuole evitare il paradosso, che x deve risultare uguale a 0.
Un'argomentazione più generale, di carattere semantico, che giustifica il rifiuto del principio di induzione è la seguente.
Consideriamo, per fissare un esempio, la logica del primo ordine di Lukasievicz in cui la struttura di valutazione è l'intervallo [0,1] ed in cui
- il quantificatore universale e quello esistenziale vengono interpretati rispettivamente tramite l'estremo inferiore e l'estremo superiore
- l'implicazione $\to$ è interpretata ponendo $x\to y = (-x)\oplus y=(y - x+1)\vee 1$.
Allora vale la seguente proposizione che mostra come l'accettazione dei concetti vaghi (e quindi delle logiche a più valori) comporta l'esclusione del principio di induzione.

Proposizione 5.1. Se si accettano concetti vaghi il principio di induzione è falso.

Dim. La valutazione della formula del secondo ordine che esprime il principio di induzione è l'estremo inferiore delle valutazioni della formule $P(1)\rightarrow((\forall n(P(n)\rightarrow P(n+1))\rightarrow \forall nP(n))$ al variare del predicato P.
Fissiamo un qualunque valore $\delta\neq 0$ nell'intervallo $[0,1]$ e supponiamo che P sia un predicato vago interpretato dal fuzzy sottoinsieme $s:N\rightarrow[0,1]$ definito dalle equazioni: $s(1)=1$ e $s(n+1)=(s(n)-\delta)\vee 0$.
Allora, se denotiamo con $|\alpha|$ la valutazione di una formula α, tramite un semplice calcolo si vede che
$|\forall n(P(n)\rightarrow P(n+1))| = 1-\delta$; $|P(1)| = 1$; $|\forall n(P(n)| = 0$,
e quindi che
$|P(1)\rightarrow((\forall n(P(n)\rightarrow P(n+1))\rightarrow \forall nP(n))|$
$= -|P(1)|+|((\forall n(P(n)\rightarrow P(n+1))\rightarrow \forall nP(n))|+1$
$= -|((\forall n(P(n)\rightarrow P(n+1))|+|\forall nP(n))|+1 = -1+\delta+1 = \delta.$
Ne segue che la valutazione della formula $\forall P\ [P(1)\rightarrow((\forall n(P(n)\rightarrow P(n+1))\rightarrow \forall nP(n))]$ è zero.

È da notare che, in accordo con il significato che assume il quantificatore universale nelle logiche a più valori, non è stata dimostrata la esistenza di un particolare predicato che rende falso il principio di induzione. Piuttosto si è visto che, comunque si consideri un δ, è possibile trovare un predicato (vago) per cui il principio di induzione vale con grado minore o uguale a δ.
Una tale negazione del principio di induzione può essere vista anche da un altro punto di vista. Riprendendo una idea di Esenin-Volpin, Dummett introduce le nozioni di debolmente finito e debolmente infinito.

Definizione 5.2. Chiamiamo *debolmente finita* una totalità C di oggetti per cui esista un buon ordinamento ed un intero n in modo che non esista un n-esimo elemento in C. Diciamo che C *è debolmente infinita* se esiste un buon ordinamento in C senza un ultimo elemento.

Normalmente si sarebbe portati a pensare che non possa esistere un insieme che sia allo stesso tempo debolmente finito e debolmente infinito e questa è anche l'opinione di Dummett. Tuttavia vale la seguente proposizione.

Proposizione 5.3. Esiste una collezione C che è sia debolmente finita che debolmente infinita.

Dim. Negando il principio di induzione si ottiene la formula
$$\exists P\ [P(1)\wedge(\forall n(P(n)\rightarrow P(n+1))\wedge \exists n\text{-}P(n)].$$
Tale formula, che per la proposizione 5.1 risulta vera, afferma appunto l'esistenza di una collezione C che è allo stesso tempo debolmente finita e debolmente infinita.

Sul rapporto tra infinito e vaghezza si veda anche Sanford [1975] e Ferrante e Gerla [1998].

6. Paradossi nella teoria dell'evoluzione e nella meccanica.

Anche all'interno dei settori della scienza più solidi il coinvolgimento di concetti vaghi è inevitabile e quindi in essi possono nascere paradossi simili a quello del mucchio di grano.

Teoria dell'evoluzione. Per prima cosa proviamo che la teoria dell'evoluzione è sbagliata (una forma non molto diversa di tale paradosso è esposta da D. H. Sanford nell'articolo citato). Scriviamo $Ant(x',x,n)$ per indicare che x' è un antenato di x nato almeno n anni prima di x e consideriamo le due asserzioni

(a) $Ant(x'',x',p) \wedge Ant(x',x,q) \rightarrow Ant(x'',x,p+q)$

(b) $\forall x \exists x' Ant(x',x,10)$

La verità della prima asserzione è una immediata conseguenza del significato che si è dato al predicato *Ant*. La verità di (b), che afferma che ogni persona x ha un antenato x' vissuto almeno 10 anni prima, risulta provata dal porre x' uguale al padre (o alla madre) di x.

Proposizione 6.1. Sia c (il nome di) una persona ed n un qualunque intero positivo, allora è possibile provare la formula $\exists x' Ant(x',c,10\times n)$.

Dim. Procediamo per induzione sull'intero n. Per n=1 l'asserto è una immediata conseguenza di (b). Supposto il teorema vero per n, sia c' tale che $Ant(c',c,10\times n)$ e sia c'' tale che $Ant(c'',c',10)$. Allora, se particolarizziamo (a)

$$Ant(c'',c',10) \rightarrow (Ant(c',c,10\times n) \rightarrow Ant(c'',c,10\times(n+1)),$$

Applicando due volte *MP*, otteniamo $Ant(c'',c,10\times(n+1))$ e quindi $\exists x' Ant(x',c,10\times(n+1))$.

Corollario 6.2. Io ho avuto un antenato extraterrestre.

Dim. Dalla proposizione 6.1 segue che io ho avuto un antenato vissuto almeno $10\times10^{100.000}$ anni fa. Conseguentemente, poiché la terra $10\times10^{100.000}$ anni fa non esisteva ancora, $10\times10^{100.000}$ anni fa esisteva un mio antenato extraterrestre.

Possiamo allora concludere con il seguente teorema.

Teorema 6.3. La teoria dell'evoluzione di Darwin è sbagliata.

Naturalmente non è una tale conclusione ad essere paradossale; quello che è paradossale è il fatto che si possa giungere ad essa per via puramente logica (logica classica, ovviamente) senza alcun riferimento ad osservazioni di fatto. Non tenterò di dare una soluzione formale di tale paradosso tramite la logica fuzzy, cosa che comunque non dovrebbe essere difficile. Tuttavia sicuramente è la nozione di uomo (che è sottintesa nell'accettazione dell'assioma $\forall x \exists x' Ant(x',x,10)$) che deve essere considerata vaga non appena la si consideri nella scala temporale della teoria dell'evoluzione. Per meglio dire, accettare la possibilità che la nozione di uomo sia vaga è un presupposto filosofico della teoria dell'evoluzione stessa. Sembrerebbe in definitiva che la logica classica non sia capace di padroneggiare situazioni in cui gli oggetti del discorso subiscono una evoluzione.

Meccanica classica e relativistica. Passiamo ora ad un settore della scienza che viene considerato il paradigma della razionalità: la meccanica. Partiamo dal seguente (paradossale) dato di fatto:

La meccanica classica è una teoria in cui (*quasi*) *tutti i teoremi sono errati e tuttavia è una teoria verificata dall'esperienza quotidiana.*

Che la teoria classica, in breve *TC*, sia verificata nell'esperienza quotidiana viene giustificato spesso tramite un "principio di inclusione" secondo cui la teoria della relatività speciale, in breve *TR*, "contiene" la meccanica classica. Ciò nel senso che, per velocità "molto piccole" (e qui compare una nozione vaga!) le proposizioni provate dalla meccanica relativistica coincidono con quelle provate nella meccanica classica. Tuttavia è alquanto difficile dare un senso logico preciso a tale affermazione. Ad esempio, poniamoci in un contesto in cui tutte le velocità sono molto piccole e supponiamo che, data una grandezza fisica g:
- in *TR* si provi l'asserzione α_R = "g=0.567662220".
Allora è presumibile che
- in *TC* sia dimostrabile un asserzione del tipo α_C = "g=0.567662221".
Poiché α_R e α_C sono in contraddizione tra di loro, appare chiaro che la meccanica classica, ben lontana dall'essere contenuta in quella relativistica, appare in totale contraddizione con questa (almeno se si usa la logica classica). Le cose migliorano un poco se si decide di limitare il grado di precisione delle nostre asserzioni, accettando, ad esempio, solo numeri reali con otto cifre decimali. In tale caso infatti le formule α_R e α_C coincidono con la formula "g=0.56766222". Allora si potrebbe formulare il principio di inclusione al modo seguente:
Sia α un teorema in TC, e supponiamo che l'ambito considerato coinvolge solo velocità piccole. Allora, se ci si accontenta di misure approssimate, α è un teorema di TR.
Purtroppo un tale principio coinvolge una nozione vaga come quella di "velocità piccola" e non è pertanto difficile trovare paradossi corrispondenti ad un tale coinvolgimento. Ad esempio si può immaginare una successione finita di riferimenti R_1, R_2, . . .,R_n tale che ciascun R_{i+1} si muova molto lentamente rispetto ad R_i e tuttavia l'ultimo riferimento R_n si muova con velocità paragonabile a quella della luce rispetto ad R_1. Allora, poichè il comportamento di ciascun R_{i+1} rispetto a R_i è regolato dalla meccanica classica, anche il comportamento di R_n rispetto a R_1 dovrebbe essere regolato dalla meccanica classica.
Sembra allora che anche per la meccanica, paradigma della razionalità scientifica, nasca l'esigenza di ricorrere agli strumenti della logica fuzzy. Il modo come questo debba essere fatto non è certamente immediato e tale impresa potrebbe non essere concretamente fattibile.

Il paradosso di Poincaré. Un paradosso simile a quello del mucchio di grano o, più precisamente, simile a quello che abbiamo asssociato al predicato *molto_vicino*, è stato esaminato dal famoso matematico H. Poincaré. Supponiamo che A, *B* e C siano i pesi di tre corpi e che io non sia capace di distinguere il peso A dal peso *B* ed il peso *B* dal peso C ma che tuttavia percepisca una differenza tra *A* e *C*. Allora avremmo che

$$A = B \quad , \quad B = C, \quad C \neq A,$$

cioè la non transitività della relazione di indiscernibilità. La paradossale compresenza di tali formule viene vista da Poincaré come una caratteristica fondamentale di ogni esperienza che possiamo avere del continuo fisico. Possiamo chiamare "paradossi della indiscernibilità" paradossi di tale tipo. Il seguente è un paradosso della indiscernibilità che appare più sorprendente.

Theorema 6.4. Non è possibile leggere un orologio (non è possibile percepire il movimento).

Dim. Indico con *Ind* la relazione di indiscernibilità e con $P(m)$ la posizione della lancetta delle ore dopo m secondi. Allora non sono in grado di distinguere $P(m)$ da $P(m+1)$ ed esprimo tale fatto tramite la formula:
i) $Ind(P(m),P(m+1))$.
Inoltre ammetto la formula che esprime la transitività della relazione di indiscernibilità:
ii) $Ind(P(0),P(m)) \rightarrow (Ind(P(m),P(m+1)) \rightarrow Ind(P(0),P(m+1)))$.
Sotto tali ipotesi è possibile provare la formula:
iii) $\forall m\ Ind(P(0),P(m))$
che esprime la impossibilità di leggere l'orologio. Infatti, procedendo per induzione su m, osserviamo che la formula $Ind(P(0),P(m))$ è vera per m=1. Supposta vera per m, poiché
$Ind(P(m),P(m+1))$ e $Ind(P(0),P(m))(\rightarrow(Ind(P(m),P(m+1)) \rightarrow Ind(P(0),P(m+1)))$
per MP ricaviamo $Ind(P(0),P(m+1))$.

Naturalmente è possibile risolvere tali paradossi tramite la logica fuzzy interpretando l'indiscernibilità con una relazione fuzzy del tipo "essere vicino" ed attribuendo alla formula che esprime la transitività un valore di verità diverso da 1. In tale senso si muove ad esempio U. Höhle. Tuttavia non è chiaro se una tale operazione sia giustificata e non sono sicuro che sia corretto assumere che la indiscernibilità sia un predicato vago. D'altra parte la soluzione del paradosso proposta da Poicarè non ha a che fare con la vaghezza. Dal momento che la indiscernibilità non è transitiva ma la identità lo è, bisogna semplicemente distinguere la possibilità di discernere due grandezze fisiche dal loro essere diverse (cioè distinguere l'identità dalla indiscernibilità). Ciò spinge a creare un modello la realtà tramite un continuo matematico basato sulla teoria dei numeri reali.

7. La logica fuzzy del minimo.

Le due logiche sino ad ora considerate hanno in comune la seguente caratteristica:
più volte si applica MP in un ragionamento, meno questo ragionamento è da considerare affidabile.
In altre parole, il coinvolgimento di concetti vaghi, sembra comportare l'esigenza di limitarsi solo a ragionamenti "brevi". È interessante osservare come ciò sia istintivamente sentito dalle persone. Ad esempio, consideriamo il predicato di "essere molto vicino" ed ammettiamo, come sembra naturale, che per tale predicato valga la proprietà transitiva. Supponiamo inoltre di sapere che:
- un certo bar è molto vicino al tale albergo
- l'albergo è molto vicino alla stazione,

allora non abbiamo difficoltà ad inferire che il bar è molto vicino alla stazione. Ma se apprendiamo ancora che:
- la stazione è molto vicina al Duomo,
- il Duomo è molto vicino all'Università
- l'Università e molto vicina al Municipio,
allora preferiamo non andare più in avanti nel ragionamento e non concludere che il bar è molto vicino al Municipio.

Dal punto di vista tecnico, l'opportunità a limitarsi alle dimostrazioni brevi è conseguenza del fatto che $x\otimes y$ risulta essere strettamente minore sia di x che di y, cioè strettamente minore del minimo $x\wedge y$. È interessante allora considerare la logica fuzzy corrispondente al caso in cui $x\otimes y= x\wedge y$. Per giustificare una tale logica, supponiamo ad esempio di volerci riferire alla nozione vaga di "non difficile" riferito a possibili percorsi per escursioni in montagna e scriviamo $ND(p)$ per indicare che il percorso p non è difficile. Un modo ragionevole di valutare la formula $ND(p)$ si può ottenere al modo seguente:

i) valutiamo ogni possibile tipo di difficoltà con un numero tra 0 ed 1,

ii) assumiamo che il grado di difficoltà di un percorso p sia misurato dal massimo dei gradi delle difficoltà che si trovano su p

iii) assumiamo che il grado di verità di $ND(p)$ sia 1 meno il grado di difficoltà di p.

Naturalmente, ciò comporta che nel dichiarare un percorso più o meno difficile non si tiene conto della sua lunghezza, o, se si vuole, del tempo necessario a percorrerlo. Se indichiamo con x e y due percorsi successivi e con $x+y$ l'unione di tali percorsi, allora appare naturale accettare l'assioma $ND(x)\wedge ND(y)\rightarrow ND(x+y)$ o quello, equivalente,

$$ND(x)\rightarrow(ND(y)\rightarrow ND(x+y)).$$

Supponiamo ora di avere una successione di piccoli percorsi p_1, p_2, . . . ciascuno successivo all'altro ed indichiamo con q_i l'unione dei percorsi $p_1, p_2, \ldots, p_i$. Supponiamo inoltre che tutti i percorsi siano facili e che quindi tutte le formule $ND(p_i)$ siano, ad esempio, valutate 0.9. Allora:
- utilizzando $ND(q_1)\rightarrow(ND(p_2)\rightarrow ND(q_2))$ ed applicando due volte *MP* riusciamo a provare $ND(q_2)$ con grado $0.9\otimes 0.9$.
- utilizzando $ND(q_2)\rightarrow(ND(p_3)\rightarrow ND(q_3))$ riusciamo a provare $ND(q_3)$ con grado $0.9\otimes 0.9\otimes 0.9$

. . .

- utilizzando $ND(q_{99})\rightarrow(ND(p_{100})\rightarrow ND(q_{100}))$ riusciamo a provare $ND(q_{100})$ con grado $(0.9)^{100}$.

Ora nel caso delle due T-norme sino ad ora considerate tale valore è o uguale o quasi uguale a zero mentre sappiamo che il grado di difficoltà di q_{100} è sempre 0.9. Appare allora chiaro che in questo caso la migliore interpretazione di $\otimes$ risulta essere il minimo. A sua volta, come è ovvio, il minimo non è adatto a spiegare il paradosso del mucchio di grano o quello dell'uomo calvo. Infatti, in tale caso, poiché ogni potenza di γ coincide con γ, l'affermazione "un mucchio con n chicchi è piccolo" sarebbe provata con grado $(0.9)^{n-1}=0.9$ qualunque sia il valore di n.

La situazione si complica ulteriormente se si osserva che lo stesso predicato *ND* potrebbe a volta richiedere una interpretazione di $\otimes$ come prodotto. Infatti, indichiamo con $p_0 ,\ldots,p_n,\ldots$ una successione di percorsi tali che
- la pendenza do p_0 è nulla

- la pendenza di p_n è lievemente superiore alla pendenza di p_{n-1}
- a parte l'eventuale pendenza, non presentano ulteriori difficoltà.

In tali ipotesi allora è lecito assumere

$$ND(p_0) \quad e \quad ND(p_n)\rightarrow ND(p_{n+1}).$$

Siamo esattamente nella stessa situazione del paradosso del mucchio di grano e $\otimes$ non può più essere il minimo!

Una via di uscita a tale problema potrebbe essere quella di avere più tipi di implicazioni a ciascuna delle quali è associato un *MP* con una diversa T-norma.

Prima di concludere questo paragrafo osserviamo che quando la parte valutativa di *MP* utilizza il minimo si ottiene una logica fuzzy molto semplice (*necessity logic*) in cui non ha importanza quante volte si utilizza un'ipotesi in una dimostrazione (si vedano Biacino, Gerla [1992], Gerla [1994]a, Dubois, Lang, Prade [1994]). Quello che conta è solo se tale ipotesi sia stata usata o meno e, naturalmente, quale sia il suo grado di validità. Indichiamo con *For* l'insieme delle formule di un dato linguaggio, con $v:For \rightarrow [0,1]$ un fuzzy sottoinsieme di formule da interpretarsi come *fuzzy insieme delle ipotesi*. Ciò significa che per ogni formula α il numero $v(\alpha)$ rappresenta il grado con cui è lecito considerare α come ipotesi Allora, avendo indicato con $\vdash$ l'usuale relazione di deduzione logica, non è difficile dimostrare che il grado $D(v)(\alpha)$ con cui è possibile provare una formula α a partire da v è dato dalla seguente semplice formula:

$$D(v)(\alpha)=Sup\{v(\alpha_1)\wedge...\wedge v(\alpha_n) \mid \alpha_1, ... ,\alpha_n \vdash \alpha\}.$$

In tale modo definiamo un fuzzy insieme di formule $D(v):For\rightarrow[0,1]$ che deve essere visto come il fuzzy insieme dei teoremi dedotti da v. Inoltre $D:For^L\rightarrow For^L$ può essere visto come un operatore di deduzione logica che associa ad ogni fuzzy insieme v di ipotesi il fuzzy sottoinsieme $D(v)$ dei relativi teoremi. Si noti che di una stessa formula α possono esistere dimostrazioni diverse e ciascuna dimostrazione può provare α con grado diverso. Naturalmente deve essere scelta la dimostrazione che fornisce una prova migliore, cioè una dimostrazione che utilizzi ipotesi $\alpha_1, ... ,\alpha_n$ più affidabili. Questo è il motivo per cui nella formula viene coinvolto l'estremo superiore.

Esempio. Supponiamo che
$v(\alpha) = 0.8$ se α è l'assioma delle parallele
$v(\alpha)=1$ se α è un assioma della geometria di Euclide
$v(\alpha) = 0$ altrimenti.
Allora
$D(v)(\alpha)=1$ se è possibile dimostrare α senza l'assioma delle parallele.
$D(v)(\alpha)=0.8$ se l'assioma delle parallele è essenziale per provare α
$D(v)(\alpha)=0$ se α non è un teorema della geometria euclidea.

8. Logica fuzzy come logica dei constraints.

Le cose sono comunque un poco più complicate di quanto abbiamo detto fino ad ora ed il tipo di formalizzazione ora proposto sembra dare luogo ad un altro tipo di paradosso. Infatti, riferendoci al paradosso del mucchio di grano, è ragionevole ammettere che anche la formula $P(2)$ sia assolutamente vera, e quindi valutata 1. In questo caso possiamo considerare, oltre alla dimostrazione esposta sopra che prova $P(3)$ con grado 0.81, anche la dimostrazione

$P(2)$ [vale con grado 1]

$P(2)\rightarrow P(3)$ [vale con grado 0.9]

allora

$P(3)$ [è provata con grado $1\times 0.9=0.9$]

che prova $P(3)$ con grado 0.9. In tale modo l'apparato deduttivo proposto sembra dare risposte differenti (e quindi incompatibili) circa il grado di validità di una formula, risposte dipendenti dal tipo di dimostrazione fatta. Questa è una caratteristica non eliminabile della logica fuzzy. Potrebbe darsi, infatti, che dopo avere effettuato una dimostrazione di una formula α utilizzando ipotesi non molto affidabili se ne trovi un'altra che, utilizzando ipotesi più sicure, conduce ad un più alto grado di verità per α.

La soluzione di tale apparente paradosso si ottiene interpretando meglio la parte valutativa di una regola di inferenza. Infatti i valori di verità che compaiono in una regola fuzzy vanno considerati come "soglie inferiori" ai possibili valori di verità delle corrispondenti formule. Ad esempio, lo schema

$$\frac{\alpha\ ;\ \alpha\rightarrow\beta}{\beta} \quad ; \quad \frac{x\ ;\ y}{x\otimes y}$$

deve essere interpretato dicendo che

- se il valore di verità di α è <u>almeno</u> x (appartiene all'intervallo $[x,1]$)
- e quello di $\alpha\rightarrow\beta$ è <u>almeno</u> y (appartiene all'intervallo $[y,1]$)
  - allora il valore di verità di β è almeno $x{\cdot}y$ (appartiene all'intervallo $[x{\cdot}y,1]$.

Allora il fatto che le due dimostrazioni della formula $P(3)$ forniscono due numeri diversi non costituisce una contraddizione. Semplicemente di tali numeri bisogna scegliere quello più grande poiché rappresenta una informazione maggiore (cioè un constraint migliore) sui possibili valori di verità di $P(3)$.

Indichiamo con $Val(d,v)$ la soglia determinata da una dimostrazione d stante le ipotesi v, allora bisogna assumere come grado con cui una formula α può essere provata stante l'informazione v il valore $D(v)(\alpha)$ dato da:

$$D(v)(\alpha)=\mathrm{Sup}\{Val(d,v) \mid d \text{ è una dimostrazione di } \alpha\}.$$

Il significato di $D(v)(\alpha)$ è che il valore di verità di α è maggiore o uguale ad $D(v)(\alpha)$ e che tale limite inferiore è il migliore possibile che può essere ricavato dall'informazione v. Naturalmente, data una logica fuzzy, è possibile dare una definizione rigorosa di dimostrazione e di soglia $Val(d,v)$ associata ad una dimostrazione. In proposito si rimanda a Pavelka [1979], Novak [1987], Novak, [1990], Gerla [1994]b.

In termini più generali, la parte valutativa di una regola di inferenza dovrebbe stabilire come determinati "constraints" (cioè condizioni, vincoli) sui possibili valori di verità delle premesse consentano di calcolare "constraints" sui possibili valori di verità della conclusione. Un tale punto di vista è ancora da sviluppare. Ad esempio si potrebbero considerare constraints del tipo intervallo. Dovremmo allora considerare regole di inferenze del tipo

$$\frac{\alpha_1, \dots , \alpha_n}{\alpha} \quad , \quad \frac{C_1, \dots ,C_n}{f(C_1, \dots , C_n)}$$

con $C_1, \dots, C_n$ e $f(C_1,\dots,C_n)$, *vincoli* sui possibili valori di verità, cioè sottoinsiemi dell'insieme dei valori di verità. L'interpretazione in questo caso sarebbe che:

- se si è provato che le formule $\alpha_1,\dots,\alpha_n$ hanno valori di verità negli insiemi $C_1,\dots,C_n$
- allora la formula α ha valore di verità nell'insieme $f(C_1,\dots,C_n)$.

Inoltre un sistema di assiomi dovrebbe essere una funzione $V{:}For \rightarrow P(L)$ il cui significato è che, per ogni formula α, il valore di verità di α appartiene $V(\alpha)$ (cioè verifica il constraint $V(\alpha)$). Infine, ogni dimostrazione d di α determinerebbe un vincolo $Val(d,v)$ ed il migliore constraint possibile per il valore di verità di α si otterrebbe tramite l'uguaglianza

$$D(v)(\alpha)=\bigcap\{Val(d,v) \mid d \text{ è una dimostrazione di } \alpha\}.$$

9. Difficoltà della logica fuzzy

Nonostante il fatto che la logica fuzzy sia ormai passata da una fase iniziale *"ingenua"* ad una fase di rigorosa matematizzazione, esiste ancora qualcosa di insoddisfacente proprio nell'idea di base e non è detto che le formalizzazioni proposte non debbano subire ulteriori evoluzioni. Ad esempio un primo problema si presenta nell'uso di un fuzzy set per rappresentare un concetto vago. Infatti l'assegnazione di un fuzzy set ad un concetto vago non può essere un processo preciso. La nozione di "grande" può essere rappresentata da diversi fuzzy sets, alcuni dei quali sembrano buoni modelli ed altri meno buoni. In un certo senso si dovrebbe interpretare la nozione di "grande" con una fuzzy classe di i fuzzy insiemi. Ma anche questo non sembra bastare, infatti una interpretazione che sembra buona (ad esempio accettabile con grado 1) in un contesto potrebbe essere non buona in un altro contesto. Nella nostra mente esiste una nozione di grande che cambia continuamente. La stessa interpretazione sembra potersi evolvere nel tempo in un gioco linguistico difficilmente dominabile dal formalismo matematico.

Esistono anche veri e propri paradossi che non sembrano facilmente risolvibili. Ad esempio supponiamo di dover scegliere tra due persone A e B per una assunzione e che per questo scopo ci si voglia servire di una serie di test. Supponiamo inoltre che i due candidati abbiano conseguito entrambi punteggio 50 in una scala da 1 a 100. Allora, in base alla logica fuzzy l'asserzione Supera_test(A)$\wedge\neg$(Supera_test(B)) dovrebbe essere valutata 0.5. Ma ciò appare paradossale e chiunque direbbe che tale asserzione è completamente falsa (se si vuole, completamente ingiusta) in quanto le capacità dei candidati si sono mostrate uguali.

Bibliografia

1. Biacino L., Gerla G., [1992], Generated necessities and possibilities, *International Journal of Intelligent Systems*, 7, 445-454.
2. Biacino L., Gerla G., [1998], Logics with approximate premises, in c.s. su, *International Journal of Intelligent Systems*.
3. Boldrin L., Sossai C., [1995], An algebraic semantics for possibilistic logic. *Proceedings of the International Conference on Uncertainty in Artificial Intelligence* (UAI 95), Montreal, Canada.
4. Chakraborty M. K., [1994], Graded consequence: further studies, in c.s. su *Journal of Applied non-classical Logic.*
5. Cignoli R., D'ottaviano I., Mundici D., [1995], *Algebras of many-valued sentential calculi of Lukasiewicz* (Portuguese), Collection CLE, Center of Logic, Epistemology and History of Science vol 12, State University of Campinas, SP, Brazil.
6. De Witt R., [1993], Vagueness, Semantics and the Language of Thought, *PSYCHE*, 1, December 1993. 039-723
7. Di Nola A., Gerla G., [1986], Fuzzy models of first order languages*, Zeitschr. f. math. Logik und Grundlagen d. Math.,* Bd. 32, 331-340.

8. Dubois D., Lang J., Prade H., [1991], Fuzzy sets in approximate reasoning, part 2: logical approaches, *Fuzzy Sets and Systems* 40, 203-244.
9. Dubois D., Lang J., Prade H., [1994], Possibilistic logic, In: D.Gabbay, C.Hogger and J.Robinson (eds.) *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 3. Clarendon Press.
10. Dummett M., [1975], Wang's paradox, Synthese 30, 301-324.
11. Ferrante F., Gerla G., [1998], Grasping Infinity by Finite Sets, to appear on *Mathematical Logic Quarterly.*
12. Fine, K., [1975], Vagueness, truth and logic, *Synthese*, 30, 265-300.
13. Frege G., [1986], Scritti postumi (a cura di Eva Picardi), Bibliopolis, Napoli.
14. Gerla G., [1994]a, An extension principle for fuzzy logics, *Mathematical Logic Quarterly*, 40, 357-80.
15. Gerla, G., [1994]b, Inferences in probability logic, Artificial Intelligence, 70, 33-52.
16. Gerla G., [1997], Graded consequence relations and closure operators, *Journal of Applied non-classical Logic*, 6, 369-379.
17. Gerla G., [1998], Closure operators, fuzzy logic and constraints, in c.s. su *Fuzzy Sets, Logics and Knowledge-based reasoning*, D. Dubois Editor, Kluwer.
18. Gerla G., Tortora R., [1990], Fuzzy natural deduction, *Zeitschr. F. math. Logik und Grundlagen d. Math.*, Bd. 36, 67-77.
19. Goguen, J.A., [1968/69], The logic of inexact concepts, *Synthese*, 19, 325-373.
20. Hahnle R., [1993], *Automated Deduction in Multiple-valued Logics*, Clarendon Press, Oxford.
21. Hajek P., [1993], On logics of approximate reasoning, Neural Network Word 6/1993, 733-744. Vagueness, Semantics, and the Language of Thought ???
22. Höhle U., [1991], Foundations of fuzzy sets, *Fuzzy Sets and Systems*, 40, 257-296.
23. Novak V., [1987], First-Order fuzzy logic, *Studia Logica*, 46, 87-109.
24. Novak V., [1990], On the syntactico-semantical completeness of first-order fuzzy logic I e II, *Kybernetika*, 26, 47-66 e 134-152.
25. Pavelka J., [1979], On fuzzy logic I: Many-valued rules of inference*, Zeitschr. F. math. Logik und Grundlagen d. Math.,* Bd.25, 45-52.
26. Poincaré, [1992], *La science et l'hypothèse*, Flammarion Paris.
27. Poincaré H., [1904], *La valeur de la science*, Flammarion Paris.
28. Russell, B., Vagueness, [1923], *The Australian Journal of Philosophy and Psychology*, 1, 84--92.
29. Sanford, D. H., [1975], Infinity and vagueness, *Philosophical Review*, 84, 521-522.
30. Termini S., [1998], Su alcuni sviluppi delle scienze dell'informazione, Atti Sem. Mat. Fis. Univ. Modena, Supplemento al Vol. XLVI, 971-989.

31. Ying M.S., [1994], A logic for approximate reasoning, *The Journal of Symbolic Logic,* 59, 830-837.
32. Zadeh, L.A., [1965], Fuzzy Sets, *Information and Control,* 12, 338-353.
33. Zadeh, L.A., [1975], Fuzzy logic and approximate reasoning, *Synthese*, 30, 407-428.